MINISTERO DELL'AMBIENTE
E DELLA SICUREZZA ENERGETICA

Sistema Integrato di Monitoraggio (SIM)

Progetto Esecutivo

ALLEGATO _ Infrastruttura e servizi di supporto tecnico

Finanziato dall'Unione europea
NextGenerationEU

Storia del documento

Versione	Data	Autore	Autorizzato da	Descrizione delle modifiche
1.0	24/11/2023	RTI PSN	MASE	Rilascio prima versione

Sommario

1 Infrastruttura e servizi di supporto tecnico

1.1 Servizi di Supporto Tecnico

La modalità con la quale si andranno a descrivere i servizi di supporto tecnico del SIM viene declinata secondo la suddivisione dei moduli funzioni del PSN rappresentati in figura.

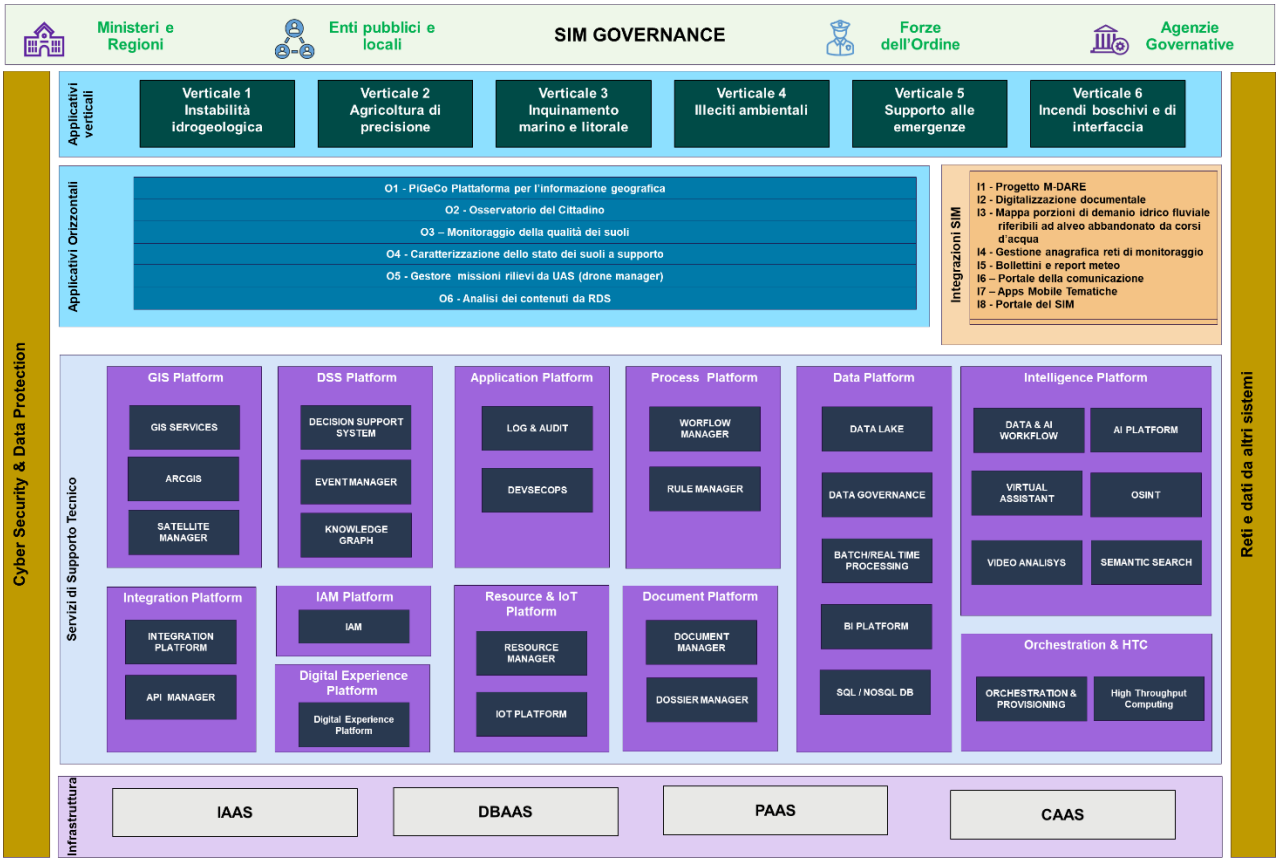


Figura 1 – SIM Suddivisione moduli funzionali

1.1.1 GIS Platform

1.1.1.1 Architettura tecnica

1.1.1.1.1 Introduzione

La piattaforma GIS eroga le funzionalità cartografiche del SIM. Ha una struttura ibrida, integra due soluzioni nell'ottica di fornire una piena fruibilità e flessibilità: la componente ArcGIS, leader nel settore, è affiancata alla componente GIS Service basata su GeoServer. Il patrimonio dei dati geospaziali della piattaforma è condiviso tra le due componenti; in questo modo, quale che sia la modalità di fruizione, è garantita la completa disponibilità del contenuto cartografico.

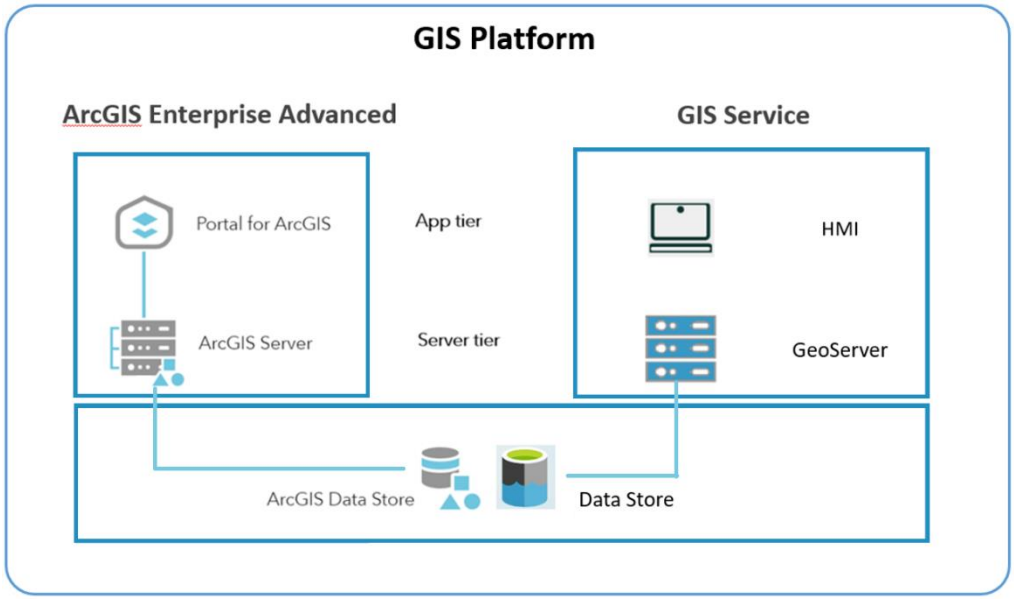


Figura 2 – GIS Platform, soluzione ibrida

Le due soluzioni hanno finalità diverse e quindi risultano complementari. La componente basata su GeoServer offre maggiore flessibilità e personalizzazione, ma richiede competenze tecniche avanzate; diversamente, la componente ArcGIS è più orientata all'utente finale e offre una suite completa di strumenti di più facile utilizzo.

La natura ibrida della soluzione non pone vincoli ad aspetti quali

- La sicurezza: entrambe le componenti offrono funzionalità di sicurezza per la protezione dei dati geospaziali; consentono di definire ruoli utente, autorizzazioni e restrizioni di accesso per garantire la riservatezza e la sicurezza dei dati geospaziali.
- La scalabilità: entrambe le componenti sono progettate per essere scalabili, consentendo l'espansione delle risorse e delle capacità in base alle esigenze dei progetti GIS.
- L'aderenza allo standard OGC: entrambe le componenti sono aderenti alle specifiche più recenti dell'Open Geospatial Consortium (OGC) per garantire l'interoperabilità tra i servizi GIS e altri sistemi.

- La capacità di collaborazione e condivisione: entrambe le componenti consentono la collaborazione e la condivisione dei dati geospaziali, sia all'interno di un'organizzazione che con altre organizzazioni, attraverso servizi web standard e protocolli.
- L'integrabilità con i database geospaziali: entrambe le componenti supportano l'integrazione con una varietà di database geospaziali, consentendo di accedere e gestire in modo efficiente i dati geografici in essi immagazzinati. GeoServer supporta diverse tipologie di database abilitati alla gestione geospaziale, come ad esempio PostgreSQL con estensione PostGIS, mentre la componente ArcGIS offre l'interoperabilità con diversi DB mediante l'estensione geospaziale proprietaria ArcSDE.
- La disponibilità di strumenti di analisi: sia GeoServer che la suite ArcGIS includono strumenti di analisi che consentono agli utenti di effettuare analisi complesse su dati geospaziali. Tuttavia, cambia la modalità con cui tali strumenti vengono resi disponibili: in ArcGIS essi costituiscono modelli o processi di analisi spaziale modificabili e configurabili, mentre in GeoServer vengono erogati mediante servizi web secondo lo standard WPS.

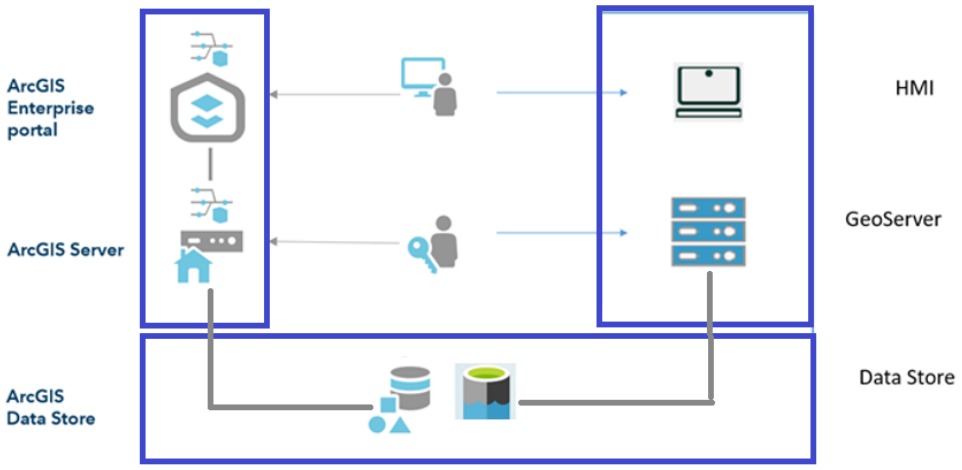


Figura 3 – GIS Platform, componente operativa

Dal punto di vista della fruizione delle funzionalità cartografiche, entrambe le componenti permettono sia l'accesso ai servizi web sia l'utilizzo di client cartografici.

Per quanto concerne l'accesso ai servizi, la componente GIS Service espone i servizi web tramite lo standard OGC mentre la componente ArcGIS affianca all'interfaccia OGC un'interfaccia proprietaria (ArcGIS REST API).

Per quanto riguarda la componente client (User Interface), l'approccio delle due soluzioni è diverso. In entrambi i casi, si ha la piena compatibilità con i client che fanno uso dei servizi OGC quale ad esempio QGIS. La componente ArcGIS prevede nativamente una suite completa e integrata di

strumenti mentre la componente GIS Service è accompagnato da HMI, un client cartografico integrato con i diversi servizi infrastrutturali del PSN quali ad esempio il PAAS IAM.

Confronto tra la piattaforma ArcGIS Enterprise e GIS Services

Nella seguente tabella sinottica vengono presentate le funzionalità offerte dalle due soluzioni, ArcGIS Enterprise e GIS Service, in relazione alle diverse caratteristiche richieste per le soluzioni software GIS.

Caratteristica	ArcGIS Enterprise	GIS Service
Tipo di Software	Software GIS proprietario	Piattaforma GIS open source
Mappatura e Visualizzazione, Analisi e Gestione dei Dati	ArcGIS Pro, QGIS	HMI, QGIS
Scalabilità	Orizzontale e verticale	Orizzontale e verticale
Integrazione con ArcGIS Pro	Totalmente integrato	Richiede configurazione
Integrazione con ArcGIS Online	Supportato	Supportato
Flessibilità nella Configurazione	Personalizzabile, ma con alcune restrizioni	Altamente personalizzabile, tramite accesso al codice sorgente dei componenti
Hosting di Servizi di Immagini	Supportato	Supportato
Analisi di Big Data	Supportato	MinIO e Spark
Implementazione in Cloud	Supporto per cloud pubblici e privati	Architettura a microservizi su tecnologia Kubernetes
Supporto per Sistemi Operativi	Windows, Linux, Kubernetes	Windows, Linux, Kubernetes
Gestione dei Dati Spaziali	ArcGIS Data Store	Database spaziali (PostgreSQL/PostGIS), HDM MinIO
Componenti Chiave	ArcGIS Server, Portal for ArcGIS, ArcGIS Web Adaptor, ArcGIS Data Store, ArcGIS Pro	GeoServer, PostgreSQL/PostGIS, Web Application Server, Tile Service, QGIS
Sicurezza dei Dati	Ampia gamma di strumenti di sicurezza	Pienamente configurabile a livello di amministratore e di utente
Accesso Multiutente e Multiruolo	Supporto integrato	Supportato
Condivisione Contenuti	Portal for ArcGIS	OGC Services, Tile Service
Condivisione Mappe	Portal for ArcGIS	Necessita di configurazione
Sviluppo di Applicazioni Web	SDK, API REST, Web AppBuilder	API REST, Librerie, Servizi OGC, Framework
Integrazione con QGIS	Richiede configurazione	Supportato
Strumenti di Geoprocessing e Analisi Spaziale	Ampia gamma di strumenti e servizi, Model Builder	Ampia gamma di strumenti e plugin customizzabili

1.1.1.1.2 ArcGIS Enterprise

ArcGIS Enterprise è la soluzione software integrata di ESRI per la gestione dei Sistemi Informativi Geografici, al livello di organizzazione. È costituita da componenti integrati che offrono funzionalità di creazione, visualizzazione e condivisione di mappe, sviluppo di applicazioni, analisi e gestione di dati geospaziali. ArcGIS Enterprise è strettamente integrato con ArcGIS Pro per la fase di authoring, e si interfaccia in modo efficace ad ArcGIS Online e Portal for ArcGIS per condividere i contenuti.

ArcGIS Enterprise è una piattaforma altamente collaborativa e flessibile, può essere installata su tecnologia Windows, Linux e Kubernetes, e consente di organizzare e condividere contenuti, dati e processi su qualsiasi dispositivo (desktop, tablet, mobile). Supporta implementazioni su infrastrutture cloud (pubbliche e private), così come on premises (su hardware fisico o virtualizzato).

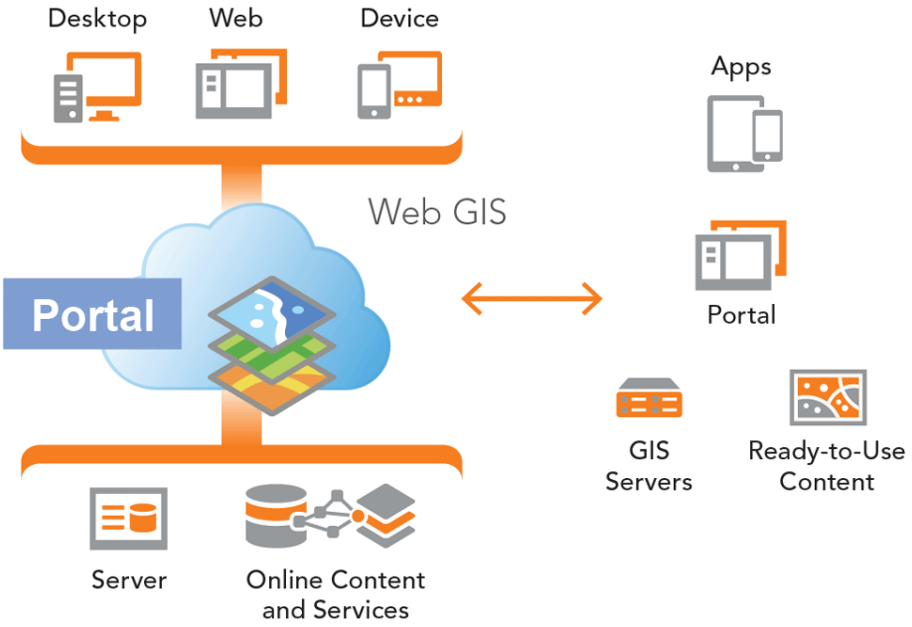


Figura 4 – componenti di ArcGIS

Nella sua configurazione standard la piattaforma ArcGIS Enterprise si configura come una suite di componenti che costituiscono un ecosistema completo per l'utente in grado di supportare l'intero ciclo di vita dei dati geospaziali, dalla fase iniziale di creazione e acquisizione dei dati (authoring) alla loro gestione, analisi e infine alla disseminazione all'interno e all'esterno dell'organizzazione tramite la condivisione di mappe interattive e applicazioni personalizzate.

In questo modo, ArcGIS Enterprise si rivela fondamentale per le organizzazioni che necessitano di un approccio integrato e scalabile per il loro lavoro con dati spaziali, garantendo un flusso continuo dall'origine dei dati, all'accesso semplice, efficace e sicuro alle informazioni geografiche.

I componenti fondamentali di ArcGIS Enterprise sono i seguenti:

- **ArcGIS Server**: è un componente software server di ArcGIS Enterprise che fornisce informazioni geografiche disponibili attraverso servizi web GIS di varia natura;
- **Portal for ArcGIS**: è l'hub principale per la condivisione dei contenuti geografici, nel quale gli utenti creano, gestiscono e condividono dati spaziali, mappe e applicazioni;
- **ArcGIS Data Store**: fornisce lo storage dei dati per il server di hosting;
- **ArcGIS Web Adaptor**: integra ArcGIS Server e Portal for ArcGIS con i web-server esistenti e con l'infrastruttura di sicurezza dell'organizzazione.

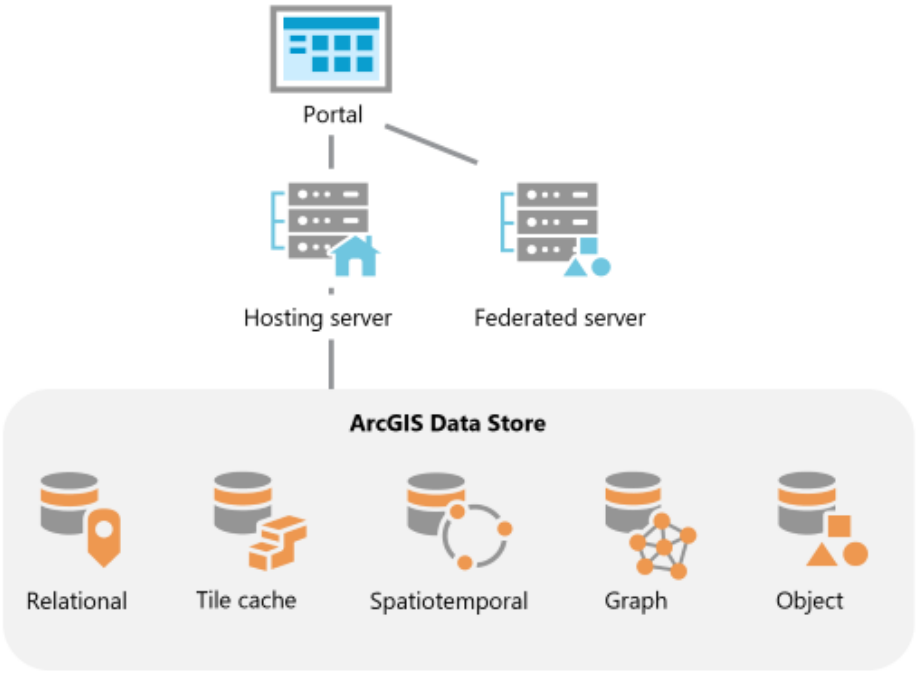


Figura 5 –ArcGIS Data Store

ArcGIS Server

ArcGIS Server è una componente fondamentale nell'ecosistema di ArcGIS Enterprise e svolge il ruolo cruciale di rendere i dati geografici disponibili ai client che li utilizzano per diversi scopi. ArcGIS Server consente la pubblicazione di servizi web GIS, che permettono di condividere layer, dati geografici e funzionalità avanzate come l'analisi geospaziale (geoprocessing) in tempo reale. ArcGIS Server offre scalabilità, sicurezza dei dati e flessibilità, permettendo alle organizzazioni di sfruttare appieno i vantaggi dei dati geografici nelle loro operazioni e applicazioni.

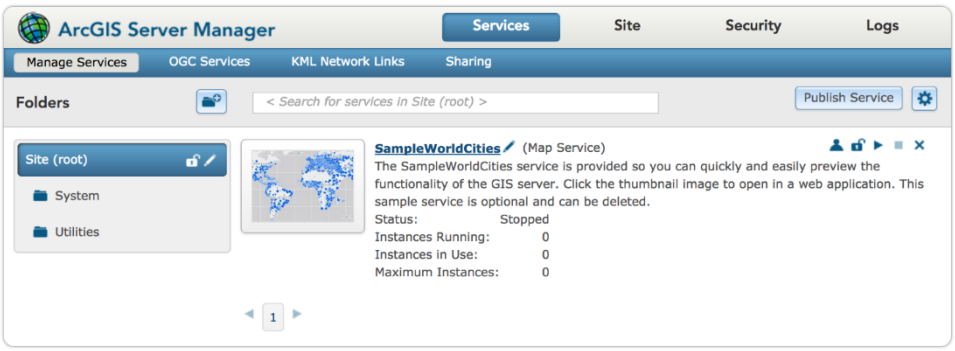


Figura 6 –ArcGIS Server manager

ArcGIS Server offre tipologie di servizi proprietari, affiancati dai protocolli standard OGC. In particolare tali tipologie di servizi includono:

- **ArcGIS MapServer**: servizio REST che consente di pubblicare mappe e layer sotto forma di immagini, secondo la stessa logica di funzionamento dei servizi OGC WMS (Web Map Service).
- **ArcGIS FeatureServer**: servizio REST che consente di pubblicare dati geografici vettoriali, in diversi formati, seguendo la logica di funzionamento dei servizi OGC WFS (Web Feature Service). Consente inoltre l'aggiornamento e la modifica dei dati da parte di utenti autorizzati, analogamente a quanto accade per il servizio OGC WFS-T (Transazionale).
- **ArcGIS Geoprocessing Service**: consente di eseguire analisi geospaziali avanzate e operazioni di geoprocessing su dati geografici, utilizzando direttamente le risorse del server.
- **ArcGIS Image Service**: consente di pubblicare e condividere dati GIS di tipo raster, tra cui immagini satellitari e foto aeree, e offre funzionalità avanzate come l'analisi raster.
- **ArcGIS Geocoding Service**: Fornisce funzionalità di geocoding per tradurre indirizzi testuali in coordinate geografiche e viceversa (geocoding e reverse geocoding).
- **ArcGIS Routing Service**: offre funzionalità di network analysis (analisi di rete) per calcolare percorsi ottimali tra luoghi specifici, e per effettuare algoritmi di service area e closest facilities.

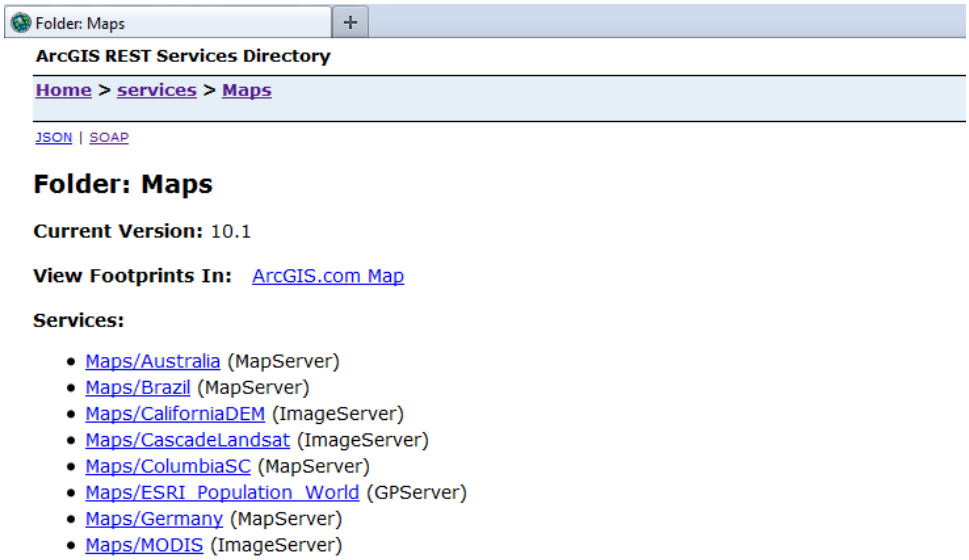


Figura 7 –ArcGIS Rest Services Directory

Portal for ArcGIS

Portal for ArcGIS è un componente chiave nell'ecosistema di ArcGIS Enterprise e fornisce un ambiente Web collaborativo che permette agli utenti di condividere, pubblicare e accedere a mappe, dati e applicazioni geografiche all'interno dell'organizzazione.

Nella configurazione standard sono disponibili le seguenti funzionalità:

- Galleria di mappe e app disponibili
- Visualizzatore di mappe che permette di creare e condividerle mappe
- Scene Viewer, che consente di visualizzare e condividere layer 3D e 2D in una Web Scene
- Funzionalità di creazione e gestione di gruppi, per controllare l'accesso ai contenuti
- Funzione di ricerca che consente di trovare informazioni geografiche, mappe, app e scene

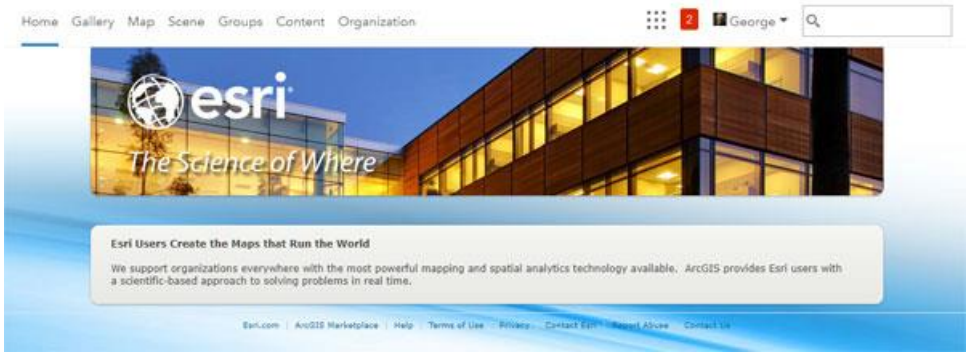


Figura 8 –Portal for ArcGIS

Il portale è composto principalmente da alcune pagine:

Pagina "Home": Questa è la pagina principale del portale ed è il punto di accesso per gli utenti. Fornisce un riepilogo delle informazioni più rilevanti per l'organizzazione, inclusi collegamenti rapidi a mappe, app, dati e contenuti rilevanti. Gli utenti possono personalizzare questa pagina per visualizzare le risorse più importanti per il loro lavoro.

Pagina "I miei Contenuti": Questa pagina è un luogo centrale per la gestione dei contenuti. Gli utenti possono visualizzare, caricare, modificare e organizzare mappe, dati e applicazioni. È possibile anche controllare l'accesso ai contenuti e condividerli con altri utenti all'interno dell'organizzazione.

Pagina "Gruppi": Questa pagina permette agli utenti di creare e gestire gruppi di collaborazione. I gruppi consentono agli utenti di lavorare insieme su progetti specifici, condividere risorse e collaborare in modo efficace. Gli utenti possono unirsi a gruppi rilevanti per il loro lavoro o creare gruppi personalizzati.

Pagina "Galleria": La Galleria è una vetrina di mappe, app e dati condivisi all'interno dell'organizzazione. Questa pagina permette agli utenti di scoprire e accedere a risorse geografiche pertinenti e applicazioni pronte all'uso.

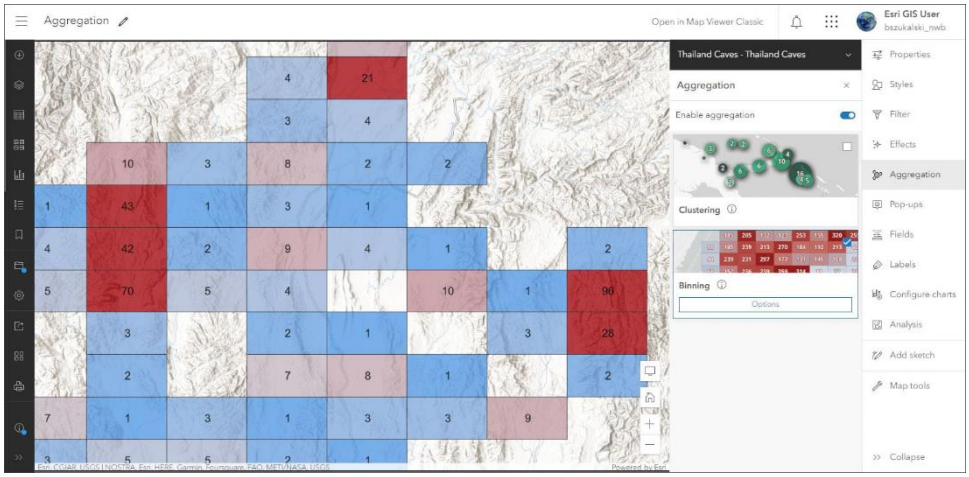


Figura 9 –Portal for ArcGIS: galleria

Portal for ArcGIS offre una serie di servizi che vanno oltre la semplice condivisione di mappe.

I servizi del Portal consentono di creare un ambiente collaborativo in cui gli utenti possono esplorare e analizzare i dati, creare mappe personalizzate, eseguire analisi geospaziali e sviluppare applicazioni geografiche. Inoltre, il portale fornisce strumenti di sicurezza avanzati per controllare l'accesso ai contenuti e garantire la protezione dei dati sensibili.

In sintesi, Portal for ArcGIS è una piattaforma essenziale per la gestione, la collaborazione e l'accesso ai dati geografici all'interno di un'organizzazione.

Figura 10 –Portal for ArcGIS: esempio 3D

Differenza tra Portal for ArcGIS e ArcGIS Online

ArcGIS Enterprise e ArcGIS Online sono due offerte di Esri per la gestione e l'uso di dati geografici, ma presentano differenze nel modo in cui sono implementati e nelle caratteristiche a disposizione.

ArcGIS Online è un servizio basato su cloud offerto come software-as-a-service (SaaS) da ESRI. È gestito completamente dal vendor, per quanto riguarda il monitoraggio, gli aggiornamenti e la manutenzione del sistema. Ciò garantisce che tutti gli utenti di ArcGIS Online abbiano accesso alle ultime funzionalità e ai dati più aggiornati. Inoltre, ArcGIS Online scala automaticamente per gestire carichi di lavoro crescenti, eliminando la necessità di utenti di fornire server o infrastrutture aggiuntive.

ArcGIS Enterprise, al contrario, è una piattaforma che le organizzazioni installano su infrastrutture che controllano, che possano essere nel cloud, on premises o su macchine virtuali. Ciò offre un controllo completo e la possibilità di personalizzare il sistema per soddisfare le esigenze specifiche dell'organizzazione. Gli utenti possono installare tutti i componenti di base di ArcGIS Enterprise su una singola macchina o distribuirli su più macchine. Questo approccio consente di gestire strategie di alta disponibilità, recupero da disastri e installazioni completamente isolate dalla rete Internet.

ArcGIS Data Store

ArcGIS Data Store è un componente di ArcGIS Enterprise che offre funzionalità di archiviazione e gestione dei dati. Questa tecnologia consente di archiviare e gestire vari formati di dati geospaziali e tabellari, all'interno dell'infrastruttura di ArcGIS Enterprise. È utilizzato per supportare le funzionalità di archiviazione dei dati necessarie per la pubblicazione di servizi GIS all'interno dell'organizzazione.

Le funzionalità principali di ArcGIS Data Store includono:

Archiviazione dei Dati: ArcGIS Data Store viene utilizzato per archiviare dati di tipo feature degli hosted feature layer o dei layer di output degli strumenti di analisi.

Tile Cache Data Store: consente di archiviare le cache degli hosted scene layer;

Big data store spazio-temporale: archivia i dati real-time provenienti da GeoEvent Server;

Scalabilità: è altamente scalabile e può essere implementato su server singoli o in configurazioni distribuite per gestire grandi volumi di dati.

Sicurezza: Offre funzionalità avanzate di sicurezza per proteggere i dati archiviati, inclusi controlli di accesso basati su ruoli e autorizzazioni.

In sintesi, ArcGIS Data Store è uno strumento fondamentale all'interno di ArcGIS Enterprise per la gestione e l'archiviazione di dati geografici, consentendo alle organizzazioni di sfruttare appieno le funzionalità di pubblicazione e condivisione dei dati geografici all'interno della loro infrastruttura GIS.

ArcGIS Web Adaptor

ArcGIS Werb Adaptor è un componente di ArcGIS Enterprise che consente l'esposizione delle funzionalità erogate dal Portal for arcGIS verso le utenze esterne; esso inoltra le richieste al modulo ArcGIS server gestendo gli aspetti tecnologici: la complessità della rete, il load balancing ed il failover.

L'utilizzo di ArcGIS Web Adaptor fornisce un singolo endpoint di accesso all'infrastruttura ArcGIS, rende accessibili le funzionalità mediante porte e protocolli standard, si integra con l'identity store e consente una gestione integrata della sicurezza e degli accessi (SSO – Single Sign On).

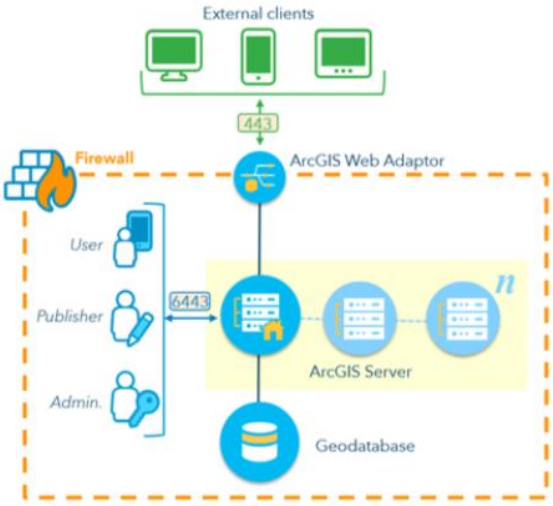


Figura 11 – ArcGIS Web Adaptor

ArcGIS PRO

Le funzionalità di ArcGIS Enterprise trovano una piena attuazione nella sinergia con lo strumento Desktop ArcGIS Pro. Sebbene la suite di ArcGIS Enterprise possa essere utilizzata in modo indipendente per diverse attività sui dati geospaziali, l'aggiunta di ArcGIS Pro al set di strumenti dell'organizzazione offre vantaggi significativi che consentono di sfruttare appieno le funzionalità dell'intera piattaforma. Per questo motivo ArcGIS Pro è considerato un complemento fondamentale per ArcGIS Enterprise.

ArcGIS Pro è uno strumento desktop GIS avanzato che integrato perfettamente con ArcGIS Enterprise, che consente agli utenti di accedere ai dati geografici aziendali, eseguire analisi avanzate e condividere informazioni con altri utenti in un ambiente di sicurezza controllato.

Affiancare ArcGIS Pro ad ArcGIS Enterprise offre una soluzione completa per la gestione dei dati geografici e l'analisi geospaziale all'interno dell'organizzazione.

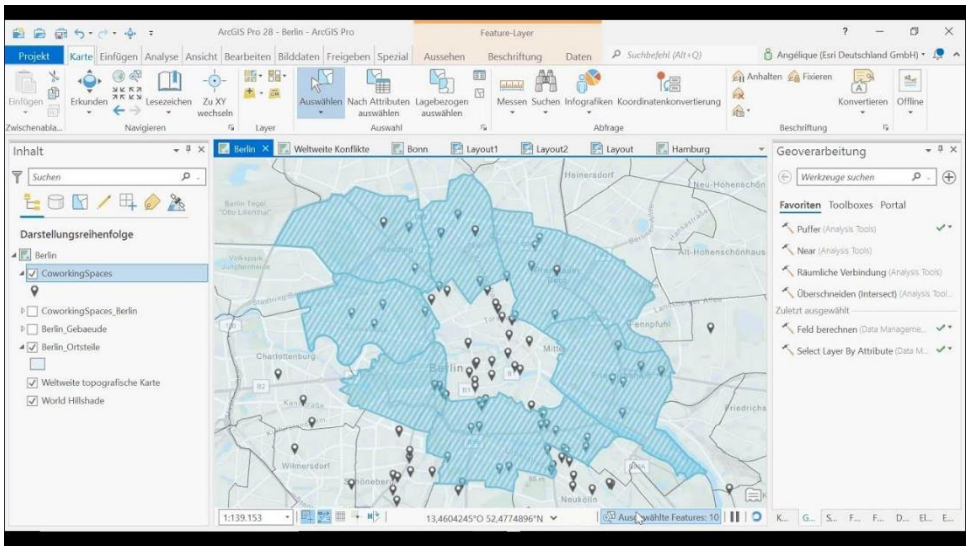


Figura 12 –ArcGIS PRO

ArcGIS Pro è un'applicazione progettata per l'analisi e la visualizzazione dei dati geografici. È progettata per sostituire e migliorare le funzionalità di ArcMap e ArcCatalog, le applicazioni desktop tradizionali di Esri. Le sue funzionalità chiave comprendono:

Interfaccia Utente Moderna: ArcGIS Pro offre un'interfaccia utente moderna e intuitiva, che lo rende facile da imparare e utilizzare. L'interfaccia è personalizzabile, consentendo agli utenti di organizzare gli strumenti e le finestre in base alle proprie preferenze.

Visualizzazione 2D e 3D: Supporta sia la visualizzazione 2D che 3D, consentendo agli utenti di creare mappe e scene tridimensionali interattive. Questa funzionalità è utile per la visualizzazione di dati geografici complessi e la creazione di mappe dettagliate.

Strumenti di Analisi Avanzata: ArcGIS Pro offre una vasta gamma di strumenti di analisi geospaziale avanzata. Gli utenti possono eseguire operazioni di geoprocessing, effettuare analisi di rete, eseguire analisi di spaziotemporali e altro ancora. Questi strumenti consentono di ottenere insight dettagliati dai dati geografici.

Editing dei Dati: È possibile modificare e aggiornare dati geografici direttamente in ArcGIS Pro. Gli strumenti di editing consentono di apportare modifiche precise ai dati, mantenendo l'integrità spaziale.

Integrazione con ArcGIS Online e Portal fro ArcGIS: ArcGIS Pro si integra perfettamente con ArcGIS Online e Portal for ArcGIS, consentendo di accedere e condividere facilmente dati, mappe e app tramite queste piattaforme online.

Elaborazione Immagini: Supporta l'elaborazione e l'analisi di dati raster, comprese immagini satellitari e foto aeree. È possibile eseguire analisi di immagini, classificazione e analisi del terreno.

Flussi di Lavoro Avanzati: ArcGIS Pro consente di creare flussi di lavoro personalizzati utilizzando il ModelBuilder, un'applicazione che consente di automatizzare complesse sequenze di operazioni di geoprocessing.

Gestione dei Dati: Fornisce strumenti avanzati per la gestione dei dati geografici, inclusi strumenti di geodatabase, catalogazione e gestione dei metadati.

Mappe Stampabili e Layout: Gli utenti possono creare layout di stampa personalizzati per produrre mappe di alta qualità per la pubblicazione o la condivisione. È possibile aggiungere titoli, legende, frecce nord e altri elementi di layout alle mappe.

Collaborazione e Condivisione: ArcGIS Pro facilita la condivisione di mappe, dati e progetti con altri utenti all'interno dell'organizzazione. Le mappe possono essere pubblicate come servizi web per l'accesso online o esportate in vari formati.

Sicurezza e Controllo dell'Accesso: Offre strumenti per la gestione della sicurezza e dei permessi per garantire che i dati siano accessibili solo agli utenti autorizzati.

ArcGIS Pro, per concludere, è una potente applicazione GIS ampiamente utilizzata in una varietà di settori, compresi governo, istituzioni accademiche, industria, e scienze ambientali, per l'analisi dei dati geografici e la creazione di mappe di alta qualità. La sua interfaccia moderna, le funzionalità avanzate di analisi e la facilità di integrazione con altre piattaforme Esri lo rendono una scelta estremamente popolare tra i professionisti GIS.

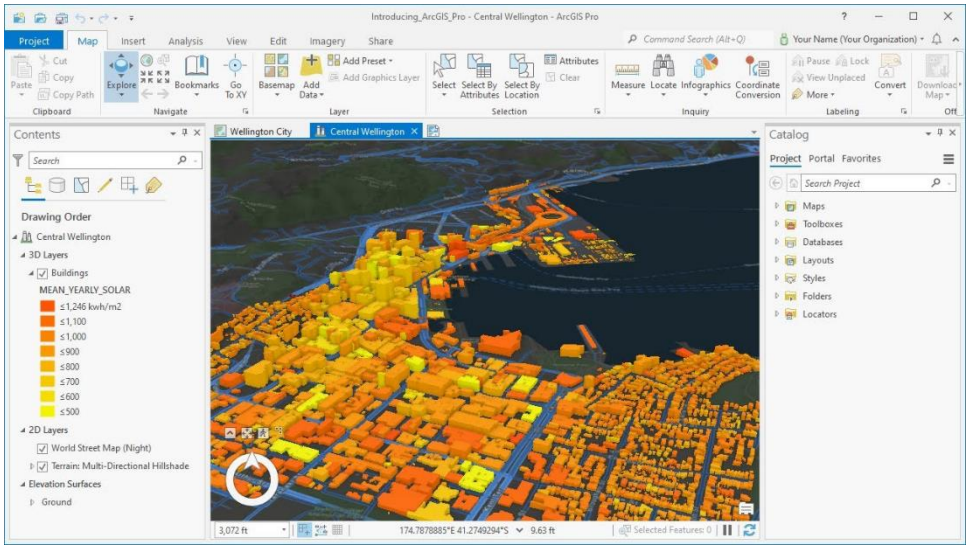


Figura 13 –ArcGISPRO: dashboard

1.1.1.1.3 GIS Service

La componente GIS Service basata su GeoServer è un sistema che espone, i servizi GIS attraverso lo strato GeoServer e tutta una serie di servizi accessori, tramite uno strato di micro servizi suddivisi per dominio/funzionalità. L'integrazione con le altre fonti di dati cartografici, la componente ArcGIS della Piattaforma in primis, avviene a livello di Tile Server.

Dal punto di vista delle funzionalità, il GIS Service rende fruibili tramite Web Services i dati geospaziali 2D e 3D provenienti dalle diverse fonti. In particolare, eroga:

- cartografia di base, proveniente da servizi online (Bing, OpenStreetMap, etc.), o da Tile Service locali;
- layer tematici di diversa natura, provenienti da varie fonti, negli standard OGC (WMS e WFS);
- layer di dati georeferenziati riguardanti la posizione di oggetti fissi (telecamere, asset, punti di interesse, device IoT, etc.) e il tracciamento di elementi mobili (risorse, terminali mobili, smartphone, eventi etc.);
- servizi di geocoding diretto e inverso, e di routing;
- modelli digitali tridimensionali generati mediante tecniche di rilievo laserscanner, LiDAR o fotogrammetrico e fruibili come nuvole di punti o mesh.

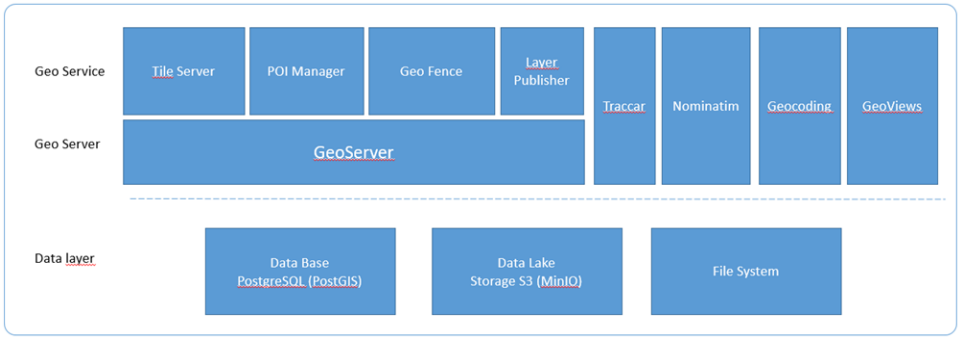


Figura 14 – GIS Service, componente operativa

Panoramica delle funzionalità

Di seguito alcune schermate di esempio che descrivono possibili utilizzi del componente GIS Service relativamente a tematiche di interesse ambientale.

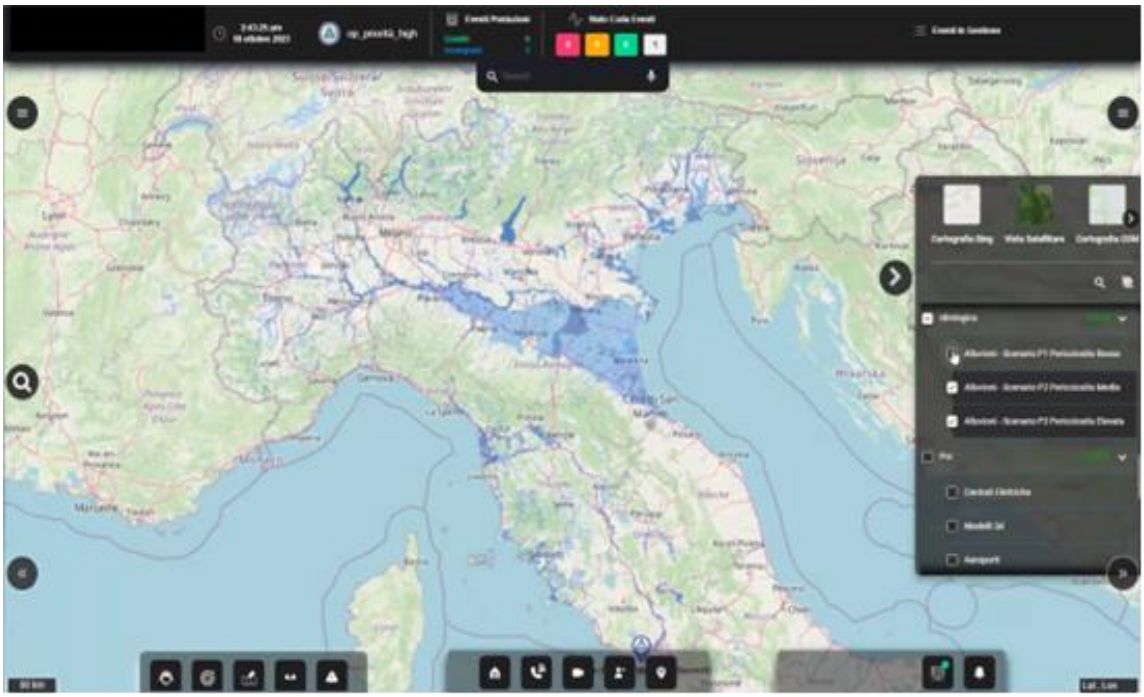


Figura 15 – GIS Service, esempio cartografia aree alluvionabili

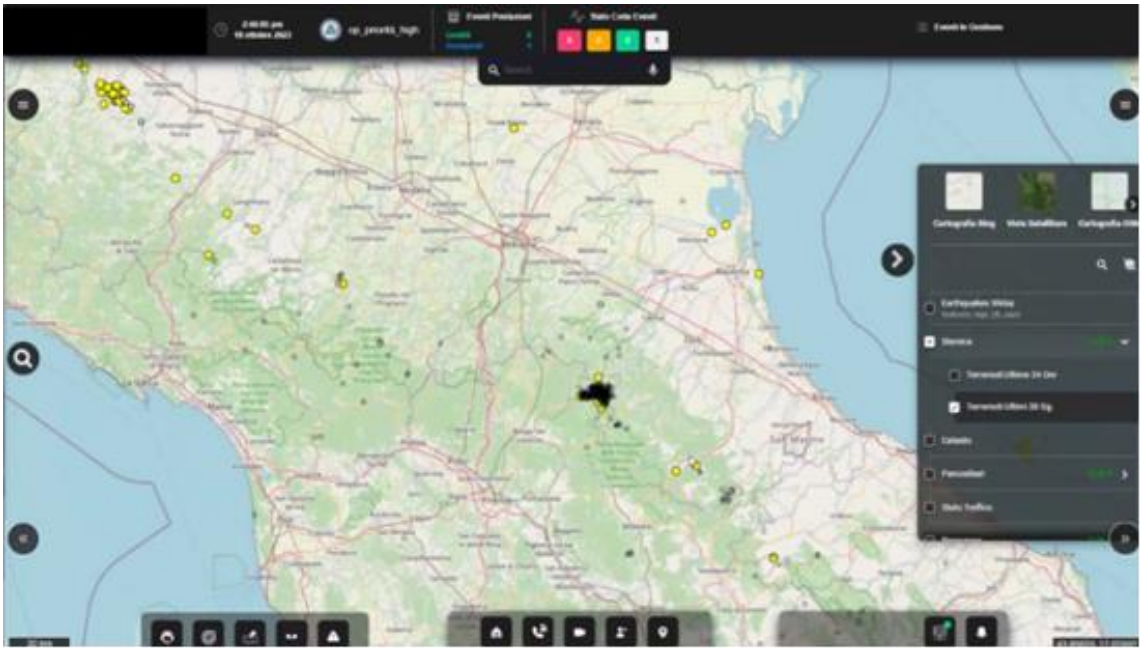


Figura 16 – GIS Service, esempio cartografia criticità geologiche

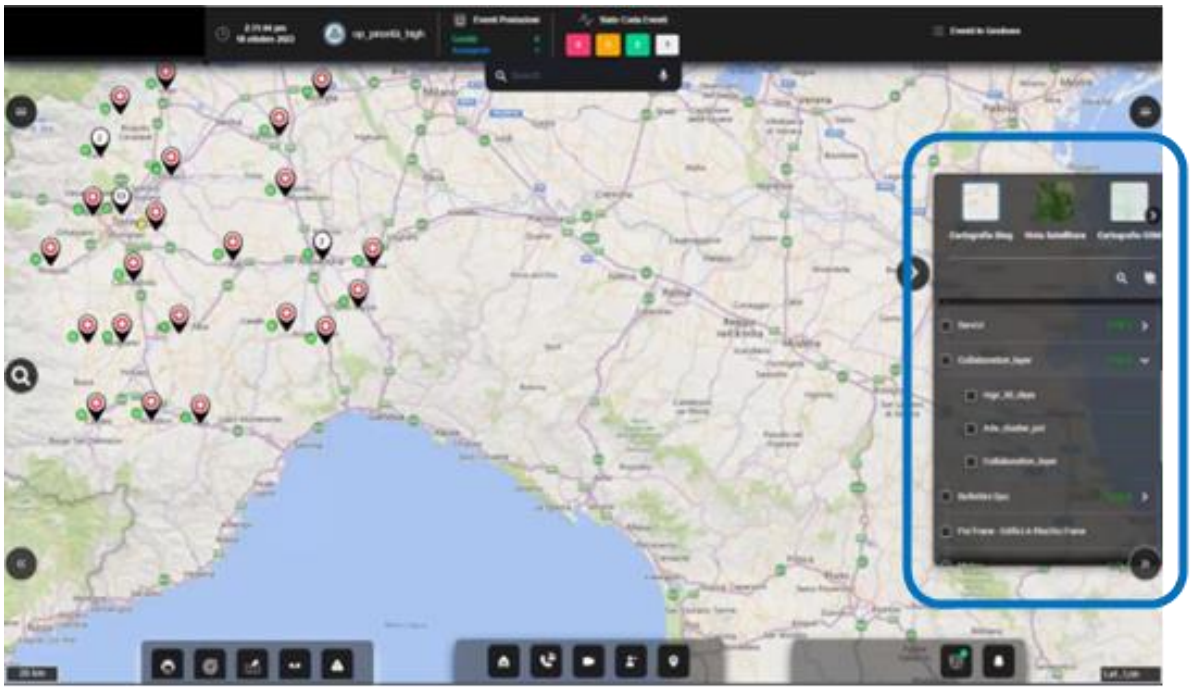


Figura 1 – GIS Service, esempio di tile render.

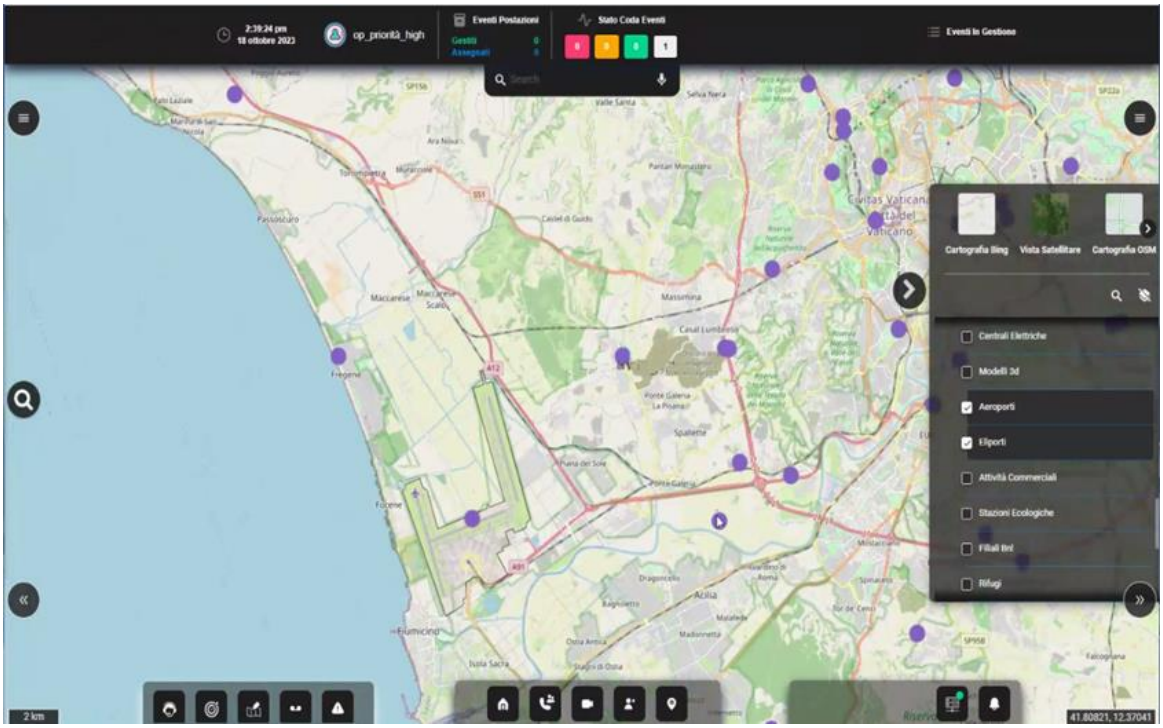


Figura 1 – GIS Service, esempio di POI (Point Of Interest).

Architettura di dettaglio

A livello di Geo Service, il sistema è composto da diversi micro-servizi, suddivisi per dominio, ognuno dei quali fornisce una funzionalità specifica e logicamente più semplice. Ogni micro-servizio espone

una I/F REST descritta tramite file in formato YAML in accordo allo standard Open Api Specification (OAS); tali file rappresentano il 'contratto' tra i client e i micro-servizi.

La figura seguente illustra schematicamente i microservizi che compongono il sottosistema GIS e le loro interazioni.

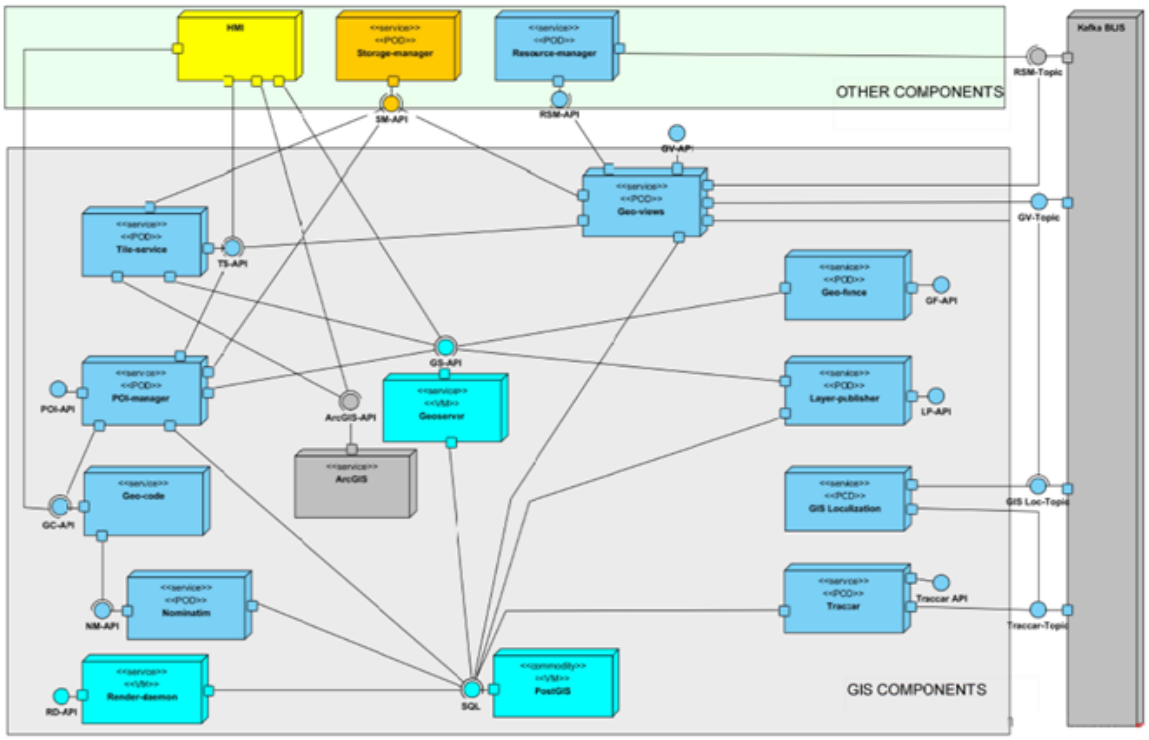


Figura 17 – GIS Schema Microservizi

I microservizi GIS disponibili sono dettagliati nei paragrafi che seguono.

Tile Render (Tile Server, render daemon)

È il componente FOSS che esegue il rendering delle tiles (immagini raster pre-renderizzate) utilizzate come base map, ovvero come 'sfondo' della cartografia; il rendering è basato sui dati di Open Street Map (OSM) ed è richiamabile in accordo allo standard 'de facto' denominato Slippy Map.

In fase di configurazione del sistema, i dati scaricati da OSM, vengono caricati, previa elaborazione, sul database (PostGIS) e tramite questo utilizzati dal demone per realizzare il rendering delle informazioni; le tiles risultanti sono immagini 256x256 in formato PNG e sono memorizzate localmente in metatile (caching) al fine di ottimizzare l'utilizzo del disco; ogni metatile contiene una matrice di 8x8 tiles. Il livello di dettaglio del rendering aumenta all'aumentare del livello di zoom. La cache viene invalidata, e quindi le tiles rigenerate, all'aggiornamento del database.

Le caratteristiche principali sono:

- Sorgente dati OSM. Il Render Daemon accede ai dati di OpenStreetMap memorizzati in un database. Questi dati comprendono informazioni dettagliate sulla struttura delle strade, edifici, punti di interesse e altri elementi geografici.
- Configurazione delle Regole di Rendering. Il daemon è configurato con regole di rendering che definiscono come i dati OSM dovrebbero essere rappresentati visivamente. Queste regole includono stili, colori, simboli e altre caratteristiche estetiche delle mappe e possono essere modificate in accordo a requisiti specifici.
- Generazione delle Tiles. Utilizzando le informazioni OSM e le regole di rendering, il daemon elabora le tiles per ciascun livello di zoom richiesto. Queste tiles rappresentano diverse porzioni della mappa geografica.
- Pyramid di Tiles. Le tiles vengono organizzate in una struttura a piramide e memorizzate sul file system in metatiles, in cui ogni livello successivo contiene tiles di dettaglio crescente. Questo consente una visualizzazione fluida delle mappe mentre si eseguono operazioni di zoom in avanti o indietro.
- Cache delle Tiles. Le tiles generate vengono memorizzate in una cache (file system) in modo che il carico elaborativo del rendering sia pagato soltanto alla prima richiesta ed avere un accesso rapido e una visualizzazione istantanea delle mappe agli utenti finali. Il rendering avviene a livello di metatiles ovvero il demone renderizza in realtà una metatile di 2048x2048 pixel al fine di ottimizzare accesso al DB e la risposta alle richieste.
- Pre-renderizzazione delle Tiles. è possibile forzare il daemon ad effettuare un pre-rendering in modo che le tiles siano disponibili sin dalla prima richiesta e non si avvertano rallentamenti nell'uso delle tiles.
- Sincronizzazione dei Dati. Il Render Daemon è in grado di gestire la sincronizzazione dei dati OSM nel database, in modo da riflettere eventuali modifiche apportate dai contributori di OSM. Questo può comportare la necessità di aggiornare o rigenerare le tiles interessate.
- Gestione delle Richieste Utente. Quando un utente richiede una mappa per una determinata area e livello di zoom, il Render Daemon è in grado di fornire le tiles corrispondenti dalla cache o, se necessario, genera nuove tiles in tempo reale.
- Scalabilità e Prestazioni. Il Render Daemon è progettato per gestire carichi di lavoro elevati, poiché i servizi di mappe basati su OSM possono avere un gran numero di utenti.
- Personalizzazione. A seconda delle esigenze, il Render Daemon può essere personalizzato per adattarsi a stili di mappa specifici, dati aggiuntivi e altre funzionalità uniche richieste dal progetto.
- Hosting: virtual machine Ubuntu server
- Interfaccia. L'interfaccia del Tile Render è esposta tramite semplici GET utilizzando la URL nel formato seguente http[s]://host[:port]/tiles/osm/{z}/{x}/{y}.png (dove z è il livello di zoom; [x, y] le coordinate della tile per quello zoom). Per esempio: http://localhost:8080/tiles/osm/11/1085/757.png

Figura 18 – GIS Esempio Tiles Zoom

Nominatim

È il componente FOSS che implementa il servizio di geocodifica e ricerca geografica. Questo software permette di convertire indirizzi o nomi di luoghi in coordinate geografiche (latitudine e longitudine) e viceversa. È spesso utilizzato in applicazioni e servizi che richiedono la capacità di cercare luoghi o mostrare mappe basate su dati geografici.

Nominatim è sviluppato come parte del progetto OpenStreetMap (OSM), che è una mappa del mondo collaborativa e basata su dati aperti. Il software Nominatim utilizza i dati di OSM per effettuare le operazioni di geocodifica e ricerca geografica.

In fase di configurazione del sistema, i dati scaricati da OSM, vengono caricati, previa elaborazione, sul database (PostGIS) e tramite questo utilizzati dal servizio

Hosting: container kubernates

Interfaccia. L'API esposta dal servizio è composta da due endpoint:

/search: l'API search permette di ottenere le coordinate geografiche corrispondenti ad un indirizzo passato in input come query parameter della URL, che ha il seguente formato: https://host[:port]/search?q=<address>&format=jsonv2

/reverse: l'API reverse permette di ottenere l'indirizzo relativo alle coordinate passate in input come query parameter della URL, che ha il seguente formato: https://host[:port]/reverse?lat=lat-value&lon=lon-value&format=jsonv2

Geocoding

È il servizio che permette la ricerca di un indirizzo (geocoding) o la traduzione di coordinate in un indirizzo (reverse geocoding), astraendo dal reale motore utilizzato e normalizzando la richiesta e la risposta. Espone una API che permette la ricerca delle coordinate corrispondenti da un indirizzo (geocoding) o la risoluzione di un indirizzo date le coordinate in input, ed effettua la normalizzazione della risposta, permettendo una reale astrazione rispetto al motore di geocoding effettivamente utilizzato.

I motori di ricerca supportati al momento sono gli analoghi servizi di Bing (Microsoft) e Nominatim (OSM).

Hosting: container kubernates

Interfaccia. L'API esposta dal servizio è composta da due endpoint:

/search: l'API search permette di ottenere le coordinate geografiche corrispondenti ad un indirizzo passato in input come query parameter della URL, che ha il seguente formato: https://host[:port]/api/search?query=<address>

/reverse: l'API reverse permette di ottenere l'indirizzo relativo alle coordinate passate in input come path parameter della URL, che ha il seguente formato: https://host[:port]/api/reverse/lon,lat

GeoServer

GeoServer è un server di mappe open source, sviluppato interamente in JAVA, che fornisce una piattaforma per la pubblicazione, la condivisione e la gestione di dati geospaziali attraverso il web. È uno strumento ampiamente utilizzato nella comunità GIS (Sistemi Informativi Geografici) per creare servizi di mappe interoperabili e accessibili tramite standard web.

Il server (o anche più istanze) GeoServer potrà essere fornito nel modulo GIS o si potrà utilizzare una (o più) istanza già in essere presso il cliente.

Di seguito una descrizione delle sue caratteristiche principali:

- **Pubblicazione di Dati Geospaziali**: GeoServer consente di pubblicare dati geografici in vari formati, tra cui:
  - Shapefile: un formato comune per dati vettoriali.
  - File Raster: ad esempio, immagini georeferenziate (geotiff).
  - Database Spaziali: è possibile pubblicare dati direttamente da database spatiali come PostgreSQL/PostGIS, MySQL, Oracle Spatial, MongoDB ed altri.
  - Dati vettoriali e raster remoti: GeoServer può anche accedere a dati geospaziali remoti tramite servizi web come WFS, WCS, e WMS di altri server e pubblicarli a sua volta.
- **Servizi Standard OGC**: GeoServer rispetta gli standard OGC (Open Geospatial Consortium) e offre servizi come:
  - WMS (Web Map Service): consente di distribuire mappe raster georeferenziate.
  - WFS (Web Feature Service): interfaccia per la distribuzione di dati vettoriali georeferenziati.

  - WCS (Web Coverage Service): utilizzato per la distribuzione di dati raster o di copertura.
  - CSW (Catalogue Service for the Web): per la ricerca e la distribuzione di metadati geospaziali.
- **Interoperabilità**: seguendo gli standard OGC (Open Geospatial Consortium), è compatibile con una vasta gamma di strumenti GIS, applicazioni e clienti web che rispettano gli stessi standard. Ciò favorisce l'interoperabilità tra diverse piattaforme e software GIS.
- Configurazione Layer: con GeoServer, è possibile configurare diversi layer o strati di dati all'interno di un singolo servizio. Ogni layer può essere personalizzato con regole di stile, limiti di accesso e altro ancora.
- **Filtraggio dei Dati**: è possibile applicare filtri ai dati durante la pubblicazione per consentire agli utenti di recuperare solo le informazioni rilevanti. Questo è particolarmente utile quando si hanno dati voluminosi e si desidera limitare ciò che viene visualizzato o scaricato.
- **Gestione delle Proiezioni**: GeoServer gestisce le proiezioni cartografiche in modo efficace, consentendo agli utenti di visualizzare e proiettare dati geografici in diversi sistemi di coordinate.
- **Gestione della Sicurezza**: GeoServer offre funzionalità di gestione della sicurezza per controllare l'accesso ai dati geospaziali. Gli amministratori possono definire ruoli e permessi per garantire che solo utenti autorizzati possano accedere e utilizzare determinati servizi o dati.
- **Styling e Symbology**: è possibile personalizzare la visualizzazione dei dati geografici attraverso la definizione di stili e simbologie. GeoServer supporta SLD (Styled Layer Descriptor) per creare regole di stile complesse che determinano l'aspetto dei dati sulle mappe come colori, simboli, etichette e altro ancora.
- **Cache di Mappe**: GeoServer supporta la creazione di cache di mappe per migliorare le prestazioni. Questa funzionalità consente di memorizzare in modo temporaneo le immagini delle mappe, riducendo così il carico del server e accelerando il recupero delle mappe.
- **Plugin e Estensibilità**: GeoServer è altamente estensibile tramite l'uso di plugin; è pertanto possibile scrivere nuovi plugin per estendere le funzionalità di base del server o personalizzare il comportamento del server in base alle esigenze specifiche del progetto.
- **Monitoraggio e Logging**: GeoServer offre strumenti di monitoraggio e logging che consentono agli amministratori di tenere traccia delle richieste dei clienti, delle prestazioni del server e degli eventuali errori.
- **Comunità Attiva**: GeoServer è supportato da una comunità attiva di sviluppatori e utenti. Ciò significa che è continuamente migliorato e aggiornato con nuove funzionalità e correzioni di bug.

Hosting: virtual machine Ubuntu server

Web Map Service

Il Web Map Service (WMS) è uno standard OGC (Open Geospatial Consortium) che definisce un protocollo per la distribuzione di mappe georeferenziate su Internet. Un servizio WMS fornisce mappe come immagini raster (gif, png, jpg) che possono essere visualizzate da client GIS (Sistemi Informativi Geografici) o semplicemente da browser web. I client interagiscono con un servizio WMS attraverso richieste HTTP standard chiamate "GetMap". Queste richieste includono parametri come la dimensione dell'immagine, l'estensione geografica desiderata, le proiezioni e le opzioni di stile.

Un servizio WMS può includere diversi "layer" o strati di dati geospaziali, ognuno dei quali può essere attivato o disattivato individualmente. Ogni layer rappresenta un insieme specifico di dati geografici. È inoltre possibile raggruppare i layer in gruppi (layer group) utilizzabili come se fossero un layer semplice.

Un servizio WMS può supportare diverse proiezioni cartografiche, consentendo agli utenti di visualizzare le mappe in una proiezione di loro scelta, o di riproiettare i dati da una proiezione all'altra

Come già accennato, i servizi WMS possono utilizzare stili personalizzati per definire l'aspetto delle mappe. Questi stili sono spesso definiti utilizzando SLD (Styled Layer Descriptor), CSS (tramite apposito plugin) o altre notazioni di stile.

Web Feature Service

Il Web Feature Service (WFS) è uno standard OGC (Open Geospatial Consortium) che consente di distribuire dati geografici vettoriali (come punti, linee e poligoni) su Internet.

A differenza del servizio WMS che fornisce mappe come immagini raster, i servizi WFS distribuiscono dati geografici vettoriali grezzi. Questi dati possono rappresentare oggetti geografici come edifici, strade, confini amministrativi, ecc.

I client possono interrogare i dati vettoriali tramite richieste HTTP standard chiamate "GetFeature". Queste richieste consentono agli utenti di recuperare dati specifici in base a criteri di ricerca definiti, come l'area geografica o gli attributi degli oggetti, e di elaborare in autonomia i dati ricevuti.

Il servizio WFS consente operazioni 'transazionali', come, ad esempio, l'aggiunta, la modifica e la rimozione di dati vettoriali direttamente sul server. Questo rende possibile l'editing dei dati geografici in tempo reale attraverso il servizio.

Il servizio supporta filtri che permettono agli utenti di definire criteri complessi per recuperare solo i dati che soddisfano determinate condizioni. Tali filtri sono espressi in CQL (Common Query Language), linguaggio di interrogazione comune utilizzato per filtrare e interrogare dati geospaziali all'interno di servizi GIS: Alcune caratteristiche del CQL sono:

- **Sintassi simile a SQL (Structured Query Language)**: lo rende familiare a coloro che hanno esperienza con i database relazionali. Tuttavia, è progettato specificamente per interrogare dati geospaziali.
- **Filtraggio dei Dati**: consente di definire condizioni di filtro per selezionare gli oggetti geografici desiderati in base ad attributi, posizione geografica o altre caratteristiche. Ad esempio, è possibile cercare tutti gli edifici con una certa altezza o tutte le strade in una determinata città.
- **Operatori Logici**: il linguaggio CQL supporta operatori logici come AND, OR e NOT, che consentono di combinare condizioni di ricerca in modo flessibile.

- **Operatori di Confronto**: è possibile utilizzare operatori di confronto come "=", "<>", "<", ">" per confrontare valori degli attributi.
- **Funzioni di Ricerca**: CQL fornisce funzioni di ricerca specializzate per dati geospaziali, come la ricerca per distanza (ad esempio, "trova tutti i punti entro 1 km da un certo punto"). Alcuni esempi sono:
  - LIKE: Restituisce risultati in cui il valore del campo corrisponde a un modello di stringa specificato (ad esempio, field LIKE 'pattern').
  - IS NULL: Restituisce risultati in cui il campo ha un valore nullo (ad esempio, field IS NULL).
  - IS NOT NULL: Restituisce risultati in cui il campo ha un valore non nullo (ad esempio, field IS NOT NULL).
  - INTERSECTS: Restituisce risultati in cui l'oggetto geospaziale interseca un'altra geometria (ad esempio, INTERSECTS (geometry, otherGeometry)).
  - BEFORE: Restituisce risultati in cui il valore del campo data è precedente a una data specificata (ad esempio, date BEFORE 'yyyy-MM-dd').
  - AFTER: Restituisce risultati in cui il valore del campo data è successivo a una data specificata (ad esempio, date AFTER 'yyyy-MM-dd').
  - DISTANCE: Restituisce risultati in cui la distanza tra due geometrie è inferiore a un valore specificato (ad esempio, DISTANCE(geometry1, geometry2) < distanceValue).
  - BBOX: Restituisce risultati in cui l'oggetto geospaziale si trova all'interno di una bounding box specificata (ad esempio, BBOX(geometry, minLon, minLat, maxLon, maxLat)).

Infine, il servizio supporta molteplici schemi dei dati in output, alcuni di seguito elencati:

- KML (Keyhole Markup Language) e GML (Geography Markup Language): KML e GML sono formati di output, XML based, ampiamente utilizzati per la visualizzazione di dati geospaziali su Google Earth e altre applicazioni. Contiene informazioni sulla geometria e gli attributi degli oggetti geografici.
- Shapefile: anche se i dati vettoriali in formato Shapefile non sono un formato nativo per i servizi WFS, GeoServer offre la possibilità di esportare dati in questo formato, che è ampiamente utilizzato nei sistemi GIS desktop.
- GeoJSON: è un formato testuale di dati leggibile dalle macchine ed è utilizzato in molte applicazioni web.
- CSV (Comma-Separated Values): semplice formato, molto comune per l'importazione in fogli di calcolo o per l'analisi dati.
- WKT (Well-Known Text): WKT è un formato di testo che rappresenta geometrie geometriche in modo leggibile dalle persone. È utilizzato principalmente per la definizione delle geometrie.

Traccar

Traccar è una piattaforma open source per il monitoraggio e la gestione di veicoli, dispositivi GPS e asset. Offre una vasta gamma di funzionalità e caratteristiche per aiutare le organizzazioni a tracciare e gestire i loro veicoli e risorse in modo efficiente.

Di seguito, una descrizione delle features principali supportate da Traccar:

- **Tracciamento in Tempo Reale**: Traccar offre un tracciamento in tempo reale dei veicoli e delle risorse, consentendo di visualizzare la loro posizione esatta su una mappa interattiva. Questa funzionalità è utile per il monitoraggio in tempo reale e per l'ottimizzazione delle operazioni.
- **Storico delle Posizioni**: è possibile accedere a un registro storico delle posizioni, consentendo di visualizzare le tracce passate dei veicoli e analizzare il loro comportamento nel tempo.
- **Notifiche**: supporta allarmi e notifiche personalizzabili. È possibile impostare notifiche per eventi come eccesso di velocità, ingresso/uscita da una zona geografica specifica (geofencing), perdita di connessione GPS e altro ancora. Le notifiche possono essere inviate tramite email, SMS, su BUS Kafka o altre forme di comunicazione.
- **Geofencing**: consente di definire aree geografiche virtuali, note come "geofence", e di ricevere notifiche quando un veicolo o una risorsa entra o esce da queste aree, o si discosta da un certo percorso oltre una data soglia. Questo è utile per il monitoraggio di zone specifiche o per il controllo degli accessi.
- **Report e Analisi**: la piattaforma offre funzionalità di generazione di report personalizzabili, che consentono di analizzare dati storici, comportamento dei veicoli, efficienza dei percorsi e altro ancora. È possibile esportare questi report per l'analisi ulteriore.
- **Gestione dei Dispositivi**: Traccar supporta la gestione dei dispositivi GPS e dei veicoli associati. È possibile aggiungere, configurare e monitorare dispositivi in modo centralizzato.
- **Accesso Multiutente**: la piattaforma offre un sistema di gestione degli utenti con ruoli personalizzabili, consentendo di definire chi può accedere a determinate funzionalità e dati.
- **API e Integrazioni**: offre API per l'integrazione con altre applicazioni e sistemi. Questo consente di personalizzare ulteriormente la piattaforma e di collegarla ad altri software aziendali.
- **Supporto Multipli Protocolli**: supporta una vasta gamma di protocolli GPS per la comunicazione con i dispositivi di tracciamento e la ricezione dei dati di posizione. Questa flessibilità consente a Traccar di essere compatibile con molti tipi diversi di dispositivi GPS. Di seguito un elenco non esaustivo dei protocolli GPS supportati da Traccar:
  - Wialon: è un protocollo proprietario utilizzato da molti dispositivi GPS. Traccar supporta il protocollo Wialon, consentendo l'integrazione di dispositivi che utilizzano questo protocollo.
  - Teltonika: è un produttore di dispositivi GPS noto, e Traccar offre un supporto completo per i protocolli Teltonika, inclusi FMXXXX e FMBXXXX. Questo consente di integrare dispositivi Teltonika nella piattaforma Traccar.
  - Meitrack: è un altro produttore di dispositivi GPS popolare. Traccar supporta i protocolli Meitrack, come il protocollo MVT600.
  - K103: è utilizzato da una serie di dispositivi GPS economici. Traccar supporta questo protocollo per l'integrazione di tali dispositivi.
  - Concox: è un altro produttore di dispositivi GPS, e Traccar è in grado di comunicare con dispositivi basati su protocolli Concox.
  - Xexun: è noto per i suoi dispositivi GPS economici. Traccar offre supporto per il protocollo Xexun, il che consente l'integrazione di dispositivi Xexun nella piattaforma.
  - Garmin: Traccar supporta anche dispositivi Garmin attraverso il protocollo FMI (Fleet Management Interface), consentendo il monitoraggio e la comunicazione con questi dispositivi.

- OBD-II: Alcuni veicoli moderni sono dotati di porte OBD-II (On-Board Diagnostics) che consentono di ottenere dati GPS direttamente dal veicolo. Traccar può comunicare con tali dispositivi utilizzando il protocollo OBD-II.
- Protocolli Standard GPS: Traccar supporta anche protocolli GPS standard come NMEA, che è comunemente utilizzato per dispositivi GPS marini e di navigazione.
- OsmAND: semplice protocollo HTTP based

Traccar espone un'API RESTful che consente di interagire con il sistema tramite richieste HTTP per accedere a dati e funzionalità. Questa API è progettata per consentire l'integrazione e l'automazione di diverse operazioni all'interno della piattaforma di monitoraggio di Traccar.

Nel sistema GIS, Traccar è il componente a cui convogliare la posizione di tutte le risorse mobili per le quali sia necessario seguire il movimento e storicizzare le posizioni. I cambi di posizioni ricevuti da Traccar sono propagati nel sistema (tramite i servizi GisLocalization, Geo Views) fino al sottosistema ResourceManager al fine di mantenere coerenti le informazioni sulla posizione delle singole risorse. Inoltre, i cambi di posizioni possono essere utilizzati da client, come ad esempio un client cartografico grafico, per mostrare in tempo reale la posizione di ciascuna risorsa.

Tile Service

È il componente che permette all'amministratore del sistema di configurare e rendere disponibili (publishing) ai client (es una HMI) i layers cartografici siano essi base o overlay: tramite questo componente è possibile 'agganciare' sia layers proprietari, ovvero esposti sul servizio WMS/WFS interno, che layer di terze parti, come ad esempio layers pubblicati da amministrazioni regionali o provinciali, o anche base map come quelle di Bing, OpenStreetMap o altri.

In particolare, consente l'integrazione anche la pubblicazione dei layer resi disponibili da una o più istanze locali di ArcgGIS –ESRI sia con le istanze esterne (terze parti).

La configurazione dei layers permette di definire diversi parametri come il titolo con cui i layers sono mostrati all'utente, il Bounding Box, la proiezione dei livelli minimo e massimo di zoom a cui visualizzare il layer, il timer di refresh, il clustering (per quelli vettoriali), il raggruppamento logico e altri ancora.

L'amministratore del sistema deve quindi configurare i layer attraverso l'esecuzione di 2 passi:

- **Configurazione di map server**: tramite la configurazione di nome, URL, path, abilitazione, eventuali parametri custom (come chiavi utilizzo), query parameters e tipologia tra cui:
  - 'slippy map' server, come ad esempio Bing, OSM o OpenWeather e altri
  - WebMap (OGC) server, come GeoServer, MapServer, ArcGIS, QGIS, Mapmik, Mapbox e altri
  - External server, come ad esempio ESRI ArcGIS
- **Configurazione dei layer supportati da ciascun map server**: nel caso di server OGC compliant o dei server ESRI, la configurazione si avvale di una funzionalità di discovery (se non inibita sul

server) che permette di leggere i layer pubblicati dal map server e di affettuare una veloce configurazione, eventualmente affinabile in un secondo momento.

I layer sono caratterizzati da molteplici attributi che permettono una configurazione estremamente fine. Gli attributi sono di seguito riportati:

- Available services (Read Only): indica se il layer supporta il servizioe WMS e/o il servizio WFS
- BBOX: indica l'estensione del layer
- Base: indica se il layer è base o overlay, quetsa informazione è utile per i client cartografici grafici
- CRS: indica la proiezion del layer
- Description: sintetica descrizione del layer
- Enabled: indica se il layer è abilitato oppure no; solo i layer abilitati saranno riportati sulle API /layers e /gisLayers
- Grouping: permette la definizione di un raggruppamento logico dei layer, ad esempio per tipologia o tema
- Max/Min zoom: il livello minimo/massimo di zoom a cui il layer deve essere visualizzato (utile per client cartografici grafici)
- Name: il nome asseganto al layer dall'amminitratore
- NativeName: il nome che il layer ha sul server di appartenenza
- Opacity: livello di opacità del layer (utile per client cartografici grafici)
- Order: ordine logico di visualizzazione del layer in eventuali controlli grafici (utile per client cartografici grafici)
- Parametes: parametri addizionali associalti al layer (es formato dell'immagine nel caso di layer WMS, filrti CQL etc)
- Path: eventuale path supplementare da aggiungere a quello specificato nel server di riferimento
- Queries: eventuali query parameters addizionali (es chiave per uso del servizio)
- Queriable: indica se il layer supporta la chiamata OGC GetFetureInfo
- Refresh: time di refresh del layer, utile per i layer che mostrano informazioni tempo varianti (es un layer di traffico o di Weather)
- Tile: il titolo del layer da usare in eventuali controlli grafici (utile per client cartografici grafici)
- Visible: indica se il layer deve essere visibile o no allo start di un client grafico (utile per client cartografici grafici)
- Z-Index: indica l'ordine di rendering del layer (utile per client cartografici grafici)

È possibile pubblicare un layer esposto da un map server più di una volta, configurando ad esempio in modo diverso alcuni attributi (es CQL filter, min/max zoom).

Hosting: container kubernates

Di seguito una breve descrizione dell'API esposta dal servizio:

/mapServers: permette la gestione dei map servers:

- GET: permette la lettura delle configurazioni dei map server, filtrando eventualmente quelli non abilitati tramite query parameter

- POST: permette la creazione di una nuova configurazione di map server passando gli attributi nel body della richiesta;

/mapServers/{serverId}: permette la gestione della singola configurazione con id specificato da serverId:

- GET: permette la lettura del singolo server
- PUT: permette la modifica (replace) della configurazione
- PATCH: permette la modifica parziale di attributi della configurazione
- DELETE: permette la cancellazione di un map server; la cancellazione comporta la rimozione automatica dei layer ad esso appartenenti

/mapServers/{serverId}/layerConfigurations: permette la gestione delle configurazioni di layer afferenti al map server specificato con serverId:

- GET: riporta le configurazioni dei layer appartenenti al map server, eventualmente filtrati per abilitazione o native name, utilizzando query parameters
- POST: permette la creazione di una nuova configurazione di layer passando gli attributi nel body della richiesta;

/mapServers/{serverId}/layerConfigurations/{layerId}: permette la lettura del layer appartenente ad un certo server

/mapServers/{serverId}/availableLayers: effettua il discovery (se supportato) dei layer pubblicati sul map server indicato; di fatto questa API causa l'esecuzione di una richiesta verso il server al fine di recuperare i layer che sono su di esso disponibili, ad esempio utilizzando la chiamata OGC GetCapabilities nel caso di server OGC compliant.

/layerConfigurations: permette la gestione delle configurazioni di layers

- GET: riporta le configurazioni dei layers, eventualmente filtrati per abilitazione, native nameo server id,. utilizzando query parameters

/layerConfigurations/{layerId}: permette la gestione della configurazione del layer specificato

- GET: legge la configurazione dei layer
- PUT: permette la modifica (replace) della configurazione
- PATCH: permette la modifica parziale di attributi della configurazione
- DELETE: permette la cancellazione del layer

/layers: permette la lettura di tutti i layers configurati sul servizio; questa API fornisce la vista finale dei layers in cui sono composte le informazioni memorizzate nelle configurazioni del server e del layer stesso. Per esempio, la URL riportata è composta utilizzando la URL e l'eventuale path del server più la parte di path del layer con tutti i query parameters (sia del server che specifici del layer). L'informazione ottenuta permette, ad esempio, la configurazione del layer su un client cartografico. Utilizzando i query parameters è possibile ottenere un sotto insieme dei layers o un diverso ordinamento della lista risultante.

/layers/{layerId}: permette la lettura di uno psecifico layer

/gisLayers: analoga alla API /layers, ma restituisce solo il layers 'public', ovvero quelli che hanno l'attributo public settato a true. Questa API è pensata per dare accesso ai layer configurati a terze parti, permettendo un filtraggio.

POI Manager

È il servizio che permette la configurazione, la ricerca e l'editing di Point Of Interest (POI), sia forniti da terze parti (per esempio da OSM o da uno specifico progetto) che inseriti tramite il servizio stesso; tramite configurazione e utilizzando i servizi di GeoServer e Tile Service, il servizio può pubblicare i POI su layers dedicati e renderli disponibili ai client.

La struttura dei POI permette la loro categorizzazione utilizzando la coppia di attributi [class, type] dove type specifica meglio il tipo di POI all'interno della class class. Ad esempio, definendo una classe shop, usando type è possibile specificare meglio la tipologia di negozio (es bakery, butcher, coffee etc). La categorizzazione dei POI è completamente libera.

Di seguito una descrizione degli attributi disponibili:

- id: id del POI, assegnato dal servizio
- poi_id: custom id, specificato dal client che crea il POI
- native_id: id native del POI sul sottosistema originario
- class: come detto, rappresenta la classe di POI
- type: come detto, rappresenta il tipo di POI, all'interno della classe
- geometry: è la geometria del POI: sono valide le geometrie Point, LineString e Polygon
- centroid: geometria di tipo Point del centroide della geometria
- heading: eventuale orientamento del POI rispetto al Nord
- label: etichetta da utilizzare nella visualizzazione del POI, ad esempio su un client cartografico
- status: eventuale stato del POI; la definizione degli stati fa parte della configurazione del servizio e sono lasciati agli specifici progetti; la definizione degli stati deve essere effettuata per le specifiche categorie di POI (coppia class, type)
- address: JSON con la risoluzione (geocode) della posizione del POI, se disponibile; il POI manager utilizza il servizio di geo-code per valorizzare questo campo. Da notare che la risoluzione dell'indirizzo avviene quando si accede ad un POI specifico se l'indirizzo non è ancora stato risolto.
- native_info: in questo attributo, di tipo JSON, possono essere memorizzate le info di dettaglio fornite dal sistema da cui è stato 'estratto' il POI, permettendo in tal modo di non perdere i dettagli
- additional_info: in questo attributo, di tipo JSON, possono essere memorizzate le info di dettaglio specifiche del progetto e modificabili attraverso le API esposte dal servizio
- poi_symbol: permette la memorizzazione di una URL di una immagine (simbolo) specifica per l'istanza di POI
- status_symbol: permette la memorizzazione di una URL di una immagine (simbolo) specifica per lo stato corrente del POI

Hosting: container kubernates

Interfaccia

Il servizio espone una interfaccia API REST che permette di gestire i POI, gli stati e i simboli associati; in particolare consente di:

- Listare i POIs
- Effettuare operazioni CRUD sui POIs
- Listare i simboli configurati per le classi di POIs
- Effettuare operazioni CRUD sui simboli
- Listare gli stati configurati per le classi di POIs
- Effettuare operazioni CRUD sugli stati

/pois: permette la gestione dei POIs

- GET: effettua la lettura paginata di tutti i POIs, permettendo di filtrare per classe, tipo, id e BoundingBox tramite i query parameters
- POST: permetta la creazione di un POI, specificando gli attributi nel body della request

/pois/{id}: permette la gestione del POI identificato da id

- GET: permette la lettura del singolo POI; qualora l'attributo address non sia stato risolto e sia stato configurato il servizio di geo code, il POI manager provvede alla risoluzione dell'indirizzo; tramite i query parameter è possibile forzare la risoluzione o nascondere il campo address
- PATCH: permette la modifica degli attributi del POI
- DELETE: permette la cancellazione del POI

/pois/{id}/additional-info: permette la lettura del solo attributo additional_info del POI (GET)

/pois/{id}/native-info: permette la lettura del solo attributo native_info del POI (GET)

/classifications: permette la lettura delle classificazioni dei POI presenti nel sistema (GET)

/symbols: permette la gestione dei simboli associati alle categorie di POI

- GET: effettua la lettura paginata di tutti i simboli associate alle categorie dei POI, permettendo di filtrare per classe e tipo tramite i query parameters
- POST: permetta la creazione di un nuovo simbolo, specificando le informazioni nel body della request

/symbols/{id}: permette la gestione del simbolo identificato da id

- GET: permette la lettura del singolo simbolo
- PATCH: permette la modifica degli attributi del simbolo
- DELETE: permette la cancellazione del simbolo

/statuses: permette la gestione degli stati associati alle categorie di POI

- GET: effettua la lettura paginata di tutti gli stati associate alle categorie dei POI, permettendo di filtrare per classe e tipo tramite i query parameters
- POST: permetta la creazione di un nuovo stato, specificando le informazioni nel body della request

/statuses/{id}: permette la gestione dello stato identificato da id

- GET: permette la lettura del singolo stato
- PATCH: permette la modifica degli attributi dello stato
- DELETE: permette la cancellazione dello stato

Layer Publisher

È il servizio che permette la pubblicazione (publish) e la rimozione (unpublish) di informazioni (features) su un layer cartografico, attraverso una API dedicata; il servizio utilizza un GeoServer dedicato e riservato per la pubblicazione del layer e il Tile Service, per renderne disponibile la configurazione.

Per la pubblicazione è necessario fornire la lista delle features che compongono il layer nel formato standard GeoJSON, i metadati per la pubblicazione del layer (title, name, description, CRS etc) ed un eventuale stile per il rendering grafico. Lo stile può essere specificato in SLD o in CSS e sarà memorizzato sul GeoServer ed associato al layer.

Hosting: container kubernates

Interfaccia

Il servizio espone una interfaccia API REST che permette di pubblicare su un layer cartografico le informazioni fornite in input:

/geoInfo: permette di leggere i layer configurati tramite questo servizio

- GET: effettua la lettura di tutti gli layer configurati tramite questo servizio, permettendo di filtrare per abilitazione tramite i query parameters
- POST: permetta la creazione di un nuovo stato, specificando le informazioni nel body della request

/geoInfo/{layerId}: permette di leggere il layer identificato da id

- GET: permette la lettura del layer
- PUT: permette la sostituzione delle informazioni contenute nel layer, ovvero la modifica dei metadati, delle features contenute e dello stile
- DELETE: permette la cancellazione del layer; la cancellazione comporta la rimozione del layer e dello stile da GeoServer e la rimozione della sua configurazione da Tile Service

GeoViews service

È il servizio che riporta su layers dedicati le risorse che sono configurate/caricate sul modulo ResourceManager. Il resource manager è il componente che permette la modellizzazione di risorse tramite la definizione di tassonomie, attributi e capabilities. Alcune risorse possono essere localizzate staticamente (si pensi a TLC sul territorio o a sensori IOT) o dinamicamente (se pensi, ad esempio, a droni, elicotteri o dispositivi mobili in generale). E' possibile caricare le risorse direttamente sul Resource Manager oppure caricarle da sottosistemi interni o di cui il progetto richieda la gestione tramite un livello di adattamento/astrazione (AbstractionLayer).

Il servizio GeoView materializza le risorse geolocalizzate su layers in accordo alla configurazione specifica del progetto, permettendo la loro visualizzazione, ad esempio, su client cartografici grafici. A tal fine il servizio utilizza il modulo Resource Manager da cui legge le risorse e le modifiche su di esse effettuate, espone una interfaccia WFS per la lettura delle features vettoriali in accordo allo standard OGC, e rende disponibile la configurazione dei layer tramite il servizio Tile Service.

La distribuzione delle geo views sui layer è altamente configurabile, si possono così ottenere, ad esempio, layer tematici (es tutte la risorse di tipo telecamera) o layer dedicati ai sottosistemi.

Data la natura delle risorse e la loro variabilità temporale, in particolare per quelle dinamicamente localizzate (mobili), la scelta progettuale è stata di non rendere disponibili i layer tramite WMS ma solo tramite WFS.

Di concerto con i servizi Traccar e GIS Localization, le posizioni delle risorse mobili sono aggiornate in tempo reale.

Hosting: container kubernates

Interfaccia

Il servizio espone una interfaccia API REST sinteticamente di seguito descritta.

/geo-views: permette la lettura paginata delle geoviews presenti nel sistema (GET)

/geo-views/{geoViewId}: permette la lettura della geoview identicata da geoViewId (GET)

/synch: forza il resync del servizio (POST), causando la ricarica delle risorse dal ResourceManager

/systems: restituisce la lista dei sottosistemi configurati sul Resource Manager (GET)

/systems/{systemId}: restituisce le informazioni relative al sottosistema identificato da systemId (GET)

/systems/{systemId}/synch: forza il resync delle soloe risorse appartenenti al sottosistema identificato da systemId (POST)

/wfs: API conforme allo standard OCG, restituisce le geo views nel formato standard GeoJSON; questa API può essere utilizzata da client cartografici che supportino WFS.

GIS localization service

Questo è un servizio di puro back-end e pertanto non espone API, se non quelle di monitoring del servizio stesso. Viene riportato per completezza.

Lo scopo del servizio è gestire la configurazione dei device (in gergo Traccar) sul servizio Traccar (via API) mappando le sole geo views dinamicamente tracciabili su device permettendo quindi a Traccar di gestire le notifiche di posizione ad esso inviato.

Inoltre, il servizio si registra sui topic di Traccar per ricevere le notifiche da esso generato; per ogni notifica di cambio posizione ne viene generata una normalizzata e arricchita con informazioni specifiche dei servizi GIS.

Tali notifiche alimentano 2 differenti topic configurati sul BUS Kafka: il primo topic, ad alto rate, permette ai sottoscrittori di ricevere tutte le modifiche di cambio di posizione (per esempio il servizio GeoViews); il secondo alimentato a basso rate per i sottoscrittori per i quali non ha rilevanza la posizione istantanea (es il Resource Manager).

Hosting: container kubernates

GeoFence service

È il servizio che permette di trovare le features (dati vettoriali) che sono in relazione con una geometria geografica (punto, linea, poligono etc) fornita in input; le relazioni disponibili sono, per esempio, intersects, touches, crosses, within, contais e altre.

Il servizio viene configurato con un GeoServer di riferimento e permette di ottenere la lista dei layer pubblicati; per ciascun layer si può ottenere poi la lista degli attributi presenti sulle features ed, eventualmente, configurare l'attributo di riferimento il cui valore sia ritornato in output nella ricerca delle features che soddisfano uno o più predicati in relazione ad una geometria fornita in input.

Ad esempio, un client potrebbe essere interessato a trovare le provincie su cui cade una certa area definita da un poligono, utilizzando le API del servizio potrà ottenere le features (in GeoJSON) che rappresentano le province risultanti o solo il loro nome, previa configurazione dell'attributo di riferimento per il layer.

Per ottenere la lista dei layer si potrà quindi invocare l'API

http://localhost:8080/api/geoFences

Individuato il layer (es italy:ProvCM01012021_WGS84) delle provincie il client potrà chiedere l'elenco degli attributi definiti per ciascuna provincia usando l'API

http://localhost:8080/api/geoFences/italy:ProvCM01012021_WGS84/attributeNames

output:

```
[
 "the_geom",
 "COD_RIP",
 "COD_REG",
 "COD_PROV",
 "COD_CM",
 "COD_UTS",
 "DEN_PROV",
 "DEN_CM",
 "DEN_UTS",
 "SIGLA",
 "TIPO_UTS",
 "Shape_Leng",
 "Shape_Area"
]
```

Per ciascun attributo è possibile chiedere la lista dei possibili valori tramite l'API (es usando l'attributo DEN_PROV, denominazione provincia)

http://localhost:8080/api/geoFences/italy:ProvCM01012021_WGS84/attributeNames/DEN_PROV/labels

output:

```
[
  "Vercelli",
   "Pisa",
   "Modena",
   "Cremona",
   "Isernia",
   "Campobasso",
   "Ragusa",
   "Sassari"
]
```

Se l'attributo riporta i valori di interesse da usare come label per ciascuna feature, allora si potrà configurarlo come attributo di default tramite l'API (POST senza body)

http://localhost:8080/geoFences/italy:ProvCM01012021_WGS84/attributeNames/DEN_PROV/labelsFrom

A questo punto si possono cercare le provincie la cui geometry soddisfa un determinato predicato in relazione ad una geometry fornita in input. Per esempio, le provincie che intersecano un poligono (es un'area di un terremoto o di una esondazione).

L'API da usare è (POST) per avere la lista delle features (provincie)

http://localhost:8080/api/selections/italy:ProvCM01012021_WGS84

per avere la lista delle solo label si userà invece

http://localhost:8080/api/selections/italy:ProvCM01012021_WGS84/labels

in entrambi i casi il body potrebbe essere come il seguente esempio


```
[
  {
    "predicate": "INTERSECTS",
    "geometry": {
      "type": "Polygon",
      "coordinates": [
        [
          [
            10.923156738281249,
            43.73736766145917
          ],
          [
            11.26922607421875,
            43.73736766145917
          ],
          [
            11.26922607421875,
            43.9349893800415
          ],
          [
            10.923156738281249,
            43.9349893800415
          ],
          [
            10.923156738281249,
            43.73736766145917
          ]
        ]
      ]
    }
  }
]
```


Hosting: container kubernates

Interfaccia

Il servizio espone una interfaccia che consente di:

- Listare le geo fences (layers) disponibili
- Listare i campi (attributi) disponibili su una certa geo fence
- Listare tutti i valori per un certo campi (attributo) su una certa geo fence
- Configurare il campo della feature da utilizzare come label
- Selezionare le features che soddisfano la ricerca con geometria e predicato

/geoFences: lista le geo fences (layers) disponibili sul GeoServer configurato (GET), la risposta riporta per quali geofences sia stata già configurato l'attributo da usare per estrarre le labels, e in caso affermativo quale sia

/geoFences/{geoFenceName}: legge la geo fence indicata con geoFenceName (GET)

/geoFences/{geoFenceName}/attributeNames: elenca gli attributi disponibili per la geo fence indicata con geoFenceName (GET)

/geoFences/{geoFenceName}/attributeNames/{attributeName}/labels: elenca tutti i valori disponibili per l'attributo indicato con attributeName per la geo fence indicata con geoFenceName (GET)

/geoFences/{geoFenceName}/attributeNames/{attributeName}/labelsFrom: setta l'attributo attributeName come attributo da utilizzare per l'estrazione delle label nelle operazioni di selezione per la geofence geoFenceName (POST)

/geoFences/{geoFenceName}/labels: elenca tutte le labels disponibili per la geo fence indicata con geoFenceName (GET); questa API ritorna errore se non fosse stato settato l'attributo da usare per le labels sulla geo fence.

/selections/{geoFenceName}: ritorna le features che soddisfano il predicato specificato nel body della richiesta (POST); i predicati sono quelli supportati dai filtri CQL (EQUALS, DISJOINT, INTERSECTS, TOUCHES, CROSSES, WITHIN, CONTAINS, OVERLAPS, RELATE): È possibile indicare più regole di selezione all'interno della stessa richiesta.

/selections/{geoFenceName}/labels: ritorna le labels delle features che soddisfano il predicato il predicato specificato nel body della richiesta (POST); sostanzialmente è come la API precedente, ma funziona solo per le geo fence per cui sia stato configurato l'attribuito da usare per le label. questa API ritorna errore se non fosse stato settato l'attributo da usare per le labels sulla geo fence.

/filters/{geoFenceName}: ritorna le features che soddisfano il filtro specificato nel body della richiesta (POST); i filtri possono combinare operatori gli booleani AND, OR e NOT con i predicati logici (EQUALS, LESS, GREATHER, LESS_OR_EQUALS, GREATHER_OR_EQUALS, LIKES) ed essere applicati a qualsiasi attributo delle features della geo fence (layer).

Mission Editor

È il componente grafico che permette la definizione, tramite disegno su client cartografico, di un'area di interesse, di inserire informazioni come nome e descrizione e di scegliere una risorsa, tra quelle disponibili, da inviare sull'obiettivo con scopi di sorveglianza; si avvale del servizio Layer Publisher per la pubblicazione del layer.

Hosting: container kubernates

Collaborative MapEditor

È il servizio di editing multiutente in real time, per l'editing collaborativo su HMI Cartografica: questo servizio permette ad un operatore di 'disegnare' su uno o più layers ad hoc, con lo scopo di evidenziare alcune informazioni di interesse, e di condividerle con gli altri operatori; si avvale del servizio GeoServer per la pubblicazione del layer.

Hosting: container kubernates

1.1.1.1.4 Approccio dualistico nella gestione integrata delle risorse GIS

La scelta di integrare in un'unica soluzione due piattaforme GIS, ossia ArcGIS Enterprise (tecnologia proprietaria Esri) e GIS Services (piattaforma basata su tecnologie completamente open source), è guidata dall'intento di sfruttare al massimo i vantaggi specifici offerti da ciascuna tecnologia.

ArcGIS si distingue per le sue potenti funzionalità GIS, un'interfaccia utente semplice ed intuitiva e una completa suite di prodotti integrati. Queste caratteristiche sono preziose per compiti di analisi geospaziale avanzata, gestione dei dati geografici e condivisione di risorse, dati e applicazioni. Esri ha costruito nel tempo un marchio solido nel settore GIS, soprattutto grazie al suo impegno nell'ambito della ricerca scientifica e dalla formazione specialistica, creando una base di clienti fidelizzati a lungo termine che contribuisce alla sua stabilità di mercato e garantisce la presenza di una ricca rete di utenti e professionisti che offrono un supporto molto prezioso, che va oltre quello ufficiale del prodotto.

D'altra parte, le tecnologie open source come GeoServer, PostGIS e QGIS offrono una grande flessibilità, utilissima nell'implementazione di soluzioni verticali, a costi contenuti e con il supporto di una comunità globale di esperti del settore.

L'integrazione tra le due soluzioni è fondamentale per sfruttare appieno i vantaggi di entrambe le piattaforme e creare un ambiente GIS sinergico e altamente efficace. Questo approccio dualistico permette alla soluzione integrata di adattarsi in modo agile a contesti diversificati e mutevoli, garantendo una gestione dei dati geografici in linea con le esigenze dell'organizzazione.

Flussi di dati tra ArcGIS e GIS-Services

Una delle fondamentali componenti dell'integrazione è la configurazione di flussi di dati bidirezionali tra ArcGIS e GIS Services. Questo processo permette di condividere dati e servizi tra le due piattaforme in modo tale da garantire che le informazioni siano costantemente accessibili e aggiornate in entrambi i sistemi. Ad esempio, i dati geospaziali raccolti o modificati in ArcGIS possono essere condivisi con i componenti di GIS Services e viceversa. Ciò assicura che il flusso di informazioni rimanga costante e coerente.

Interoperabilità dei dati

Un aspetto fondamentale dell'integrazione è l'interoperabilità dei dati. Infatti, entrambe le piattaforme supportino i medesimi formati di dati (proprietari e open standard), rendendo estremamente efficace lo scambio di informazioni. In particolare, l'adozione di standard aperti per i dati (ad es. GeoJSON) e per i servizi (OGC WMS e WFS) garantisce la piena interoperabilità tra ArcGIS e GIS Services, riducendo in modo drastico i problemi di incompatibilità che spesso vengono

riscontrati nei processi di integrazione, e permettendo uno scambio fluido di risorse tra le due piattaforme senza la necessità di complesse operazioni di conversione o traduzione.

API e Servizi Web

Le API e i servizi web offerti da entrambe le piattaforme giocano un ruolo chiave nella soluzione integrata. Utilizzando l'approccio dei microservizi, infatti, contribuiscono in maniera determinante alla creazione di processi di gestione ed elaborazione dei dati e allo sviluppo di applicazioni personalizzate che possono interagire con entrambe le piattaforme in modo flessibile, massimizzando i vantaggi di ognuna a seconda del tipo di operazione necessaria, e sfruttandone le peculiarità.

Condivisione delle mappe di entrambe le piattaforme

Un ulteriore contributo verso la completa integrazione della soluzione è la possibilità di incorporare su entrambe le piattaforme i layer cartografici e le mappe generate su ciascuna di esse. Ad esempio, è possibile visualizzare mappe create in ArcGIS all'interno delle applicazioni verticali di GIS Services. Analogamente, i layer creati sulla piattaforma GIS Services sono utilizzabili all'interno di mappe e applicazioni condivise nel Portal di ArcGIS Enterprise.

Questo consente, da una parte di poter disporre di un set di risorse geografiche condivise tra gi ambienti, dall'altra di poter implementare applicazioni verticali pienamente trasversali che attingono a tali fonti condivise.

Aspetti di sicurezza e di manutenzione

La soluzione integrata è dotata di un sistema centralizzato per la gestione degli utenti e dei ruoli all'interno delle due piattaforme per poter garantire l'accesso sicuro alle risorse geospaziali e la conformità con le policy dell'organizzazione. Questa pratica semplifica notevolmente la gestione degli utenti in un ambiente integrato.

Infine, è presente un sistema di monitoraggio integrato che consente di tenere traccia delle prestazioni e della disponibilità delle due piattaforme. Questo aiuta a risolvere tempestivamente eventuali problemi e a garantire un funzionamento continuativo dell'ambiente GIS integrato.

Punti di forza delle singole piattaforme

La scelta di utilizzare la piattaforma ArcGIS Enterprise o GIS Services nella gestione dei dati geografici e nello sviluppo di processi e soluzioni verticali dipende da diverse considerazioni.

Come primo aspetto bisogna considerare la necessità di poter garantire la continuità operativa, l'accesso ai dati e il mantenimento delle applicazioni già esistenti e consolidate sulle due piattaforme. Per diverse ragioni e decisioni aziendali, ma anche per venire incontro a specifiche necessità e esigenze dei clienti, nel corso del tempo si sono, infatti, sviluppati processi che si basano

esclusivamente su una delle due tecnologie, portando di fatto ad una situazione di coesistenza di dati e soluzioni accessibili da un particolare ambiente.

Un altro aspetto riguarda la necessità di utilizzare volumi di dati geospaziali residenti su una piattaforma specifica, che specialmente nel caso di grandi volumi, potrebbero rendere eventuali procedure di migrazione complesse e dispendiose. In questi casi, potrebbe essere preferibile mantenere entrambe le piattaforme e progettare soluzioni di condivisione e scambio di dati tra le due piattaforme.

In alcuni scenari, potrebbe essere utile sviluppare soluzioni o componenti di middleware che fungano da ponti tra le due piattaforme, consentendo la condivisione e l'interscambio di dati in modo efficiente. In altri casi invece la condivisione di risorse e soluzioni potrebbe essere gestita in base alle funzionalità di interoperabilità che offrono le singole piattaforme.

Per quanto riguarda la progettazione e implementazione di nuove soluzioni, sempre nell'ottica di massimizzazione dei vantaggi offerti da ognuna delle piattaforme GIS potrebbe essere conveniente orientarsi alla soluzione ArcGIS nel caso in cui risultino strategici i seguenti obiettivi:

- Effettuare Analisi Geospaziali Complesse in maniera semplice
- Avere un insieme di strumenti e applicazioni pronte all'uso, che coprano una gran parte delle esigenze degli utenti e operatori
- Gestione di dataset dettagliati e ricchi di informazioni
- Facilità d'Uso per Utenti non Esperti nella creazione di contenuti geografici e di applicazioni
- Disponibilità di Supporto e Documentazione dettagliati ed esaustivi
- Formazione professionale specifica sugli strumenti

La scelta della piattaforma GIS Services è strategica invece nei casi in cui ricoprano un ruolo fondamentale i seguenti aspetti:

- Necessità di utilizzare Standard Aperti
- Garantire funzioni di piena interoperabilità con altri sistemi esistenti
- Personalizzazione e Flessibilità, ossia consentire un alto grado di personalizzazione e flessibilità nell'implementazione e nella gestione dei servizi GIS
- Riduzione dei costi

Pertanto, la scelta di utilizzare un sistema unico che integri le due piattaforme GIS, ArcGIS e GIS Services, in modo sinergico rappresenta un approccio strategico che consente, come già detto, di massimizzare i vantaggi offerti dalle singole tecnologie. Questa coesistenza consapevole di due ambienti GIS offre la flessibilità necessaria per affrontare una vasta gamma di sfide e obiettivi dell'organizzazione.

1.1.1.1.5 Satellite Manager

Una componente abilitante necessaria agli scopi di monitoraggio che integra diverse piattaforme satellitari è fondamentale per affrontare le sfide ambientali e ottenere una visione completa e dettagliata del nostro pianeta.

L'obiettivo è quindi unire le potenzialità delle diverse piattaforme satellitari, consentendo di raccogliere dati provenienti da sensori ottici, termici, radar e altri strumenti specializzati e valorizzando gli asset attuali quali:

- Copernicus: Sentinel 1, Sentinel 2, Sentinel 3, Sentinel 5p con accesso ai prodotti a vario livello di processamento secondo le necessità
- Cosmo-SkyMed e Cosmo Second Generation
- Piattaforme IRIDE attraverso l'adattamento alle API di accesso

E quelli futuri quali i satelliti della futura costellazione IRIDE di cui è stato assegnata la realizzazione ad ESA e che fornirà capacità multi-sensore (SAR, Multi-spettrale, Iper-spettrale) e multi-risoluzione garantendo anche un revisit delle acquisizioni su territorio nazionale italiano come mai prima d'ora.

Oltre a questo, il programma IRIDE svilupperà anche servizi innovativi di generazione di contenuti a valore aggiunto fra i quali si prevedono:

- Servizi di monitoraggio su temi ambientali quali il rischio idrogeologico e le pratiche agricole
- Servizi di supporto alla gestione del territorio: land cover su vari tematismi, detection dei cambiamenti, delineazione di eventi di cambiamento improvvisi (ex. per supporto alla gestione degli incendi o di altri fenomeni che determinano rischi ambientali)
- Servizi di supporto alle emergenze tramite delineazione degli impatti di eventi quali alluvioni, incendi, terremoti, eruzioni vulcaniche etc.
- Servizi di monitoraggio del traffico e dell'inquinamento marittimo
- Servizi di tipo Digital Twin con l'obiettivo di creare una base di dati da diverse fonti non solo satellitari e sfruttando la modellistica fisica dei fenomeni

L'integrazione di tali dati in una struttura facilmente fruibile ed ottimizzata per il riutilizzo da parte dei servizi verticali è l'obiettivo del Satellite Manager. Le basi dati saranno fruibili dai sistemi informativi geografici (GIS) al fine di analizzare e visualizzare in modo efficace le informazioni raccolte e integrarle con le informazioni provenienti dalle altre sorgenti di dati (e.g. sensori, IoT, modelli, cartografia).

L'architettura del Satellite Manager prevede l'implementazione a micro-servizi di tipo worker per l'accesso alle varie piattaforme e una scalabilità basata su eventi. I micro-servizi, ospitati su una piattaforma di orchestrazione (Kubernetes), sono poi gestiti da uno scheduler che ne controlla l'esecuzione tramite l'analisi dei messaggi scambiati con il sistema di code (queue manager). Lo scheduler è in grado di combinare i diversi worker di estrazione/trasformazione dei dati, il polling verso le sorgenti e lìinterfaccia con le differenti API di accesso garantendo sia la ridondanza delle sorgenti ove possibile sia la componibilità dei vari step di recupero dati.

L'architettura è riportata nel seguente schema dove sono mostrate le componenti della piattaforma centrale. Nel seguito sono anche elencati i worker disponibili che possono essere combinati e anche usati:

- Per integrare i servizi disponibili quali Copernicus Dataspace e Copernicus Services
- Partendo dai template, per sviluppare facilmente nuovi worker da adattare alle piattaforme satellitari quali quelle sviluppate nel programma IRIDE (progetto in corso di sviluppo).

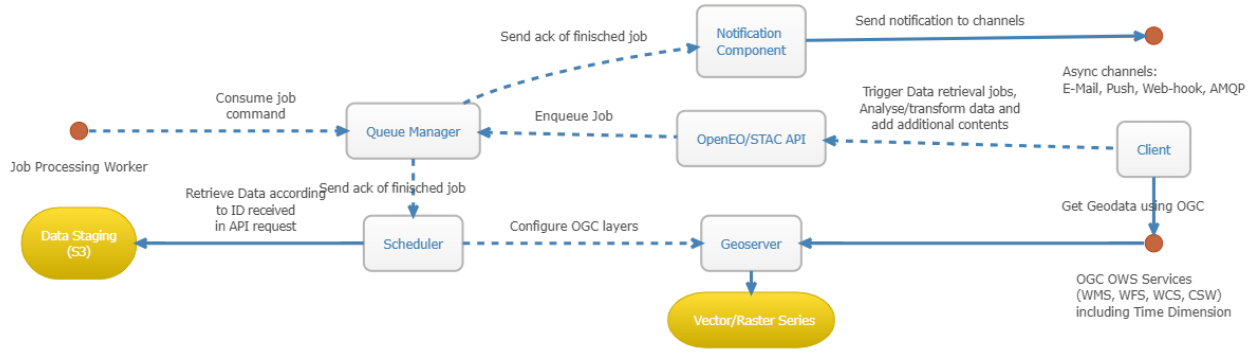


Figura 1 – Architettura di alto livello della piattaforma Satellite Manager.

Il core della piattaforma Satellite Manager espone una API unica di accesso a collezioni virtuali di dati e funzioni di processing basate sullo standard OpenEO (https://openeo.org/). Le API OpenEO costituiscono uno standard aperto che facilita l'interazione e l'accesso a servizi di elaborazione e analisi di dati geospaziali ed è orientato a garantire un accesso secondo una struttura dati di tipo DataCube sia per i dati Satellitari che per i prodotti a valore aggiunto derivanti da elaborazioni. Oltre ad OpenEO, sarà messa a disposizione l'API STAC (https://stacspec.org/), ovvero SpatioTemporal Asset Catalogue, ossia uno standard progettato per organizzare e descrivere dati geospaziali distribuiti nel tempo e nello spazio. Questo catalogo fornisce metadati strutturati e facilmente accessibili, consentendo agli utenti di individuare e recuperare dati geospaziali specifici in modo efficiente. STAC è particolarmente utile in scenari dove è cruciale tenere traccia delle variazioni spaziali e temporali dei dati, come nel caso delle immagini satellitari o di altri prodotti geospaziali.

Questi standard sono attualmente utilizzati da Copernicus come base per l'accesso ai servizi e rappresentano lo stato dell'arte al quale tutte le piattaforme satellitari si stanno adeguando. La piattaforma è inoltre dotata di un sistema di organizzazione e caricamento dei dati ottimizzato su Object Storage che può essere esposto tramite un Geoserver secondo gli standard OGC (https://www.ogc.org/).

Oltre all'architettura di base che consente l'esecuzione dei diversi job, la piattaforma è dotata di una serie di worker specializzati per il trattamento di dati di Earth Observation da varie piattaforme. Le funzionalità sono:

- Interfaccia e recupero dati ottimizzata dalle sorgenti dati dalle varie repliche di dati presenti dagli owner delle piattaforme esterne con il supporto a vari standard quali: STAC, Opensearch, OData, OGC CSW.
- Catalogazione del metadato (ove non fosse sufficiente il semplice brokering)
- Generazione di ARD (Analysis Ready Data) ove questi non siano presenti nelle piattaforme originali (e.g. compositi multi-temporali, mosaici)
- Ottimizzazione in formati Cloud-native sia Raster che Vettoriali fruibili tramite interfacce API o da software cartografici
- Pubblicazione automatica (e configurazione granulare degli accessi) su server cartografico Geoserver (https://geoserver.org)

Fornendo già i client verso molti standard di accesso e grazie alla presenza di un sistema a template dei worker, è inoltre possibile adattare tali worker alle future diverse piattaforme consentendo di aggiungere dati all'ecosistema SIM in maniera flessibile e garantendo una uniformità di gestione del dato recuperato che sia poi utilizzabile nei servizi verticali per ulteriori elaborazioni attraverso l'integrazione con i dati in-situ.

A livello di integrazione, la piattaforma espone i suoi servizi attraverso OpenEO/STAC con la possibilità di semplice auto-discovery fornendo quindi una modalità di accesso che non richiede integrazioni successive. Questo si adatta al tipo di integrazione secondo Data Mesh che accedono a API esterne per garantire il recupero dei dati utili alla logica applicativa.

A livello applicativo, in base alle descrizioni pubbliche dei servizi IRIDE, si riporta la lista dei contenuti che saranno resi disponibili dagli stessi e che saranno alimentati sia da dati Copernicus sia dalle missioni nazionali quali Cosmo-SkyMed and the future costellazioni IRIDE.

Prodotti a Valore Aggiunto	**Descrizione**
Rapid Damage mapping/delineation for Earthquake, Flood, Volcanoes, Fires, Extreme Meteo	Fornisce una mappatura rapida e una delimitazione delle aree danneggiate in seguito a un evento disastroso, consentendo una risposta immediata e mirata da parte degli enti di gestione emergenze.
Detailed Damage mapping/grading for Earthquake, Flood, Volcanoes, Fires, Extreme Meteo	Mappatura dettagliata e una valutazione dei danni, permettendo una gestione precisa della ricostruzione e favorendo la pianificazione di ripristino e ricostruzione.
Oil spills and sea pollution management	Detezione di sversamenti di petrolio e inquinamento marino basato su acquisizioni satellitari al fine di facilitare l'attuazione di misure di risposta e ripristino ambientale.
Support to S&R actions	Localizzazione di imbarcazioni speditiva a supporto delle operazioni di ricerca e soccorso in situazioni di emergenza in mare, ottimizzando l'efficacia degli interventi di soccorso attraverso l'integrazione con altre sorgenti dati quali AIS (Automatic Identification System).
Port Activity Monitoring	Monitoraggio delle attività portuali per analisi storica dei trend e supporto alla gestione delle infrastrutture portuali.

Prodotti a Valore Aggiunto	Descrizione
Environmental intelligence\fighting environmental crimes	Estrazione di informazioni per combattere i reati ambientali, supportando l'attuazione di azioni legali e misure di tutela dell'ambiente.
Land Cover/Land Use mapping & monitoring	Mappatura e monitoraggio dell'uso del suolo in supporto alla pianificazione e gestione del territorio.
Land consumption/soil sealing mapping & monitoring	Mappatura e monitoraggio del consumo di suolo consentendo di valutare rischi di diverso tipo da quello di inondazione alla perdita di bio-diversità e habitat.
Habitat mapping	Mappatura dettagliata degli habitat per la conservazione della biodiversità e la gestione ecosistemica.
Urban heat island monitoring	Monitoraggio delle isole di calore urbane per consentire una migliore pianificazione urbana e del verde urbano e la gestione dell'emergenza connessa con le ondate di calore.
Green urban areas characterization	Mappatura delle aree urbane a supporto della pianificazione e gestione sostenibile delle città.
National Forest mapping	Mappatura dettagliata delle foreste a livello nazionale per la gestione sostenibile delle risorse forestali e la conservazione della biodiversità.
Fire burnt area mapping	Mappatura delle aree bruciate dagli incendi per varie finalità.
Fire damage assessment	Mappa di valutazione dei danni causati dagli incendi che integra le informazioni sulle aree bruciate con gli asset coinvolti.
Forest health assessment	Mappa di valutazione dello stato di salute delle foreste per la gestione e la conservazione sostenibile delle risorse forestali.
Carbon stock indexes	Mappatura dell'accumulo di carbonio nelle foreste al fine di monitorare le politiche di mitigazione del cambiamento climatico.
Soil Organic Carbon (SOC) monitoring	Monitoraggio dei livelli di carbonio organico nel suolo per la gestione sostenibile delle risorse terrestri.
Erosion risk assessment	Valutazione del rischio di erosione del suolo in supporto alla pianificazione e l'implementazione di misure di conservazione del suolo.
Crop Production Areas (CPA) mapping & monitoring	Mappatura e monitoraggio delle aree destinate alla produzione agricola.
Water need and water volumes mapping	Mappatura e monitoraggio del fabbisogno idrico e dei volumi d'acqua.
Identification of indexes for crop health assessment	Valutazione dello stato di salute delle colture agricole.
Common Agriculture Policy (CAP) Support	Prodotti a supporto alla politica agricola comune (CAP) in sostegno all'implementazione di politiche agricole sostenibili.
Mapping of Ground Motion National coverage	Mappatura del movimento del suolo a livello nazionale per la gestione del rischio sismico e la pianificazione urbana.

Prodotti a Valore Aggiunto	Descrizione
Landslide monitoring (da Ground Motion)	Monitoraggio delle frane per la gestione del rischio geologico e degli asset coinvolti.
Cultural Heritage monitoring (da Ground Motion)	Monitoraggio del patrimonio culturale in supporto alle attività di conservazione e la gestione delle risorse culturali quali edifici e monumenti storici.
Critical Infrastructure monitoring (da Ground Motion)	Monitoraggio delle infrastrutture critiche.
Seismic wide area monitoring (da Ground Motion)	Monitoraggio sismico su vasta area per la valutazione e una risposta tempestiva in caso di eventi sismici.
Volcanic areas monitoring (da Ground Motion)	Monitoraggio delle aree vulcaniche per la gestione del rischio collegato alle attività vulcaniche quali i bradisismi.
National Territory DSM\DTM 1 mt geometric resolution	Modello Digitale di Superficie (DSM) e un Modello Digitale del Terreno (DTM) a risoluzione geometrica di 1 metro per l'intero territorio nazionale.
Coastal Mapping, Monitoring and Forecast	Mappatura, monitoraggio e previsioni costiere per la gestione delle zone costiere e la sicurezza delle comunità.
Air quality monitoring and forecast	Monitoraggio e previsioni della qualità dell'aria per la tutela della salute pubblica e la gestione dell'inquinamento.
Monitoring and assessment of pollutant emissions	Monitoraggio e valutazione delle emissioni inquinanti in supporto alla gestione dell'ambiente.
Re-analysis of air quality at national scale	Re-processing dei dati di qualità dell'aria al fine di garantire un dataset a livello nazionale in tema air-quality.
Hydro-meteorological mapping and monitoring atmospheric structure	Mappatura e monitoraggio idro-meteorologico per la gestione delle risorse idriche e la previsione meteorologica.
Monitoring of greenhouse gases and other Essential Climate Variables (ECVs)	Monitoraggio dei gas serra e di altre Variabili Climatiche Essenziali (ECVs) per la comprensione e la gestione del cambiamento climatico.
Classification of herbaceous agricultural crops	Classificazione delle coltivazioni agricole erbacee per la gestione agricola e la produzione alimentare.
Lightening Monitoring	Monitoraggio dei fulmini.
Hydrological and Hydraulic modelling, flood forecasting and sediment management	Modellazione idrologica e idraulica al fine di supportare la gestione del rischio da inondazioni, la gestione dei sedimenti e la gestione delle risorse idriche.

Focus sulle analisi delle immagini SAR (Synthetic Aperture Radar)

L'analisi interferometrica è una tecnica molto potente per rilevare e monitorare movimenti lenti e deformazioni di terreni e manufatti. Un'analisi approfondita di questi dataset permette di estrarre molte informazioni sulle condizioni dell'asset/terreno e sulle rispettive evoluzioni.

La tecnica interferometrica è in grado di ottenere misure in corrispondenza di punti di misura chiamati PS (Persistent Scatterer) che vengono automaticamente selezionati attraverso algoritmi specializzati. La tecnica, totalmente non invasiva, evidenzia punti PS del terreno o di manufatti soggetti a movimento, indicando per ogni PS velocità di spostamento e la sua evoluzione nel tempo (in corrispondenza di ciascuna osservazione satellitare) permettendo così di ricostruire il comportamento nel tempo. Le misure di spostamento sono dell'ordine di pochi millimetri, con densità di punti di misura che può raggiungere le decine di migliaia di rilevazioni per chilometro quadrato in corrispondenza di aree densamente urbanizzate.

L'analisi interferometrica è quindi uno strumento di analisi preliminare, monitoraggio sistematico e di supporto alla pianificazione di interventi, utile al fine dell'individuazione di luoghi e strutture a rischio, che mostrano la necessità di monitoraggi in tempo reale.

A valle dell'analisi interferometrica verranno individuate delle aree (identificate da poligoni) contenenti i punti con evoluzioni temporali delle deformazioni simili tra loro, che si definiscono CDZ (Common Deformation Zone). Per ognuna di queste aree verranno fornite le misure di deformazione ottenute dalla media delle misure dei PS che si trovano all'interno del poligono.

La tecnica dei Persistent scatterers si basa sull'identificazione di punti all'interno delle immagini SAR che mantengono una firma radar stabile nel corso dell'intervallo di osservazione, e sull'estrazione dell'informazione di spostamento dalla fase interferometrica del PS identificato. La tecnica è in grado di misurare gli spostamenti relativi tra i punti, e tale misurazione dipende dalla distribuzione, densità e rumore di fase dei PS stessi.

Per ottenere delle misurazioni affidabili, solo i punti con un rapporto segnale/rumore sufficientemente elevato sono presi in considerazione. La valutazione della bontà del PS dipende dall'analisi dell'evoluzione della loro fase e della loro ampiezza col tempo. La capacità di discriminare i PS dai non-PS aumenta col numero di immagini analizzate e con l'intervallo temporale analizzato: l'aumento delle immagini permette di ottenere una capacità di rilevamento migliore.

Misure PSP-IFSAR: geometria ascendente

La lista delle immagini COSMO-SkyMed acquisite in geometria ascendente e usata per le analisi interferometriche effettuate. I parametri associati sono la componente della baseline interferometrica ortogonale alla Linea di Vista (LOS), indicata con BN, e baseline temporale, indicata con BT.

Misure PSP-IFSAR: geometria discendente

La lista delle immagini COSMO-SkyMed acquisite in geometria discendente e usata per le analisi interferometriche effettuate. I parametri associati sono la componente della baseline interferometrica ortogonale alla Linea di Vista (LOS), indicata con BN, e baseline temporale, indicata con BT.

Analisi della decomposizione dei risultati nelle componenti est-ovest e verticale

Per quanto riguarda le misure di spostamento si è scelto di restituire il risultato della scomposizione del moto nelle direzioni verticale e orizzontale est-ovest in corrispondenza di tutte le date di acquisizione ascendenti e discendenti relative ad un periodo temporale

Le componenti verticali ed est-ovest del movimento si riferiscono ad una cella a terra di dimensione fissata e pari alla risoluzione delle misure, al cui interno sono disponibili almeno una misura ascendente ed una discendente. Ad ogni cella viene associato sia il valor medio delle misure ascendenti sia quello delle misure discendenti presenti al suo interno. Tali misure mediate rappresentano i valori di input del processo di separazione della componente verticale ed est-ovest. Dato che i dati SAR ascendenti e discendenti non sono sotto-campionati ma elaborati a risoluzione massima, essi possono essere usati per derivare le componenti verticali ed est-ovest delle deformazioni relativi a celle di risoluzione a terra di 10m × 10m.

1.1.1.1.6 Infrastruttura

Come accennato nei capitoli precedenti, la GIS Platform prevede sia componenti ospitate da un Container Platform sia componenti residenti su Virtual Machine.

I POD previsti per le due soluzioni del modulo GIS sono allocati sull'infrastruttura CaaS del PSN, in questo modo è gestito, ad esempio, l'autoscaling.

Sono altresì previste macchine virtuali per ottimizzare la capacità di calcolo e quindi gestire al meglio i carichi di lavoro in presenza di un elevato utilizzo. Il risultato è ottenuto grazie all'adozione di macchine virtuali con un elevato rapporto tra CPU e memoria.

Unitamente alla componente server, lo strato Data Store delle due soluzioni, è ospitato dal blocco logico Data System (RDS) del SIM e quindi dalle piattaforme e dai servizi del PSN. Sono coinvolte il PaaS Data Lake, il PaaS DB e le eventuali componenti infrastrutturali quali, ad esempio, il file system.

1.1.2 DSS Platform

Ogni evento generato dal sistema SIM è corredato di uno specifico contenuto informativo; oltre ai dati generici sulla data e l'ora, sulla località e sulla tipologia di evento, vi sono tutta una serie di informazioni relative al contesto operativo, agli elementi e alle tematiche inerenti l'evento. Queste informazioni sono estratte mediante analisi testuale della descrizione dell'evento, delle note e di eventuali allegati. Il contenuto informativo costituisce una mole di preziose indicazioni per il sistema ai fini della gestione dell'evento, sia in termini di valutazione delle azioni e delle strategie da intraprendere sia in merito alle risorse da impiegare.

L'insieme delle informazioni viene riportato in un'opportuna struttura dati che a sua volta è riportato su un **database a grafo**, questo per sfruttare al meglio l'interconnessione tra le informazioni ed ottenere prestazioni più efficienti nel contesto di un'interrogazione veloce del dato.

I dati messi a disposizione sul database a grafo, per struttura e per tipologia di informazioni offerta, ben si prestano per rivestire il ruolo di **knowledge base** rispetto ad un sistema di supporto alle decisioni che permetta la gestione dell'evento in maniera rapida ed efficiente.

Sfruttando gli attuali strumenti offerti dall'intelligenza artificiale, attraverso gli algoritmi di machine learning/deep learning allo stato dell'arte, la soluzione propone all'operatore strumenti e suggerimenti e **decision support system** e per la conduzione delle attività, offrendo capacità di:

- valutazione dell'insieme di **eventi correlati** o correlabili ad un evento di riferimento, in termini di correlazione spazio-temporale, tematica o con coinvolgimento di entità comuni;
- valutazione delle **strategie di gestione** degli eventi, anche tramite suggerimenti di azioni specifiche quali presa in carico dell'evento, chiamate verso contatti utili, invio di risorse opportunamente valutate, invio di informazioni sul campo e altro;
- valutazione del **trend di serie temporali** di eventi, con la possibilità di intraprendere azioni a carattere preventivo rispetto all'insorgere di eventi analoghi futuri laddove l'andamento della serie superasse una soglia di allerta;
- predizione dell'**evoluzione temporale** degli eventi, anche in termini di probabili date, fasce orarie e zone in cui gli eventi potrebbero ripetersi o in cui una catena di eventi potrebbe proseguire il suo corso;
- di suggerimento delle **risorse più idonee** da impiegare negli eventi, con criteri di valutazione basati sulla competenza e sull'esperienza delle risorse rispetto alla specifica tipologia di evento e sulla potenziale prontezza operativa di ciascuna risorsa.

1.1.2.1 DECISION SUPPORT SYSTEM

1.1.2.1.1 Architettura tecnica

Di seguito l'architettura tecnica dei componenti del DSS.

Resource Usage Manager

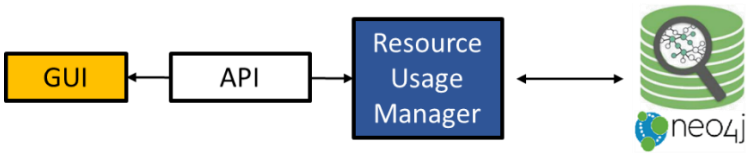


Figura 19 – DSS, utilizzo risorse

Il Resource Usage Manager espone un servizio interrogabile da interfaccia grafica o ad uso di altri moduli; il servizio, sottoponendo le informazioni su un evento in esame, reperisce le informazioni sulle risorse sul database a grafo Neo4J di Knowledge Graph. A seguito della valutazione delle informazioni reperite secondo criteri di prossimità, esperienza e utilità di capabilities delle risorse rispetto all'evento in esame, restituisce le risorse più idonee da inviare in campo.

Modulo Analisi Temporale

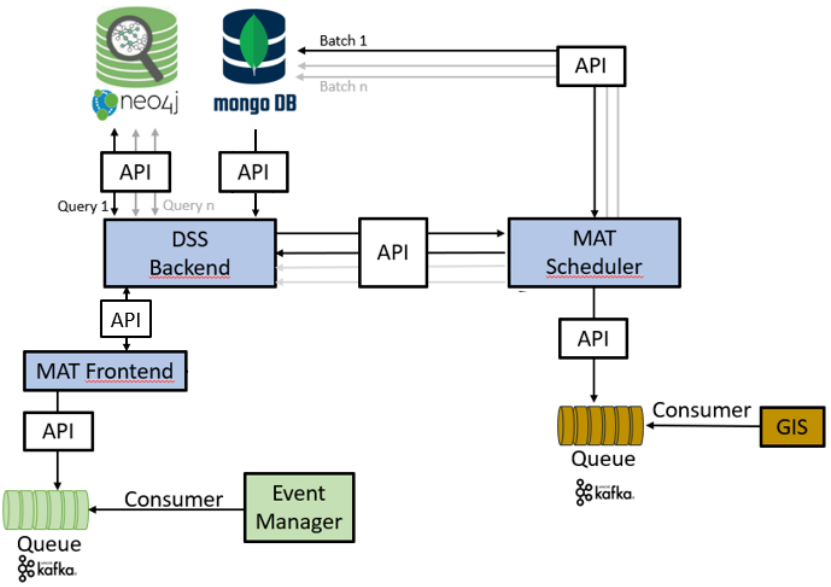


Figura 20 – DSS, integrazione modulo Event Manager

Al momento dell'installazione, vengono effettuate delle procedure di inizializzaizone per il recupero di alcune informazioni relative a categorizzazioni necessarie per i servizi del DSS e per la conseguente scrittura su Neo4j di nodi ancillari rappresentanti le dette informazioni.

L'utente può creare un job di analisi temporale attraverso la pagina frontend dedicata. Un payload con i parametri di ricerca selezionati viene inviato allo scheduler, il quale lancia periodicamente dei task di analisi costituenti il job complessivo; ciò viene effettuato inviando le informazioni relative alla ricerca degli eventi da effettuare al Backend del DSS. Questo genera un'opportuna query cypher e la sottopone al db a grafo Neo4j. I risultati della ricerca eventi del singolo task (costituenti un batch) vengono dunque reinviati allo scheduler e sottoposti ad analisi statistica. I risultati dell'analisi statistica del singolo task sul relativo batch e sul cumulativo dei dati vengono scritti su Mongo db.

In dettaglio, il componente scheduler utilizza al suo interno Celery, cui si sottopone in input la definizione dei task periodici costituenti un job attraverso uno strato api interno. Le attività e i parametri dei task vengono salvati su un db PostgresSQL, il quale implementa una logica "heartbeat" di controllo del rilancio delle attività confrontando costantemente il tempo corrente con il tempo di rilancio successivo previsto per ciascuna attività. Celery istanzia inoltre l'esecuzione concreta dei task al trigger. La gestione delle risorse rispetto ai task sottomessi viene gestita tramite coda Redis.

Ogni qual volta

- un'analisi, cioè un job, viene creata
- lo status di un'analisi cambia, in quanto soggetta ad aggiornamento su un nuovo batch, o in quanto giunta a fine vita o perché cancellata manualmente tramite frontend

avviene una segnalazione tramite la scrittura di un messaggio con standard opportuno su una coda Kafka su cui è presente un consumer di GIS. Quest'ultimo modulo è responsabile della visualizzazione delle heatmap, mostra un layer per ciascuna analisi.

Da frontend, per ciascun job che presenta la condizione *True* nel campo "stato allarme", cioè per ciascuna analisi il cui trend sia in aumento al di sopra della soglia selezionata al momento della creazione del job, è possibile inviare una notifica per la creazione di un evento opportuno per il presidio del territorio; di fatto viene scritto un messaggio con uno standard opportuno su una coda Kafka su cui è presente un consumer di Event Manager.

Data Filler

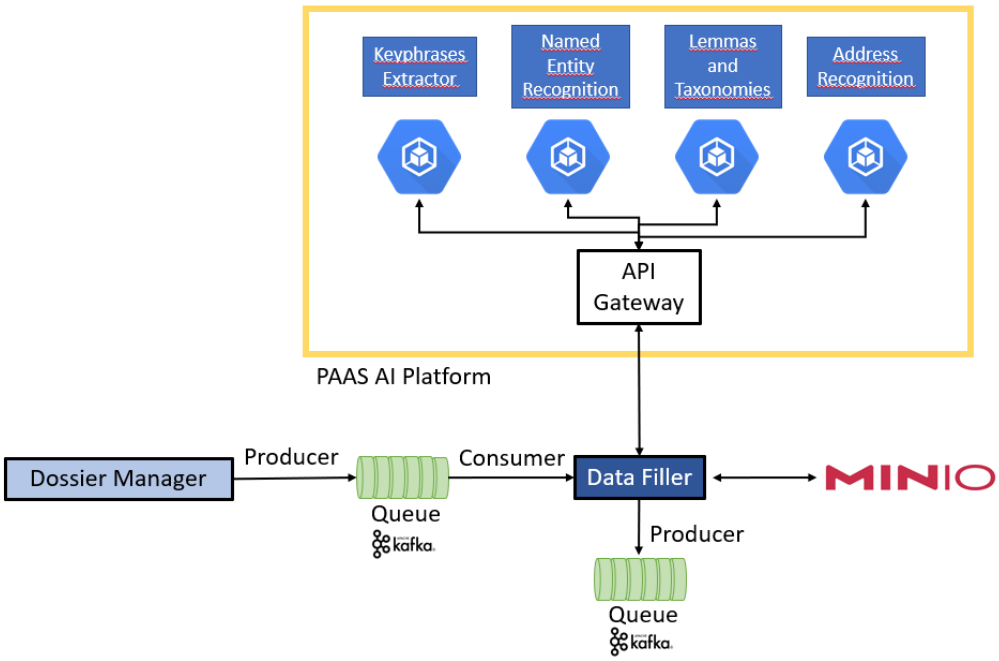


Figura 21 – DSS, data filter model

Quando viene allegato un file ad un dossier, il Dossier Manager scrive un messaggio con i riferimenti al file su un topic Kafka; questo scatena un'azione sul Data Filler. Quest'ultimo, dopo aver consumato il messaggio, recupera il file su PaaS Data Lake e lo sottopone ad un'analisi NLP mediante interrogazione dei modelli che risiedono su Paas AI Platform. I risultati dell'analisi sull'allegato vengono infine scritti su un topic Kafka dedicato, ad uso dei moduli che necessitano di tale informazione (es. Dossier Manager).

Simul (Contemporaneità Temporale)

Il servizio di contemporaneità temporale ricerca gli eventi correlati sia temporalmente, sia spazialmente rispetto all'evento relativo ad un dossier di riferimento; per far questo, sottopone un'opportuna query cypher a Neo4j, restituendo i risultati categorizzati mediante uno score di rilevanza.

Questo modulo è esposto sia come servizio API REST, sia come componente scatenato automaticamente alla creazione di un dossier (scrittura di Dossier Manager su coda Kafka). Esso cerca gli eventi correlati per prossimità spaziale e temporale, riportando se eventuali risultati sono già correlati (in quanto dossier padre, figli, etc.).

Figura 22 – DSS, esposizione API Simul

Nel caso del servizio scatenato, i risultati vengono scritti tramite messaggio su un topic dedicato su coda Kafka.

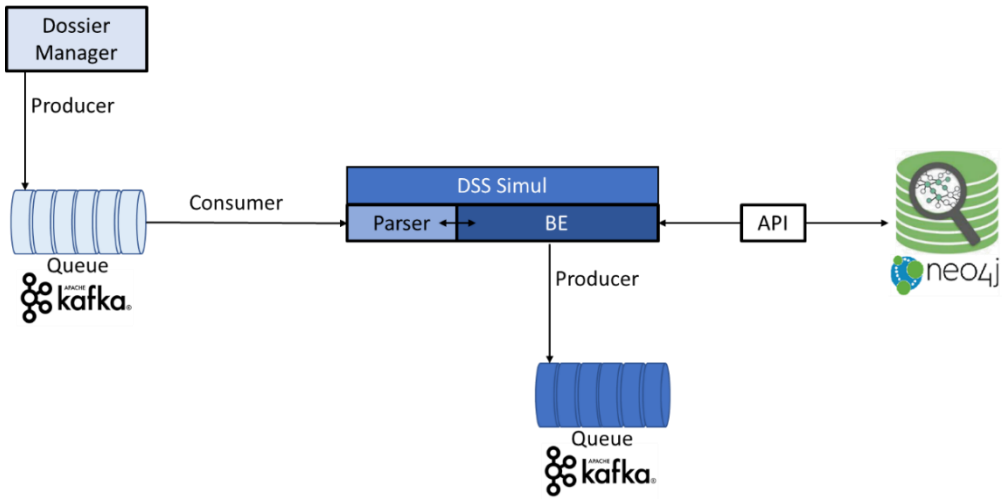


Figura 23 – DSS, utilizzo tpoic Kafka in Simul

Suggested Operation Resolver

Il servizio Suggested Operation Resolver permette di interpretare una query in linguaggio naturale per comandi da eseguire sulla GUI o per la ricerca di opportune risorse, restituendo poi i risultati della ricerca alla GUI.

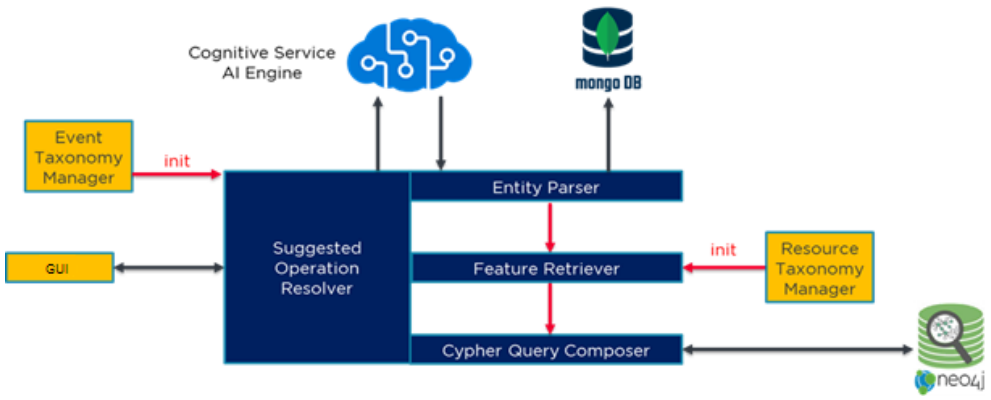


Figura 24 – DSS, suggested operation resolver

All'inizializzazione del servizio, il componente recupera tutte le suggested operation presenti sulle tassonomie di Event Manager e le analizza per estrarre intenti ed entità. In particolare, il servizio invia l'espressione in linguaggio naturale al motore AI, questo interpreta la richiesta ed estrae i parametri della ricerca; le entità estratte vengono confrontate per similarità con le label e le description del componente Taxonomies di Resource Manager, recuperando il taxonomyName della risorsa corrispondente o il name della feature corrispondente (property, capability). Il json response frutto di tale analisi viene salvato su MongoDB a vantaggio della rapidità di risoluzione di suggested operation sottomesse al sistema e già precedentemente memorizzate.

La GUI interroga un servizio REST API fornendo la posizione dell'evento e la suggested operation con la ricerca di risorse in linguaggio naturale. Se la suggested operation è stata già sottoposta ad analisi, viene restituito il risultato a seguito di opportuni processamenti delle informazioni memorizzate, altrimenti, il sistema effettua le analisi e le ricerche del caso sulla nuova suggested operation, memorizzandole infine su MongoDB.

Sul backend viene dunque composta dinamicamente una query cypher con i parametri estratti; la query viene inviata al db Neo4j del Knowledge Graph per la ricerca dei risultati; questi vengono infine restituiti sulla response dell'API.

Modulo Riconoscimento Contesto

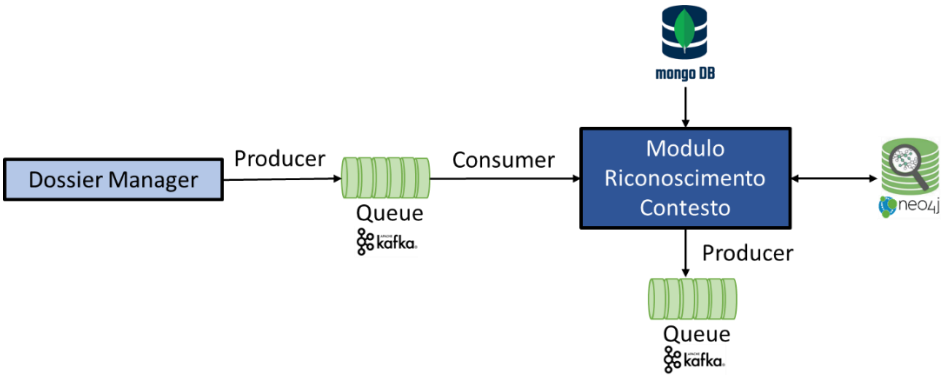


Figura 25 – DSS, modulo riconoscimento contesto

Il componente prevede le fasi di inizializzazione descritte di seguito.

All'inizializzazione del Modulo Riconoscimento Contesto, vengono caricate da MongoDB delle tassonomie di contesto ambientale (che prevedono per ciascun contesto ambientale una lista di termini ad esso correlati) da MongoDB.

Il servizio di riconoscimento ambientale può essere triggerato dal Dossier Manager tramite Kafka o eventualmente essere invocato come API REST. A seguito dell'analisi testuale di documenti ed allegati inerenti a un dossier relativo ad un evento, vengono estratte informazioni utili a definire la tipologia di ambientazione presso cui l'evento ha luogo.

Risulta utile, tra le altre cose, come ausilio per il Resource Usage Manager in merito alla ricerca delle risorse più idonee per intervenire su un evento, sulla quale i criteri in taluni casi non possono prescindere dal contesto ambientale.

1.1.2.2 EVENT MANAGER

1.1.2.2.1 Architettura tecnica

I componenti dell'Event Manager interagiscono tra loro e con i moduli esterni mediante messaggi asincroni su bus Kafka oppure tramite chiamate API REST. Ciascun microservizio di BE mantiene le informazioni su una propria base dati, un database non relazionale MongoDB.

L'operatore, il Call Taker, il Gestore Eventi, ovvero qualsiasi operatore/applicazione autorizzato ad inserire eventi, finalizza la scheda evento e la sottomette al sistema tramite interfaccia del Event Reporting (microservizio nell'ambito del Event Manager).

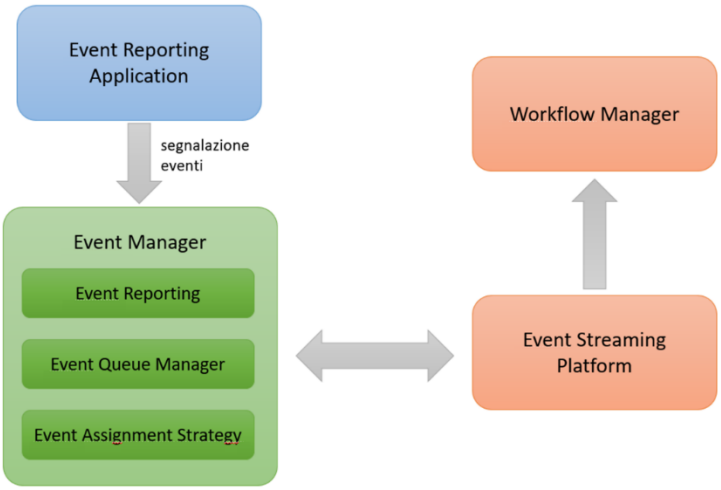


Figura 26 – Event Reporting

La scheda evento percorre una pipeline che interessa diversi elementi architetturali con lo scopo di validare ed inizializzare la gestione dell'evento. Alla fine della pipeline l'evento è inserito nella coda degli eventi che è gestita dal Event Queue Manager – EQM (microservizio nell'ambito del Event Manager).

Servizi

Event Taxonomy Manager (ETM)

- Classification Level: Gestisce i livelli di classificazione
- Classification-levels-hierarchy: Gestisce la gerarchia nei livelli di classificazione
- Custom-templates: Gestione dei template
- Event-template-icons: Inquiry delle icone disponibili
- Generated-from-custom-template: Inquiry custom template
- Import-export: Gestione import/export struttura della taxonomia
- Priorities: Gestione delle priorità
- Process-definiton-before-close: Ritorna la llsta di sottoprocessi disponibili per la fase before e on close del processo base
- Process-definition-event-management: Ritorna la lista degli id delle definizioni di processo presenti
- Process-definition-on-close: Ritorna la llsta di sottoprocessi disponibili per la fase on-close del processo base
- Questions: Gestione delle domande
- Suggested-operations: API per recuperare I permessi da precaricare su IAM

Event Queue Manager (EQM)

- Permissions: API per recuperare i permessi da precaricare su IAM
- Queue: Recupera informazioni sullo stato GLOBALE degli eventi accodati

- Events: Gestione degli eventi
- Resource allocation: Allocazione di risorse
- Resources: Gestione delle risorse dell'evento
- Allocated resources: Aggiunge un time reference alla risorsa di un evento

Event Assignment Strategy (EAS)
- Algorithms: Gestione degli algoritmi di assegnazione operatore
- Configuration-parameters: Gestione dei parametri di configurazione
- Import-export: Importa/esporta la struttura di EAS
- Pipelines: Gestione delle Pipelines
- Priority-weights: Gestione del peso delle priorità
- Status: Visualizza lo stato di EAS

Event Reporting Manager (ERM)
- Event-report-templates: Gestione dei templates
- Event-templates-icons: Gestione delle icone associate ai templates
- Event-reports: Gestion dei reports

Event Report Generator (ERG)
- v1-get-report: Esporta tutti i dati
- v2 -publish-report: Pubblicazione dei reports
- v2-publish-report-osint: Pubblicazione dei report OSINT

Event Manager Tools (EMT)
- Export: Esposrtazione del databasae in base ad alcuni filtri
- Import: Importazione dati da file
- Event Suggested Operations (ESO)
- Operations/Routines: Gestione delle operazioni

Infrastruttura

L'Event Manager è costituito da componenti suddivise in containers che utilizzano il database MongoDB per la persistenza dei dati. Inoltre, è previsto l'interfacciamento con il cluster Apache Kafka per quanto riguarda l'utilizzo delle code per scambio messaggi sia tra componenti interni che esterni.

1.1.2.3 Knowledge Graph

Il componente Knowledge Graph è votato alla memorizzazione su un database a grafo delle informazioni riguardanti gli eventi e le risorse in esame durante l'utilizzo della soluzione. Le informazioni vengono memorizzate in termini di nodi e relazioni di tipi specifici, che possono contenere delle proprietà (attributi) quali dati secondari.

Il Knowledge Graph contiene non soltanto i dati grezzi relativi ad eventi e risorse, ma un insieme di dati arricchiti a seguito di analisi da parte di altri componenti della soluzione; a titolo esemplificativo ma non esaustivo: correlazioni derivanti da analisi testuale, dal riscontro della presenza di entità o tematiche in comune a più nodi, valutazioni di similarità.

La particolare struttura dati di siffatti database, con un'architettura opportunamente ponderata, permette un netto miglioramento nelle performance ed una rappresentazione delle relazioni tra i dati più efficiente, in special modo rispetto a interrogazioni che coinvolgono grandi moli di dati.

Una rappresentazione a grafo dei dati ben si presta all'applicazione di algoritmi di graph data science, ad esempio, per analisi su centralità, community detection, similarità, shortest path, predizione di relazioni e altro.

Un consumer acquisisce i messaggi su una coda, in particolare su specifici topic foraggiati rispettivamente dai producer. Ciascun messaggio consumato viene sottoposto ad un'operazione di parsing al fine di estrarre i dati necessari e comporre correttamente una query cypher atta a scrivere sul db a grafo Neo4J o a modificare taluni nodi o relazioni all'interno dello stesso.

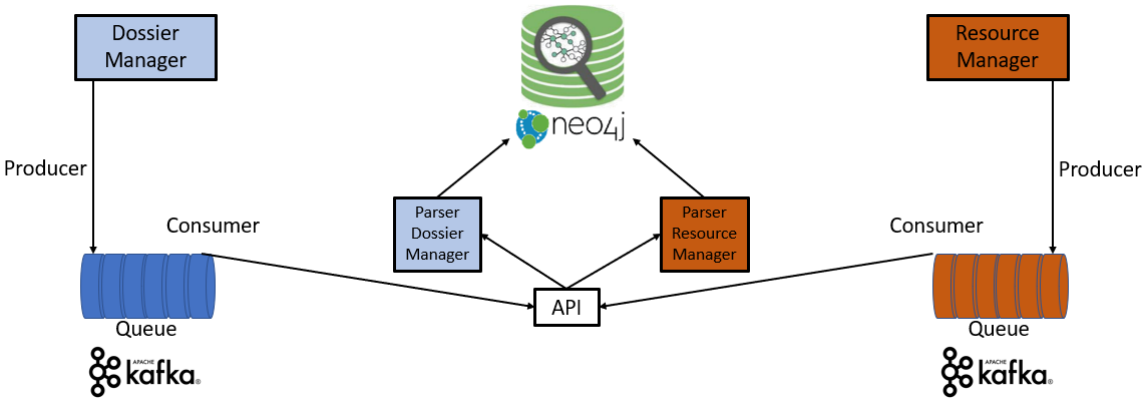


Figura 27 – Knowledge Graph, estrazione dati

Nell'oggetto header all'interno del messaggio json consumato, sono presenti i riferimenti di provenienza (producer) del messaggio, mentre nell'oggetto body del messaggio, il campo type indica se il messaggio deve produrre un'operazione di scrittura o di aggiornamento. In tal modo, dopo aver consumato un messaggio dalla coda, viene invocata un'API che provvede ad effettuare il parsing dei campi relativi al messaggio, a seconda che lo schema sia quello di un messaggio dal producer, per costruire l'apposita query cypher e sottoporla al db a grafo Neo4j.

Il Knowledge Graph permette di effettuare le tipiche operazioni CRUD, ovvero di creazione, lettura, aggiornamento e cancellazione sugli elementi del database a grafo quali nodi, relazioni, attributi degli uni e delle altre.

1.1.2.3.1 Servizi

RUM: servizio per l'individuazione di risorse con le skill più appropriate per la gestione e l'attuazione delle attività relative ad un dato evento.

MAT Analyzer: servizio per la trend analysis su un set di eventi individuato da una lista di parametri inseriti nel payload.

MAT Scheduler: servizio per la schedulazione di job di trend analysis secondo frequenza parametrica, con monitoraggio di una soglia di allerta, anch'essa parametrica.

Data Filler: servizio per l'arricchimento di un dossier mediante analisi NLP dei suoi allegati testuali.

Simul: servizio per la ricerca di dossier correlati spazialmente e temporalmente

MRC: servizio per la valutazione del contesto ambientale di un evento tramite analisi NLP di un dossier

1.1.2.3.2 Infrastruttura

Il Decision Support System è costituito da componenti containerizzate sottoposte all'azione di un orchestratore.
A ciò si aggiungono il database MongoDB e il data lake MinIO. È previsto anche un interfacciamento con un cluster Apache Kafka e con il database a grafo Neo4J (del Knowledge Graph).

1.1.3 Application Platform

1.1.3.1 LOG & AUDIT

1.1.3.1.1 Architettura tecnica

La tecnologia utilizzata è quella dello stack Elastic e comprende:

- Database NOSQL Elastic che è possibile demandare come servizio PAAS esterno.
- Console Kibana per la visualizzazione dei dati all'interno di Elastic.
- Agent sugli host dove vengono erogati i servizi.

1.1.3.1.2 Servizi

Gli Agent all'interno del cluster sono abilitati a raccogliere:

- Metriche degli host, deployment e tutti gli altri componenti kubernetes
- Log dagli host e da tutti i containers.
- Dati di APM da tutti i microservices
- Audit Log emessi dai gateway
- Dati di usabilità (RUM – Real user Monitoring) , inviati dal frontend.

1.1.4 Process Platform

1.1.4.1 Workflow Manager

1.1.4.1.1 Architettura Tecnica

L'architettura del Workflow Manager è costituita dai seguenti componenti secondo il seguente schema:

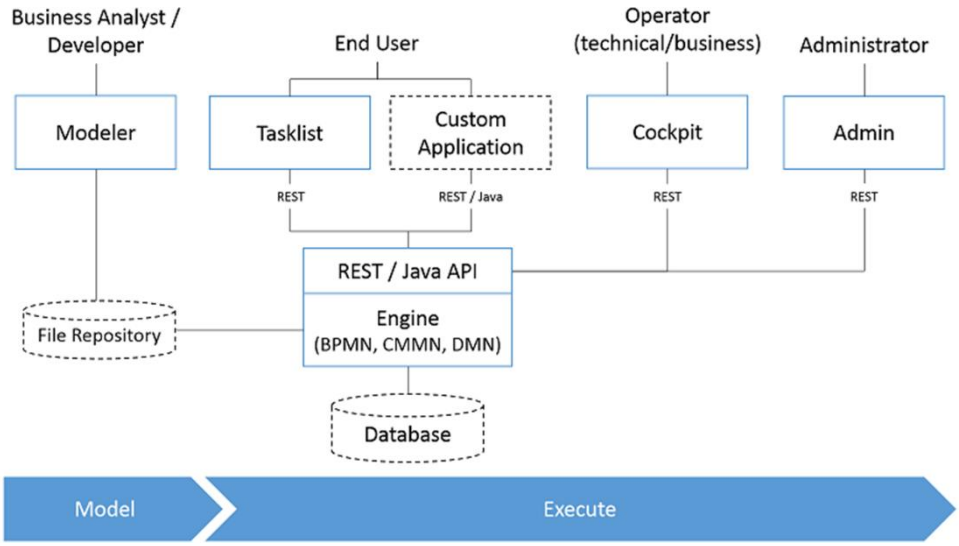


Figura 28 - WorkflowManager, ciclo di vita di un processo

Workflow Manager (WFM)

WFM è il micro-servizio Workflow Manager che consiste in una serie di definizioni di processi definiti all'interno del motore BPMN e relative classi a loro supporto.

Il micro-servizio Workflow Manager non espone interfacce proprie all'esterno ma ha il compito di integrarsi con la gestione dei processi dell'engine BPMN. Le componenti del Workflow Manager interagiscono tra loro e con i moduli esterni mediante messaggi asincroni su bus oppure tramite chiamate API REST.

Il micro-servizio Workflow Manager non ha una vera base dati dedicata ma si basa su quella integrata nel motore di gestione processi. Durante il funzionamento del processo vengono effettuate chiamate REST verso altri moduli, come ad esempio al Dossier Manager per la creazione di un dossier,

WFM Middleware

il middleware WFM ha il compito di elaborare il messaggio associato all' evento pubblicato sulle code e richiamare le REST API dell'engine BPMN per lanciare il processo richiesto a cui viene allegato l'event sheet contenente tutti i dati dell'evento.

Il micro-servizio WFM Middleware non espone servizi verso gli altri componenti, in quanto ha come scopo quello di estrarre i messaggi dalla coda e lanciare il processo ad essi corrispondente.

WFM Workers

Il microservizio WFM Workers sfrutta gli External Task di Camunda per ottenere una modularizzazione ancora più spinta, consentendo di eseguire i task richiesti dal processo da componenti esterne.

Ciò consente di realizzare un ulteriore disaccoppiamento tra mappa di processo e microservizi che eseguono le operazioni.

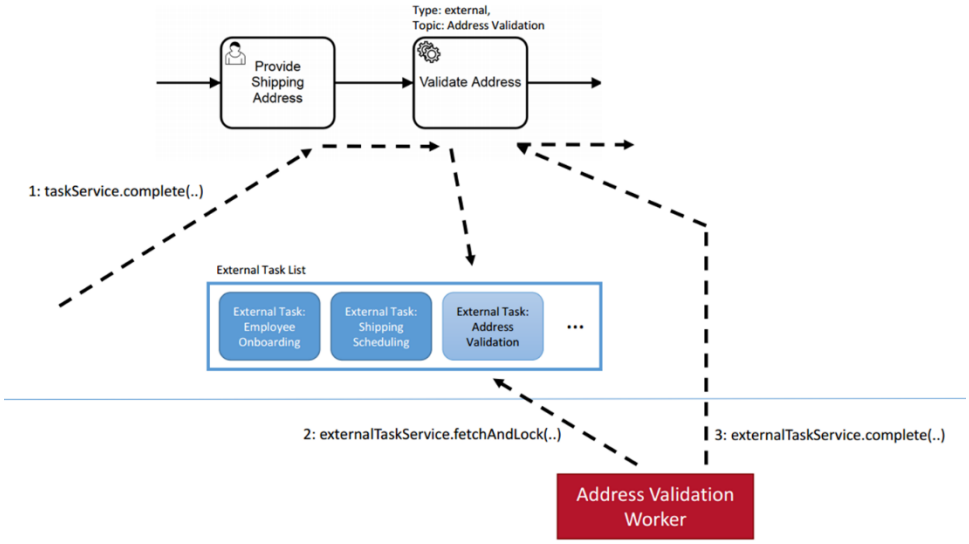


Figura 29 - WorkflowManager, WFM Worker

Sostanzialmente questi external tasks sono task che vengono lanciati al di fuori della piattaforma e l'engine verifica la loro esecuzione e terminazione tramite una procedura di polling, riportando il risultato all'interno.

Questa strategia consente di alleggerire in motore di processi principale e permette una ulteriore modularizzazione e scalabilità del sistema.

Camunda Modeler

Il Modeler è un'applicazione desktop che consente di modellare i processi BPMN direttamente sul proprio sistema locale e poi di caricarli sul motore di processo.

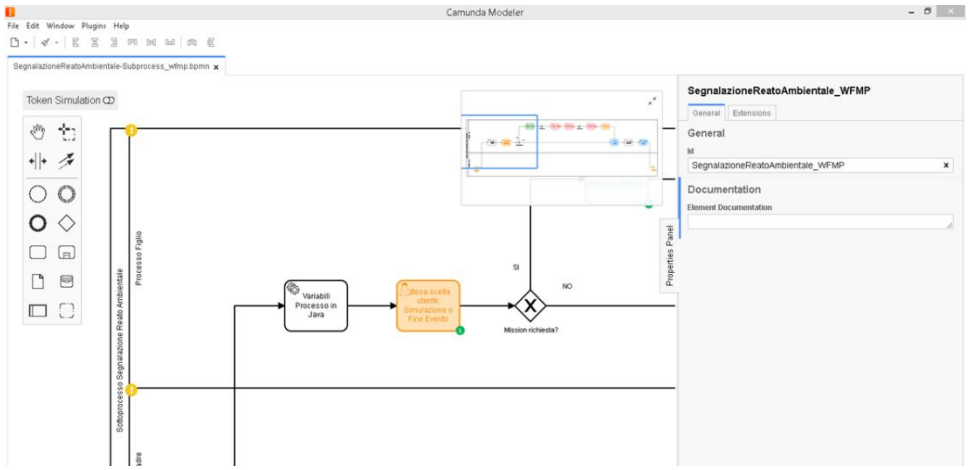


Figura 30 - WorkflowManager, modellazione processi con Camunda Modeler

I processi, essendo in notazione BPMN standard sotto forma di file XML, possono anche essere importati e modificati a piacimento.

L'applicativo mette a disposizione tutta una serie di elementi utilizzabili per costruire il processo che meglio si adatta alle necessità di business.

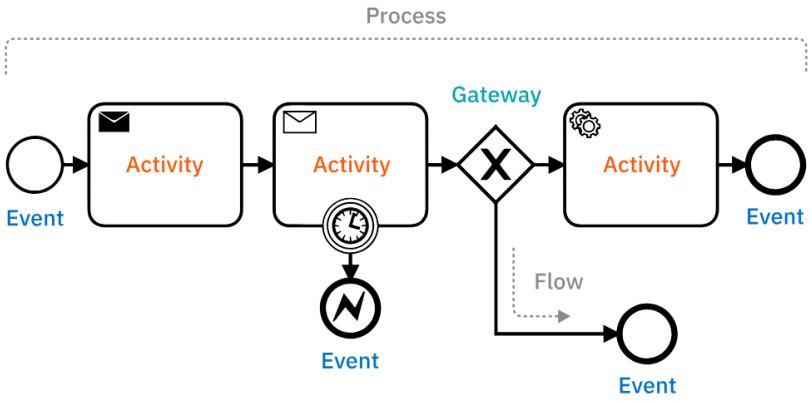


Figura 31 - WorkflowManager, processo BPMN

Gli elementi principali sono:

- Eventi: sono cose che accadono rispetto al processo, quali ad esempio l'inizio oppure la sua terminazione.

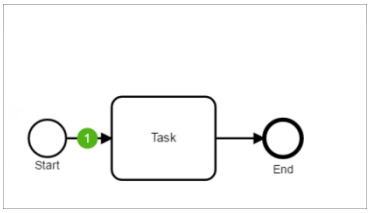


Figura 32 - WorkflowManager, processo BPMN

- Task: sono attività che un utente deve eseguire oppure attività automatiche che possono coinvolgere, ad esempio, chiamate a webservices.

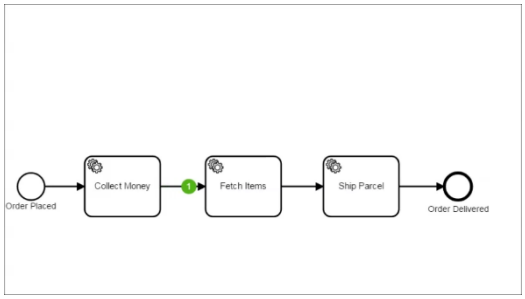


Figura 33 - WorkflowManager, esempio di processo BPMN

- Gateways: sono elementi in cui viene presa una decisione che fa, ad esempio, scegliere di eseguire un ramo di un processo rispetto ad un altro.

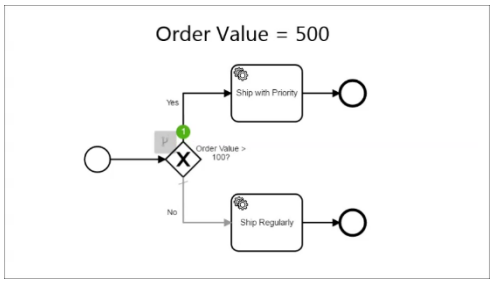


Figura 34 – WorkflowManager, gateway BPMN

Sono comunque possibili gateway paralleli che muovono il processo verso due o più attività in contemporanea.

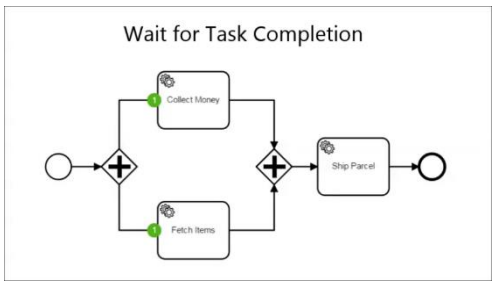


Figura 35 – WorkflowManager, gateway paralleli BPMN

Possono anche utilizzare variabili dell'istanza di processo e sfruttare espressioni per accedere a variabili e calcolarne il valore.

- Sotto-processi: sono processi specifici che possono essere richiamati dal processo principale

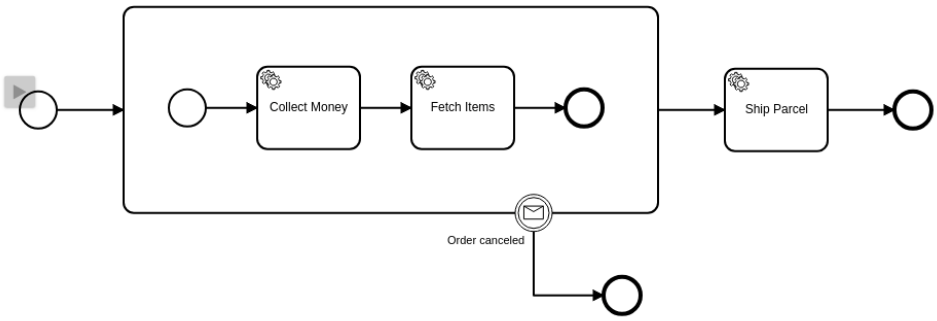


Figura 36 – WorkflowManager, sottoprocesso BPMN

Sul pannello dell'applicazione è possibile comporre quindi una sequenza di blocchi che rappresentano le attività da portare a termine da parte del processo:

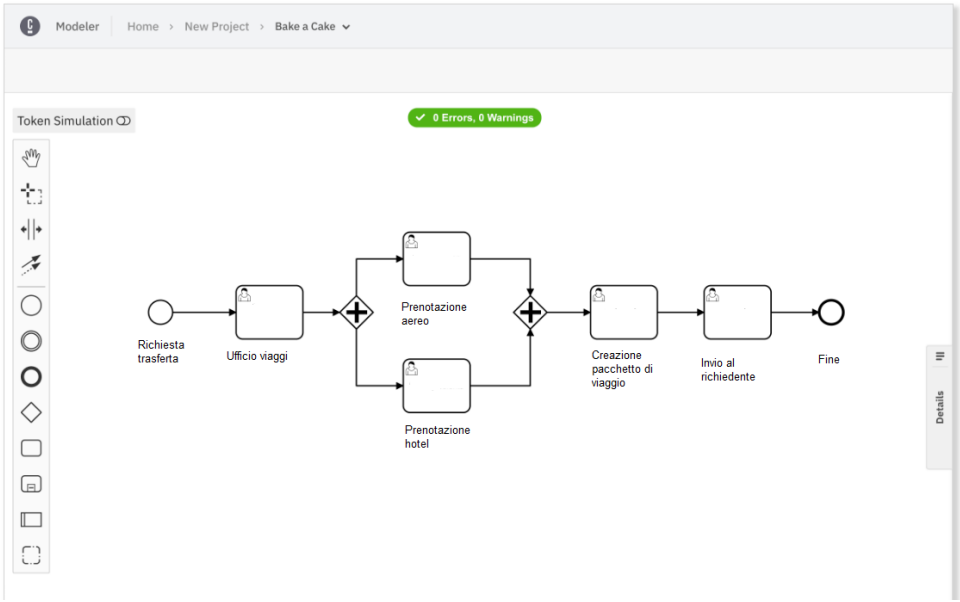


Figura 37 – WorkflowManager, composizione di un processo

Il processo può essere validato per verificare che gli elementi non contengano errori e, in caso positivo, caricato direttamente sulla piattaforma ed eseguito.

Ovviamente gli elementi possono ad esempio avere parti di codice che vengono eseguite, connettori HTTP che chiamano servizi esterni, timer che scatenano altri eventi se il task non viene portato a termine in un certo lasso di tempo prestabilito.

Camunda Cockpit

Il Camunda Cockpit è una web application che consente di monitorare le istanze dei processi e gestire eventuali problemi. Essa consente di accedere a tutti i processi caricati, eseguire ricerche tra le istanze attive e quelle terminate ed eseguire operazioni su di esse.

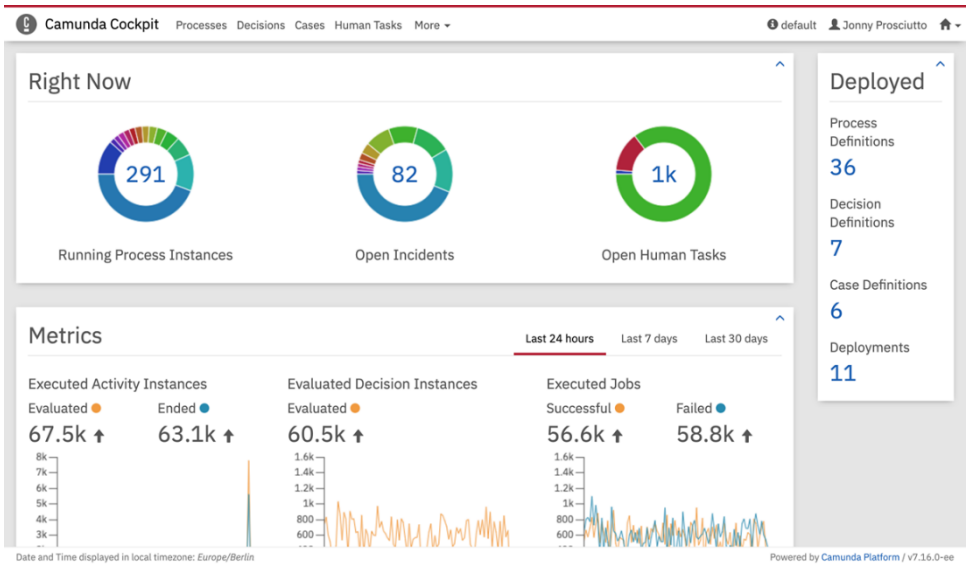


Figura 38 – WorkflowManager, schermata Camunda Cockpit

Sono messi a disposizione anche alcune metriche che possono aiutare a capire lo "stato di salute" della piattaforma, in modo da poter mettere in campo gli opportuni correttivi in caso di problemi.

Camunda Cockpit mette a disposizione anche un pannello per tenere sotto controllo tutti i cosiddetti *deployments* ovvero tutte le risorse quali BPMN, classi e immagini che sono state caricate sulla piattaforma.

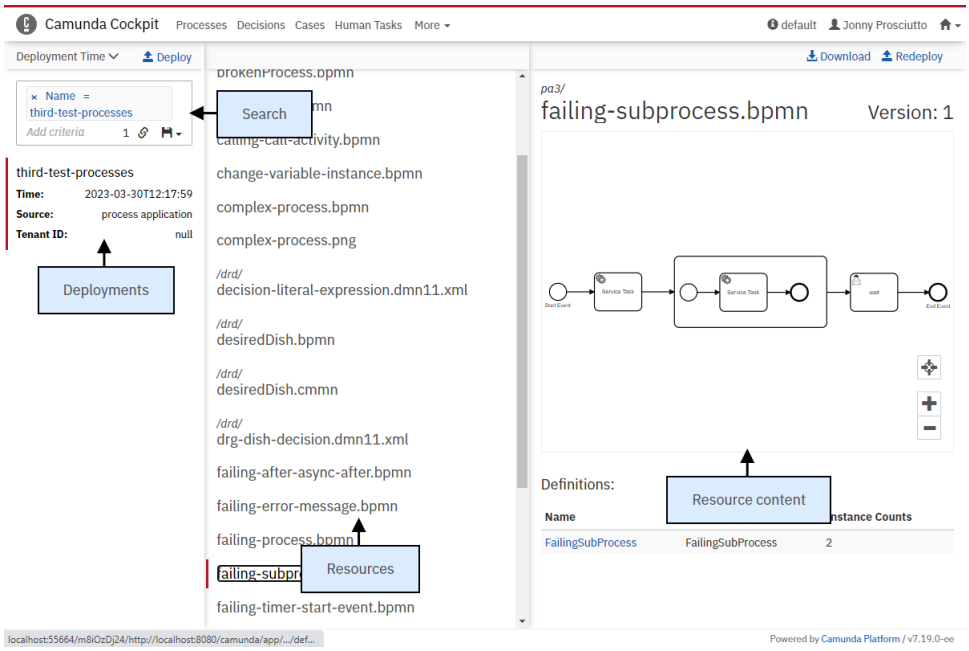


Figura 39 – WorkflowManager, sezioni della schermata di Camunda Cockpit

La pagina elenca tutti i deployments presenti, permette di ricercarli, visualizza il contenuto delle risorse e mostra la rappresentazione grafica dei BPMN come compare anche sul Modeler.

Camunda mette inoltre a disposizione plugin sia per la parte Cockpit che per la parte Modeler:

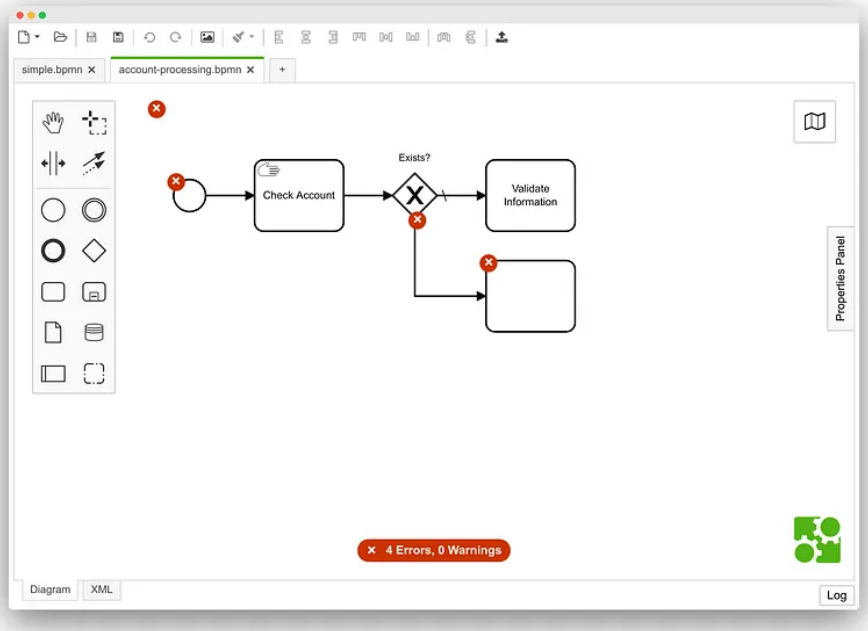


Figura 40 – WorkflowManager plugin Camunda per il modeler

API e Eventi

I componenti del Workflow Manager interagiscono tra loro e con i moduli esterni mediante messaggi asincroni su bus Kafka oppure tramite chiamate API REST.

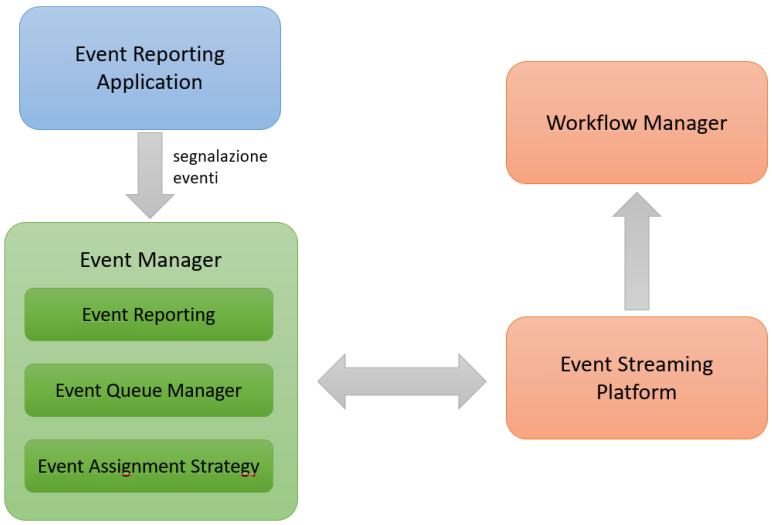


Figura 41 – Schema Workflow Manager

Il microservizio Workflow Manager non ha una vera base dati dedicata ma si basa su quella del motore di gestione processi Camunda.

Infatti, esso richiama le API REST di Camunda per eseguire le operazioni relative ai processi.

L'elenco di queste chiamate viene messo a disposizione attraverso un endpoint Swagger che permette di accedere ad un'interfaccia per eseguire chiamate verso l'engine.

Da quest'interfaccia si hanno a disposizione per ogni API REST:

- Una descrizione
- URL
- I parametri utilizzabili nonché quelli obbligatori o meno
- Una descrizione dettagliata della response e dei suoi contenuti
- I possibili codici di risposta
- Un breve esempio di request e response

1.1.4.1.2 Infrastuttura

Il Workflow Manager prevede componenti ospitati da un Container Platform.

I POD previsti per la soluzione del modulo sono allocati sull'infrastruttura CaaS del PSN, in questo modo è gestito, ad esempio, l'autoscaling.

Il componente Camunda BPM Engine è disponibile sia come immagine Docker che come Helm Chart per l'installazione su Kubernetes.

A livello di storage tutte le informazioni sono contenute all'interno del database schema di Camunda.

1.1.4.2 Rule Manager

1.1.4.2.1 Architettura tecnica

Di seguito uno schema di architettura tecnica con i componenti principali del modulo.

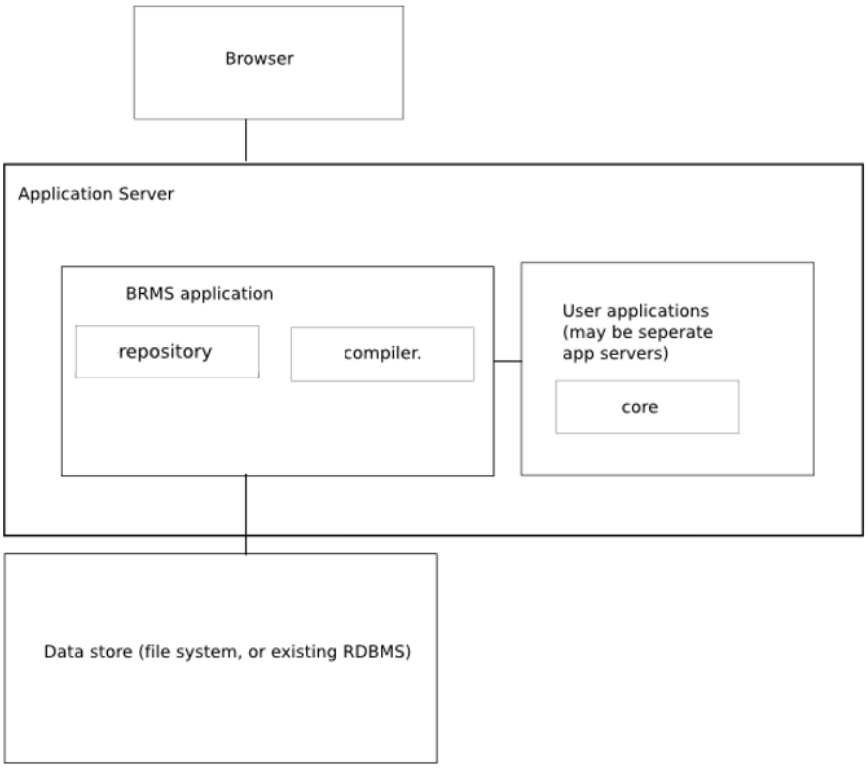


Figura 42 – Rule manager, architettura tecnica

L'architettura del modulo utilizza due application server: uno che fornisce servizi per la user application utilizzando un business rule engine (core), l'altro con un sistema BRMS (business rule management system) responsabile dello storing, del versioning e della validazione delle regole di business. I componenti core, compiler e repository sono costituiti da librerie responsabili dell'applicazione, dell'interpretazione e della traduzione delle rules.

Servizi

Le principali funzionalità messe a disposizione da questo modulo sono:

- capacità di valutare le tabelle decisionali DMN;
- definizione input/output e condizioni delle decisioni;

- motore di linguaggio di espressione FEEL;
- descrizione di tipi di dati di input/output;
- creazione di decision test come test unitari;
- possibilità di essere utilizzato come libreria incorporata in un'applicazione;
- capacità d'implementazione dello standard OMG DMN;
- utilizzo DMN in ambienti Big Data con Apache Spark (dmn4spark);
- possibilità di archiviazione flusso di regole utilizzando JSON;
- disponibilità editor basato su browser per la definizione delle rules.

1.1.5 Data Platform

1.1.5.1 PaaS Data Lake

1.1.5.1.1 Architettura tecnica

Queste caratteristiche sono rese disponibili su un sistema orizzontalmente scalabile progettato per gestire grandi quantità di dati e throughput elevati sia in lettura che in scrittura.

Il servizio fornisce una piattaforma di archiviazione pronta all'uso, con una soluzione certificata gestita e manutenuta dal fornitore, garantendo una riduzione del costo di archiviazione in relazione al volume di dati gestiti.

Inoltre, poiché questo servizio sarà erogato da infrastrutture di PSN ospitate nei propri Data Center, il MASE avrà la possibilità di archiviare dati Critici e Strategici, oltre a quelli Ordinari, nel rispetto dei requisiti imposti dall'Agenzia di Cybersecurity Nazionale.

Lo strato applicativo Data Lake sarà basato su un object storage distribuito. Tale strato utilizzerà un layer sottostante di block storage di tipo bare metal o software defined.

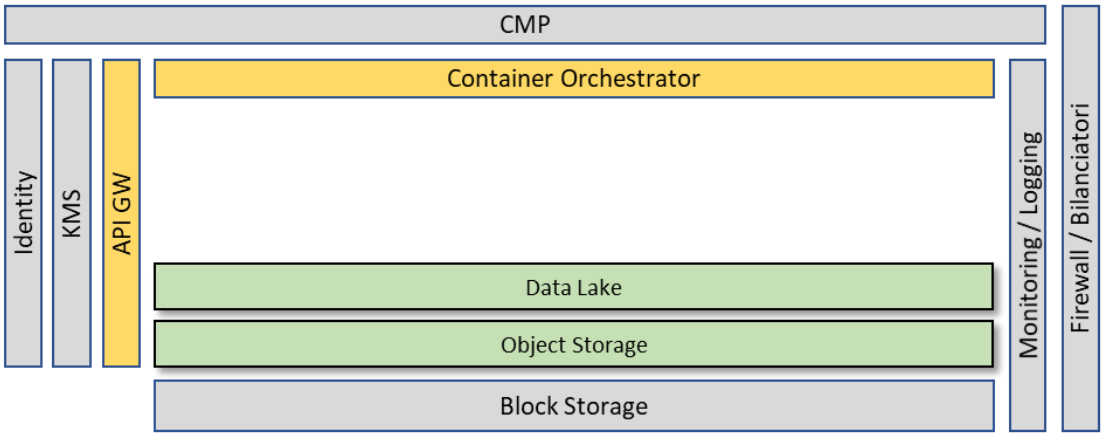


Figura 43 - Il servizio PaaS Data Lake

L'object storage sarà accessibile dai servizi che lo utilizzeranno tramite l'utilizzo del protocollo S3. Questo, essendo uno standard-de-facto, garantirà la compatibilità con l'engine Apache Spark e con la maggior parte dei tool dell'ecosistema Big Data.

L'object storage garantirà scalabilità orizzontale adeguata ai workload di tipo Big Data tramite una architettura distribuita basata sulla distribuzione dei blocchi e la replicazione degli stessi su diversi nodi dello storage cluster al fine di garantire sia un elevato throughput in termini di I/O che una elevata disponibilità e resilienza in caso di indisponibilità di una certa percentuale di nodi del cluster. Al fine di ottimizzare l'efficienza di utilizzo del raw block storage la replicazione dei blocchi non sarà reale (x2, x3, etc.) ma implementata tramite erasure coding.

Per soddisfare i requisiti di sicurezza più stringenti, verrà implementata la crittografia dei dati tramite chiavi archiviate su dispositivi HSM. Questo verrà reso possibile tramite interfacciamento con il modulo KMS comune a tutti i servizi del PaaS.

Modello multitenancy

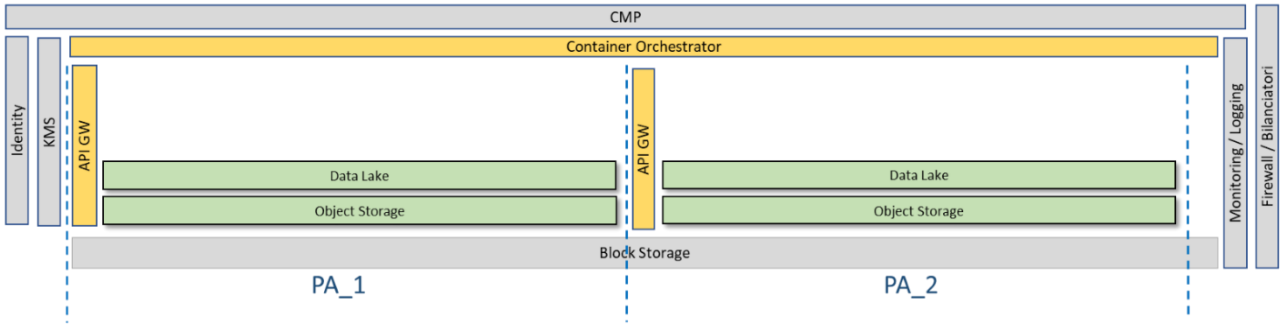


Figura 44 - Multitenancy per il servizio PaaS Data Lake

Ogni cliente sarà rappresentato da un "tenant" che avrà un suo spazio allocato (quota) e potrà utilizzarlo per creare i suoi bucket e dentro questi i suoi files/folder (n.b. in realtà negli object storage il concetto di folder è "virtuale" in quanto il path altro non è che parte integrante del nome file, considerato l'identificativo -chiave- dell'oggetto secondo un modello di tipo chiave-valore dove il contenuto binario è considerato il "valore").

I dati di un tenant saranno completamente isolati da quelli degli altri tenant.

Identity

La componente Identity, esterna al modulo e trasversale rispetto a tutti i servizi PaaS, si occuperà di verificare le credenziali ed emettere i token di autenticazione che verranno validati dai servizi soggetti a richieste di interazione.

Preferred Technology Platform: Minio

1.1.5.1.2 Servizi

I servizi offerti dal moulo Data Lake sono i seguenti.

Gestione Bucket: Un bucket è un contenitore di dati del Data Lake al quale vengono applicate policy omogenee (es: accesso, retention, replicazione, crittografia, etc.). All'interno del bucket è possibile

creare oggetti ai quali poter accedere con una semantica di tipo filesystem (struttura folders ad albero e files).

La gestione dei Bucket include, a titolo esemplificativo e non esaustivo, la gestione delle quete, la gestione del versioning dei files, la gestione del locking dei files nonché la retention policy di quest'ultimi.

Gestione Folders e File: All'interno dei bucket del nostro Data Lake sono memorizzati oggetti identificati all'interno del sistema da una semantica di ti object storage (key-value). Il sistema consente tuttavia di poter manipolare logicamente questi oggetti utilizzando una semantica di tipo filesystem, ovvero operare su folders (con una struttura ad albero) e files.

Monitoring: Oltre alla gestione ed alla modifica degli aspetti visti nei punti precedenti, è possibile monitorare e modificare i vari aspetti (policy, percentuali utilizzo della memoria, fattore di replica ecc) tramite il servizio dedicato.

1.1.5.2 Data Governance

1.1.5.2.1 Architettura tecnica

L'architettura è composta da diversi moduli di seguito descritti:

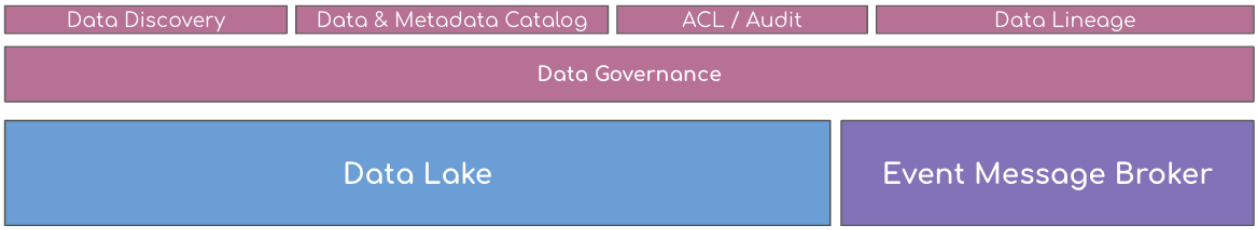


Figura 45 – Moduli del servizio Data Governance

Data & Metadata Catalog

Essendo possibile gestire dati in ambito Big Data di varia tipologia (variety: structured, semi-structured, unstructured) essi sono memorizzati nel Data Lake sotto forma di folder e files in vario formato archiviati su un filesystem distribuito. Al fine di poter manipolare in modo semplice queste folder tramite processing engine o query engine (es: tramite SQL) è necessario fornire a questi ultimi delle informazioni aggiuntive. Ad esempio, immaginiamo di avere una ipotetica folder /path/to/customers con dentro una serie di files in formato CSV che usano il carattere "|" come separatore. Volendo eseguire una ipotetica query "SELECT * from customers" avremo bisogno di conoscere come minimo:

- il percorso dove sono memorizzati i files della tabella "customers"
- il formato dei files dentro il folder di cui sopra (es: CSV)
- il separatore di campo utilizzato (es: "|")

- il separatore di riga utilizzato (es: "\n")
- numero e nome dei campi presenti nei files (es: nome, cognome, indirizzo, email, etc.)
- etc.

Queste informazioni sono memorizzate su un "metastore" (archivio di metadati). Quello più utilizzato in ambiente Big Data è HCatalog (https://cwiki.apache.org/confluence/display/Hive/HCatalog+UsingHCat) componente di Hive. Una volta memorizzate queste informazioni per tutti i percorsi del FileSystem (o object storage) che compone il Data Lake è possibile gestire i dataset in maniera semplice e consistente.

Data Search & Discovery

Per "Data Discovery" si intende un processo di esplorazione automatico dei dataset del Data Lake alla ricerca di (meta)dati che possano arricchire o approfondire la conoscenza delle informazioni disponibili. Nel contesto SIM, infatti sono attese numerose fonti di alimentazione e procedure ETL e i dati possono crescere velocemente. Infatti i tool di data discovery scansionano in modo sistematico tutti i contenuti del Data Lake e tramite sistemi basati su regole (es: regex) o intelligenza (es: classificazione tramite algoritmi di machine learning) memorizzando per ciascun dataset una serie di metadati. Al fine di creare un catalogo di metadati quanto più affidabile e completo possibile, si prevede di eseguire una scansione automatica da un processo di approvazione/etichettatura di tipo "supervisionato" da un operatore umano. Una volta popolato il database dei metadati sarà possibile utilizzare questo per poter effettuare ricerche, ad esempio per "tipologia" di dato. Ecco spiegato il concetto di "search & discovery". Utilizzando questo approccio, le funzionalità in questione verranno svolte dal tool "Data Hub".

Data Lineage

Il data lineage è l'operazione che consente di tracciare l'intero ciclo di vita dei dati, dalla loro origine a tutte le trasformazioni che intervengono durante l'operatività. Esistono varie definizioni, ma nella sostanza è sufficiente sapere che il data lineage si identifica con la derivazione e la tracciabilità dei dati nel tempo. Ciò consentirà di avere a disposizione tutte le informazioni necessarie per identificarli e gestirli nel modo migliore. La presenza dei metadati contribuisce a creare una cultura del dato, essenziale per valorizzare il patrimonio informativo disponibile.

I due elementi principali del data lineage sono pertanto l'origine e il cambiamento dei dati, utile, quest'ultimo, a descrivere come, dove e perché i dati sono stati soggetti a determinate trasformazioni. Tecnicamente, stiamo parlando di un database che contiene record di tracciatura delle operazioni intervenute sul Data Lake implementato principalmente tramite dei "filtri" che si interpongono tra gli engine utilizzati per l'ingestion/ETL dei dati che tracciano automaticamente (metodo PULL) le operazioni sul database prima descritto. Laddove non sarà possibile "intercettare" automaticamente queste operazioni (es: utilizzo di tools non standard o non supportati) è possibile alimentare manualmente il database invocando delle API, metodo PUSH (anche questa funzionalità verrà svolta dal tool "Data Hub").

ACL/Audit

Una corretta implementazione di processi di Data Governance comprende solitamente anche una robusta gestione dei permessi granulari di accesso ai dati, e di auditing rispetto all'utilizzo degli stessi (saper rispondere in qualsiasi momento alla domanda "chi ha acceduto quali dati e quando?"). Essendo i dati su Big Data essenzialmente folder e files, la componente fondamentale da "proteggere" (ACL) e "monitorare" (audit) è il Data Lake, ovvero il FS/object storage. Tali funzionalità verranno implementate a livello di MinIO. E in ogni caso, qualora fosse necessario inibire anche l'accesso ai metadati (es: nomi delle tabelle "virtuali" del metastore o struttura delle stesse), le ACL verranno implementate a livello HCatalog.

DataHub è uno strumento che estende il concetto di catalogo dati offrendo funzioni di data discovery, data observability e data governance.

Grazie a questo strumento si può verificare i cambiamenti effettuati sui dei dati nel tempo all'interno del catalogo distinguendo le varie sorgenti che hanno popolato il Data Lake, la tipologia di dato inserito (dati anagrafici, dati finanziari, etc) e identificare quelli che sono sensibili a specifiche leggi o procedure di compliance interni o esterni all'organizzazione.

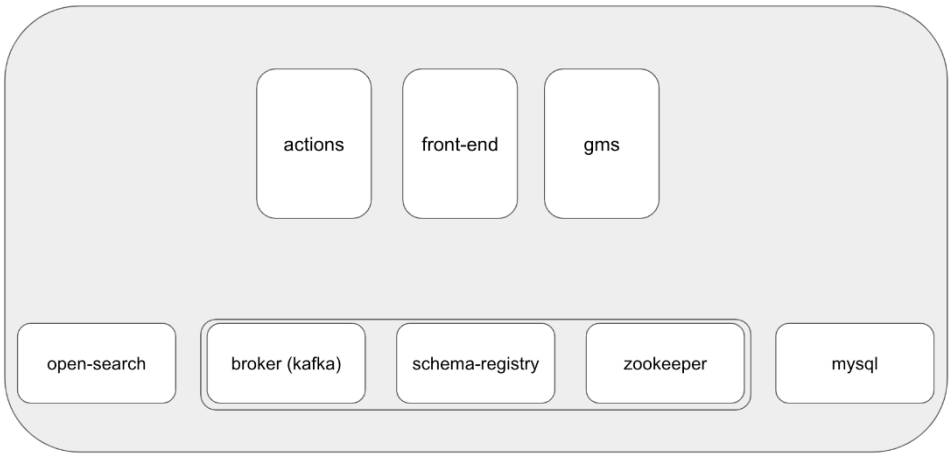


Figura 46 – Software coinvolti all'interno di DataHub

Nello specifico, la struttura è la seguente:

- Gms, DataHub Metadata Service, che contiene una serie di servizi per:
  - Interfaccia GraphQL API per fetching e gestione dei metadati (graph)
  - Interfaccia Rest per gestire i metadati
- Front-end, Interfaccia visuale
- Actions, modulo che comprende dei componenti per filtrare, trasformare e gestire i dati in real-time

Oltre ciò, DataHub si basa sulle seguenti componenti opensource:

- Apache Kafka + Apache Zookeeper + Schema Registry, per gestione degli eventi da registrare
- Opensearch, come supporto alla ricerca

MySQL, relazionale di utilità

1.1.5.2.2 Servizi

I principali servizi offerti dal modulo sono:

- Ricerca sull'intero set di metadati
- Visualizzazione del ciclo di vita dei dati
- Identificazione delle dipendenze dei dati ed eventuali incompatibilità

1.1.5.3 PAAS BATCH/REAL TIME PROCESSING

1.1.5.3.1 Architettura tecnica

Il servizio Batch/Real time Processing fornisce funzionalità per il processing dei dati archiviati sul Data Lake. Attraverso l'utilizzo dell'engine open source Apache Spark è possibile eseguire pipelines di tipo DAG (Directed Acyclic Graph) distribuendo il carico di lavoro su più nodi worker, approccio che rende la soluzione adatta all'analisi di qualsiasi volume di dati tramite scalabilità orizzontale. Il motore Apache Spark messo a disposizione consente l'esecuzione di workload sia di tipo batch, attraverso Dataframe API, che near-realtime mediante Structured Streaming API. Il tutto con la possibilità di utilizzare uno dei linguaggi supportati: Python, R, Java, Scala, SQL e C#.

Apache Spark è un framework di elaborazione parallela open source che supporta l'elaborazione in memoria per migliorare le prestazioni delle applicazioni che analizzano Big Data. Le soluzioni Big Data sono progettate per gestire i dati troppo grandi o complessi per i database tradizionali. Spark elabora grandi quantità di dati in memoria, molto più veloci rispetto alle alternative basate su disco.

Apache Spark include tre componenti principali: driver, executors e gestione cluster. Le applicazioni Spark vengono eseguite come set indipendenti di processi in un cluster, coordinata dal programma driver.

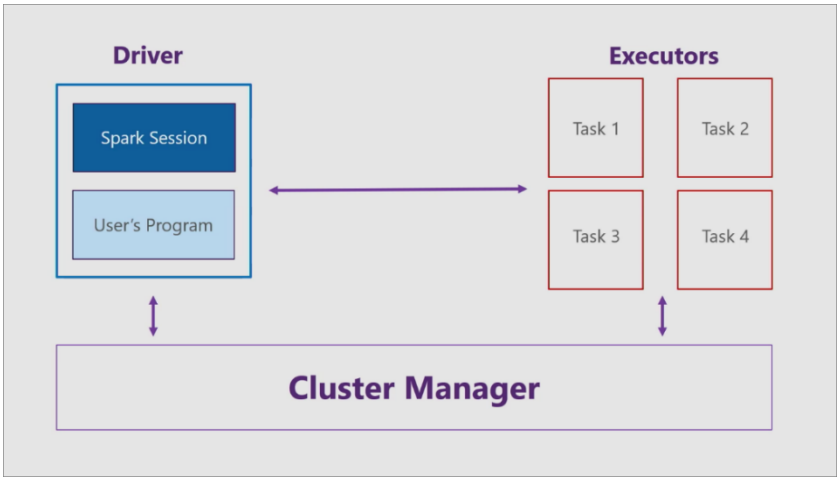


Figura 47 – Modello di esecuzione delle applicazioni Batch/Real Time Processing

Il driver è costituito dal programma, ad esempio un'app console C# e una sessione Spark. La sessione Spark accetta il programma e la divide in attività più piccole gestite dagli executors.

Ogni executor o nodo di lavoro riceve un'attività dal driver ed esegue tale attività. Gli executors risiedono in un'entità nota come cluster.

Il gestore del cluster comunica sia con il driver che con gli executors per:

- Gestire l'allocazione delle risorse
- Gestire la divisione del programma
- Gestire l'esecuzione del programma

Apache Spark supporta i seguenti linguaggi di programmazione:

- Scala
- Python
- Java
- SQL
- R
- Linguaggi .NET (C#/F#)

L'ecosistema Spark comprende cinque componenti chiave:

- Spark Core è un motore di elaborazione dati distribuito e per uso generico. Comprende librerie per SQL, l'elaborazione dei flussi, il machine learning e il calcolo dei grafici: tutti elementi che possono essere utilizzati insieme in un'applicazione. Spark Core è la base di un intero progetto, che fornisce l'invio, la programmazione e le funzionalità I/O di base di attività distribuite.
- Spark SQL è il modulo Spark per lavorare con dati strutturati che supporta un modo comune per accedere a una varietà di origini dati. Consente di eseguire query di dati strutturati all'interno dei programmi Spark, utilizzando SQL o un'API DataFrame familiare. Spark SQL supporta la sintassi HiveQL e consente l'accesso ai warehouse Apache Hive esistenti. Una modalità server fornisce la connettività standard attraverso la connettività di database Java o la connettività di database aperti.
- Spark Streaming semplifica la creazione di soluzioni per flussi di dati scalabili e a tolleranza di errore. Utilizza l'API integrata nel linguaggio di Spark per l'elaborazione dei flussi, in modo da poter scrivere job in flussi allo stesso modo dei job in batch. Spark Streaming supporta Java, Scala e Python e dispone di una semantica "exactly-once" di tipo stateful.
- MLlib è la libreria Spark scalabile per il machine learning con strumenti che rendono il ML pratico scalabile e facile da usare. MLlib contiene molti algoritmi di apprendimento comuni, come la classificazione, la regressione, i suggerimenti e il clustering. Contiene inoltre il flusso di lavoro e altre utilità, comprese le trasformazioni delle funzionalità, la costruzione di pipeline ML, la valutazione dei modelli, l'algebra lineare distribuita e le statistiche.
- GraphX è l'API Spark per i grafici e il calcolo grafico parallelo. È flessibile e funziona perfettamente sia con i grafici che con le raccolte: unifica estrazione, trasformazione, caricamento, analisi esplorativa e calcolo iterativo dei grafici all'interno di un unico sistema. Oltre a essere un'API altamente flessibile, GraphX fornisce diversi algoritmi grafici. In termini di prestazioni, compete con i sistemi di grafici più rapidi pur mantenendo la flessibilità, la tolleranza di errore e la facilità d'uso di Spark.

I vantaggi di Apache Spark sono:

- Velocità
  - Puoi eseguire i carichi di lavoro 100 volte più rapidamente di Hadoop MapReduce. Spark raggiunge elevate prestazioni sia per i dati in batch che per quelli in flusso utilizzando uno scheduler per grafico aciclico diretto (DAG) all'avanguardia, un ottimizzatore di query e un motore di esecuzione fisica.
- Facilità di utilizzo
  - Spark mette a disposizione più di 80 operatori di alto livello che facilitano la creazione di app parallele. Puoi utilizzarlo in modo interattivo dalle shell Scala, Python, R e SQL per scrivere rapidamente le applicazioni.
- Flessibilità
  - Spark supporta uno stack di librerie, tra cui SQL e DataFrames, MLlib per il machine learning, GraphX e Spark Streaming. È possibile combinare perfettamente queste librerie nella stessa applicazione.
- Modello Open Source
  - Spark è supportato da community globali unite per l'introduzione di nuovi concetti e funzionalità in modo più rapido ed efficace rispetto ai team interni che lavorano su soluzioni proprietarie. Il potere collettivo di una community open source genera un numero maggiore di idee, offre uno sviluppo più rapido e assicura la risoluzione dei problemi, contribuendo ad accelerare il time to market.
- Spark vs. motore SQL classico
  - Apache Spark è un veloce motore di calcolo per cluster per uso generico di cui è possibile eseguire il deployment in un cluster Hadoop o in modalità autonoma. Con Spark, i programmatori possono scrivere rapidamente applicazioni in Java, Scala, Python, R e SQL, diventando così accessibile a sviluppatori, data scientist e professionisti esperti di statistica. Con Spark SQL, gli utenti possono connettersi a qualsiasi origine dati e presentarla come tabella da utilizzare per i client SQL. Inoltre, gli algoritmi interattivi di machine learning sono facilmente implementabili in Spark.
- Connettori
  - Apache Spark è in grado di leggere e scrivere interfacciandosi con la maggior parte dei distributed storage, object storage, RDBMS, NoSQL, Message Broker e altre tipologie di storage adatto all'archiviazione di dati in modalità classica o streaming.

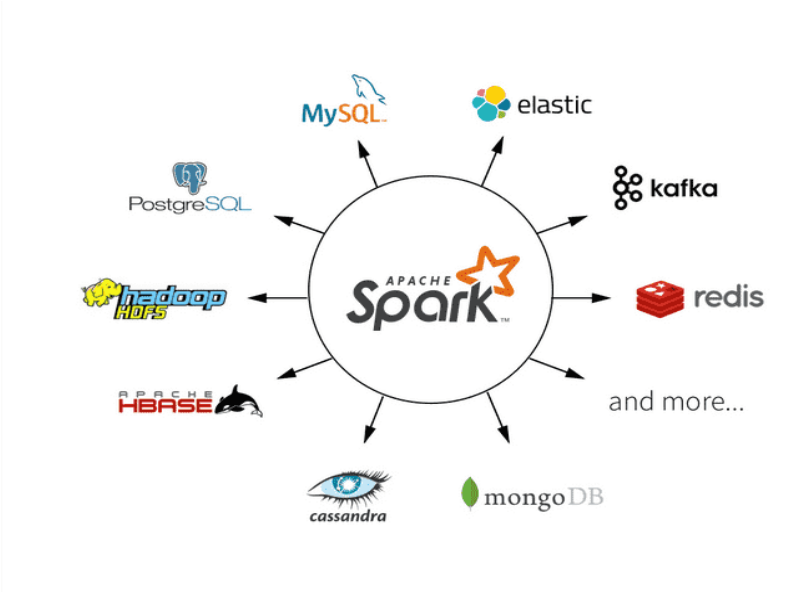


Figura 48 – Connettori di Spark

In particolare, è garantita la possibilità di leggere e scrivere, tramite protocollo S3, dati archiviati sul servizio PaaS Data Lake.

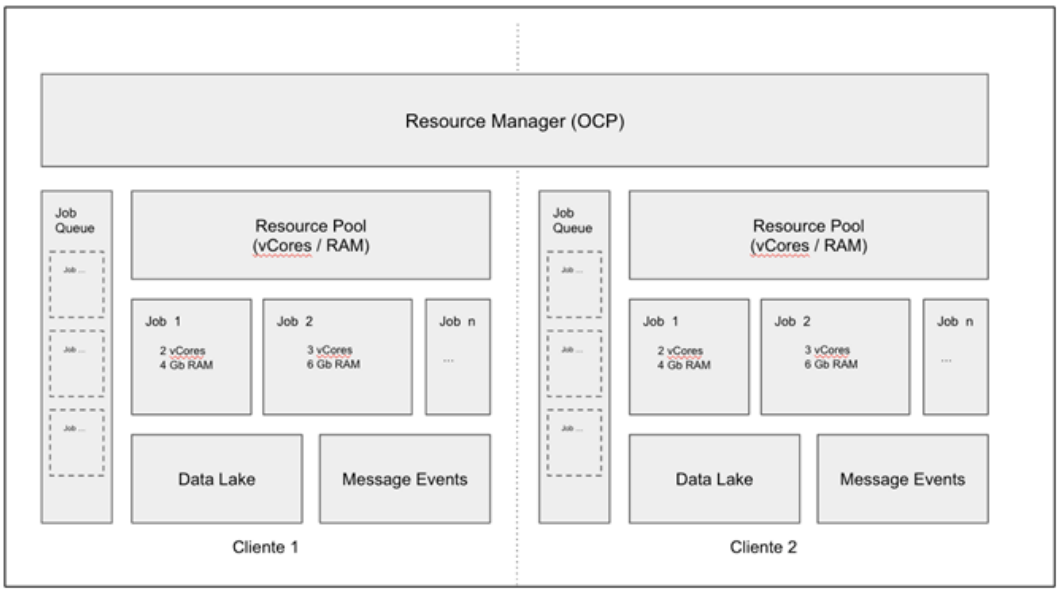


Figura 49 – Soluzioni di Alto livello e Multitenancy per il servizio di Batch/Real Time Processing

- Dettaglio implementazione multitenancy
  - Per ogni cliente attivo sul PSN saranno riservate un pool di risorse quantificate in "workers", ovvero un certo numero di cores e quantitativo di RAM
  - I workers disponibili si traducono in una "resource quota" legata al namespace (k8s) del cliente

- Il cliente avrà la possibilità di creare cluster spark (driver + n spark executors) per poter eseguire applicazioni di tipo interattivo (es: notebooks o shell) o non interattivo (es: job di tipo batch o streaming)
- In assenza di condizioni contrattuali concordate ad hoc, normalmente, qualora il cliente abbia superato la propria quota di risorse, eventuali applicazioni che verranno lanciate verranno posizionate in una coda in attesa che le risorse richieste diventino nuovamente disponibili (modalità pending)

1.1.5.3.2 Infrastrutture

L'infrastruttura si basa sul PaaS erogato dal PSN di Batch & Real time Processing.

Il PSN, in qualità di provider, si fa carico di gestire l'infrastruttura sottostante, comprensiva degli strumenti di automation e orchestration. L'offerta dei servizi PaaS prevede un approccio strutturato in cui ogni componente della soluzione PaaS, come il sistema operativo, solution stack ed altri software necessari, è strettamente controllata in termini di utilizzo e configurazione e gestita dal PSN. In questo caso le soluzioni vengono "create" al momento della necessità. Una rappresentazione di questa strutturazione vede quattro livelli di componenti:

- sistema operativo;
- run-time e librerie necessarie;
- soluzione caratterizzante – tipicamente un database, middleware, web server, etc.;
- un'interfaccia programmatica con cui controllare gli aspetti operazionali della soluzione.

1.1.5.4 BI Platform

1.1.5.4.1 Architettura tecnica

Tableau

L'architettura del modulo è costituita da una installazione di **Tableau Server On-Premise**, prodotto di riferimento e piattaforma di Visual Analytics su cui vengono realizzate tutte le visualizzazioni orientate sia al monitoraggio che al supporto alla presa di decisioni in riferimento alle applicazioni verticali, oltreché alla dashboard Executive Summary. Per comprendere appieno come Tableau riesce ad ottimizzare la gestione e l'analisi dei dati all'interno dell'organizzazione, è fondamentale conoscere le sue due principali componenti: Tableau Desktop e Tableau Server. Questi due strumenti lavorano in tandem per consentire ai professionisti di elaborare, visualizzare e condividere dati in modo efficiente e sicuro.

Tableau Desktop è un'applicazione software progettata per la creazione e la progettazione di visualizzazioni e report interattivi basati su dati. Gli utenti utilizzano Tableau Desktop per connettersi a varie fonti di dati, esplorare e analizzare i dati, quindi creare grafici, dashboard e report interattivi. Una volta creati i contenuti, possono essere pubblicati su Tableau Server per essere condivisi e collaborati con altri utenti.

Tableau Server è una piattaforma basata su server che consente agli utenti di pubblicare, condividere e collaborare su report e dashboard creati con Tableau Desktop. È progettato per facilitare la distribuzione delle visualizzazioni dei dati in un'organizzazione in modo sicuro e scalabile. Gli utenti possono accedere a Tableau Server tramite un browser web o un'app mobile e interagire con i contenuti condivisi, eseguire ricerche nei dati e collaborare con altri utenti. Tableau Server offre anche funzionalità avanzate per la gestione della sicurezza, dell'autenticazione e del monitoraggio dei contenuti, garantendo che i dati rimangano protetti e accessibili solo alle persone autorizzate.

In breve, Tableau Desktop è lo strumento per la creazione di visualizzazioni e report, e Tableau Server è la piattaforma che consente di condividere, distribuire e collaborare su tali contenuti in modo sicuro all'interno di un'organizzazione.

Schematicamente, l'architettura del modulo si può rappresentare come nella figura seguente:

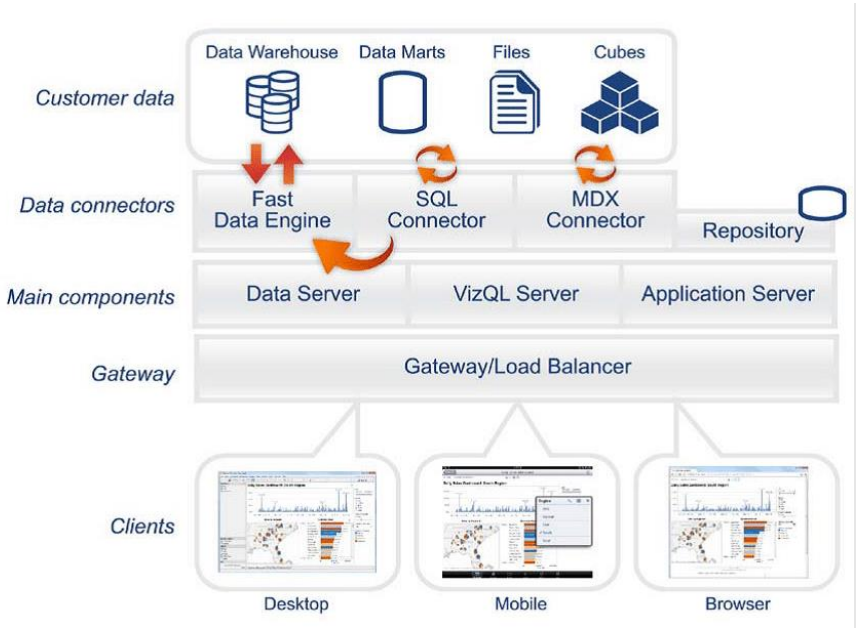


Figura 50 – BI Platform, architettura BI Platform Server

Data Server

Il Data Server consente di controllare e archiviare a livello centrale le origini dati di Tableau. Gestisce anche i metadati, come calcoli, definizioni e gruppi.

VizQL Server

Se una vista viene rilasciata, il client invia una richiesta al processo VizQL. Il processo VizQL invia quindi le query direttamente all'origine dati, restituendo un set di risultati che viene visualizzato come

immagini e presentato all'utente. Ogni server VizQL dispone della propria cache che può essere condivisa tra più utenti.

Application Server

Elabora la navigazione e le autorizzazioni per le interfacce Web e mobili di Tableau Server. Quando un utente apre una vista in un dispositivo client, avvia una sessione su Tableau Server. Quindi inizia il thread del server delle applicazioni e verifica le autorizzazioni per quell'utente e quella vista.

Gateway/Load balancer

Le richieste che arrivano dai client colpiscono prima il server gateway e poi vengono instradate alle procedure appropriate. Se sono configurate più procedure per qualsiasi componente, il Gateway funzionerà come bilanciatore del carico e condividerà le richieste con le procedure. In una configurazione a server singolo, ogni procedura si trova sul gateway o sul server primario. Quando si esegue in ambiente distribuito, una macchina fisica viene designata come server primario e le altre come server di lavoro che possono eseguire un numero qualsiasi di altre procedure.

A questa componente tecnologica si aggiungono funzionalità specifiche utili a:

- integrazione con il Data Lake, tramite predisposizione connettore JDBC o ODBC, per accesso a dati storici e real time.
- incorporamento di dashboard/report in applicazione web.
- integrazione con modulo IAM (Identity and Access Management) per profilazione verticale (per le funzionalità del modulo) e orizzontale (per l'accesso al dato).

Distribuzione processi

Tableau Server si può installare in locale, nel proprio cloud o data center privato e funziona anche sulle piattaforme di virtualizzazione. La figura seguente mostra i processi rilevanti di Tableau Server e come sono distribuiti nell'architettura di riferimento.

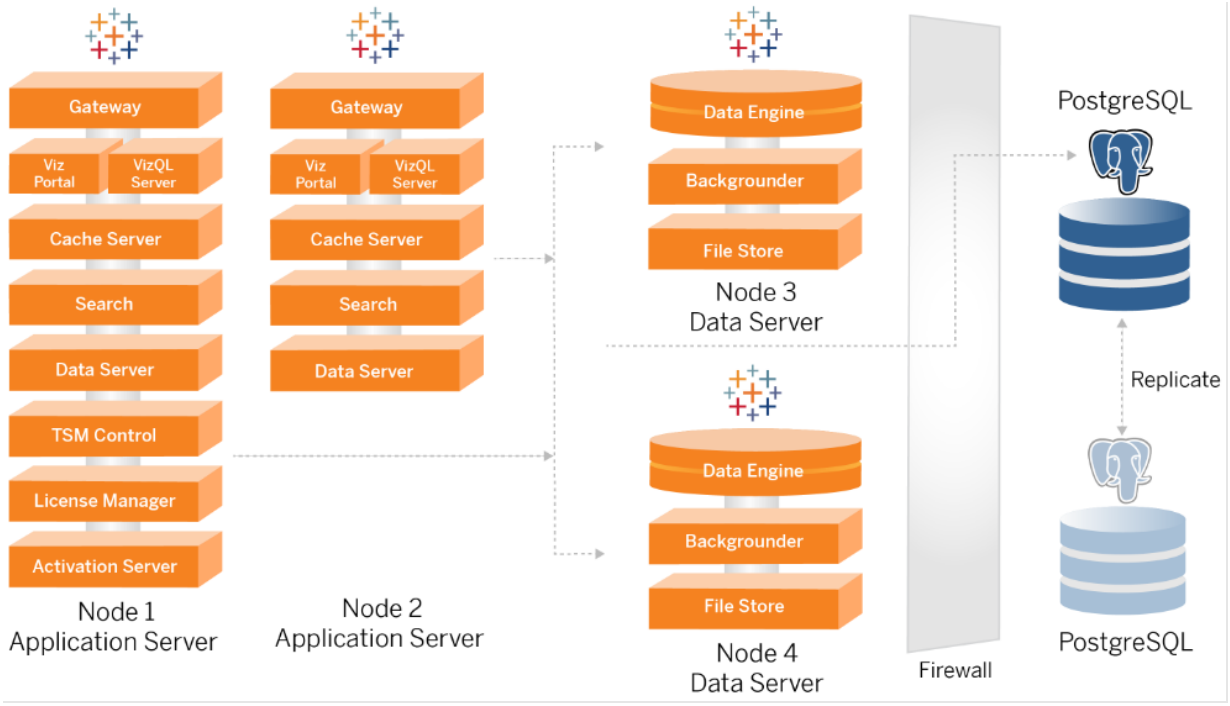


Figura 51 – BI Platform, architettura relativa ai processi più rilevanti di una installazione Tableau Server On-Premise

L'architettura di riferimento di Tableau Server è una distribuzione cluster di Tableau Server a quattro nodi con il repository esterno in PostgreSQL:

- **Nodo iniziale di Tableau Server (Nodo 1):** esegue i servizi amministrativi e di licenza di TSM richiesti che possono essere eseguiti solo su un singolo nodo nel cluster. Nel contesto aziendale, il nodo iniziale di Tableau Server è il nodo primario del cluster. Questo nodo esegue anche servizi applicativi ridondanti con il Nodo 2.
- **Nodi dell'applicazione Tableau Server (Nodo 1 e Nodo 2):** i due nodi elaborano le richieste client, si connettono a origini dati e nodi di dati ed eseguono query.
- **Nodi di dati di Tableau Server (Nodo 3 e Nodo 4):** due nodi dedicati alla gestione dei dati.
- **PostgreSQL esterno:** questo host esegue il processo di repository di Tableau Server. Per la distribuzione a disponibilità elevata è necessario eseguire un host PostgreSQL aggiuntivo per la ridondanza attiva/passiva.

Il repository di Tableau Server è un database che archivia i dati del server. Questi dati includono informazioni sugli utenti di Tableau Server, i gruppi e le assegnazioni di gruppo, le autorizzazioni, i progetti, le origini dati, l'estrazione di metadati e le informazioni di aggiornamento.

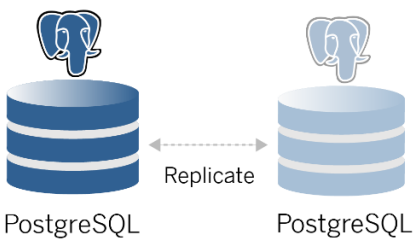


Figura 52 – BI Platform, il repository Tableau consiste in una replica PostgreSQL

La distribuzione predefinita di PostgresSQL utilizza quasi il 50% delle risorse di memoria del sistema. In base al relativo utilizzo (per distribuzioni di produzione e di grandi dimensioni), l'utilizzo delle risorse può aumentare. Per questo motivo, è consigliabile eseguire il processo Repository in un computer che non esegue altri componenti server con un utilizzo intensivo delle risorse, come VizQL, Gestione componenti in background o Motore dati. L'esecuzione del processo Repository insieme a uno di questi componenti creerà conflitti di I/O o vincoli di risorse e ridurrà le prestazioni complessive della distribuzione.

Grafana

Grafana utilizza l'architettura server-client.
Il backend Grafana (server) può essere autogestito tramite Grafana OSS Stack. Il server Grafana raccoglie dati di telemetria (metriche, log e tracce) e altro da una infrastruttura o applicazioni utilizzando data source esterne o tramite API. Il server fornisce diversi servizi per questi dati, tra cui analisi, grafici, avvisi, reporting e governance.
Grafana non raccoglie né archivia dati, quindi necessita di fonti dati esterne che possono essere facilmente integrate tramite installazione di plugin sviluppati da Grafana Labs o dalla community.

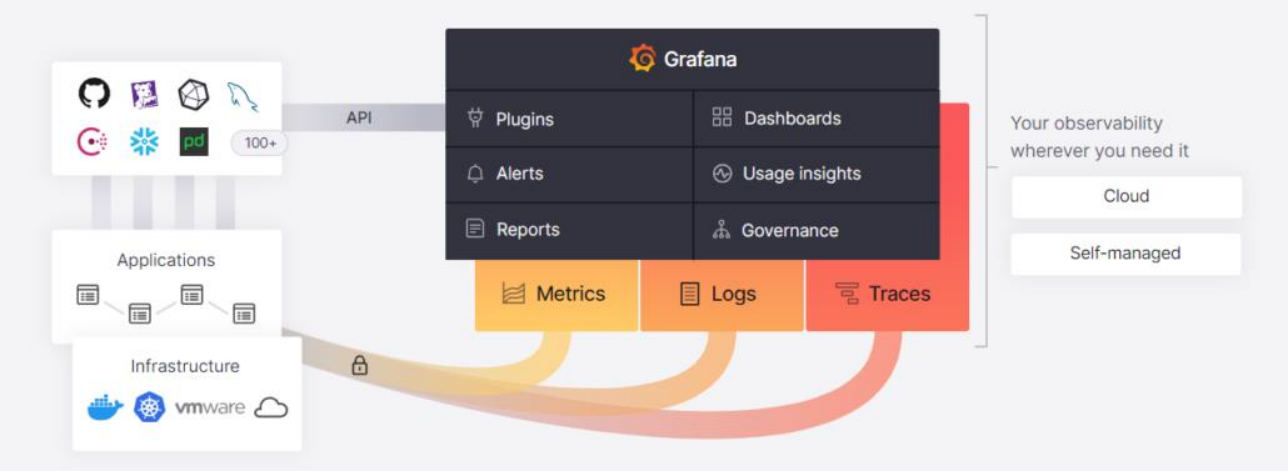


Figura 1: Architettura tecnica Grafana

Grafana server viene installato su una infrastruttura congeniale alle esigenze della piattaforma, sono consentite installazioni su Kubernetes, Docker, ecc. Successivamente, i plugin, che fungono da connettori, permettono di collegare Grafana alle fonti dati senza che questi vengano immagazzinati sul server Grafana.
La UI di Grafana permette poi all'utente la creazione di dashboard interattive per la rappresentazione delle telemetrie e KPI.

1.1.5.4.2 Tableau generative AI

Tableau può integrare nella propria piattaforma l'intelligenza artificiale per consentire a tutti di far emergere previsioni e informazioni, e di comprenderne la rilevanza dei dati aiutandoti a prendere decisioni più intelligenti direttamente nel flusso dell'analisi. Le seguenti funzionalità non sono indirizzate solo agli "addetti ai lavori" come data scientist o data analyst, ma anche a business user che intendo creare modelli predittivi intuitivi senza dover necessariamente scrivere una riga di codice. Tableau, quindi, permette di sfruttare tutti i vantaggi di produttività dell'IA generativa senza compromettere la sicurezza e la privacy dei dati.
In particolare, le due funzionalità sono **Einstein Copilot** e **Tableau Pulse**:

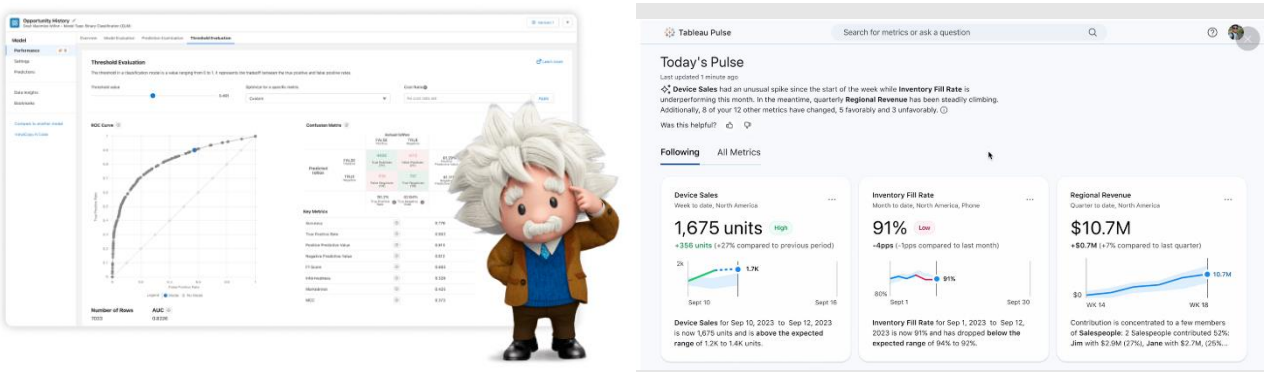


Figura 53 – BI Platform, Einstein Copilot a sinistra, e Tableau Pulse a destra

Einstein Copilot

Tramite il linguaggio naturale puoi chiedere ad Einstein Copilot di generare una nuova colonna estraendo una stringa secondo una certa regola da una colonna già esistente, o puoi chiedere di fare un calcolo avanzato, o verificare se esiste una qualche relazione tra i campi del dataset. Questo strumento funziona come una vera chat con cui puoi comunicare tramite il normale linguaggio.

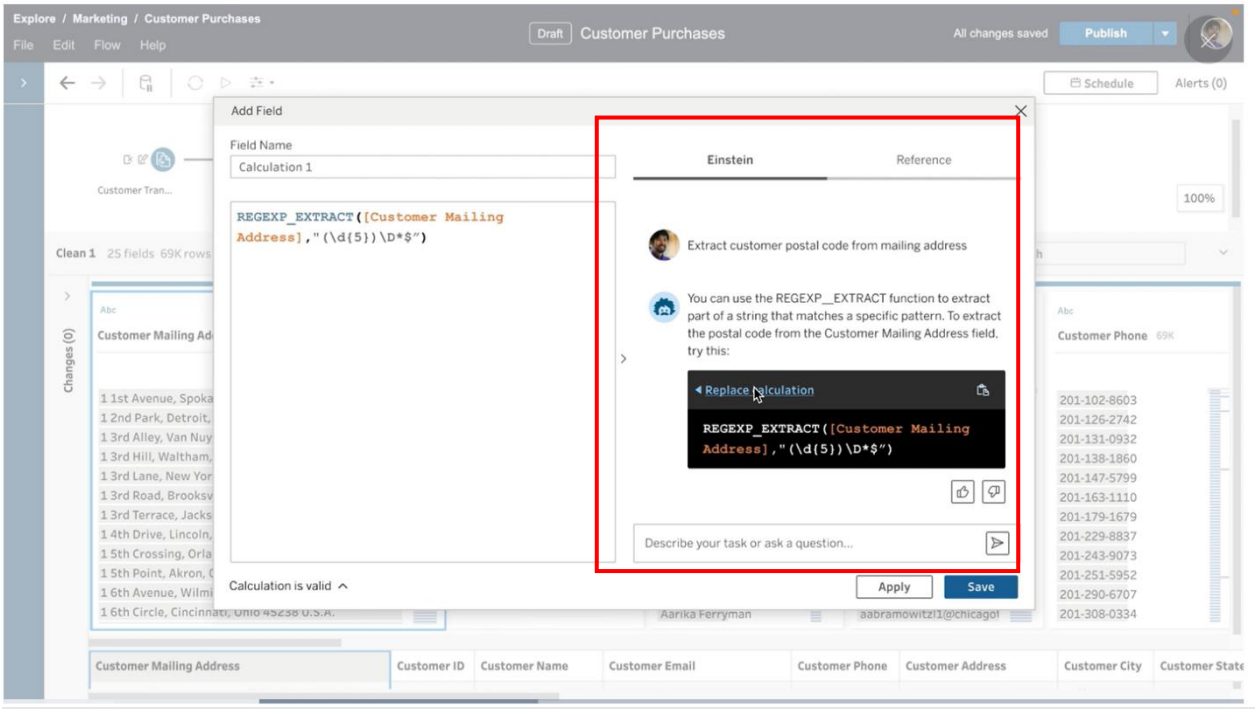


Figura 54 – BI Platform, esempio di Einstein Copilot

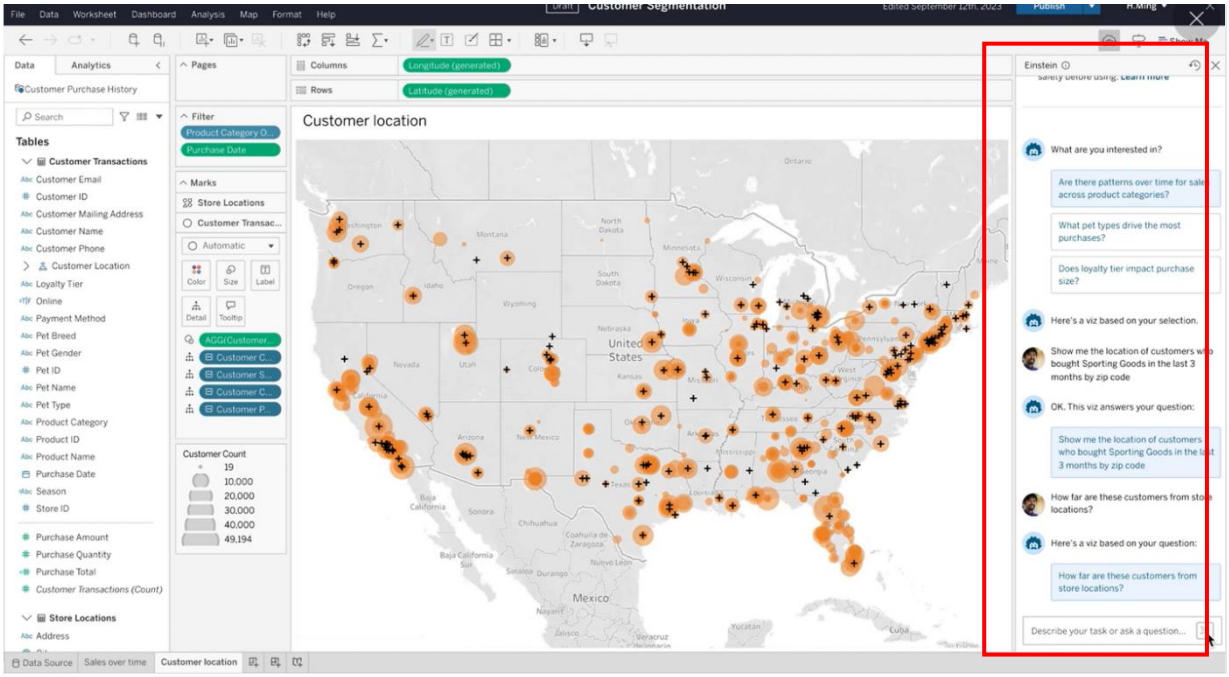


Figura 55 – BI Platform, mappa generata tramite richiesta ad Einstein Copilot

Tableau Pulse

Presenta analisi automatizzate in linguaggio semplice, anticipa proattivamente le domande degli utenti e suggerisce persino domande che potrebbero non aver considerato diversamente. Tableau Pulse trasforma il modo in cui le persone interagiscono con i loro dati, aiutando tutti nell'organizzazione a diventare orientati ai dati.

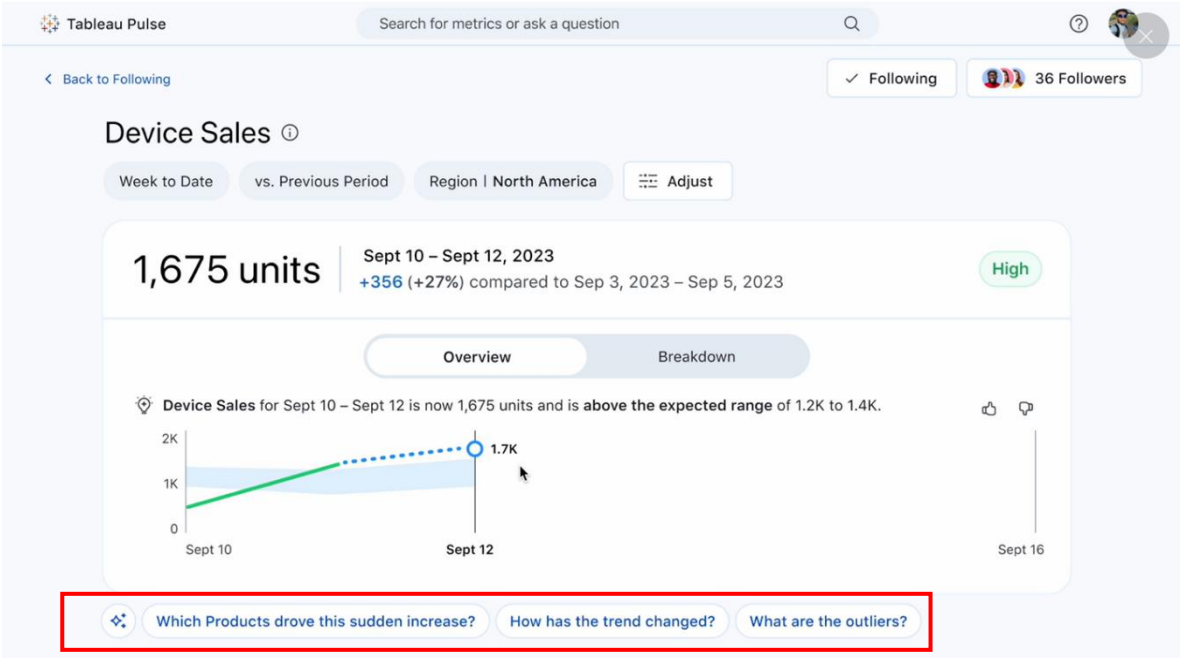


Figura 56 – BI Platform, Tableau pulse

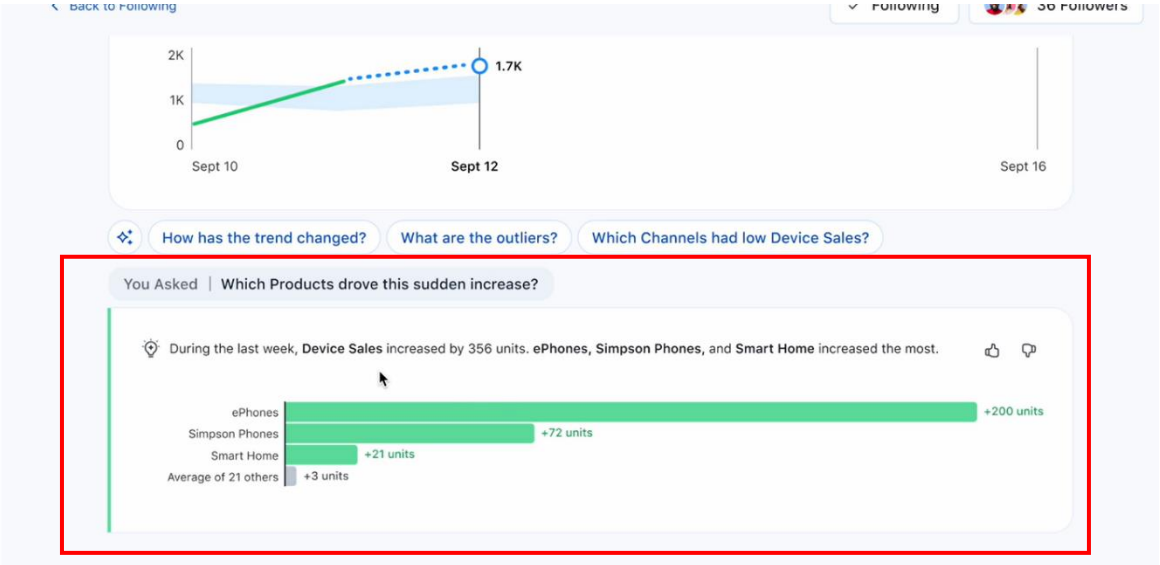


Figura 57 – BI Platform, grafico suggerito da Tableau Pulse per via di outlier nel dataset

1.1.5.4.3 Infrastruttura

La BI Platform prevede sia componenti ospitate da un Container Platform sia componenti residenti su Virtual Machine.

I POD previsti per la soluzione del modulo sono allocati sull'infrastruttura CaaS del PSN, in questo modo è gestito, ad esempio, l'autoscaling.

Sono altresì previste macchine virtuali per installazione dei tool cliente quali Tableau Desktop in cui gli utilizzatori della piattaforma possono creare dashboard e visualizzazioni.

Unitamente alla componente server, lo strato Data Store della soluzione è ospitato dal blocco logico Data System (RDS) del SIM e quindi dalle piattaforme e dai servizi del PSN. Sono coinvolte il PaaS Data Lake, il PaaS DB e le eventuali componenti infrastrutturali quali, ad esempio, il file system.

1.1.5.5 SQL/NO SQL DB

1.1.5.5.1 Architettura tecnica

Avendo diverse esigenze nell'ambito della gestione dei dati, questi sono i database utilizzati con caratteristiche distintive e che soddisfino specifici casi d'uso:

PostgreSQL: Un DBMS relazionale open-source altamente estensibile con un forte supporto per l'estensione delle funzionalità. È conosciuto per la sua affidabilità, sicurezza e conformità agli standard. il DB sarà incluso dell'estensione PostGIS che aggiunge il supporto per gli oggetti spaziali e la geolocalizzazione,

MongoDB: Un database non relazionale (NoSQL) orientato ai documenti, ideale per gestire dati non strutturati o semi-strutturati. È noto per la sua flessibilità e scalabilità orizzontale.

1.1.5.6 Time Series Database

il Time Series DB (TSDB) è parte di una particolare categoria di database specializzati per gestire elevate quantità di dati espresse come serie temporali. Questi database sono specificamente progettati per immagazzinare e analizzare sequenze ordinate di valori, misurati a intervalli regolari o irregolari lungo l'asse temporale.

Uno dei principali vantaggi di un Time Series DB risiede nella sua capacità di gestire enormi volumi di metriche, spesso provenienti da svariate fonti di dati con un'elevata cardinalità (questo ampiamente descritto di seguito). Questa peculiarità lo rende particolarmente efficace nell'elaborazione di dati provenienti da dispositivi IoT, che generano flussi continui di informazioni in tempo reale. Grazie alla sua struttura ottimizzata, un Time Series DB può archiviare, aggregare e recuperare questi dati con una velocità e un'efficienza notevoli.

Ma non si tratta solo di velocità e capacità. La natura stessa di un Time Series DB permette di rilevare modelli, anomalie o tendenze nei dati nel corso del tempo. Questo tipo di analisi temporale diventa fondamentale quando si considera l'importanza della previsione e della reazione tempestiva in contesti come le rilevazioni sismiche, la prevenzione delle frane o l'analisi dell'inquinamento marino.

Attraverso una gestione efficace delle serie temporali, è possibile non solo comprendere il presente, ma anche anticipare il futuro, attraverso elaborazioni predittive basate su modelli che fanno uso dei dati storici archiviati.

La cardinalità in un database rappresenta il numero di valori unici per un determinato campo o insieme di campi. In un contesto di Time Series DB, dove l'enfasi è sulla tracciatura dei dati nel tempo, la cardinalità può avere implicazioni significative, soprattutto quando ci si riferisce a dati con una varietà di misurazioni e metadati associati.

Prendendo come esempio i dati provenienti dalle centraline meteo, è possibile comprendere pienamente la complessità legata alla cardinalità. Ogni centralina, in un singolo momento, può registrare una serie di variabili meteorologiche come temperatura, umidità, velocità e direzione del vento, livello di precipitazione e intensità di irradiazione solare. Ogni misurazione è caratterizzata da un *timestamp* che indica il momento preciso della registrazione.

Oltre a queste misurazioni, vi è anche un elemento critico che amplifica ulteriormente la cardinalità: la posizione geografica. Contrariamente ad altri dati, come il codice identificativo del dispositivo che può essere un valore discreto e limitato, la latitudine e la longitudine sono dati continui. Ciò significa che, teoricamente, anche una minima variazione nelle coordinate può produrre un valore unico. Quando si considera l'ampio spettro di possibili coordinate geografiche, si comprende che la cardinalità potenziale può diventare virtualmente illimitata.

Le implicazioni della cardinalità

Come già accennato la gestione di un elevata cardinalità nei dati rappresenta una sfida considerevole per i database tradizionali. La necessità di indicizzare e interrogare un numero così vasto di valori unici può influire notevolmente sulle prestazioni e sull'efficienza del sistema. Ecco perché i Time Series DB sono stati progettati e ottimizzati specificamente per affrontare questa sfida.

Un Time Series DB, grazie alle sue caratteristiche intrinseche, è in grado di gestire enormi volumi di dati con una cardinalità elevata, offrendo al contempo prestazioni efficienti in termini di scrittura, lettura e interrogazione. La sua architettura e le tecniche di compressione dei dati gli permettono di archiviare in modo efficiente sequenze di misurazioni, riducendo l'overhead associato alla gestione di valori unici e garantendo tempi di risposta rapidi alle query, anche in presenza di una cardinalità estremamente elevata.

La soluzione di Time Series DB per il SIM

Nel panorama dei Time Series DB (TSDB), una soluzione, open source, particolarmente efficace e apprezzata per gestire dati con elevata cardinalità è TimeScaleDB. Si tratta di un'estensione del noto database relazionale PostgreSQL, ma ottimizzato per la gestione di serie temporali. Questa integrazione tra le caratteristiche dei database relazionali e le specifiche necessità dei Time Series DB rende TimeScaleDB una scelta adatta per una vasta gamma di applicazioni, compresi i sistemi che gestiscono dati ambientali e territoriali.

Le caratteristiche rilevanti di TSDB sono:

- **Scalabilità:** Una delle principali sfide nel trattare enormi quantità di dati è la capacità del sistema di scalare in modo efficiente. TimeScaleDB, sfruttando la natura di PostgreSQL, offre una scalabilità nativa. Questo significa che è possibile gestire un aumento progressivo del volume dei dati senza sacrificare le prestazioni. Che si tratti di terabyte o petabyte di dati, TimeScaleDB può gestirli, rendendolo ideale per ambienti in cui i dati vengono acquisiti continuamente da dispositivi come sensori e centraline.
- **Resilienza:** La sicurezza e la robustezza dei dati sono essenziali, specialmente quando si tratta di monitorare e proteggere l'ambiente. TimeScaleDB eredita le caratteristiche di resilienza di PostgreSQL. Ciò garantisce che i dati siano protetti da perdite accidentali e che il sistema possa recuperare rapidamente da eventuali interruzioni.
- **Prestazioni:** Per quanto riguarda l'efficienza nelle operazioni di lettura e scrittura, TimeScaleDB è ottimizzato per fornire tempi di risposta rapidi. Grazie a specifiche tecniche di indicizzazione e compressione, riesce a gestire grandi volumi di dati temporali in modo molto più efficiente rispetto ai database relazionali tradizionali.
- **Flessibilità:** Essendo basato su PostgreSQL, TimeScaleDB offre una grande flessibilità in termini di tipi di dati e struttura delle query. Questo permette di sfruttare sia le funzionalità avanzate di un database relazionale sia quelle specifiche dei Time Series DB. Per esempio, la possibilità di eseguire join complessi o di utilizzare funzioni analitiche avanzate.

TSDB combina il meglio dei mondi dei database relazionali e dei Time Series DB. Questa sinergia lo rende particolarmente adatto per applicazioni dove la cardinalità elevata, la scalabilità e la resilienza sono cruciali, come nel caso del trattamento dell'enorme mole di dati, su base temporale, relativa si servizi a tutela dell'ambiente e del territorio.

Ingestion dei dati e memorizzazione

L'ingestion di dati in un Time Series DB, come TimeScaleDB, rappresenta un momento critico nell'architettura di un sistema. Quando si tratta di dati provenienti da vari dispositivi IoT, come sensori meteorologici, droni, apparati di rilevazione e altri, l'efficienza e la velocità di ingestion sono essenziali per garantire l'aggiornamento tempestivo delle informazioni e la capacità del sistema di rispondere in tempo reale agli eventi.

Nel contesto di TimeScaleDB, ci sono due approcci principali all'ingestion: modalità **push** e modalità **pull**.

- **Modalità Push:** in questo approccio, i dispositivi IoT o i sistemi che raccolgono dati inviano attivamente le loro rilevazioni al database. Questo metodo è spesso preferito quando i dispositivi hanno la capacità di stabilire connessioni di rete in modo indipendente e di inviare dati a intervalli regolari o in risposta a particolari eventi. L'ingestione in modalità push è immediata e consente di avere dati sempre aggiornati, ma richiede una gestione attenta delle connessioni e del traffico di rete per evitare congestioni o perdite di dati.
- **Modalità Pull:** contrariamente alla modalità push, in questo caso è il Time Series DB, o un servizio associato, che richiede periodicamente dati dai dispositivi. La modalità pull è particolarmente adatta quando i dispositivi non possono o non devono stabilire connessioni in modo autonomo. Questo approccio può anche aiutare a ridurre il traffico di rete, poiché le richieste possono essere programmate e ottimizzate in base alle esigenze del sistema.

Nel contesto del sistema SIM viene utilizzata la piattaforma IoT Hono che di fatto funziona come elemento di intermediazione per alimentare il TSDB.

Nell'ambito dei servizi a tutela dell'ambiente e del territorio, la scelta tra push e pull dipenderà dalla natura e dalle caratteristiche dei dispositivi di rilevazione. Ad esempio, un sensore fisso che monitora la qualità dell'aria potrebbe inviare dati in modalità push ogni volta che rileva una variazione significativa, mentre un sistema di monitoraggio della crescita della vegetazione potrebbe essere configurato in modalità pull, con il database che richiede dati a intervalli regolari.

L'importante è che l'ingestion dei dati sia fluida, efficiente e affidabile. TimeScaleDB, offre gli strumenti necessari per gestire entrambi gli approcci, garantendo che i dati siano sempre disponibili e pronti per l'analisi non appena vengono raccolti.

Uso dei dati memorizzati nel TSDB

Nel contesto dell'utilizzo di un Time Series DB, come TimeScaleDB, una volta che i dati sono stati correttamente ingeriti e archiviati, il passo successivo è capire come questi possono essere interrogati ed utilizzati in modo efficace. La gestione di volumi ingenti di dati temporali, specialmente quando sono correlati a fenomeni complessi come quelli ambientali, richiede strumenti e tecniche particolari per garantire tempi di risposta rapidi e risultati accurati.

Con TimeScaleDB, le query possono essere ottimizzate per interrogare grandi quantità di dati temporali. Grazie alla sua base su PostgreSQL, TimeScaleDB eredita la potenza del linguaggio SQL, ma con estensioni e funzionalità specifiche per la gestione dei dati temporali. Questo significa che è possibile eseguire operazioni complesse come aggregazioni temporali, interpolazioni e correlazioni su grandi set di dati in modo efficiente.

Inoltre, TimeScaleDB supporta formati di dati avanzati, tra cui il dato di tipo **BJSON (Binary JSON)**. Questo formato consente di gestire dati strutturati e semi-strutturati, offrendo una grande flessibilità nel modo in cui le informazioni possono essere rappresentate e memorizzate. Ad esempio, i dati provenienti da sensori IoT potrebbero includere sia misure numeriche (come temperature o livelli di umidità) sia metadati (come l'ID del dispositivo o la sua posizione geografica). Utilizzando BJSON, questi dati possono essere memorizzati in un unico record, semplificando le query e migliorando le prestazioni.

Esempi di query relative a dati ambientali tipicamente memorizzate in TimeScaleDB

- Top of Form

Ecco alcuni esempi di query per interrogare TimeScaleDB in relazione a dati di misurazioni provenienti dai dispositivi dispiegati nel territorio:

1. Recuperare le ultime dieci misurazioni di temperatura e umidità per un determinato dispositivo in una specifica posizione:

```
SELECT timestamp, temperature, humidity
FROM measurements
```

```
WHERE device_type = 'centralina_meteo' AND longitude = 12.4924 AND latitude = 41.8902
ORDER BY timestamp DESC
LIMIT 10;
```

2. Calcolare la media di temperatura e umidità per ogni giorno dell'ultimo mese per un determinato dispositivo in una specifica posizione:

```
SELECT time_bucket('1 day', timestamp) AS day,
       AVG(temperature) AS avg_temperature,
       AVG(humidity) AS avg_humidity
FROM measurements
WHERE device_type = 'centralina_meteo' AND longitude = 12.4924 AND latitude = 41.8902
AND timestamp > NOW() - INTERVAL '1 month'
GROUP BY day;
```

3. Trovare il valore massimo e minimo di temperatura e umidità nell'ultima settimana per una specifica posizione:

```
SELECT MAX(temperature) AS max_temperature,
       MIN(temperature) AS min_temperature,
       MAX(humidity) AS max_humidity,
       MIN(humidity) AS min_humidity
FROM measurements
WHERE device_type = 'centralina_meteo' AND longitude = 12.4924 AND latitude = 41.8902
AND timestamp > NOW() - INTERVAL '1 week';
```

4. Contare il numero di misurazioni al di sopra di una certa soglia di temperatura nell'ultimo giorno per una specifica posizione:

```
SELECT COUNT(*)
FROM measurements
WHERE temperature > 35 AND timestamp > NOW() - INTERVAL '1 day'
AND device_type = 'centralina_meteo' AND longitude = 12.4924 AND latitude = 41.8902;
```

5. Rilevare eventuali anomalie nella misurazione in una specifica posizione:

```
SELECT timestamp, temperature, humidity
FROM measurements
WHERE (humidity < 10 OR humidity > 90)
AND device_type = 'centralina_meteo' AND longitude = 12.4924 AND latitude = 41.8902
ORDER BY timestamp DESC;
```

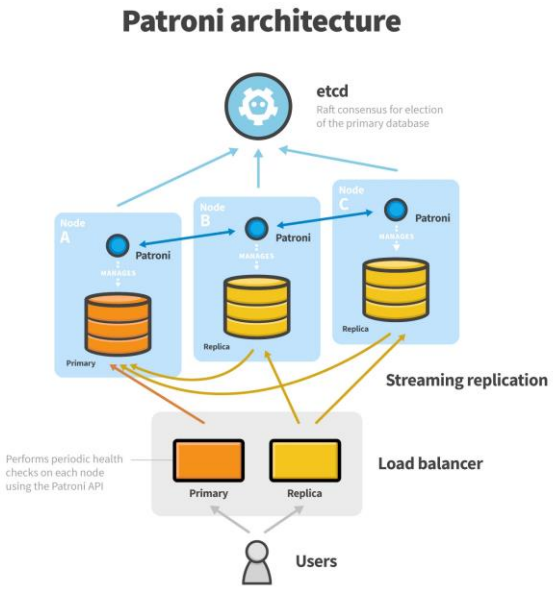
Queste query assumono che lo schema dell'hypertable TimeScaleDB includa le coordinate: `latitude` e `longitude` e che vi sia un campo `device_type` per specificare la tipologia di apparato.

Uso di TimeScaleDB in configurazione HA

Considerando che il TimeScaleDB è utilizzato per implementare il servizio che gestisce i principali data set dei dati IoT e in generale di qualsiasi altra metrica su base temporale del SIM, associato al

RdS (Repository di Sistema) è alquanto importante garantire un elevato livello di resilienza di questo servizio.

A tale scopo di seguito viene descritta la soluzione raccomandata da TimeScale (azienda che ha creato e offre il supporto di TimeScaleDB).



Architettura di TimeScaleDB in modalità HA

La figura mostra la tipica architettura di un'installazione di PostgreSQL in modalità high-availability (HA) utilizzando Patroni, una soluzione che può sfruttare ZooKeeper, etcd o Consul per il consensus e la gestione delle configurazioni. Quando parliamo di TimeScaleDB, stiamo parlando di un'estensione di PostgreSQL che è ottimizzata per le serie temporali. Pertanto, l'architettura di HA di PostgreSQL si applica anche a TimeScaleDB.

Descrizione dell'architettura:

1) **Nodi**: l'architettura mostra tre nodi (Node A, Node B, Node C) che ospitano installazioni di Patroni e istanze di database PostgreSQL (o TimeScaleDB).

2) **etcd**: questo componente è responsabile del consensus e della gestione delle configurazioni. Utilizza l'algoritmo Raft per eleggere il nodo primario del database. Gli altri nodi diventano repliche.

3) **Streaming replication**: questa è la tecnica utilizzata per replicare i dati dal nodo primario alle repliche. Garantisce che tutte le repliche siano sincronizzate con il primario.

4) **Load balancer**: questo componente distribuisce le richieste degli utenti tra il nodo primario e le repliche, garantendo che la maggior parte delle operazioni di lettura venga distribuita alle repliche per bilanciare il carico.

5) **Patroni API**: Patroni esegue controlli di salute periodici su ogni nodo attraverso la sua API. Se il nodo primario fallisce, Patroni utilizza etcd (o ZooKeeper o Consul, a seconda della configurazione) per eleggere una nuova istanza primaria tra le repliche disponibili.

Benefici dell'adozione di questa architettura:

1) **Alta disponibilità**: grazie alla replica e all'*election* automatica di un nuovo nodo primario in caso di guasto, l'architettura garantisce che il database sia sempre disponibile.
2) **Scalabilità di lettura**: utilizzando un *load balancer*, è possibile distribuire le richieste di lettura tra diverse repliche, migliorando le prestazioni e la capacità di gestione delle richieste.
3) **Failover automatico**: In caso di guasto del nodo primario, Patroni assicura che una delle repliche diventi il nuovo primario, garantendo una minima interruzione del servizio.
4) **Gestione semplificata**: Patroni offre una soluzione pronta all'uso per la gestione di configurazioni, failover e health checks, semplificando la gestione dell'installazione di PostgreSQL/TimeScaleDB in modalità HA.
5) **Flessibilità**: Patroni supporta diverse soluzioni di consensus come etcd, ZooKeeper e Consul, offrendo flessibilità nella scelta della soluzione più adatta alle esigenze specifiche.

1.1.6 Intelligence Platform

1.1.6.1 DATA & AI Workflow

1.1.6.1.1 Una Piattaforma per Data Science a Livello Industriale e per la Realizzazione di Progetti basati su analisi dei dati e modellazione ML/DL AI

Dataiku rappresenta una soluzione avanzata e completa progettata per supportare la Data Science e la realizzazione di progetti basati sull'Intelligenza Artificiale (AI) a livello industriale. Questa piattaforma offre una vasta gamma di strumenti e funzionalità concepiti per affrontare tutte le sfide che le organizzazioni devono affrontare nel ciclo di vita dei progetti di Data Science.

Con Dataiku, è attuale un percorso agevole verso la realizzazione di soluzioni avanzate di AI. La piattaforma fornisce un ambiente integrato che copre tutti gli aspetti chiave della Data Science e dell'AI, consentendo un processo di sviluppo efficiente e scalabile.

La seguente figura mostra l'architettura a blocchi di Dataiku da un punto di vista funzionale dove è presente al centro il nucleo con le funzionalità esposte dal DSS (Digital Science Studio) raggruppate nel contesto della Progettazione e Sperimentazione e Produzione e Orchestrazione.

È graficamente evidente anche la capacità della piattaforma di offrire un uso molto assistito nello sviluppo (Low code) e ove necessario il supporto per un utilizzo rivolto a utenti molto esperti che possono lavorare sfruttando appieno la capacità massima (Full code).

Allo stesso tempo si può notare il modello di sicurezza e governance dell'ambiente che consente di applicare politiche di acceso e di uso delle funzionalità con un elevata granularità. Infine, la caratura enterprise grade della soluzione per garantire un elevato livello di resilienza nell'uso continuativo dei servizi offerti.

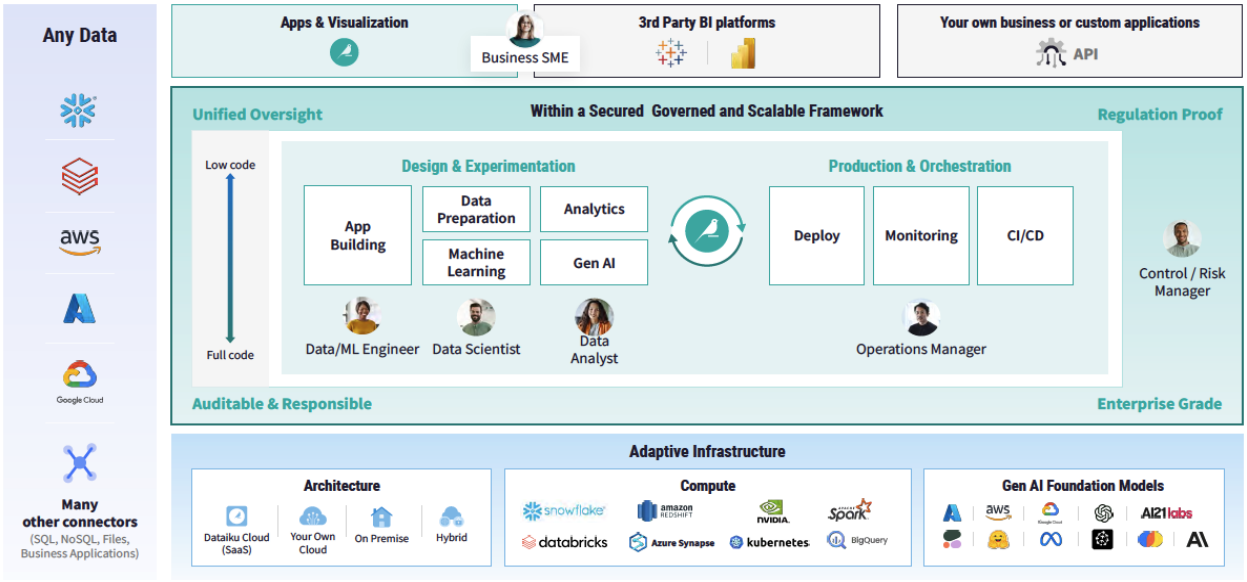


Figura 58 - Architettura a blocchi di Dataiku

Le caratteristiche salienti sono:

- **Approccio visuale**: uso di concetti rappresentativi delle attività svolte come Flows (flussi) e Recipes (ricette) per strutturare e semplificare il lavoro svolto con la piattaforma.
- **Integrazione Semplificata dei Dati**: Dataiku semplifica l'acquisizione dei dati da sorgenti eterogenee, tra cui database, servizi cloud e file locali. La piattaforma offre una vasta libreria di connettori preconfigurati che rendono la connessione alle fonti dati un processo fluido.
- **Preparazione dei Dati Intuitiva**: una volta acquisiti, i dati possono essere preparati, puliti e arricchiti utilizzando strumenti intuitivi all'interno dell'interfaccia utente di Dataiku. Gli utenti possono applicare trasformazioni complesse senza scrivere codice, migliorando l'efficienza della preparazione dei dati.
- **Visualizzazione dei Dati Interattiva**: permette agli utenti di esplorare e visualizzare i dati in modo interattivo. Grafici, tabelle e dashboard personalizzabili consentono di identificare facilmente pattern e tendenze nei dati.
- **Collaborazione Efficiente**: promuove la collaborazione tra gruppi, consentendo agli utenti di condividere progetti, analisi e risultati. Le funzionalità di commento e revisione facilitano il feedback e migliorano la produttività.
- **Sviluppo di Modelli Avanzati**: fornisce un ambiente completo per la creazione e la gestione di modelli di Machine Learning avanzati. Gli utenti possono eseguire analisi complesse, sviluppare modelli predittivi e valutarne le prestazioni, il tutto attraverso un'interfaccia visuale user-friendly.
- **Supporto per il Coding**: il DSS offre un approccio Low code per chi deve fare prototyping rapido e sperimentazione. Allo stesso per coloro che preferiscono scrivere codice, Dataiku supporta l'integrazione di script Python, R e SQL. Questa flessibilità consente agli sviluppatori, data scientist e ricercatori di personalizzare ulteriormente le analisi e l'elaborazione dei dati.
- **MLOps e Operationalization**: la soluzione promuove una solida pratica di MLOps, facilitando il rilascio e la gestione dei modelli in produzione. Gli utenti possono automatizzare i processi di implementazione, monitoraggio e aggiornamento dei modelli, garantendo la continuità delle operazioni aziendali basate sull'AI. Dataiku supporta direttamente anche l'uso di modelli in formato MLFlow. MLflow è una piattaforma open source per la gestione del ciclo di vita dell'apprendimento automatico e ha definito un formato standard per il packaging di modelli di apprendimento automatico addestrati: i cosiddetti MLflow Models. È possibile importare i modelli MLflow in DSS, come modelli salvati da DSS. Ciò consente di beneficiare di tutte le funzionalità di gestione ML di DSS sui modelli MLflow esistenti:
  - Data sets di scoring utilizzando una "ricetta" di scoring;
  - Distribuzione del modello in un bundle su un nodo di automazione (Distribuzioni e pacchetti di produzione);
  - Distribuzione del modello per lo scoring in tempo reale, utilizzando il nodo API
  - Gestione di più versioni dei modelli;
  - Valutare le prestazioni di un modello di classificazione o regressione su un set di dati etichettato, comprese tutte le schermate dei risultati;
  - Confronto di più modelli o più versioni del modello, utilizzando i confronti dei modelli;
  - Analizzare le prestazioni e valutare i modelli su altri set di dati;
  - Analisi della deriva sul modello MLflow;
  - Gestire il modello MLflow utilizzando il Govern Node.

Nel complesso Dataiku è una soluzione all'avanguardia per la realizzazione di progetti di Data Science e AI su larga scala e particolarmente adeguata agli scenari applicativi *mission-critical*

affrontati nel SIM. La ricchezza di funzionalità consente chiunque opera nel SIM di sfruttare il potenziale dei dati e ottenere risultati significativi in modo rapido, efficiente e collaborativo.

Flows e Recipes: un modo per facilitare il lavoro con dati e modelli

Nell'ambito dell'interfaccia visuale di Dataiku, sono presenti due concetti fondamentali: Flows e Recipes (flussi e ricette), ciascuno dei quali svolge un ruolo cruciale nell'ambito del processo di data science e analisi dei dati. La figura seguente mostra un tipico esempio di Flow con diverse Ricette.

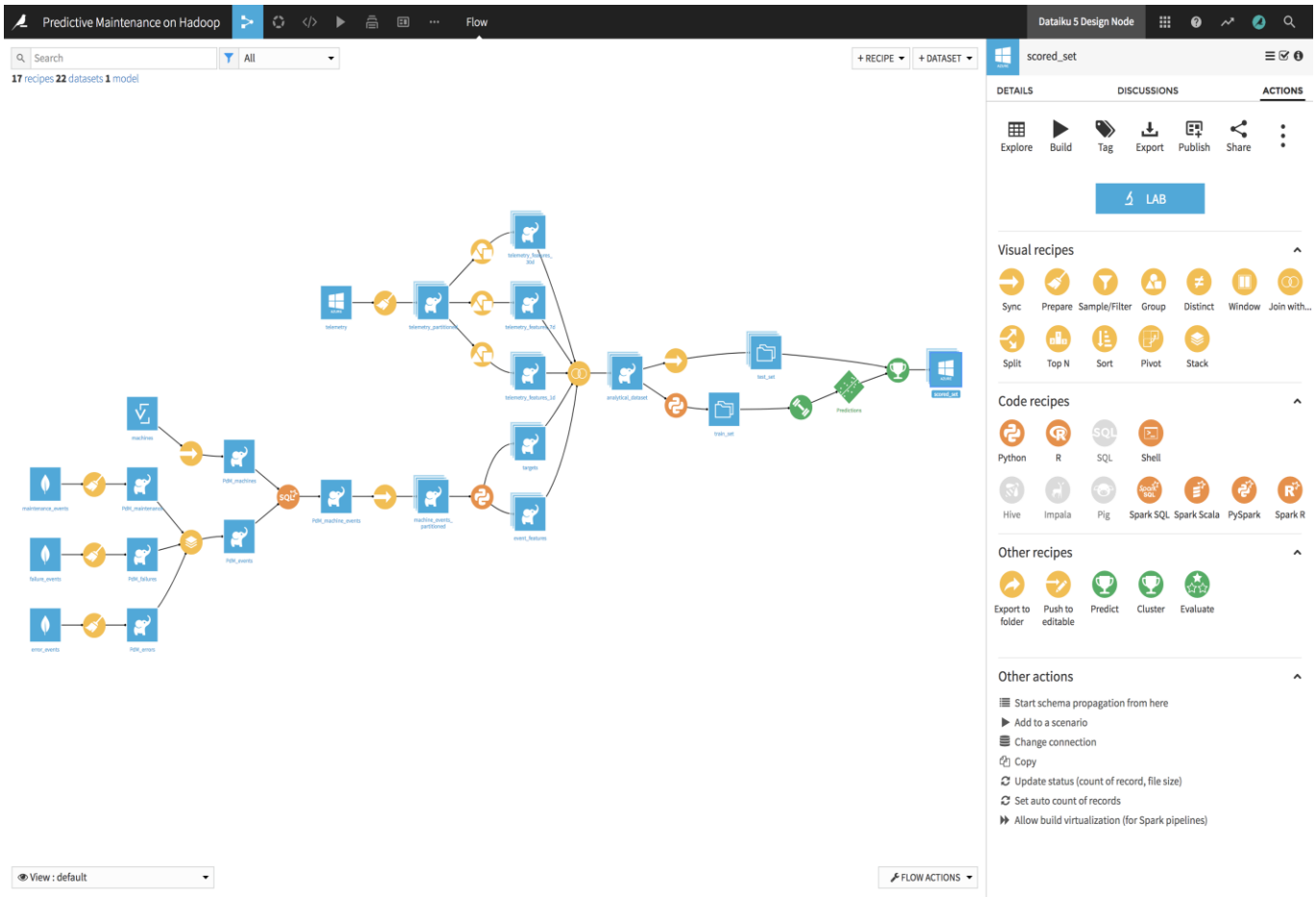


Figura 2 - Flows e Recipes

Di seguito, vengono trattati questi concetti e i vantaggi che ne derivano.

Flows in Dataiku

I Flows rappresentano il cuore dell'ambiente di lavoro di Dataiku. Sono rappresentazioni grafiche dei processi di data science e analisi dei dati. I Flows consentono agli utenti di definire la sequenza di operazioni e trasformazioni da eseguire sui dati, dalla loro acquisizione fino alla creazione di modelli predittivi. Alcuni vantaggi dei Flows includono:

- **Trasparenza Visiva**: l'interfaccia grafica dei Flows consente agli utenti di visualizzare chiaramente il flusso di lavoro e le interazioni tra le diverse fasi. Questa trasparenza aiuta a comprendere e comunicare il processo in modo efficace.

- **Modularità**: i Flows consentono la suddivisione del processo in moduli o passaggi separati. Questa modularità agevola la gestione e la manutenzione dei processi, oltre a facilitare l'aggiunta o la rimozione di passaggi quando necessario.
- **Collaborazione**: i Flows promuovono la collaborazione all'interno del gruppo di data science. Gli utenti possono lavorare in parallelo su diverse parti del flusso di lavoro, migliorando l'efficienza complessiva.
- **Tracciabilità**: ogni passaggio all'interno del Flow è documentato in modo automatico, consentendo una tracciabilità completa delle operazioni eseguite sui dati. Questo è fondamentale per scopi di audit e per garantire la ripetibilità del processo.

Recipes in Dataiku

Le Ricette (Recipes) sono componenti essenziali dei Flows e rappresentano le singole trasformazioni e operazioni eseguite sui dati. Ogni Ricetta definisce come i dati vengono elaborati in un determinato passaggio del flusso di lavoro. Alcuni vantaggi delle Ricette includono:

- **Facilità di Utilizzo**: le Ricette sono progettate per essere intuitive ed è possibile crearle utilizzando un'interfaccia visuale. Questo rende l'elaborazione dati accessibile anche a utenti senza competenze di programmazione.
- **Riusabilità**: le Ricette possono essere facilmente riutilizzate in diversi progetti e Flows. Ciò consente di risparmiare tempo e impegno di sviluppo, poiché le trasformazioni precedentemente create possono essere applicate nuovamente senza la necessità di ricodificare.
- **Catalogo di Preparazione Dati**: il DSS offre un catalogo di Ricette predefinite per una vasta gamma di operazioni di preparazione dati. Gli utenti possono sfruttare queste Ricette come punto di partenza per le proprie esigenze.
- **Tracciabilità e Riproducibilità**: ogni modifica apportata a una Ricetta è tracciata, consentendo di monitorare e ripetere facilmente le operazioni eseguite sui dati.

i concetti rappresentati da Flows e Recipes in Dataiku offrono un approccio visuale e modulare per la definizione e l'esecuzione dei processi di data science e analisi dei dati, contribuendo a migliorare la trasparenza, la collaborazione, la tracciabilità e l'efficienza complessiva dei progetti di data science.

Data Sourcing

Il punto cruciale di qualsiasi progetto di Data Science o AI è l'acquisizione di dati di alta qualità da varie fonti. Dataiku semplifica notevolmente questo processo, consentendo agli utenti di connettersi e acquisire dati da una vasta gamma di fonti, tra cui:

- **Database Relazionali**: Dataiku supporta il recupero di dati da database relazionali, come MySQL, PostgreSQL, Microsoft SQL Server e molti altri. Gli utenti possono configurare facilmente connessioni per estrarre dati da tabelle specifiche o eseguire query personalizzate.
- **Sorgenti NoSQL**: È possibile acquisire dati da database NoSQL, come MongoDB o Cassandra, per gestire informazioni non strutturate o semi-strutturate.
- **File Locali e Remoti**: Dataiku consente di importare dati da file locali o situati su servizi cloud, come Dropbox o Amazon S3. Questo è utile per dati in formato CSV, Excel, JSON o altri.

- **Servizi Cloud**: La piattaforma offre connettori preconfigurati per i principali servizi cloud, come AWS, Azure e Google Cloud. Questi connettori semplificano l'accesso ai dati archiviati nelle piattaforme cloud.
- **Sorgenti Web**: Dataiku supporta il web scraping per l'acquisizione di dati da siti web. Gli utenti possono configurare raccolte di dati da pagine web specifiche o RSS feeds.
- **API Esterne**: La piattaforma può consumare dati da API esterne, consentendo l'integrazione di dati provenienti da servizi di terze parti.

Dataiku fornisce un'interfaccia intuitiva per configurare le connessioni alle fonti dati. Gli utenti possono definire le sorgenti, specificare i parametri di connessione e testare la connessione in tempo reale. Inoltre, Dataiku offre un ambiente di sviluppo sandbox per eseguire rapidamente prove e prototipi di acquisizione dati.

Una volta configurate le sorgenti dati, Dataiku offre strumenti avanzati per l'esplorazione e la comprensione dei dati. Gli utenti possono visualizzare campioni di dati, esaminare schemi e statistiche riassuntive e identificare eventuali problemi di qualità dati.

Dataiku semplifica notevolmente il processo di data sourcing, consentendo agli utenti di accedere a dati provenienti da una varietà di fonti in modo efficiente e organizzato. Ciò fornisce una solida base per il successo di progetti di Data Science e AI, garantendo che i dati siano pronti per la preparazione e l'analisi.

Data preparation

La preparazione dei dati è spesso la fase più laboriosa e critica in un progetto di Data Science o AI. Dataiku semplifica notevolmente questo processo, offrendo una serie di strumenti e funzionalità avanzate per la pulizia, la trasformazione e l'arricchimento dei dati. Ecco come Dataiku affronta la data preparation:

- **Data Cleaning**: Dataiku consente agli utenti di individuare e gestire i dati mancanti o errati. Puoi definire regole per la gestione dei dati mancanti, ad esempio, scegliendo di eliminarli o di sostituirli con valori appropriati. La piattaforma offre anche funzionalità per rilevare e correggere duplicati nei dati.
- **Data Transformation**: è possibile applicare trasformazioni complesse ai dati con facilità. ci sono diverse opzioni per attivare operazioni di trasformazione, tra cui il ridimensionamento delle variabili, la normalizzazione dei dati e la creazione di nuove feature attraverso la combinazione o l'estrazione di informazioni dalle colonne esistenti.
- **Data Enrichment**: per arricchire ulteriormente i dati, Dataiku consente di connettersi a fonti esterne. Ad esempio, puoi arricchire un dataset con dati demografici o meteorologici da fonti pubbliche o private.
- **Data Profiling**: la piattaforma offre strumenti di data profiling che forniscono una panoramica dettagliata dei dati, comprese statistiche, distribuzioni e insight sui valori univoci nelle colonne. Ciò aiuta gli utenti a comprendere meglio la qualità e la struttura dei dati.
- **Automazione delle Ricette**: è possibile creare "ricette" di preparazione dati che registrano le trasformazioni applicate. Queste ricette possono essere riutilizzate su nuovi dati o condivise con

altri membri del gruppo. Inoltre, Dataiku offre la possibilità di automatizzare le ricette, consentendo l'elaborazione automatica dei dati in batch o in tempo reale.

- **Esplorazione Interattiva**: gli utenti possono esplorare interattivamente i dati durante la fase di preparazione. Ciò significa che è possibile vedere immediatamente l'effetto delle trasformazioni applicate e apportare modifiche in tempo reale.

Tutto il processo di trattamento dei dati vede applicata una metodologia iterativa per la data preparation, consentendo agli utenti di eseguire rapidamente prove, apportare modifiche e ottimizzare il processo. Inoltre, la piattaforma mantiene una traccia completa delle trasformazioni applicate ai dati, garantendo la ripetibilità e la documentazione delle operazioni.

Data visualization

La visualizzazione dei dati è un aspetto cruciale nella comprensione e nella comunicazione dei risultati di un progetto di Data Science o AI. Dataiku offre una serie di strumenti e funzionalità per creare visualizzazioni ricche e informative dei dati. Ecco come Dataiku affronta la data visualization:

- **Dashboard Interattive**: gli utenti possono creare dashboard interattive personalizzate per visualizzare i risultati delle analisi. Puoi trascinare e rilasciare componenti come grafici, tabelle e filtri per costruire dashboard che forniscono una panoramica completa dei dati e dei risultati del modello.
- **Esplorazione dei Dati**: prima di creare visualizzazioni avanzate, Dataiku permette di esplorare i dati in modo interattivo. Puoi esaminare le distribuzioni delle variabili, identificare tendenze e outlier, e ottenere una visione dettagliata dei dati.
- **Libreria di Visualizzazioni**: Dataiku offre una libreria di visualizzazioni predefinite che coprono una vasta gamma di tipi di grafici e tabelle. Gli utenti possono selezionare rapidamente la visualizzazione più adatta per i loro dati, risparmiando tempo prezioso nella creazione di grafici personalizzati.
- **Personalizzazione delle Visualizzazioni**: se le visualizzazioni predefinite non soddisfano le esigenze specifiche del progetto, Dataiku consente la creazione di visualizzazioni personalizzate. Gli utenti possono scrivere codice Python o R per generare grafici personalizzati e integrarli nelle dashboard.
- **Esportazione e Condivisione**: le visualizzazioni create in Dataiku possono essere facilmente esportate in diversi formati, come immagini o file PDF, per la condivisione con altri membri del gruppo o la presentazione dei risultati. Inoltre, le dashboard possono essere condivise in modo interattivo con gli stakeholder del progetto.
- **Integrazione con Altre Piattaforme**: Dataiku può essere integrato con altre piattaforme di visualizzazione dei dati, consentendo agli utenti di sfruttare le proprie soluzioni preferite per la visualizzazione, se necessario. Un tipico esempio di integrazione possibile nel contesto del SIM è l'uso combinato di Dataiku con Tableau.
- **Monitoraggio delle Prestazioni**: nel contesto delle applicazioni AI e di Machine Learning, Dataiku offre strumenti per monitorare le prestazioni dei modelli in tempo reale. Gli utenti possono creare visualizzazioni per tracciare metriche come l'accuratezza del modello, il tempo di risposta e altro ancora.

In pratica viene resa disponibile una suite completa di strumenti per la data visualization, consentendo agli utenti di creare visualizzazioni efficaci e interattive per comunicare i risultati dei

progetti di Data Science e AI. Ciò favorisce una migliore comprensione dei dati e supporta la presa di decisioni informate.

Collaboration

La collaborazione è fondamentale in qualsiasi progetto di Data Science e AI, in quanto coinvolge spesso gruppi multidisciplinari e richiede un flusso di lavoro fluido. Dataiku mette a disposizione una serie di strumenti e funzionalità per agevolare la collaborazione all'interno delle organizzazioni. Il modello di supporto alla collaborazione include le seguenti funzionalità:

- **Progetti condivisi**: Dataiku consente agli utenti di lavorare su progetti condivisi, che possono includere l'accesso a set di dati, modelli, notebook e altro. Questo permette ai gruppi di collaborare in tempo reale, riducendo la duplicazione del lavoro e migliorando l'efficienza complessiva.
- **Controllo delle Versioni**: è disponibile un sistema di controllo delle versioni integrato che tiene traccia delle modifiche apportate ai progetti. Questo è fondamentale per garantire la coerenza e la tracciabilità delle attività di sviluppo e analisi.
- **Flussi di Lavoro condivisi**: gli utenti possono creare flussi di lavoro condivisi che includono diverse fasi del processo di Data Science, come l'acquisizione dei dati, la preparazione, l'addestramento dei modelli e altro. Questi flussi di lavoro possono essere eseguiti da membri del gruppo con le autorizzazioni appropriate.
- **Collaborazione Interfunzionale**: Dataiku favorisce la collaborazione tra gruppi diversi, inclusi data scientist, analisti, ingegneri dei dati e stakeholder aziendali. Questo è essenziale per garantire che le competenze di tutti siano sfruttate al massimo.
- **Commenti e Annotazioni**: gli utenti possono aggiungere commenti e annotazioni direttamente all'interno di progetti e dataset. Ciò favorisce la comunicazione all'interno del gruppo e la documentazione delle decisioni prese.
- **Notifiche e Allerte**: Dataiku consente di impostare notifiche e allerte per tenere traccia delle attività importanti all'interno del progetto. Ad esempio, è possibile ricevere una notifica quando un modello è pronto per la produzione o quando un flusso di lavoro è stato completato con successo.
- **Accesso basato su Ruoli**: Dataiku offre un sistema di gestione degli accessi basato su ruoli che consente di controllare chi può accedere e modificare i diversi componenti del progetto. Questo garantisce la sicurezza e la conformità con le normative aziendali.
- **Integrazione con Strumenti di Comunicazione**: Dataiku può essere integrato con strumenti di comunicazione come Slack o Teams per facilitare la comunicazione tra i membri del gruppo, consentendo loro di ricevere aggiornamenti e notifiche direttamente nelle chat di lavoro.

Dataiku pone una forte enfasi sulla collaborazione tra i gruppi di lavoro, consentendo loro di lavorare insieme in modo efficace e integrato. Questa collaborazione migliora la produttività, la qualità dei risultati e la condivisione delle conoscenze all'interno dell'organizzazione.

Machine Learning e Analytics

Dataiku offre un ampio set di strumenti e funzionalità per la creazione, l'addestramento e l'implementazione di modelli di machine learning e per l'analisi dei dati. Ecco come affronta il processo di machine learning e analytics:

- **Interfaccia Utente Intuitiva**: una delle caratteristiche distintive di Dataiku è la sua interfaccia utente intuitiva che consente a utenti di diverse competenze, inclusi data scientist e analisti non tecnici, di creare e addestrare modelli senza dover scrivere codice. Questo favorisce la democratizzazione dell'AI all'interno dell'organizzazione.
- **Libreria di Algoritmi**: include una vasta libreria di algoritmi di machine learning pronti per l'uso. Gli utenti possono selezionare dagli algoritmi più comuni e iniziarne l'uso immediato. Inoltre, è possibile integrare algoritmi personalizzati.
- **Flussi di Lavoro Visuali**: gli utenti possono creare flussi di lavoro visuali per orchestrare il processo di machine learning, che può includere l'acquisizione dei dati, la preparazione, l'addestramento del modello e la valutazione delle prestazioni. Questi flussi di lavoro sono altamente personalizzabili e consentono un controllo completo sul processo.
- **Automazione dell'esecuzione ML**: semplifica l'automazione ML attraverso funzionalità come l'autoML, che consente di selezionare rapidamente il miglior modello per un dato problema, e il monitoraggio delle prestazioni dei modelli in produzione.
- **Esecuzione Scalabile**: consente di sfruttare l'architettura distribuita di Apache Spark per eseguire il training di modelli su grandi dataset in modo scalabile ed efficiente. Ciò è fondamentale per progetti di machine learning su larga scala.
- **Valutazione delle Prestazioni**: fornisce strumenti per valutare le prestazioni dei modelli, inclusi indicatori chiave come l'accuratezza, la precisione e il recall. Questi strumenti aiutano gli utenti a comprendere quanto bene i loro modelli stanno facendo previsioni.
- **Implementazione in Produzione**: semplifica l'implementazione di modelli in produzione attraverso il supporto di API per il rilascio di modelli addestrati. Questo consente di integrare facilmente i modelli nei sistemi aziendali esistenti.
- **Analisi dei Dati Avanzata**: oltre al machine learning, Dataiku offre strumenti per l'analisi avanzata dei dati. Gli utenti possono eseguire query, aggregazioni e analisi statistica direttamente all'interno dell'ambiente Dataiku.
- **Coding**

Il codice svolge un ruolo fondamentale nell'analisi dei dati e nella creazione di modelli avanzati. Pertanto, la piattaforma offre diverse funzionalità legate al coding per soddisfare le esigenze degli sviluppatori e dei data scientist che preferiscono scrivere codice personalizzato. Le caratteristiche più rilevanti in questo contesto sono:

- **Integrazione con Linguaggi di Programmazione**: sono supportati diversi di linguaggi di programmazione, tra cui Python, R, SQL, Scala e altri. Gli utenti possono scrivere script personalizzati utilizzando il linguaggio che preferiscono e integrarli nei flussi di lavoro di Dataiku.
- **Scripting Flessibile**: gli utenti possono scrivere script personalizzati all'interno dell'ambiente Dataiku utilizzando un'interfaccia flessibile. Questo consente di eseguire analisi avanzate e manipolazioni dei dati utilizzando il proprio codice.
- **Integrazione con IDE Esterni**: per coloro che preferiscono utilizzare ambienti di sviluppo integrati (IDE) esterni come Jupyter Notebook o RStudio, Dataiku consente una integrazione con queste

soluzioni. È anche possibile connettersi a un IDE esterno e lavorare su script direttamente all'interno di questi ambienti.

- **Version Control**: Dataiku supporta il controllo di versione per il codice, consentendo agli utenti di tenere traccia delle modifiche apportate agli script nel corso del tempo. Questo è cruciale per la collaborazione e per garantire la riproducibilità dei risultati.
- **Pubblicazione di API**: gli utenti possono pubblicare facilmente API utilizzando il proprio codice all'interno di Dataiku. Ciò consente di esporre funzionalità personalizzate come servizi web che possono essere utilizzati in altri sistemi o applicazioni.
- **Estendibilità**: Dataiku è altamente estensibile, il che significa che è possibile integrare librerie e pacchetti personalizzati. Questo è utile quando si lavora con funzionalità specifiche del settore o modelli personalizzati. Inoltre, è sempre possibile attivare servizi esterni richiamabili tramite API.
- **Testing e Debugging**: Dataiku fornisce strumenti per testare e debuggare il codice personalizzato. Gli utenti possono eseguire test unitari e identificare e correggere eventuali errori.
- **Documentazione**: è possibile documentare il codice all'interno dell'ambiente Dataiku, fornendo spiegazioni, note e commenti per garantire che il codice sia comprensibile e riutilizzabile.
- **MLOps, Operationalization e Automazione**

L'attuazione degli aspetti di "operations" dei modelli (Operationalization) e la gestione delle operazioni machine learning (MLOps) presentano una criticità nell'implementazione di modelli di intelligenza artificiale in un ambiente di produzione. Dataiku si distingue per la sua capacità di semplificare e automatizzare questi processi, garantendo che i modelli siano pronti per essere utilizzati in scenari reali. Ecco come Dataiku affronta l'MLOps, l'"operationalization", l'automazione e l'uso di cluster Kubernetes e Apache Spark:

- **Deployment semplificato**: Dataiku offre un processo semplificato per il deployment dei modelli, consentendo agli utenti di passare agevolmente dalla fase di sviluppo alla produzione. È possibile distribuire i modelli su varie piattaforme, tra cui server dedicati, cloud o sistemi edge. Grazie all'integrazione con cluster Kubernetes, il deployment può essere automatizzato e orchestrato in modo efficace.
- **Versioning dei Modelli**: Dataiku supporta il versioning dei modelli, consentendo agli utenti di tenere traccia delle diverse iterazioni di un modello e di tornare a versioni precedenti se necessario.
- **Automazione dei Workflow di MLOps**: Dataiku automatizza gran parte dei compiti ripetitivi legati all'MLOps, come il monitoraggio delle prestazioni dei modelli in produzione, la gestione dei dati di input e output, e l'aggiornamento dei modelli quando nuovi dati sono disponibili. L'automazione può essere potenziata ulteriormente sfruttando la potenza di Apache Spark per l'elaborazione distribuita dei dati.
- **Integrazione con Strumenti Esterni**: Dataiku può essere integrato con strumenti di MLOps di terze parti, consentendo agli utenti di utilizzare le loro soluzioni preferite per la gestione dei modelli. L'integrazione con cluster Kubernetes consente una gestione centralizzata di container e risorse, semplificando l'interoperabilità con altre piattaforme.
- **Monitoraggio delle Prestazioni in Tempo Reale**: gli utenti possono monitorare le prestazioni dei modelli in tempo reale e ricevere avvisi in caso di degrado delle prestazioni o di anomalie. Apache Spark può essere utilizzato per l'analisi in tempo reale di grandi volumi di dati.

- **Gestione della Sicurezza**: Dataiku offre funzionalità avanzate per garantire la sicurezza dei modelli e dei dati, incluso il controllo degli accessi e la crittografia dei dati sensibili. L'uso di cluster Kubernetes può contribuire a rafforzare la sicurezza dell'infrastruttura.
- **Collaborazione e Governance**: la piattaforma consente una collaborazione efficace tra gruppi di data scientist, sviluppatori e responsabili delle operazioni. Le funzionalità di governance garantiscono che i modelli siano conformi alle normative e alle politiche aziendali.
- **Ritorno di Investimento (ROI) Misurabile**: fornisce metriche e dashboard per valutare il ROI dei modelli implementati, consentendo alle aziende di valutare l'efficacia degli investimenti in AI.
- **Pianificazione degli Aggiornamenti**: consente agli utenti di pianificare gli aggiornamenti dei modelli in modo da garantire un passaggio fluido tra versioni. L'uso di cluster Kubernetes semplifica la gestione delle operazioni di aggiornamento.
- **Retraining Automatico**: la piattaforma offre funzionalità di "re-training" automatico dei modelli, garantendo che rimangano accurati nel tempo.

Automazione, Schedulazione, uso di Kubernetes e Apache Spark

L'automazione è un elemento chiave nella gestione dei processi di data science e machine learning, e Dataiku offre un solido supporto per l'automazione dei workflow, la schedulazione delle attività e l'utilizzo dei cluster Kubernetes e Apache Spark. Ecco come Dataiku affronta questi aspetti:

- **Automazione dei Workflow**: consente agli utenti di automatizzare i workflow di data preparation, model building e deployment. È possibile creare flussi di lavoro complessi che includono molteplici passaggi e attività. Questi workflow possono essere eseguiti in modo automatico, riducendo al minimo l'intervento manuale e garantendo la coerenza dei processi.
- **Schedulazione delle Attività**: offre una schedulazione flessibile delle attività. Gli utenti possono pianificare l'esecuzione di workflow e task in base a una programmazione temporale specifica o a eventi scatenanti. Ad esempio, è possibile pianificare l'aggiornamento periodico dei modelli o l'esecuzione di analisi ricorrenti.
- **Utilizzo di Kubernetes**: integra nativamente la gestione di cluster Kubernetes. Questo consente di distribuire e orchestrare facilmente i container Docker contenenti i modelli e i processi di data science. Kubernetes offre scalabilità orizzontale, gestione dei failover e una distribuzione affidabile delle risorse computazionali. Dataiku semplifica la configurazione e la gestione dei cluster Kubernetes, consentendo agli utenti di sfruttare al massimo questa tecnologia.
- **Utilizzo di Apache Spark**: la piattaforma consente una piena integrazione con Apache Spark, un framework di elaborazione dati distribuito. Apache Spark è ideale per l'elaborazione di dati su larga scala e per l'addestramento di modelli di machine learning su dataset complessi. Con Dataiku, è possibile utilizzare Apache Spark per l'elaborazione parallela dei dati e per eseguire analisi avanzate su grandi volumi di informazioni.
- **Monitoraggio e Log**: fornisce strumenti di monitoraggio avanzati per tenere traccia delle attività automatizzate. Gli utenti possono accedere a log dettagliati e dashboard di monitoraggio per verificare lo stato delle attività e identificare eventuali problemi. Questo è fondamentale per garantire l'affidabilità dei processi automatizzati.
- **Gestione delle Risorse**: la piattaforma consente agli utenti di assegnare e gestire le risorse computazionali in modo efficiente. È possibile stabilire limiti di utilizzo delle risorse e monitorare l'allocazione delle CPU e della memoria nei cluster Kubernetes e durante l'esecuzione di processi Apache Spark.

- **Sicurezza e Governance**: Dataiku garantisce che i processi automatizzati rispettino le politiche di sicurezza e governance aziendali. Le autorizzazioni e i controlli di accesso vengono applicati automaticamente, evitando accessi non autorizzati o modifiche non autorizzate ai dati e ai modelli.

Gestione utenze, ruoli, accessi e profilazione dell'uso

La gestione degli utenti e dei ruoli in Dataiku è una parte fondamentale per garantire un utilizzo sicuro ed efficiente dei servizi offerti agli utenti e della sostenibilità in un contesto enterprise della piattaforma stessa. Dataiku offre un sistema flessibile per la creazione e l'amministrazione degli account degli utenti, consentendo di assegnare ruoli e permessi specifici in base alle esigenze dell'organizzazione.

Creazione e Amministrazione degli Utenti

un amministratore della piattaforma può creare nuovi account utente direttamente dall'interfaccia amministrativa di Dataiku. Qui è possibile specificare informazioni come nome, cognome ed e-mail dell'utente. Inoltre, è possibile definire il ruolo iniziale dell'utente, che determinerà i permessi di accesso e le funzionalità disponibili.

Assegnazione di Ruoli e Permessi

È disponibile una gamma di ruoli predefiniti che coprono le esigenze comuni delle organizzazioni. Questi ruoli includono amministratori, sviluppatori, analisti e collaboratori. Tuttavia, è possibile personalizzare ulteriormente i ruoli per adattarli alle esigenze specifiche dell'organizzazione e per le necessità anche temporanee che possono sorgere nel contesto dei progetti collaborativi con i vari stakeholders del SIM.

Gli amministratori hanno il massimo controllo sulla piattaforma e possono assegnare ruoli e permessi agli utenti in modo granulare. Per esempio, è possibile definire chi può creare progetti, chi può accedere ai dati sensibili o chi può eseguire analisi avanzate.

Gestione degli Accessi e Integrazione con Sistemi IAM

Dataiku è progettato per essere compatibile con i sistemi di gestione delle identità e degli accessi (IAM) esistenti delle organizzazioni. Questo consente un'efficace gestione degli accessi e una maggiore sicurezza.

L'integrazione con i sistemi IAM consente di sincronizzare gli utenti e i ruoli tra Dataiku e il sistema IAM aziendale. In questo modo, quando un nuovo utente viene aggiunto o rimosso dal sistema IAM, le sue autorizzazioni su Dataiku vengono automaticamente aggiornate, riducendo il rischio di errori o accessi non autorizzati.

I meccanismi di autenticazione possono essere:

- **Locale**: consente agli utenti di accedere utilizzando le credenziali nome utente/password memorizzate direttamente in DSS.
- **SSO**: delega l'autenticazione a un gestore di identità utilizzando i protocolli SAML o OpenID Connect.
- LDAP: consente agli utenti di accedere utilizzando le credenziali nome utente/password memorizzate in un server LDAP.
- **PAM**: consente agli utenti di accedere utilizzando le credenziali nome utente/password tramite PAM (Pluggable Authentication Modules).
- **Personalizzato**: fornisce la flessibilità di creare un autenticatore personalizzato per autenticare gli utenti utilizzando le credenziali nome utente/password.

Le possibili basi dati delle utenze supportate sono:

- **SSO**: il protocollo SSO scelto restituisce un'identità (token ID, asserzione SAML, ecc.) dall'autenticatore SSO che viene utilizzato per fornire o sincronizzare l'utente in DSS durante l'accesso.
- **LDAP**: Recupera le informazioni utente dal server LDAP e le esegue il provisioning/sincronizza come utente DSS.
- **Azure AD**: Recupera le informazioni utente da Azure AD e le fornisce/sincronizza come utente DSS.
- PAM: funzionalità limitata; PAM può fornire/sincronizzare un utente solo con le informazioni sul nome utente e solo per gli utenti autenticati utilizzando PAM.
- **Personalizzato**: consente alla creazione di un fornitore utente personalizzato di convertire l'identità ottenuta durante l'autenticazione in un utente DSS.

Profilazione dell'Uso della Piattaforma

Dataiku offre strumenti di profilazione avanzati che consentono di monitorare e analizzare l'uso della piattaforma. Questi strumenti forniscono informazioni dettagliate sull'attività degli utenti, sui progetti in corso e sulle risorse utilizzate.

La profilazione dell'uso della piattaforma è utile per identificare tendenze, ottimizzare le risorse e garantire che la piattaforma sia utilizzata in modo efficiente. per esempio, è possibile identificare i progetti più intensivi in termini di risorse e assegnare risorse aggiuntive quando necessario.

Profili Legati all'Uso della Piattaforma in Funzione del Costo delle Sottoscrizioni

Dataiku offre opzioni di sottoscrizione che consentono alle organizzazioni di scegliere il livello di servizio più adatto alle proprie esigenze. Ogni livello di sottoscrizione ha un costo associato e offre una serie di funzionalità e risorse.

In base alla sottoscrizione sottoscritta, gli utenti possono accedere a diverse funzionalità e risorse. Per esempio, un utente con una sottoscrizione di livello superiore avrà accesso a funzionalità avanzate come l'addestramento di modelli avanzati, mentre un utente con una sottoscrizione di livello inferiore potrebbe avere accesso solo a funzionalità di base. La gestione dei profili legati alle sottoscrizioni consente alle organizzazioni di massimizzare il valore della piattaforma e di controllare i costi in base alle esigenze specifiche.

1.1.6.1.2 Integrazioni nella piattaforma per il MASE

Nell'ambito del progetto MASE il modulo DataIKU DSS verrà integrato con i moduli PaaS Big Data & AI secondo lo schema architetturale mostrato in figura, che può essere dettagliata nelle seguenti integrazioni:

- **Event Message**
- **Spark**
- **Metastore (HCatalog)**
- **Delta Lake**
- **AI Platform**
- **IAM SSO**
- **API**

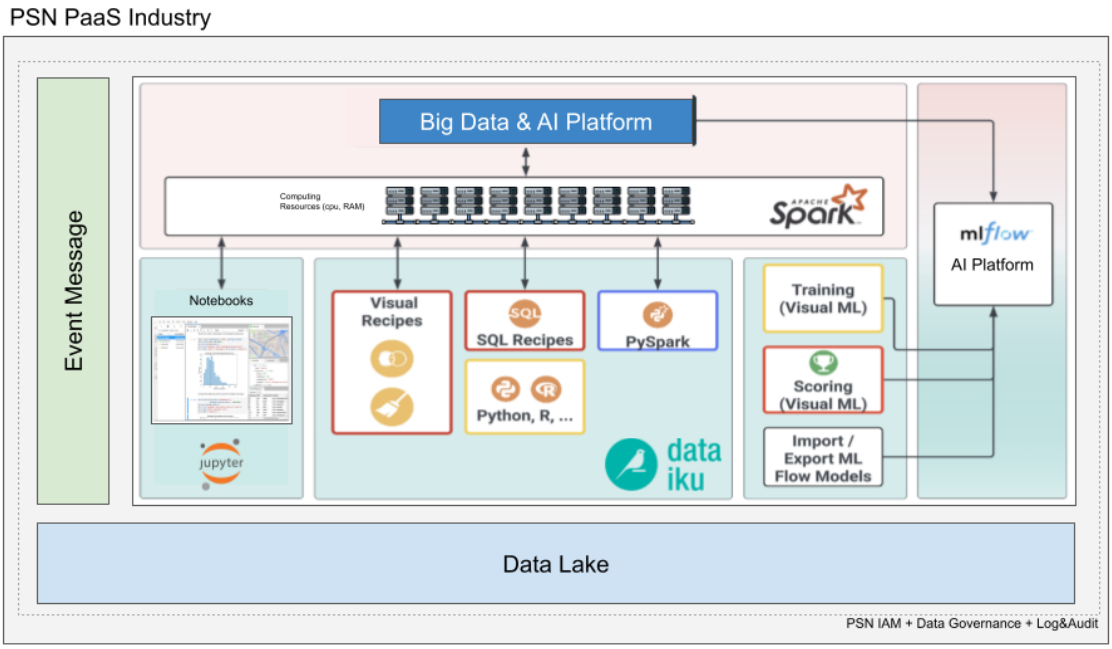


Figura 59 – Architettura Tecnica e Integrazioni

Di seguito verrà dettagliato in che cosa consiste l'integrazione nella Data Platform.

Data Lake

Il modulo PaaS Data Lake espone una interfaccia pienamente compatibile con il protocollo S3.

DataIKU DSS può interagire con il modulo Data Lake per:

- Leggere e scrivere set di dati
- Leggere e scrivere cartelle gestite

S3 è un servizio di archiviazione di oggetti: permette di creare "bucket" in grado di archiviare contenuti binari arbitrari e metadati testuali sotto una chiave specifica, univoca nel contenitore.

Sebbene non sia tecnicamente un file system gerarchico con cartelle, sottocartelle e file, tale comportamento può essere emulato utilizzando chiavi contenenti `/`. Ad esempio, puoi archiviare i tuoi registri giornalieri utilizzando chiavi come `2015/01/24/app.log`. Il protocollo S3 consente di elencare tutti gli oggetti con un prefisso specifico, ad esempio `2015/01/` e i client S3 possono "esplorare" i bucket attraverso questo meccanismo.

DataIKU DSS utilizza lo stesso meccanismo simile a un file system quando si accede a S3: quando l'utente specifica un bucket, può esplorarlo per trovare rapidamente il tuo set di dati oppure può impostare il prefisso mediante il quale DSS può identificare i set di dati. I set di dati possono essere in uno dei seguenti formati supportati:

Format	Read	Write
Delimited values (CSV, TSV, ...)	yes	yes
Fixed width	yes	no
Excel (from Excel 97)	yes	only via export
Avro	yes	yes
Custom format using regular expression	yes	no
XML	yes	no
JSON	yes	no
ESRI Shapefiles	yes	no
MySQL Dump	yes	no
Apache Combined log format	yes	no

N.B. i formati basati su file possono essere letti compressi: ZIP, GZIP, BZ2.

Sono inoltre supportati i seguenti formati specifici del mondo Hadoop/Spark:

Format	Read	Write
Parquet	yes	yes
Hive ORCFile	yes	yes
Delta Lake	yes	no

Per creare una connessione S3, è necessario disporre di una coppia di chiavi (chiave di accesso e chiave segreta). L'integrazione del server DataIKU DSS con il PaaS Big Data & AI permette di recuperare automaticamente le credenziali dall'ambiente.

L'accesso specificato in una connessione Data Lake deve avere le seguenti autorizzazioni:

- Per leggere i dati da un bucket, è necessario disporre almeno delle autorizzazioni di elenco e lettura su quel bucket:
  - s3:ListBucket arn:aws:s3:::examplebucket
  - s3:GetObject arn:aws:s3:::examplebucket/*

- Per scrivere dati su un bucket, è necessario disporre anche delle autorizzazioni di scrittura, eliminazione e interruzione del caricamento in più parti su quel bucket:
  - s3:ListBucket arn:aws:s3:::examplebucket
  - s3:GetObject arn:aws:s3:::examplebucket/*
  - s3:PutObject arn:aws:s3:::examplebucket/*
  - s3:DeleteObject arn:aws:s3:::examplebucket/*
  - s3:AbortMultipartUpload arn:aws:s3:::examplebucket/*
- Non obbligatorio, ma l'autorizzazione all'elenco dei bucket ti consente di scegliere un bucket da un elenco a discesa invece di digitarne manualmente il nome durante la creazione di un set di dati:
  - s3:ListAllMyBuckets arn:aws:s3:::*
- Inoltre, non è obbligatorio, l'autorizzazione per la localizzazione del bucket può migliorare le prestazioni consentendo a DataIKU DSS di accedere a un bucket tramite il suo endpoint preferito:
  - s3:GetBucketLocation arn:aws:s3:::*

Per ulteriori dettagli su come creare policy e aggiungere queste autorizzazioni, consultare la documentazione MinIO qui:

https://min.io/docs/minio/linux/administration/identity-access-management/policy-based-access-control.html

Dopo aver creato la connessione S3, puoi creare/manipolare i tuoi set di dati.

Dal DataIKU DSS e in particolare dall'elenco del flusso o dei set di dati, fare clic su Nuovo set di dati > S3.

- Seleziona la connessione in cui si trovano i tuoi file
- Se disponibile, seleziona il bucket (elencandolo o inserendolo)
- Fare clic su "Sfoglia" per individuare i file.

Event Message

Il modulo Event Message del PaaS Big Data PSN espone una interfaccia pienamente compatibile con il protocollo Kafka.

DataIKU DSS può sfruttare i topic Kafka come endpoint di streaming.

Per leggere o scrivere da topic Kafka, è necessaria una connessione a un cluster Kafka. La connessione è definita da un elenco di server bootstrap e impostazioni di sicurezza. DSS supporta i protocolli PLAINTEXT e SASL per la comunicazione con i broker e l'autenticazione Kerberos che è un caso speciale di SASL. Poiché entrambi i protocolli non sono crittografati in modo nativo, di solito richiedono l'attivazione di SSL. In questo caso, nella connessione vengono impostati anche il truststore e il keystore che contengono i certificati. Tieni presente che se il tuo server DataIKU DSS è in esecuzione sul PaaS Big Data & AI stesso, DataIKU DSS non richiederà alcuna configurazione aggiuntiva per poter funzionare.

I messaggi Kafka comprendono una chiave, un valore e un timestamp. Sia la chiave che il valore sono trattati dai broker Kafka come dati binari ed è dovere del producer e del consumer del messaggio leggere e scrivere questi dati binari.

Quando si converte un messaggio in una riga, DataIKU DSS legge prima la chiave (se è impostato un formato), quindi il valore. Le colonne presenti nella chiave hanno la precedenza sulle colonne presenti nel valore. Allo stesso modo, la colonna timestamp (se definita) ha la precedenza sulle colonne presenti in chiave o valore.

Tutti i messaggi in un topic Kafka hanno un timestamp, solitamente impostato dal broker quando il messaggio viene aggiunto al topic. Quando il nome di una colonna di timestamp è impostato nelle impostazioni dell'endpoint di streaming, dopo la lettura in DataIKU DSS, il valore del timestamp viene recuperato in quella colonna. Al contrario, se la casella di controllo fornisci timestamp è spuntata, durante la scrittura da parte di DSS il valore della colonna timestamp viene utilizzato per il timestamp del messaggio (se il broker consente l'impostazione del timestamp).

Spark

DataIKU DSS supporta l'utilizzo del motore Spark per la distribuzione dei carichi di lavoro e l'utilizzo delle risorse di un cluster. Il modulo PSN Big Data Processing è basato sul motore opensource Apache Spark.

In particolare, l'installazione DataIKU DSS integrata sul PaaS Big Data & AI del PSN è configurata in modo da poter eseguire recipes di tipo Spark facendo leva sulle risorse computazionali del cluster di processing.

Tale modalità di utilizzo consente di sfruttare la completa integrazione con tutti gli altri moduli del PaaS Industry PSN.

Metastore (HCatalog)

DSS supporta l'utilizzo di un metastore compatibile HiveServer2 come quello integrato nei servizi PaaS Big Data & AI del PSN.

In questo modo sarà possibile far leva sui database e sulle tabelle "virtuali" disponibili sul modulo Data Governance.

Delta Lake

Il servizio Data Lake del PaaS Big Data & AI del PSN utilizza Delta Lake come formato di default per la memorizzazione dei dati. Delta Lake rappresenta l'evoluzione dello storage dati, permette di preservare l'integrità dei dati originali senza sacrificare le prestazioni e l'agilità richiesti per le applicazioni di analisi in tempo reale, intelligenza artificiale (AI) e machine learning (ML).

Le caratteristiche più importati di Delta Lake sono:

Formato aperto: Un Delta Lake utilizza il formato Apache Parquet open source ed è pienamente compatibile con il motore di analisi unificata Apache Spark per operazioni potenti e flessibili.

Transazioni ACID: Delta Lake supporta le transazioni ACID (Atomicity, Consistency, Isolation, Durability) per i carichi di lavoro dei Big Data. Acquisisce tutte le modifiche apportate ai dati in un log delle transazioni serializzato, proteggendo l'integrità e l'affidabilità dei dati e fornendo audit trail completi e precisi.

Supporto per audits, rollbacks, o reproduce: il log delle transazioni di Delta Lake fornisce un record master di ogni modifica apportata ai dati consentendo di ricreare lo stato preciso di un set di dati in qualsiasi momento. Il controllo delle versioni dei dati supporta la riproduzione completa di analisi ed esperimenti.

Applicazione dello schema: Delta Lake protegge la qualità e l'uniformità dei dati con una solida applicazione dello schema, garantendo che i tipi di dati siano corretti e completi e impedendo il danneggiamento dei processi critici da parte di dati non validi.

Unificazione, aggiornamento, eliminazione: Delta Lake supporta operazioni Data Manipulation Language (DML) per compliance e casi d'uso complessi come lo streaming degli upsert, change-data capture, slowly-changing-dimension (SCD) e altro ancora.

DataIKU DSS può leggere i file Delta Lake ed elaborarli utilizzando Spark o qualsiasi altro tipo di recipe.

Delta Lake è un formato di dati open-source che è stato progettato per fornire una solida infrastruttura per l'archiviazione e la gestione dei dati in ambienti di data lake. È costruito su Apache Parquet per integrarsi perfettamente con l'ecosistema Big Data (es: Hadoop e Spark) e offre una serie di caratteristiche tecniche che lo distinguono.

Una delle principali caratteristiche di Delta Lake è la gestione transazionale dei dati, che consente transazioni atomiche, coerenti, isolate e durature. Questo garantisce l'integrità dei dati e la possibilità di eseguire operazioni di scrittura e lettura simultaneamente senza conflitti. Delta Lake supporta anche schemi evoluti, consentendo l'aggiunta, la modifica e la rimozione di colonne senza dover riscrivere i dati esistenti. Inoltre, offre un sistema di versioning che consente di tracciare le modifiche ai dati nel tempo e di eseguire il rollback a versioni precedenti. Queste caratteristiche tecniche rendono Delta Lake una scelta potente per la gestione dei dati in ambienti complessi e altamente distribuiti.

Le librerie per la gestione di questo formato sono sviluppate da una vasta community gestita da Linux Foundation e tra i maggiori contributor c'è Databricks che è anche tra i principali committers del progetto Apache Spark.

AI Platform

Il ciclo di vita degli esperimenti e dei modelli AI è gestito dal modulo PaaS AI del PSN tramite il software opensource MLFlow.

È possibile importare modelli MLFlow in DataIKU DSS come modelli nativi. Questo consente di poter beneficiare di tutte le capacità di gestione AI di DataIKU DSS sui modelli MLFlow creati in AI Platform. In particolare, è possibile:

- Assegnare un punteggio ai set di dati utilizzando una ricetta di punteggio
- Distribuzione del modello in un bundle su un nodo di automazione. Vedi Distribuzioni e bundle di produzione
- Distribuzione del modello per il punteggio in tempo reale, utilizzando il nodo API
- Gestione di più versioni dei modelli
- Valutazione delle prestazioni di un modello di classificazione o regressione su un set di dati etichettato, comprese tutte le schermate dei risultati
- Confronto di più modelli o più versioni del modello, utilizzando i confronti dei modelli
- Analisi delle prestazioni e valutazione dei modelli su altri set di dati
- Analisi della deriva sul modello MLflow
- Governare il modello MLflow utilizzando il Govern Node

L'importazione dei modelli MLflow viene eseguita:

- tramite l'API
- utilizzando l'azione "Distribuisci" disponibile per i modelli nelle esecuzioni del monitoraggio degli esperimenti

IAM SSO

DataIKU DSS è integrato con il sistema IAM di PaaS Industry PSN. In particolare, il SSO è implementato tramite il protocollo OpenID Connect (OIDC).

Per meglio comprendere i termini dell'integrazione partiamo da un glossario:

- OIDC: OpenID Connect
- IDP: un provider di identità è un'entità di sistema che crea, mantiene e gestisce le informazioni sull'identità per i principali e fornisce anche servizi di autenticazione alle applicazioni che fanno affidamento all'interno di una federazione o rete distribuita
- Utente finale: l'utente finale è l'entità per la quale richiediamo informazioni sull'identità. Nel nostro caso, è l'utente DSS che deve effettuare il login per accedere a DSS.
- Client OIDC: chiamato anche Relying party RP nelle specifiche OIDC, il client OIDC è l'applicazione che si basa sull'IDP per autenticare l'utente finale. Nel nostro caso si tratta di DSS.
- Token ID: simile a una carta d'identità o a un passaporto, contiene molti attributi o affermazioni richiesti sull'utente. Questo token viene quindi utilizzato da DSS per mappare le attestazioni a un

utente DSS corrispondente. Firmato digitalmente, il token ID può essere verificato dai destinatari previsti (DSS).

- Attestazione: nel contesto DSS, le attestazioni sono coppie nome/valore che contengono informazioni su un utente.
- Ambito: nel contesto di OIDC, l'ambito fa riferimento a una serie di attestazioni di cui il client OIDC ha bisogno. Esempio: posta elettronica
- Codice di autorizzazione: durante il protocollo OIDC, il codice di autorizzazione viene generato dall'IDP e inviato all'utente finale, che lo trasmette al client OIDC. Viene quindi utilizzato dal client OIDC, che invia il codice di autorizzazione all'IDP e riceve in cambio un token ID. L'utilizzo di un codice di autorizzazione intermedio consente all'IDP di incaricare il client OIDC di autenticarsi per recuperare il token ID.
- Client confidenziale: un client OIDC con la capacità di scambiare il codice di autorizzazione con un token ID in un canale posteriore protetto. Questo è il caso del DSS.
- Client pubblico: un client OIDC non è in grado di archiviare il segreto in modo sicuro e deve scambiare il codice di autorizzazione con un token ID in un canale pubblico. DSS non è un cliente pubblico.
- PKCE: Proof Key for Code Exchange è un'estensione del protocollo OIDC, per consentire ai client pubblici di scambiare il codice di autorizzazione in un canale pubblico.

Funzionalità OIDC supportate da DSS

- authorization code grant flow
- simple string claims in the ID token
- non encrypted or signed authentication requests
- ID token signed with RSA or EC
- DSS dietro proxy
- response mode supported: query or fragment
- token endpoint auth method supported: client secret basic or client secret POST
- confidential OIDC client only

Flusso di integrazione:

1) Quando l'utente finale tenta di accedere a DataIKU DSS e non è ancora autenticato, DataIKU DSS lo reindirizzerà all'IDP. L'URL utilizzato sarà l'endpoint di autorizzazione dell'IDP, che contiene alcuni parametri GET specifici della configurazione DataIKU DSS.
2) L'IDP convaliderà i parametri GET e presenterà una pagina di accesso all'utente. Il percorso di autenticazione ora dipende dalle capacità del tuo IDP. A volte, quando è già effettuato l'accesso dal lato IDP, la pagina di accesso viene saltata e l'utente potrebbe non vedere il reindirizzamento all'IDP.
3) L'IDP ha autenticato l'utente finale e reindirizzerà l'utente a DataIKU DSS con un codice di autorizzazione. A seconda della configurazione del client OIDC nel tuo IDP, il codice può essere passato tramite i parametri di query o il frammento.
4) Il front-end di DataIKU DSS analizzerà e invierà i parametri, incluso il codice di autorizzazione, al backend DataIKU DSS.

5) Il backend DataIKU DSS scambierà questo codice di autorizzazione con un token di accesso, chiamando l'endpoint del token con le credenziali precedentemente configurate nelle impostazioni DataIKU DSS SSO. In caso di esito positivo, l'IDP restituirà un token ID corrispondente all'utente finale.
6) DataIKU DSS utilizza il token ID per associare l'utente finale a un utente DataIKU DSS. L'impostazione della mappatura fa parte della configurazione SSO di OIDC.
7) DataIKU DSS crea una sessione utente corrispondente all'utente DataIKU DSS. A questo punto, l'OIDC è completato e la sessione utente è indipendente dal protocollo di autenticazione utilizzato.

1.1.6.1.3 Infrastruttura

Il Data&AI Workflow prevede sia componenti PAAS, ospitati sulla Container Platform del PSN, sia componenti residenti su Virtual Machine (DataIku DSS).

Per ogni Paas erogato all'interno del PSN verranno garantiti:

- La disponibilità dei servizi PaaS sarà garantita (secondo i parametri RTO e RPO indicati nel piano di Continuità Operativa) avvalendosi della possibilità di replica del servizio tra due DC della stessa Region.
- Supporto alla replicazione del servizio su scala geografica (multi-Region)
- Funzionalità di backup trasparente per gli utenti finali
- Esposizione metriche di vario tipo per l'integrazione con sistemi di monitoraggio e notifica
- Possibilità di deployare la piattaforma in modalità automatica

Sono altresì previste macchine virtuali per installazione dei tool client come DataIKU DSS.

1.1.6.2 AI PLATFORM

1.1.6.2.1 Architettura tecnica

Di seguito il disegno di alto livello della soluzione:

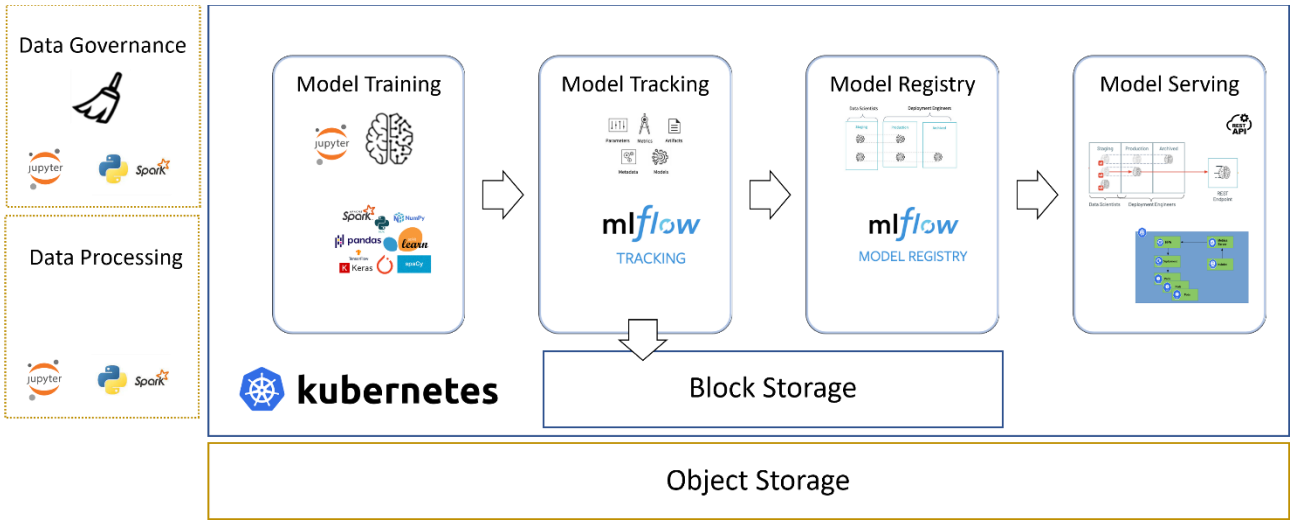


Figura 60 - Soluzione di alto livello dell'AI Platform

DI seguito una descrizione dei principali componenti delineati nella soluzione.

Jupyter Lab
Il nome del progetto Jupyter deriva dai tre linguaggi di programmazione di base, Julia, Python e R.

È un'applicazione basata sul modello client-server dell'organizzazione no profit Progetto Jupyter fondata nel 2015. Permette la creazione e la condivisione di documenti web nel formato JSON, che seguono uno schema e una lista ordinata di celle input/output. Queste celle offrono lo spazio per codice scritto direttamente nei linguaggi di programmazione supportati, testi descrittivi in linguaggio markdown (che permette la formattazione del testo in modo semplice e struttura), formule matematiche ed equazioni o contenuti multimediali.

L'elaborazione funziona su un'applicazione client basata sul web che si avvia con un browser standard. È sufficiente che sul sistema sia installato e venga eseguito anche il server del Notebook Jupyter. I documenti Jupyter creati si possono esportare come documenti HTML, PDF, Markdown o Python o in alternativa si possono condividere con altri utenti tramite e-mail, Dropbox, GitHub o il proprio Notebook Jupyter.

I due componenti centrali di Notebook Jupyter sono un set di diversi kernel (interpreti) e la dashboard. I kernel sono piccoli programmi che elaborano richieste ("request") specifiche nel linguaggio e reagiscono con relative risposte. Un kernel standard è IPython, un interprete della riga di comando che permette di lavorare con Python. Oltre 50 kernel forniscono supporto per altri linguaggi come C++, R, Julia, Ruby, JavaScript, CoffeeScript, PHP o Java. La dashboard serve da una parte come interfaccia di gestione per i singoli kernel e dall'altra come centrale per la creazione di nuovi documenti Notebook o per aprire progetti già esistenti. Notebook Jupyter è open source con licenza BSD modificata.

MLflow
MLflow è la piattaforma open source per la gestione del ciclo di vita dei modelli di machine learning. Esso semplifica le complesse procedure che gli sviluppatori devono seguire per realizzare l'apprendimento automatico.

A parte i problemi standard nello sviluppo del software, lo sviluppo dell'apprendimento automatico introduce una serie di ostacoli aggiuntivi come il saper utilizzare molte librerie dedicate allo sviluppo dei modelli, gestire tanti parametri personalizzabili per identificare il tuning corretto su un particolare task di ML, garantire la riproducibilità dell'esperimento, identificare un processo snello per portare un modello in produzione e renderlo disponibile all'utente finale.

Il tool ML Flow riesce a superare tutti questi ostacoli, attraverso una piattaforma open source per la gestione del ciclo di vita di ML che include: sperimentazione, riproducibilità, distribuzione e un unico registro di modelli.

Principalmente MLflow usa due moduli:

MLflow Tracking, che registra e tiene traccia delle metriche e degli artefatti (modelli più dipendenze dei medesimi) del processo di training. Gli artefatti verranno memorizzare nel Data Lake.

MLflow Model Registry archivia i modelli in un registro centralizzato, mettendo a disposizione sia un set di API che un'interfaccia utente per gestire in modo collaborativo l'intero ciclo di vita di un modello (controllo delle versioni del modello, transizioni di fase, ad esempio dalla gestione temporanea alla produzione e annotazioni).

Infine, ML Flow oltre a supportare diversi linguaggi di programmazione, inclusi Python e R, può tracciare, salvare e confrontare rapidamente diverse versioni di modelli in maniera semplice e intuitiva.

La soluzione prevede i seguenti servizi:

- Model Serving: la piattaforma facilita, attraverso un tool di model serving, la distribuzione dei modelli ML su larga scala negli ambienti di produzione. In altre parole, offre un servizio REST API di inferenza con bassa latenza e ad alte prestazioni.
- DBMS Metadata: il tool MLflow conserva tutti i metadati prodotti dagli esperimenti in un DB relazionale in modo da tenere traccia di tutti i flussi che hanno portato allo sviluppo di un determinato modello di ML.
- Object Storage: il tool MLflow conserva tutti i modelli sviluppati con le relative dipendenze su un object storage al fine di facilitarne il successivo processo di model serving in produzione. Per questo motivo il servizio ha come prerequisito l'attivazione del servizio Data Lake, del catalogo dei servizi del PaaS Big Data.
- Block Storage: il tool MLflow conserva tutti i metadati relativi agli esperimenti (per la produzione dei modelli) su un block storage al fine di facilitarne la riproducibilità, il versioning, transizioni di fase, e annotazioni di tali modelli.

BentoML

BentoML è una piattaforma open source per il deployment e la gestione di modelli di machine learning in produzione. Fornisce un'infrastruttura completa per l'organizzazione, il monitoraggio e la distribuzione di modelli ML in ambienti di produzione.

Una delle caratteristiche principali di BentoML è la sua flessibilità nel supportare diversi framework di machine learning, come TensorFlow, PyTorch, scikit-learn e molti altri. Questo significa che è possibile confezionare i modelli utilizzando un qualsiasi framework e BentoML si occuperà della creazione di un'API pronta per l'uso che può essere facilmente distribuita.

BentoML semplifica anche il processo di creazione di container Docker per il deployment dei modelli. È possibile confezionare un modello con le sue relative dipendenze in un container Docker autonomo e distribuirlo facilmente su diverse piattaforme senza dover gestire manualmente le complessità di configurazione.

Inoltre, BentoML offre funzionalità di monitoraggio per tenere traccia delle prestazioni dei modelli in produzione. È possibile raccogliere e visualizzare metriche chiave, come l'accuratezza del modello, i tempi di risposta delle richieste e altro ancora. Queste informazioni sono fondamentali per il monitoraggio e l'ottimizzazione continua dei modelli.

BentoML supporta anche la gestione delle versioni dei modelli. È possibile creare, tenere traccia e passare tra diverse versioni dei tuoi modelli per supportare il processo di sviluppo e deployment iterativo.

1.1.6.3 VIRTUAL ASSISTANT

1.1.6.3.1 Architettura tecnica

Il Virtual Assistant permette di integrare in un portale Web un modulo di sintesi vocale con lo scopo di eseguire dei comandi con la sola voce. L'architettura del modulo è mostrata nella figura seguente:

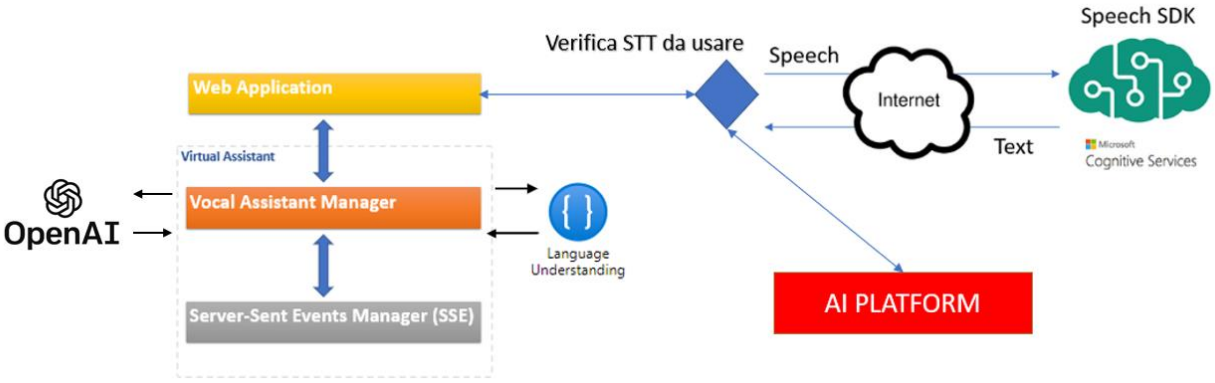


Figura 61 - Architettura tecnica Virtual Assistant

- La Web Application è il componente esterno che utilizza il Virtual Assistant e comunica con lo speech to text.
- Il Vocal Assistant Manager è il componente di Backend del Virtual Assistant che riceve riconosce la trigger work e interpreta il testo in comandi vocali sfruttando il servizio di comprensione del testo.
- Il modulo di BackEnd SSE Manager riceve l'interpretazione del comando ed esegue l'indirizzamento dei messaggi verso il canale corretto sulla base della natura del comando.

DESCRIZIONE DEI FLUSSI

Su un frontend è presente un tool di audio detection che rimane sempre in ascolto aprendo una socket; il tool individua le espressioni di interesse da sottoporre ad analisi a seguito del riconoscimento della triggerword "SIM" che deve essere pronunciata all'inizio del comando vocale.

L'audio del comando vocale in esame viene sottoposto a speech-to-text; la trascrizione testuale del comando viene dunque analizzata da un motore di intelligenza artificiale, il cui modello è stato progettato per riconoscere se l'espressione contiene una domanda o un comando; effettuata questa distinzione, il modello prosegue distinguendo – a seconda dei casi – la natura della richiesta e le informazioni da cercare o la natura del comando vocale (intento) e le entità contenute nello stesso.
Il file json contenente le dette informazioni estratte dal testo viene inviato verso il Server-Sent Events (SSE) del Virtual Assistant; all'ingresso è presente un layer costituito da un dispatcher che, a seconda della richiesta o della natura del comando, smista i json verso il canale SSE del Dossier Manager per

i comandi relativi a navigazione/filtraggio sul frontend di Dossier Manager oppure allo stesso canale SSE del Virtual Assistant per le richieste e gli altri comandi (interrogazione chatbot, ricerca eventi/risorse su Knowledge Graph, visualizzazione risorse, invio informazioni sul campo, …).

Le richieste e i comandi sull'SSE vengono infine eseguiti, a seconda della loro natura, tramite azioni sul frontend di Dossier Manager, o invocando i microservizi del DSS per la ricerca sul Knowledge Graph o per l'interrogazione di baseline o basi documentali, visualizzando a schermo la risposta alla domanda sottoposta, la risorsa desiderata o inviando informazioni sul campo.

In seguito, i flussi relativi:

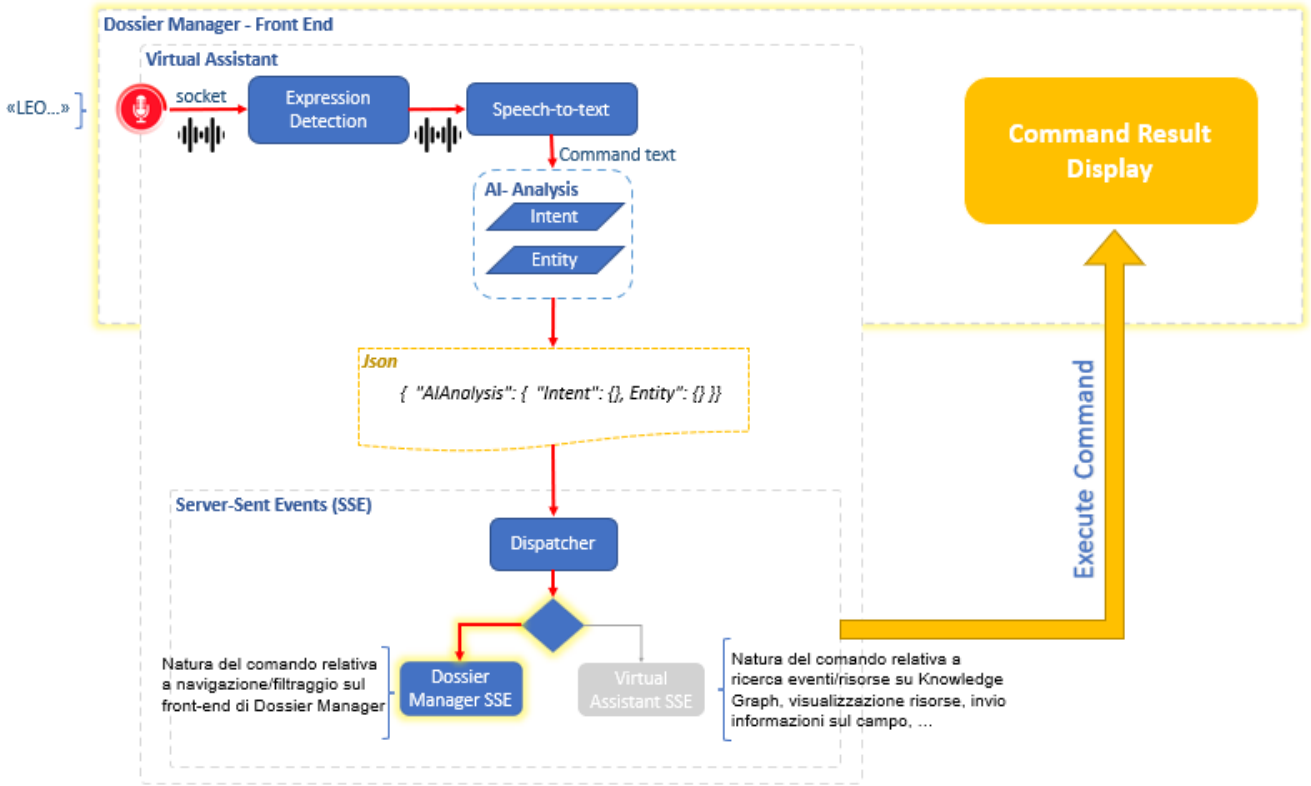


Figura 62: Scenario Comando per il Dossier Manager SSE

Il Virtual Assistant permette di integrare in un portale Web un modulo di sintesi vocale con lo scopo di eseguire dei comandi con la sola voce.

SSE-Manager

Il servizio SSE-Manager ha lo scopo di trasmettere i messaggi inviati da una sorgente verso tutte le destinazioni che si registrano al canale; in particolare questo servizio viene utilizzato per comunicare al frontend la decodifica delle domande e dei comandi vocali relativi, alla ricerca con filtri di eventi

e risorse, comandi destinati al Resource Manager e qualsiasi altro componente che potenzialmente debba eseguire delle azioni tramite comandi vocali.

Infine, può smistare tutti i messaggi su un topic dedicato di una piattaforma di data streaming; in particolare quando riceve i comandi vocali elaborati dal virtual assistant trasmette nel canale sse solo se sono destinati al frontend del componente di destinazione (che è legato al canale al proprio canale sse), gli altri comandi vocali e le richieste che non sono di competenza di un determinato componente non vengono trasmessi nel canale sse dedicato.

In altre parole, dentro questo servizio esiste una componente di dispatching dei messaggi che prefiltra il comando e lo spedisce nel canale sse di destinazione (se specificato in fase di configurazione) oppure nel canale sse generico su cui qualsiasi componente può mettersi in ascolto

Vocal-Assistent-Manager

Il servizio riceve i comandi vocali per

- navigazione su interfaccia grafica
- inserimento di campi per il filtraggio delle ricerche su interfaccia grafica (quali ad esempio date, stato del fascicolo, etc)
- ricerca testuale
- ricerca di eventi e risorse rispetto a filtri temporali (data, intervallo temporale), spaziali (prossimità geografica da un punto nell'intorno di un raggio dichiarato o elementi appartenenti ad una località geografica definita) e relativi ad altre proprietà degli elementi cercati (tipo evento, stato/priorità evento, località evento, tipo risorsa, proprietà/equipaggiamento risorsa, raggio di interesse delle risorse da un indirizzo, nome risorsa, etc.)
- invio di informazioni sul campo
- sottomissione di domande al sistema per ottenere indicazioni su procedure o interrogazioni di basi documentali

A seguito dell'attivazione del microfono mediante click su un apposito tasto sul frontend, viene trascritto il testo dall'audio; la trascrizione viene analizzata da un motore di intelligenza artificiale in grado di estrarre dal parlato il comando e le specifiche dello stesso.

1.1.6.3.2 8.1.1.4 Infrastruttura

Il modulo Virtual Assistant prevede componenti ospitati su Container Platform.

I POD previsti per la soluzione del modulo sono allocati sull'infrastruttura CaaS del PSN, in questo modo è gestito, ad esempio, l'autoscaling.

Unitamente alla componente server, la soluzione si integra con il modulo PaaS AI del SIM e quindi delle piattaforme e dei servizi del PSN nell'ottica di utilizzare le componenti Speech To Text, Question Answering e Intent Recognition da esso esposte.

1.1.6.4 OSINT

1.1.6.4.1 Architettura tecnica

A seguire una descrizione schematica dell'architettura OSINT.

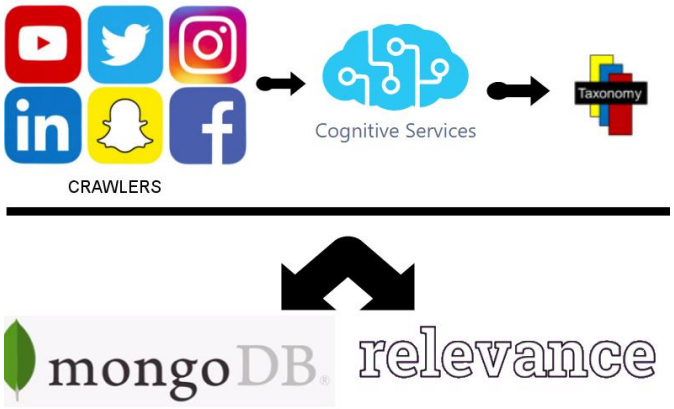


Figura 63: OSINT, architettura logica.

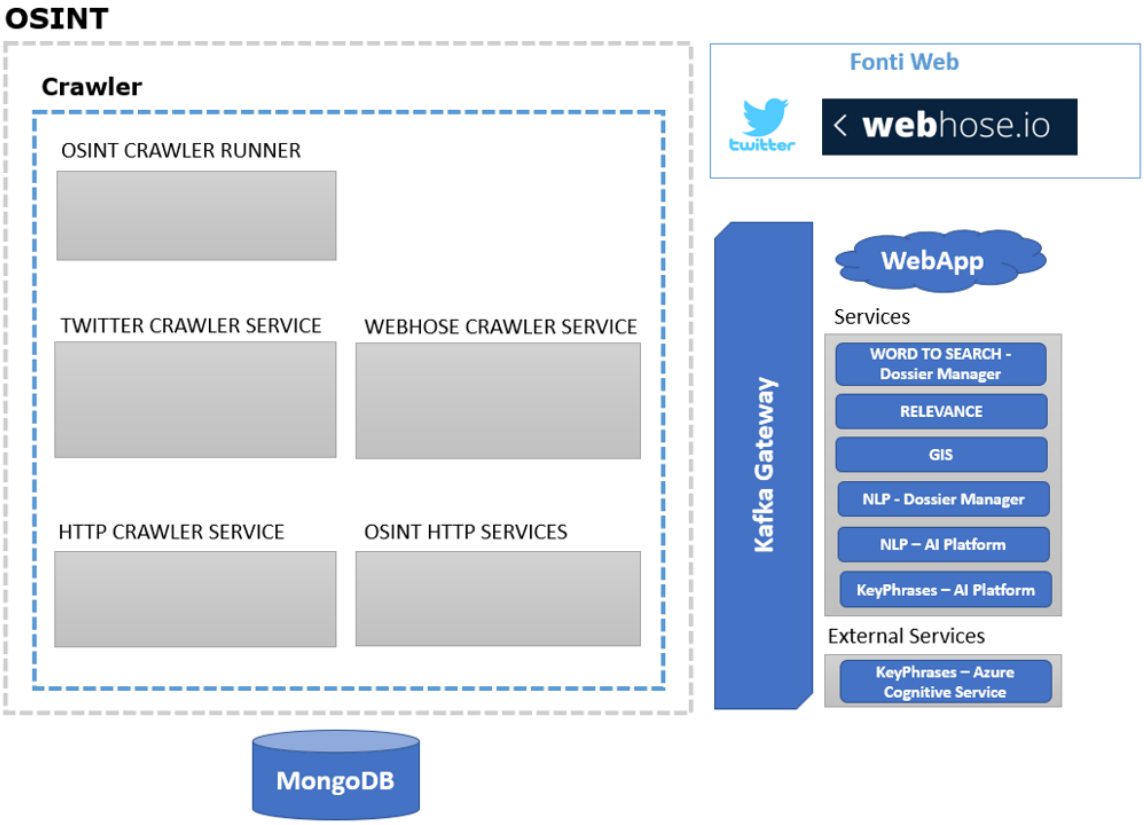


Figura 64: OSINT, schema componenti.

Segue la descrizione dei componenti della soluzione

- **Crawler**: componete principale di gestione delle informazioni estratte dai social web, in particolare il servizio si compone dei seguenti moduli software che sono descritti in seguito:
  - OSINT CRAWLER RUNNER

  - CRAWLER SERVICE
  - HTTP CRAWLER SERVICE
  - OSINT HTTP SERVICES
- **Fonti Web**: le API esposte dai social network come Twitter, Webhose, ecc, attraverso le quali il crawler può accedere allo stream di news da analizzare. Per maggiori dettagli fare riferimento alle documentazioni online (https://developer.twitter.com/en/docs/twitter-api, https://webz.io/data-apis/news-api).
- **Kafka Gateway**: componente di gestione delle comunicazioni, maschera i riferimenti fisici dei servizi, tramite una configurazione memorizzata in una collection su MongoDB. Per maggiori dettagli fare riferimento alla documentazione del Kafka Gateway presente in Dossier Manager.
- **External Services**: servizi esterni ad OSINT esposti da altri componenti, sono utilizzati dal processo di analisi delle news:
  - *Word To Search*: servizio facente parte del Dossier Manager che restituisce le keyword provenienti dai dossier. Questo servizio è usato nel processo di visualizzazione delle keyword all'utente finale. Per maggiori dettagli fare riferimento alla documentazione del Relevance presente in Dossier Manager.
  - *Relevance*: componente facente parte del Dossier Manager, gestisce la rilevanza delle notizie rispetto ai dossier in esame. Per maggiori dettagli fare riferimento alla documentazione del Relevance presente in Dossier Manager.
  - *GIS*: il componente GIS consente di erogare servizi di localizzazione, i layer cartografici, tecniche per il geocoding diretto e inverso e un sistema per la tracciatura delle posizioni. Il componente GIS consente di erogare vari servizi attinenti alla parte GIS come la cartografia, la posizione di dispositivi fissi (telecamere, punti di interesse, asset) e mobili (terminali professionali, smartphone, etc.), layer GIS informativi (come percorsi, aree di attenzione etc).
  - *NLP*: servizi che permettono l'analisi semantica del testo contenuto nell'informazione secondo algoritmi di elaborazione del linguaggio naturale.
  - *KeyPhrases*: servizi che eseguono il confronto del testo con delle frasi chiave e ne determina la corrispondenza, usando degli algoritmi di elaborazione del linguaggio naturale. Sono usati due servizi di KeyPhrases, ***Web App***: applicazione web per l'utente finale, attraverso la quale l'utente può definire i criteri di filtro personalizzati delle informazioni da ricercare con il crawler. Per maggiori dettagli fare riferimento alla documentazione di Dossier Manager.
- ***MongoDB***: database documentale utilizzato per immagazzinare le notizie estratte dal web e le configurazioni del crawler, come utenze, servizi, keyword, tag, ecc.

La progettazione del Crawler permette l'esecuzione di multi-istanze, ognuna con una configurazione diversa, in modo tale da estrarre le informazioni da diverse fonti web (Twitter, WebHose, Facebook, ecc) simultaneamente.
In base alla fonte oggetto di analisi, viene istanziato il relativo componente, e per questo l'architettura del crawler contiene una componente dinamica. Nelle successive immagini sono mostrati gli scenari per le due fonti: Twitter e WebHose.

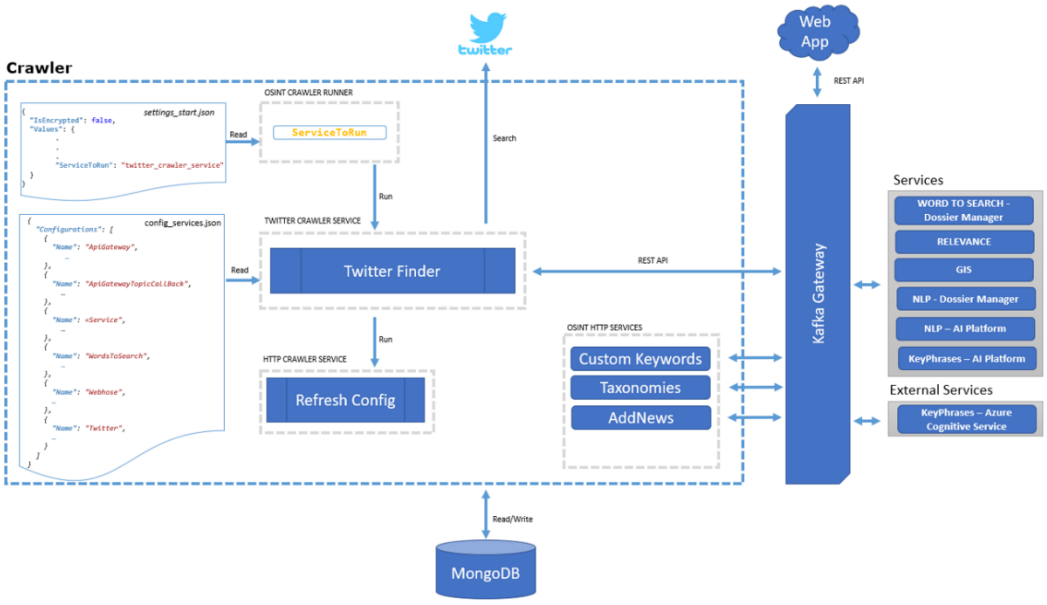


Figura 65: OSINT, scenario Twitter.

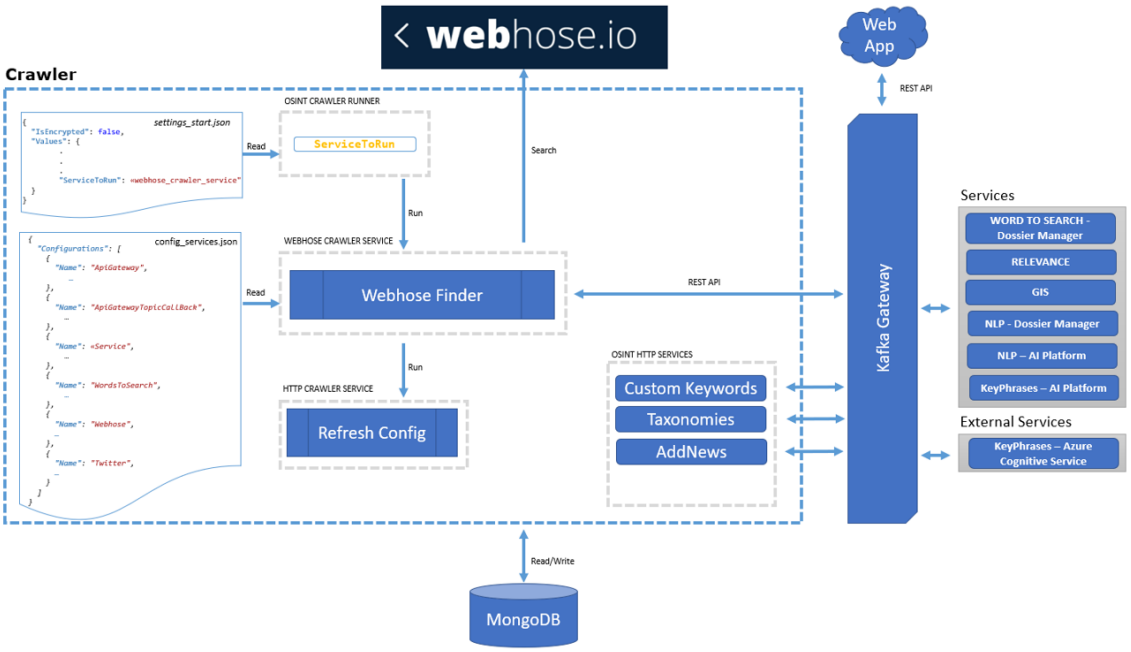


Figura 66: OSINT, scenario webhose.io.

L'avvio del componente OSINT richiede che sia up and running un'istanza MongoDB in versione>4.2, inoltre è necessario aver rilasciato il servizio kafkaproxy manager presente nel Dossier Manager e devono essere attivi anche i servizi di AI Platform e/o i servizi Azure per poter eseguire l'analisi semantica delle notizie estratte.

La figura seguente mostra il processo di Startup di OSINT.

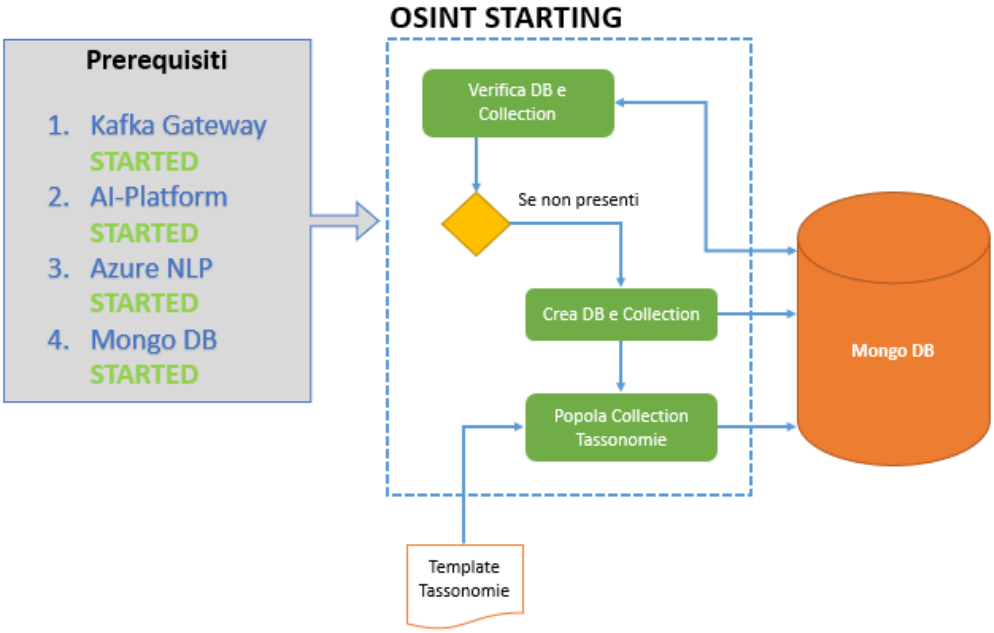


Figura 67: OSINT, prerequisiti e popolamento MongoDB all'avvio.

OSINT dispone di default una collezione di parole da ricercare, provenienti da un file di configurazione; l'utente finale può modificare a suo piacimento attraverso l'interfaccia fornita dalla Web App. Si evidenzia di seguito l'interazione tra diversi componenti durante il processo di aggiornamento delle Keywords.

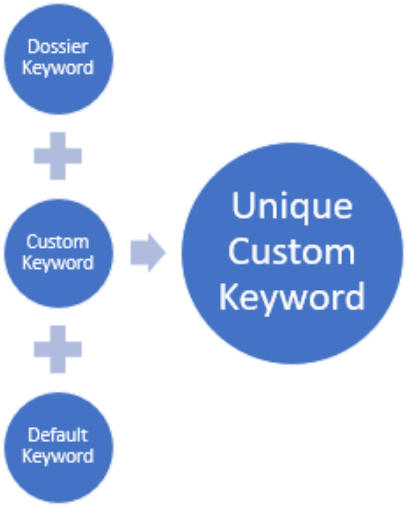


Figura 68: OSINT, Unique Custom Keyword.

OSINT_CRAWLER_SERVICE

Il componente OSINT_CRAWLER_SERVICE si occupa di eseguire un'istanza del crawler. In questo modo, configurando n istanze del servizio, per tutte le fonti web in esame, è possibile eseguire i rispettivi crawler in modo simultaneo. In particolare, i vari crawler eseguiranno le ricerche sul web utilizzando le parole chiavi inserite manualmente per lingua, insieme alle parole chiavi derivanti da tutti i dossier aperti (senza distinzione di lingua).

Le figure seguenti mostrano i due scenari di start dei crawlers per le fonti Twitter e WebHose.

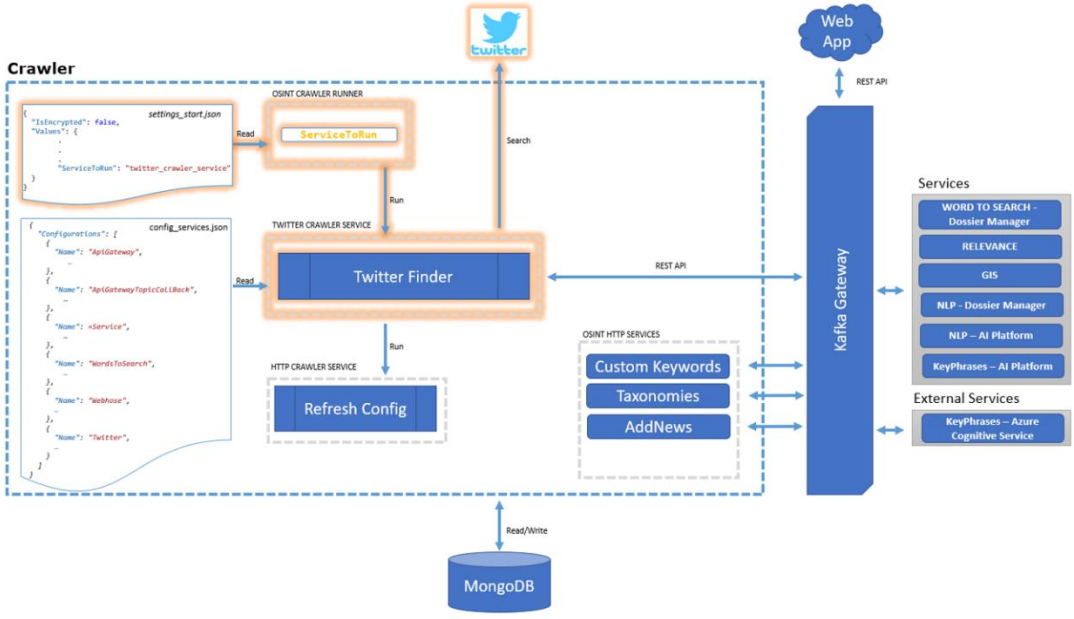


Figura 69: OSINT, avvio del crawler twitter.

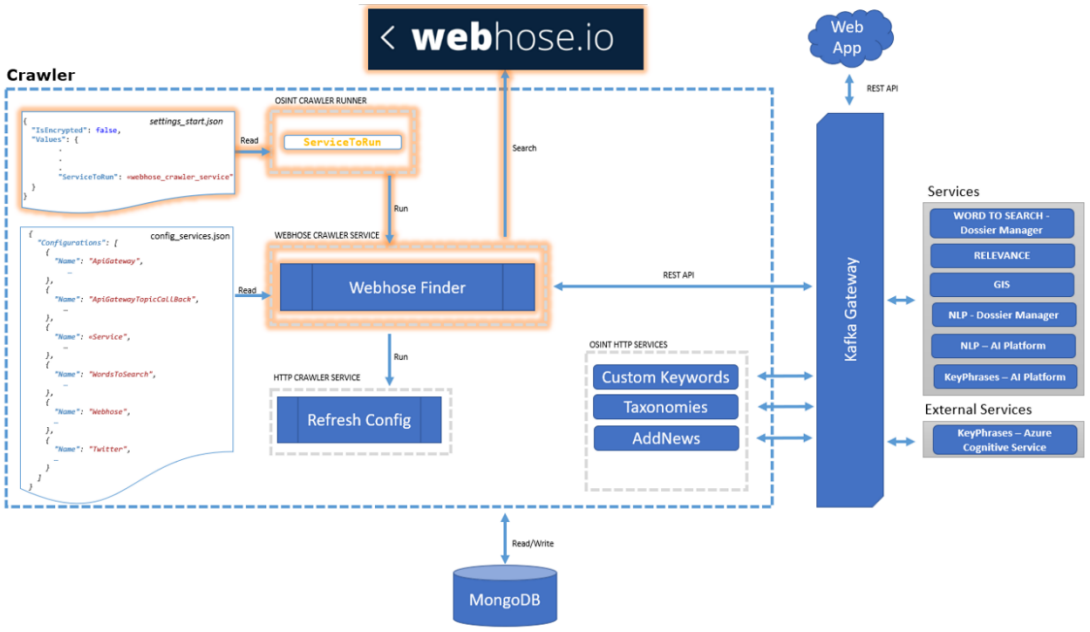


Figura 70: OSINT, avvio del crawler webhose.io.

TWITTER_CRAWLER_SERVICE

Il componente TWITTER_CRAWLER_SERVICE, al suo start istanzia la classe TwitterFinder e subito dopo istanzia il componente HTTP_CRAWLER_SERVICE, quest'ultimo è utilizzato per la gestione del riavvio del crawler. La classe TwitterFinder contiene la logica necessaria per estrarre ed analizzare lo stream di tweet in tempo reale:

- Legge i parametri dalla configurazione per accedere a twitter.
- Recupera le Parole da cercare nel testo.
- Recupera le Lingue con il quale cercare.
- Recupera i Tag con il quale filtrare i tweet.
- Recupera la modalità di ricerca, KEYWORDS E HASHTAGS: KEYWORDS O HASHTAGS.
- Esegue l'analisi del testo del tweet e, in caso di match positivo con i filtri impostati, viene salvato per un'analisi successiva, che verrà eseguita dal servizio AddNews del componente OSINT_HTTP_SERVICES descritto in un paragrafo successivo.

La figura seguente mostra lo scenario in cui i componenti sono coinvolti.

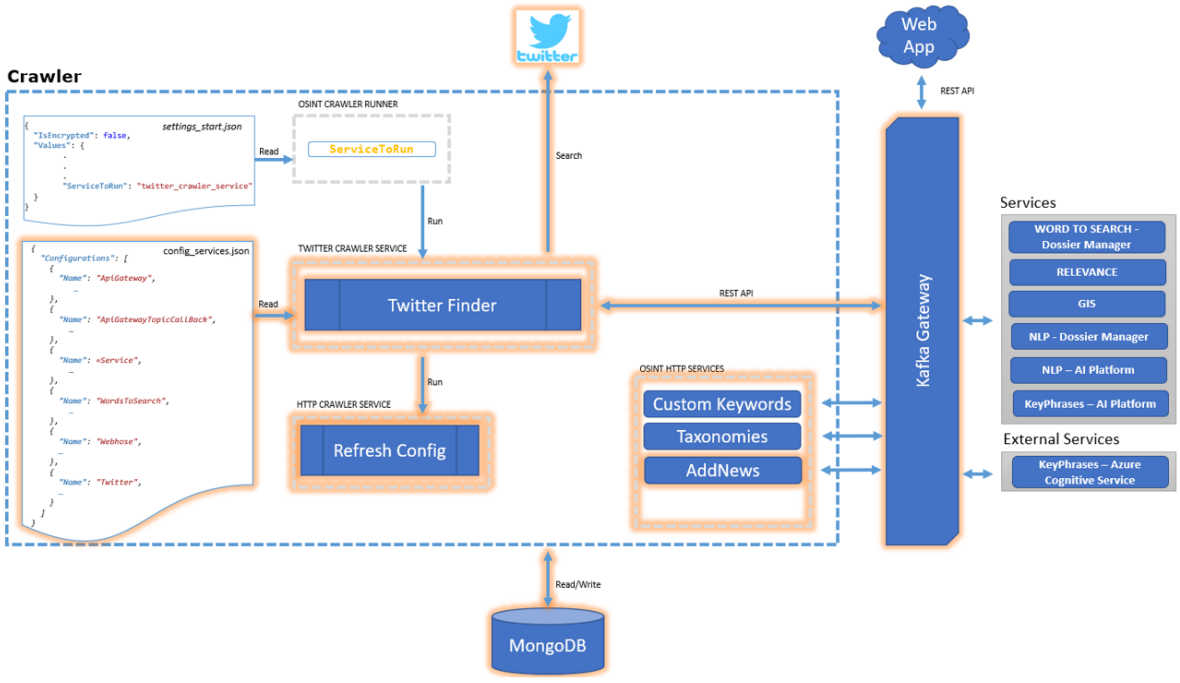


Figura 71: OSINT, flusso proveniente da twitter.

WEBHOSE_CRAWLER_SERVICE

Il componente WEBHOSE_CRAWLER_SERVICE, al suo avvio istanzia la classe WebHoseFinder e subito dopo istanzia il componente HTTP_CRAWLER_SERVICE, quest'ultimo è utilizzato per la gestione del

riavvio del crawler. La classe WebHoseFinder contiene la logica necessaria per eseguire le query al servizio Webhose.io impostando i parametri di ricerca. La classe esegue i seguenti passi:

- i parametri dalla configurazione per accedere a webhose.io.
- Recupera le Parole da cercare nel testo.
- Recupera le Lingue con il quale cercare.
- Esegue la query al servizio con i filtri impostati, in caso di match positivo, il risultato viene salvato per un'analisi successiva, che verrà eseguita dal servizio AddNews del componente OSINT_HTTP_SERVICES.

In particolare, è possibile impostare l'access_key + il numero massimo di notizie da estrarre ad ogni interrogazione, l'intervallo di tempo in ore in cui cercare le notizie ed infine il timeout in ore, entro cui il crawler eseguirà una estrazione interrogando webhose.

La figura seguente mostra lo scenario.

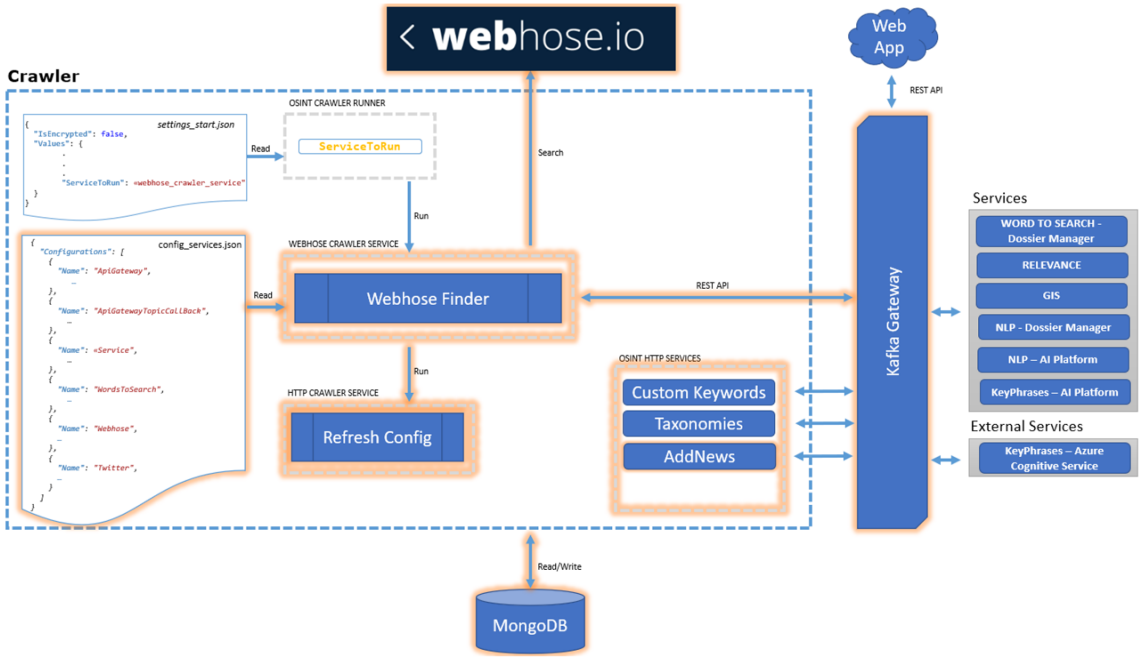


Figura 72: Flusso proveniente dal webhose.ioc

HTTP_CRAWLER_SERVICE

Il crawler legge le keyword in fase di avvio quindi, se i parametri cambiano mentre il crawler è attivo, è necessario riavviare il crawler per far in modo che i nuovi parametri vengano considerati dall'istanza. Il componente HTTP_CRAWLER_SERVICE è incaricato a tale scopo, e la figura seguente mostra i componenti coinvolti:

- L'utente cambia attraverso la WebApp le keyword.

- Il servizio CustomKeywords è predisposto a ritornare e aggiornare le keyword da parte della WebApp. Tale servizio CustomKeywords del componente OSINT_HTTP_SERVICES è descritto in un paragrafo successivo.
- Il servizio HTTP_CRAWLER_SERVICE, già attivo perché istanziato all'avvio dal crawler corrispondente, si accorge che il servizio CustomKeyword ha modificato i parametri di configurazione, ed esegue il restart del Crawler corrispondente.

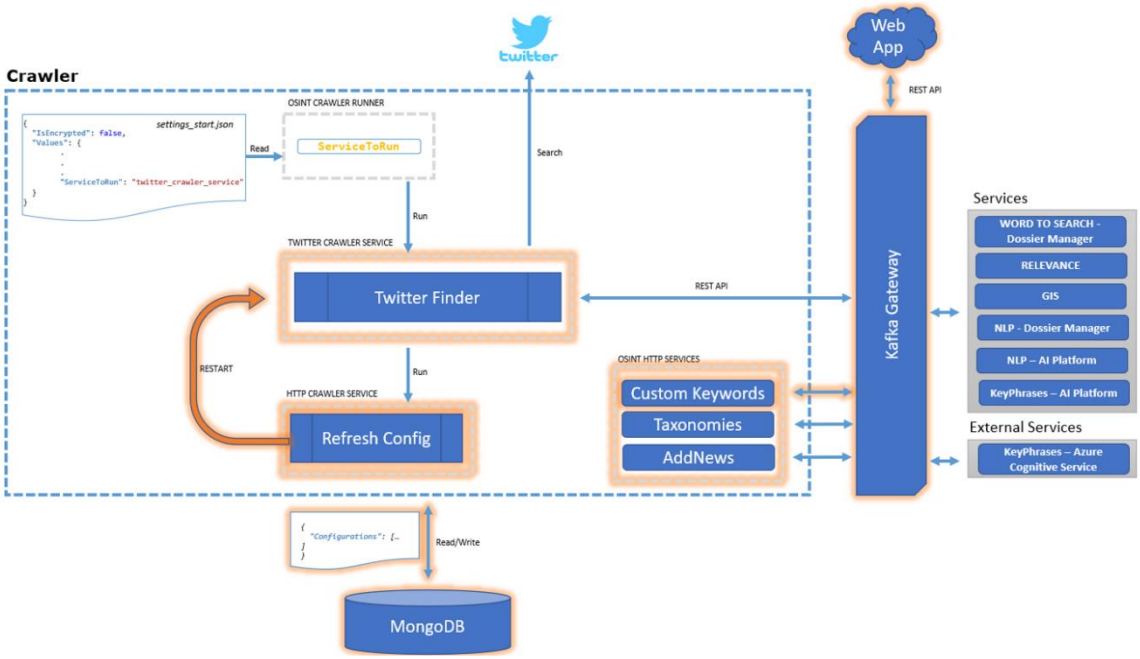


Figura 73: OSINT, aggiornamento configurazione del crawler.

OSINT_HTTP_SERVICES

Il componente OSINT_HTTP_SERVICES è un contenitore di servizi che permette una migliore distribuzione del carico e delle responsabilità. Il componente contiene dei servizi che vengono esposti attraverso il KafkaGateway, sono quindi raggiungibili anche dagli altri componenti della soluzione e da servizi che possono essere considerati come utility al Crawler, il quale li richiama direttamente senza l'ausilio del Gateway, ed infine un servizio di controllo per la verifica dello stato del componente.

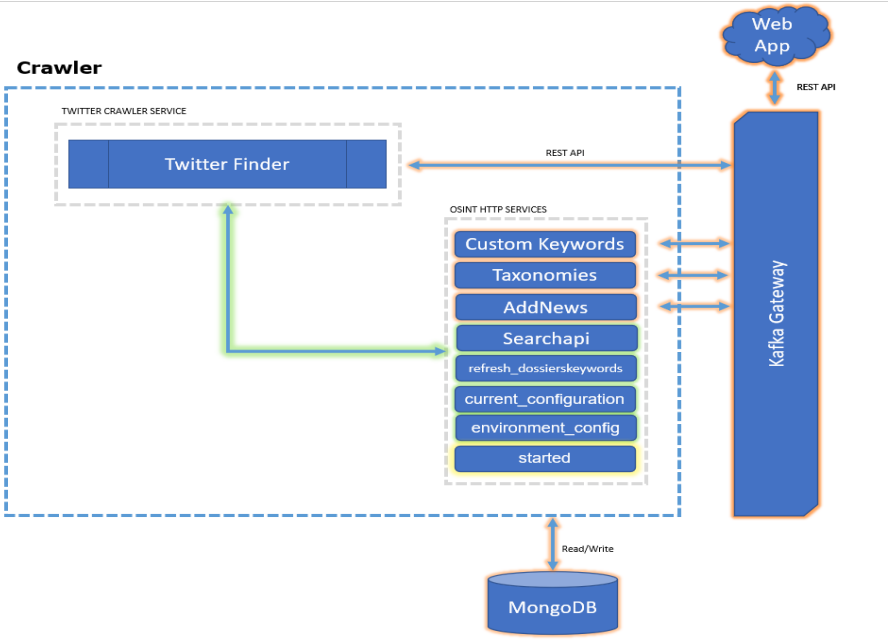


Figura 74: OSINT, servizi esposti dal crawler.

1.1.6.4.2 Infrastruttura

Come accennato in precedenza il componente OSINT si basa su diversi servizi rilasciati in modalità Container riutilizzabili in qualsiasi piattaforma che riesce ad eseguire tali Container.

Molti dei Container previsti hanno necessità di comunicare con un'istanza mongodb versione >= 5.0.15 da cui leggere e/o scrivere dei dati. La componente crawling si basa su dei container ad hoc che comunicano con provider esterni sfruttando la piattaforma TIS

Oltre Mongodb è richiesta la presenza di un broker Kafka per spedire e/o leggere messaggi. La componente di analisi notizie estratte ha necessità di poter comunicare con il GIS per estrarre indirizzi ed eventualmente geolocalizzarli e i servizi di AI Platform o i servizi Cognitivi di Azure on premises per estrazione del contesto semantico di un determinato testo

Ogni pod è replicabile in modo da garantire un'alta affidabilità dei servizi nel caso in cui si verifichino degli errori di sistema e/o applicativi.

1.1.6.5 VIDEO ANALYSIS

1.1.6.5.1 Architettura tecnica

Questa soluzione fornisce una piattaforma pronta all'uso in grado di applicare algoritmi basati su AI su fonti video.

La piattaforma sarà in grado di gestire flussi video al fine di elaborare i relativi task di inferenza.

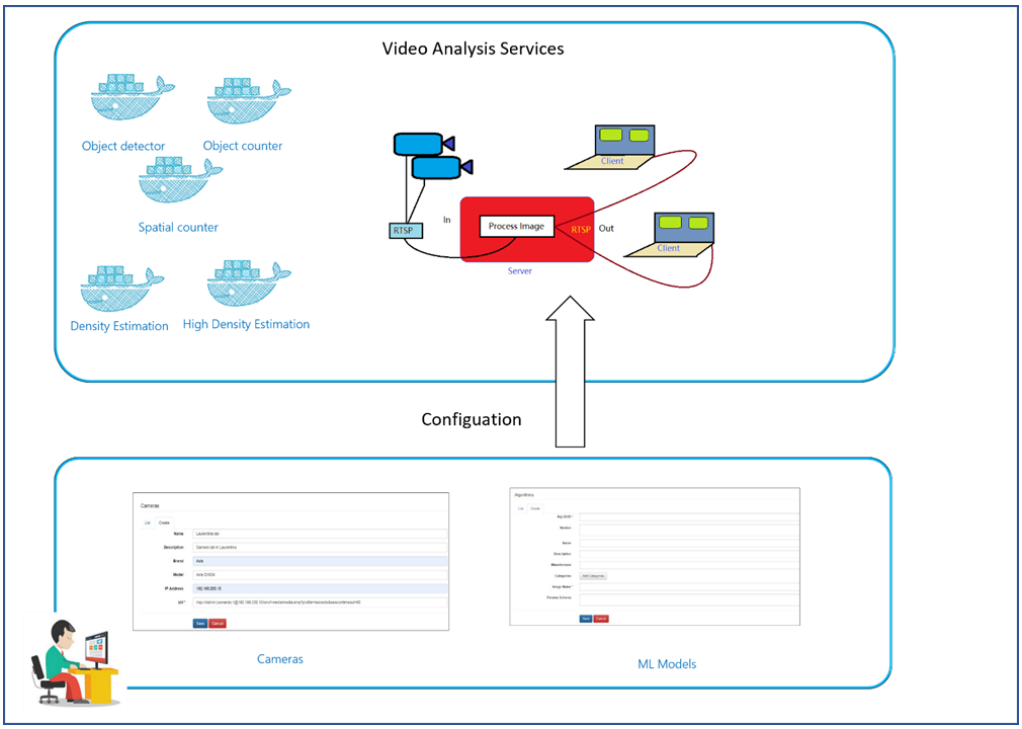


Figura 75 - Soluzione di alto livello PaaS Video Analytics

L'operatore può configurare un determinato processo (selezionando un flusso video da un dispositivo ed un modello di Machine Learning cui vuole sottoporre il flusso) attraverso l'interfaccia grafica. In seguito, l'operatore può avviare il processo che andrà ad arricchire il flusso video con i risultati dell'inferenza del modello sui frame del video stesso. Ciascun processo è gestito attraverso tecnologia a microservizi come container (docker).

Architettura Multi-Layer

L'architettura del sistema Video Analytics è disegnata secondo i paradigmi di flessibilità, scalabilità e modularità. Tale scelta predispone il sistema Video Analytics alla successiva integrazione di ulteriori algoritmi specializzati a nuove specifiche di analisi del flusso multimediale, e driver d'integrazione con le GPU disponibili sull'HW di esercizio. L'architettura modulare multi-layer è costituita da cinque livelli principali interconnessi tra di loro mediante un livello trasversale detto Infrastructure Layer.

Tali livelli sono:

- **Presentation Layer**: strato applicativo di front-end verso gli utenti per la gestione e la visualizzazione integrata delle informazioni rilevanti del sistema, dei sottosistemi e dei sensori integrati. Appartengono a tale livello per esempio: il client di amministrazione per la configurazione del sistema e il servizio di compressione e manipolazione del flusso video in uscita
- **Software Service Layer**: questo livello comprende i servizi API esposti per la configurazione del sistema. Video Analytics Admin si appoggia a questi servizi per leggere e scrivere ai livelli inferiori.

- **Business Layer**: questo livello comprende tutte le applicazioni di back-end necessarie per l'elaborazione dei dati e l'interfacciamento da e verso i sistemi di gestione dei dispositivi. Appartengono a tale livello per esempio: il servizio di analisi del flusso multimediale che alimenta di informazioni il Data Layer e il servizio di gestione degli eventi e contatori basati dal flusso di metadati raccolti.
- **Data Layer**: questo livello comprende tutte le applicazioni di back-end necessarie per la storicizzazione delle informazioni e la gestione di sistemi di memorizzazione. Appartengono a tale livello per esempio: il servizio di gestione della base dati MongoDB (*System Manager*), il servizio di archiviazione dei metadati alimentato dal VCA Controller (*AI Logic*) e letto dall'Event Controller (*Message Bus*).
- **Device Layer**: strato applicativo adibito all'interfacciamento alla capacità di calcolo delle GPU e al monitoraggio del consumo di risorse di sistema (*RAM, CPU, GPU_1, GPU_2, …. GPU_N*)

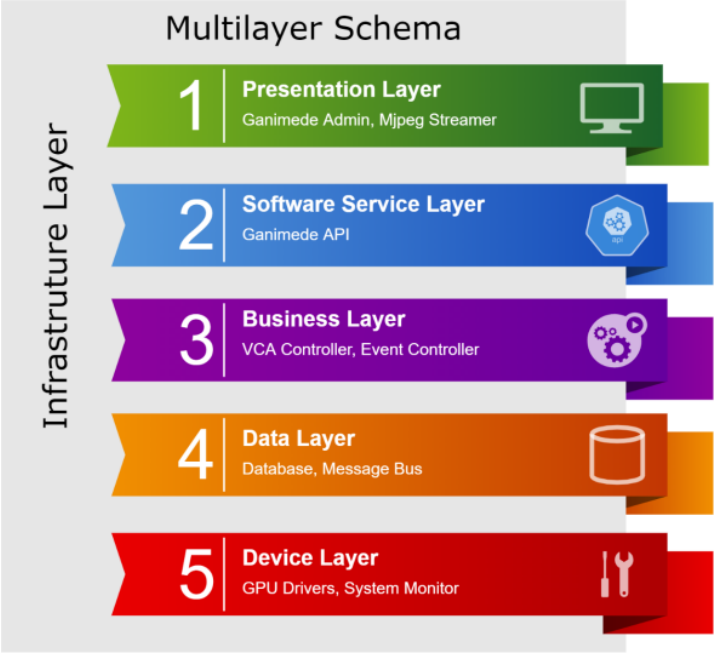


Figura 76 - Organizzazione di Video Analytics in livelli logici

Microservice Architecture

Video Analytics si basa su una architettura a micro-servizi. Una Micro-service Architecture è un modello architetturale per la creazione di sistemi residenti su una rete che separa e isola ogni singola funzionalità / servizio applicativo. Un sistema costruito seguendo la filosofia a micro-servizi è costituito da piccole applicazioni isolate ed interconnesse in rete.

In particolare, un servizio dovrà avere le seguenti caratteristiche:

- **Detectionable**: un servizio deve poter essere ricercato in base alla sua interfaccia e richiamato a tempo di esecuzione. La definizione del servizio in base alla sua interfaccia rende quest'ultima

(e quindi l'interazione con altri servizi) indipendente dal modo in cui è stato realizzato il componente che lo implementa.

- **Atomic & Modular**: ogni servizio deve essere ben definito, completo ed indipendente dal contesto o dallo stato di altri servizi. Si è adottato l'utilizzo di container virtuali per isolare ogni servizio garantendo l'utilizzo di interfacce diservizio per la concatenazione e interoperabilità modulare degli algoritmi VCA.
- **Service Interface**: essere definito da un'interfaccia indipendente dall'implementazione: deve cioè essere definito in termini di ciò che fa, astraendo dai metodi e dalle tecnologie utilizzate per implementarlo.
- **Closed**: essere debolmente accoppiato con altri servizi (loosely coupled). Un'architettura è debolmente accoppiata se le dipendenze fra le sue componenti sono in numero limitato.
- **Public Interface:** essere reso disponibile sulla rete attraverso la pubblicazione della sua interfaccia (in un Service Directory o Service Registry) ed accessibile in modo trasparente rispetto alla sua allocazione.
- **Simple**: fornire un'interfaccia possibilmente a "grana grossa" (coarse-grained): deve mettere a disposizione un basso numero di operazioni, cioè poche funzionalità, in modo tale da non dover avere un programma di controllo complesso.

Progetto Architetturale del sistema

Il presente capitolo descrive la progettazione architetturale del sistema Video Analytics identificando i componenti per ogni Layer.

Infrastructure Layer

La componente server di Video Analytics sarà gestita avviando ogni micro-servizio in specifici container isolati e interconnesso agli altri attraverso una rete virtuale privata (atomicità e isolamento).

I containers rappresentano l'unita funzionale base per ogni servizio del sistema e sono dichiarati e gestiti attraverso un container-engine (es. d*ocker o kubernetes*).

I container realizzati per ospitare un algoritmo VCA saranno vincolati da interfacce formalizzate e documentate per garantire la massima integrazione con il sistema ganimede.

Le dipendenze tra container/servizi saranno esplicitate in fase di configurazione del sistema e il numero di porte esposte in rete sarà ridotto al minimo necessario alle piene funzionalità applicative.

I containers contenenti gli algoritmi di elaborazione del flusso multimediale saranno estensione dell'algoritmo VCA base o ACA base e gestiti dal VCA Controller o ACA Controller durante l'intero ciclo di vita del processo che le utilizza (modularità). Ogni algoritmo disporrà di un JSON Schema che descriverà i parametri di configurazione specifici per il servizio.

Per specifica peculiarità tecnologica docker, attraverso il Dockerfile di ogni container sarà possibile interpretare i prerequisiti relativo per la sua esecuzione.

Il front-end dedicato alla configurazione dei servizi sarà full-web based e si appoggerà a API Rest. La logica di virtualizzazione a containers dovrà essere implementata per garantire:

- Configurabilità e avvio dei soli container VCA necessari all'utilizzo
- Configurabilità e avvio di più flussi distinti in parallelo
- Configurabilità e avvio di distinti moduli VCA sul medesimo flusso
- Configurabilità e avvio del medesimo modulo VCA su flussi distinti

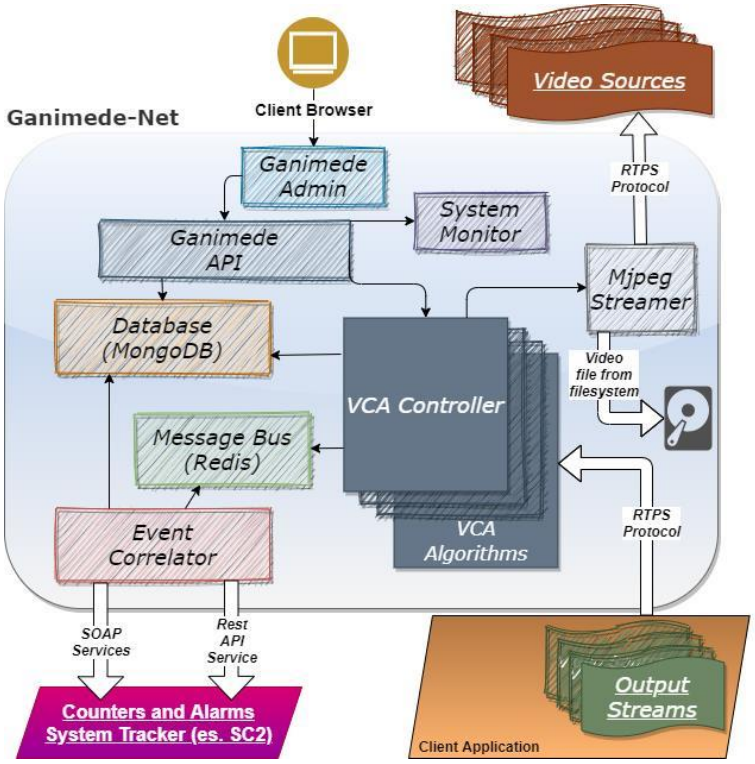


Figura 77 - Infrastructure Layer (Containers' network)

Presentation Layer

Il Presentation Layer espone esternamente due servizi principali:

- Video Analytics Admin: applicazione web per la gestione e monitoraggio dello stato dei processi attivi
- Mjpeg Streamer: servizio di manipolazione del flusso di streaming)

Software Service Layer

Il sistema Video Analytics espone dei servizi API REST finalizzati alla configurazione amministrativa dell'ambiente di esecuzione.

Dal Presentation Layer l'applicazione web Video Analytics Admin utilizzerà questi servizi fornendo una interfaccia utente semplice e intuitiva.

REST (Representational State Transfer) è uno stile architetturale per sistemi software distribuiti, che si caratterizza con l'identificazione delle risorse tramite URI, la manipolazione delle risorse tramite

operazioni HTTP (PUT, GET, POST, DELETE), l'utilizzo di messaggi auto-descrittivi (le risorse sono disaccoppiate dalle loro rappresentazioni; ogni richiesta client contiene tutte le informazioni sufficienti per l'elaborazione della richiesta da parte del server così come ogni risposta del server contiene le informazioni sufficienti a descrivere come il client possa elaborare la risposta stessa).

I servizi REST (detti anche RESTful web service) consentono di realizzare servizi web più "leggeri" rispetto alla tecnologia SOAP, offrendo la medesima interoperabilità, migliori garanzie su prestazioni e scalabilità e maggior semplicità di programmazione.

Il sistema GND prevede anche dei web services esposti in REST API per permettere la configurazione applicativa da una interfaccia web dedicata. Sinteticamente la piattaforma GND è rappresentabile come una architettura di sottosistemi e sotto-servizi integrati:

- **Video Source**: Sorgente video in streaming proveniente da una generica videocamera o da un flusso streaming esterno
- **Audio Source:** Sorgente audio in streaming proveniente da una generica sorgente streaming esterno
- **VCA Controller:** Raccoglie la logica computazionale del flusso video e sarà esteso con algoritmi AI dedicati allo specifico contesto di utilizzo
- **ACA Controller:** Raccoglie la logica computazionale del flusso audio e sarà esteso con algoritmi AI dedicati allo specifico contesto di utilizzo
- **Message Bus:** Applicativo di archiviazione dei metadati raccolti dal flusso multimediale e resi disponibili per la verifica e misurazione di eventi sensibili per lo specifico contesto di utilizzo
- **System Monitor:** Servizio integrato con le risorse locali come RAM, CPU e GPU per misurare costantemente il carico di utilizzo. Il grafo di utilizzo in tempo reale sarà disponibile da Video Analytics Administrator
- **Video Analytics API:** REST API esposte per leggere e aggiornare la configurazione del sistema. Queste API sono utilizzate da
- **Video Analytics Administrator:** I dati letti e memorizzati da questo servizio sono gestiti da Database
- **Database:** Servizio di archiviazione e indicizzazione dei dati di configurazione del sistema Video Analytics
- **Event Correlator:** Servizio di monitoraggio dei metadati raccolti dal Message Bus che si integra con servizi esterni per inviare notifiche o aggiornamenti di contatori. Sarà integrabile con protocollo SOAP o REST API
- **MJPG Streamer:** Servizio di manipolazione del flusso streaming input/output
- **Video Analytics Administrator:** Semplice web app per gestire la configurazione applicativa. Si appoggia ai servizi di Video Analytics API
- **Video output:** Flusso video modificato che contiene le informazioni aggiuntive definite dal VCA Controller.

Figura 78 – Schema di contesto

Business Layer

Il sistema Video Analytics, al fine di identificare elementi da ogni frame analizzato è in grado di coinvolgere uno o più algoritmi VCA.
Durante i processi VCA vengono tracciate informazioni **overlay** sul flusso in uscita e/o storicizzate come metadati sul Message Bus.

Tutti i cambiamenti di stato rilevanti all'interno del flusso multimediale devono essere tracciabili analizzando e applicando opportune regole di interpretazione e correlazione dei metadati archiviati dal Message Bus.

Solo gli eventi che soddisfano le regole di correlazione specificate in fase di configurazione devono essere trasformati in notifiche o variazioni dei contatori.

A questo scopo si distinguono queste informazioni misurabili da una alterazione del contenuto del flusso multimediale gestite dal servizio **Event Correlator** e i VCA coinvolti durante l'elaborazione del flusso.

L'Event Correlator gestirà gli aggiornamenti di contatori e la segnalazione di eventi sensibili connettendosi ad endpoint esterni utilizzando rispettivamente chiamate REST API e Soap.

Data Layer

Il Data Layer gestisce due tipologie di informazioni:

- I parametri di configurazione del sistema Video Analytics

- I metadati informativi raccolti dagli algoritmi di analisi del flusso multimediale

I parametri di configurazione saranno accessibili in lettura e scrittura dai servizi API chiamati direttamente dall'applicativo web di configurazione.

I metadati informativi saranno gestiti dal Business Layer:

- in scrittura dagli algoritmi VCA
- in lettura dall' Event Correlator

MongoDB

Il DBMS (Database Management System) scelto per la storicizzazione dei dati di configurazione è MongoDB: un sistema gestionale di basi di dati non relazionali, orientato ai documenti, di tipo NoSQL.

Il linguaggio utilizzato per la gestione dei dati è Javascript, del quale sfrutta in particolare la notazione BSON (JSON).

MongoDB è disponibile per sistemi Linux e Windows, mette a disposizione API ufficiali per l'integrazione e rende possibile inserire nel database qualsiasi tipo di oggetto (es. array, date, stringhe, numeri, etc).
MongoDB è un gestore di database schemaless, il che significa che i dati, anche all'interno di una stessa tabella (chiamata *collection*), possono essere strutturati in qualsiasi modo e non devono rispettare alcuna regola.
MongoDB raggiunge prestazioni estremamente elevate anche con grandi quantità di dati, ma naturalmente per farlo rinuncia ad altre caratteristiche che sono tipiche dei database relazionali. Ad esempio, non supporta alcun tipo di vincolo di integrità né le transazioni, e la gestione della concorrenza ha grossi limiti.
Inoltre, la durabilità dei dati è garantita solo a intervalli relativamente lunghi: i dati sono conservati nella memoria, e a intervalli scritti su disco attraverso il *logging* e il *journaling*

Redis

Il Message Bus Engine scelto per la l'archiviazione dei metadati informativi dei flussi multimediali è REDIS: un datastore in memoria rapido, open source e di tipo chiave-valore per l'utilizzo con database, caching, broker di messaggistica e code.
Tra i vantaggi offerti da Redis è l'utilizzo della memoria principale del server, a differenza di quanto avviene nei database tradizionali, che memorizzano la maggior parte dei dati su disco.

Nei database tradizionali basati su disco, i dati vengono trasferiti da e verso lo storage per ogni operazione, mentre Redis non prevede questa operazione supportando una quantità di operazioni di ordine di grandezza superiore e tempi di risposta più rapidi.

Device Layer

L'ambiente di esecuzione dei containers dovrà garantire servizi e interfacce necessarie per l'utilizzo delle capacità di calcolo della scheda grafica e di monitoraggio del carico computazionale delle risorse HW: GPU, CPU, RAM

Modelli di ML

La soluzione mette a disposizione algoritmi ML per l'estrazione di informazioni da fonti video:

- **Object Detector**: algoritmo di ML capace di riconosce e localizzare oggetti (persone, macchine, truck, etc.) all'interno di un determinato frame. Con questo servizio per ogni oggetto identificato in un frame si riesce ad estrarre i metadati contenti la sua classificazione e la sua posizione nello spazio
- **Spacial Counter**: estensione del modello Object Detector. Il modello per ogni frame riesce anche ad elaborare un single shot counting per classe di oggetti Object Counter: modello di ML capace non solo di localizzare un oggetto (persone, macchine, truck, etc.) ma anche di tracciarlo attraverso uno o più virtual gate in una sequenza di frame. Con questo servizio per ogni oggetto identificato si riesce ad estrarre i metadati contenti la sua classificazione, posizione nello spazio e traiettoria. Inoltre, per ogni classe di oggetti e per ogni virtual gate l'algoritmo riesce ad ottenere il counting degli oggetti rilevati.
- **Density Estimation:** modello di ML capace di stimare la densità di persone su un dato frame video; tale densità viene rappresentata attraverso una heatmap o mappa di calore. Questo modello è adatto nel monitoraggio di zone poco affollate. Per ogni persona rilevata il modello estrae la sua classificazione e posizione.
- **High Density Estimation**: modello di ML capace di stimare la densità di persone su un dato frame video; tale densità viene rappresentata attraverso una heatmap o mappa di calore. Questo modello è adatto nel monitoraggio di zone molto affollate. Il modello è in grado di estrarre una stima delle persone presenti in un determinato frame video.

1.1.6.6 SEMANTIC SEARCH

1.1.6.6.1 Architettura tecnica

La piattaforma basa la sua metodologia di ricerca semantica su un database composto da fonti informative interne opportunamente selezionate, nonché sul feedback degli utenti utilizzatori del sistema. In questo modo, i risultati prodotti si dimostreranno significativamente più efficaci, in quanto verranno combinati gli output di un tool informatico con le valutazioni di esperti di dominio.

La piattaforma consentirà agli utenti:

- la possibilità di sottoporre query in linguaggio naturale in diverse lingue
- una significativa riduzione dei tempi di ricerca delle informazioni, che non sarà più basata sulla consultazione manuale della documentazione, bensì trarrà vantaggio dall'efficienza dell'AI
- l'ottimizzazione dello strumento e la condivisione delle esperienze dei singoli operatori attraverso il sistema di feedback

Di seguito uno schema di alto livello della soluzione Semanthic Search.

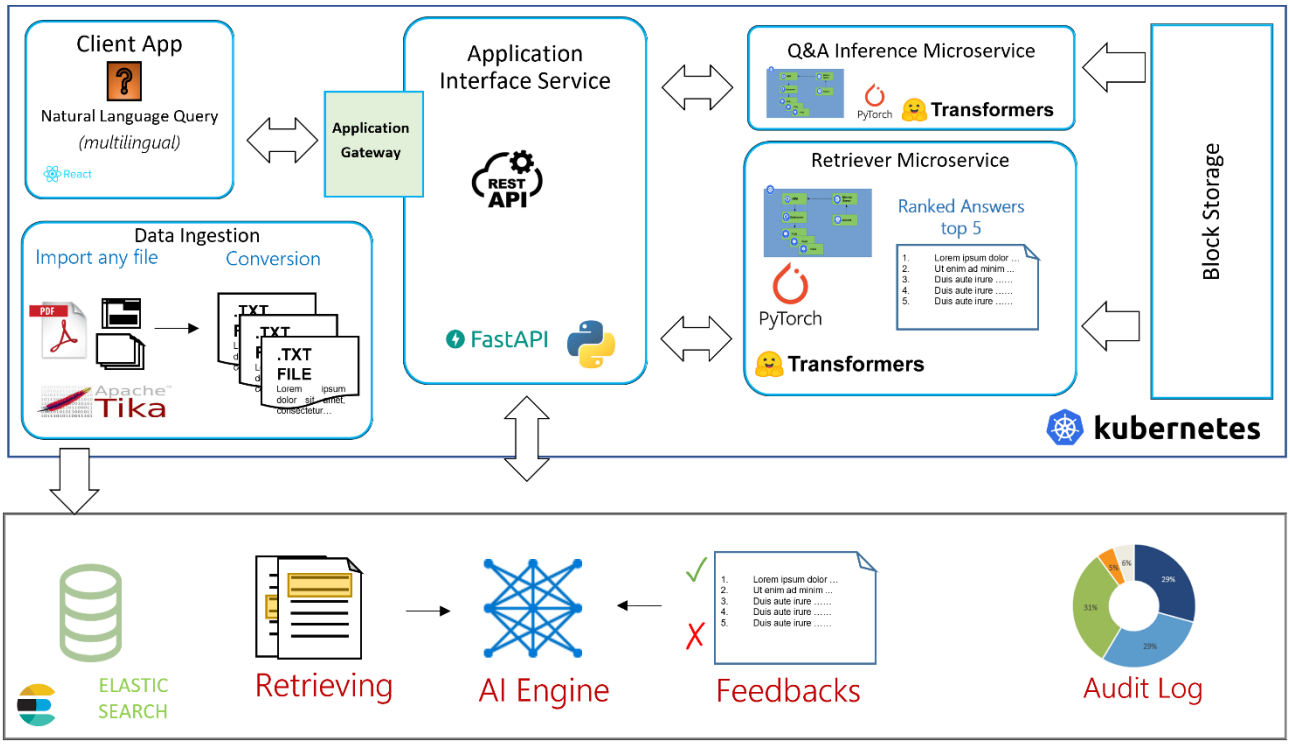


Figura 79 - Soluzione di alto livello Semantic Knowledge Search

Nel seguito si descrivono i componenti principali della soluzione.

Client App
Un frontend minimale user friendly attraverso il quale l'utente può:

- trovare documenti pertinenti alla domanda posta in linguaggio naturale
- sottoporre domande in diverse lingue
- restringere il campo di ricerca attraverso dei metadati d'interesse
- sottomettere dei feedback al fine di generare risultati più rilevanti
- indicizzare i propri documenti attraverso l'upload di uno o più file

FastAPI Framework
FastAPI è un framework Web moderno, veloce (ad alte prestazioni) per la creazione di API con Python

Le caratteristiche principali sono:

- Veloce: prestazioni molto elevate, alla pari con NodeJS e Go (grazie a Starlette e Pydantic). Uno dei framework Python più veloci attualmente disponibili
- Veloce da codificare: aumenta la velocità di sviluppo delle funzionalità
- Meno bug: riduce circa il 40% degli errori indotti dall'uomo (sviluppatore)
- Breve: minimizza la duplicazione del codice, di conseguenza meno bug
- Robusto: ottieni codice pronto per gli ambienti di produzione. Con documentazione interattiva automatica
- Basato su standard: basato sugli standard open per le API: OpenAPI (precedentemente noto come Swagger) e JSON Schema

Bidirectional Encoder Representations from Transformers

I Bidirectional Encoder Representations from Transformers o BERT sono modelli NLP, descritti dai ricercatori di Google AI Language, che hanno suscitato scalpore nella comunità del machine learning presentando risultati pionieristici su 11 diverse attività di elaborazione del linguaggio naturale.

Queste attività comprendono: sentiment analysis, named entity recognition, coinvolgimento testuale (ovvero la predizione della frase successiva), etichettatura dei ruoli semantici, classificazione del testo e coreference resolution. Inoltre, sono stati raggiunti ottimi risultati analizzando parole con significati multipli, dette parole polisemiche, questo grazie ad una comprensione più profonda del contesto di riferimento.

I Bert solitamente sono descritti come modelli di deep learning pre-addestrati; tuttavia, sono dei veri e propri framework, che forniscono ai professionisti del ML, una base su cui costruire le proprie versioni simil-Bert tramite le quali è possibile soddisfare una vasta gamma di task.

In particolare, nella soluzione Semantic Knowledge abbiamo due modelli simil-Bert per i task di Semantic Similarity Search e Q&A (Question and Answering).

Apache Tika

Apache Tika è un software per l'estrazione di dati e analisi dei contenuti, scritto in Java, gestito dalla Apache Software Foundation.

È in grado di trovare ed estrarre testo e metadati da oltre un migliaio di formati di file. In origine apparteneva al progetto Apache Nutch, per l'identificazione di contenuti e l'estrazione di dati da internet per i web crawler. Successivamente divenne un sotto progetto di Lucene. Nel 2007 divenne un progetto autonomo, per diventare una libreria richiamabile da qualunque sistema di gestione dei contenuti (Content Management System) e motore di ricerca.

Tika ha la capacità di analizzare oltre 1400 tipi di file tra quelli elencati dalla Internet Assigned Numbers Authority nei tipi MIME. Tika fornisce l'estrazione del contenuto, dei metadati e l'identificazione della lingua. Tika è scritto in Java, ma è usato da moltissimi altri linguaggi. In particolare, il server REST e la versione CLI consentono agli altri linguaggi di agganciarsi e sfruttare le potenzialità della libreria.

Open Search

OpenSearch è un motore di ricerca open source distribuito basato su una libreria Java, Apache Lucene, che fornisce funzionalità di ricerca e indicizzazione molto performanti su tutti i tipi di dati.

È basato sul concetto di indici, cioè collezioni di documenti in formato JSON collegati tra loro. Ogni documento è caratterizzato da un insieme di chiavi in cui vengono definiti campi o proprietà, con i rispettivi valori.

La struttura di dati che consente di eseguire ricerche full-text in maniera estremamente veloce è chiamata “inverted index”. Tramite questa tecnologia, viene creata una lista per ogni singola parola

contenuta nei documenti dell'indice, dove vengono definiti tutti i documenti per cui quella parola è stata utilizzata. Al momento della memorizzazione dei dati viene quindi creato anche l'indice invertito. L'interazione con il motore di ricerca avviene tramite la tecnologia REST API

1.1.7 Integration Platform

1.1.7.1 Enterprise Integration Platform

1.1.7.1.1 Implementazione dell'Integration Platform

Implementare una soluzione di Integration Platform, in particolare per un sistema complesso e distribuito come quello del SIM impone di considerare più tecnologie che sono integrate al fine di soddisfare i seguenti criteri:

1) supporto per creare e operare processi pienamente federati di interoperabilità e allineati con le direttive e le linee guida definite dalla Comunità Europea e dagli organi competenti operanti nella P.A. Italiana quali AgID e Dipartimento per la Trasformazione Digitale;
2) utilizzare le tecnologie più recenti per rendere disponibili le funzionalità utili a quanto espresso al punto 1;
3) permettere l'integrazione di servizi e dati basati sia su standard aperti e formalmente approvati e permettere anche la gestione di situazioni in cui i dati e i servizi possono essere non allineati a tali standard, sfruttando specifiche tecniche di trasformazione/adattamento; questo caso può presentarsi in situazioni di sistemi datati o comunque realizzati originariamente senza aver considerato esigenze di interoperabilità;
4) modularità e flessibilità della soluzione in modo che offra la possibilità di essere ampliata nel tempo per includere future necessità non identificabili nell'idea del progetto attuale del SIM; questa necessità può considerare un percorso temporale fino ai prossimi 5 o 6 anni;
5) le tecnologie utilizzate per creare l'Integration Platform devono essere pienamente operabili sull'infrastruttura scalabile prevista per il SIM (containers e Kubernetes).

In questo senso è stato primariamente esplorato l'ampio ambito di soluzioni che considerano questi aspetti:

- realizzate in modalità open source, in questo caso devono mostrare un adeguato livello di qualità e sostenibilità nel periodo dei prossimi 5-6 anni;
- offrano un insieme di funzionalità che includano:
  - capacità di creazione e gestione di scenari di integrazione (si veda di seguito la definizione di questo termine);
  - gestione completa di API management;
  - disponibilità di funzionalità di data transformation per rispondere adeguatamente al criterio 3, descritto precedentemente; questa capacità deve poter operare nel contesto dell'elaborazione in tempo reale delle interazioni tramite chiamate API;
  - offra la capacità di estendere le funzioni base offerte in modo da poter affrontare situazioni non previste in fase iniziale; tipicamente è riscontrabile con un'architettura modulare e flessibile con l'impiego di plug-in e connettori;

- metta a disposizione gli strumenti utili per una gestione completa su aspetti funzionali e non funzionali come amministrazione degli accessi, policy di gestione, autenticazione per le funzioni amministrative e sull'uso delle API, piena scalabilità, controllo del traffico API, ecc.
- metta a disposizione strumenti per lo sviluppo facilitato (low-code/no-code) di scenari di integrazione e API;
- la soluzione deve operare su piattaforma container Kubernetes;
- sia disponibile in modalità ibrida o on-prem; una soluzione completamente offerta in modalità SaaS non è adeguata in funzione dell'uso del Cloud Nazionale (PSN)

Da una completa analisi, sia di soluzioni open source che commerciali è emerso che la soluzione più adeguata è **MuleSoft Anypoint**. A oggi non esistono soluzioni open source che rispondo pianamente ai criteri di identificazione di Integration Platform. Ci sono diverse soluzioni di API Manager, di ottima qualità e pienamente sostenibili ma non offrono tutte le funzionalità che una soluzione di Integration Platform deve fornire.

In questo senso la seguente tabella riporta l'elenco delle soluzioni esaminate e riscontrabili sull'offerta odierna del mercato:

	Requisiti	MuleSoft Anypoint	Dell Boomi	Informatica App. Integration	SAP	Oracle
Requisiti generali	Single Unified platform	buono	buono	nella media	buono	nella media
	Low-code/No-code	buono	buono	buono	nella media	scarso
	User Interface	buono	buono	buono	nella media	scarso
	Product or Service	buono	scarso	nella media	buono	buono
	Rischio lock-in	basso	alto	alto	alto	alto
	Collaboration	buono	buono	nella media	nella media	nella media
	Scalability	buono	buono	buono	buono	nella media
	Standards support	buono	buono	buono	buono	buono
	Governance	buono	buono	nella media	buono	nella media
	Scenarios development	buono	buono	buono	nella media	nella media
Requisiti tecnici	Erogabile su qualsiasi cloud e su PSN	si	no	no	no	si
	Attivabile su piattaforma container	si	no	no	no	si

Tabella scorecard su soluzioni di Integration Platform

MuleSoft Anypoint in ogni caso include il modulo principale (Mule Runtime Engine) disponibile in modalità open source.

Architettura e Componenti Principali

Al centro della piattaforma Anypoint si trova il Mule Runtime Engine. Questo motore runtime, finalizzato alle tematiche di integrazione applicative, è responsabile dell'esecuzione di flussi di integrazione e API seguendo una tipica filosofia, guidata da eventi: "*Receive, Transform, Process*". Con un'architettura orientata agli eventi, è in grado di offrire prestazioni elevate e una capacità di elaborazione in tempo reale. La sua scalabilità è anche una delle sue caratteristiche principali, garantendo che le organizzazioni possano adattarsi alle esigenze in evoluzione senza sacrificare le prestazioni.

Il *Design Center di Anypoint* facilita la creazione di API attraverso un'interfaccia intuitiva e visuale. Gli sviluppatori possono sfruttare una serie di strumenti per progettare, testare e documentare le API in un unico luogo. Questo centro offre anche modelli e best practice per accelerare il processo di sviluppo e assicurare che le API siano costruite seguendo le migliori norme del settore.

Flex Gateway è il gateway/API Manager ottimizzato per offrire elevate prestazioni e mantenendo al contempo elevati standard di sicurezza e performance.

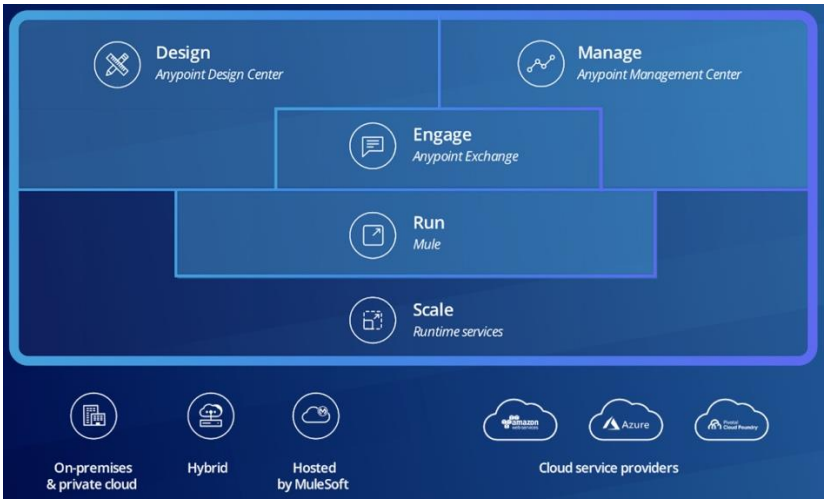
Un altro componente cruciale è l'*Anypoint Exchange*. Funzionando come un hub centrale, l'Exchange permette alle organizzazioni di condividere e scoprire risorse quali API, connettori, modelli e best practice. Questa condivisione di risorse accelera la realizzazione di progetti e promuove una cultura di riutilizzo all'interno dell'organizzazione.

La gestione delle API è ulteriormente potenziata dal *Management Center*. Questa suite di strumenti fornisce una visione completa del ciclo di vita delle API, dalla creazione alla pubblicazione e monitoraggio. Con capacità avanzate di analisi e reporting, le organizzazioni possono ottenere intuizioni preziose sulle prestazioni delle API e identificare rapidamente eventuali problemi.

Integrazione e Connettività

Un aspetto distintivo della piattaforma Anypoint è la sua vasta libreria di connettori precostruiti. Questi connettori facilitano l'integrazione con una moltitudine di sistemi, dai database tradizionali alle moderne applicazioni SaaS. Invece di scrivere codice personalizzato per ogni integrazione, gli sviluppatori possono sfruttare questi connettori per velocizzare il processo e garantire l'affidabilità.

La piattaforma non si limita a offrire strumenti per integrare sistemi esistenti ma adottando nuove tecnologie e nuovi standard come GraphQL, Service Mesh è possibile creare soluzioni integrate *cloud native*.



Schema architetturale funzionale di MuleSoft Anypoint

Scenari di integrazione

L'interazione prettamente digitale nel SIM non si può più basare su applicazioni e servizi isolati, ma piuttosto su ecosistemi interconnessi che necessitano di scambiarsi informazioni in modo fluido e sicuro. In questo panorama, gli scenari di integrazione rappresentano la struttura su cui poggia questa rete di comunicazioni. Essi non sono altro che sequenze coordinate di interazioni, spesso

realizzate attraverso chiamate API, che permettono di orchestrare il dialogo tra le varie entità di un sistema, come il SIM.

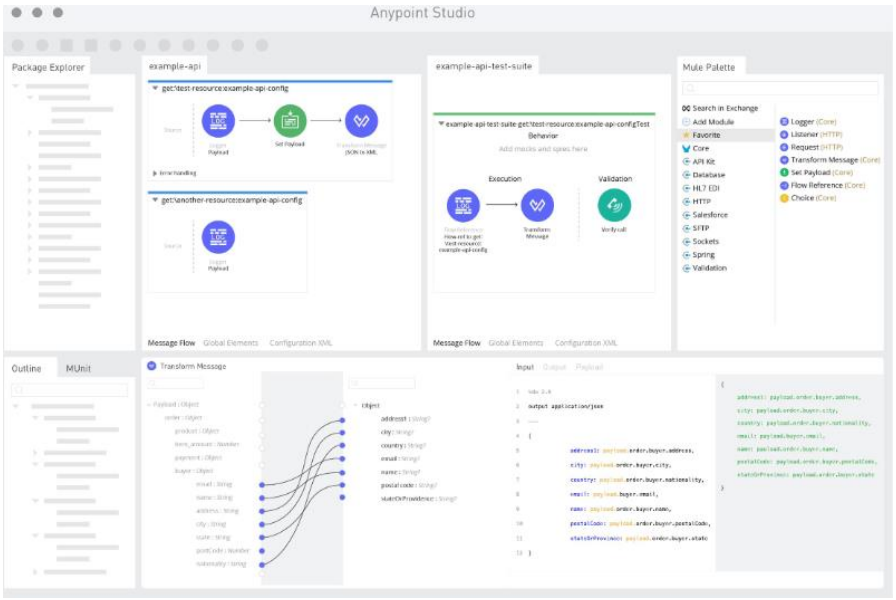
Immaginiamo un'applicazione che necessiti di informazioni provenienti da più stakeholder all'interno del SIM. Invece di effettuare chiamate API separate per ogni singola informazione e poi cercare di amalgamare i risultati, l'applicazione può avviare uno scenario di integrazione predefinito, che si prenderà cura di orchestrare tutte le chiamate necessarie, gestendo le interazioni tra i diversi stakeholders e restituendo un risultato sintetico.

L'Integration Platform gioca un ruolo chiave in questo contesto, offrendo gli strumenti per definire, ottimizzare e monitorare questi scenari. La soluzione tecnologica candidata in questo ruolo, MuleSoft Anypoint, mette a disposizione una funzionalità denominata "flows".

Flows in MuleSoft Anypoint

MuleSoft Anypoint Studio, con il suo editor visuale, consente la creazione di scenari di integrazione un'esperienza intuitiva e flessibile. Ogni *flow* rappresenta uno scenario di integrazione, ed è composto da una serie di componenti o "building blocks" che definiscono come i dati vengono ricevuti, elaborati e trasmessi.

1) **Avvio**: Ogni flow ha un punto di avvio, spesso una specifica chiamata API o un evento, che innesca l'intero scenario di integrazione.
2) **Componenti di Elaborazione**: Dopo l'avvio, i dati possono passare attraverso una serie di componenti che li trasformano, li arricchiscono o li filtrano secondo le esigenze specifiche dello scenario.
3) **Chiamate verso Stakeholders**: All'interno del flow, si possono definire chiamate API specifiche verso i vari stakeholders, gestendo autonomamente autenticazione, errori e retries.
4) **Gestione degli Errori**: Anypoint Studio permette di definire comportamenti specifici in caso di errori, garantendo che ogni scenario di integrazione sia resiliente e capace di gestire situazioni impreviste.
5) **Conclusione**: Alla fine del flow, il risultato aggregato viene restituito all'applicazione chiamante o ad un altro destinatario predefinito.



Anypoint Studio

L'editor visuale di Anypoint Studio rende questo processo accessibile anche a chi non ha una profonda conoscenza tecnica, di fatto si tratta di una soluzione *low-code/no-code*, permettendo di "disegnare" il flusso di integrazione trascinando e collegando i vari componenti. Inoltre, grazie alla sua natura basata su standard open, garantisce che ogni scenario creato sia flessibile, scalabile e sicuro.

Data transformation nel contesto dell'esecuzione di uno scenario di integrazione

In un ambiente digitalmente avanzato e interconnesso, ogni volta che si effettua una chiamata API per ottenere o inviare dati, si è spesso di fronte alla sfida di garantire che questi dati siano nel formato corretto e strutturato come ci si aspetta. Questa operazione di "traduzione" o "trasformazione" dei dati diventa ancor più critica quando si parla di scenari di integrazione, in cui diverse fonti e destinatari possono avere bisogno di dati in formati diversi.

La trasformazione dei dati, nella sua essenza, si preoccupa di operazioni come la conversione di date da un formato all'altro, la normalizzazione delle unità di misura, o la mappatura tra differenti rappresentazioni di una stessa informazione. Un esempio concreto può riguardare la necessità di convertire dati da XML a JSON, una trasformazione molto comune dato che XML è stato a lungo lo standard dominante, ma JSON è diventato sempre più popolare per la sua leggerezza e facilità di lettura. Ancora, un'operazione come il "forward geocoding", che traduce un indirizzo in coordinate geografiche, rappresenta una forma di trasformazione dei dati che combina sia la conversione di formato che un'operazione di elaborazione più avanzata.

La capacità di eseguire queste trasformazioni in modo efficace e affidabile è fondamentale per garantire la fluidità e la coerenza degli scenari di integrazione. E qui entra in gioco MuleSoft Anypoint con la sua soluzione DataWeave.

DataWeave in MuleSoft Anypoint

DataWeave è il linguaggio di trasformazione e mappatura dei dati nativo di MuleSoft. È stato progettato per semplificare la trasformazione dei dati, permettendo agli utenti di convertire facilmente tra i vari formati e di elaborare i dati secondo le esigenze specifiche del flusso di integrazione.

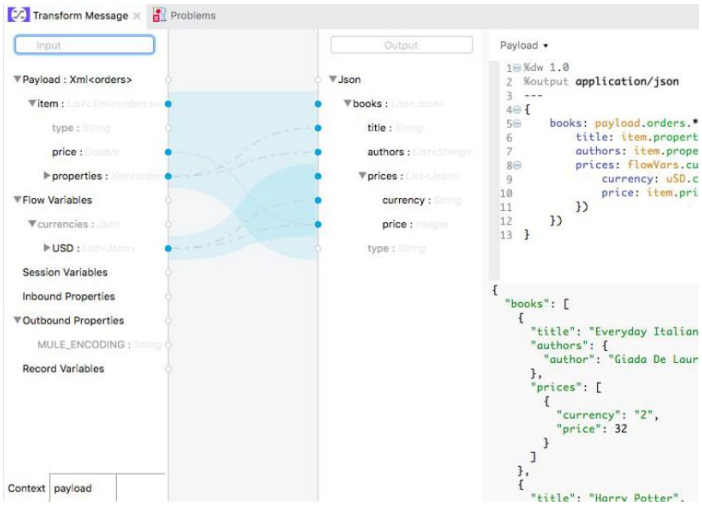
Alcune delle principali caratteristiche di DataWeave sono:

1) **Supporto per vari formati**: DataWeave può lavorare con una vasta gamma di formati di dati, tra cui JSON, XML, CSV, Java, e molti altri.

2) **Sintassi espressiva**: La sintassi di DataWeave è allo stesso tempo completa (come funzionalità) e facilmente usabile, consentendo trasformazioni complesse con poche righe di codice.

3) **Integrazione profonda con Anypoint Studio**: all'interno di Anypoint Studio, gli utenti possono sfruttare un editor visuale per definire le trasformazioni, rendendo il processo accessibile anche a chi non è esperto di programmazione.

4) **Funzioni predefinite**: DataWeave include una vasta libreria di funzioni predefinite che coprono le esigenze più comuni, dall'elaborazione delle stringhe alla manipolazione delle date.



Anypoint DataWeave

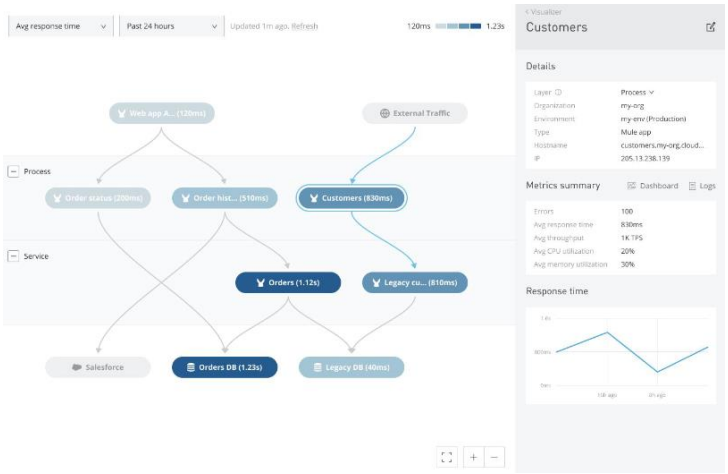
Il linguaggio *DataWeave* e relativo "engine" è in corso di rilascio in modalità open source mentre il DataWeave Client Interface è già pubblicato come progetto open source: https://github.com/mulesoft-labs/data-weave-cli.

Funzionalità relative all'API Manager

L'API Manager nel contesto di un'Integration Platform, è fondamentale per gestire, controllare e monitorare l'utilizzo delle API, facilitando così l'integrazione tra diversi servizi e piattaforme.

Ciclo di vita delle API

Un'efficace gestione del ciclo di vita delle API è essenziale per garantire che le interfacce siano sempre aggiornate, sicure e in linea con le esigenze degli stakeholder. L'API Manager supporta tutte le fasi del ciclo di vita, dalla progettazione, al testing, alla produzione e alla deprecazione, garantendo che ogni API sia sempre nella sua versione ottimale.



Anypoint Visualizer

Per aiutare nel contesto della comprensione complessiva dei flussi e scenari di integrazione esiste la funzionalità Visualizer che mostra le interazioni tra API in termini di:

- Dipendenze
- Integration flows
- Impact Analysis
- Change Management

MuleSoft Anypoint e Flex Gateway offrono una suite integrata di strumenti e servizi per semplificare la gestione del ciclo di vita delle API, dalla progettazione alla distribuzione, passando per la gestione e il monitoraggio. In particolare, questo sono i contesti tipici nel processo di gestione del ciclo di vita e l'uso delle funzionalità messe a disposizione dalla soluzione.

Le API create con MuleSoft Anypoint possono essere pubblicate e gestite attraverso il componente chiamato Anypoint Exchange. Anypoint Exchange è un'area in cui gli sviluppatori e gli utenti possono scoprire, condividere e riutilizzare API, connettori, modelli e altri asset aziendali. Ovviamente le stesse API possono essere pubblicate anche su:

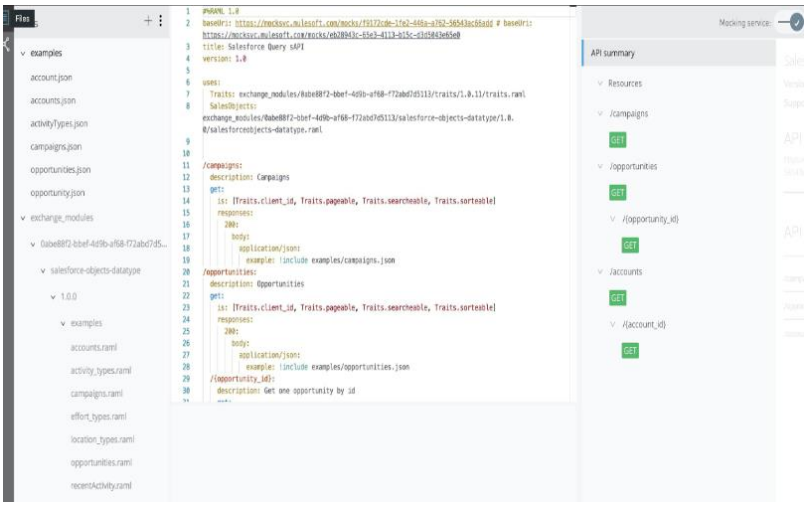
- Catalogo API dell'infrastruttura di interoperabilità PDND (Piattaforma Digitale Nazionale Dati).
- Catalogo API/Servizi previsto nel SIM.

MuleSoft Anypoint Design può generare dinamicamente la documentazione del "contratto API" durante la progettazione, in modo che non solo gli utenti possano utilizzare la documentazione una volta pubblicata, ma anche i progettisti possano rivedere rapidamente la documentazione del loro contratto.

La documentazione generata automaticamente fornisce informazioni sulla REST API. Tutti gli end-point sono descritti in una pagina web con i dettagli riguardanti il metodo HTTP, il nome della risorsa,

i tipi di dati di richiesta e risposta insieme ai tipi di errori che possono essere intercettati e al meccanismo di sicurezza utilizzato. Inoltre, vengono visualizzati esempi di codice in vari linguaggi.

Con *MuleSoft Anypoint API Designer* è inoltre possibile generare un servizio di simulazione (*mocking service*). Ciò offre il vantaggio di consentire al gruppo responsabile dello sviluppo del software che utilizza le API di scrivere il proprio codice rispetto all'implementazione fittizia in parallelo con il gruppo che deve implementare l'API effettiva. In questo modo, il gruppo responsabile della creazione del software può completare il proprio lavoro prima ancora che le API vengano sviluppate, consentendo la collaborazione sin dalle primissime fasi del processo di progettazione, il che amplia l'opportunità di progettare un'API con granularità esponendo al contempo le risorse e i tipi di dati corretti.



API Designer

Progettazione e Sviluppo

- **Designer API**: è disponibile un designer API integrato che permette di progettare, testare e simulare API in un ambiente drag-and-drop. Gli sviluppatori possono definire e modellare API usando RAML (RESTful API Modeling Language), facilitando così la creazione di API RESTful e basate su standard OpenAPI e RAML.
- **Connettività senza soluzione di continuità:** gli sviluppatori possono connettere facilmente le API a una vasta gamma di sistemi e servizi, grazie ai numerosi connettori precostruiti disponibili.

Gestione dell'accesso

L'API Manager offre funzionalità avanzate per controllare chi può accedere alle API e in che modo. Questo comprende autenticazione, autorizzazione e la possibilità di emettere, revocare o rinnovare token di accesso.

Routing

Grazie al routing, è possibile indirizzare le chiamate API al giusto endpoint o versione, garantendo che le richieste siano sempre gestite in modo efficiente.

Monitoring

Monitorare le prestazioni delle API e analizzare il traffico in tempo reale è fondamentale per garantire la massima affidabilità e tempestività nel risolvere eventuali problemi. L'API Manager fornisce strumenti per il monitoraggio in tempo reale, la generazione di report e/o notifiche in caso di malfunzionamenti.

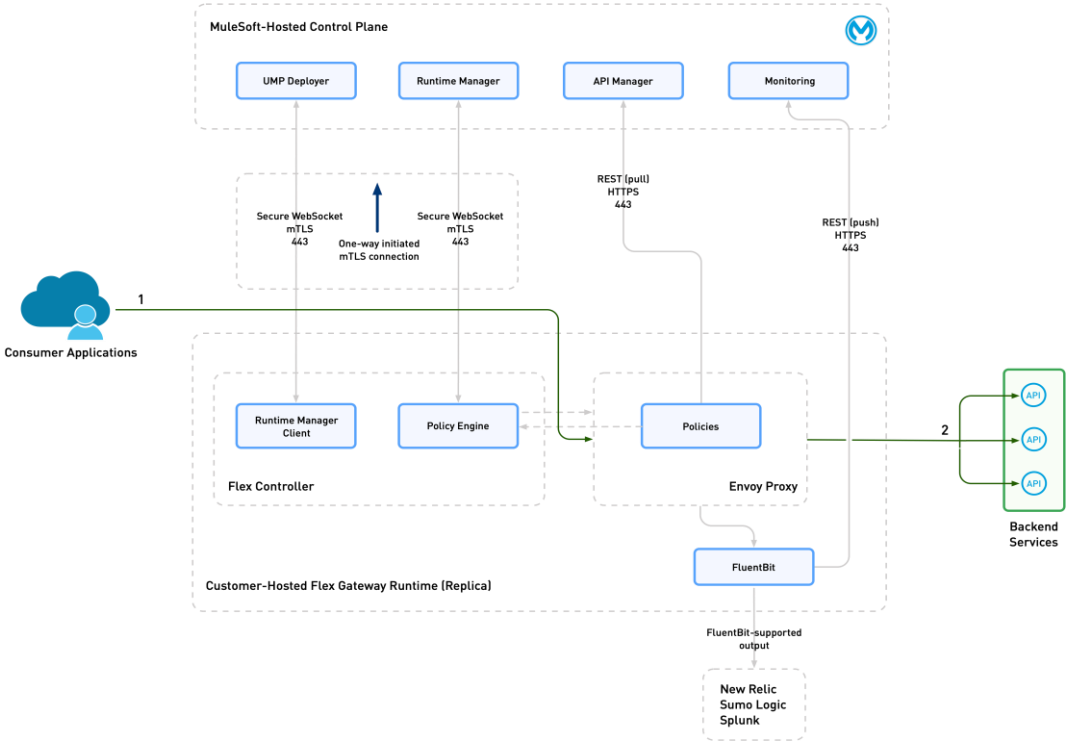
Policy e sicurezza

Oltre a garantire l'accesso sicuro, l'API Manager permette di definire policy specifiche per limitare, ad esempio, il numero di chiamate per utente, oppure per assicurare che le API siano utilizzate nel rispetto delle condizioni stabilite.

API Manager

L'API Manager di MuleSoft, noto anche come *Flex Gateway*, mette in primo piano la sicurezza quando si tratta di gestire e distribuire API.

Anypoint Flex Gateway è un gateway API leggero e ultraveloce basato su Envoy, soluzione di high-performance edge/middle/service proxy (runtime scritto in C++ come progetto open source) e progettato per gestire e proteggere le API in esecuzione ovunque. Costruito per integrarsi perfettamente con i flussi di lavoro DevOps e CI/CD, Flex Gateway offre le prestazioni necessarie per le applicazioni e i microservizi più esigenti, garantendo al contempo sicurezza e gestibilità a livello aziendale in qualsiasi ambiente.



Schema funzionale di Flex Gateway

Gli aspetti di sicurezza dell'API Manager sono progettati per garantire che solo le parti autorizzate possano accedere alle risorse, oltre a proteggere contro minacce esterne e interne. Ecco una panoramica dettagliata delle principali funzionalità di sicurezza offerte:

Autenticazione e Autorizzazione

- **Autenticazione basata su token:** l'API Manager supporta OAuth 2.0, che è uno standard industriale per l'autenticazione basata su token. Questo garantisce che solo gli utenti o i sistemi autorizzati possano accedere alle API.
- **Autenticazione basata su client:** oltre al token, è possibile utilizzare l'autenticazione basata su client per un ulteriore livello di sicurezza.
- **Interoperabilità secondo standard MODI (AgID)**: ai fini di garantire la piena interoperabilità in ottica AgID/DTD viene utilizzata un'estensione appositamente realizzata per l'API Manager che garantisce il pieno supporto della negoziazione a livello Trust tra soggetti interessati nell'interazione tramite API (https://docs.italia.it/AgID/documenti-in-consultazione/lg-sicurezza-interoperabilita-docs/it/bozza/index.html).

Supporto funzionalità IAM

La piattaforma MuleSoft Anypoint fornisce una solida funzionalità di Identity Management, essenziale per le organizzazioni che desiderano semplificare e centralizzare la gestione delle identità dei propri utenti. Il supporto per il Single Sign-On (SSO) è uno degli elementi chiave di questa funzionalità.

Uno degli standard principali supportati per l'implementazione del SSO è il **SAML 2.0** (Security Assertion Markup Language). Attraverso SAML 2.0, le organizzazioni possono facilmente integrare Anypoint con soluzioni di Identity Provider (IdP) esterne, consentendo un'esperienza di accesso unificata attraverso vari sistemi e applicazioni. La conformità a SAML 2.0 garantisce che le informazioni di autenticazione e autorizzazione vengano scambiate in modo sicuro e affidabile tra i servizi.

In aggiunta, Anypoint supporta anche **OpenID Connect**, un protocollo di autenticazione moderno basato su OAuth 2.0. Questo standard, essendo largamente adottato, assicura una facile integrazione con una vasta gamma di soluzioni di autenticazione presenti sul mercato.

Gestione dei token

È disponibile una soluzione integrata per la gestione dell'API e dei tokens associati, garantendo sicurezza, interoperabilità e scalabilità. La gestione dei tokens è fondamentale per garantire che solo le parti autorizzate abbiano accesso alle risorse. Le possibili tipologie di Tokens sono:

- **Access Token:** un token temporaneo che rappresenta l'identità dell'utente e le sue autorizzazioni. Esso ha una durata limitata e viene utilizzato per accedere alle API.
- **Refresh Token:** un token che può essere utilizzato per ottenere nuovi access tokens senza richiedere all'utente di autenticarsi nuovamente.

L'autenticazione e autorizzazione (basate su OAuth 2.0) prevede quattro flussi di autorizzazione:

- **Authorization Code Grant:** adatto per applicazioni client-server.
- **Implicit Grant:** ideale per applicazioni client-side.
- **Resource Owner Password Credentials Grant:** adatto per applicazioni di fiducia.
- **Client Credentials Grant:** utilizzato quando l'applicazione agisce per suo conto.

l'implementazione dei Tokens con MuleSoft prevede:

- **Policy di Sicurezza:** policy predefinite che facilitano l'implementazione dei tokens. Ad esempio, la policy "OAuth 2.0 protected" protegge l'API richiedendo un access token valido.
- **Anypoint Security:** sono funzionalità per la generazione, la validazione e la revoca dei tokens.
- **Token Storage:** i tokens generati e rilasciati sono conservati in modo sicuro, garantendo che siano al sicuro da accessi non autorizzati.

mentre la gestione e rinnovo dei Tokens:

- **Scadenza e Rinnovo:** gli access tokens hanno una durata limitata, dopodiché diventano non validi. Utilizzando un refresh token, gli utenti possono ottenere un nuovo access token senza doversi autenticare nuovamente.
- **Revoca dei Tokens:** gli amministratori possono revocare i tokens in qualsiasi momento attraverso il portale di gestione di Anypoint Platform.

sul tema della sicurezza e relative best practices:

- **SSL/TLS:** la raccomandazione è di utilizzare sempre SSL/TLS per proteggere le comunicazioni, in particolare quando si trasmettono tokens.
- **Rotazione dei Tokens:** è una buona pratica ruotare i tokens regolarmente per ridurre il rischio di esposizione.
- **Limitare le Autorizzazioni:** i tokens dovrebbero avere solo le autorizzazioni strettamente necessarie per la loro funzione specifica.

Throttling e Rate Limiting

Sono disponibili meccanismi avanzati per gestire e controllare il traffico delle API mediante Throttling e il Rate Limiting. Questi strumenti permettono alle organizzazioni di garantire prestazioni ottimali e proteggere le risorse backend da sovraccarico. Nel contesto di Anypoint il **Throttling** è considerato come la regolazione dinamica della velocità di richiesta in base alle condizioni di sistema, come l'utilizzo della CPU; mentre il **Rate Limiting** è l'impostazione di un numero massimo di richieste consentite per un determinato periodo.

Le funzionalità di Throttling e Rate Limiting in MuleSoft prevedono:

- **Policy predefinite:** sono policy per entrambe le funzionalità, consentendo una facile implementazione e personalizzazione.
- **Ambito:** gli utenti possono applicare limiti a livello di API, singolo endpoint o persino a livello di metodo specifico all'interno di un'API.

L' implementazione del *Throttling* prevede:

- **Policy di Throttling:** Permette di limitare le richieste in base alle condizioni di sistema. È possibile definire soglie specifiche oltre le quali le richieste verranno messe in coda o rifiutate.

- **Throttling adattivo:** MuleSoft permette di adattare dinamicamente il Throttling in base al comportamento del sistema, proteggendo le risorse in tempo reale.

mentre quella relativa al *Rate Limiting*:

- **Policy di Rate Limiting:** gli utenti possono definire un numero massimo di richieste per un intervallo di tempo specifico (ad es., 1000 richieste al minuto).
- **Rate Limiting basato sull'identità:** è possibile personalizzare i limiti in base all'identità del chiamante, come un'applicazione specifica o un token di utente.

La Gestione e Monitoraggio prevede:

- **Notifiche:** vengono generate notifiche automatiche quando si avvicina o si supera un determinato limite, prefissato a priori.
- **Dashboard:** è disponibile una dashboard dedicata per monitorare il traffico, le violazioni dei limiti e altre metriche correlate.
- **Log e Report:** gli utenti, dell'Integration Platform possono visualizzare dettagli specifici su quale chiamante ha raggiunto o superato i limiti.

in termini di vantaggi e adozione delle best practices in questo contesto:

- **Prevenzione dell'abuso:** protezione delle risorse da attacchi DDoS e abuso da parte di utenti o applicazioni malintenzionate.
- **Ottimizzazione delle prestazioni:** garanzia che le risorse siano disponibili per tutti gli utenti, evitando colli di bottiglia.
- **Pianificazione della capacità:** gli amministratori possono pianificare e scalare le risorse in base al traffico previsto.
- **Personalizzazione:** possibilità di personalizzare i limiti in base alle esigenze specifiche di ciascuna API e al comportamento previsto degli utenti.

CORS (Cross-Origin Resource Sharing)

L'API Manager permette di configurare le policy CORS, garantendo che le API siano accessibili solo dai domini approvati. Questo previene attacchi di tipo Cross-site request forgery.

Protezione contro attacchi comuni

L'API Manager offre protezione contro una serie di attacchi comuni, tra cui SQL Injection, XML External Entity Attacks e altri attacchi basati sulle API.

SSL/TLS

La comunicazione tra client e API può essere protetta utilizzando SSL/TLS, garantendo che i dati trasmessi siano crittografati e sicuri da intercettazioni.

Logging e Monitoraggio

Ogni accesso e utilizzo delle API viene registrato, permettendo un'analisi dettagliata in caso di sospette attività malevole. Questo consente anche di avere un'idea chiara delle prestazioni delle API e di identificare possibili problemi prima che diventino critici.

Policy personalizzate

Da un punto di vista operativo sono disponibili strumenti per controllare il traffico delle API applicando automaticamente policy predefinite (rate limit, throttling, OpenID Connect) o personalizzate. È possibile aggiungere o rimuovere i criteri per la sicurezza delle API, la limitazione della velocità, la memorizzazione nella cache e la gestione delle identità in fase di esecuzione senza tempi di inattività.

Tra le policy più usate disponibili ci sono: Cross-origin resource sharing, Throttling, Throttling (SLA based), Rate limiting, Client ID enforcement, HTTP Endpoint invoker, Auth header with SOAP header.

Oltre alle policy predefinite, l'API consente agli amministratori di creare policy personalizzate per rispondere a esigenze specifiche di sicurezza.

GraphQL e benefici nella gestione del payload

La crescente popolarità di GraphQL, un linguaggio di query per le API, ha portato molte piattaforme a integrare questa tecnologia all'interno dell'API Manager. GraphQL offre una flessibilità senza precedenti permettendo ai client di richiedere esattamente i dati di cui hanno bisogno. Questo riduce il traffico inutile e ottimizza le prestazioni. I principali benefici derivanti dall'adozione di GraphQL sono:

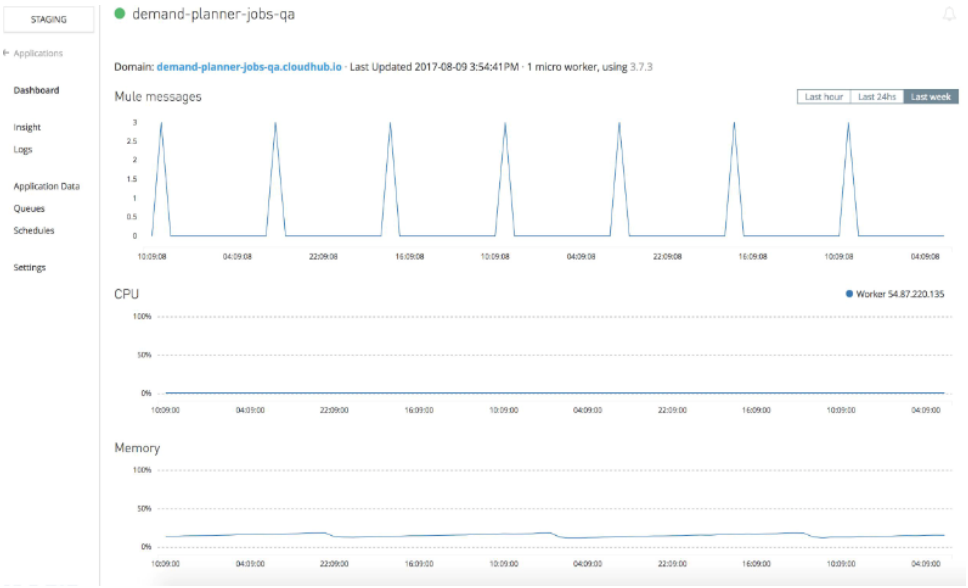
- **Efficienza**: richiedere solo i dati necessari significa minor traffico di rete e risposte più veloci.
- **Maggiore flessibilità per gli sviluppatori**: gli sviluppatori possono adattare le chiamate alle loro esigenze specifiche senza dover attendere modifiche all'API.
- **Evoluzione senza interruzioni**: GraphQL consente di deprecare campi in modo elegante, aggiungere nuove funzionalità senza interrompere le versioni precedenti e mantenere la compatibilità all'indietro.

Distribuzione

- **Deploy semplificato:** la soluzione facilita il deploy delle API sia on-premise che nel cloud, grazie a runtime integrati e opzioni flessibili di distribuzione. L'esecuzione relativa al codice delle API avviene sui componenti del Runtime Fabric MuleSoft su cluster Kubernetes a garanzia di scalabilità e resilienza.
- **Ambienti isolati:** è possibile creare ambienti isolati (come sviluppo, test e produzione) per garantire che le modifiche possano essere testate in modo sicuro prima di essere spostate in produzione.

Una caratteristica rilevante è l'uso del Runtime Fabric che permette il deployment di Anypoint su cluster Kubernetes che è uno dei requisiti dell'architettura tecnologica del SIM. Le principali caratteristiche di questa modalità di implementazione includono:

- Isolamento tra le applicazioni grazie all'esecuzione di un server runtime Mule separato per ogni applicazione.
- Possibilità di eseguire più versioni del server runtime Mule sullo stesso insieme di risorse.
- Scalare le applicazioni su più repliche.
- Failover automatico delle applicazioni.
- Gestione delle applicazioni con Anypoint Runtime Manager.



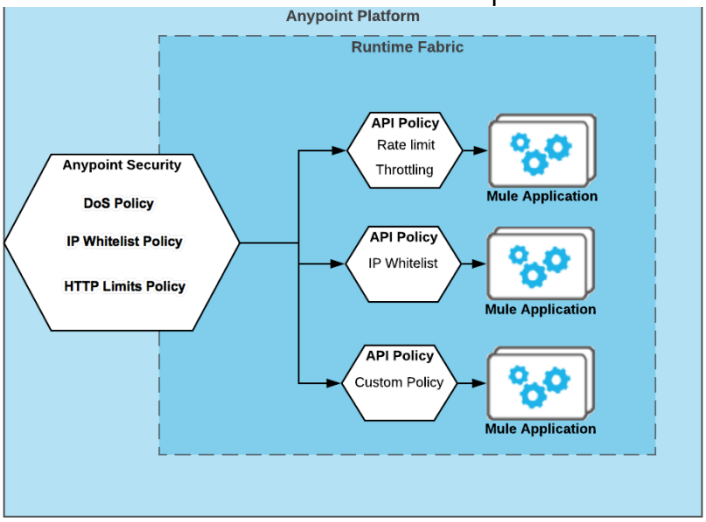
Runtime Manager

Gestione

- **API Manager:** questa componente consente di gestire, monitorare e proteggere le API. Si possono applicare policy di sicurezza, controllare l'accesso, configurare il "throttling" e il "rate limiting", già descritti precedentemente e molto altro.
- **Monitoraggio in tempo reale:** sono disponibili strumenti di monitoraggio e analisi per tenere traccia delle prestazioni delle API, degli errori, dei tempi di risposta e dell'uso. Ciò permette di identificare rapidamente i problemi e di agire di conseguenza.

Sicurezza

- **Policy di sicurezza predefinite:** è possibile applicare facilmente policy di sicurezza come già descritto precedentemente senza dover scrivere codice personalizzato.



Security Policy

Versioning e Ciclo di Vita

- **Gestione delle versioni:** la creazione e la gestione di diverse versioni delle API, è facilitato dalle funzioni integrate per sfruttare repository software esterni che permettono di apportare modifiche senza interrompere i servizi esistenti.
- **Automazione:** grazie all'integrazione con strumenti CI/CD, è possibile automatizzare il ciclo di vita delle API, dalla creazione al deploy, al monitoraggio e alla manutenzione.

Consumo

- **Portale per sviluppatori:** MuleSoft Anypoint offre un portale per sviluppatori integrato, dove gli sviluppatori di terze parti possono ricercare, testare e consumare le API in modo semplice e sicuro.

1.1.7.1.2 Message Queueing Applicativo abbinabile agli scenari di integrazione

L'Integration Platform offre le funzionalità per gestire il message queueing applicativo, particolarmente utile per garantire che i messaggi tra le applicazioni siano trasmessi in modo affidabile, specialmente in ambienti distribuiti. Le soluzioni per il message queueing abbinabili alle funzionalità di integrazione dei servizi basati su API sono:

Anypoint MQ

- **Messaging completamente gestito:** un servizio di messaggistica enterprise multi-tenant basato sul cloud, progettato per fornire alta affidabilità e scalabilità.
- **Code di messaggi:** permette di inviare e ricevere qualsiasi volume di dati, a qualsiasi livello di throughput, senza perdere messaggi.
- **Funzionalità avanzate:** include la consegna garantita, l'elaborazione persistente dei messaggi e la capacità di configurare la durata dei messaggi.
- **Integrazione con API:** è facilmente integrabile con altre API grazie all'Anypoint Platform, consentendo, ad esempio, di inoltrare messaggi da code a endpoint API.

1.1.7.1.3 Mule Runtime Engine

- **Integrazione con broker di messaggistica esterni:** supporta l'integrazione con broker di messaggistica popolari come Apache Kafka, RabbitMQ e altri attraverso connettori precostruiti.
- **Transaction Management:** garantisce che le operazioni su code e sulle risorse siano completate in modo affidabile.

Connettori precostruiti

- **Connettività plug-and-play:** i connettori facilitano l'integrazione con sistemi di coda di messaggi di terze parti, come Amazon SQS, JMS, Apache Kafka, e altri. Questi connettori possono essere utilizzati per trasferire dati tra diverse applicazioni e servizi in modo affidabile.

Flow Control

- **Elaborazione asincrona:** MuleSoft supporta la gestione di flussi asincroni, permettendo l'elaborazione in parallelo e l'uso di code temporanee per gestire grandi volumi di messaggi.

1.1.7.1.4 MuleSoft Anypoint Studio

MuleSoft Anypoint Studio si presenta come un ambiente integrato di sviluppo (IDE) specializzato nell'integrazione di applicazioni e nello sviluppo di API. Basandosi sull'IDE Eclipse, Anypoint Studio si integra nell'ecosistema di sviluppo familiare per molti sviluppatori, garantendo un'esperienza utente omogenea e riducendo la curva di apprendimento per gli sviluppatori già abituati a Eclipse.

L'interfaccia di Anypoint Studio è progettata per semplificare il processo di sviluppo. Con un insieme di strumenti chiaramente organizzati, gli sviluppatori possono facilmente navigare tra varie funzioni, dalla codifica all'integrazione. In particolare, il focus sull'integrazione di applicazioni è ben supportato da una serie di connettori predefiniti, che offrono soluzioni pronte all'uso per collegarsi a una gamma di servizi e applicazioni.

Nel contesto delle API, Anypoint Studio offre strumenti dedicati alla creazione e gestione delle specifiche delle API. L'editor di specifiche API, ad esempio, consente di definire endpoint, parametri e risposte attraverso un'interfaccia intuitiva. Questo approccio facilita la creazione delle specifiche, riducendo la complessità e il tempo necessari per stabilire e documentare le funzionalità delle API.

Mocking delle API

Una delle funzionalità di Anypoint Studio che merita particolare attenzione è la capacità di "mocking". Questo strumento consente di emulare l'effettivo comportamento delle API, permettendo agli sviluppatori di testare le loro soluzioni in un ambiente controllato, senza la necessità di un sistema esterno. Questa funzionalità è particolarmente utile per garantire che le API funzionino come previsto prima della loro effettiva implementazione, contribuendo a ridurre errori e incoerenze.

Strumenti di test e debug

Sono disponibili strumenti integrati per il testing e il debugging delle API, consentendo agli sviluppatori di verificare e correggere eventuali errori o problemi di performance.

1.1.7.2 API Manager

1.1.7.2.1 Architettura tecnica

Per l'Api Gateway è stato utilizzato Kong.

Kong dal punto di vista implementativo è un'applicazione Lua che gira su Nginx. È distribuito insieme a OpenResty, che è un insieme di moduli che estendono il modulo lua-nginx-module.

Questo crea le basi per un'architettura modulare, in cui i plugin possono essere abilitati ed eseguiti durante l'esecuzione. Nel suo nucleo, Kong Gateway implementa l'astrazione del database, il routing e la gestione dei plugin. I plugin possono risiedere in code base separate ed essere inseriti ovunque nel ciclo di vita della richiesta, tutto con poche righe di codice.

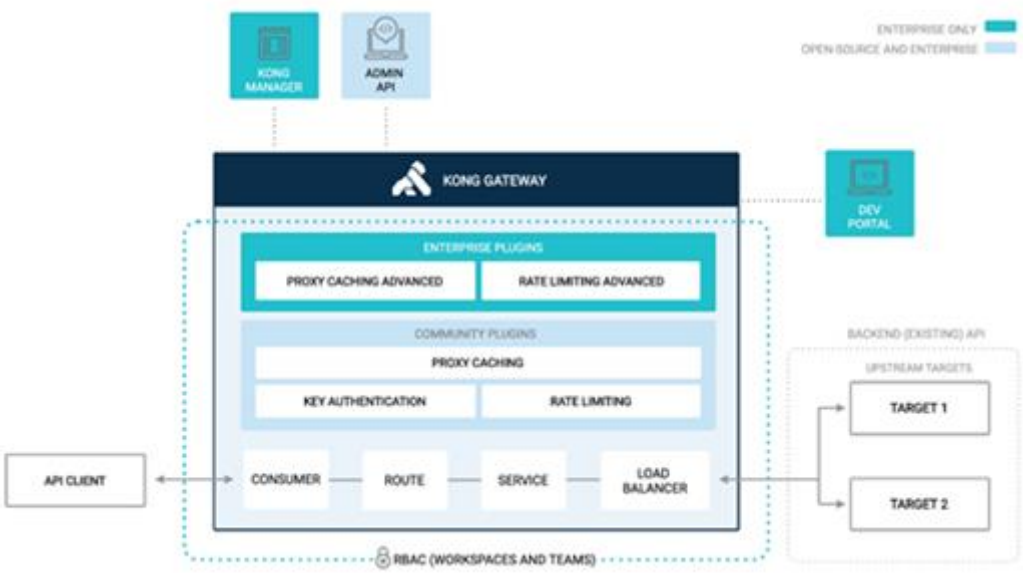


Figura 80 – API Manager, dettaglio Kong Gateway

Kong fornisce molti plugin che possono essere utilizzati nelle diverse implementazioni.

Un'istanza di plugin singola viene sempre eseguita una volta per richiesta. La configurazione con cui viene eseguita dipende dalle entità per cui è stata configurata. I plugin possono essere configurati per varie entità, combinazioni di entità o persino globalmente.

Figura 81 – API Manager, blocchi logici di Kong

Il tool comprende una serie di plugin custom, scritti in nodejs, che permettono l'interazione con lo IAM e con la parte di analitycs e monitoring.

Altri plugin, rilasciati invece ufficialmente da kong, vengono utilizzati il controllo del throttling, cors e versioning.

L'installazione comprende inoltre un'interfaccia grafica (Konga) che permette la modifica della configurazione del gateway, aggiungendo plugin o modificando rotte.

1.1.7.2.2 Infrastruttura

L'Infrastruttura di API Manager è coerente col resto dei moduli e prevede componenti (microservices) progettate per girare all'interno di containers ospitati da un Container Platform.

I POD previsti sono allocati sull'infrastruttura CaaS del PSN, in questo modo è gestito, ad esempio, l'autoscaling.

1.1.8 IAM Platform

1.1.8.1 IAM

1.1.8.1.1 Architettura Tecnica

Soluzione di alto livello

La soluzione è basata su due elementi principali:

- ***Keycloak***: per le funzionalità IAM; in particolare, si utilizza una customizzazione di prodotto.
- ***Eid-gateway***: prodotto interno per la gestione dei Federation Services.

La soluzione può essere distribuita tramite container e non prevede licenze. Di seguito sono indicati i moduli che sono presenti nella soluzione:

1) Identity Management (**Keycloak)**: supporta la gestione del ciclo di vita rilascio dell'identità.
2) Access Control & Management (**Keycloak)**: consente la gestione dei diritti di accesso e convalida degli stessi.
3) Credential Manager (**Keycloak):** consente la gestione delle credenziali e del ciclo di vita delle stesse.
4) Multy Factor Authentication (**Keycloak** e **eid-gateway)**: keycloak supporta nativamente l'integrazione con TOTP/HOTP via Google Authenticator o FreeOTP. In caso poi di integrazione con SPID, i vari identity provider mettono a disposizione differenti meccanismi di MFA.
5) Logging e Reporting **(Keycloak)**: supporta nativamente un ricco set di funzioni di audit.
6) Identity Provider Federation **(Keycloak)**: supporta nativamente l'identity brokering consentendo l'autenticazione con identity provider esterni tramite protocollo OpenID o SAML 2.
7) Spid enabling **(Eid-gateway)**: gestisce integrazione con SPID e CIE (in futuro anche eIDAS e CNS).

Componenti principali della soluzione.

Ciascun box rappresentato nella figura seguente può essere visto come un container/POD

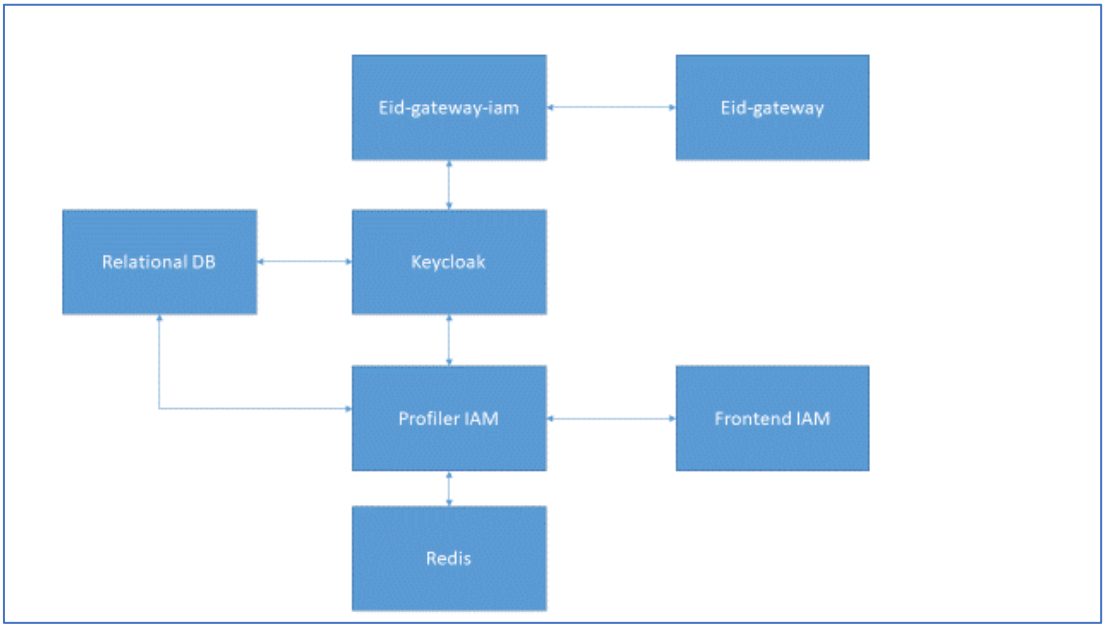


Figura 82 – Profilatore, le componenti principali della soluzione tecnologica

La customizzazione di keycloak è composta dai seguenti blocchi:

- Keycloak
- Frontend IAM
- Profiler IAM
- Redis
- Relational DB

L'Eid-gateway è composto dai seguenti blocchi

- eid-gateway
- eid-gateway-iam

Nel dettaglio la descrizione dei blocchi

- eid-gateway: gestisce le integrazioni con SPID e CIE;
- eid-gateway-iam: è il componente applicativo che fa da ponte da eid-gateway e la soluzione iam; quindi, nel nostro caso si occupa di verificare l'esistenza dell'utenza (tramite CF) su keycloak ed eventualmente ne effettua la creazione, una volta trovata la corrispondenza crea una sessione di sso su keycloak;
- Keycloak: identity and access management;
- Relational DB: Database di appoggio per keycloak, IAM e per il logging dell'eid-gateway;
- Redis: cache dei dati;
- Profiler IAM: si interfaccia con keycloak, redis e postgres;
- Frontend IAM; frontend custom per la gestione avanzata di profili, permessi e rotte applicative.

Disegno di dettaglio architettura applicativa

Keycloak

Keycloak rappresenta il cuore IAM della soluzione, è un software open source molto diffuso per l'Identity and Access Management.

Tra le features principali che presenta:

- Single Sign On: gli utenti possono effettuare login su keycloak, piuttosto che sulle singole applicazioni.
- Identity Brokering: possibilità di autenticarsi su IdP esterni tramite protocolli OpenId Connect e SAML 2.
- Social Login: possibilità di effettuare login sui principali social Network.
- User Federation: supporto per la connessione a LDAP o Active Directory Server.
- Strong authentication: supporto di OTP con Google authenticator o Free OTP.
- Protocolli Standard: supporto di OpenID Connect, OAuth 2.0, and SAML.

La soluzione è distribuita tramite container in un ambiente Open Shift ed è composta da 2 elementi principali:

- Un container contenente Keycloak
- Un container contente il database di appoggio che in questo caso è Postgres

Figura 83 – Profilatore, le componenti della soluzione keycloak

Keycloak è installato fornendo un'utenza amministrativa "locale". Con questa è possibile, per l'utente amministratore della PA, controllare tutti gli aspetti di configurazione (realms, clients, identity providers, utenti, ecc.).

Sarà quindi possibile ad esempio creare nuovi realms, utenti, ruoli o gruppi.

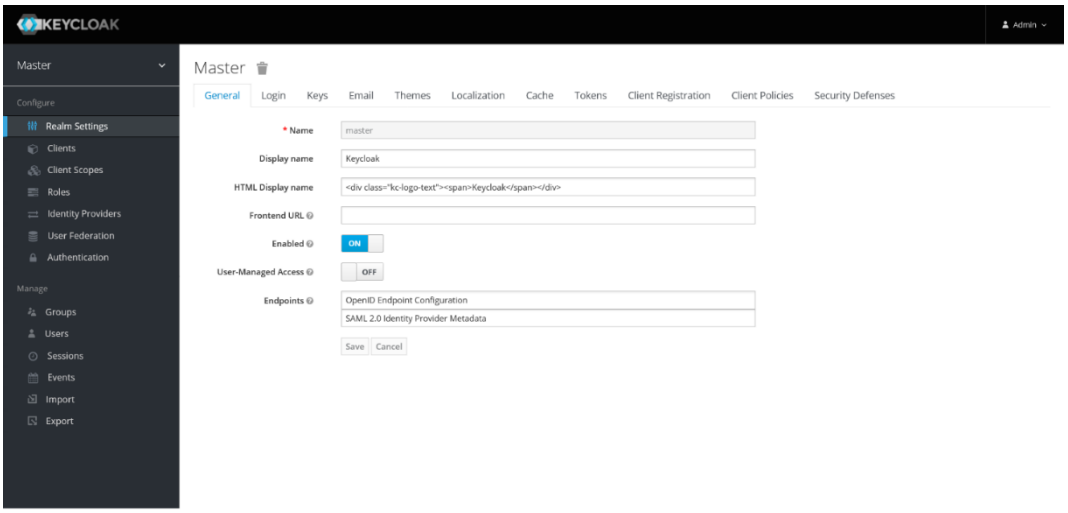


Figura 84 – Profilatore, esempio di schermata amministrativa di Keycloak

L'istanza di Postgres è installata già prepopolata con 2 database:

- keycloak database di appoggio per Keycloak
- iam database di appoggio per il backend di Profiling

Sul database keycloak sono precaricati i dati necessari all'interazione con IAM:

- un realm (kcDev01)
- un client (iam_client) configurato col protocollo OpenId connect e relativo secret
- un utente amministrativo (iam_admin) e relativo gruppo IamAdministrators

Keycloak comunica con postgres tramite un service di tipo ClusterIP sulla porta standard 5432. Postgres non è quindi visibile al di fuori della rete del cluster su cui è installata l'istanza IAM

Keycloak è esposto tramite ingress.

IAM & Profiling

È una soluzione che si propone come estensione delle funzionalità di Keycloak, fornisce una gestione avanzata della profilazione, dei permessi, delle funzioni, e delle uri accessibili.

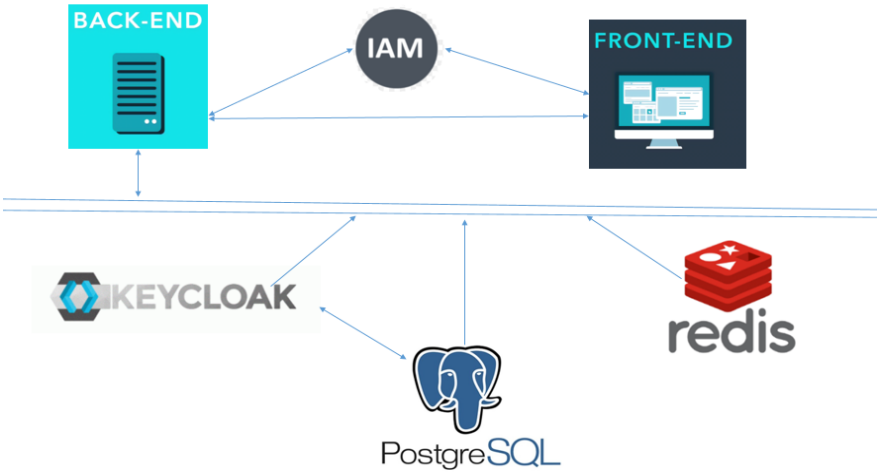


Figura 85 – Profilatore, architettura complessiva della soluzione

Lo IAM si occupa di gestire l'autenticazione e le autorizzazioni legate all'utente corrente. In particolare, l'autenticazione viene effettuata sfruttando le funzionalità del componente Keycloak che rilascia dei token e gestisce la validità degli stessi.

I token rilasciati da keycloak vengono memorizzati nella cache interna del backend e nella cache del servizio Redis e su ogni token vengono associate altre informazioni di contorno utili alla profilazione.

Il servizio di backend permette di limitare l'operabilità dell'utente in determinate entità dello IAM secondo i permessi che sono stati assegnati all'utente o al gruppo di appartenenza dell'utente.

In alternativa è possibile effettuare l'autenticazione direttamente su keycloak per sfruttare ad esempio le funzionalità di multi factor authentication.

Il token ottenuto potrà essere utilizzato per richiamare le api esposte dal backend ed ottenere informazioni aggiuntive sui permessi assegnati all'utente.

Nel complesso, l'architettura della soluzione è si basa su 3 componenti base (Keycloak e Postgres sono già stati descritti nel paragrafo precedente) e 2 componenti verticali.

Ciascun componente è distribuito come container.

Componenti di base:

POSTGRES: questo componente viene utilizzato per ospitare la base dati del servizio keycloak e del servizio di backend del profiling.

KEYCLOAK: istanza di keycloak in cui vengono creati gli account, i gruppi e rilasciati i token di accesso.

REDIS: in questo modulo viene fatta girare un'istanza di REDIS per memorizzare in cache le regole di profilazione dello IAM ed i token di accesso.

Inoltre, vengono memorizzati i token staccati da keycloak in modo da non perdere i token staccati nel caso in cui il backend venga riavviato manualmente o automaticamente a seguito di un errore.

Componenti verticali specializzati:

BACKEND: in questo modulo è presente tutta la logica del Profiling, in particolare sono presenti le api che permettono di creare/modificare/cancellare le regole di profilazione e le api che restituiscono le autorizzazioni assegnate ad un account. Gestisce inoltre la comunicazione con keycloak, redis e postgres.

FRONTEND: In questo modulo è presente l'interfaccia grafica del Profiling dove è possibile inserire le regole avanzate di profilazione utente.

Il backend comunica con Postgres, Keycloak e Redis Keycloak attraverso service di tipo ClusterIP interni al cluster.

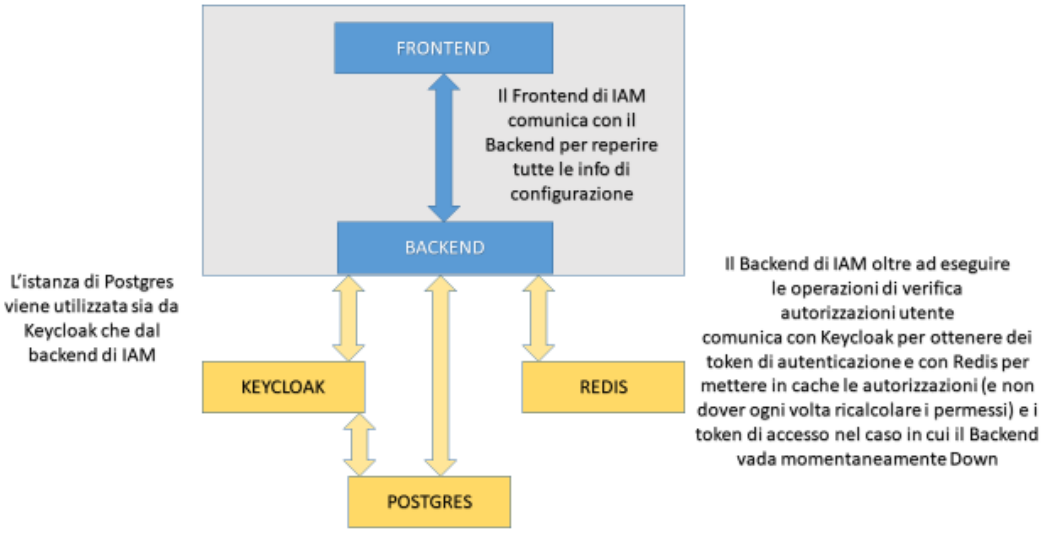


Figura 86 – Profilatore, interazione tra i componenti dello IAM

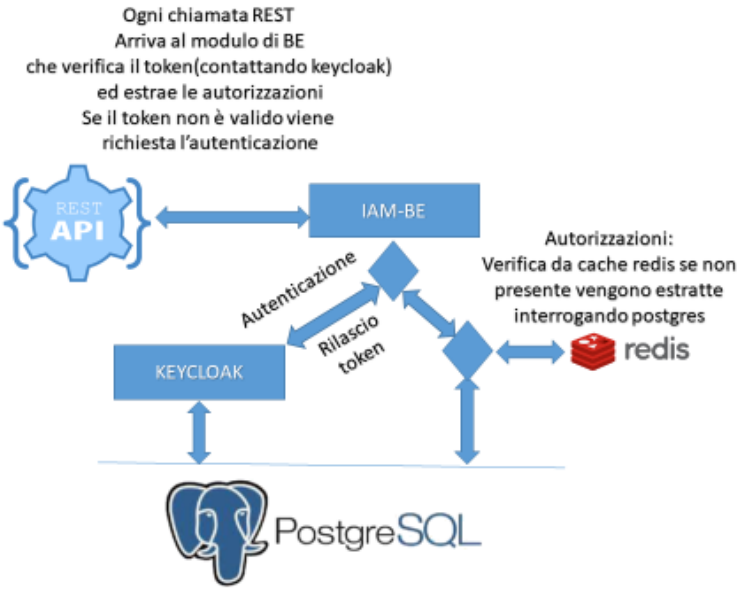


Figura 87 – Profilatore, flusso dati backend IAM

eID Gateway

Il Gateway è il componente che realizza l'integrazione tra l'infrastruttura IAM ed i sistemi, fornitori di Identità Digitali, SPID e CIE 3.0 (di seguito CieID), EIDAS e CNS e ha lo scopo di implementare le seguenti funzionalità:

- esporre verso AgID e gli Identity Provider SPID il file di metadati per SPID;
- esporre verso il Poligrafico dello Stato, gestore per conto del Ministero dell'Interno del sistema /di autenticazione CieID, il file di metadati per CieID;
- esporre un sistema di autenticazione via EIDAS;
- esporre un sistema di autenticazione via CNS basato su mutua autenticazione;
- generare secondo le specifiche del protocollo SAML e le regole tecniche di AgID le Authentication Request da inviare agli IdP SPID e CieID per inizializzare il processo di autenticazione sia dei cittadini che dispongono di una identità digitale ad uso personale che dei professionisti che dispongono di una identità digitale ad uso professionale, o di una carta d'identità elettronica;
- gestire le Authentication Response inviate dagli Identity Provider in modo da verificarne l'autenticità e controllarne la validità secondo quanto riportato dalle regole tecniche di AgID per SPID e del Ministero dell'Interno per CieID;
- effettuare l'estrazione e la decodifica e la normalizzazione delle informazioni utente inviate dagli IdP SPID o da CieID;
- gestire gli errori inviati dagli Identity Provider secondo quanto definito nelle regole tecniche di AgID per SPID e del Ministero dell'Interno per CieID;
- gestire gli errori inviati dal Poligrafico dello Stato secondo quanto definito nel manuale operativo del Ministero dell'Interno;
- permettere l'integrazione con una qualsiasi applicazione o sistema che voglia implementare l'accesso con SPID o CIE tramite l'invio dei dati dell'identità dell'utente sottoforma di token JWT;

- effettuare il tracciamento di tutti gli accessi SPID o CIE utilizzando gli appositi servizi di logging e auditing messi a disposizione dall'infrastruttura.

Il sistema garantisce l'accesso secondo il livello di autenticazione 2 (SPID Livello 2) che garantisce un alto grado di affidabilità all'identità accertata nel corso dell'attività di autenticazione. A tale livello è associato un rischio ragguardevole e compatibile con l'impiego di un sistema di autenticazione informatica a due fattori, non necessariamente basato su certificati digitali (password e OTP associati alla digitazione di una UserID).

Il sistema CieID nativamente garantisce l'accesso secondo il livello di autenticazione 3 che pertanto, essendo maggiore di quello richiesto, è perfettamente utilizzabile.

Oltre alle funzionalità implementate dal eID-Gateway sono gestite tutte le fasi previste dalla procedura tecnica di certificazione prevista da AgID ed al termine della procedura tecnica è supportata l'amministrazione anche nella procedura amministrativa per la stipula della convenzione.

Analoga gestione deve essere effettuata per la procedura di OnBoarding con il Ministero dell'Interno per l'adesione al sistema CieID.

Anche se i sistemi SPID e CieID si basano sul protocollo SAML v2 supportato dai prodotti di Access Management, il rispetto formale delle regole tecniche definite da AgID e dal MdI per l'integrazione, nonché la gestione della creazione delle Identità in modo diversificato in funzione della tipologia di utente che si sta collegando, fa sì che la soluzione più indicata per questo componente sia l'implementazione di un modulo software dedicato che si interfacci con il prodotto di Identity & Access Management. All'interno del modulo sono utilizzate le librerie opensource OpenSAML per la gestione di tutte le asserzioni previste dal protocollo SAML.

Il sistema è composto da 2 componenti principali:

- **eID-Gateway**: implementa l'integrazione con il Sistema Pubblico di Identità Digitale e CIE, può essere utilizzato in modo generico anche al difuori del contesto IAM per integrare l'accesso con SPID/CIE/EIDAS/CNS con altri portali pubblici.
- **eID-Gateway-IAM**: è il componente che implementa l'integrazione specifica tra eID-Gateway ed il sistema IAM.

Nella figura seguente viene descritto il modello logico del eID-Gateway e successivamente le modalità di integrazione con il sistema IAM.

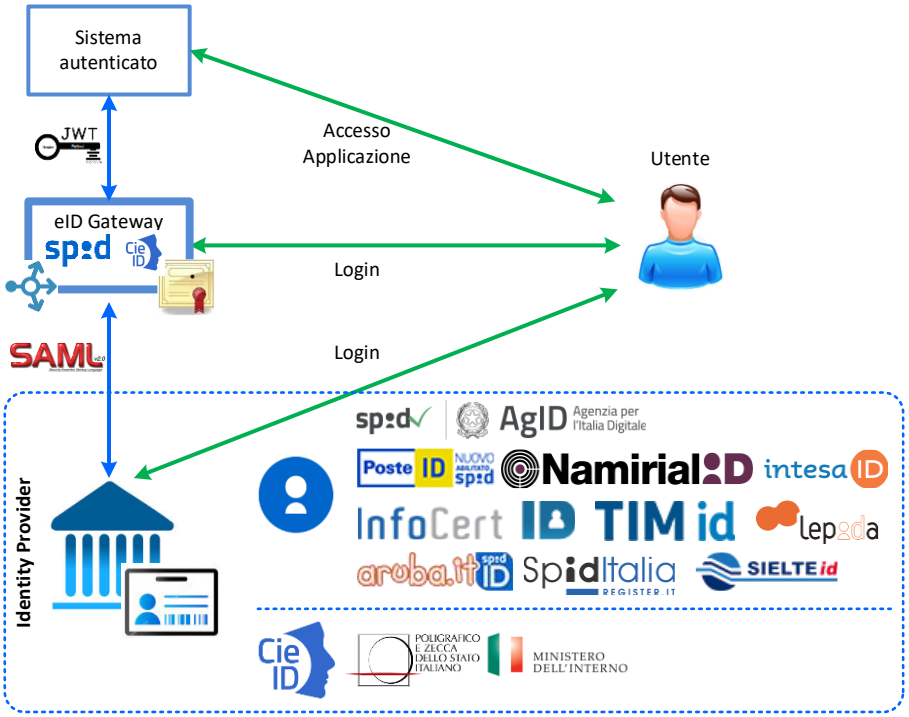


Figura 88 – Profilatore, modello logico eID-Gateway

Nello schema riportato di seguito viene descritto il flusso di integrazione tra eID-Gateway ed il resto del sistema IAM in cui dovranno essere implementate delle funzionalità specifiche per gestire il ciclo di vita dell'utenza, in particolare:

- creare un account sul sistema di Identity Management per i nuovi utenti che effettuano un primo accesso al sistema, effettuando le varie verifiche di validità in funzione della tipologia di utente con cui si sta entrando
- agganciare l'utente che ha fatto accesso con le proprie credenziali SPID o con la propria CIE al relativo account eventualmente già presente sul sistema di Identity Management;
- creare la sessione utente all'interno del sistema di SSO per gli utenti che accedono con le proprie credenziali SPID o con la propria CIE.

Per mantenere il componente eID-Gateway generico le integrazioni specifiche per IAM saranno implementate nel modulo specifico denominato eID-Gateway-IAM.

1.1.8.1.2 Infrastruttura

Come già anticipato nei capitoli precedenti la soluzione prevede un deploy completo tramite container.

La soluzione, quindi, sarà distribuita come POD Kubernetes nell'infrastruttura CaaS del PSN

Da un punto di vista logico si possono individuare sette gruppi di POD differenti:

- Keycloak
- Frontend IAM
- Profiler IAM
- Redis
- Postgres SQL
- eid-gateway
- eid-gateway-iam

Ciascun gruppo di POD sara' esposto tramite Service di tipo Cluster-IP (raggiungibili solo all'interno del cluster K8)

I servizi di Keycloak, Backend IAM, Frontend IAM, eid-gateway saranno esposti con opportuni ingress control.

Postgres e Redis saranno dotati di un PV (Persistence Volume) di supporto.

1.1.9 Digital Experience Platform

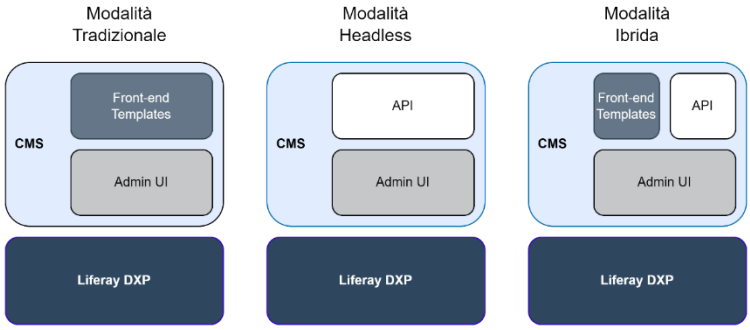
Il termine Digital eXperience Platform (DXP) è relativamente recente e identifica un nuovo modello di soluzione per la gestione integrata dell'interazione di un sistema digitale con i suoi diversi utenti o più precisamente per gestire la cosiddetta "customer journey". Liferay DXP è progettato per la scalabilità e l'integrazione, rendendolo una scelta ideale per le organizzazioni che necessitano di consolidare il loro ecosistema digitale e promuovere una cultura collaborativa e basata sui dati. Con un forte enfasi sulla sicurezza e sulla conformità, la piattaforma garantisce una navigazione sicura aderendo agli standard di settore e proteggendo i dati degli utenti.

1.1.9.1 Architettura flessibile

1.1.9.1.1 Modalità Headless

Liferay DXP può essere utilizzato in tre modalità architetturali distinte: tradizionale, headless e ibrida:

1. **Modalità Tradizionale:** in questa modalità, Liferay DXP funziona come un portale web completo, fornendo un'interfaccia utente front-end per la visualizzazione e la gestione dei contenuti.
2. **Modalità Headless**: in questa modalità, Liferay DXP funziona come una piattaforma headless, fornendo solo API per l'accesso ai dati senza un'interfaccia front-end per gli utenti finali ma con un'interfaccia utente per amministratori e redattori. Questo permette agli sviluppatori di creare applicazioni personalizzate (come applicazioni mobili o single-page applications) che utilizzano le API di Liferay per accedere ai dati gestiti (siti, contenuti, profilazioni degli utenti, etc.)
3. **Modalità Ibrida**: la modalità ibrida combina gli aspetti della modalità tradizionale e della modalità headless. In questa modalità, Liferay DXP fornisce sia un'interfaccia utente front-end che API per l'accesso ai dati. Questo permette agli sviluppatori di creare applicazioni che utilizzano le API di Liferay per accedere ai dati ed implementare funzionalità complesse ma permette l'uso del front-end di Liferay per presentare contenuto standard che non richiede particolari interazioni lato front-end.

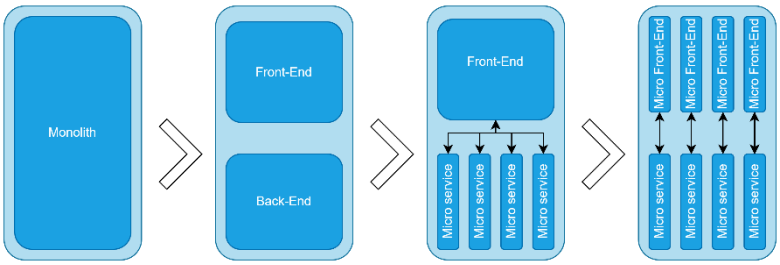


Date le caratteristiche di PIGeCO e dato, in generale, l'elevato livello di riuso che si può stabilire con i vari componenti di front-end degli applicativi del SIM la **modalità implementata sarà quella Headless** che offre una maggiore flessibilità. Questa modalità consente di eseguire lo sviluppo front-end con qualsiasi tecnologia e/o framework specializzato per gli scopi di presentation per rispondere sia alle sempre più sfidanti richieste funzionali che all'ampio insieme di tipologie dispositivi (web, mobile, tablet, kiosk, smartwatch) che l'utenza finale richiede di utilizzare.

Lo strumento che abilita questa flessibilità è costituito da una serie di API (Application Programming Interface) che permettono, in maniera sicura, di integrare facilmente il CMS con altri sistemi esterni, quali siti web, applicazioni mobile native, chatbot, ecc. oppure con sistemi interni, quali la Intelligence Platform, il sistema di Identity Access Management, ecc. A livello di API di Liferay DXP segue le

specifiche OpenAPI, il framework open source standard de facto per le API RESTful e il query language GraphQL per soddisfare le richieste più avanzate dei client.

1.1.9.1.2 Micro Front-end



L'architettura tecnologica degli applicativi del SIM e quindi di PIGeCO prevede l'utilizzo della modalità Headless di Liferay DXP integrata con componenti client sviluppati secondo il pattern Micro Front-end. Quest'ultimo applica la logica di disaccoppiamento solitamente applicata in ambito back-end per scomporre gli applicativi monoliti in microservizi alle componenti di Front-end. In questo mondo una applicazione front-end monolitica viene scomposta in componenti più piccoli, riusabili e più facilmente gestibili che possono essere facilmente sviluppati, testati e messi in esercizio in maniera indipendente.

Esattamente come i microservizi back-end anche i Micro Front-end supportano pienamente l'applicazione di metodologie agile di sviluppo iterativo, incrementale e con time-to-market ridotto. Altri vantaggi sono i seguenti:

I Micro front-end offrono diversi vantaggi rispetto ai front-end monolitici. Ecco alcuni dei principali:

- **Maggiore scalabilità**: le singole parti possono scalare in maniera indipendente
- **Rapidità e agilità di sviluppo**: i team sviluppano e rilasciano in modo indipendente le proprie porzioni di codice.
- **Indipendenza nello sviluppo**: Questo si traduce nella possibilità non solo di sviluppare parti diverse in contemporanea, ma anche di aggiungere, rimuovere o riscrivere parti del front-end senza che questo ne intacchi la stabilità o il funzionamento.
- **Codebase gestibile**: una codebase più piccola è più facilmente gestibile.
- **Framework adatti**: la possibilità di utilizzare framework diversi all'interno dello stesso applicativo selezionando quello più adatto allo scopo.
- **Maggior riuso**: componenti progettati in modo da concentrarsi su singole funzionalità possono essere più facilmente oggetto di riuso

1.1.9.1.3 Mobile

Liferay DXP 7.4 offre il supporto per lo sviluppo di applicazioni mobile sia in modalità nativa che in modalità ibrida:

- **Modalità nativa**: per questa modalità Liferay offre due soluzioni distinte e complementari:
  - un insieme di componenti visuali denominati Screenlet che costituiscono gli elementi di User Interface specificatamente progettati per il mobile e per integrarsi con istanze server di Liferay;

  - un SDK Mobile che permette di comunicare in maniera agevole e rapida con i servizi offerti dalle istanze Liferay e che consente di integrare rapidamente le API di Liferay in quanto gestisce autonomamente gli aspetti di comunicazione e interpretazione delle risposte:
- **Modalità ibrida**: Liferay supporta questa ulteriore modalità che prevede un contenitore nativo al cui interno vengono eseguite applicazioni basate su tecnologie Web. Il vantaggio principale di questa modalità consiste nel disporre di una sola "code base" per entrambe le piattaforme attualmente più diffuse (Android e iOS)

Infine, la modalità Headless di esecuzione della DXP Platform supporta pienamente le due modalità descritte abilitando anche scenari di riuso delle API messe a disposizione da Liferay stesso sia per le app realizzate per i dispositivi mobile sia per gli applicativi acceduti tramite browser web o mobile.

1.1.9.1.4 Creazione di una Comunità Tematica nazionale

Come anticipato, Liferay DXP svolge un ruolo cruciale nella creazione di una "Comunità Tematica" e nella sua operatività quotidiana, grazie alle sue funzionalità avanzate e alla sua flessibilità come piattaforma di gestione dei contenuti e di collaborazione. In aggiunta alle funzionalità di CMS, infatti, esso dispone di una Collaboration Suite di progettata per facilitare la comunicazione e la collaborazione tra gli utenti all'interno di una o più organizzazione. Ecco una panoramica delle principali funzionalità offerte:

- **Blogs**: Gli utenti possono creare, pubblicare e gestire blog, permettendo la condivisione di notizie, aggiornamenti e idee con la comunità.
- **Forums e Discussioni**: Gli utenti possono avviare discussioni, rispondere ai post e interagire con altri membri della comunità attraverso forum tematici.
- **Wiki**: Una piattaforma collaborativa dove gli utenti possono creare, modificare e organizzare contenuti in modo collettivo, ideale per la documentazione e la condivisione di conoscenze.
- **Messaggistica Istantanea e Notifiche**: Gli utenti possono inviare messaggi privati ad altri membri e ricevere notifiche in tempo reale su attività e interazioni rilevanti.
- **Calendario**: Gli utenti possono programmare, gestire e condividere eventi, riunioni e appuntamenti, facilitando la coordinazione e la pianificazione.
- **Annunci**: Gli amministratori o gli utenti autorizzati possono pubblicare annunci o notizie importanti visibili a tutti o a gruppi specifici di utenti.
- **Polls**: Gli utenti possono creare sondaggi e questionari per raccogliere feedback o opinioni dalla comunità.
- **Commenti e Valutazioni**: Gli utenti possono commentare e valutare contenuti come blog, articoli e documenti, promuovendo l'interazione e il feedback.
- **Ricerca Avanzata**: Gli utenti possono cercare contenuti, persone e risorse all'interno della piattaforma, grazie a una potente funzionalità di ricerca.
- **Personalizzazione e Dashboard**: Gli utenti possono personalizzare la loro dashboard, organizzando le informazioni e gli strumenti secondo le proprie esigenze.

Queste funzionalità, combinate con la flessibilità e la scalabilità di Liferay DXP, rendono la Collaboration Suite uno strumento potente per organizzazioni e comunità che desiderano migliorare la comunicazione, la collaborazione e la gestione delle informazioni. Per implementare la Comunità

Tematica all'interno di PIGeCo è possibile quindi utilizzare le funzionalità di creazione di nuovi contenuti, modifica degli stessi e condivisione con gli altri utenti. Questi contenuti possono essere strutturati e presentati in modalità ampliamente diffuse come blog, forums o wiki inserendo anche strumenti di coinvolgimento interattivo dei fruitori quali commenti, valutazioni, sondaggi e notifiche. Tutti ciò favorisce la condivisione e la collaborazione tra i partecipanti e consente un loro ingaggio attivo e continuativo nel tempo concorrendo, così, al raggiungimento degli obiettivi della Comunità Tecnica quali la determinazione di innovazioni di settore e la fertilizzazione incrociata delle esperienze.

1.1.10 Resource & IOT Platform

1.1.10.1 Resource Manager

1.1.10.1.1 Architettura tecnica

Si divide in tre moduli:

- Taxonomy Manager
- Asset Manager
- Abstraction Layer

Il componente Taxonomy Manager si occupa della gestione delle tassonomie all'interno del Resource Manager. Consente la definizione di librerie di tassonomie che consentano di definire le risorse necessarie ai diversi ambiti o casi d'uso. È inoltre possibile modificare ed estendere le tassonomie esistenti.

Il componente Asset Manager gestisce l'inventario delle risorse secondo le tassonomie caricate nel sistema.

Si prevede che gli assets associati all'Asset Manager provengano dai sottosistemi integrati e che possano essere creati editorialmente dagli operatori.

Il modulo si interfaccia con il modulo di abstraction layer per la gestione degli assets e di eventuali notifiche real-time, con il layer GIS per la notifica degli update sugli assets.

Il Resource Abstraction Layer consente di fornire un livello di astrazione delle risorse erogate dai sottosistemi implementando direttamente i protocolli di comunicazione specifici e uniformando l'accesso a risorse della stessa tipologia. Consentirà ad esempio di accedere ad una risorsa di tipo telecamera o drone indipendentemente dal prodotto specifico utilizzato per implementare il sottosistema di controllo video.

L'architettura del Resource Manager è costituita dai seguenti componenti:

- Asset Manager (AM): è il microservizio che gestisce delle risorse (sia interne che esterne).
- GeoView Manager (GVM): è il microservizio che gestisce una particolare vista delle risorse, adeguata alla loro visualizzazione su cartografico.

- Taxonomy Manager (TXM): è un microservizio che gestisce le tassonomie, ossia i modelli di tutte le risorse della soluzione.
- Abstraction Layer (AL): è il microservizio che gestisce l'interazione con i sottosistemi e le risorse esterne alla piattaforma.
- Resource Manager HMI (RM-HMI): è un componente micro-fe che si interfaccia con i microservizi del Resource Manager.
- Resource Manager Widget (RMW): è un web component per la visualizzazione delle risorse, che si interfaccia con i moduli di RM (in particolare AM e TXM) ed è utilizzabile anche da componenti esterni.

La comunicazione fra i servizi avviene in modo:

- **sincrono** – mediante l'esposizione di api rest che vengono registrate all'interno dell'api manager
- **asincrono** – mediante la subscription a specifiche code kafka

L'immagine seguente esemplifica quanto sopra detto.

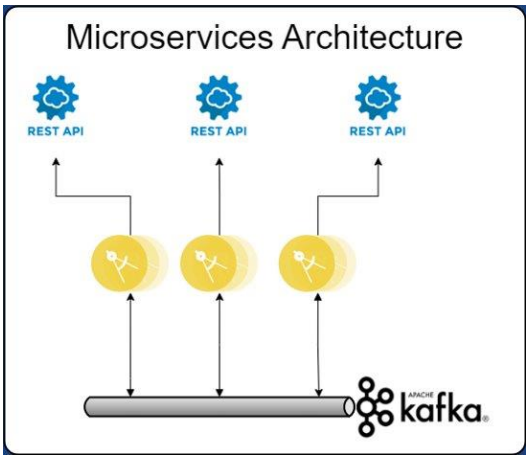


Figura 89 – Resource Manager, schema logico della comunicazione tra i servizi

Ciascun microservizio di back-end (Abstraction Layer, Asset Manager e Taxonomy Manager) mantiene le informazioni su una propria base dati. Abstraction Layer ed Asset Manager usano database non relazionali MongoDB, mentre Taxonomy Manager si appoggia ad un database relazionale Postgres.

La figura seguente mostra le interazioni tra i vari componenti. Le frecce azzurre indicano interazioni sincrone tramite REST API, quelle arancio indicano scambio di messaggi asincroni tramite bus Kafka.

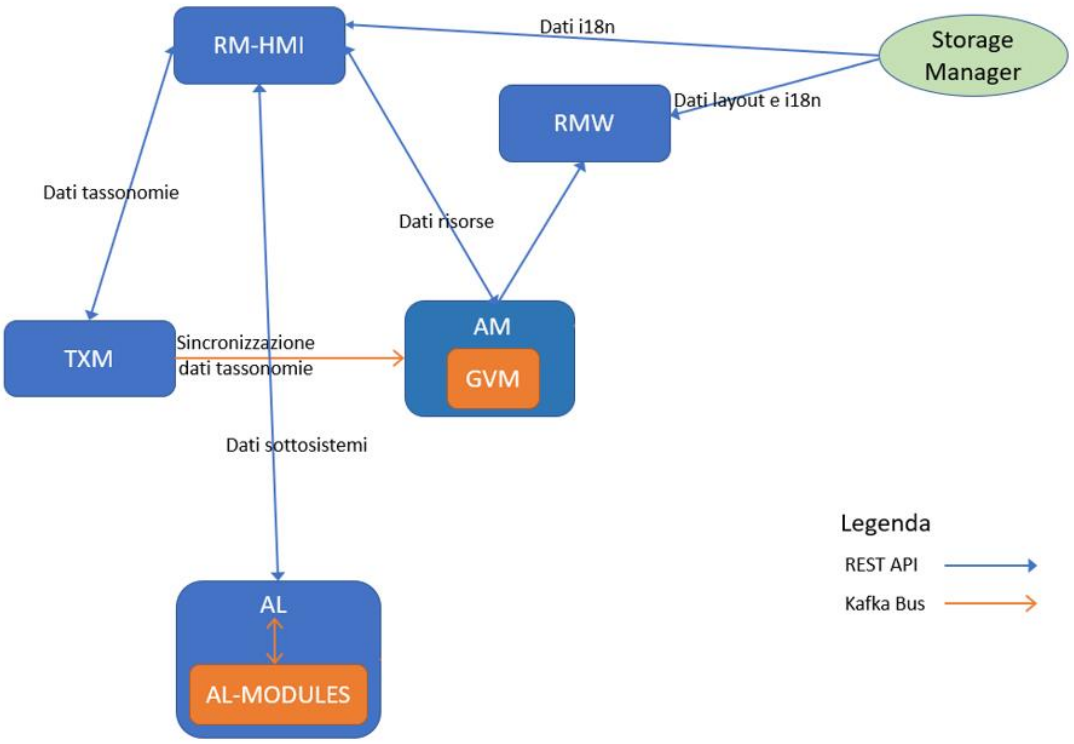


Figura 90 – Resource Manager, interazioni tra i componenti logici

I componenti di Resource Manager sono mostrati in blu e sono evidenziati in arancio gli eventuali componenti interni. In verde sono mostrati invece i componenti esterni al Resource Manager.

Il funzionamento del modulo di Abstraction layer merita un approfondimento particolare.

Questo servizio è un processo autonomo di aggiornamento dei dati di resource manager da parte dei sistemi esterni. Abstraction Layer espone una architettura a "plugin", ognuno specializzato nella comunicazione con uno specifico sistema esterno. Possono essere configurate all'interno del servizio diverse istanze di diversi plugin, ognuno per gestire uno specifico sottosistema. Questo approccio permette di gestire in modo flessibile ogni formato esposto dai sistemi esterni in termini di formato dati (json, xml, csv etc) protocollo (http/s, file…) e metodologia di trasmisisone (push, polling, scheduling…).

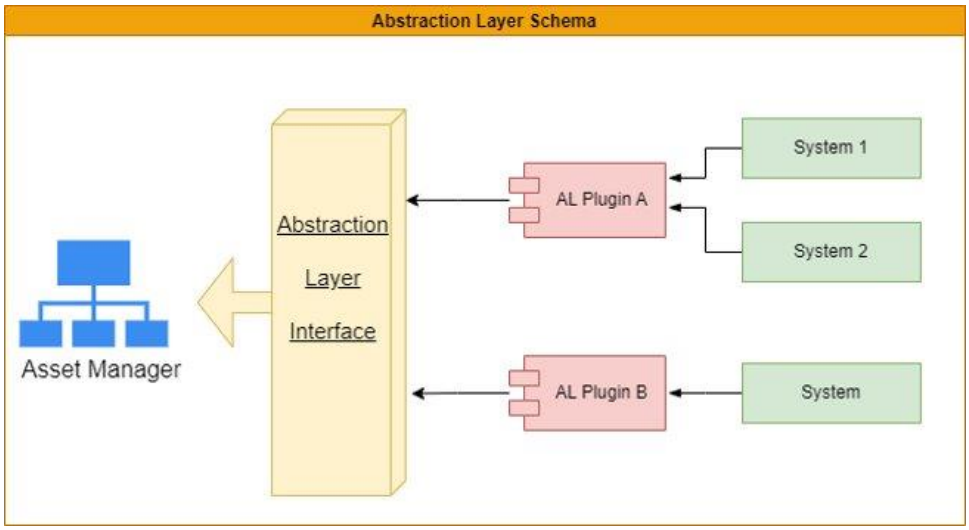


Figura 91 – Resource Manager, schema Abstraction Layer

Abstraction layer coordinerà le informazioni provenienti dai plugin per trasmettere in modo asincrono questi aggiornamenti verso il servizio di asset manager, che esporrà i dati aggiornati verso tutti gli altri servizi.

In sintesi:

- il Widget riceve tramite API REST i dati della risorsa da Asset Manager e i dati sul rendering del layout e internazionalizzazione dall'Object Storage;
- La HMI scambia dati mediante API REST con tutti i componenti di back-end per la gestione di tassonomie (Taxonomy Manager), risorse (Asset Manager) e sistemi esterni (Abstraction Layer). Inoltre riceve da Storage Manager i file per l'internazionalizzazione;
- Asset Manager ed Abstraction Layer inviano in modalità asincrona al modulo Notification Manager (NM) le notifiche riguardanti risorse e relative notifiche ed eventi;
- Abstraction Layer invia inoltre eventuali dati provenienti dai sistemi esterni, nel formato originale di ricezione, a Historical Data Manager;
- Taxonomy Manager sincronizza con Asset Manager e Abstraction Layer tramite bus kafka i dati aggiornati delle tassonomie;
- Asset Manager riceve dal Geo View Service gli aggiornamenti della posizione delle risorse.

Preferred Technology Platform: Custom Microservices, MongoDb, Kafka

Servizi

Le interfacce esposte sono:

Asset Manager Interface

In particolare, le API consentono di:

- Gestire (CRUD tramite api o topic) la georeferenziazione in formato GeoJSON

- ricercare una risorsa in base alla sua posizione (nell'intorno di un punto o entro un poligono)
- Gestire (tramite api o topic) lo stato
- Gestire (CRUD tramite api o topic) le metrics
- Eseguire una ricerca semplice per tassonomia, displayName, ambito o tipo (HUMAN, INFRASTRUCTURE, ecc)
- Ricercare una risorsa per:
  - Tipo (HUMAN, PHYSICAL, ANIMAL, TOOL, INFRASTRUCTURE, TRANSPORTATION, KIT, TEAM, ORGANIZATION)
  - Properties
  - Stato
  - Capabilities
  - Tassonomia
  - Id del sistema esterno di appartenenza
  - Coppia id-sottosistema e id-risorsa nel sottosistema

GeoView Manager Interface

Le API esposte consentono di:

- Ottenere la lista dei sottosistemi
- Ottenere una lista paginata di GeoViews
- Listare le GeoView appartenenti ad un dato sottosistema
- Ottenere le informazioni su una singola GeoView, in base al suo id o in base al sottosistema

Taxonomy Manager Interface

Le seguenti funzionalità sono disponibili tramite l'interfaccia REST API per le tassonomie:

- CRUD delle tassonomie
- CRUD delle properties
- CRUD delle capabilities e degli attributi
- Funzionalità di import/export di una libreria di tassonomie, comprensive di properties e capabilities, sia completa che parziale (dato un elenco di nomi di tassonomie)
- Funzionalità di import/export delle sole properties, capabilities (sia completa che parziale), oppure l'insieme delle due
- Funzionalità di import/export di file di configurazione in formato JSON

Abstraction Layer Interface

Il microservizio è produttore su vari topic. In particolare:

- Produttore sul topic "ral-external-metadati". Su questo topic vengono immessi attualmente i metadati prodotti dagli algoritmi del sottosistema Ganimede nel formato originale proveniente dal sottosistema.
- Produttore sul topic "ral-external-event". Su questo topic vengono immesse le notifiche generate dai sottosistemi. Il formato del messaggio è il seguente:
- Il body del messaggio dipende dal sottosistema e dal tipo di evento.

- Produttore sul topic "ral-external-resource". Su questo topic vengono immessi i messaggi per la creazione delle risorse generate dai sottosistemi. Il body del messaggio è un oggetto ResourceUpdateDTO visto in precedenza.
- Produttore sul topic "rsm-current-resource-status". Su questo topic vengono immesse le modifiche allo stato delle risorse, provenienti dai sottosistemi. Il formato del messaggio è il seguente:

```
{
"activityState": String,
"resourceState": String
}
```

- Produttore sul topic "rsm-current-resource-notification". Su questo topic vengono immesse le notifiche ricevute dal modulo abstraction, provenienti dai sottosistemi. Il formato del messaggio è il seguente:

```
{
"topic": String,
"systemUuid": "XXXXXXXX-YYYY-ZZZZ-JJJJ-KKKKKKKKKKKK",
"origin": String,
"message": Object
}
```

- L'oggetto "message" ha il formato originale del sottosistema.
- Produttore sul topic "ral-external-resource-update". Su questo topic vengono immessi i messaggi di modifica risorsa del modulo abstraction, provenienti dai sottosistemi. Il body del messaggio è un oggetto ResourceUpdateDTO.
- Produttore e Consumatore sul topic "am-txm-update". Questo topic è utilizzato per la sincronizzazione interna tra Asset Manager, Abstraction Layer e Taxonomy Manager.

Consumatore sul topic "txm-am-data". Questo topic è utilizzato per la sincronizzazione delle tassonomie tra Asset Manager, Abstraction Layer e Taxonomy Manager.

Infrastruttura

L'Infrastruttura di Resource manager è coerente col resto dei moduli e prevede componenti (microservices) progettate per girare all'interno di containers ospitati da un Container Platform.

I POD previsti sono allocati sull'infrastruttura CaaS del PSN, in questo modo è gestito, ad esempio, l'autoscaling.

Lo strato Data Store delle due soluzioni è ospitato dal blocco logico Data System (RDS) del SIM e quindi dalle piattaforme e dai servizi del PSN. Sono coinvolti il PaaS DB e le eventuali componenti infrastrutturali quali, ad esempio, il file system.

1.1.10.2 Iot Platform

1.1.10.2.1 Architettura tecnica

L'IoT Platform è costituito da una serie di componenti distribuibili anche come contenitori Docker.

Il diagramma seguente rappresenta l'architettura (High Level) proposta per soddisfare i requisiti del modulo.

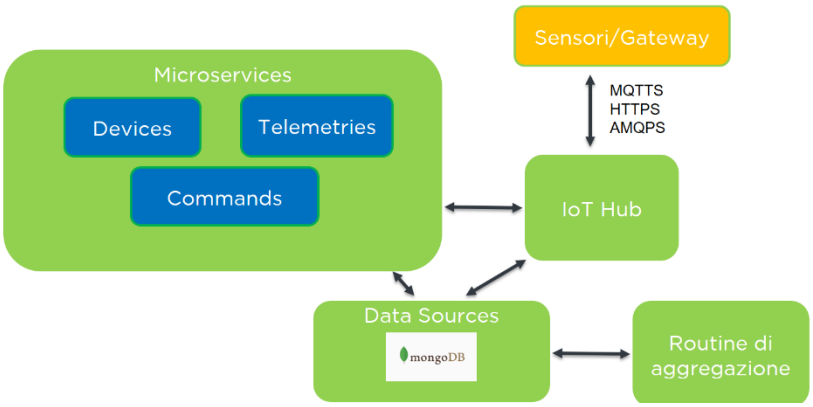


Figura 92 - Componenti IoT Platform

IoT Hub: Soluzione, basata su Eclipse Hono e Kafka, che si occupa dell'interfacciamento con i sensori e le componenti di campo (gateway). Esso fornisce interfacce di servizio remoto per la connessione di un gran numero di dispositivi IoT a un back-end e che consente di interagire in modo uniforme tramite API, indipendentemente dal protocollo di comunicazione del dispositivo. IoT Hub supporta dispositivi che comunicano tramite protocolli IoT comuni come HTTPS, MQTTS, AMQPS. Questo componente possiede anche un'istanza kafka, che è una soluzione open source in grado di ricevere ed elaborare milioni di eventi al secondo basandosi sul paradigma Producer – Consumer

Data Source: Repository in cui vengono persistite, salvate tutte le anagrafiche, le configurazioni del modulo e le telemetrie ricevute dal campo. Per questo componente si è scelto di utilizzare si è scelto di utilizzare MongoDB, un DBMS non relazionale, che mette a disposizione la possibilità di effettuare query ad hoc, potendo restituire solo certe parti di un determinato documento, garantendo l'aggregazione di dati (tramite MapReduce o Aggregation Framework) e l'alta affidabilità di mantenimento del dato, grazie ai replica set.

Custom Microservices: Micro servizi in cui sono presenti tutte le funzionalità di business del modulo che permettono di leggere, scrivere, ricercare e persistere tutti le informazioni veicolate nel modulo

Routine aggregazione: Routine di aggregazione utile al raggruppamento delle telemetrie d'interesse e per finestra temporale note

Iot HUB: Il componente vitale all'interno del modulo che si occupa dell'interfacciamento con i sensori e le parti di campo (gateway) è l'IoT Hub, basato su Eclipse Hono e Kafka.

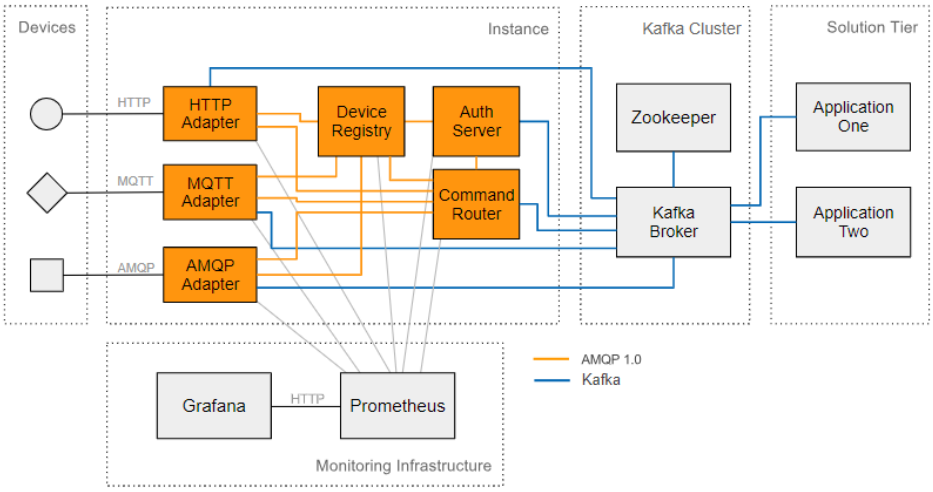


Figura 93 - IoT Hub

IoT Hub – Componenti IoT avente:

- Istanza HTTP Adapter che espone le API di telemetria ed eventi come risorse URI.
- Istanza MQTT Adapter che espone le API di telemetria ed eventi come una gerarchia di argomenti MQTT generica
- Istanza AMQP Adapter che espone le API di telemetria ed eventi come un set di indirizzi AMQP 1.0.
- Istanza Command Router che riceve messaggi di comando e controllo e li inoltra ai Protocol Adapter (HTTP, MQTT, AMQP).
- Istanza Device Registry che gestisce le informazioni di registrazione e invia asserzioni di registrazione del dispositivo ai Protocol Adapter (HTTP, MQTT, AMQP).
- Istanza Auth Server che autentica i componenti IoT ed emette token che asseriscono identità e autorità.

IoT Hub - Cluster Kafka

- Istanza del broker Apache Kafka a cui si connettono le applicazioni downstream per utilizzare dati ed eventi di telemetria dai dispositivi e per inviare messaggi di comando e controllo ai dispositivi
- Istanza di Apache Zookeeper richiesta dal cluster Kafka

A questo componente si aggiunge la componente di storage (MongoDB) e la componente a Micro servizi.

Questa componente permette di leggere, scrivere, ricercare e persistere tutti le informazioni veicolate all'interno del modulo. Di seguito si riportano i principali servizi in esso contenuti:

- Servizio REST API Devices che gestisce le operazioni in CRUD (Create, Read, Update e Delete) e la possibilità di ricerca sui dispositivi/sensori;
- Servizio REST API Telemetries che gestisce le operazioni in CRUD (Create, Read, Update e Delete) e la possibilità di ricerca sulle telemetrie associate ai dispositivi/sensori;

- Servizio REST API Commands che gestisce i comandi associati ai dispositivi/sensori

I dispositivi, quindi, si connetteranno ai *Protocol Adapter* (HTTP, MQTT o AMQP) per pubblicare dati ed eventi di telemetria. Tali dispositivi si autenticheranno utilizzando le informazioni archiviate nel *Device Registry*. I dati verranno poi inoltrati a valle alle applicazioni utilizzatrici tramite il broker Kafka.

Autenticazione/Autorizzazione

Di seguito è descritto il funzionamento dell'autenticazione e dell'autorizzazione di dispositivi, consumatori (applicazioni back-end) e componenti di sistema nell'IoT Platform.

Requisiti

- I dispositivi vengono autenticati e autorizzati quando si connettono a un Protocol Adapter
- I componenti del sistema vengono autenticati e autorizzati quando si connettono tra loro
- Le credenziali e le regole di autorizzazione possono essere gestite centralmente, ovvero non è necessario configurare manualmente le credenziali e le regole per ciascun componente

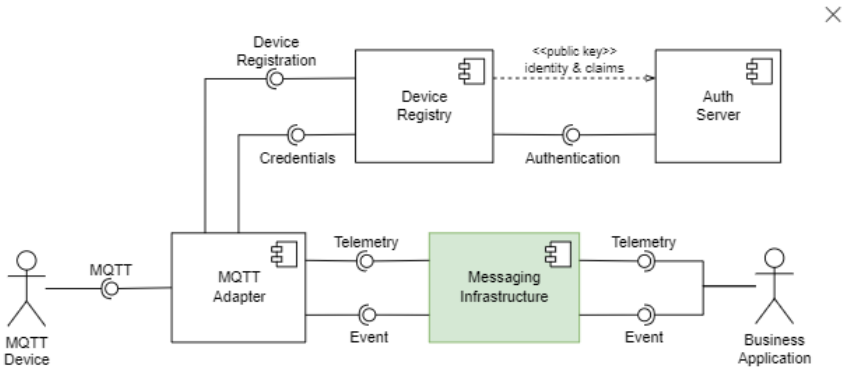


Figura 94 - Componenti IoT Platform

Autenticazione dispositivo

Per impostazione predefinita, sia *l'HTTP Adapter* che *l'MQTT Adapter* richiedono che i dispositivi eseguano l'autenticazione durante la creazione della connessione. Entrambi si basano sulle API esposte dall' *Auth Server* per facilitare la verifica delle credenziali fornite da un dispositivo.

Autenticazione componente di sistema

I componenti client che aprono una connessione AMQP a un componente server vengono autenticati utilizzando SASL PLAIN come specificato in RFC 4422. Il componente server accetta le informazioni di autenticazione fornite dal componente client e apre una connessione all'*Auth Server*, utilizzando le credenziali fornite dal client. Una volta avvenuta l'autenticazione, l'*Auth Server* emette un JSON Web Token (JWT) affermando l'identità del client e le autorità concesse al componente server. Il componente server collega quindi questo token alla sua connessione AMQP con il client e da quel momento in poi lo utilizza per prendere decisioni di autorizzazione relative alle richieste del client.

In base ai componenti mostrati sopra (figura 3), il seguente diagramma di flusso mostra come *l'MQTT Adapter* si connette al registro del dispositivo e viene autenticato in modo trasparente utilizzando l'*Auth Server*.

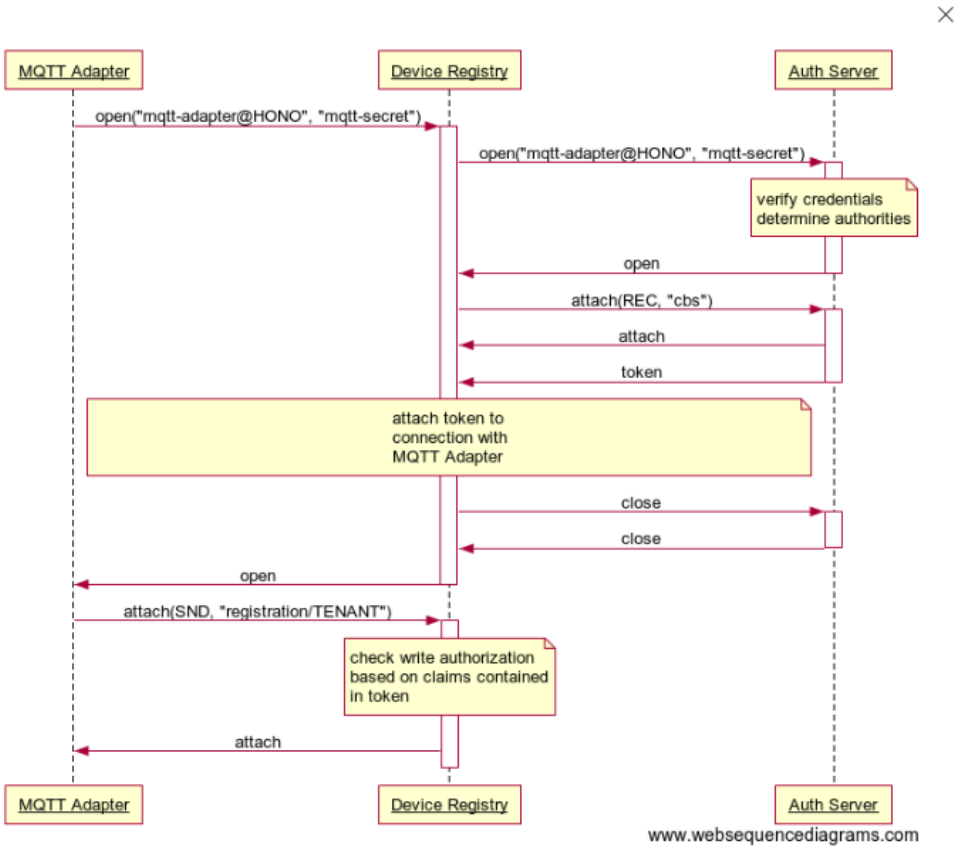


Figura 95 - Diagramma di flusso autenticazione MQTT Adapter

I componenti di sistema vengono autorizzati ogni volta che aprono un nuovo collegamento AMQP su una connessione esistente al server. Quando un client tenta di aprire un collegamento ricevente, il server controlla se il client è autorizzato a leggere dall'indirizzo di origine che il client ha specificato nel frame di collegamento AMQP. Analogamente, quando un client tenta di aprire un collegamento mittente, il server controlla se il client è autorizzato a scrivere sull'indirizzo di destinazione.

Le implementazioni del servizio possono inoltre autorizzare singoli messaggi (di richiesta) per indicare l'operazione da invocare. In tal caso il server verifica se il componente di sistema è autorizzato a eseguire l'operazione indicata dall'oggetto del messaggio sull'indirizzo di destinazione del collegamento.

Autenticazione applicazione

Le applicazioni aziendali si connettono all'infrastruttura di messaggistica per consumare dati ed eventi di telemetria e inviare comandi ai dispositivi. È pertanto responsabilità dell'infrastruttura di messaggistica basata su Apache Kafka autenticare e autorizzare adeguatamente l'applicazione.

Il broker Kafka, infatti, è responsabile dell'autenticazione delle connessioni dalle applicazioni. A questo scopo, il broker può essere configurato per autenticare le applicazioni utilizzando meccanismi SASL arbitrari. L'accesso agli indirizzi per la ricezione dei messaggi può essere limitato a determinate identità.

Apache Kafka è una piattaforma distribuita di streaming di eventi progettata per un throughput molto elevato fornendo allo stesso tempo determinate garanzie nell'ordine dei messaggi. Si adatta perfettamente ai requisiti dell'IoT Platform in merito alla messaggistica per diversi motivi. I principali vantaggi sono:

- Scalabilità: più server Kafka vengono gestiti contemporaneamente come un cluster che può essere espanso secondo necessità. Kafka è ottimizzato per consentire a molti processi di scrivere e leggere dati contemporaneamente
- Ordinamento dei messaggi: una delle caratteristiche principali di Kafka è la partizione dei dati tramite una chiave di partizione, che può essere utilizzata per selezionare i dati per i quali deve essere mantenuto l'ordine e i dati che possono essere elaborati in parallelo. All'interno del modulo, questo ci consente di garantire facilmente che tutti gli eventi provenienti da un dispositivo vengano forniti all'applicazione aziendale nell'ordine corretto. I dati di telemetria del dispositivo possono essere elaborati in parallelo e anche il loro ordine è garantito. I messaggi provenienti da tutti gli altri dispositivi possono essere consumati in parallelo, in modo completamente indipendente, anche in più istanze dell'applicazione aziendale utilizzata contemporaneamente.
- Adozione su larga scala: Kafka è utilizzato da un gran numero di aziende ed è supportato da molte altre tecnologie. Diverse aziende offrono Kafka come servizio prenotabile su diverse piattaforme cloud. Ciò semplifica la fornitura di un sistema di messaggistica gestito in modo professionale. Viene inoltre evitato il lock-in del fornitore.

Identità del dispositivo

Lo scopo principale della IOT Platform è fornire un'interfaccia uniforme affinché le applicazioni possano interagire con i dispositivi, indipendentemente dal particolare protocollo di comunicazione utilizzato nativamente dai dispositivi. Per fare ciò, IOT Platform utilizza un identificatore logico univoco per fare riferimento a ciascun dispositivo individualmente.

IOT Platform non fa alcuna ipotesi sul formato dell'identificativo di dispositivo. Fondamentalmente è una stringa definita al momento del provisioning di un dispositivo. Una volta creato, è possibile fare riferimento al dispositivo tramite questo identificativo finché il dispositivo non viene rimosso dal sistema.

Tenant

La IOT Platform supporta il partizionamento logico dei dispositivi in gruppi chiamati tenant. Ogni tenant ha un identificatore univoco e può essere utilizzata per fornire un raggruppamento logico di dispositivi appartenenti allo stesso ambito applicativo o unità organizzativa.

Registrazione del dispositivo

I componenti della IOT Platform utilizzano un'API per accedere alle informazioni di registrazione del dispositivo. L'API definisce l'operazione di asserzione della registrazione per verificare lo stato di registrazione di un dispositivo.

Oltre a ciò, la IOT Platform definisce un'API di gestione del *Device Registry*, che può essere implementata per sfruttare operazioni standardizzate per la gestione di dispositivi e credenziali.

Autenticazione del dispositivo

I dispositivi si connettono agli adattatori di protocollo per pubblicare dati o eventi di telemetria. Le applicazioni downstream che utilizzano questi dati spesso intraprendono azioni particolari in base al contenuto dei messaggi. Tali azioni possono includere semplicemente l'aggiornamento di alcune statistiche, ma può anche innescare attività più serie come lo spegnimento di una centrale elettrica. È quindi importante che le applicazioni verticali possano fare affidamento sul fatto che i messaggi che elaborano siano stati effettivamente prodotti dal dispositivo indicato dall'indirizzo di origine del messaggio.

IOT Platform si affida ad adattatori di protocollo per stabilire l'identità di un dispositivo prima che gli sia consentito pubblicare dati downstream o ricevere comandi. Concettualmente, questo distingue tra due identità

- un'identità associata alle credenziali di autenticazione (denominata identità di autenticazione o auth-id)
- un'identità con cui agire (l'identità del dispositivo o l'ID dispositivo).

Un dispositivo presenta quindi un auth-id come parte delle sue credenziali durante il processo di autenticazione che viene poi risolto in un'identità del dispositivo dall'adattatore di protocollo una volta verificata con successo le credenziali. Per supportare i *Protocol Adapter* e i segreti registrati per il dispositivo e utilizzare queste informazioni per verificare le credenziali.

L'*Auth Server* supporta la registrazione di più set di credenziali per ciascun dispositivo. Un insieme di credenziali è costituito da un ID di autenticazione e da informazioni segrete. Il particolare tipo di segreto determina il tipo di informazioni conservate. Sulla base di questo approccio, un dispositivo può essere autenticato utilizzando diversi tipi di segreti, ad es. una password con hash o un certificato client, a seconda delle funzionalità del dispositivo e/o del *Protocol Adapter*.

Una volta che il *Protocol Adapter* ha risolto l'ID dispositivo per un dispositivo, utilizza questa identità quando fa riferimento al dispositivo in tutte le successive invocazioni API, ad es. quando si inoltrano messaggi di telemetria a valle.

Ogni dispositivo che si connette alla IOT Platform deve essere registrato nell'ambito di un singolo tenant. Le API dell'*Auth Server* richiedono pertanto che un identificatore tenant venga passato alle loro operazioni. Di conseguenza, il primo passaggio che un *Protocol Adapter* deve eseguire quando autentica un dispositivo è determinare il tenant a cui appartiene il dispositivo.

I mezzi utilizzati da un dispositivo per indicare il tenant di appartenenza variano a seconda del tipo di credenziali e del meccanismo di autenticazione utilizzato.

Autenticazione basata su nome utente/password

I *Protocol Adapter* MQTT, HTTP e AMQP supportano l'autenticazione dei dispositivi con un meccanismo basato su nome utente/password. In questo caso, un *Protocol Adapter* verifica che la password presentata dal dispositivo durante la creazione della connessione corrisponda alla password registrata per il dispositivo nel registro del dispositivo.

Durante la creazione della connessione il dispositivo presenta un nome utente e una password al *Protocol Adapter,* quest'ultimo estrarrà quindi l'identificatore del tenant dal nome utente e richiamerà l'operazione di restituzione delle credenziali dell'*Auth Server* per recuperare le credenziali di tipo password con hash registrate per il dispositivo.

Autenticazione basata su chiave pre-shared

Un *Protocol Adapter* supporta l'autenticazione dei dispositivi utilizzando una chiave precondivisa (PSK) come parte di un handshake DTLS. In questo caso, l'adattatore di protocollo verifica che il PSK utilizzato dal dispositivo per generare il segreto pre-master della sessione DTLS sia lo stesso della chiave contenuta nelle credenziali registrate per il dispositivo nel registro del dispositivo.

Durante l'handshake DTLS il dispositivo fornisce una psk_identity al *Protocol Adapter*, quest'ultimo estrarrà quindi l'identificatore del tenant dall'identità PSK e richiamerà l'operazione di restituzione delle credenziali dell'*Auth Server* per recuperare le credenziali di tipo psk registrate per il dispositivo. I psk contenuti nelle credenziali verranno poi utilizzati dal *Protocol Adapter* per generare il pre-master secret della sessione DTLS in fase di negoziazione con il client. Se sia il dispositivo che il *Protocol Adapter* utilizzano la stessa chiave, l'handshake DTLS avrà esito positivo e il dispositivo sarà stato autenticato correttamente

Autenticazione basata sul certificato cliente

I dispositivi possono anche utilizzare un certificato X.509 (client) per autenticarsi sui *Protocol Adapter*. In questo caso, il *Protocol Adapter* tenta di verificare che sia possibile stabilire una catena valida di certificati a partire dal certificato client presentato dal dispositivo fino ad uno dei certificati root attendibili configurati per il tenant del dispositivo.

Durante la creazione della connessione con il dispositivo, il *Protocol Adapter* tenta di determinare il tenant a cui appartiene il dispositivo e di recuperare le informazioni di configurazione del tenant utilizzando le API specifiche. L'adattatore utilizza quindi i trust anchor configurati per il tenant per verificare il certificato client.

Dopo aver verificato il certificato client utilizzando i trust anchor, il *Protocol Adapter* estrrrà il DN dell'oggetto del certificato client e richiamerà l'operazione di restituzione delle credenziali dell'*Auth Server* per recuperare le credenziali di tipo x509-cert registrate per il dispositivo.

Autenticazione basata su token Web JSON

I *Protocol Adapter* MQTT e HTTP supportano l'autenticazione dei dispositivi con un meccanismo basato su JSON Web Token (JWT) firmato. In questo caso, il *Protocol Adapter* tenta di convalidare il token presentato dal dispositivo utilizzando una chiave pubblica registrata.

Durante la creazione della connessione, è previsto che il dispositivo presenti al *Protocol Adapter* l'identificatore del tenant, l'identificativo di autenticazione e un JWT valido e firmato.

Il *Protocol Adapter* richiamerà quindi l'operazione di restituzione delle credenziali dell'*Auth Server* per recuperare le credenziali di tipo rpk (chiave pubblica grezza) registrate per il dispositivo. La chiave contenuta nelle credenziali viene quindi utilizzata dal *Protocol Adapter* per verificare la firma JWS.

Multi-Tenancy

IOT Platform è progettato per strutturare l'insieme di tutti i dati e i flussi di dati gestiti internamente in sottoinsiemi isolati. Ciò include i dati di registrazione e le credenziali dei dispositivi, degli utenti interni utilizzati per l'autenticazione e anche delle applicazioni aziendali che fanno parte di tali sottoinsiemi.

Questo metodo di segregazione è generalmente noto come multi-tenancy, dove un tenant è il termine per tale sottoinsieme. Un tale isolamento è essenziale per consentire a un'architettura distribuita scalabile di gestire sottoinsiemi indipendenti come se ogni sottoinsieme avesse la propria installazione.

Il concetto multi-tenancy si basa sulla gestione dei tenant come entità proprie. Tutte le funzionalità dell'IoT Platform vengono fornite nel contesto di un tenant creato in precedenza, ad eccezione della creazione del tenant stesso.

Di seguito vengono descritte le diverse funzionalità della multi-tenancy implementata sulla IOT Platform fornendone una panoramica completa.

Gestore Tenant

Tramite il gestore dei tenant l'IOT Platform gestisce gli stessi come entità proprie. Il servizio definisce come recuperare i dettagli di un tenant specifico. Ciò offre la possibilità di gestire proprietà arbitrarie a livello di tenant. Per comodità, sono disponibili operazioni CRUD per la gestione dei tenant.

Protocol Adapter

Quando un dispositivo si connette a uno dei *Protocol Adapter* della piattaforma, l'adattatore determina a quale tenant appartiene questo dispositivo Dopo aver determinato il tenant, l'adattatore recupera i dettagli del tenant e solo se il tenant esiste ed è abilitato l'adattatore elabora ulteriormente i dati del dispositivo che si sta connettendo. In caso contrario la connessione verrà chiusa.

Configurazione dei Protocol Adapter

I *Protocol Adapter* recuperano parti della configurazione a livello di tenant utilizzando i dettagli del tenant determinato. Ciò include ad es. se un adattatore specifico è abilitato per questo tenant, consentendo di definire tenant con solo un sottoinsieme delle funzionalità dell'IOT Platform. Si prevede che questa funzionalità sia particolarmente importante per le configurazioni di produzione.

Un dispositivo fisico sarà solitamente rappresentato come un'entità nel *Device Registry*, avente un'identità univoca e appartenente esattamente a un tenant. Tutti i dati inviati da un dispositivo, così come dall'applicazione al dispositivo, vengono quindi trattati come appartenenti al corrispondente tenant.

Il diagramma seguente mostra la relazione tra tenant, dispositivi e relative credenziali:

Figura 96 - Relazione tra tenant, dispositivi e credenziali

Controllo del flusso basato su tenant

Un dettaglio importante nell'architettura dell'IOT Platform è che i dati inviati a valle vengono trasportati tramite i collegamenti AMQP 1.0 con ambito tenant dai *Protocol Adapter* alla rete AMQP 1.0. Ciascun tenant dispone della propria coppia di collegamenti AMQP 1.0 e viene trattato in modo indipendente dagli altri tenant per quanto riguarda il meccanismo di contropressione offerto da AMQP 1.0. Ciò consente a un'applicazione aziendale di limitare la velocità con cui utilizza i messaggi AMQP 1.0 per tenant.

Applicazioni aziendali e tenant

Le applicazioni aziendali si connettono sempre agli endpoint AMQP 1.0 della piattaforma. Per mezzo della configurazione di autenticazione e autorizzazione e del fatto che gli endpoint hanno come ambito un tenant, l'applicazione aziendale agisce solo nel contesto di un tenant.

Separazione dei tenant

I tenant sono separati gli uni dagli altri in tutti i componenti dell'IOT Platform. Ecco un riepilogo di come viene implementato:

- la registrazione dei dispositivi è strettamente riservata a un tenant
- le credenziali dei dispositivi sono strettamente limitate a un tenant
- i Protocol Adapter possono essere abilitati/disabilitati per un tenant
- il flusso di dati a valle è isolato per ogni tenant
- il flusso dati upstream (Comando e Controllo) è isolato per ogni tenant
- le applicazioni aziendali devono autenticarsi sulla rete AMQP 1.0 e tramite tale meccanismo rientrano nell'ambito del relativo tenant

Provisioning del Dispositivo

Di seguito viene descritto come viene effettuato il provisioning dei dispositivi nella IOT Platform, ovvero come viene generata la loro rappresentazione digitale. Per ciascun dispositivo vengono archiviate le informazioni di registrazione che definiscono l'identità del dispositivo. Ogni dispositivo appartiene esattamente a un tenant. Ogni dispositivo deve avere almeno un set di credenziali utilizzate per autenticarsi sulla IOT Platform.

Per poter utilizzare un dispositivo con la IOT Platform, è necessario effettuarne il provisioning. Ciò significa che le informazioni di registrazione e almeno un record delle credenziali devono essere archiviati nel *Device Registry*.

Esistono diversi modi per eseguire il provisioning dei dispositivi.

Provisioning manuale dei dispositivi

È possibile eseguire il provisioning dei dispositivi utilizzando l'API di gestione del *Device Registry* della piattaforma tramite HTTPS. Se il tenant desiderato non esiste ancora, deve essere prima creato.

Provisioning automatico dei dispositivi

Il termine Provisioning automatico denota una funzionalità in cui il *Device Registry* genera automaticamente le credenziali e le informazioni di registrazione per un dispositivo la prima volta che si connette. Il provisioning automatico è supportato dai *Protocol Adapter* per i dispositivi che si autenticano con certificati client o per i dispositivi connessi tramite un gateway.

Client certificate based Auto-Provisioning

Il *Device Registry* genera un identificativo univoco del dispositivo durante il provisioning automatico. Se il modello ID dispositivo per il provisioning automatico è configurato nella voce CA del tenant corrispondente, il *Device Registry* genera l'identificativo del dispositivo in base al modello configurato. Se non configurato, viene generato un identificativo del dispositivo univoco casuale.

Il *Device Registry* crea le credenziali del dispositivo durante il provisioning automatico. Se auth-id-template è configurato nella voce CA del tenant corrispondente, il *Device Registry* genera l'identità di autenticazione delle credenziali in base al modello configurato. Se non configurato, il DN del soggetto del certificato del dispositivo viene utilizzato come identità di autenticazione.

Provisioning automatico basato su gateway

Si riferisce al provisioning automatico dei dispositivi edge connessi tramite gateway. L'abilitazione del provisioning automatico basato sul gateway richiede la configurazione del dispositivo gateway.

Dispositivi di connessione

Una delle caratteristiche più importanti della IOT Platform è quella di astrarre i protocolli di comunicazione specifici utilizzati dai dispositivi. Prima che un dispositivo possa connettersi alla IOT Platform e caricare dati o ricevere comandi dalle applicazioni downstream, è necessario fornirlo al

sistema. Nell'ambito del provisioning del dispositivo, il dispositivo è associato al tenant a cui appartiene e gli viene assegnato un identificativo logico univoco all'interno del tenant.

I dispositivi possono essere generalmente suddivisi in due gruppi: dispositivi che supportano nativamente il protocollo Internet (IP) per la comunicazione e dispositivi che non lo supportano.

I dispositivi che rientrano nel primo gruppo possono connettersi direttamente utilizzando uno qualsiasi dei protocolli basati su IP supportati dai *Protocol Adapter* della IOT Platform. I dispositivi di quest'ultimo gruppo utilizzano spesso protocolli di comunicazione basati su radio o su linea seriale limitati a un'area locale e richiedono un gateway per connettersi a uno dei *Protocol Adapter* tramite IP.

Il diagramma di flusso seguente mostra un dispositivo che supporta il protocollo MQTT e si connette direttamente all' *MQTT Protocol Adapter* e un altro dispositivo che utilizza Bluetooth LE per connettersi localmente a un gateway che poi si connette all' *MQTT Protocol Adapter*.

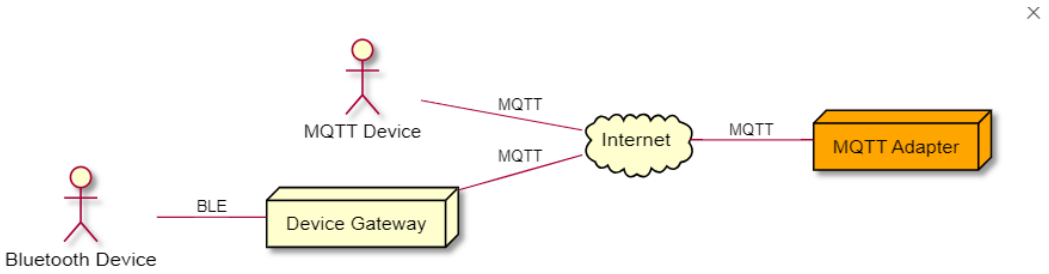


Figura 97 - Connessione tramite MQTT Protocol Adapter

Connessione diretta a un *Protocol Adapter*

Lo scenario più semplice è un dispositivo che si connette a uno dei *Protocol Adapter* della IOT Platform direttamente tramite un'infrastruttura di rete basata su IP. Affinché ciò funzioni, il dispositivo deve utilizzare un protocollo di comunicazione supportato da uno degli adattatori e deve essere in grado di utilizzare gli endpoint delle risorse esposti da quel particolare *Protocol Adapter*.

In questo caso l'identità del dispositivo connesso verrà risolta come parte dell'autenticazione durante la creazione della connessione. Affinché ciò funzioni, è necessario fornire una serie di credenziali per il dispositivo che deve essere appropriata per l'utilizzo con uno degli schemi di autenticazione supportati dall'adattatore.

Connessione tramite un gateway di dispositivi

In alcuni casi, un dispositivo potrebbe non essere in grado di connettersi direttamente a uno degli adattatori di protocollo della IOT Platform. Un esempio è un dispositivo che utilizza un bus seriale o onde radio per la comunicazione locale. Tali dispositivi possono essere collegati a un *Protocol Adapter* tramite un gateway di dispositivo che agisce per conto del/i dispositivo/i quando comunica con la IOT Platform. Un gateway di dispositivi è spesso implementato come un supporto hardware

vicino ai dispositivi, che esegue alcuni software gateway che traducono il traffico nei due versi tra il dispositivo e uno dei *Protocol Adapter* della piattaforma.

Dal punto di vista di un *Protocol Adapter*, il gateway è analogo a qualsiasi altro dispositivo con la propria identità e le proprie credenziali.

Il diagramma seguente illustra come un gateway pubblica i dati per conto di un dispositivo che utilizza Bluetooth per la comunicazione locale con il gateway.

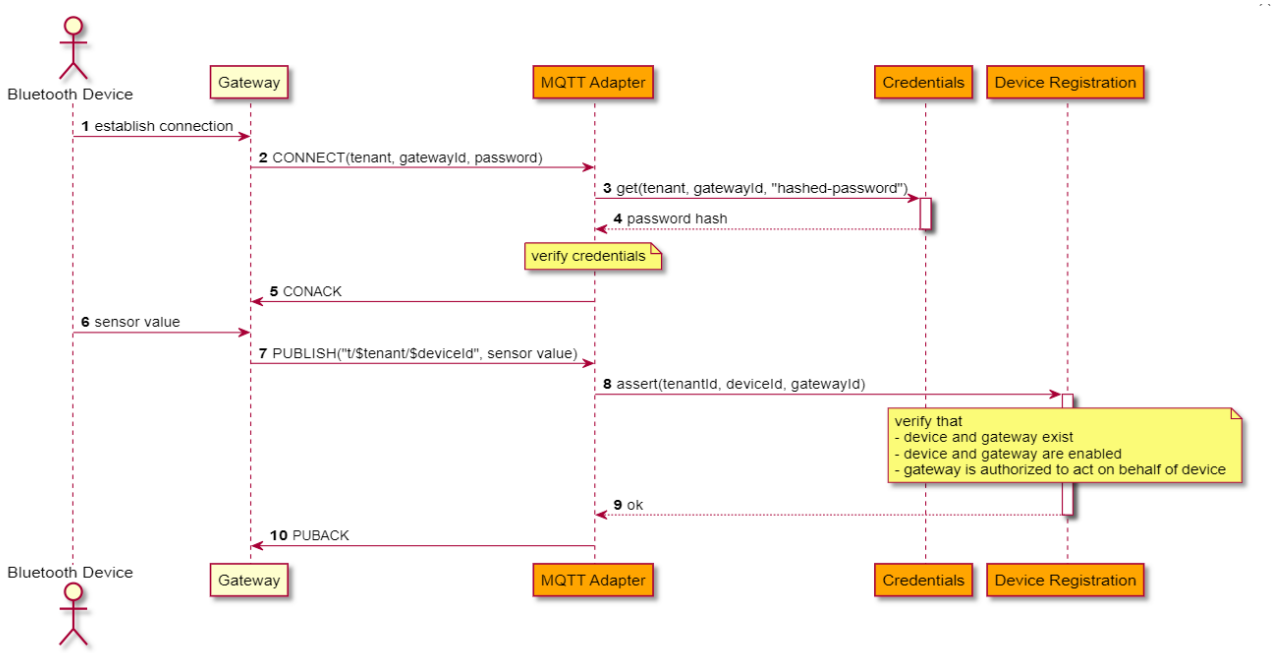


Figura 98 - Connessione tramite un gateway di dispositivi

Il dispositivo stabilisce una connessione Bluetooth con il gateway.

- Il gateway invia un pacchetto MQTT CONNECT all'adattatore MQTT di piattaforma per stabilire una connessione MQTT. Il pacchetto contiene le credenziali del gateway.
- L'adattatore MQTT determina il tenant dal nome utente contenuto nel pacchetto CONNECT e recupera la password con hash registrata per il gateway dal servizio Credenziali.
- Il servizio Credenziali dentro l'Auth Server restituisce la password con hash.
- L'adattatore MQTT controlla la password e accetta la richiesta di connessione.
- Il dispositivo invia alcune letture del sensore tramite Bluetooth al gateway.
- Il gateway inoltra i dati del sensore in un pacchetto MQTT PUBLISH all'adattatore MQTT. Il nome dell'argomento contiene l'identificatore del dispositivo per conto del quale agisce il gateway.
- L'adattatore MQTT richiama l'operazione di asserzione della registrazione del dispositivo del servizio Device Registration per verificare se il gateway è autorizzato ad agire per conto del dispositivo.
- Il servizio di registrazione del dispositivo conferma l'autorizzazione del gateway.
- L'adattatore MQTT accetta i dati del sensore dal gateway e li inoltra a valle.

In questo caso il dispositivo stesso non viene autenticato dall'adattatore MQTT. In questa configurazione la responsabilità di stabilire e verificare l'identità del dispositivo spetta al gateway. Non è quindi necessario fornire alla IOT Platform le credenziali dei dispositivi. La risorsa/dispositivo del servizio Device Registration può essere utilizzata per registrare gateway e dispositivi. I gateway autorizzati ad agire per conto di un dispositivo possono essere impostati tramite le proprietà via e viaGroups del dispositivo. Ciò è utile nei casi in cui un dispositivo può effettuare il roaming tra più gateway. Quando si inviano comandi a un dispositivo, l'IOT Platform deve determinare quale dei gateway autorizzati deve essere utilizzato per inoltrare il messaggio di comando al dispositivo. A questo scopo, i Protocol Adapter dell'IOT Platform tengono traccia dell'ultimo gateway conosciuto che ha agito per conto di ciascun dispositivo tramite il servizio Command Router.

Comando e Controllo

Le applicazioni aziendali possono inviare comandi ai dispositivi tramite servizi dentro il componente *Command Router*.

I comandi possono essere inviati seguendo una richiesta/risposta o uno schema unidirezionale. Per i comandi di richiesta/risposta, è sempre prevista una risposta dal dispositivo.

Affinché i dispositivi possano ricevere comandi, devono prima connettersi a un *Protocol Adapter* della piattaforma e indicare la loro disponibilità a ricevere comandi. Per i dispositivi che comunicano tramite AMQP o MQTT, ciò significa connettersi a un adattatore e sottoscrivere esplicitamente i comandi. Per i dispositivi che inviano messaggi tramite HTTP, ciò significa utilizzare il parametro ttd (tempo fino alla disconnessione) quando si invia un messaggio di evento o di telemetria, indicando così per quanto tempo il dispositivo attenderà un messaggio di comando.

Il *Protocol Adapter* inoltrerà quindi una notifica alle applicazioni aziendali a valle sulla capacità del dispositivo di ricevere comandi e per quanto tempo.

Un'applicazione può inviare un comando a tale dispositivo tramite l'infrastruttura di messaggistica utilizzata (Apache Kafka). L'IOT Platform riceverà il comando tramite il componente *Command Router* e lo inoltrerà all'istanza del *Protocol Adapter* a cui è connesso il dispositivo. L'adattatore invierà quindi il comando al dispositivo. In caso di comando di richiesta/risposta, è previsto che il dispositivo invii un messaggio di risposta al comando.

Quando il dispositivo termina esplicitamente la sottoscrizione del comando, il *Protocol Adapter* invierà una notifica corrispondente alle applicazioni downstream.

Comando e controllo che coinvolgono un gateway

L'IOT Platform dispone di un supporto speciale per l'invio di comandi ai dispositivi collegati all'IOT Platform tramite un dispositivo gateway. I gateway che un dispositivo può utilizzare per connettersi ai *Protocol Adapter* della piattaforma devono essere configurati nel *Device Registry* come parte del processo di provisioning del dispositivo.

Quando inviano comandi, le applicazioni non hanno bisogno di sapere a quale gateway è connesso il dispositivo di destinazione del comando. Un'applicazione invia il comando con l'indirizzo del dispositivo e la piattaforma indirizzerà il comando a un gateway che ha sottoscritto tali comandi e che è configurato per agire per conto del dispositivo di destinazione del comando. Se sono presenti più gateway corrispondenti, viene scelto quello a cui si è connesso per ultimo il dispositivo di destinazione del comando.

Flusso di messaggi utilizzando la rete di messaggistica AMQP

Le sezioni seguenti e i diagrammi di sequenza contenuti forniscono una panoramica di un dispositivo che indica la sua disponibilità a ricevere comandi, di un'applicazione che invia un comando al dispositivo e del dispositivo che restituisce una risposta al comando.

Dispositivo che indica la disponibilità a ricevere comandi

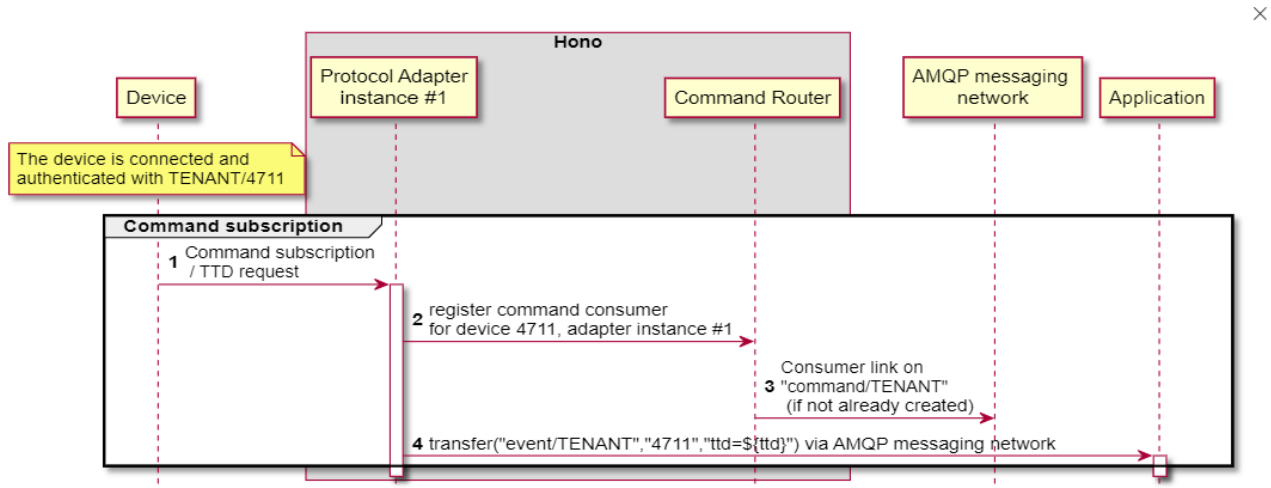


Figura 99 - Dispositivo che indica la disponibilità a ricevere comandi

Quando il dispositivo sottoscrive i comandi (1), l'adattatore di protocollo registra il consumatore del comando con il servizio Router di comando, associando il dispositivo al suo identificatore di istanza dell'adattatore di protocollo (2). Il servizio *Command Router* crea un collegamento ricevitore con ambito tenant del dispositivo (3) se non esiste ancora. Successivamente, la notifica relativa all'abbonamento del dispositivo viene inviata all'Applicazione tramite la rete di messaggistica AMQP (4).

Applicazione aziendale che invia un comando al dispositivo

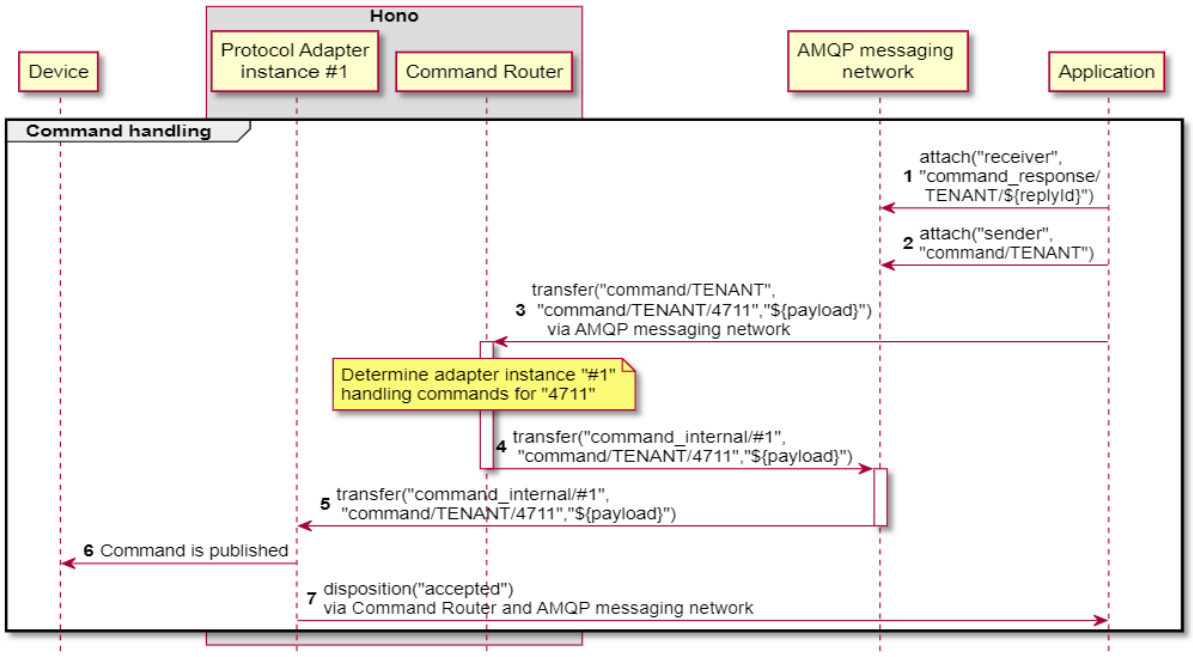


Figura 100 - Applicazione aziendale che invia un comando al dispositivo

Dopo aver ricevuto la notifica, l'Applicazione prepara i collegamenti del mittente e del destinatario della risposta al comando (1,2) e invia il messaggio di comando alla rete di messaggistica AMQP. Il messaggio viene ricevuto dal componente del servizio *Command Router* (3), che determinerà l'istanza del *Protocol Adapter* n. 1 che è in grado di gestire il messaggio di comando. Il comando viene quindi inoltrato alla rete di messaggistica AMQP all'indirizzo dell'istanza dell'adattatore n. 1 (4). L'istanza del *Protocol Adapter* n. 1 riceve il messaggio (5) e lo inoltra al dispositivo (6). Come ultimo passaggio, una disposizione accettata verrà rinviata all'Applicazione (7).

Dispositivo che invia un messaggio di risposta al comando

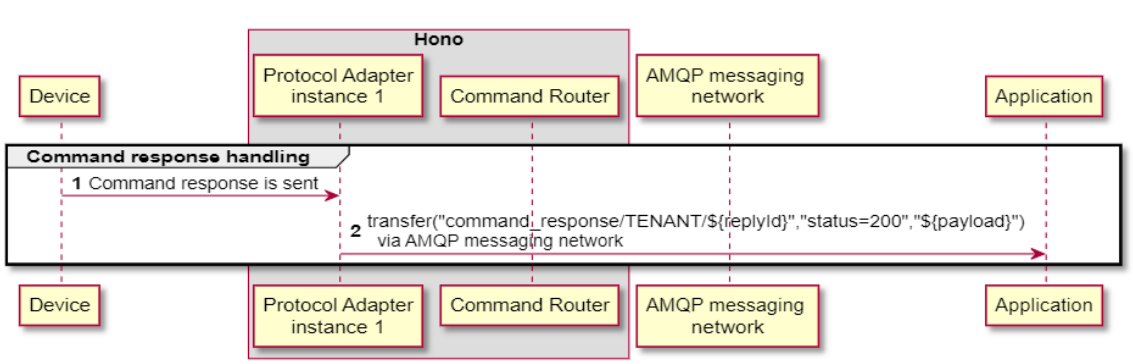


Figura 101 - Dispositivo che invia un messaggio di risposta al comando

Il messaggio di risposta al comando viene inviato all'applicazione dal *Protocol Adapter* tramite la rete di messaggistica AMQP. Il *Command Router* non è coinvolto in questo trasferimento.

Preferred Technology Platform: Eclipse Hono, MongoDb, Custom Microservices, Kafka

1.1.10.2.2 Infrastruttura

L'IoT Platform prevede componenti ospitati da un Container Platform.

I POD previsti per la soluzione del modulo sono allocati sull'infrastruttura CaaS del PSN, in questo modo è gestito, ad esempio, l'autoscaling.

Unitamente alla componente server, lo strato Data Store della soluzione è ospitato dal blocco logico Data System (RDS) del SIM e quindi dalle piattaforme e dai servizi del PSN. Sono coinvolte il PaaS Data Lake, il PaaS DB e le eventuali componenti infrastrutturali quali, ad esempio, il file system.

1.1.11 Document Platform

1.1.11.1 Document Manager

1.1.11.1.1 Architettura Tecnica

La componente di Document Manager si compone dei sottomoduli:

- DM Repository: Il modulo principale che fornisce la funzionalità di base per la gestione dei contenuti. Include funzioni come l'archiviazione, l'indicizzazione e il recupero dei documenti, nonché la gestione degli utenti e dei gruppi.
- DM Share: il modulo che consente agli utenti di condividere e lavorare insieme sui documenti e sui contenuti. Fornisce funzionalità come la condivisione di file, la creazione di siti di collaborazione, la gestione delle attività e la messaggistica integrata.
- DM Governance Services: Il modulo che fornisce funzionalità di gestione dei record e di conformità normativa. Consente di definire politiche di conservazione dei documenti, di creare audit trail e di gestire il ciclo di vita dei record.
- DM Search Services: Il modulo che offre funzionalità di ricerca avanzata per consentire agli utenti di individuare rapidamente e facilmente i contenuti desiderati. Supporta la ricerca full-text, la ricerca faceted e altre funzionalità di ricerca avanzate.
- DM Media Management: Il modulo che consente di gestire e distribuire contenuti multimediali come immagini, video e audio. Fornisce funzionalità di gestione dei metadati, di conversione dei formati multimediali e di streaming dei contenuti.
- DM Intelligence Services: Il modulo utilizza l'intelligenza artificiale e l'apprendimento automatico per fornire funzionalità avanzate come l'etichettatura automatica dei contenuti, l'analisi del sentiment e il suggerimento di contenuti correlati.

La piattaforma tecnologica di riferimento è Alfresco ECM.

I servizi rappresentati nel precedentemente dialogano con uno strato wrapper REST che fornisce una serie di interfacce generiche per interrogare il repository documentale.

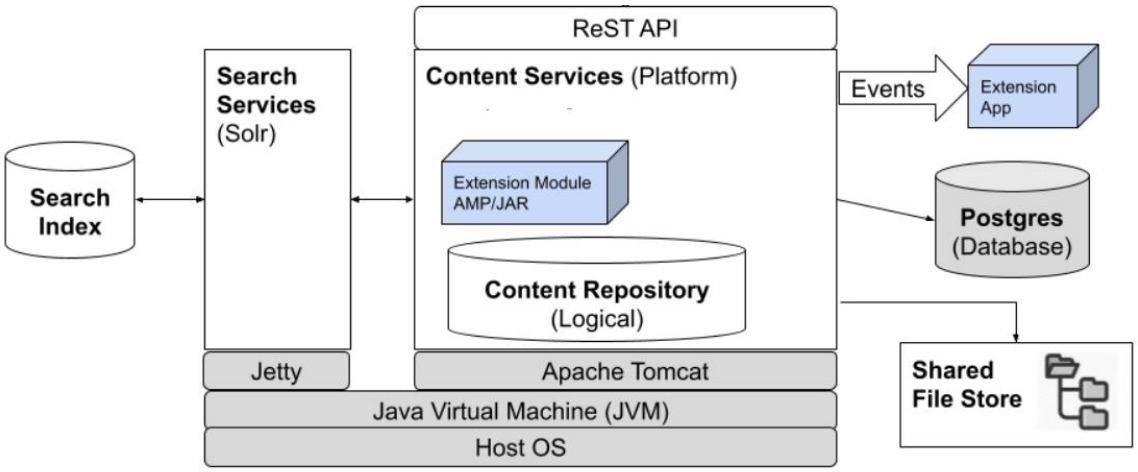


Figura 102 – Document Manager, Architettura tecnica

Le interfacce di Front End del DM Repository e del DM Share fanno riferimento alle API REST, mettendo a disposizione vari punti di estensione dove eventualmente innestare interfacce di UI custom.

Repository

Il modulo principale DM Repository (Content Services) si occupa di gestire il repository dove vengono immagazzinati i contenuti in aggiunta ad una serie di servizi per la loro gestione, quali classificazione, versionamento, nonché dell'interazione con il modulo di ricerca DM Search.

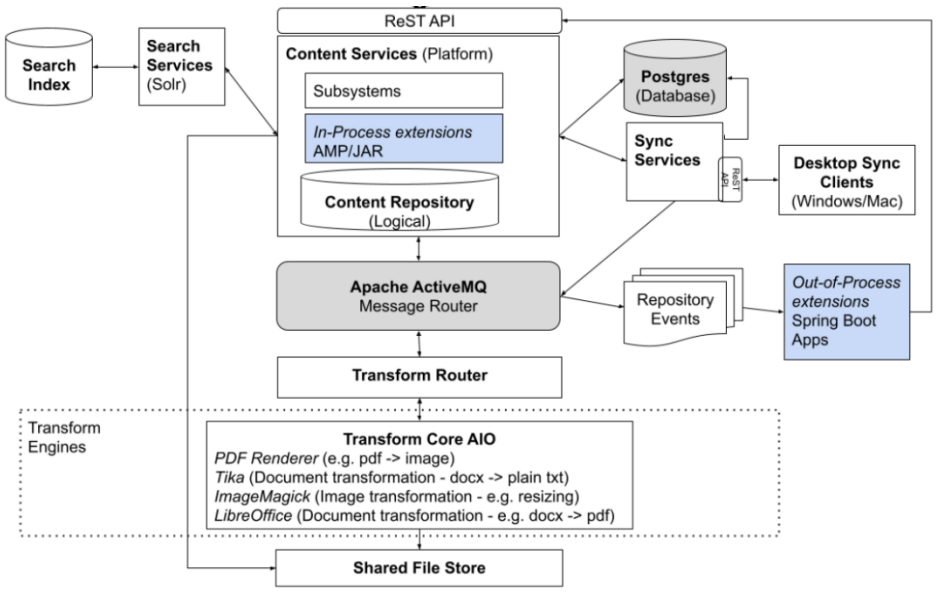


Figura 103 – Document Manager Repository

Questi servizi sono i cosiddetti *subsystems* quali, ad esempio AuthenticationService, ContentStoreService, NodeService etc.

Il sistema Content Services è implementato in Java, il che significa che può girare su tutti quei sistemi dove può girare la Java Standard Edition. I componenti della piattaforma sono stati implementati usando Spring framework, che fornisce la capacità di modularizzare le funzionalità, come ad esempio versionamento, sicurezza e regole. La piattaforma fornisce anche un ambiente di scripting per semplificare l'aggiunta di nuove funzionalità ed interfacce utente.

I contenuti sono criptati e salvati su un file store dedicato mentre i metadati sono inseriti nel database a supporto.

La struttura logica del repository è la seguente:

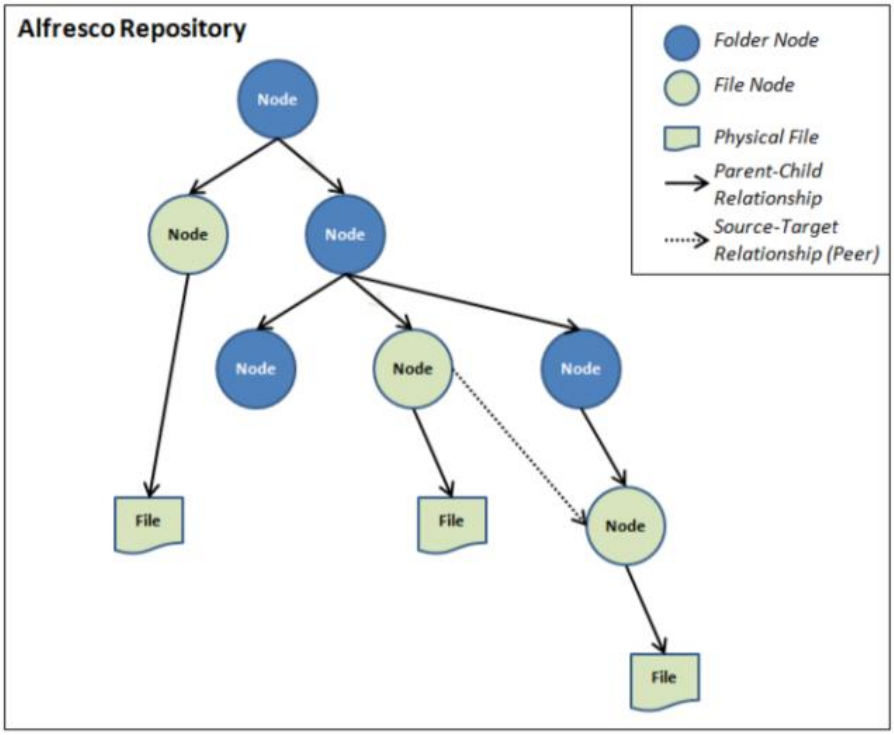


Figura 104 – Document Manager, struttura logica repository

Tutti i file e folder che vengono caricati e creati nel repository vengono trattati come nodi. Alcuni nodi, come i folder, possono contenere altri nodi e quindi esisterà una relazione padre-figlio, mentre altri avranno una relazione peer-to-peer.

Tutti i nodi coesistono in uno *Store* ed ogni store ha alla base un nodo radice da cui discendono tutti gli altri.

Un nodo ha le seguenti caratteristiche:

- **Tipo**: un nodo è di un tipo, come ad esempio Folder, File, Verbale, Rapporto, Dossier, e così via
- **Aspect**: un nodo può avere molti aspects applicati su di esso, quali Versioned, Classified, Searchable, etc.
- **Proprietà**: definiscono i metadati di un contenuto
- **Permessi**: configurazione del controllo degli accessi al nodo

- **Associazioni**: relazioni con altri nodi (peer o child)

La struttura dei *logical store* presenti nella piattaforma è la seguente:

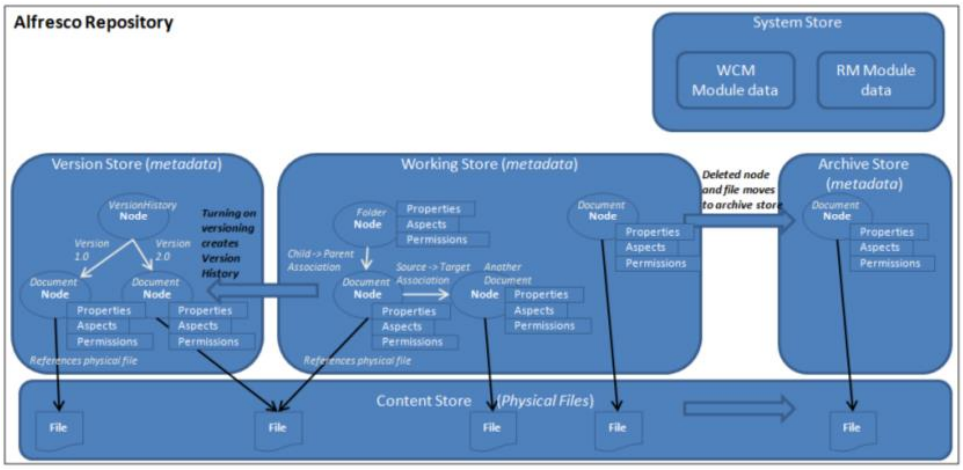


Figura 105 – Document Manager, struttura logical store

- Il **Working Store** contiene tutti i metadati per tutti in nodi attivi nel repository ed è implementato usando un database (RDBMS).
- Il **Content Store** contiene i file fisici caricati nel repository e di default si trova in una cartella del filesystem, ma potrebbe essere configurato per usare altri storage come, per esempio, Amazon S3. È anche possibile definire regole di storage per cui i file con determinate caratteristiche possono essere salvati su differenti storage.
- Ogni volta che un nodo viene cancellato, i metadati del nodo sono spostati nell' **Archive Store,** che usa in database configurato. Il file fisico legato al nodo cancellato viene spostato in un'apposita directory dove risiede indefinitamente.
- Quando l'aspect versionable è applicato ad un nodo, una version history è creata nel **Version Store** (workspace://version2Store). I metadati di un nodo versionato sono salvati nel database e il file rimane nel content store. Il versionamento non è applicabile ai nodi di tipo folder.
- Il **System Store** è usato per salvare informazioni circa i moduli di estensione installati sulla piattaforma.

Ricerca

DM Search ha il compito di indicizzare gli elementi inseriti nel sistema sia per quanto riguarda i metadati che per quanto riguarda il contenuto, ove possibile. Tale processo porta alla creazione di un Search Index che contiene le informazioni relative a tale indicizzazione e che ha lo scopo di velocizzare le ricerche.

In figura viene riportato il motore di ricerca SolR ma è possibile anche integrare altri motori come ad esempio Elastic.

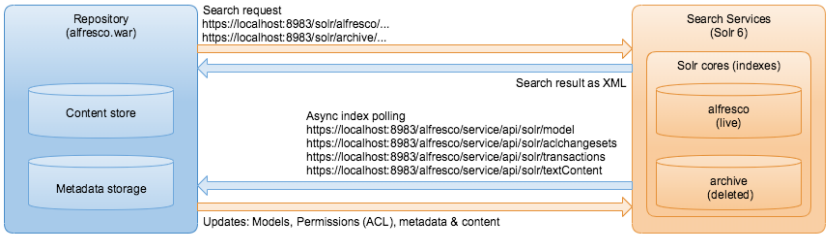


Figura 106 – Document Manager, motore di ricerca SolR

Il repository è anche in grado di gestire gli eventi in modo da scatenare chiamate verso servizi esterni al verificarsi di determinate condizioni configurabili.

Modello dati

Il modello dati è una parte costitutiva fondamentale del repository che fornisce le basi per strutturare i contenuti.

Esso ha le seguenti caratteristiche principali:

- Descrive i dati che vengono inseriti nel repository
- Permette la gestione dei metadati legati al contenuto applicando custom types o aspects ai file e folder
- È univocamente identificato da un suo Namespace, Prefix e Name be definito
- È definito usando un insieme ristretto di elementi base: custom types, aspects, proprietà e constraints

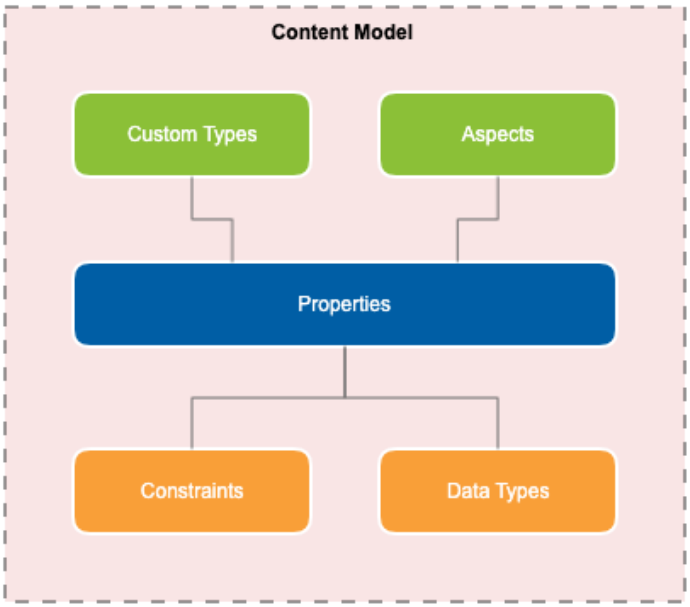


Figura 107 – Document Manager, content model Alfresco

Un **custom type** descrive le proprietà e relazioni che un nodo di un certo tipo può supportare. Content e Folder sono due importanti tipi predefiniti di questo tipo.

Un **aspect** è un insieme di proprietà che possono comprendere sia dati che comportamento. Un file deve essere di un singolo tipo ma potrebbe avere uno o più aspects collegati ad esso. Esempi di aspects sono Classifiable, Versionable, Searchable, e così via.

Una **proprietà** rappresenta i metadati che descrivono il contenuto.

Un **constraint** è usato per validare una proprietà in input. Ad esempio, un campo non deve avere una lunghezza superiore a 20 caratteri.

Il modello può definire classi documentali che discendono da quella principale in modo da specializzare ancora più precisamente la natura di un determinato contenuto.

È molto importante questo aspetto di costruzione del repository documentale nonché l'attenta segmentazione dell'archivio documentale in quanto esse, se correttamente indirizzate, portano benefici in termini di indicizzazione (e quindi maggiore velocità nelle ricerche) e soprattutto maggiore facilità di utilizzo nel caso di navigazione da parte di un utente.

I modelli possono essere creati attraverso un'interfaccia dedicata

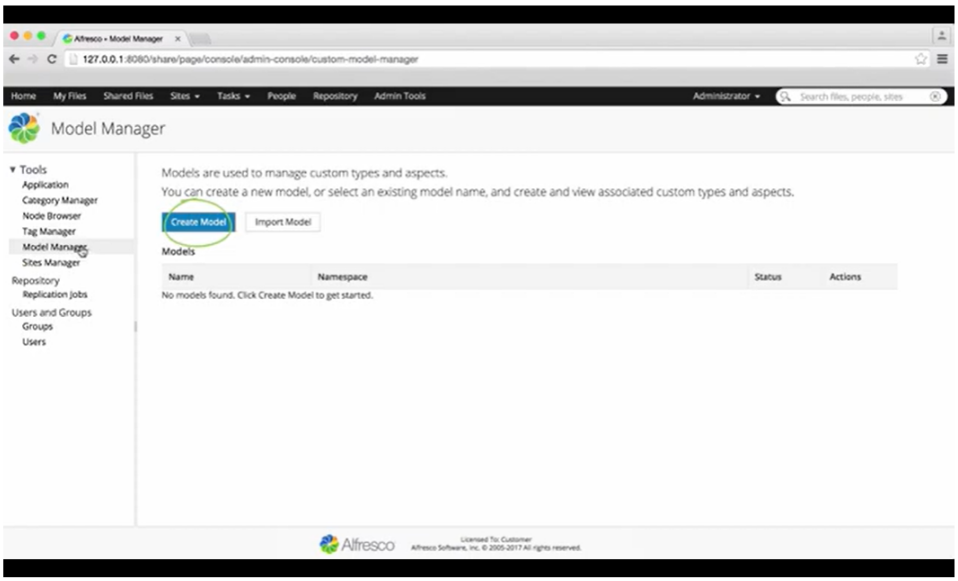


Figura 108 – Document Manager, pagina creazione modelli

Connettori verso altre applicazioni

Il Document Manager, attraverso la piattaforma Alfresco, mette a disposizione tutta una serie di connettori verso gli applicativi più importanti:

- AWS S3
- Azure
- Salesforce

- SAP
- Microsoft 365
- Teams
- Outlook
- Google Docs

Inoltre, c'è la possibilità di utilizzare servizi applicativi di migrazione per importare da altri sistemi documentali come IBM Filenet Content Manager, OpenText, Sharepoint oppure da file system, share, etc.

1.1.11.1.2 Servizi

I servizi messi a disposizione sono di tipo REST e sono solitamente costituiti da un base URL seguito dall'entità di riferimento, l'id e la relazione o operazione da fare sull'oggetto.

Di seguito sono presentati i servizi documentali messi a disposizione dalla piattaforma.

Definizione del modello

Definizione del modello della struttura dell'archivio documentale coerentemente con il dominio applicativo. Definizione di classi documentali, proprietà, constraints, regole, relazioni, etc.

Gestione contenuti

Possibilità di creazione, modifica e cancellazione di contenuti, file, folder, metadati associati e relazioni. Definizione di regole su folder.

Gestione ruoli e permessi

Gestione delle policy di accesso ai contenuti, creazione gruppi, ruoli, utenti e modifica ACL.
I permessi coinvolgono tre entità:

Figura 109 – Document Manager, relazioni ruoli, gruppi, permessi

In aggiunta esistono alcuni ruoli base predefiniti:

- Consumer
- Contributor
- Editor
- Collaborator
- Coordinator

Gestione tag e categorie

Gestione di tag e categorie per singolo contenuto.

I tag non sono strutturati, sono creabili da tutti gli utenti e agevolano le ricerche mentre le categorie sono gerarchiche, create dall'amministratore e aiutano nell'organizzazione dei contenuti.

Versionamento contenuti

Gestione del versioning dei contenuti, con creazione di una version history per il documento nonché il salvataggio dei metadati di ogni versione.

Gestione fascicoli

Integrazione del sistema documentale dentro la gestione fascicolo, creazione, modifica e organizzazione dell'entità fascicolo.

Siti e dashboards

Creazione di siti e dashboards, dedicati alla visualizzazione con widget di documenti suddivisi per aree tematiche, tag o categorie.

Ricerche per proprietà e/o per contenuto

Possibilità di effettuare ricerche di files, siti e utenti in modalità semplice o avanzata dove è possibile specificare criteri di ricerca più complessi.

Import/Export dei contenuti

Sono disponibili funzionalità per importare ed esportare in maniera massiva contenuti e metadati dalla piattaforma.

Regole e azioni

È possibile assegnare regole ad alcuni oggetti documentali, come ad esempio i folder, in modo da scatenare un'azione o una serie di operazioni scatenate, ad esempio, quando viene inserito all'interno un documento.

Gestione di ambienti multilingua

Gestione della localizzazione delle interfacce e delle proprietà associate ad un contenuto documentale.

Audits

La piattaforma consente di tenere traccia delle attività che vengono eseguite su di essa e di configurare il dettaglio con cui queste informazioni vengono tracciate nel sistema.
In figura viene mostrato un diagramma schematico della struttura del sistema di auditing

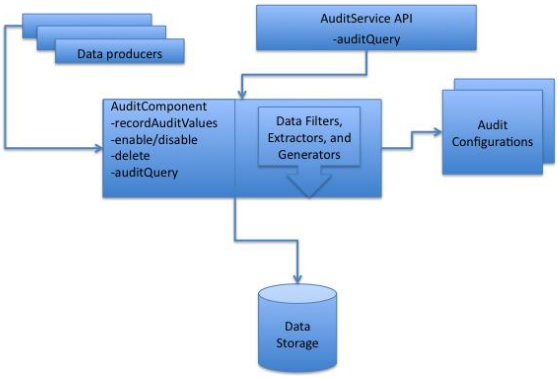


Figura 110 – Document Manager, audit

Trasformazione dei contenuti

Il servizio consente di trasformare un tipo di file (MIME type) in un altro, ovviamente quando possibile. La piattaforma mette a disposizione un elenco completo delle opzioni di trasformabilità da un formato ad un altro.

1.1.11.1.3 Infrastruttura

Il Document Manager prevede componenti ospitati da un Container Platform.

I POD previsti per la soluzione del modulo sono allocati sull'infrastruttura CaaS del PSN, in questo modo è gestito, ad esempio, l'autoscaling.

Le componenti documentali (Alfresco) sono disponibili come immagini Docker e sono anche disponibili Helm charts in modo da poter effettuare il deploy su cluster Kubernetes.

Normalmente il Document Manager usa il suo database interno per i metadati del contenuto mentre i file sono fisicamente salvati su filesystem. Ovviamente è possibile anche avere soluzioni di tipo PaaS con infrastrutture esistenti.

1.1.11.2 Dossier Manager

1.1.11.2.1 Architettura Tecnica

L'architettura del sistema garantisce alta resilienza e scalabilità, è basata su containers ed orchestrators compatibili con lo standard OCI (Open Container Initiative); mette a disposizione un sistema di archiviazione dati e documenti in modo sicuro e regolato da autorizzazioni restituite dal componente IAM sia per l'accesso in sola lettura che in scrittura.

Ogni fascicolo viene creato a partire da un evento scatenato sul modulo Event Manager e/o tramite caricamento a batch per fascicoli documentali.

Ogni modifica effettuata nel tempo ad un fascicolo, viene mantenuta e resa visibile all'operatore tramite una timeline dedicata. Per ogni evento scatenante un workflow, il sistema mantiene le informazioni principali e quelle scatenanti.

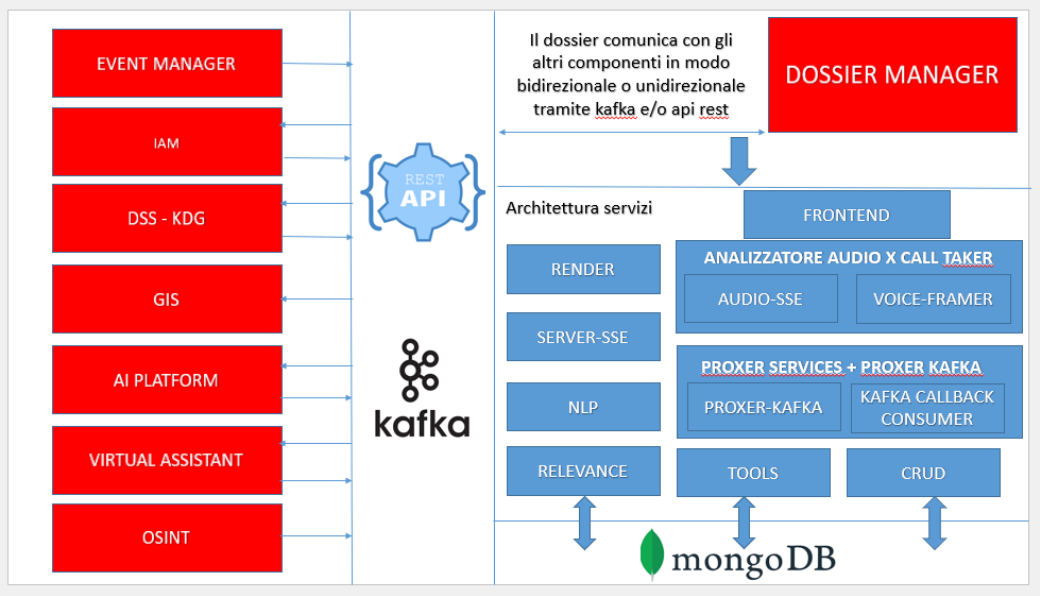


Figura 111 – Dossier Manager, architettura logica

Il Dossier Manager è costituito dai seguenti moduli in modalità container:

- Modulo PROXER.KAFKA: ha il compito di far comunicare i vari moduli del dossier tra di loro e/o di altri componenti come DSS, GIS, ecc.; inoltre permette di scrivere su topic kafka in modo del tutto trasparente sia in modo implicito che esplicito con possibilità di tracciare in code kafka.
- Modulo KAFKA CALLBACK CONSUMER: permette di consumare messaggi da una coda kafka ed eseguire in automatico una chiamata di callback (POST) verso il servizio che si è preregistrato passando il messaggio consumato nel doby della request, oppure permette di aprire un websocket verso il servizio chiamante che si è preregistrato con invio dei messaggi consumati nella socket.
- Modulo NLP: permette di analizzare un testo ed estrarre lemmi e NER per arricchire il fascicolo in esame (il servizio viene richiamato se i servizi di AI Platform non sono disponibili).
- Modulo RELEVANCE: ha il compito di analizzare le notizie estratte, filtrate ed arricchite dal componente OSINT per accoppiarle con uno o più fascicoli se i controlli di affinità vengono superati. Il modulo fa uso del proxy kafka per scrivere su kafka e/o per interpellare il servizio SSE.
- Modulo TOOLS: ha il compito di confermare l'accoppiamento notizia OSINT con il fascicolo interpellando il servizio render, inoltre esegue la scrittura dei log di sistema su una collection dedicata raccolti dal kafka proxy manager, infine permette di eseguire la chiusura forzata di tutti i fascicoli e restituisce le chiavi di configurazione IAM utilizzate poi dall'intero Dossier Manager in base all'utente loggato. Il modulo fa uso del proxy kafka per scrivere su kafka e/o per interpellare il servizio SSE.

- Modulo CRUD: ha il compito di prendere in carico qualsiasi richiesta di scrittura e/o lettura dei fascicoli verso il database mongodb, inoltre permette di ottenere le informazioni storiche di un fascicolo e comunica con i servizi di AI Platform, modulo NLP per arricchire il contesto semantico del dossier. Il modulo fa uso del proxy kafka per scrivere su kafka e/o per interpellare gli altri servizi.
- Modulo RENDER: ha il compito di renderizzare in pdf le notizie estratte dal web dal componente OSINT e/o estrarre risorse allegate alla notizia come video, immagini, audio, ecc.; le risorse scaricate vengono poi catturate dal servizio TOOLS.
- Modulo SERVER-SSE: ha il compito trasmettere ai frontend a cui si registrano gli eventi generati a seguito di una qualche operazione, ad esempio l'aggiornamento di un fascicolo, l'inserimento di una risorsa allegata, un comando elaborato dal VIRTUAL ASSISTANT.
- Modulo AUDIO-SSE: ha il compito di elaborare un file audio interpellando il servizio di speech to text, estraendo il testo, verificando la similarità rispetto ad un elenco di domande (Call Taker) e restituire quindi il comando tradotto nel canale SSE sottostante.
- Modulo VOICE-FRAMER: ha il compito di prendere in ingresso una traccia audio (ad esempio generata da un microfono), dividerla in tanti file audio utilizzando un VAD e spedendo il file generato al servizio AUDIO-SSE per l'analisi e l'elaborazione del testo estratto.
- Modulo FRONTEND: questo modulo presenta 2 modalità di visualizzazione una per il capo sala (nel portale SingleSPA) dove vengono visualizzati tutti i fascicoli presenti nella piattaforma sia aperti che chiusi in modo da poter analizzare lo storico di ogni singolo dossier e le operazioni effettuate nel tempo, la seconda modalità è integrata nel HMI operatore e permette di visualizzare e gestire il singolo dossier relativo all'evento aperto e assegnato , disabilitando alcune sezioni sia lato schema frontend che backend viene utilizzatto il dossier manager pure come archivio di risorse.

Componenti del DM

I successivi paragrafi sono dedicati alla descrizione delle funzionalità rilasciate, le modalità di integrazione da parte degli altri moduli e le attività di verifica per attestarne il corretto funzionamento a valle delle attività di deployment relative al Dossier Manager.

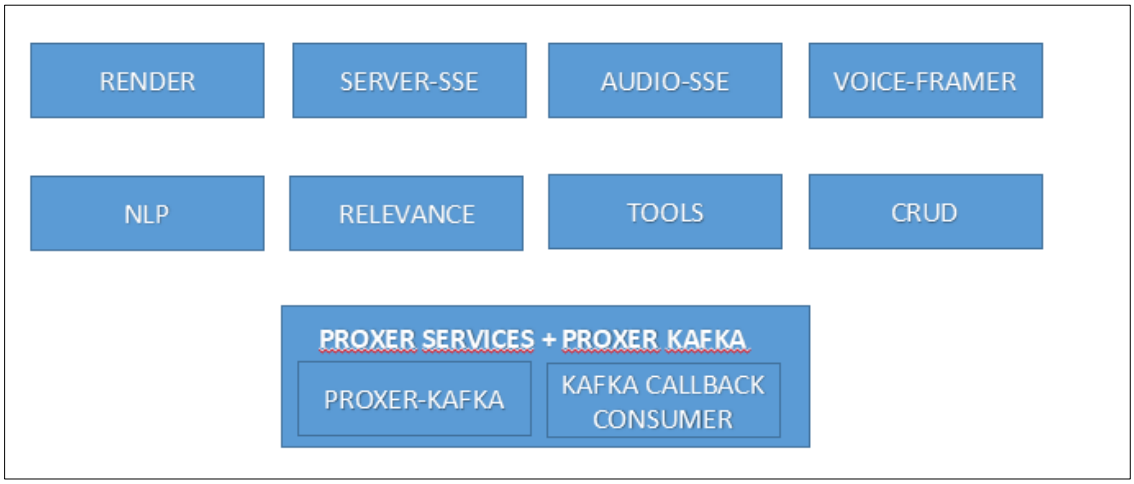


Figura 112 – Moduli del Dossier Manager

Generazione/Modifica Fascicolo

Il componente Event Manager gestisce eventi generati da un operatore tramite l'interfaccia utente HMI. Quando si verifica un'azione, il componente genera un fascicolo utilizzando l'API presente nel modulo dossier. Prima di memorizzare i dati, questi vengono normalizzati, arricchiti con AI Platform o il modulo NLP + Azure (se AI Platform non è disponibile o non risponde) e convalidati attraverso uno schema validator. Successivamente, i dati vengono archiviati in un database MongoDB. Se tutto il processo ha successo, viene generato un evento che viene inviato tramite il canale SSE utilizzando il proxy Kafka Manager. Questo evento viene quindi scritto in un topic dedicato agli eventi SSE all'interno di Kafka. Inoltre, ogni operazione di generazione del fascicolo, viene tracciata e registrata in una collezione dedicata.

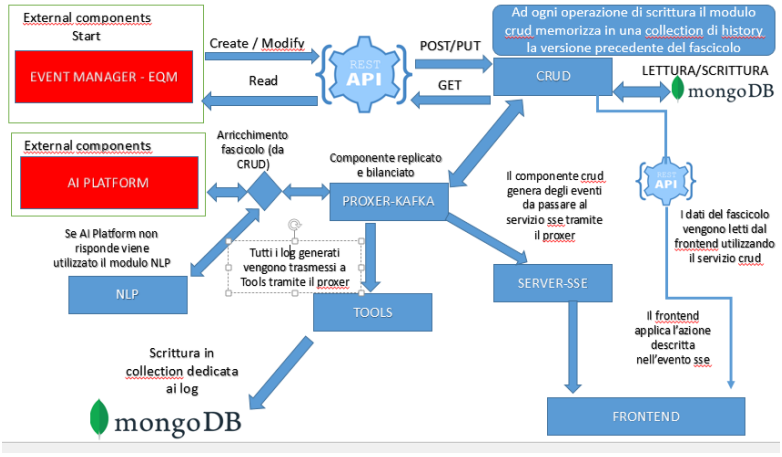


Figura 113 – Flusso creazione /modifica fascicolo

OSINT – Relevance – GIS – Conferma proposta Rilevanza

Il componente OSINT esegue una chiamata di elaborazione al modulo Relevance del DOSSIER MANAGER, che estrae tutti i dossier aperti confrontandoli con la notizia in ingresso. Se la notizia ha informazioni relative ad indirizzi, viene interpellato il componente GIS per estrarre le coordinate geospaziali dell'indirizzo, dopodiché la procedura verifica che il fascicolo in esame abbia delle informazioni geospaziali affini a quella della notizia. Vengono inoltre considerate anche altre informazioni di contesto semantico presente nel fascicolo in esame, se anche queste sono affini alla notizia, il modulo Relevance considera la notizia rilevante per il fascicolo in esame. In assenza di informazioni geospaziali nella notizia, si verificano altre informazioni di contesto semantico tra la notizia ed il fascicolo in esame. Al termine dell'elaborazione il servizio Relevance invia un evento per informare l'utente della presenza di una notizia rilevante per il fascicolo.

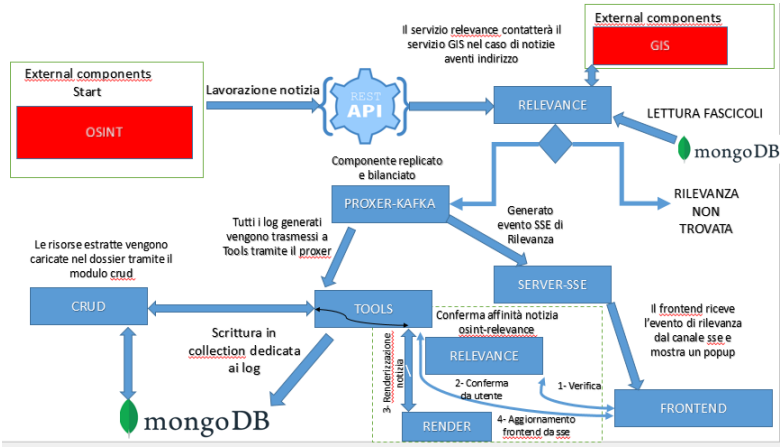


Figura 114 – Flusso creazione/Modifica fascicolo

Comunicazione con DSS/KG – Arricchimento fascicolo

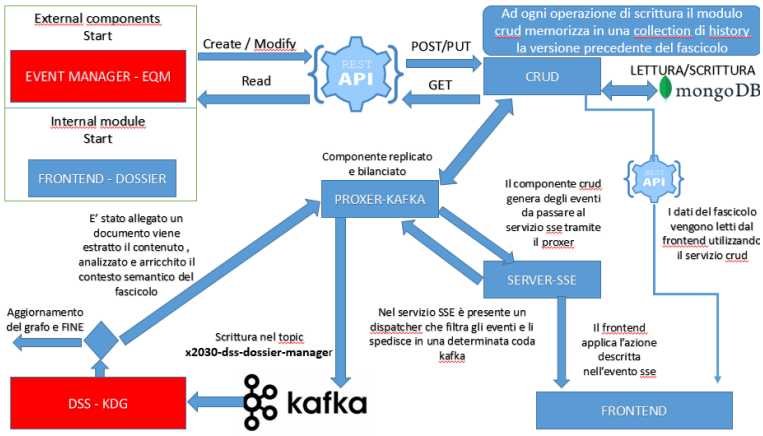


Figura 115 – Flusso interazione DM con DSS / KDG

Ogni volta che avviene una qualche genere di modifica all'interno di un fascicolo, tramite Event Manager o tramite Frontend del fascicolo stesso, viene scatenato un evento. Inoltre vengono impostati gli eventi a seguito di una modifica a livello di metadati o come inserimento di un documento esterno. Il dispatcher, scrive in una coda kafka dedicata con un messaggio indicante il tipo di azione da intraprendere: *ObjectAttachedHandler* oppure *FullDossierHandler*.

Quindi il DSS esegue o un'azione di analisi del dossier e ricalcolo di tutti i nodi di interesse all'interno del grafo, oppure esegue l'analisi dell'allegato cercando di estrarre il contenuto ove possibile; il contenuto estratto viene poi dato in pasto ai servizi di AI Platform per estrarre parole chiavi, NER ecc.

Comunicazione con Virtual Assistant

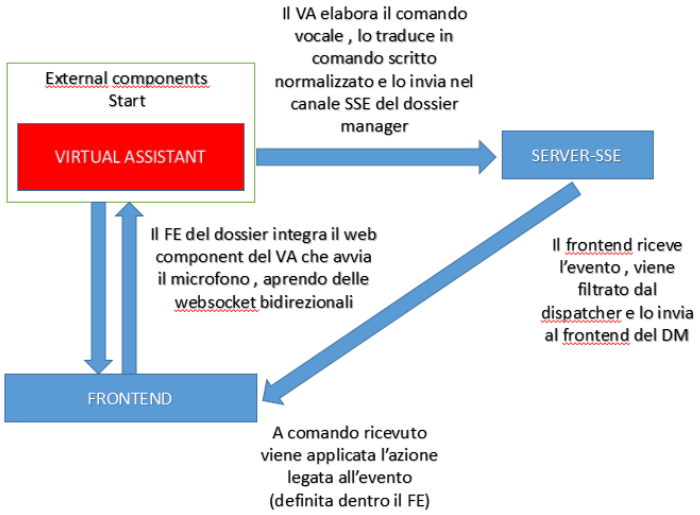


Figura 116 – Flusso interazione DM con Virtual Assistant

Il Virtual Assistant mette a disposizione un web component integrabile in un qualsiasi frontend, apre una connessione verso il proprio backend e delle websocket verso il frontend che lo utilizza; la prima connessione verso il servizio *speech to text* che analizza in tempo reale l'audio e lo invia, in formato binario, allo *speech to text*, mentre la seconda connessione serve a prenormalizzare il comando estratto dallo *speech to text* tramite un servizio di backend del Virtual Assistant.

A fine elaborazione del comando vocale, il testo tradotto, normalizzato e filtrato viene inviato al canale SEE del dossier che a sua volta lo invia ai frontend che si sono preregistrati. Non appena il comando arriva nel frontend del dossier manager, viene applicata un'azione legata al comando, ad esempio la selezione di un dossier o la navigazione all'interno di un fascicolo senza toccare il mouse.

Integrazione con AI-Platform/Azure

Il componente AI Platform/Azure viene interpellato dal Dossier Manager nella fase di estrazione del contesto semantico del fascicolo in esame, all'interno del modulo dossier-crud.

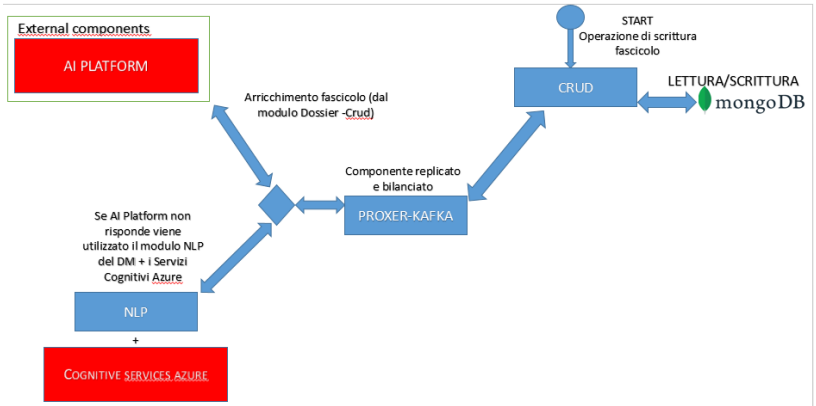


Figura 117 – Flusso integrazione con AI-Platform/Azure

A seguito di un'operazione di scrittura fascicolo all'interno del modulo Dossier-Crud, vengono interpellati i servizi di AI-Platform per estrarre le parole chiavi, le NER ed i lemmi; se i servizi AI Platform non sono raggiungibili, vengono interpellati i servizi Azure + il modulo NLP del Dossier Manager.

Negli scenari descritti sopra il componente esterno PaaS IAM viene interrogato diverse volte per ottenere le autorizzazioni necessarie ad eseguire le diverse azioni. Ogni sezione del dossier (secondo schema dati) ha un livello di riservatezza che è quindi accessibile solamente se l'utente corrente ha i relativi permessi.

Dettaglio singoli servizi

A seguire sono descritti in modo dettagliato i vari moduli presenti nel dossier manager e le interfacce di comunicazione esposte.

In ogni modulo del Dossier Manager che necessita di comunicare con lo strato di persistenza, prevede una procedura di autoconsistenza in modo da creare e popolare le collection MongoDB necessarie al corretto funzionamento.

Proxy-Kafka-Manager

Lo scopo di questo modulo è di assurgere al ruolo di gateway e collettore verso il broker kafka, inoltra le chiamate verso il servizio destinatario restituendo la risposta al chiamante; permette anche di memorizzare ogni singola chiamata all'interno di un topic kafka, sia per le chiamate in ingresso che per le risposte dei servizi chiamati.

Kafka-Consumer

Lo scopo di questo modulo è permettere in modo semplificato, il consumo di messaggi da una coda kafka sfruttando un meccanismo di callback verso il client chiamante, oppure tramite una websocket aperta dal client, in questo modo restituisce in tempo reale i messaggi consumati da una coda kafka.

Relevance

Lo scopo di questo modulo è quello di associare fonti (al momento aperte cioè prelevate dal web ma eventualmente anche classificate) ai dossier aperti.

Il servizio permette di ricevere delle notizie e cerca di associarle a dei dossier aperti. Il criterio di associazione si basa sulla distanza geospaziale tra la notizia e la mainlocation del dossier (la distanza fisica viene calcolata ricercando nelle fonti pervenute delle informazioni riguardo ad indirizzi e luoghi), combinando il risultato con il contesto semantico del dossier, oppure si basa sulla verifica del contesto semantico e sulla presenza di parole chiave o di informazioni importanti presenti nelle note del fascicolo.

Dossier-Crud

Questo modulo è il cuore pulsante del Dossier Manager: espone le funzionalità che permettono di accedere alle risorse presenti all'interno del database. Permettono anche di recuperare o filtrare i dossier archiviati, di caricare o accedere alle risorse multimediali quali video, audio, pdf e file generici che sono associate al documento. Inoltre, il servizio permette di validare i dati inviati dal client, tramite l'utilizzo di uno schema. Infine, tramite l'utilizzo di un codice di autorizzazione, riesce ad accedere e recuperare le informazioni necessarie per permette di profilare i vari dati del fascicolo.

Render

Questo modulo permette di renderizzare delle pagine html eseguendo degli screenshoot e convertendoli in pdf; esegue inoltre il download dal web di altre risorse quali immagini, audio e video.

Dossier-Tools

Questo modulo permette di eseguire il download delle risorse multimediali e di generare dei documenti pdf allegandole al dossier selezionato. Permette di eseguire la scrittura e lettura da MongoDB dei log degli eventi pervenuti attraverso il canale del kafka-proxy manager ed infine prevede delle funzionalità di supporto come, ad esempio, la chiusura di tutti i dossiers.

Dossier-NLP

Il modulo NLP-Service-Manager ha lo scopo di analizzare il testo ricevuto ed estrarre *key phrases*, entità *prebuilt* (nomi, location, organizzazioni, etc.), entità custom definite da gruppi di lemmi costituenti delle tassonomie.

SSE-Server

Il modulo SSE permette di far comunicare il backend del Dossier Manager con il frontend in tempo reale. È inoltre presente una componente di dispatching che filtra gli eventi pervenuti, esegue una redirezione del messaggio verso il componente DSS / KDG per elaborare le risorse allegate e per aggiornare il grafo di contesto.

Nel caso di messaggi destinati alla comunicazione con il DSS, viene utilizzata una procedura di dispatching che genera un json in un formato previsto in cui viene indicata il tipo di operazione.

VoiceFramer

Il modulo Voice Framer ha lo scopo di ricevere flussi audio tramite protocollo UDP, con o senza buffering. Analizza i pacchetti e genera tante tracce audio tramite un algoritmo VAD (Voice activity detection); le varie tracce vengono scritte su un file audio wav con codifica PCM con campionamento a 8khz/16bits. L'audio wav generato viene inviato al servizio AudioSSE per analizzare l'audio e gestire il testo estratto.

AudioSSE

Il modulo Audio SSE ha lo scopo di ricevere file audio di tipo wav, passarlo al servizio di *speech to text*, e mettere a confronto l'audio estratto con delle domande precaricate interrogando il Taxonomy Manager; tramite una procedura di *cosine similary* vengono restituiti tutti i match trovati ai client in listen nel canale SSE. Vengono richiamati dei servizi AI Platform che esegue il confronto per similarità tra la frase estratta dall'audio e quelle presenti nel taxonomy.

Infrastruttura

Come accennato in precedenza il Dossier Manager si basa su diversi servizi rilasciati in modalità Container riutilizzabili in qualsiasi piattaforma che riesce ad eseguire tali Container.

Molti dei Container previsti hanno necessità di comunicare con un'istanza mongodb versione >= 5.0.15 da cui leggere e/o scrivere dei dati, riguardo alla gestione degli allegati è previsto la possibilità di memorizzare e leggere tali file sfruttando PaaS Data Lake.

Oltre Mongodb è richiesta la presenza di un servizio Redis per la memorizzazione e la lettura veloce di determinate strutture dati (ad esempio lo schema dei dossier) e un broker Kafka per spedire e/o leggere messaggi. La componente Relevance ha necessità di poter comunicare con il GIS per geolocalizzare indirizzi da esaminare a runtime.

Ogni pod è replicabile in modo da garantire un'alta affidabilità dei servizi nel caso in cui si verifichino degli errori di sistema e/o applicativi.

Infine, molti servizi richiamano o i servizi di AI Platform o i servizi Cognitivi di Azure on premises per estrazione del contesto semantico di un determinato testo

1.1.12 Orchestration & HTC Platform

1.1.12.1 Orchestration & Provisioning

1.1.12.1.1 Architettura tecnica

L'architettura è suddivisa in moduli come descritto di seguito:

- **Portal**: Una soluzione di front end web e mobile flessibile, consente di comporre la soluzione verso diverse esigenze del cliente. Permette la gestione degli utenti, delle risorse come virtual

machines, storages, networks e consente di effettuare monitoring e provisioning di tutte le risorse di ogni cloud provider configurato.

- **Components**: In questo layer vengono inserite tutte le componenti della piattaforma, sviluppate per poter erogare le funzionalità di gestione della piattaforma cloud management.
- **Abstraction Layer**: L'Abstraction Layer consente di fornire un livello di astrazione delle risorse erogate dai cloud providers implementando direttamente i protocolli di comunicazione specifici e uniformando l'accesso a risorse della stessa tipologia.
- **3rd Parts Components**: In questo layer vengono inserite tutte le componenti della piattaforma gestite con soluzioni di terze parti. In questo layer sono presenti i sistemi di identificazione, la soluzione di Api gateway, il sistema di cache e la gestione dei topics.
- **Cloud providers**: In questo layer vengono inseriti tutte le piattaforme cloud che riesce a supportare, come Microsoft Azure, GCP, AWS, Openshift, Vmware vSphere.

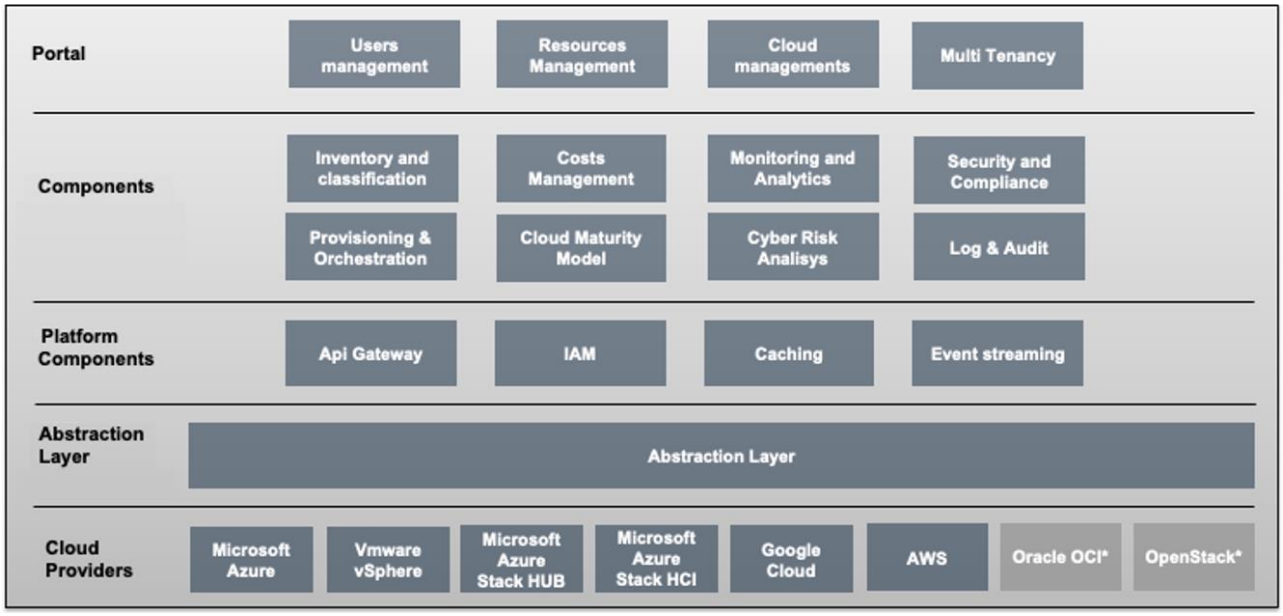


Figura 118 - Architettura Logico/Funzionale

Di seguito si riportano le principali funzionalità di ogni modulo:

Portal:

- Users management: Gestione degli utenti e dei rispettivi ruoli e risorse;
- Resources management: Gestione delle risorse (VM, storage, services);
- Cloud managements: Gestione dei vari cloud providers supportati

Components:

- Inventory e Classification
  - recuperare la lista delle risorse già acquistate dal cliente sui vari CSP;
  - catalogare ciascuna risorsa all'interno dell'inventario locale;
  - creare, modificare e cancellare le regole di classificazione dei servizi;
  - classificare ciascun servizio in funzione delle sue caratteristiche tecniche e di business in funzione delle regole definite dal cliente;

  - gestire attraverso Dashboard la visualizzazione aggregata dei dati di questo modulo;
  - produrre report sui dati visualizzati.
- Monitoring and Analytics: Il modulo permette l'analisi ed il monitoraggio delle performance e dalla capacity dei servizi Cloud, in particolare:
  - Effettua la raccolta di log e delle metriche di performance e di capacity dai servizi rilevati dal modulo Inventory and Classification;
  - Visualizza grafici di performance e di capacity per ciascun assetal fine di analizzare l'andamento storico;
  - Permette la costruzione di Dashboard personalizzate, compresa la costruzione di scenari what-if per la valutazione dell'evoluzione futura dei servizi, nell'ottica della reclamation delle risorse per ottimizzare i servizi.
- Cost management and Workload optimization: Tale modulo si occupa del governo e dell'analisi dei costi dei servizi Cloud, in maniera da permettere ai client di valutare il corretto posizionamento dei servizi sui vari Service Provider, siano essi Public Provider oppure on-premises. In particolare consente di:
  - Effettuare la raccolta di log e delle metriche di spesa relative ai servizi rilevati dal modulo Inventory and Classification;
  - Visualizzare grafici di spesa per ciascun asset al fine di analizzare l'andamento storico;
  - Permettere la costruzione di Dashboard personalizzate, basate anche sulla classificazione cliente.
- Security & compliance: Tale modulo si occupa di verificare la rispondenza dell'implementazione in cloud dei servizi cliente rispetto a diversi standard di sicurezza.
  - Il modulo permette di definire dei profili personalizzabili dall'utente che includano gli standard di security desiderati;
  - Verrà data la possibilità di poter eseguire test con strumenti di compliance di terze parti e visualizzare i risultati.
- Provisiong & Orchestration: Il modulo di Provisioning & Orchestration rappresenta una delle funzionalità core essenziale per la gestione e l'orchestrazione delle risorse e infrastrutture cloud. È progettato per promuovere un ecosistema in cui il provisioning delle risorse sia un processo rigoroso, standardizzato e immune da crescite incontrollate e disorganizzate, un fenomeno noto come "sprawling", che potrebbe compromettere la coerenza e la performance dell'ambiente cloud:
  - Uno degli aspetti salienti di questo modulo è la sua specializzazione nella gestione avanzata dei cluster Kubernetes. Esso è configurato per facilitare un array complesso di funzioni legate a Kubernetes, promuovendo un provisioning, una configurazione e una supervisione dei cluster che sono impeccabili e all'avanguardia. Ogni cluster viene gestito, assicurando una orchestratrazione, una configurazione e una ottimizzazione che sono sintonizzate finemente con le necessità operative in continua evoluzione, consolidando un paradigma di operatività che esalta la resilienza, l'efficienza e la sicurezza.
  - Inoltre, questo modulo è equipaggiato per supportare la gestione e il provisioning dei servizi Platform-as-a-Service (PaaS). Questo ampliamento delle capacità garantisce che la piattaforma possa orchestrare e allocare servizi PaaS, offrendo un servizio comprensivo che risponde efficacemente alle esigenze di un'infrastruttura cloud moderna e dinamica.

- La governance è al centro di questo modulo, implementando un controllo incisivo e una supervisione rigorosa su ogni aspetto dei servizi, sia che si tratti di Kubernetes o di soluzioni PaaS. Questo approccio metodico garantisce una gestione oculata delle configurazioni, un monitoraggio attento e una ottimizzazione continua dei costi, e una applicazione scrupolosa delle politiche e dei protocolli di sicurezza, realizzando un ambiente operativo che è sinonimo di eccellenza, affidabilità e conformità strategica e normativa.

Abstraction Layer:

- L'Abstraction Layer consente di fornire un livello di astrazione delle risorse erogate dai cloud providers implementando direttamente i protocolli di comunicazione specifici e uniformando l'accesso a risorse della stessa tipologia. Consente, ad esempio, di accedere ad una risorsa di tipo macchina virtuale indipendentemente dal cloud provider su cui la macchina è istanziata.

3rd Parts Components:

- Api Gateway:
  - La piattaforma sarà dotata di un API Gateway che si occuperà della gestione delle API e viene collocato tra il client e i servizi di backend;
  - Funge da proxy inverso per accettare le interfacce dell'applicazione, le chiamate API, aggregare i vari servizi necessari per soddisfare le richieste e restituire il risultato appropriato;
  - Gestisce attività comuni come l'autenticazione dell'utente, la limitazione della velocità e le statistiche;
- IAM
  - La piattaforma sarà dotata di un sistema di Identity and access management (IAM) per la gestione delle identità e per il controllo degli accessi basato sui ruoli (RBAC) che consentirà il single sign on (SSO) fra le diverse componenti.
  - Il sistema consentirà di identificare, autenticare e autorizzare centralmente gli utenti che desiderano accedere ad una particolare risorsa e fornirà meccanismi di sicurezza avanzati come la Two Factor authentication basata su codici OTP (One Time Password).
  - La soluzione sarà basata sui protocolli standard come OpenID Connect, OAuth2.0, and SAML e consentirà di federarsi con directory esistenti.
  - Consentirà di definire dei ruoli specifici che potranno essere associati ad utenti o gruppi di utenti per gestire in maniera efficiente l'accesso alle risorse del sistema.
- Caching
  - La piattaforma sarà dotata di un sistema di caching per salvare i dati in memoria, il che abilita l'accesso ai dati a bassa latenza e velocità effettiva elevata
  - Il sistema di cache offre prestazioni incredibilmente veloci con operazioni di lettura e scrittura medie che richiedono meno di un millisecondo e supporto per milioni di operazioni al secondo
- Event streaming
  - La piattaforma sarà dotata di un sistema di event streaming
  - Il sistema di event streaming archivia, elabora e interconnette la piattafroma con i vari cloud providers esterni attraverso flussi di "eventi"

L'architettura è suddivisa nei seguenti componenti funzionali:

- **Portal**: Portale per gli utenti finali che permetta di configurare e usufruire le funzionalità erogate dalla piattaforma sui vari cloud provider supportati
- **Portal: Portale per gli utenti finali che permetta di configurare e usufruire le funzionalità erogate dalla piattaforma sui vari cloud provider supportati API Gateway**: Utilizzato per esporre le API dei vari microservizi della piattaforma e per la gestione sicurezza
- **Microservices Layer**: Sono presenti tutte le funzionalità business della piattaforma
- **Event streming**: La comunicazione tra i microservizi e l'abstraction layer avviene attraverso lo stream di eventi asincroni per la massimizzazione delle performance e la stabilità dell'intera piattaforma
- **Abstraction Layer**: adegua i diversi formati di ricezione delle risorse erogate dai cloud providers gestiti dalla piattaforma in un formato standard verso i microservizi interessati
- **Data sources**: repository dove vengono salvate le configurazioni e i dati ricevuti dai vari cloud providers
- **IAM**: Componente che gestisce l'autenticazione e l'autorizzazione degli utenti della piattaforma
- **Cloud Provider**: Cloud Provider sia pubblici che privati che è possibile gestire tramite la piattaforma
- **Log & Audit**: Modulo per la gestione dei logs di sistema

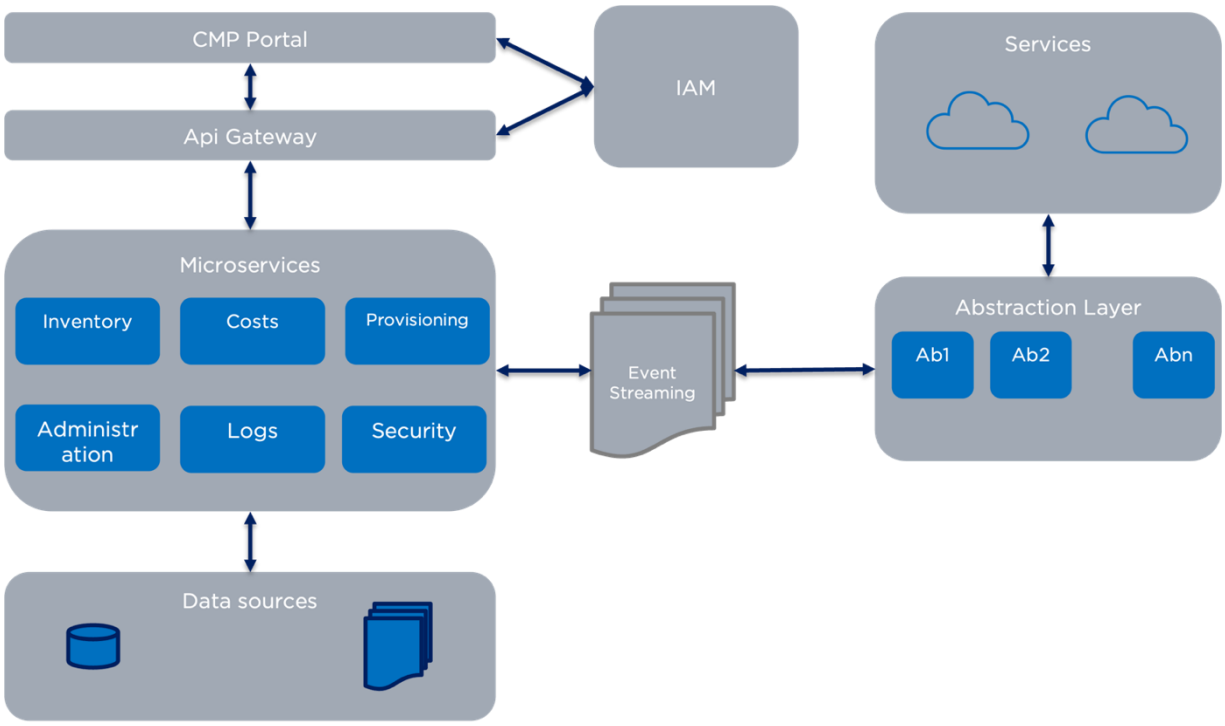


Figura 119 - Orchestration & Provisioning Platform

La capacità di integrazione con i vari Cloud providers è garantita tramite l'Abstraction Layer, che consente di fornire un livello di astrazione delle risorse erogate dai cloud providers implementando direttamente i protocolli di comunicazione specifici e uniformando l'accesso a risorse della stessa tipologia.

Le risorse rilevate dall'Abstraction Layer vengono notificate, tramite eventi, ai microservizi della piattaforma tramite l'utilizzo di topics. Tale notifica avviene in maniera asiconcrona in modo tale da disaccoppiare la rilevazione e il provisioning sui vari cloud providers, dai microservizi che espongono le API verso le interfacce utente.

Tra i microservizi della piattaforma e il front end verrà utilizzato un API gateway sia per garantire la sicurezza delle API esposte dalla piattaforma che per fornire un unico punto di accesso a tutte le API. Questo approccio permette di minimizzare il fronte esposto ai possibili attacchi alla piattaforma, migliorando la sicurezza complessiva del prodotto.

Per quanto riguarda il portale di accesso e la gestione della piattaforma è stata scelta la tecnologia SPA (Single page application) implementata in micro-front end, per poter dare una maggiore flessibilità sia nello sviluppo che nella configurazione dell'interfaccia utente e agevolare l'erogazione futura di nuove funzionalità.

Riguardo i sistemi di storage utilizzati nell'architettura, si è scelto di percorrere un approccio misto utilizzando sia basi dati relazionali che non relazionali. I prodotti integrati, in merito, sono:

- PostgreSQL: DBMS ad oggetti che tra le sue caratteristiche principali offre la data integrity, compatibilità con vari tipi di dato, estensibilità (supporta infatti linguaggi procedurali come Perl, Python etc), sicurezza e il supporto a varie caratteristiche del linguaggio SQL, dalla possibilità di offrire nested query al partizionamento delle tabelle.
- MongoDB: DBMS non relazionale, mette a disposizione la possibilità di effettuare query ad hoc, potendo restituire solo certe parti di un determinato documento, l'aggregazione di dati (tramite MapReduce o Aggregation Framework) e alta affidabilità di mantenimento del dato, grazie alla replica set.
- Elasticsearch: motore di ricerca distribuito, supporta la ricerca real-time, e con l'utilizzo di Lucene, mette a disposizione le sue caratteristiche tramite le API java o JSON.

Lo stack tecnologico si basa sui principali prodotti di riferimento per ogni area di utilizzo. Principalmente si farà uso di Angular e Single SPA per l'implementazione del frontend, di keyclock per la gestione della autenticazione del modulo IAM, Kafka per la data streaming e Kong come sistema di api management.

Di seguito un diagramma riassuntivo:

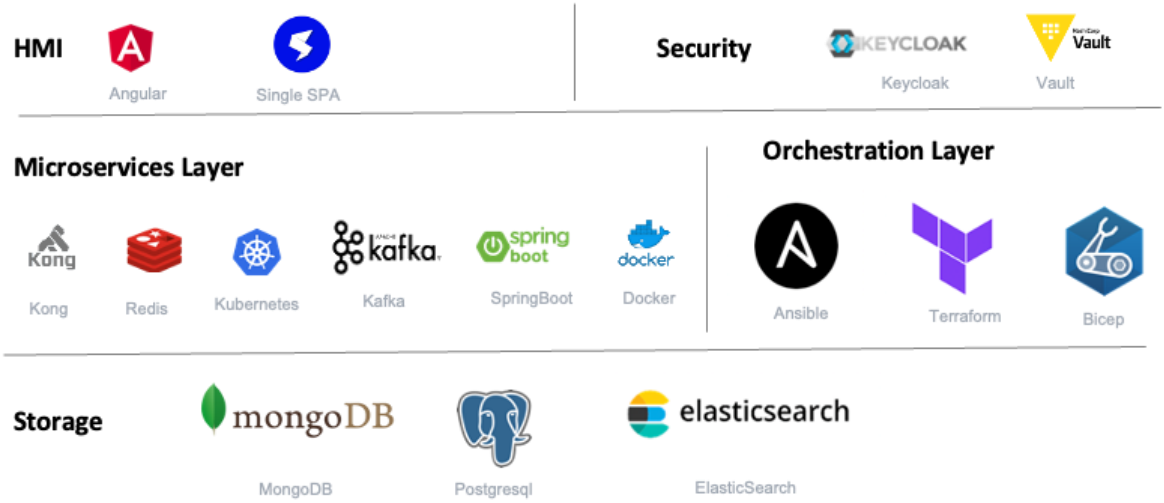


Figura 120 - Stack Tecnologico

1.1.12.2 High Throughput Computing

1.1.12.2.1 Architettura tecnica

Ogni macchina in un pool HTCondor può svolgere diversi ruoli. La maggior parte delle macchine svolge più di un ruolo contemporaneamente. Alcuni ruoli possono essere eseguiti solo da una singola macchina nel pool. La figura seguente descrive i ruoli principali.

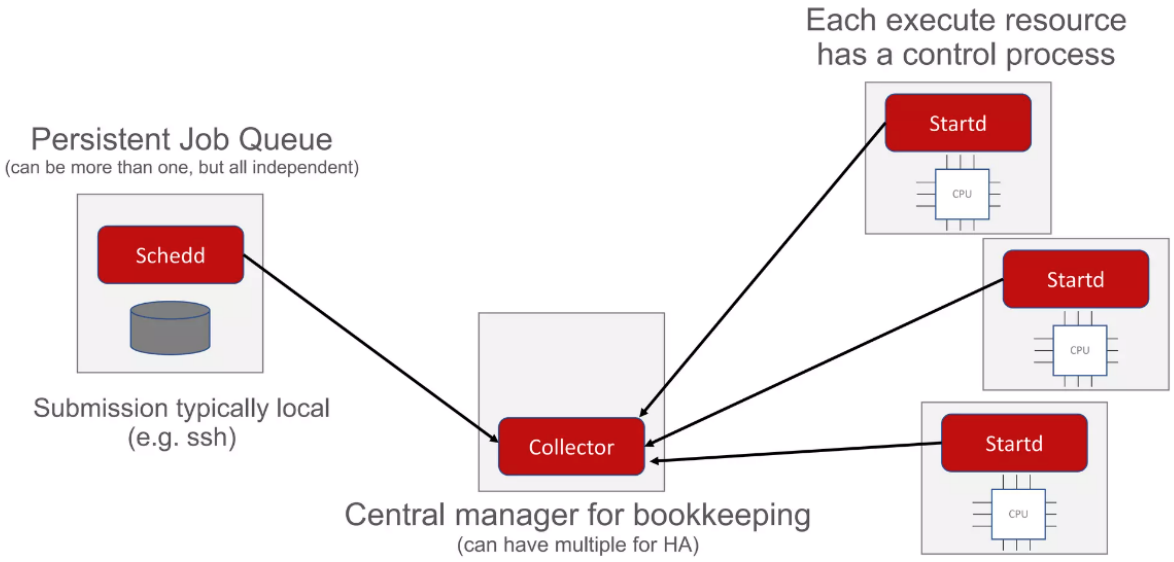


Figura 121 – Architettura HTCondor

Execute Role

Il motivo più comune per aggiungere una macchina a un pool HTCondor è far sì che un'altra macchina esegua i lavori HTCondor; il primo ruolo principale, quindi, è il ruolo di esecuzione. Questo ruolo è responsabile degli aspetti tecnici dell'effettiva esecuzione, monitoraggio e gestione del programma eseguibile del lavoro; del trasferimento dell'input e dell'output del lavoro; e della pubblicizzazione, monitoraggio e gestione delle risorse della macchina di esecuzione. HTCondor può gestire pool contenenti decine di migliaia di macchine di esecuzione, quindi questo è di gran lunga il ruolo più comune.

Il ruolo di esecuzione in sé utilizza pochissime risorse, quindi praticamente qualsiasi macchina può contribuire a un pool. Il ruolo di esecuzione può essere eseguito su una macchina con solo connettività di rete in uscita, ma la capacità di accettare connessioni in ingresso dalla (o dalle) macchine che svolgono il ruolo di invio semplificherà la configurazione e ridurrà l'onere. La macchina di esecuzione non deve consentire l'accesso degli utenti, né condividere ID utente con altre macchine nel pool (anche se ciò può essere molto conveniente, specialmente su Windows).

Submit Role

Questo ruolo è responsabile dell'accettazione, monitoraggio, gestione e pianificazione dei lavori sulle risorse assegnate; del trasferimento dell'input e dell'output dei lavori; e della richiesta e dell'accettazione degli assegnamenti delle risorse. (Una "risorsa" è una frazione riservata di una macchina di esecuzione.) HTCondor consente un numero arbitrario di ruoli di invio in un pool, ma per comodità amministrativa, la maggior parte dei pool ne ha solo uno o un piccolo numero di macchine che svolgono il ruolo di invio.

Una macchina con il ruolo di invio richiede un po' meno di un megabyte di RAM per ogni lavoro in esecuzione, e la sua capacità di trasferire dati da e verso le macchine con il ruolo di esecuzione può diventare un collo di bottiglia delle prestazioni. Di solito raccomandiamo di aggiungere un altro punto di accesso per ogni ventimila lavori in esecuzione contemporaneamente. Un punto di accesso deve avere connettività di rete in uscita, ma una macchina di invio senza connettività di rete in ingresso non può utilizzare le macchine con il ruolo di esecuzione senza connettività di rete in ingresso. Poiché le macchine di esecuzione sono più numerose, di solito i punti di accesso consentono connessioni in ingresso. Sebbene tu possa consentire agli utenti di inviare lavori tramite la rete, ti consigliamo di permettere agli utenti l'accesso SSH al punto di accesso.

Central Manager Role

Nel pool HTCondor, solo una macchina può svolgere questo ruolo (tranne nella modalità ad alta disponibilità, in cui solo una macchina può svolgere questo ruolo alla volta). Un gestore centrale associa le richieste di risorse - generate dal ruolo di invio in base ai suoi lavori - alle risorse descritte dalle macchine di esecuzione. Ci riferiamo all'invio di queste descrizioni (generate automaticamente) al gestore centrale come "pubblicizzazione" perché è il modo principale in cui le macchine di esecuzione ottengono i lavori da eseguire.

Un gestore centrale deve accettare connessioni da ciascuna macchina di esecuzione e da ciascun punto di accesso in un pool. Tuttavia, gli utenti non dovrebbero mai avere bisogno di accedere al gestore centrale. Ogni macchina nel pool aggiorna il gestore centrale ogni pochi minuti e risponde a query di sistema e utente sulla disponibilità delle risorse del pool, quindi una rete veloce è importante. Per pool molto grandi, la memoria potrebbe diventare un fattore limitante.

HTCondor on k8s

Qui di seguito l'architettura ad alto livello di HTCondor su k8s

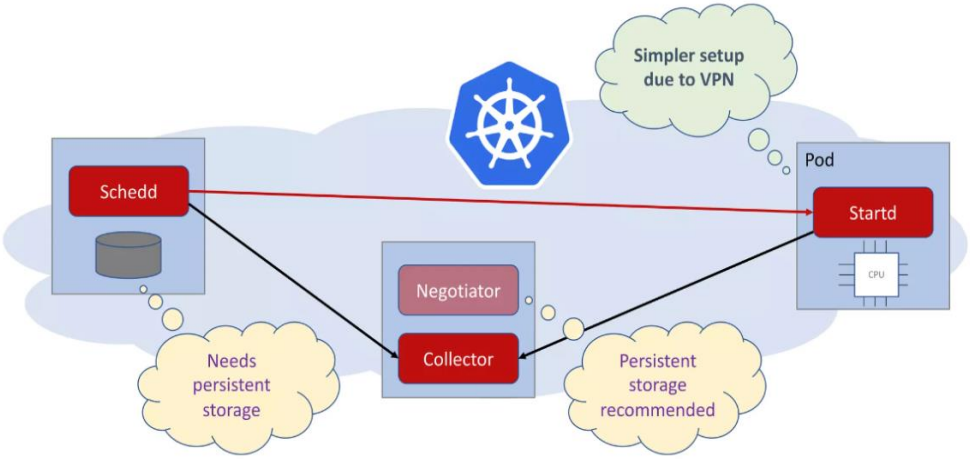


Figura 122 – HTCondor su Kubernetes

L'uso di HTCondor su Kubernetes può offrire diversi vantaggi:

- **Scalabilità**: Kubernetes è noto per la sua capacità di scalare orizzontalmente in modo efficiente. Integrandolo con HTCondor, è possibile gestire il bilanciamento del carico e la distribuzione dei lavori in modo più dinamico e scalabile, aumentando o diminuendo le risorse in base alle esigenze.
- **Isolamento dei lavori**: Kubernetes offre un isolamento efficace dei contenitori. Questo significa che i lavori eseguiti con HTCondor su Kubernetes sono ben separati, riducendo il rischio di interferenze tra lavori diversi e garantendo che ciascun lavoro abbia accesso alle risorse richieste.
- **Orchestrazione dei contenitori**: Kubernetes semplifica l'orchestrazione dei contenitori, consentendo la definizione delle risorse necessarie per i lavori in termini di container e risorse di sistema. Questo semplifica la gestione e la distribuzione dei lavori.
- **Riproducibilità**: Con Kubernetes, è possibile definire in modo dichiarativo l'ambiente in cui verranno eseguiti i lavori, garantendo una maggiore riproducibilità tra diverse esecuzioni di lavori.

- **Gestione delle risorse**: Kubernetes offre un'ottima gestione delle risorse, consentendo un controllo preciso sull'allocazione di CPU, memoria e altri asset. Ciò può essere particolarmente utile per lavori che richiedono risorse specifiche.
- **Monitoraggio e logging**: Kubernetes fornisce strumenti per il monitoraggio dei contenitori e la registrazione delle attività. Questo è utile per tenere traccia delle prestazioni dei lavori e per il debug in caso di problemi.
- **Community e supporto**: Kubernetes è supportato da una vasta community e da numerosi strumenti di gestione. Questo facilita la ricerca di soluzioni per problemi specifici e l'ottenimento di supporto dalla community.
- **Portabilità**: L'utilizzo di Kubernetes può aumentare la portabilità dei lavori. Poiché Kubernetes è ampiamente supportato da molti servizi cloud e ambienti on-premise, è possibile spostare i lavori tra diversi ambienti con relativa facilità.
- **Gestione dei lavori multipli**: Kubernetes semplifica la gestione di un gran numero di lavori, garantendo che siano distribuiti in modo efficiente tra le risorse disponibili.

1.1.12.2.2 Infrastruttura

Il modulo HTC prevede un deploy ospitato su Container Platform; questo consente di sfruttare tutte le caratteristiche di scalabilità, isolamento dei lavori evidenziate in precedenza. La piattaforma CaaS del PSN è completamente gestita consentendo così vantaggi in termini di autoscaling, sicurezza, resilienza ecc.

Unitamente alla componente server, lo strato Data Store della soluzione è ospitato dal blocco logico Data System (RDS) del SIM e quindi dalle piattaforme e dai servizi del PSN. Possono di volta in volta essere coinvolte il PaaS Data Lake, il PaaS DB e le eventuali componenti infrastrutturali quali, ad esempio, il file system.

1.2 Infrastruttura

L'infrastruttura a supporto del SIM prevede la realizzazione di 3 ambienti:

- Sviluppo, basato su Secure Public Cloud Azure
- Collaudo, basato su soluzioni industry standard
- Produzione, basato su soluzioni industry standard

Per ognuno dei 3 ambienti saranno presenti risorse IaaS, PaaS, DBaaS, CaaS.

Di seguito la descrizione delle risorse divise per tipologia.

1.2.1 IaaS

1.2.1.1 Secure Public Cloud

Il Secure Public Cloud è un servizio PSN Core che si basa su Region pubbliche di Azure a cui vengono aggiunti tutti gli elementi di sicurezza (Chiavi esterne, backup, template, servizi professionali).

L'architettura del servizio "Secure Public Cloud" è basata su due componenti principali:

- **Public Cloud**: La componente Hyperscale Public Cloud, erogata da una Region collocata sul territorio nazionale, ai cui servizi vengono applicate configurazioni, policy e controlli di sicurezza, al fine di garantire ai clienti ambienti di elaborazione segregati aventi una sicurezza di base adeguata agli scopi del PSN;
- **Security & Governance**: Una componente, erogata dai Data Center del PSN distribuiti sul territorio Nazionale, nella quale verranno configurati servizi atti a garantire l'adeguato livello di sicurezza dei servizi erogati sul Public Cloud (Gestione Chiavi e Backup).

Tale scenario prevede la presenza dei seguenti attori:

- Fornitore dei servizi di Public Cloud (CSP):
  - fornisce la piattaforma su cui è costruita la componente Hyperscale Public Cloud dell'architettura
- PSN:
  - si occupa di progettare, erogare, gestire e controllare i servizi cloud ed in modo particolare la componente di sicurezza e governo di base adeguati agli scopi del PSN;
  - fornisce servizi di sicurezza opzionali a "valore aggiunto" integrati ai servizi base tramite servizi professionali per la securizzazione.

Il Secure Public Cloud è un servizio core del PSN che garantisce alti standard di sicurezza:

GESTIONE DELLE CHIAVI. Relativamente alla gestione delle chiavi la soluzione comprende:

- Impiego di terze parti (e.g., Thales CipherTrust) con grande livello di autonomia nella gestione delle chiavi crittografiche per soluzioni in cloud con il modello Bring Your Own Key (BYOK).
- Soluzione di key management replicata nei due datacenter HA e territorialmente nelle due Region.
- Controllo on-premise per ciascuna fase del ciclo vita delle chiavi, consentendo di eseguire in autonomia:
  - generazione delle chiavi ON-PREMISE tramite l'utilizzo di dispositivi crittografici certificati;
  - esecuzione dei backup delle chiavi;
  - installazione diretta delle chiavi sui Key Vault in cloud;
  - monitoraggio degli accessi alle chiavi;
  - rotazione manuale o periodica delle chiavi;
  - revoca delle chiavi.
- On-Prem HSM certificato FIPS 140-2 L3 con partizioni multiple per la corretta gestione del materiale crittografico (chiavi simmetriche e asimmetriche, generazione entropia…).
- CipherTrust Manager per la gestione del ciclo di vita delle chiavi on-premise e in Cloud.
- CipherTrust Cloud Key Manager come orchestratore dei processi di gestione delle chiavi in Cloud. Generazione delle chiavi on-premise per importazione sicura sul cloud provider per tutto il ciclo di vita.

GOVERNANCE MODEL. Per ogni cliente viene creato un ambiente standard segregato e auto-consistente in cui, tramite servizi di delega dei privilegi (ad esempio Azure Lighthouse e Privileged

Identity Management) è possibile proiettare i servizi di monitoraggio e sicurezza dello specifico ambiente cliente verso l'ambiente del gestore del PSN che quindi avrà:

- Visibilità di tutti gli ambienti
- Capacità di intervento automatizzato su larga scala
- Possibilità di enforcement delle policy definite

I Privilegi di amministrazione sono disabilitati per default e vengono attribuiti agli operatori a valle di un processo di autorizzazione: questo meccanismo garantisce il mutuo controllo da parte del cliente e del provider con intrinseco innalzamento del livello di sicurezza.
Le caratteristiche di questo modello di gestione forniscono:

- Gestione uniforme e standardizzata dei tenant cliente;
- Creazione, distribuzione e aggiornamento, tramite sistemi di automazione, di set di regole di sicurezza predefinite in linea con best practices internazionali;
- Creazione, distribuzione e aggiornamento, tramite sistemi di automazione, dei ruoli standard per ogni funzione (Ruoli PSN, Ruoli PA, Ruoli terze parti);
- Disponibilità di template securizzati ed integrati a strumenti di sicurezza;
- Gestione unificata dell'identità;
- Gestione degli eventi di sicurezza;

CONFIDENTIAL COMPUTING. L'obiettivo del PSN è rafforzare il livello di confidenzialità e sicurezza del dato in uso tramite i seguenti metodi:

- Ridurre al minimo le cosiddette Trusted Compute Bases (TCB) sui piani hardware, software e operations.
- Usare tecniche di enforcement basate su componenti tecnologiche piuttosto che su processi organizzativi.
- Fornire trasparenza sulle garanzie, i rischi residui e le mitigazioni che si possono implementare.
- I modelli di attacco contro le applicazioni cloud si basano su tecniche diverse per prendere di mira codice o dati in uso, ad esempio:
    - breakout di hypervisor e container;
    - compromissione del firmware ed altre minacce interne, ognuna delle quali si basa su tecniche diverse per prendere di mira codice o dati in uso.

Confidential Computing (per VM, K8S, HSM) è la protezione dei dati in uso utilizzando ambienti di esecuzione attendibili basati su hardware

SOLUZIONI HUB & SPOKE. Per quanto riguarda l'ambiente Secure Public Cloud è previsto l'uso di un modello Hub & Spoke per consentire al PSN il controllo del traffico e la gestione delle DMZ per l'ambiente cloud.

Le Amministrazioni potranno creare reti virtuali spoke nei segmenti, dove saranno attive Policy che forzeranno la connessione con Virtual Network Hub e impediranno la creazione di tipologie di risorse controllate centralmente, come, ad esempio, gli indirizzi IP pubblici.

BACK UP. Per esercitare la sovranità del dato, il Secure Public Cloud prevede l'esistenza e la fruibilità di una copia di tale dato in maniera indipendente dai servizi del CSP tramite ulteriore livello di archiviazione.

Tale servizio sarà fornito attraverso l'integrazione delle risorse in Public Cloud con il Backup del PSN in modo che lo Storage su cui risiede il dato protetto sia gestito dal personale PSN.

L'integrazione prevede l'uso di tecniche di backup snapshot o stream-based e la cifratura dei dati "at rest" e in transito per garantire la protezione e il ripristino delle macchine virtuali a cui è rivolto il servizio, anche di quelle che implementano meccanismi di encryption del disco di sistema e dei dischi dati.

Di seguito la descrizione delle principali risorse IaaS presenti

- **Virtual Machine**

Le macchine virtuali di Azure sono istanze che forniscono risorse di calcolo on demand e scalabili.

Azure Virtual Machine è un servizio cloud che consente di creare e gestire macchine virtuali (VM) in un ambiente scalabile e sicuro. Con Azure VM, è possibile scegliere tra una vasta gamma di dimensioni, sistemi operativi, immagini e configurazioni per soddisfare le esigenze del proprio progetto. Azure VM offre anche funzionalità avanzate come il bilanciamento del carico, la replica, il backup, il monitoraggio e l'automazione. Azure VM si integra facilmente con altri servizi cloud di Azure, come il database, lo storage, la rete e l'identità. Azure VM è una soluzione ideale per eseguire applicazioni ad alte prestazioni, workload complessi.

- **Network Security Group**

È possibile usare un NSG di Azure per filtrare il traffico di rete tra le risorse di Azure in una rete virtuale. Un NSG contiene regole di sicurezza che consentono o negano il traffico in ingresso o il traffico in uscita da diversi tipi di risorse di Azure, o da rete pubblica.

Per ogni regola è possibile specificare e gestire l'accesso alla base dell'origine, la destinazione, la porta, il protocollo, singolo IP o range di IP, o tipologia di risorsa Azure.

Le regole di sicurezza vengono valutate e applicate in base alle informazioni su cinque tuple (origine, porta di origine, destinazione, porta di destinazione e protocollo)

- **Virtual Network**

La rete virtuale di Azure è il blocco fondamentale per la rete privata in Azure. Una rete virtuale consente a molti tipi di risorse di Azure, ad esempio Azure Virtual Machine (VM), di comunicare in modo sicuro tra loro, internet e reti locali.

Una rete virtuale è simile a una rete tradizionale che si potrebbe operare nel proprio data center. Un Rete virtuale di Azure offre vantaggi aggiuntivi dell'infrastruttura di Azure, ad esempio scalabilità, disponibilità e isolamento

Per impostazione predefinita, tutte le risorse in una singola rete virtuale possono comunicare fra di loro. Per poter communicare con delle risorse presenti in una rete virtuale diversa, è necessario

attivare il servizio di Peering fra le due reti virtuali. Le reti virtuali connesse possono essere in aree di Azure uguali o diversi

- **Azure Bastion**

Azure Bastion è un servizio Azure che consente la connettività RDP/SSH sicura e trasparente alle macchine virtuali direttamente tramite TLS, dal portale di Azure. Azure Bastion viene configurato su una SubNet dedicata all'interno della rete virtuale nella quale sono presenti le VM.

Azure Bastion apre la connessione RDP/SSH alla macchina virtuale di Azure usando un indirizzo IP privato nella macchina virtuale, per cui non è necessario un indirizzo IP pubblico o un agent.

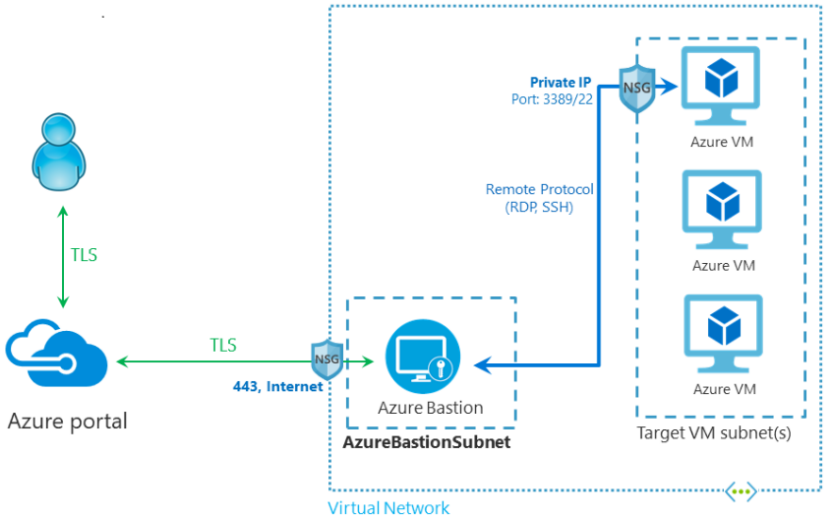


Figura 123 - Funzionalità Azure Bastion

- **Storage Account**

Uno Storage Account di Azure contiene tutti gli oggetti dati di Archiviazione di Azure, inclusi BLOB, condivisioni file, code, tabelle e dischi.

Lo Storage Account di Azure usa la crittografia lato servizio (SSE) per crittografare automaticamente i dati quando viene mantenuta nel cloud. La crittografia protegge i dati e consente di soddisfare gli impegni di sicurezza e conformità dell'organizzazione.

Un Private Endpoint consente alle risorse di Azure, ad esempio le macchine virtuali, di comunicare privatamente e in modo sicuro con lo Storage Account.

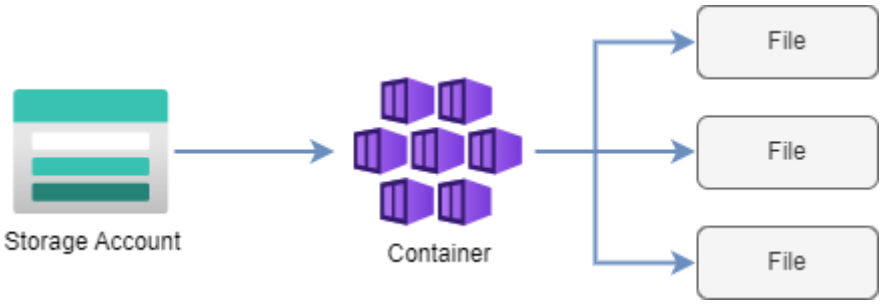


Figura 124 - Funzionalità Azure Storage Account

- **Azure Firewall**

Firewall di Azure è un servizio di sicurezza firewall di rete nativo del cloud Azure, che offre una protezione ottimale dalle minacce per i Workload eseguiti in Azure.

Azure Firewall fornisce filtri L3-L7 e feed di intelligence sulle minacce direttamente da Microsoft Cybersecurity. I filtri basati sulle minacce possono avvisare e negare il traffico da/verso indirizzi IP dannosi noti e/o domini aggiornati in tempo reale per proteggersi da attacchi nuovi ed emergenti.

Consente l'ispezione sia del traffico orizzontale destra-sinistra, sia del traffico verticale alto-basso.

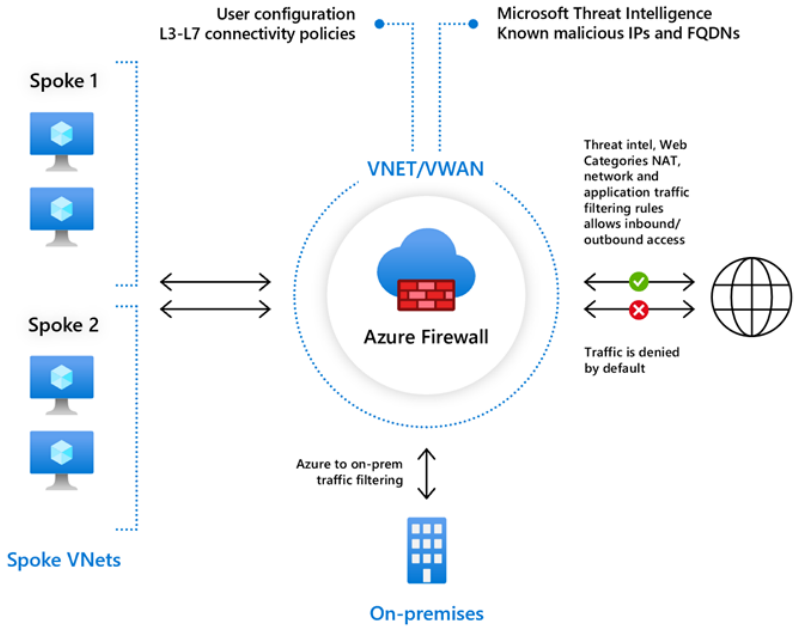


Figura 12516 - Funzionalità Azure Firewall

La disponibilità elevata è integrata essendo veicolato su Azure come servizio PaaS.

Azure Firewall è conforme agli standard internazionali, PCI (Payment Card Industry), SOC (Service Organization Controls) e ISO (International Organization for Standardization).

Tutti gli eventi sono integrati con Azure Monitor, consentendo di archiviare i log in uno Storage Account, trasmettere eventi all'Event Hub, o inviarli ai log di Azure Monitor.

- **Azure Monitor**

Azure Monitor è una soluzione di monitoraggio completa per la raccolta, l'analisi e la risposta ai dati di telemetria dagli ambienti cloud. Azure Monitor raccoglie e aggrega i dati da ogni livello e componente del sistema in una piattaforma dati comune. Poiché questi dati vengono archiviati insieme, possono essere correlati e analizzati usando un set comune di strumenti. I dati possono quindi essere usati per l'analisi e per comprendere il funzionamento delle applicazioni e rispondere automaticamente agli eventi di sistema.

Azure Monitor può essere usato per monitorare questi tipi di risorse in Azure:

- Applicazioni
- Macchine virtuali
- Sistemi operativi guest
- Container
- Database
- Eventi di sicurezza in combinazione con Azure Sentinel
- Eventi di rete in combinazione con Network Watcher

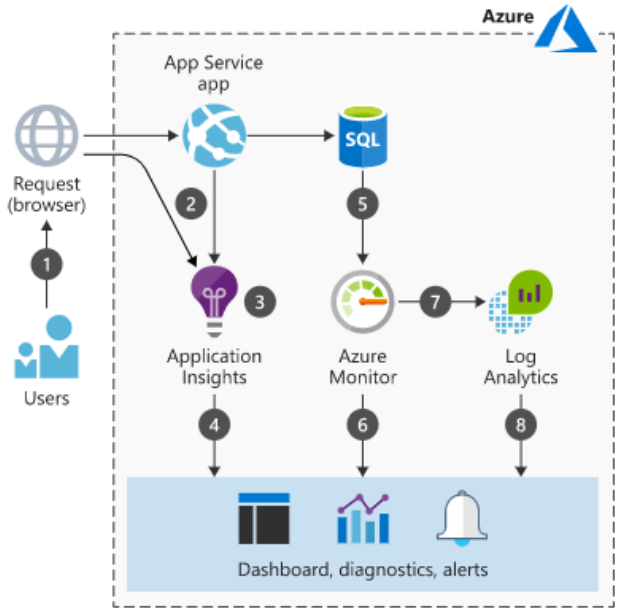


Figure 117 - Esempio di Azure Monitor

Oltre a monitorare e raccogliere i log, possono essere configurate delle risposte automatiche che alla base di eventi specificati, rispondono in modo proattivo agli eventi critici.

La risposta potrebbe essere un testo o un messaggio di posta elettronica a un amministratore o un processo automatizzato che tenta di correggere una condizione di errore.

- **Key Vault**

Azure Key Vault è un servizio cloud per archiviare i segreti e accederli in modo sicuro.

Un segreto è qualsiasi elemento per cui si vuole controllare rigorosamente l'accesso, ad esempio chiavi API, password, certificati o chiavi crittografiche.

Azure Key Vault applica il protocollo Transport Layer Security (TLS) per proteggere i dati quando si viaggia tra l'insieme di credenziali delle chiavi di Azure e i client. I client negoziano una connessione TLS con Azure Key Vault. Il protocollo TLS offre autenticazione avanzata, riservatezza dei messaggi e integrità (abilitando il rilevamento di manomissioni, intercettazioni e falsificazioni di messaggi), interoperabilità, flessibilità degli algoritmi e facilità di distribuzione e di utilizzo.

- **Web Application Firewall**

Web Application Firewall (WAF) di Azure, presente nel Application Gateway Azure offre protezione centralizzata delle applicazioni Web da exploit e vulnerabilità comuni. Ad esempio, gli attacchi SQL injection e quelli tramite scripting intersito.

Per abilitare Web Application Firewall nel Application Gateway, è necessario creare un criterio di WAF, che può essere associato ad uno o più Application Gateway. Questo criterio include tutte le regole gestite, le regole personalizzate, le esclusioni e altre personalizzazioni, come ad esempio il limite per il caricamento di file.

È possibile creare criteri completamente personalizzati che soddisfino specifici requisiti di protezione delle applicazioni combinando regole gestite e personalizzate

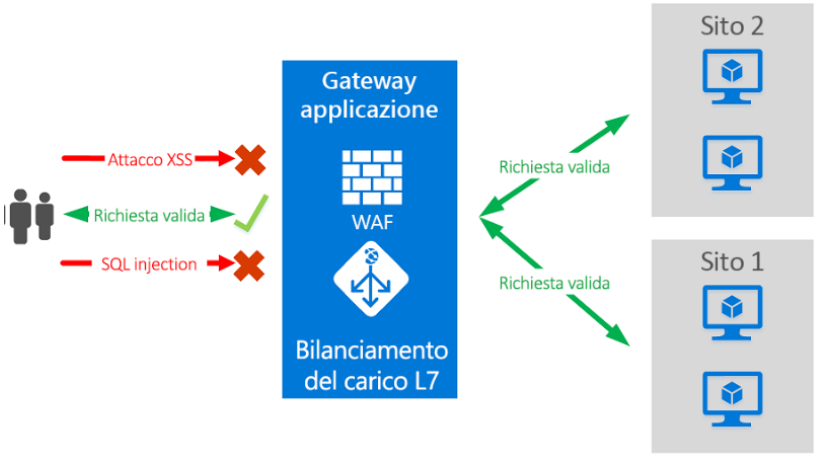


Figura 126 - Web Application Firewall

Tutti gli eventi sono integrati con Azure Monitor, consentendo di archiviare i log in uno Storage Account, trasmettere eventi all'Event Hub, o inviarli ai log di Azure Monitor.

1.2.1.2 Industry Standard Descrizione Servizio IaaS

I servizi di tipo Infrastructure as a Service (IaaS) sono servizi *Core* e prevedono l'utilizzo, da parte dell'Amministrazione, di risorse infrastrutturali virtuali erogate in remoto. Infrastructure as a Service (IaaS) è uno dei tre modelli fondamentali di servizio di cloud computing. Come tutti i servizi di questo tipo, fornisce l'accesso a una risorsa informatica appartenente a un ambiente virtualizzato tramite una connessione Internet. La risorsa informatica fornita è specificamente un hardware virtualizzato, in altri termini, un'infrastruttura di elaborazione. La definizione include offerte come lo spazio virtuale su server, connessioni di rete, larghezza di banda, indirizzi IP e bilanciatori di carico.

Figura 127 - Infrastructure as a Service

Il servizio IaaS è suddiviso in:

- **IaaS Shared:** Fondamentalmente, il servizio IaaS Shared garantisce delle risorse elaborative al cliente finale e tali risorse sono individuate attraverso dei Pool di Risorse "elastiche" che comprendono vCPU, vRAM e Storage Space. Le risorse sono definite elastiche perchè i Pool possono essere scelti in differenti sizing in funzione delle esigenze e, una volta allocati, possono essere pur sempre oggetto di resizing. Grazie alla disponibilità di questo Pool di Risorse, il cliente finale potrà autonomamente creare e gestire VMs e relativo vNetworking per consentire l'erogazione di un determinato modello di servizio applicativo installato all'interno delle VM sempre in modo del tutto autonomo. Le risorse elaborative incluse nel Pool di Risorse sono ricavate su Bare Metal Hypervisors server condivisi con altri Pool di Risorse di altri clienti ma ad ogni modo, come già descritto, ogni cliente avrà una netta separazione logica rispetto al contesto/workload di ogni altro cliente. I Pool di Risorse possono essere allocati in modalità "Local Only" in una specifica HA Zone oppure in modalità "Stretched" e quindi con span in due HA Zone. All'interno del proprio contesto, il cliente finale disporrà anche di un Catalogo di VM template da poter utilizzare per avviare appunto istanze di VM nelle proprie risorse elaborative disponibili. Il Catalogo conterrà VM template generati dal PSN come fornitore del servizio ma potrà anche avere una sezione privata e quindi gestita autonomamente dal cliente finale per la registrazione di VM template "proprietari" da poter mettere a disposizione dei propri utenti finali.

1.2.2 DBaaS

1.2.2.1 Secure Public Cloud Descrizione Servizio DbaaS

Il Servizio PaaS consiste nella messa a disposizione di una piattaforma in grado di erogare elementi applicativi e middleware come servizio, come ad esempio i Data Base. Il SCP Azure, in qualità di provider, si fa carico di gestire l'infrastruttura sottostante, comprensiva degli strumenti di automation e orchestration. L'offerta dei servizi PaaS prevede un approccio strutturato in cui ogni componente della soluzione PaaS, come il sistema operativo, solution stack ed altri software necessari, è strettamente controllata in termini di utilizzo e configurazione e gestita dal SCP Azure. In questo caso le soluzioni vengono "create" al momento della necessità. Una rappresentazione di questa strutturazione vede quattro livelli di componenti:

- sistema operativo;
- run-time e librerie necessarie;
- soluzione caratterizzante – tipicamente un database, middleware, web server, etc.;
- un'interfaccia programmatica con cui controllare gli aspetti operazionali della soluzione.

- **Azure Database for PostgreSQL**

Azure Database for PostgreSQL è un servizio di database relazionale basato sul motore open source PostgreSQL. Questo servizio offre scalabilità, affidabilità, sicurezza e prestazioni elevate per le applicazioni che richiedono PostgreSQL. Con Azure Database for PostgreSQL, è possibile creare e gestire database PostgreSQL nel cloud senza doversi occupare della manutenzione, del backup, del ripristino o del patching dell'infrastruttura sottostante.

Azure Database for PostgreSQL offre numerosi vantaggi per le applicazioni che usano PostgreSQL come database relazionale. Alcuni di questi vantaggi sono:

- Scalabilità: è possibile modificare le dimensioni del database in base alle esigenze, sia in termini di spazio di archiviazione che di potenza di calcolo, senza interruzioni del servizio.
- Affidabilità: il servizio garantisce una disponibilità elevata e una continuità operativa, grazie a funzionalità come il failover automatico, il backup automatico e il ripristino.
- Prestazioni: il servizio ottimizza le prestazioni del database con funzionalità come l'indicizzazione iperveloce, le query intelligenti e il tuning automatico.
- Integrazione: il servizio si integra facilmente con altri servizi cloud di Azure, come Azure Monitor, per il monitoraggio e l'analisi delle metriche del database.
- Backup: Il servizio crea automaticamente backup del server e li archivia nell'archiviazione con ridondanza ZRS. I backup possono ripristinare il server in qualsiasi punto entro il periodo di conservazione dei backup. Tutti i backup vengono crittografati con crittografia AES a 256 bit.
- Monitor: Il servizio è dotato di funzionalità predefinite di monitoraggio delle prestazioni e di avviso. È possibile configurare avvisi in base alle metriche. Tutti gli eventi sono integrati con Azure Monitor, consentendo di archiviare i log in uno Storage Account, trasmettere eventi all'Event Hub, o inviarli ai log di Azure Monitor.

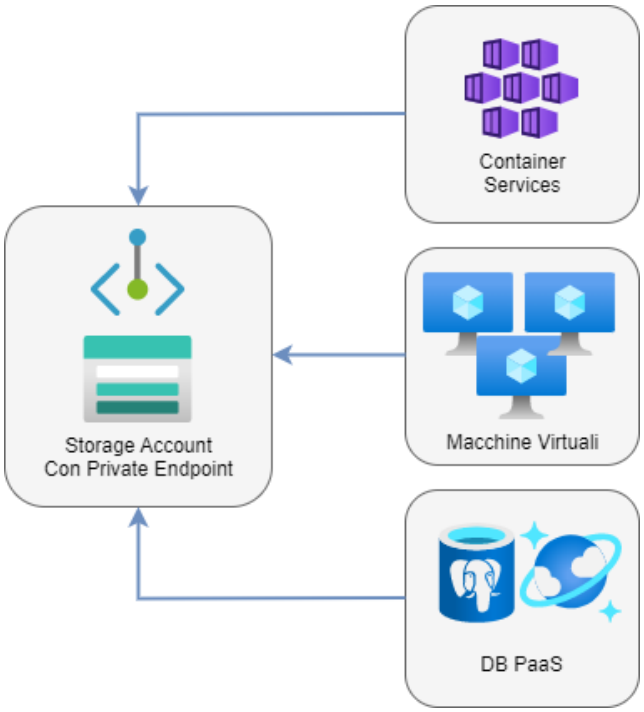


Figura 128 - Azure Database for PostgreSQL

- **Azure Cosmos DB**

Azure Cosmos DB è un database NoSQL completamente gestito e relazionale per lo sviluppo di app moderne.

Azure Cosmos DB è un servizio di database multi modello distribuito globalmente che offre scalabilità orizzontale, coerenza flessibile e garanzie di throughput elevate. Azure Cosmos DB consente di creare applicazioni in grado di gestire dati eterogenei e adattarsi facilmente ai cambiamenti di schema e di carico. Azure Cosmos DB offre inoltre funzionalità avanzate come indicizzazione automatica, query SQL, trigger, stored procedure e funzioni definite dall'utente.

Cosmos DB per MongoDB offre numerosi vantaggi rispetto ad altre offerte di servizi MongoDB, ad esempio MongoDB Atlas:

- Scalabilità istantanea: con la funzionalità di scalabilità automatica, il database viene ridimensionato istantaneamente con zero periodo di "riscaldamento".
- Distribuzioni serverless: Cosmos DB per MongoDB UR offre una modalità di capacità serverless.
- Backup: Azure Cosmos DB per MongoDB UR offre backup continui
- Monitor: Cosmos DB per mongoDB UR si integra in modo nativo con Monitoraggio di Azure per offrire funzionalità di monitoraggio approfondite.

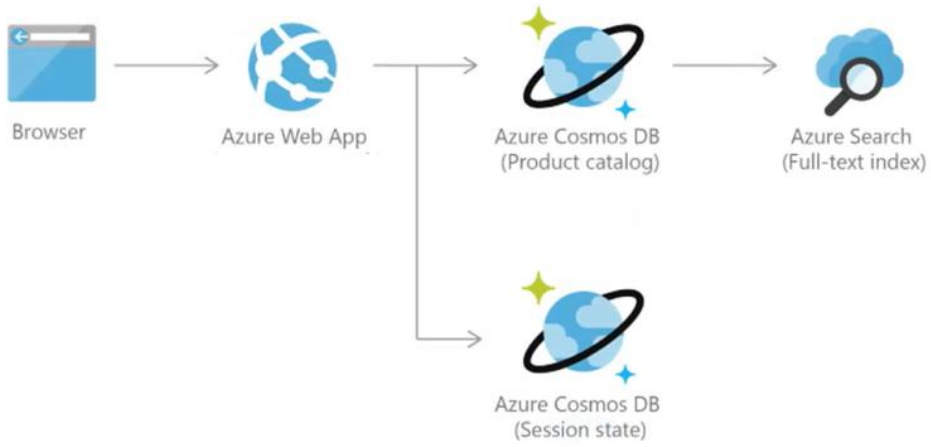


Figura 129 - Azure Cosmos DB

1.2.2.2 Industry Standard Descrizione Servizio DbaaS

Il Database-as-a-Service è un servizio che consente agli utenti di configurare, gestire e ridimensionare database utilizzando un insieme comune di astrazioni secondo un modello unificato, senza dover conoscere o preoccuparsi delle esatte implementazioni per lo specifico database. Viene demandato al provider tutto quanto relativo all'esercizio e alla gestione dell'infrastruttura sottostante, comprese le operazioni di riconfigurazione della capacità elaborativa e delle repliche, mentre gli utenti possono così focalizzarsi sulle funzionalità applicative ed estrarre valore dai dati.

Tramite la console di gestione del servizio vengono messe a disposizione del cliente in particolare le funzionalità di:

- Creazione (o cancellazione) di un database;
- Modifica delle principali caratteristiche infrastrutturali dell'istanza DB e ridimensionamento ove non automatico;
- Configurazione di alcuni parametri del database;
- Attivazione di funzionalità aggiuntive, come ad esempio la replica dei dati su istanze passive (ove applicabile);
- Attivazione di funzionalità di backup od esportazione dei dati (ove applicabile).

Altre funzionalità avanzate di configurazione delle specifiche istanze database sono demandate alle relative interfacce di amministrazione native.

Il catalogo del servizio comprende:

- **Database relazionali** (**Oracle DB Enterprise e Standard, MySQL, PostgreSQL, Maria DB, …**) che supportano il modello dati relazionale e lo standard SQL di interrogazione. Sono quindi adatti a spostare carichi di lavoro di DB SQL preesistenti a casa del cliente su ambienti moderni e sicuri, in grado di garantire l'elevata affidabilità e le possibilità di crescita offerte dal Cloud;

- **Database NoSQL** (**MongoDB**, …) ottimizzati per trattare dati non strutturati, con volumi elevati o con caricamento di grandi quantità di informazioni in modelli dati flessibili e con bassa latenza.

1.2.3 PaaS

I Servizi PaaS Industry in generale consistono nella messa a disposizione di piattaforme in grado di erogare elementi applicativi e/o middleware sottoforma di servizio, per cui viene affidata alla responsabilità del fornitore (PSN) la gestione dell'intera pila della infrastruttura sottostante la piattaforma, comprensiva degli strumenti di automation e orchestration.

Il Platform as a Service è un modello di delivery di servizi Cloud Computing in cui il fornitore mette a disposizione degli utenti una piattaforma cloud completa (infrastruttura, hardware e piattaforme software) per lo sviluppo, l'esecuzione e la gestione di applicazioni, accedendo alle tecnologie in logica "*as a Service*".

Il PaaS si configura come un servizio che offre all'utente la possibilità di accedere agli strumenti messi a disposizione dalla piattaforma a diversi livelli: da un livello più basilare per il funzionamento di un particolare software ad un livello più complesso che potrebbe avvicinarsi ad un vero e proprio servizio applicativo di tipo SaaS.

Gli strumenti tipicamente inclusi in una piattaforma PaaS (oltre alle risorse di infrastruttura come networking, capacità di storage, capacità elaborativa, …) possono essere sistemi operativi, sistemi di sicurezza, sistemi di gestione dei dati, servizi di integrazione, etc…

Nel caso specifico, l'offerta dei servizi PaaS Industry del PSN prevede un approccio strutturato in cui ogni componente della soluzione PaaS è strettamente controllato in termini di utilizzo e configurazione e gestito dal PSN. Le soluzioni sono istanziate al momento della necessità, secondo il modello *on-demand*, sfruttando strumenti di automazione ed orchestrazione. La struttura del servizio può essere rappresentato da una serie di livelli come nella figura a lato: sono mostrate, dal basso verso l'alto, prima le componenti infrastrutturali (Networking, Storage, Server, OS e piattaforma di containerizzazione), successivamente le componenti software che completano/costituiscono la soluzione tecnologica (tipicamente, database, middleware, tool applicativi, ecc.) per ogni specifica piattaforma PaaS ed infine la rappresentazione dell'interfaccia con cui controllare gli aspetti operazionali dei servizi di piattaforma offerti.

La gestione dell'infrastruttura HW e SW rappresentata in figura è totalmente demandata al PSN che fornisce la piattaforma: l'utente è responsabile solamente di configurare e, in alcuni casi sviluppare, la soluzione applicativa, utilizzando gli strumenti applicativi che saranno messi a disposizione per il servizio.

- **Architettura di Alto Livello del PaaS Industry**

Questo paragrafo fornisce una rappresentazione ad alto livello dell'architettura generale dell'intero servizio PaaS Industry, elencando tutti i servizi di piattaforma inclusi nel catalogo del PaaS Industry e di alcuni servizi trasversali di supporto.

Le piattaforme PaaS Industry oggetto di questo documento sono:

- Big Data
- AI (Artificial Intelligence)
- IAM

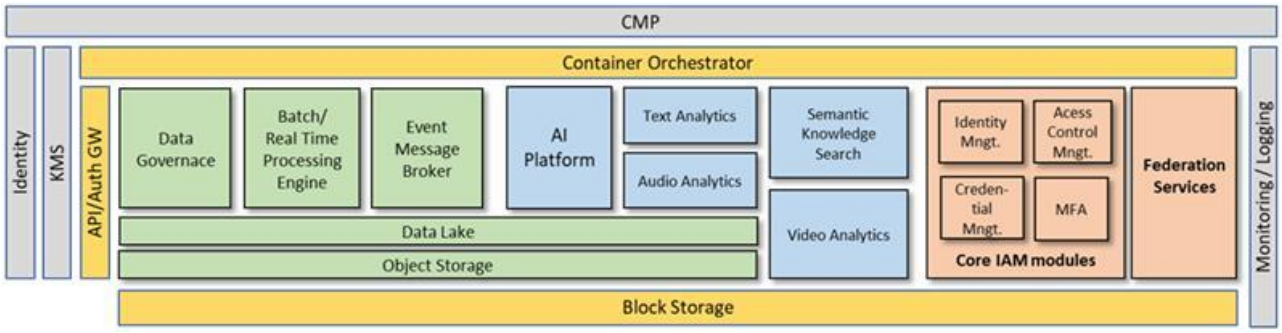


Figura 130 – PaaS Industry, rappresentazione schematica

La figura sopra vuole dare una rappresentazione schematica unitaria di tutti i servizi:

- In colore ***verde*** sono rappresentati i servizi di piattaforma del *PaaS Big Data*:
  - Data Lake
  - Batch/Real Time Processing
  - Event Message Broker
  - Data Governance
- In colore ***azzurro*** sono rappresentati i servizi di piattaforma del *PaaS AI*:
  - AI Platform
  - Semantic Knowledge Search
  - Text Analytics
  - Audio Analytics
  - Video Analytics
- in colore ***salmone*** è rappresentato il servizio di piattaforma unico del *PaaS IAM* il quale è composto da vari moduli riportati in figura (Federation Services e Core IAM modules)

In Figura 1, con il colore grigio sono rappresentati alcuni servizi esterni alla soluzione PaaS Industry ed erogati dall'infrastruttura del PSN, come il Cloud Management System (CMP), il sistema di accesso (Identity), la soluzione per la gestione delle chiavi di crittografia (KMS), il sistema di Log Management/Monitoring.

In arancione sono rappresentate alcune componenti dedicate ai servizi PaaS come la piattaforma infrastrutturale kubernetes per l'erogazione dei servizi rappresentato dal Container Orchestrator, la soluzione di File&Block Storage e i servizi di API Gateway e Authorization Gateway.

Si vuole far notare che in Figura 1 è rappresentato graficamente il fatto che i servizi di piattaforma *Batch/Real Time processing*, *Event Message* e *Data Governance* del PaaS Big Data e i servizi di piattaforma *AI Platform* del PaaS AI, Text Analytics e Audio Analytics si *"appoggiano"* sul servizio *Data Lake* messo a disposizione dal *PaaS Big Data* (come descritto in dettaglio nel paragrafo 2.3 "*Modello di erogazione*").

Il presente documento fornisce una vista di alto livello dei servizi rappresentati in modo da rendere una visione d'insieme del servizio PaaS Industry. Le descrizioni dei singoli servizi di piattaforma saranno invece sviluppate nei documenti HLD dei relativi servizi PaaS Industry.

1.2.4 CaaS

Il CaaS è un particolare modello IaaS in cui le risorse di calcolo, invece che essere messe a disposizione come macchine virtuali (VM) o server bare metal, sono messe a disposizione in modalità Container che rappresentano l'unità di risorsa primaria con la quale si eroga il servizio.

Nell'ambito dei servizi di cloud computing, il CaaS quindi offre un'estensione delle funzionalità rispetto all' Infrastructure as a Service (IaaS), gestendo le risorse Container. Esso non possiede tutte le funzionalità offerte dai servizi di Platform as a Service (PaaS).

Come estensione dei servizi IaaS, il CaaS prevede che sia soddisfatto il prerequisito che servizi Infrastrutturali di base (come servizi di connettività, rete, sicurezza, bilanciamento, backup, crittografia, etc…) siano erogati da preesistenti servizi già disponibili.

Nel servizio CaaS, il fornitore mette a disposizione:

- l'infrastruttura per ospitare la piattaforma del servizio CaaS
- la piattaforma per l'erogazione delle funzionalità CaaS che include anche gli strumenti di automation e orchestration

Il servizio CaaS prevede che sia di totale responsabilità del fornitore la gestione dell'infrastruttura sottostante, della piattaforma di erogazione dei container, dei tool di orchestrazione e automazione e dei vari tool di gestione. Il cliente è responsabile dell'esecuzione dei propri container.

Il servizio CaaS permette all'utente di gestire:

- Il deploy dei container
- L'avvio dell'esecuzione dei container
- L'arresto dell'esecuzione dei container
- Il ridimensionamento dei container
- Lo scale-up e lo scale-down del numero di container
- L'organizzazione dei workloads a container, la configurazione di task e servizi
- Il servizio fornisce anche strumenti che permettono all'utente di monitorare lo stato dei propri container.

L'utente ha inoltre la possibilità di gestire alcune policy di connessioni tra container dello stesso namespace (area segregata logicamente, dedicata al cliente per il proprio progetto) attraverso i tool di Software Define Network fornite dalla piattaforma e di esporre i servizi associati al container su rete esterna alla piattaforma di erogazione dei container in modalità http, https, udp, tcp.

Nel caso di Container che richiedono la persistenza del dato, vi è la possibilità di eseguire delle Persistent Volume Claim (PVC) per utilizzare storage persistente di tipo Block&File in modalità RWO e RWX.

A tutte queste funzionalità offerte dal CaaS l'utente accede attraverso uno unico strumento di gestione sia in modalità CLI sia Web Console.

In servizio CaaS è integrato con soluzione di Object storage di tipo S3, per permettere all'applicazione a Container di utilizzare storage ad oggetti.

Il servizio è predisposto pensando anche che il cliente possa essere libero di utilizzare processi e strumenti già attivi per il DevSecOps, CI/CD e che dal servizio CaaS richieda solo la piattaforma per eseguire i propri workload applicativi a Container. In questo caso il servizio parte dal presupposto che esistono le immagini presenti su un registry del cliente.

La soluzione tecnologica sottostante al servizio CaaS garantisce un livello di resilienza elevata a livello Regional mantenendo la continuità operativa attraverso la distribuzione della piattaforma di erogazione dei Container su 3 Availabilty Zone (AZ). All'interno della Region questa soluzione tollera l'indisponibilità di una AZ con valore del parametro RPO uguale a zero ed il valore del parametro RTO dell'ordine dei minuti.

Il servizio CaaS si basa su tecnologia Red Hat: RH OCP (Red Hat OpenShift Contanier Platform) al quale vengono aggiunte dei strumenti di gestione avanzata in ambiente multi cluster e per la sicurezza.

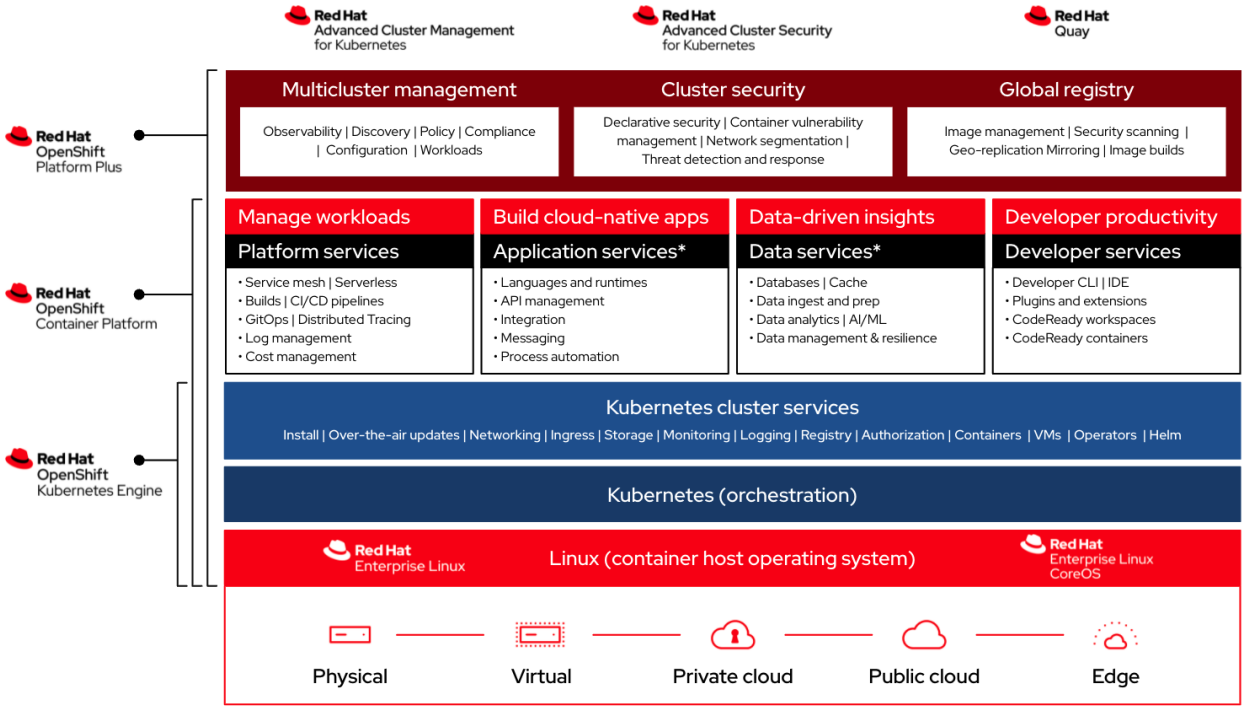


Figura 131 – Red Hat OpenShift Contanier Platform

Il Cluster Kubernetes basato su tecnologia Red Hat OpenShift Contanier Platform, gli strumenti di gestione forniti da Red Hat adottati sono: Red Hat Advanced Cluster Manager, Red Hat Advanced

Cluster Security ed il registry Red Hat Quay. Per la componente di Block&File Storage utilizzo di una soluzione Software Defined Storage (SDS) fornita dal prodotto Red Hat Openshift Data Foundation su dischi interni SSD ai server dedicati al Block&File Storage

OpenShift Container Platform (RH OCP) è una piattaforma per lo sviluppo e l'esecuzione di applicazioni containerizzate con la quale le organizzazioni possono creare applicazioni con maggior velocità, distribuire, eseguire e gestire le applicazioni ovunque, in sicurezza e su larga scala. È progettato infatti per consentire alle applicazioni e ai Data Center che le supportano di scalare facilmente ad un grande numero di applicazioni e sistemi per erogare servizi contemporaneamente a milioni di clienti.

Basato su Kubernetes, OpenShift Container Platform incorpora la stessa tecnologia che funge da motore per tutte le applicazioni anche business critical: consente di supportare una varietà di casi d'uso, come artificiale intelligenza e apprendimento automatico (AI/ML) che gestisce i big data. etc.

La sua implementazione in tecnologie Red Hat opensource offre una gestione operativa coerente dell'infrastruttura Kubernetes sottostante in qualsiasi ambiente e consente di estendere le applicazioni containerizzate oltre un singolo cloud a on-premise e multi-cloud.

- **Kubernetes**

Sebbene le immagini dei container ed i container siano gli elementi costitutivi principali per lo sviluppo di applicazioni moderne, eseguirle su larga scala richiede un sistema di distribuzione affidabile e flessibile. Kubernetes è lo standard di fatto per l'orchestrazione dei container.

Kubernetes è un motore di orchestrazione di container open source per automatizzare la distribuzione, il dimensionamento e la gestione delle applicazioni containerizzate.

In pochi anni, Kubernetes ha visto un'enorme adozione sia in ambito cloud che on-premise. Il modello di sviluppo open source consente a molte persone di estendere Kubernetes implementando diverse tecnologie per componenti come rete, storage e autenticazione.

- **Applicazioni Containerizzate**

L'utilizzo di applicazioni containerizzate offre molti vantaggi rispetto all'utilizzo dei metodi di distribuzione tradizionali. Laddove una volta ci si aspettava che le applicazioni fossero installate su sistemi operativi che includono tutte le loro dipendenze, i container consentono a un'applicazione di portare con sé le loro dipendenze. La creazione di applicazioni containerizzate offre molti vantaggi:

- Sistema operativo: I container utilizzano piccoli sistemi operativi Linux dedicati senza kernel. Il loro file system, rete, cgroup, process table e namespace sono separati dal sistema Linux host, ma i container possono integrarsi perfettamente con gli host quando necessario. Essendo basato su Linux permette ai container di sfruttare tutti i vantaggi che derivano dal modello di sviluppo open source di rapida innovazione. Poiché ogni container utilizza un sistema operativo dedicato, le applicazioni che richiedono dipendenze software in conflitto possono essere distribuite sullo stesso host. Ogni container trasporta il proprio software e gestisce le proprie interfacce, come reti e file system, per cui le applicazioni non devono mai competere per tali risorse.
- Deployment & Scaling: Nel ciclo di vita delle applicazioni, queste possono essere continuamente migliorate senza interruzioni e mantenere comunque la compatibilità con la

versione corrente. E' possibile implementare e testare una nuova versione insieme alla versione esistente, distribuendola in aggiunta alla versione corrente. Analogamente è possibile scalare le applicazioni, ridimensionando i singoli microservizi.

- **OpenShift Container Platform**

OpenShift Container Platform rende Kubernetes enterprise-ready, includendo:

- Deployment su cloud ibridi: i cluster OpenShift Container Platform possono essere distribuiti su una varietà di piattaforme cloud pubbliche o in Data Center.
- Tecnologia Red Hat integrata: la maggior parte delle componenti di OpenShift Container Platform vengono da Red Hat Enterprise Linux e altre tecnologie Red Hat. OpenShift Container Platform beneficia delle iniziative di test e certificazione previste per tutto il software fornito da Red Hat.
- Modello di sviluppo open source: il codice sorgente è disponibile da repository software pubblici favorendo una rapida innovazione e sviluppo.

Le sezioni seguenti descrivono alcune caratteristiche di OpenShift Container Platform.

Sistema operativo

OpenShift Container Platform utilizza Red Hat Enterprise Linux CoreOS (RHCOS), un sistema operativo immutabile e orientato ai container che combina alcune delle migliori caratteristiche e funzioni dei sistemi operativi CoreOS e Red Hat Atomic Host.

RHCOS è progettato specificamente per l'esecuzione di applicazioni containerizzate da OpenShift Container Platform e funziona con nuovi strumenti per fornire un'installazione rapida, una gestione basata sugli operator e aggiornamenti semplificati.

RHCOS include:

- Ignition: OCP lo utilizza come configurazione del sistema durante la fase di firstboot per avviare e configurare i nodi.
- CRI-O, container runtime nativo di Kubernetes che implementa l'interfaccia CRI (Container Runtime Interface) e si integra strettamente con il sistema operativo per offrire un'esperienza Kubernetes efficiente e ottimizzata. CRI-O fornisce strutture per l'esecuzione, l'arresto e il riavvio dei container. Sostituisce completamente il Docker Container Engine, utilizzato nelle versioni di OpenShift Container Platform 3.
- Kubelet, l'agente eseguito sui compute node responsabile dell'avvio e del monitoraggio dei container tramite l'interazione con il container runtime, nonché del monitoraggio dello stato del nodo in termini di carico e saturazione

Sebbene sia possibile utilizzare RHEL come sistema operativo per i nodi compute, il PSN utilizzerà RHCOS per tutte le macchine del cluster vista la modalità di installazione. Inoltre non sarebbe possibile utilizzare RHEL per installazioni di tipo IPI.

Processo di installazione e aggiornamento

Con OpenShift Container Platform è possibile creare un cluster di produzione, eseguendo un singolo comando. Il processo di installazione è ridotto notevolmente rispetto alla versione precedente.

Per i cluster che utilizzano RHCOS per tutti i nodi, come nel caso del PSN, l'aggiornamento di OpenShift Container Platform è un processo semplice e altamente automatizzato. Poiché OpenShift Container Platform controlla completamente i sistemi e i servizi in esecuzione su ogni macchina, incluso il sistema operativo stesso, gli aggiornamenti sono progettati per diventare eventi automatici.

Il processo di installazione sarà descritto più avanti nel documento e terrà conto dei prodotti selezionati per il PSN, come ad esempio RHACM per la gestione centralizzata dei cluster e delle applicazioni ospitate.

Operator

Gli Operator sono un framework che semplificano il provisioning e la gestione automatizzata del ciclo di vita di uno stack software su OpenShift, rendendo facile e conveniente distribuire ed aggiornare componenti software containerizzati.

In OpenShift Container Platform, gli operator fungono da base della piattaforma ed eliminano la necessità di aggiornamenti manuali dei sistemi operativi e delle applicazioni della control plane, infatti operator come Cluster Version Operator e Machine Config Operator consentono una gestione semplificata a livello di cluster di questi componenti critici.

Operator Lifecycle Manager (OLM) e Operator Hub forniscono strutture per contenere e distribuire

gli operator a chi si occupa di sviluppare e distribuire applicazioni.

Telemetry

In OpenShift Container Platform, si richiede l'accesso a Internet per installare il cluster, direttamente o tramite proxy. Anche il servizio di Telemetria, che viene eseguito di default per fornire metriche sull'integrità del cluster e sull'esito positivo degli aggiornamenti, richiede l'accesso a Internet.

Una volta installato il primo cluster, il servizio Telemetry automaticamente registra il cluster in Red Hat OpenShift Cluster Manager (OCM). Confermata la presenza del cluster, è possibile visualizzare lo stato e la scadenza delle sottoscrizioni per monitorarlo. In generale il servizio Telemetry astrae la gestione delle sottoscrizioni a livello di account o multi-cluster.

Il PSN deve garantire l'accesso a internet direttamente o tramite proxy per:

- Accedere alla pagina Red Hat OpenShift Cluster Manager per scaricare il programma di installazione ed eseguire la gestione della sottoscrizione.
- Accedere a Quay.io per ottenere i pacchetti necessari per installare il cluster.
- Ottenere i pacchetti necessari ad eseguire aggiornamenti.

Red Hat OpenShift Data Foundation (**RH ODF**) è un Software-Defined Storage (SDS) per container.

Progettato come piattaforma di servizi di archiviazione e dati per Red Hat OpenShift, Red Hat OpenShift Data Foundation facilita lo sviluppo e la distribuzione di applicazioni in modo rapido ed efficiente tra i cloud.

L'architettura disegnata per il PaaS del PSN sfrutta Openshift Plaftform Plus - ODF Advanced anche per le funzionalità di Crittografia e di gestione del Disaster Recovery (Regional DR).

Red Hat Advanced Cluster Management (**RH ACM**) for Kubernetes fornisce visibilità e controllo end-to-end per gestire i cluster e il ciclo di vita delle applicazioni stesse da un'unica console. Include strumenti avanzati di sicurezza e conformità per la totalità dei cluster gestiti, ospitati su più data center e cloud pubblici.

È uno strumento qui di che estende le funzionalità di Red Hat OpenShift distribuendo applicazioni, gestendo più cluster e applicando policy su più cluster su larga scala, garantisce la conformità, monitora l'utilizzo e mantiene la coerenza.

Le principali caratteristiche di RH ACM sono:

- Gestione centralizzata e unificata dei cluster:
  - Creazione, aggiornamento ed eliminazione cluster OCP su più cloud pubblici e privati
  - Gestione centralizzata delle risorse Kubernetes
  - Possibilità di eseguire troubleshooting orizzontalmente su tutti i domini gestiti o federati
- Gestione del rischio e criteri di conformità basati su policy:
  - Possibilità di impostare e applicare criteri per la sicurezza di applicazioni e infrastruttura
  - Visualizzare più velocemente audit dettagliati sulla configurazione di applicazioni e cluster
  - Ottenere visibilità immediata sulla conformità in base agli standard definiti
- Gestione avanzata del ciclo di vita delle applicazioni:
  - Distribuire applicazioni su larga scala
  - Distribuire applicazioni da più origini
  - Visualizzare più rapidamente le relazioni tra le applicazioni tra i cluster
- Osservabilità multicluster dello stato di consistenza dei cluster e ottimizzazione:
  - Ottenere una panoramica dell'integrità e dell'ottimizzazione dei cluster utilizzando dashboard personalizzate e pronte all'uso utilizzando le metriche collezionate
  - Ordinare, filtrare e analizzare le prestazioni di singoli cluster o multicluster aggregati
  - Risolvere i problemi più velocemente utilizzando la ricerca dinamica e le funzionalità della console web

In generale mentre OCP si concentra sul modello di applicazione a cluster singolo e fornisce un eccellente framework per la CI/CD, RH ACM modella le applicazioni per un'architettura multicluster con funzionalità enterprise che aiutano a garantire sia la rapida implementazione di un'applicazione che la resilienza di tutta l'infrastruttura.

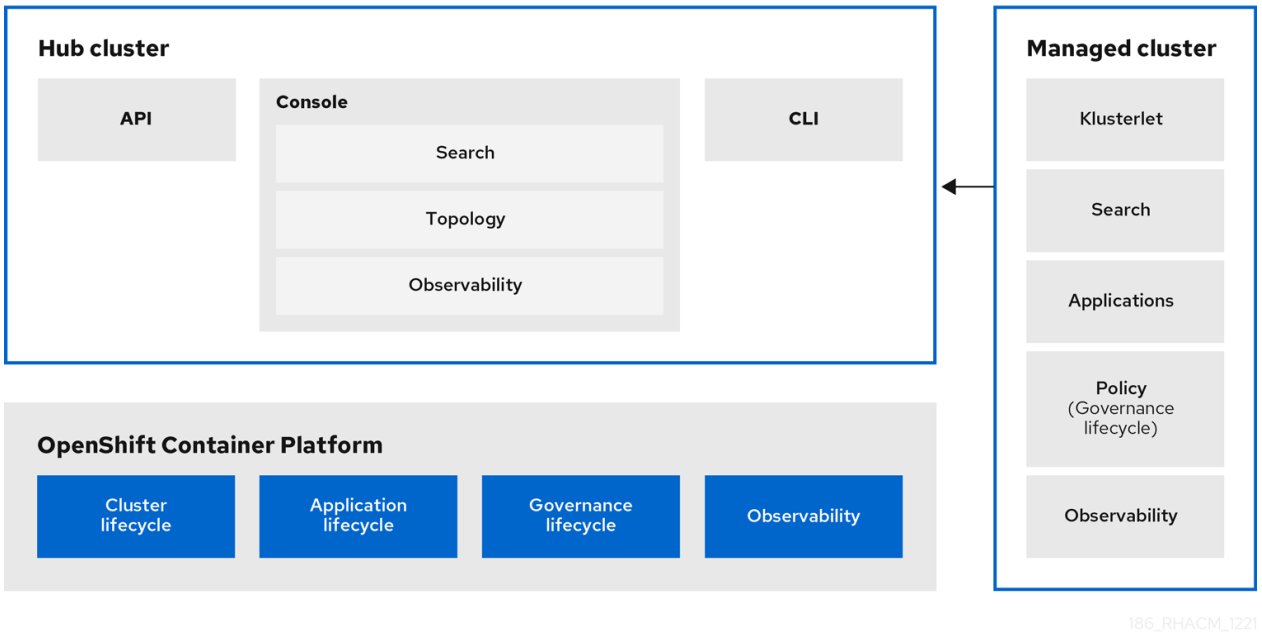


Figura 132 – RedHat ACM e ODF

Tra le funzionalità più interessanti nell'utilizzo congiunto di RH ACM e ODF proprio nella soluzione disegnata per il PSN, sono sicuramente quelle che permettono le soluzioni di DR, Backup&Restore e Volume Replication.

- Hub Backup&Restore: è possibile eseguire backup e restore delle configurazioni dei siti gestiti anche su hub differenti, utilizzando la soluzione di backup che utilizza le api OADP.
- Regional DR: ODF e RHACM offrono una strategia robusta e multisite di DR per applicazioni stateful. Mentre ODF fornisce volumi persistenti alle applicazioni, e ne gestisce la replica, gli operator di DR di RHACM automatizzano le operazioni di failover e failback, assicurando un recovery veloce e libero da interventi manuali. La replica asincrona dei volumi è basata su rbd mirroring, e la copia del dato avviene tramite i tunnel Submariner orchestrati da RH ACM tramite i ClusterSet.
- VolSync : esiste la possibilità di effettuare la replica asincrona dei volumi persistenti di tipo CSI tramite VolSync su volumi erogati da storage esterni enterprise, garantendo una strategia di migrazione pianificata tra i vari siti, delle applicazioni stateful.

Red Hat Advanced Cluster Security è la soluzione Enterprise utilizzata per la verifica della sicurezza dei container in ambito cluster Openshift per i servizi PaaS.

Include funzionalità di sicurezza Kubernetes native che permettono di individuare le vulnerabilità critiche negli ambienti OCP gestiti, supportando i processi di verifa, monitoraggio e correzione di problematiche di sicurezza individuate.

Di seguito riportiamo alcune delle principali funzionalità che è possibile attivare tramite RH ACS:

- Vulnerability Management:

  - Individuazione delle vulnerabilità nelle immagini dei container in esecuzione sulla piattaforma OCP:
  - Possibilità di eseguire uno scanning delle immagini al fine di individuarne vulnerabilità note.
  - Identificazione dei deployment in cui sono utilizzate immagini che hanno evidenziato vulnerabilità
- Network Segmentation:
  - Possibilità di verificare e confrontare il traffico di rete effettivo rispetto a quello consentito, applicando criteri di rete e aumentando la segmentazione utilizzando controlli Kubernetes nativi.
  - Simulazione dell'applicazione di network policy prima dell'effettiva applicazione al fine di ridurre il rischio di errori che potrebbero portare a problemi nell'erogazione dei servizi.
- Compliance:
  - Possibilità di verificare la conformità dei sistemi ai benchmark di sicurezza adotattati globalmente in ambito IT Security.
  - Approfondire i dettagli di conformità per individuare cluster o namespace che non sono conformi a standard e controlli specifici.
- Detection and Response:
  - Possibilità di configurare regole di allow list e baseline per individuare le attività sospette ed adottare le necessarie contromisure per gestire eventuali incident di sicurezza.
  - Monitoraggio degli eventi di sistema che avvengono all'interno dei container al fine di identificare attività anomale che potrebbero indicare una minaccia di sicurezza

1.3 Cyber Security – Professional Services

La messa in produzione di un sistema sul cloud è un processo complesso e un cambiamento rilevante sia per il modello di fruizione, sia per la gestione dei workload e dei dati. È bene specificare che trasferendo le informazioni nel cloud non si trasferisce anche la responsabilità della sicurezza di tali informazioni. Il PSN metterà a disposizione molti strumenti nativi, all'interno delle diverse tipologie di cloud scelte, per gestire la sicurezza dei dati, ma questi devono essere in ogni caso previsti ed implementati dalle Amministrazioni. La responsabilità della sicurezza di tutti i dati trasferiti su cloud rimane sempre e comunque del cliente finale. Il fatto che le infrastrutture cloud siano intrinsecamente dotate di un livello di sicurezza elevato, di per sé non elimina il principio di precauzione che deve sempre essere adottato.

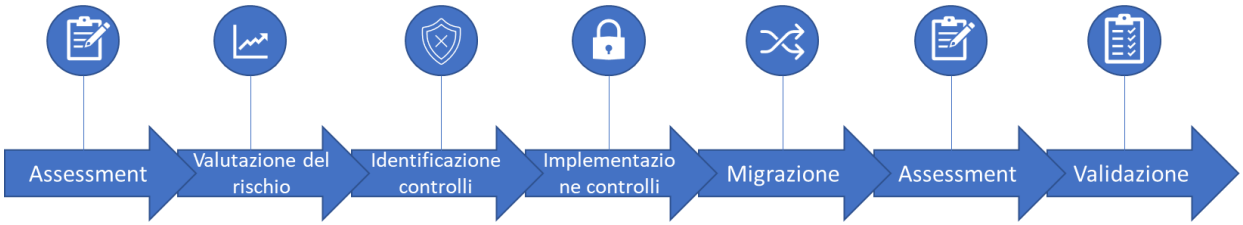
I servizi professionali di sicurezza sono quindi necessari, sinergici e parte integrante dei servizi di set-up e attivazione degli ambienti cloud nelle diverse fasi del progetto e servono sia a valutare lo stato di sicurezza dei workload in perimetro, prevedendo in un approccio security-by-design l'analisi del rischio, l'identificazione, l'implementazione e la gestione dei controlli di sicurezza.

I servizi sono necessari per:

- Garantire la conformità ai requisiti normativi e cogenti.
- Valutare e applicare le best practice di cloud security.
- Mitigare il rischio cyber.

- Valutare rischi e vulnerabilità prima e dopo il processo di migrazione.
- Prevedere, progettare e implementare i controlli di sicurezza
- Supportare l'Amministrazione nella gestione della cybersicurezza.

Di seguito vengono illustrati i diversi step previsti per la gestione della sicurezza implementabili tramite i servizi professionali in oggetto



Il processo di attivazione dei Security Professional Services è caratterizzato da due step distinti:

- Primo step: messa in opera di quanto necessario per attivare il servizio nella sua interezza,
- Secondo step: avvio del "servizio ricorrente" finalizzato ad estendere i servizi di sicurezza man mano che tutta la piattaforma SIM viene dispiegata.

I servizi che saranno erogati per ciascuna delle fasi di progetto sono i seguenti.

Servizi security core (Security Core Services)
Nel periodo iniziale di ogni fase sono previsti dei servizi professionali a supporto delle attività di Assessment, per indirizzare correttamente, dal punto di vista della sicurezza, l'allestimento degli ambienti oggetto del presente progetto del piano dei fabbisogni. Di seguito sono elencati tali servizi:

- Cyber Assessment / Maturity Level Assessment.

Successivamente verranno attivati i servizi di Sicurezza di Base ed Avanzati, elencati di seguito.

- Servizi sicurezza di base (Security Base Services)
  - servizio di Consulting, finalizzato a:
  - verifica della compliance normativa al Framework Nazionale Cyber Security (FNSC) con identificazione di eventuali attività successive di definizione di procedure e processi (prevista in fase avanzata);
  - analisi in ottica cyber&strategic risk management
  - verifica del Cyber Maturity Model
- Protezione perimetrale con secure device management, security policy review/advisory (NGFW/WAF)
- Attività di Vulnerability Management, in sinergia con le attività previste in altri ambiti contrattuali riferibili al SIM
- Compliance Assessment FNCS (per perimetro di migrazione) con identificazione di eventuali attività successive di definizione di procedure e processi;

- Dynamic Application Security Testing per infrastruttura migrata
- Security Event Monitoring & Notification con Log Management e Continuous improvement SIEM per l'ambiente Secure Public Cloud e Industry Standard, erogato in modalità a servizio;
- Supporto di Secure Design & Planning, per l'esecuzione di progettualità d'insieme in ottica Security by Design in funzione della gap analysis e dei controlli di sicurezza indirizzati alla opportuna protezione della rete, dei servizi e degli endpoint;
- Sicurezza hosts con Managed Detection & Response
- Attivazione di servizi di Cyber Threat Intelligence, in particolare:
  - Early Warning & Data Breach, Brand Abuse, Anti-phishing;
  - Pre Planned Attack + Black Market Monitor
- Incident Response & Crisis Management
- Service Assurance

Servizi sicurezza avanzati (Security Advanced Services)

- Servizio di Web Application Penetration Test
- Servizi di cyber threat intelligence, early warning, data breach (IOC)
- Boundle 2: Brand Abuse + Antiphishing / Site Takedown
- Servizio di Security Orchestration, Automation, and Response (SOAR);
- Security Policy review/advisory sui sistemi di protezione perimetrale
- servizio di Consulting: revisione policy e procedure di cyber security
- Servizio di Threat Hunting

Nei paragrafi seguenti è riportata una descrizione di dettaglio dei principali servizi professionali di sicurezza.

1.3.1 Maturity Level Assessment ("security core services")

Il Servizio è erogato as a service ed ha lo scopo di effettuare una gap analysis preliminare dell'attuale contesto infrastrutturale ed applicativo al fine di definire il livello di sicurezza esistente e notificare un report operativo che descrive le necessità per il raggiungimento della conformità rispetto le normative vigenti e le best practices di riferimento, in particolare lo scopo del checkup di sicurezza è analizzare lo stato di maturità di tutti gli ambiti di sicurezza definiti dal Framework Nazionale per Cyber Security e la Data Protection (di seguito per brevità anche "FNCS") integrato con le raccomandazioni dettate dal DPCM 14 aprile 2021 n. 81/2021 in tema di Perimetro di Sicurezza Nazionale Cibernetica.

Verranno proposte una serie di domande attraverso le quali l'Amministrazione potrà acquisire gli elementi utili all'identificazione del miglior approccio cloud, specifico per il proprio contesto. Al completamento delle attività saranno consegnati i seguenti deliverable denominati:

- GA Results Executive Summary: Il report contiene una overview di tipo executive ad alto livello relativo al processo di valutazione che considera quattro aree 'chiave': Business, Functional, Technical, Implementation

- GA Results Assessment Report: Il report contiene i dettagli del processo di valutazione finalizzato ad indirizzare il corretto approccio alla migrazione relativamente alle quattro aree 'chiave' indicate: Business, Functional, Technical, Implementation.

1.3.2 Security Event Monitoring Notification & Log Management e Continuous improvement ("security base services")

Alla luce delle crescenti minacce informatiche per le organizzazioni, diventa fondamentale rivedere l'approccio alla gestione del rischio e individuare strategie per ridurre la vulnerabilità delle infrastrutture informatiche. Quindi per garantire l'adeguato livello di protezione delle reti, dei dati e dei servizi, diventa un fattore di primaria importanza l'individuazione e la gestione immediata degli incidenti di sicurezza.

In tale ottica il presente servizio, erogato remotamente da un Centro Servizi presidiato H24 per 365 giorni l'anno, garantisce un'attività di monitoraggio tramite un team di specialisti (Security Analyst, Security Solution architect, Information Security Consultant) in ambito sicurezza.

A valle del ongoing della soluzione vengono eseguite attività di Continuous Improvement per consentire un tangibile miglioramento della risposta dei sistemi di Security Event Monitoring a fronte dell'insorgere di nuove minacce ed in coerenza con le politiche di sicurezza implementate e nel rispetto delle modalità organizzative adottate dall'Amministrazione. Il servizio è erogato in modalità H24x7 e si articola nelle seguenti fasi:

- Onboarding/Startup: è la fase che precede l'avvio del servizio vero e proprio, con la presa in carico degli accessi alle piattaforme deputate alla "Detection", l'analisi degli allarmi configurati sulle stesse.
- In caso di condivisione, da parte dell'Amministrazione, di regole di correlazione preesistenti, inerenti il perimetro di Cyber Security e finalizzate al miglioramento dalla capacità di Detection del servizio, potranno essere opportunamente integrate all'interno della soluzione prevista.
- Continous Monitoring: è la fase il cui avvio coincide con l'avvio del servizio, è a carattere continuativo ed è costituita da attività di monitoraggio degli allarmi (servizio Live/Running) ed eventi prodotti dalle piattaforme di sicurezza o di ticketing e dalle quali saranno estratte e analizzate le informazioni necessarie all'espletamento delle fasi successive.
- Identification: è la fase in cui l'analista prende in carico un allarme di Sicurezza o una segnalazione e ne identifica i connotati principali al fine di procedere con la fase successiva. A titolo di esempio per ogni allarme preso in gestione vengono estratti se pertinenti i seguenti dati:
  - La tipologia e/o regola di correlazione ad esso associata
  - L'indirizzo IP della sorgente di attacco e della destinazione
  - L'utente o gli utenti coinvolti
  - Indirizzi email o caselle di posta compromessi
  - Il nome e la tipologia del malware usato nell'attacco
  - La vulnerabilità sfruttata e/o l'exploit utilizzato
  - I riferimenti temporali dell'accaduto
  - Lo stato del traffico e/o dell'azione (e.g. bloccato/non bloccato/non noto)
- Classification: è la fase in cui l'analista dopo aver raccolto tutte le evidenze ed aver fatto una prima analisi dell'accaduto procede con la classificazione dell'evento in termini di categoria di

minaccia e di livello di gravità/pericolosità. L'assegnazione del livello di criticità ad un allarme dipende da diversi fattori, tra i quali ad esempio:

- La tipologia di allarme/ anomalia;
- La criticità puntuale dell'asset coinvolto, ove per asset si intende non solo un PC/Server ma anche un utente o casella di posta o dispositivo di rete.
- La frequenza dell'allarme stesso.

Si propone a titolo di esempio la seguente matrice:

INCIDENT PRIORITY LEVELS		IMPACT (Asset)		
		Low	Medium	High
SEVERITY (Attack)	Low	Low	Low	Medium
	Medium	Low	Medium	High
	High	Medium	High	High

Tabella 1 – Tabella di correlazione tra gravità incidenti e impatto sugli asset

INCIDENT PRIORITY	
Priority Levels	Descrizione
LOW	Gli incidenti non rappresentano un rischio immediato. Un workaround risolutivo è già disponibile o un piano di remediation è facilmente realizzabile con azioni basilari.
MEDIUM	L'incidente riguarda le attività classificate come a medio impatto. Gli incidenti presentano una discreta probabilità di provocare danni all'infrastruttura, soprattutto se le azioni di remediation non vengono implementate nel breve termine.
HIGH	Questo tipo di incidenti ha un'alta probabilità di causare, o ha già causato, una o più interruzioni dei servizi aziendali. La classificazione High solitamente riguarda gli incidenti su asset classificati come "business-critical".

Tabella 2 – Descrizione dei livelli di incidente

- Notification: è la fase di produzione dei deliverable previsti dal servizio ossia la fase in cui le informazioni estratte dalle piattaforme tecnologiche vengono normalizzate ed inserite in elementi di notifica.
- Tuning: fase di supporto operativo verso i gestori delle piattaforme tecnologiche deputate alla "Detection" attivata nel caso di tuning necessario sulle stesse per limitare o azzerare l'incidenza di falsi positivi e del conseguente "rumore" da essi generato.

Il servizio di TRIAGE (identification, classification, notification) ha l'obiettivo di facilitare la messa a punto dei falsi positivi e di segnalare all'Amministrazione le anomalie reali.

Processo di Analisi ed Incident Notification

Il processo di Incident Notification ha come obiettivo la rapida e corretta comunicazione agli attori interessati. Il processo alla base è lo standard previsto dall'incident management per le comunicazioni e le escalation. A tale proposito, nel corso della fase di avvio del servizio saranno identificate le opportune interfacce competenti per la ricezione delle notifiche in funzione della classe degli asset coinvolti e della criticità dell'incidente.

Reporting

Il servizio produce due tipologie di report:

- Executive Summary, un rapporto di sintesi destinato prevalentemente al management e al personale non tecnico per una comprensione immediata degli attacchi riscontrati. Si tratta di un elaborato in excel contenente tutti i dati relativi ai KPI di servizio.
- Technical Report una scheda incidente con tutte le indicazioni necessarie per la comprensione dei problemi riscontrati, per la loro classificazione in termini di severità e con un suggerimento relativo alle misure più idonee da adottare per la loro risoluzione. Tale rapporto fornirà il dettaglio delle principali vulnerabilità/minacce riscontrate.

1.3.3 Vulnerability Management ("security base services")

Il servizio sarà erogato in modalità continuativa da remoto e prevederà una fase di preparazione in funzione del perimetro target indentificato. La soluzione consente una gestione delle vulnerabilità basata sull'attribuzione di un indice di rischio in funzione della vulnerabilità rilevata, della risorsa impattata ed il ruolo aziendale dell'asset. Di seguito i punti di attenzione ed attività eseguite:

- Asset Discovery: rilevamento e classificazione di tutte le risorse conosciute e sconosciute che si collegano all'ambiente IT (cloud, applicazioni, gli endpoint e risorse mobili, container, etc..);
- Inventario delle risorse: crea un inventario aggiornato in tempo reale di tutte le risorse IT
- Rilevamento continuo delle vulnerabilità tramite un set opportuno di feature automatizzate:
  - misurazione e quantificazione del rischio informatico, con disamina delle vulnerabilità delle risorse e di gruppi di risorse, con identificazione delle misure pratiche che riducano l'esposizione e migliorino l'efficacia remediation plan;
  - valutazione, segnalazione e monitoraggi degli errori di configurazione basandosi su benchmark interni al tool di scansione e issue cases trattati;
  - valutazione dei certificati digitali interni ed esterni ed analisi delle configurazioni TLS per rilevamento di problemi e vulnerabilità
  - patch detection: correlazione delle vulnerabilità rilevate e remediation plan con indicazione di specifiche patch da applicare su host specifici, con diminuzione di tempi di risposta ed opportune azioni di mitigazione del problema;
- Le attività oggetto di test saranno eseguite a valle della formalizzazione dei seguenti documenti:
  - Legal Agreement (Manleva): Un accordo stabilito tra le parti che autorizza il Security Assessment Team a svolgere le attività specifica e che lo scarica da responsabilità per eventuali danni o disservizi creati.

- Regole di Ingaggio: Documento che contiene indicazione di inizio e durata delle singole fasi, le finestre orario in cui verranno erogate le attività, l'elenco dei deliverable, l'assegnazione dei ruoli e delle responsabilità per il fornitore e l'Amministrazione e il perimetro oggetto di analisi.

Tali documenti costituiscono perimetro e modalità di esecuzione dei test e devono essere sottoposti ad accettazione e firma dal cliente, in mancanza delle quali non sarà possibile procedere all'esecuzione dei test.

Al completamento delle stesse saranno consegnati i seguenti deliverable denominati:

- VA Results Executive Summary: Il report contiene una overview di tipo executive ad alto livello delle vulnerabilità individuate, ordinate per livello di rischio;
- VA Results Technical Report: Il report contiene i dettagli delle vulnerabilità segnalate, ordinate per criticità (utilizzando il sistema CVSS), incluse gli entry-point e le contromisure suggerite.

I deliverable, in base alla complessità del perimetro, possono far parte di un unico documento di report. Il servizio è limitato all'analisi di una quantità massima di 5000 host, nell'ambito della componente migrata ed on premise.

1.3.4 Dynamic Application Security Testing ("security base services")

Il servizio sarà erogato in modalità a task da remoto e consente l'identificazione delle vulnerabilità all'interno delle applicazioni Web e l'analisi dell'esposizione al rischio di attacchi informatici ai Sistemi Informativi mediante l'utilizzo di tecniche di analisi dinamica.

L'attività ha lo scopo di rilevare e gestire le vulnerabilità applicative che insistono sui sistemi informativi in ambiente WEB di pre-collaudo/pre-produzione e loro relative classificazione e prioritizzazione.

Il servizio prevede l'esecuzione dei test dinamici di sicurezza per le applicazioni per la verifica delle vulnerabilità tenendo conto dell'esposizione e dell'ambiente operativo in cui l'applicazione è in esecuzione. L'input è rappresentato dalle informazioni relative ai target da analizzare e le relative modalità attuative che dovranno essere concordate con l'Amministrazione.

L'analisi comprenderà almeno i seguenti ambiti:

- Configurazione (es. directory traversing);
- Autenticazione (cifratura degli accessi, password policy, dictionary attack);
- Autorizzazione (Privilege escalation);
- Input Validation.

A seguito delle scansioni effettuate sarà prodotto un report indicante le vulnerabilità individuate e la relativa classificazione.

Il report costituirà il Detailed Software Security Assessment Report contenente i dettagli tecnici del livello di sicurezza dell'istanza a run-time applicazione:

- Riferimenti ai tipi di attacco e vulnerabilità;

- Vulnerabilità/rischi identificati e la gravità di ognuno in termini di potenziale impatto sul sistema software oggetto dell'analisi;
- Notazioni e classificazione dei bugs sulla sicurezza secondo gli standard applicabili.

Le attività oggetto di test saranno eseguite a valle della formalizzazione dei documenti riportati sotto.

- Legal Agreement (Manleva): Un accordo stabilito tra le parti che autorizza il Security Assessment Team a svolgere le attività specifica e che lo scarica da responsabilità per eventuali danni o disservizi creati.
- Regole di Ingaggio: Documento che contiene indicazione di inizio e durata delle singole fasi, le finestre orario in cui verranno erogate le attività, l'elenco dei deliverable, l'assegnazione dei ruoli e delle responsabilità per il fornitore e l'Amministrazione e il perimetro oggetto di analisi.

Tali documenti costituiscono perimetro e modalità di esecuzione dei test e devono essere sottoposti ad accettazione e firma dal cliente, in mancanza delle quali non sarà possibile procedere all'esecuzione dei test.

1.3.5 Servizio di supporto per attività di Security Device Management (Protezione Perimetrale) ("security base services")

Il servizio professionale richiesto è orientato a supportare l'Amministrazione nella gestione e nel monitoraggio continuativo delle piattaforme di protezione perimetrale previste sulla nuova infrastruttura ICT. Il servizio è erogato "as a service" remotamente ed include la gestione degli apparati di protezione perimetrale con una finestra di servizio H8x5 e prevede:

- Definizione del perimetro di servizio: definizione della baseline dei sistemi di sicurezza che saranno oggetto del servizio Security Device Management;
- Definizione delle politiche di sicurezza: un'analisi globale dell'infrastruttura dei sistemi di sicurezza oggetto del servizio; lo scopo è quello di analizzare l'as-is della configurazione dei firewall, delle policy già configurare e dell'architettura complessiva nella quale i firewall sono posizionati
- Presa in carico dei sistemi, in RW, nello specifico le attività di presa in carico prevedono:
  - pianificazione temporale delle attività;
  - completa raggiungibilità dei devices e delle relative piattaforme di management ove presenti;
  - configurazione di utenze nominali per gli specialisti del SOC;
- Gestione a regime:
  - ogni richiesta viene validata ed implementata secondo le best practice di sicurezza ed in conformità a quanto definito con il Cliente in relazione anche alle policy aziendali vigenti.
  - i change, ad esempio, possono riguardare aggiunta/rimozione/modifica di policy firewall, creazioni tunnel vpn, modifica routing, /creazione/modifica profili UTM etc

1.3.6 Sicurezza hosts - servizio di Managed Detection & Response ("security base services")

Il servizio è erogato remotamente ed include le licenze e la gestione degli Endpoint, la cui distribuzione ed installazione è da attivare a carico del MASE e quindi non oggetto del presente servizio.

La gestione centralizzata della soluzione viene fatta attraverso una piattaforma di management presente su cloud. Tale piattaforma di fatto raccoglie tutte le informazioni di telemetria (metadati) inoltrate dagli agent installati sugli endpoint dell'Amministrazione tramite opportuno collegamento Internet, di cui è richiesta la visibilità continuativa (tra agent e piattaforma di management) in carico all'infrastruttura di accesso Internet del MASE. Il servizio è erogato as a service ed include un monitoraggio continuativo con finestra di servizio H24 per 365 giorni con notifica degli eventi ritenuti di interesse.

Il modello di servizio consente di:

- Ridurre al minimo le possibili finestre d'esposizione a eventuali attacchi informatici per gli endpoint in perimetro (con agent installato);
- Remediation automatica (ove applicabile) per gli incident riconosciuti come "veri positivi" ed a criticità massima;
- Garantire la protezione degli endpoint anche in assenza momentanea di connessione ad internet;
- Isolare dalla rete endpoint compromessi conservandone il controllo dalla piattaforma in cloud internet;
- Proteggere in tempo reale il perimetro da attacchi sconosciuti e che non utilizzano metodologie e/o indicatori noti internet (limitatamente alle caratteristiche della soluzione tecnologica impiegata).

1.3.7 Compliance Assessment Framework Nazionale Cyber Security (FNSC) ed eventuali attività successive di definizione di procedure e processi ("security base services")

L'esigenza dell'Amministrazione è quella di avvalersi di servizi professionali finalizzati alla fornitura del supporto specialistico per lo svolgimento di un assessment, basato su Framework Nazionale Cyber Security, in modo da:

- indicare il grado di adeguamento dell'Amministrazione ai livelli standard di sicurezza;
- individuare le possibili azioni correttive e soluzioni rispetto agli standard vigenti nell'organizzazione.

Il presente servizio è finalizzato ad eseguire un assessment per indicare la copertura e maturità dei controlli di sicurezza inerenti ai seguenti ambiti:

- Governance, che indirizza l'insieme delle pratiche volte a definire le politiche e l'organizzazione necessarie per poter reagire e prevenire, in maniera efficace, alle minacce di sicurezza, in modo da minimizzare l'impatto di possibili danni alle finalità istituzionali dell'Amministrazione dovuti ad incidenti di sicurezza di natura informatica.
- Prevent, che esprime la capacità di attuare pratiche e misure di sicurezza per la protezione delle informazioni, delle infrastrutture e dei servizi digitali presso l'Amministrazione.
- Detect, che esprime la capacità di individuare tempestivamente potenziali violazioni o eventi che possono influenzare o compromettere la sicurezza dell'Amministrazione.
- Respond & Recovery, che esprime la capacità di rispondere efficacemente ad un incidente di sicurezza e possibilmente di ripristinare i servizi impattati dallo stesso.

Nella definizione dei suddetti controlli si prenderanno in considerazione i requisiti previsti dal Framework Nazionale sulla Cyber Security (FNCS) e dalla normativa europea sul General Data Protection Regulation (GDPR). Il perimetro delle attività descritte è da intendersi limitato alle attività del perimetro oggetto della migrazione sul PSN.

I servizi oggetto della presente fornitura hanno carattere consulenziale tecnico-organizzativa ed ogni decisione in merito alle soluzioni proposte è in capo al MASE.

Come deliverable del servizio è previsto il rilascio di un documento che conterrà i risultati dell'assessment e che sarà condiviso con i referenti dell'Amministrazione. Verrà prodotta inoltre una presentazione di sintesi sulle attività svolte.

Nello specifico possono essere identificate eventuali attività di consulenza erogate on-demand per la definizione di procedure e/o processi legati al contesto cyber security.

1.3.8 Metodologia di Valutazione dei Maturity Level ("security base services")

L'esigenza dell'Amministrazione è quella di avvalersi di servizi professionali finalizzati alla produzione di un documento per la valutazione dei Maturity Level delle Misure di Sicurezza. Consente la rilevazione del livello di della copertura e maturità dei controlli di sicurezza inerenti i seguenti ambiti:

- Identify: che indirizza l'insieme delle pratiche volte a definire le politiche e l'organizzazione necessarie per poter reagire e prevenire, in maniera efficace, alle minacce di sicurezza, in modo da minimizzare l'impatto di possibili danni alle finalità istituzionali dell'Amministrazione dovuti ad incidenti di sicurezza di natura informatica.
- Protect: che esprime la capacità di attuare pratiche e misure di sicurezza per la protezione delle informazioni, delle infrastrutture e dei servizi digitali presso l'Amministrazione.
- Detect: che esprime la capacità di individuare tempestivamente potenziali violazioni o eventi che possono influenzare o compromettere la sicurezza dell'Amministrazione.
- Respond & Recovery: che esprime la capacità di rispondere efficacemente ad un incidente di sicurezza e possibilmente di ripristinare i servizi impattati dallo stesso.

Nella valutazione dei suddetti controlli si prenderanno in considerazione i requisiti previsti dal Framework Nazionale sulla Cyber Security (FNCS) e dalla normativa europea sul General Data Protection Regulation (GDPR). L'attività si svolge secondo il seguente schema:

- Avvio del progetto e pianificazione: In avvio di progetto si procederà alla composizione di un gruppo di lavoro misto formato da personale specialistico e personale dell'Amministrazione, al fine di individuare le figure da coinvolgere nell'esecuzione dell'assessment e procedere ad una pianificazione di dettaglio (sotto-fasi 1 e 2, cfr. Figura 133). Dette figure dovranno essere in grado di fornire informazioni in ordine allo stato di implementazione dei controlli di sicurezza presi in considerazione e che potranno essere di tipo logico, fisico e organizzativo.
- Esecuzione dell'assessment: La presente attività consiste nella raccolta delle informazioni rilevate mediante gli incontri con i responsabili e all'analisi della documentazione ottenuta (sotto-fase 3).
- Elaborazione dei risultati dell'assessment: A seguito degli incontri/interviste verranno analizzati, gestiti ed elaborati i dati acquisiti, che avranno particolarmente valore e qualità, in quanto aggiornati allo stato dell'arte. Tutte le informazioni acquisite saranno riportate in un documento,

che conterrà i risultati dell'assessment e che sarà condiviso con i referenti dell'Amministrazione (sotto-fasi 4 e 5).

Di seguito un diagramma di sintesi delle fasi progettuali per l'assessment sui processi di sicurezza

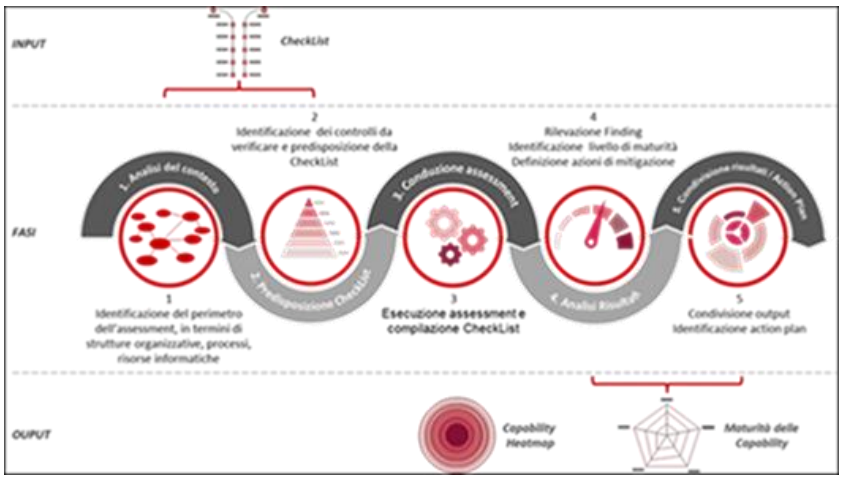
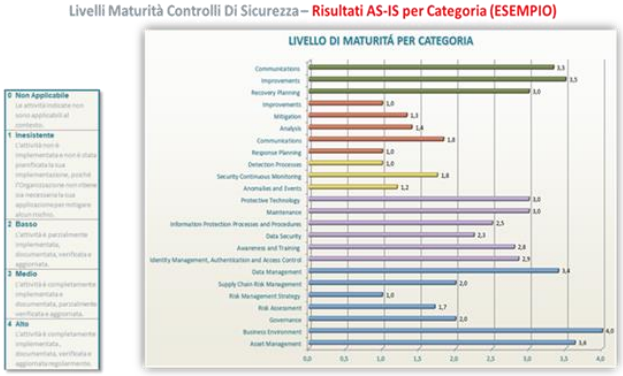


Figura 133: Sotto-fasi per l'assessment sui processi di sicurezza

I risultati saranno riportati secondo il seguente schema:



Tali risultati daranno seguito a un piano di trattamento necessario all'esecuzione delle azioni volte ad innalzare il livello di maturità laddove lo stesso risulti inferiore al target stabilito di concerto con l'organizzazione.

Tra gli interventi primari, di norma previsti dal piano di trattamento, vengono si citano i seguenti:

- Definizione di un Modello Organizzativo Cyber;
- Definizione e implementazione di una Policy di Cybersecurity;
- Procedure prioritarie:
  - Gestione e Classificazione delle informazioni;
  - Gestione degli Incidenti cyber;
  - Vulnerability Management.

Questi interventi di tipo consulenziale saranno propedeutici a:

- Introdurre o innalzare la consapevolezza della sicurezza dell'informazione all'interno dell'organizzazione.
- Iniziare un percorso strutturato di riorganizzazione dei processi inerenti alla sicurezza dell'informazione.
- Aiutare l'organizzazione a comprendere ed utilizzare in modo adeguato gli strumenti tecnologici salvaguardando la sicurezza dell'informazione.

Il perimetro delle attività descritte è da intendersi limitato al perimetro MASE. I servizi oggetto della presente fornitura hanno carattere consulenziale tecnico-organizzativa ed ogni decisione in merito alle soluzioni proposte è in capo all'azienda committente. Ove non specificato diversamente, il perimetro delle attività descritte è sempre da intendersi come limitato ai dati la cui titolarità ricada in capo all'Amministrazione contraente ed ai soggetti che hanno relazione con la stessa.

1.3.9 Cyber Strategic Risk Management ("security base services")

Il servizio di Cyber Strategic Risk Management mira a misurare la postura di sicurezza informatica di un'azienda in termini di risorse, tecnologie, processi e politiche attraverso una analisi statica di rischi cyber situazionali.

Questo servizio mette insieme diversi servizi (ad es. valutazione ed analisi del rischio strategico, valutazione dell'intelligence sulle minacce più verticali al contesto) in modo strutturato per valutare e mitigare i rischi specifici e con un approccio "data driven".

L'output finale del servizio è costituito da documenti che indicheranno chiaramente la situazione "AS IS" ed un piano strategico cyber con azioni evolutive misurato in modo chiaro da quanto emerso dall'analisi della situazione. I rischi, legati all'attuale contesto di cyber security ed al contesto industriale, saranno presi in considerazione ed analizzati in funzione del core business della PA, con una valutazione fruibile sia a livello di management che operativamente.

Il servizio si compone delle seguenti fasi:

- Fase 1 - Studio preliminare: in questa fase del servizio verrà eseguito uno studio sia sul settore industriale che sul dominio in termini di rischi. Questa fase è fondamentale per adattare correttamente le fasi successive.
- Fase 2 - Cyber Strategic Risk: questa è la fase principale, nel quale:
  - Definire il perimetro in cui opera l'azienda, sia in termini cyber che di business;
  - Definire una lista di controllo e interviste che il cliente deve compilare e che costituiranno l'input principale della valutazione strategica del rischio;
  - Preparare una valutazione delle minacce open source che definirà quali informazioni sono disponibili sull'azienda su Internet per definire meglio le minacce informatiche che potrebbero derivare da questo;
  - Creare un piano di mitigazione e rimedio su misura per ridurre il rischio complessivo a cui è esposta l'azienda;
  - Definire un Piano Strategico Cyber per il mantenimento e raggiungimento degli obiettivi di business.

Deliverables finale:

- documento pdf di un Piano Strategico Cyber;
- Consegna dei Report di dettaglio:
  - Strategic Risk Assessment & Analysis;
  - Report verticale di Threat Intelligence;

1.3.10 Security By Design in funzione della gap-analysis e dei controlli di sicurezza indirizzati alla protezione della rete, dei servizi e degli endpoint ("security base services")

Il servizio consiste nella predisposizione di un team di specialisti con le competenze e l'esperienza necessarie ad effettuare l'attività di supporto al design ed implementazione della protezione perimetrale oggetto di migrazione al fine di incrementarne il livello di sicurezza. Tale servizio, erogato durante la fase di setup della migrazione, prevede un'analisi preliminare volta a comprendere le tecnologie utilizzate e le specifiche caratteristiche al fine di poter predisporre le opportune linee guida o best practice in ambito security by design secondo una metodologia articolata in tre step di seguito descritti:

- Step 1 – Analisi preliminare: In questa fase verrà eseguita un'analisi preliminare dello scenario proposto, svolgendo le seguenti attività:
  - raccolta delle informazioni sulle tecnologie utilizzate dall'Amministrazione contraente;
  - analisi del contesto specifico e classificazione del rischio Cyber sulla base dei livelli di criticità dei servizi a cui sono associate le specifiche tecnologie in esame;
  - analisi degli impatti di indisponibilità dei servizi, per l'individuazione delle aree problematiche e contromisure tecnologiche da adottare.
- Step 2 – Disegno delle linee guida di security by design: In questa fase verranno identificate le linee guida di security by design, propedeutica alla fase di progettazione, svolgendo le seguenti attività:
  - predisposizione delle linee guida di security by design (sulla base della classificazione delle tecnologie fatta nella fase di assessment);
  - ipotesi di progettazione dell'infrastruttura sulla base delle contromisure suggerite;
  - condivisione della documentazione predisposta ai referenti coinvolti.
- Step 3 – Fase implementativa di Delivery e migrazione soluzioni esistenti: in questa fase verranno condotte le attività di predisposizione delle nuove soluzioni perimetrali WAF nonché il refining/migrazione dell'attuale soluzione perimetrale esistente NGFW sulla nuova infrastruttura Cloud.

I deliverable prodotti consistono in attività operative secondo quanto di seguito rappresentato in ottica di definire una mappatura tra servizi, tecnologie e contromisure.

1.3.11 Gestione degli Incidenti di Sicurezza e Crisis Management ("security base services")

I servizi professionali previsti sono finalizzati a supportare l'Amministrazione nella Gestione degli Incidenti di Sicurezza e Crisis Management previsti secondo due fasi: Tattica e Strategica. Il servizio in fase Tattica sarà erogato in modalità "a task" per incidente su richiesta del MASE e con valorizzazione da definire con l'Amministrazione sulla base dell'effettiva necessità. Nello specifico

verranno messe in atto tutte le attività necessarie all'identificazione delle dinamiche associate all'incidente e alla definizione di opportune azioni di contenimento e risposta alla minaccia.

Qualora l'evento si riveli impattante sugli obiettivi strategici, le funzioni vitali o la reputazione dell'Organizzazione, il servizio viene supportato dal Crisis Management che sarà in grado di far fronte a situazioni anomale, che esulano dalla gestione standard degli incidenti, organizzando le attività in modo da ricondurre tali eccezionalità all'interno di opportune best practice.

Su richiesta del MASE e come attività on-demand potranno essere attivati ulteriori servizi di supporto in Fase Strategica, in modo da definire i passi operativi, da programmare nel breve e medio periodo, relativi alla definizione delle attività necessarie ad innalzare le capacità difensive del perimetro oggetto della comprovata azione malevola. L'attività sarà svolta in modalità a task ed afferirà ad un basket di giornate rese disponibili dal presente contratto.

Di seguito sono riportate, a titolo esemplificativo e non esaustivo, le principali attività eseguite dal supporto richiesto:

- Isolamento dei sistemi compromessi dalla rete del Cliente.
- Indagini sull'entità e la tipologia degli eventi.
- Indicazione delle più efficienti modalità di sanitizzazione e ripristino dei sistemi coinvolti.
- Supporto nelle comunicazioni verso le Autorità competenti.
- Raccolta e sintesi delle Lesson Learned;
- Piano di rientro breve e medio periodo nell'ambito del programma di consolidamento delle misure di sicurezza intraprese dal MASE

Il servizio di risposta all'incidente comprende anche le attività legate al contenimento, all'eradicazione e al ripristino della normale operatività dei servizi del Cliente che hanno subito l'incidente, al fine di fornire tempestivamente, una prima risposta di contrasto alle problematiche di sicurezza riscontrate dall'Amministrazione.

Figura 134 - Incident Response & Response Process

Come deliverable del servizio è previsto il rilascio di un report di alto livello ed uno di natura tecnica, con cui ripercorrere e comprendere gli step legati all'incidente come di seguito rappresentato:

- Technical Report di dettaglio
- Executive Report di sintesi

1.3.12 Cyber Threat intelligence: Early Warning e Data Breach ("security base services")

Il servizio professionale richiesto è orientato a supportare l'Amministrazione nella erogazione di servizi di Cyber Threat Intelligence. Nello specifico i servizi verranno erogati attraverso una soluzione tecnologica che consente l'attivazione delle seguenti tipologie di servizio:

- Early Warning

- Data Breach Discovery

Tale capacità di cyber threat intelligence garantisce una migliore efficacia dei servizi di detection, grazie alla verifica continua su fonti aperte di eventuali nuove vulnerabilità e/o possibili vettori di attacco che possano impattare la nuova infrastruttura su Cloud.

- **Early Warning**

La componente di servizio Early Warning ha lo scopo di acquisire, da fonti aperte, elementi informativi tali da individuare nuove vulnerabilità applicative con l'obiettivo di segnalare proattivamente le vulnerabilità rilevate, al fine di prevenire attacchi informatici che possano sfruttare malware evoluti nonché zero-day che possono mettere a rischio le tecnologie in uso presso l'Amministrazione. Le principali caratteristiche e funzionalità offerte riguardano:

- integrazione di feed esterni di sicurezza, potenziando in tal modo le capacità native della piattaforma tecnologica proprietaria;
- monitoraggio in tempo reale delle fonti aperte per la ricerca di possibili nuove vulnerabilità informatiche non ancora note;
- motore di cross correlazione per analizzare le informazioni raccolte rispetto all'elenco delle tecnologie in esame;
- generazioni di report basati sulle configurazioni definite.

- **Data Breach Discovery**

La componente di servizio denominata Data Breach Discovery ha lo scopo di rilevare attività che mirano a trafugare dati e/o divulgare e rendere pubbliche informazioni da parte di soggetti non autorizzati, relativi a target di interesse, attraverso il monitoraggio continuativo della rete (surface e deep/dark web).

Le principali caratteristiche e funzionalità riguardano:

- controllo continuo in tempo reale di fonti aperte alla ricerca di elementi di interesse quali, ad esempio, indirizzi e-mail, documenti, nomi macchina ecc., citati o individuati all'interno di determinate aree della rete;
- generazione di allarmi in base alle evidenze derivanti dall'analisi dei dati;
- produzione di report basati sulle evidenze derivanti dall'analisi dei dati.

L'erogazione del servizio prevede una fase di avvio e una fase di esecuzione (gestione, Maintenance e miglioramento continuativo). All'emergere di informazioni di particolare rilevanza per l'Amministrazione sarà redatto un Intelligence Alert ed un Intelligence Report con il quale il sistema di Cyber Threat intelligence segnalerà tempestivamente e con il dovuto dettaglio le evidenze rilevate e suggerirà eventuali misure da adottare per la risoluzione della problematica. Le due tipologie di deliverable conterranno informazioni relativamente a:

- categoria di interesse;
- severità;
- data di rilevazione;
- Traffic Light Protocol – TLP;
- dettaglio degli elementi raccolti e dei risultati delle analisi effettuate;

- indicazioni di eventuali raccomandazioni da porre in essere per la risoluzione degli opportuni «case».

Durante la fase di avvio sarà attivata un'istanza dedicata all'ambito MASE sul sistema di Cyber Threat Intelligence, sulla quale verranno eseguiti i processi di analisi che realizzano il servizio specifico. Inoltre, verrà abilitato un portale Web dedicato (TIS Disclosure) ad accesso sicuro sul quale saranno rese disponibili le interfacce per la consultazione dei risultati individuati.

1.3.13 Cyber Threat intelligence: Brand Abuse, Anti-phishing con Site takedown ("security base services")

Il servizio professionale richiesto è orientato a supportare l'Amministrazione, nella erogazione di servizi di Cyber Threat Intelligence, richiesti dal MASE. Nello specifico i servizi verranno erogati attraverso una soluzione tecnologia che consente l'attivazione delle seguenti tipologie di servizio:

- Brand Abuse
- Anti-phishing con Site takedown

Il servizio è inoltre orientato a fornire conoscenza riguardo attori malevoli operanti in contesti illeciti piuttosto che ad evidenziare elementi di contesto potenzialmente utili a prevenire o mitigare azioni informatiche malevole mirate.

- **Brand Abuse**

La componente di servizio Brand Abuse ha lo scopo di acquisire, da fonti aperte, elementi informativi tali da individuare contenuti illegittimi legati all'Amministrazione, come ad esempio il logo o il marchio presenti in siti web o URL di terze parti, individuati attraverso il monitoraggio dei typosquatted domain. Le principali caratteristiche e funzionalità offerte riguardano:

  - Identificazione di possibili attacchi di dirottamento del traffico dati relativi all'utilizzo indebito di domini web afferenti al brand o sue variazioni (typosquatting);
  - Analisi in tempo reale dell'utilizzo improprio del brand o senza diretta autorizzazione dell'Amministrazione;
  - Rilevazione di contenuti "typosquatted" per l'individuazione di domini o contenuti su Social Media non autorizzati e pubblicati con esplicito riferimento al brand (tipicamente finalizzati a perpetrare tentativi di frode);
  - monitoraggio di potenziali siti web fraudolenti relativi alla propria supply chain tramite set di keywords di interesse;
  - generazioni di report basati sulle configurazioni definite.

- **Anti-phishing con Site takedown**

La componente di servizio di Anti-Phishing con site takedown (max qty 5/anno) ha lo scopo di generare allarmi a fronte dell'utilizzo improprio dei domini web dell'Amministrazione per perpetrare attività di phishing fraudolente. Il servizio abbraccia in generale i seguenti ambiti:

  - Phishing: "rilevazione" di domini e siti web fraudolenti creati appositamente per simulare ed impersonare una organizzazione a fini malevoli;

- Rilevazione di contenuti legati all'organizzazione, come ad esempio il logo o il marchio, presenti in siti web o URL di terze parti, individuati attraverso il monitoraggio dei typosquatted domain, senza l'autorizzazione dell'organizzazione stessa;
- motore di cross correlazione per analizzare le informazioni raccolte rispetto all'elenco delle tecnologie in esame;
- generazioni di report basati sulle configurazioni definite.

Il servizio prevede l'identificazione preliminare dei contenuti legittimi dell'organizzazione da proteggere attraverso la stesura di una scheda informativa contenente i seguenti dati:

- nome del marchio registrato;
- logo ufficiale;
- domini legittimi;
- profili e pagine legittime sui social media network (se presenti);
- certificato con dettagli del copyright e/o trademark registrato;
- pagine clonate e malware mail pervenute solo da fonti OSint, se note;
- individuazione dei phish-kit utilizzati per gli attacchi, se noti
- elenco terze parti costituenti la supply chain.

Il contenuto di tale scheda sarà utilizzato per il monitoraggio del web e del dark-web, come aspetto proattivo del servizio, ossia l'identificazione preventiva di possibili contenuti fraudolenti non ancora inseriti o adottati dagli attaccanti in campagne malevole.

L'erogazione dei servizi Brand Abuse ed Anti-Phishing con takedown è costituita da una fase di avvio e una fase di esecuzione (gestione, maintenance e miglioramento continuo). All'emergere di informazioni di particolare rilevanza per l'Amministrazione sarà redatto un Intelligence Alert ed un Intelligence Report con il quale il servizio segnalerà tempestivamente e con il dovuto dettaglio le evidenze rilevate e suggerirà eventuali misure da adottare per la risoluzione della problematica. Le due tipologie di deliverable conterranno informazioni relativamente a:

- categoria di interesse;
- severità;
- data di rilevazione
- Traffic Light Protocol - TLP

Il servizio è organizzato in modo da effettuare un monitoraggio continuo di nuovi domini di tipo typo-squatted in maniera proattiva attraverso l'identificazione di uno o più domini di nuova costituzione e che possono essere considerati di tipo typo-squatted rispetto a quelli legittimi. Pertanto, verrà prevista una fase di analisi preliminare del contenuto in esame e, qualora la verifica confermasse da subito l'effettiva pericolosità del contenuto, si procederà, previa autorizzazione dell'Amministrazione, alla stesura di un report dettagliato da utilizzare per l'avvio della fase finale del processo, ovvero l'azione di takedown.

Se la fase preliminare di analisi non fornisse risultati esaustivi oppure determinasse che il contenuto non è ancora «armato» (ovvero non rappresenta una minaccia diretta per l'Amministrazione) si procederà alla fase di monitoraggio; questa fase avrà l'obiettivo di tenere sotto controllo i contenuti

e la loro potenziale pericolosità per avviare, nel momento in cui essi risultino effettivamente armati e previa autorizzazione dell'Amministrazione, alla fase finale del processo, ovvero la stesura del report dettagliato e l'avvio dell'azione di takedown vera e propria.

Riassumendo vengono comprese all'interno del servizio le attività di:

- blocco dell'accesso ai siti di Phishing, attraverso il takedown dei domini typosquatted malevoli;
- segnalazione del dominio malevolo all'ente responsabile dell'hosting del sito di phishing;
- generazione di allarmi e report
- Il takedown dei siti è effettuato contattando diverse terze parti coinvolte nell'esposizione del sito di phishing, principalmente l'hosting provider e il registrar del dominio.

Contestualmente alla fase di avvio sarà attivata un'istanza dedicata, sulla quale verranno eseguiti i processi di analisi, che realizzeranno il servizio specifico, ed abilitato un portale Web dedicato (TIS Disclosure) ad accesso sicuro sul quale saranno rese disponibili le interfacce per la consultazione dei risultati individuati (report).

1.3.14 Web Application Penetration Testing ("security advanced services")

Il presente paragrafo descrive il servizio professionale di Web Application Penetration Testing che sarà svolto operativamente da remoto One Shot. Le attività si compongono da un insieme di test manuali ed automatici, volti ad effettuare tentativi di intrusione sui sistemi Web in scope e delle applicazioni concordate. Su richiesta dell'Amministrazione vengono previste attività di test anche sfruttando le vulnerabilità emerse dal servizio di Vulnerability Management o Dynamic Application Security Testing. Le attività oggetto di test saranno eseguite a valle della formalizzazione dei documenti riportati sotto.

- Legal Agreement (Manleva): Un accordo stabilito tra le parti che autorizza il Security Assessment Team a svolgere le attività specifica e che lo scarica da responsabilità per eventuali danni o disservizi creati.
- Regole di Ingaggio: Documento che contiene indicazione di inizio e durata delle singole fasi, le finestre orario in cui verranno erogate le attività, l'elenco dei deliverable, l'assegnazione dei ruoli e delle responsabilità per il fornitore ed Ente e il perimetro oggetto di analisi.

Tali documenti costituiscono perimetro e modalità di esecuzione dei test e devono essere sottoposti ad accettazione e firma dal cliente, in mancanza delle quali non sarà possibile procedere all'esecuzione dei test.

Operativamente, sono previste le seguenti attività:

- Tentativi d'intrusione sui sistemi WEB sfruttando le vulnerabilità identificate in eventuali attività precedentemente svolte (VM, DAST);
- Tentativi di escalation dei privilegi, nel caso l'accesso ottenuto non fornisca privilegi amministrativi;
- In caso di penetrazione in un sistema, produzione delle relative evidenze al fine di dimostrare l'intrusione effettuata;
- Descrizione dei rischi esistenti relativi alle possibilità di accesso non autorizzato ai suddetti sistemi.

Le attività sono condotte applicando metodologie globalmente riconosciute come standard de-facto per la conduzione di attività di penetration test, e in particolare le metodologie OSSTMM (Open Source Security Testing Methodology Manual) e OWASP (The Open Web Application Security Project) che definiscono le modalità per la conduzione di test completi, accurati, ripetibili e verificabili. Il test è eseguito per la ricerca di vulnerabilità applicative ed in base alle tecnologie utilizzate dalle applicazioni, consentirà l'identificazione di tutte le categorie di vulnerabilità top 10 OWASP.

Al termine delle attività verrà prodotto un documento denominato PT Results Technical Report che conterrà il report di dettaglio delle attività eseguite durante la fase di testing e le evidenze degli attacchi e delle eventuali compromissioni rilevate.

1.3.15 Security Orchestration, Automation and Response ("security advanced services")

I servizi professionali previsti sono finalizzati a supportare l'Amministrazione nell'attuazione di una soluzione di Security Orchestration, Automation and Response (SOAR) che sarà messa in campo con lo scopo di fornire funzionalità e capacità di orchestrazione ed automazione, mediate issue cases contestualizzate in casi specifici simili a quelli necessari al MASE. La soluzione tecnologica assume un ruolo centrale nell'erogazione di tale servizio ed ha lo scopo di:

- esaltare tutto il know-how tecnico per implementare workflow automatici che, sfruttando le integrazioni tecnologiche, permettano la gestione di use case in ambito IT Security;
- implementare una componente di dashboarding e di reporting che permettano di rendere visibili i benefici delle automazioni fornite ai clienti e le statistiche ad esse legate.

I principali benefici dei punti di forza derivanti dall'attivazione dello stesso sono:

- incrementare il proprio livello di maturità nell'ambito dell'information security, con l'adozione di processi di SOC automation;
- ridurre i tempi di risposta agli incidenti automatizzando la remediation e/o il containment;
- aumentare il livello di integrazione delle tecnologie in uso per massimizzarne il ritorno economico e operativo;
- migliorare la gestione degli incidenti di sicurezza con l'implementazione di processi automatici frutto delle competenze del Fornitore;
- accentrare la visibilità di tutte le informazioni legate all'operatività quotidiana in ambito IT security e non solo.

Il servizio viene erogato in due fasi distinte:

- Avvio del servizio, dove vengono raccolte tutte le informazioni di contesto utili alla definizione dei requisiti specifici di automazione ed utili alla creazione della baseline iniziale di use case da implementare. Durante la fase di avvio saranno definite anche le componenti architetturali della soluzione SOAR, si procederà al disegno dei meccanismi di automazione offerti dalla piattaforma, alle attività di integrazione delle varie piattaforme di terze parti con la relativa configurazione e, infine, al test di collaudo per passare alla messa in produzione e rendere la piattaforma pronta all'utilizzo durante la fase di conduzione;
- Conduzione del servizio, dove vengono:

- manutenuti e ottimizzati i playbook esistenti in ottica di miglioramento continuo, per garantire azioni di enrichment, dissemination, containment e remediation più puntuali;
- manutenuta ed ottimizzata la configurazione generale della piattaforma in termini di dashboarding e reporting;
- implementati nuovi playbook e nuovi use case che possono derivare sia da attività di gestione degli incidenti sia da service review.

Dal punto di vista operativo l'erogazione del servizio consentirà di:

- mitigare tentativi d'intrusione sui sistemi WEB sfruttando le vulnerabilità identificate in eventuali attività precedentemente effettuate (VM, DAST);
- attivare use case di interesse per la Committente;
- analizzare tipologie di incidenti che il SOAR deve gestire;
- provvedere ad effettuare integrazioni tecnologiche necessarie allo svolgimento delle attività previste;
- definire gruppi di lavoro coinvolti nei workflow da implementare;
- consentire gli accessi allo strumento in termini di ruoli e permessi;
- consentire la raccolta dei metadati che lo strumento deve indicizzare ed elaborare
- produrre reportistica necessaria;
- raccogliere contenuti da mostrare attraverso le dashboard dello strumento;
- data retention policy.

Il servizio produce due tipologie di report:

- Executive Summary, un rapporto di sintesi destinato prevalentemente al management e al personale non tecnico per una comprensione immediata degli attacchi riscontrati ed automaticamente risolti.
- Technical Report una scheda incidente con tutte le indicazioni necessarie per la comprensione dei problemi riscontrati, per la loro classificazione in termini di severità e con un suggerimento relativo alle misure più idonee da adottare per la loro risoluzione.

1.3.16 Security Policy review/advisory ("security advanced services")

Il servizio di Policy Review consiste in attività di analisi dei flussi perimetrali (inbound ed outbound) per una valutazione rapida ed efficace dello stato di conformità di un singolo firewall o di architetture complesse di firewall attuando una maggiore coerenza delle politiche implementate nel rispetto delle modalità organizzative oggi adottata dal MASE.

Questo servizio è attualmente conforme allo standard PCI-DSS e al framework NIST e si riassume nelle seguenti macro-attività:

- Rilevazione delle problematiche di sicurezza e conformità ai requisiti;
- Ottimizzazione delle regole e configurazione del firewall, con relativi flussi di traffico

Nel contesto specifico dell'Amministrazione, il servizio proposto si configura come un'attività di analisi-statica e dove possibile dinamica, eseguita sui flussi di traffico che attraversano i firewall oggetto di verifica.

Il servizio prevede le seguenti fasi operative:

- Definizione della baseline
- Information Gathering
- Definizione del modello di riferimento
- Analisi dei flussi

Il processo prevede che, con l'analisi dei log di traffico, si possano individuare elementi di miglioramento nelle configurazioni agendo sui seguenti parametri: service, source e destination. Verranno cioè individuate restrizioni dei parametri indicati sulla base di quello che effettivamente si rileva dai log, eliminando permission non necessarie (es. individuazione degli effettivi host source/ host destination/services utilizzati realmente nelle policy analizzate).

Questo processo sarà strutturato sulla falsa riga del servizio generale sopra descritto, con fasi di raccolta informazioni, analisi, reporting, ottimizzazione e reportistica finale a conclusione dell'attività.

Per il presente servizio è necessario che gli apparati di sicurezza perimetrale a cui afferiscono i servizi del cliente siano sempre raggiungibili e che vengano fornite al personale dedicato le credenziali di accesso.

Il cliente dovrà autorizzare l'invio dei log dai firewall ai sistemi di analisi presenti presso il centro servizi.

Come deliverable del servizio è previsto il rilascio di un report di analisi che riporti tutte le indicazioni riguardo le ottimizzazioni apportate sugli apparati di sicurezza oggetto di analisi.

- Technical Report di dettaglio
- Executive Report di sintesi

1.3.17 Threat Hunting as a service ("security advanced services")

Il presente paragrafo descrive il servizio professionale di Threat Hunting che sarà svolto operativamente da remoto con una figura specializzata in varie sessioni di hunting le cui attività saranno concordate operativamente negli obiettivi con il cliente.

Il servizio si compone di un insieme di analisi manuali ed automatiche, supportate dall'intelligence, volte alla ricerca proattiva di tecniche, tattiche e procedure (TTP) riconducibili ad attività malevole, sospettose e rischiose che potrebbero avere eluso gli strumenti di sicurezza esistenti.

Il servizio prevede l'utilizzo degli strumenti erogati e presenti presso l'Amministrazione purché raggiungibili da remoto tramite collegamento opportuno (VPN). Su richiesta, le sessioni di hunting potranno essere indirizzate alla ricerca di TTP riconducibili a tecniche definite nel framework Att&ck Mitre, o di TTP associati a incidenti intercorsi o di attori oggetto di report di threat intelligence.

Operativamente, saranno previste le seguenti attività:

- Definizione con l'Amministrazione degli obiettivi della sessione di Hunting (TTP, Attori, Incidenti pregressi)

- Definizione dell'ipotesi di dettaglio da verificare. Analizzando le TTP selezionato si effettua un elenco di indicatori da verificare, e si associa ogni TTP alle diverse fonti informative a disposizione e al tipo di evidenza che in esse un attaccante può aver generato.
- Analisi ed elaborazione dei dati. Si procedere a verificare la presenza delle fonti dati ipotizzate, e ove presenti a verificare se queste contengono le evidenze ricercate.
- Stesura della reportistica, che contiene un Executive Summary e un dettaglio tecnico di tutte i TTP verificati, le fonti dati usate e quelle eventuali mancanti, possibili riscontri ed evidenze di compromissione, indicazioni per un miglioramento della security posture (ad esempio di abilitare alcuni log, nuove regole di correlazione possibili, etc.)

Ci si riserva la possibilità di utilizzare ulteriori strumenti alternativi.

Per il servizio è previsto il rilascio di un report di Threat Hunting, che rispecchia i seguenti contenuti:

- Executive Summary
- Elenco delle tecniche, tattiche e procedure (TTP) oggetto di hunting
- Eventuali problemi di visibilità o assenza di telemetriche riscontrate rispetto al singolo TTP
- Eventuali riscontri osservati nell'infrastruttura
- Ove possibile, suggerimento di tecniche o regole di detection a colmare i problemi di visibilità

1.3.18 Consulting per revisione policy e procedure di cyber security ("security advanced services")

L'esigenza dell'Amministrazione è quella di avvalersi di servizi professionali finalizzati alla fornitura del supporto specialistico per lo svolgimento di attività di revisioni di policy e procedure di governance/sicurezza in ottica Cyber Security e in ottemperanza con i requisiti previsti dal Framework Nazionale sulla Cyber Security (FNCS) e dalla normativa europea sul General Data Protection Regulation (GDPR).

Il perimetro delle attività descritte è da intendersi limitato alle attività del perimetro oggetto della migrazione sul PSN.

I servizi oggetto della presente fornitura hanno carattere consulenziale tecnico-organizzativa ed ogni decisione in merito alle soluzioni proposte è in capo al MASE.

Come deliverable del servizio è previsto il rilascio di un documento che conterrà i risultati dell'assessment e che sarà condiviso con i referenti dell'Amministrazione. Verrà prodotta inoltre una presentazione di sintesi sulle attività svolte.

1.3.19 Cyber Threat intelligence: Pre Planned Attack + Black Market Monitor ("security advanced services")

servizio professionale richiesto è orientato a supportare l'Amministrazione, nella erogazione di servizi di Cyber Threat Intelligence, richiesti dal MASE. Nello specifico i servizi verranno erogati attraverso una soluzione tecnologica che consente l'attivazione delle seguenti tipologie di servizio:

- Pre Planned Attack
- Black Market Monitoring

- **Pre Planned Attack**

Il servizio ha lo scopo di raccogliere e correlare in tempo reale informazioni di tipo OSINT da molteplici fonti aperte (Surface, Deep e Dark Web), al fine di individuare eventuali "segnali deboli" indici di potenziali minacce che possano trasformarsi in attacchi informatici. Tale approccio consente di individuare eventuali "segnali deboli" in aree di interesse del cliente che siano indici di potenziali minacce e possano trasformarsi in attacchi informatici. Questo consentedi effettuare una analisi di contesto dell'ambito di interesse del Cliente (settori di business, tecnologie di riferimento, area geografiche e relativo contesto geo-politico, esposizione mediatica ecc.). Ad esempio sono rilevati e identificati:

- scenari di attacco cibernetico in corso nel mondo o nel paese che possono in qualche modo interessare il cliente per natura, area di mercato, area geografica o altro
- agenti di minaccia che hanno in corso o hanno avuto sospetti di attacco verso infrastrutture di interesse del cliente, per natura, area di mercato, area geografica o altro.
- Agenti di minaccia o attacchi specificatamente mirati al cliente stesso, anche di natura non cibernetica, se presenti (es. minacce fisiche ad asset o persone apicali del gruppo)
- Informazioni relative a eventi naturali, sociali e politici, se in zone/aree geografiche di interesse del cliente.

Le principali caratteristiche del servizio di Pre-Planned Attack sono:

- Il monitoraggio continuo delle fonti aperte alla ricerca di possibili eventi ed elementi che possano riguardare direttamente o indirettamente il Cliente;
- Analisi di contesto dell'ambito di interesse del Cliente (settori di business, tecnologie di riferimento, area geografiche e relativo contesto geo-politico, esposizione mediatica, eventuali minacce di natura non cibernetica ecc.);
- La creazione di una base dati continuamente aggiornata che consenta, attraverso algoritmi di Machine Learning e strumenti di visual link analysis, di correlare le informazioni raccolte con il perimetro tecnologico ed operativo del Cliente;
- Capacità di identificazione e protezione su aspetti relativi domini IP;
- Generazione di allarmi e report

- **Black Market Monitoring**

Il servizio consente l'analisi in tempo reale di grandi quantità di informazioni da fonti aperte (compresi Dark e Deep Web) per identificare nuovi "black-market" liberamente accessibili e relative frodi al loro interno al fine di rilevare prontamente attività illegali su argomenti selezionati realizzati nei mercati neri. Questo consente di individuare tempestivamente attività illecite su temi di interesse specifici perpetrate nell'ambito dei black market sul dark e deep web, ove tali ambiti siano liberamente accessibili. Attraverso tool di analisi semantica c'è la possibilità di riconoscere sui black market analizzati luoghi, numeri di carte di credito, targhe di veicoli o altre entità di interesse riconducibili ad attività illecite (entity extraction).

Le informazioni che il servizio è in grado di monitorare sono di seguito elencate[1]:

[1] Durante la fase di avvio del servizio saranno concordate con il cliente le informazioni necessarie per il riconoscimento e la classificazione di dette informazioni

- Riconoscimento, attraverso analisi sintattiche/semantiche e nei mercati neri monitorati, delle posizioni, carte di credito, targhe, entità legate a frodi e attività illegali.
- E-mail Address, numeri di carte di credito/debito all'interno del market monitoring, numeri di telefono, numeri di conto corrente, credenziali di accesso a siti web;
- Analisi relative a carte di credito, conti bancari e"riciclaggio di denaro", per identificare frodi nell'ambito del crimine informatico.
- Ricerca continua dei mercati neri e delle informazioni pertinenti al contesto specifico del Cliente
- Generazione di allarmi e report