

Infrastruttura BDP del CNAF

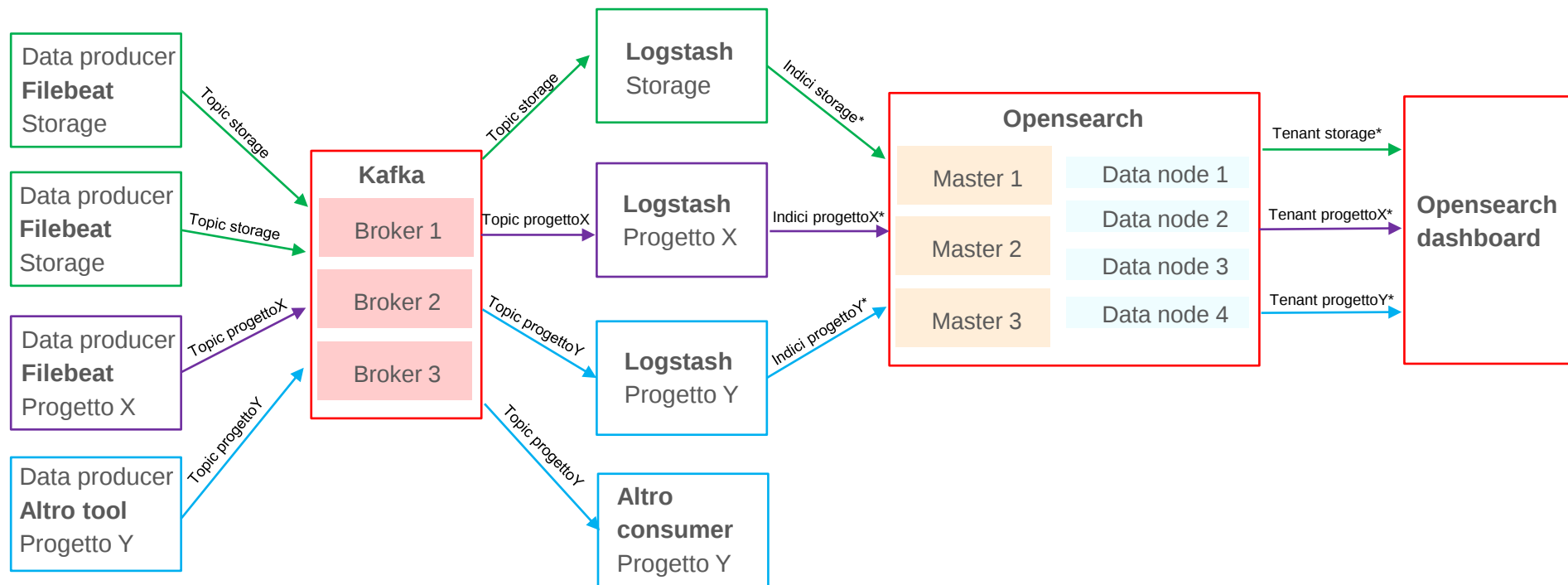
Enrico Fattibene

Censimento ML@CNAF
21 dicembre 2022

Introduzione

- Infrastruttura di gestione e analisi dati eterogenei a servizio di amministratori ed utenti del CNAF
- A partire dal lavoro già fatto al CNAF su Big Data Platform
 - Esperienza del gruppo di lavoro su BDP
 - Esperienza del Sistema Informativo
- Attività nell'ambito del progetto CNAF Reloaded
 - Task Operations
 - Infrastruttura di base general purpose + caso d'uso analisi dei log (ELK o simile)

Schema infrastruttura



- Segregazione dati per reparto / progetto
 - Ognuno può scrivere solo sul proprio topic Kafka e accedere via Dashboard proprio tenant Opensearch
 - Un utente può avere visibilità su più tenant
- Comunicazione via SSL

Stato e prossimi passi

- Tutti i servizi sono installati e running
 - Al momento unico Logstash (sufficiente per fase iniziale di produzione)
- Dati di diversi host Storage da fine novembre
 - Dashboard Storage importate da infrastruttura attuale (all in one)
- Studio di policy Opensearch
 - Per retention dei dati e quote
- In fase di studio integrazione con Wazuh
- Da fare
 - Ingest di altri dati (farming, servnaz, etc.)
 - Documentazione per utenti e amministratori
 - Puppettizzazione dei vari componenti
- In produzione a inizio 2023