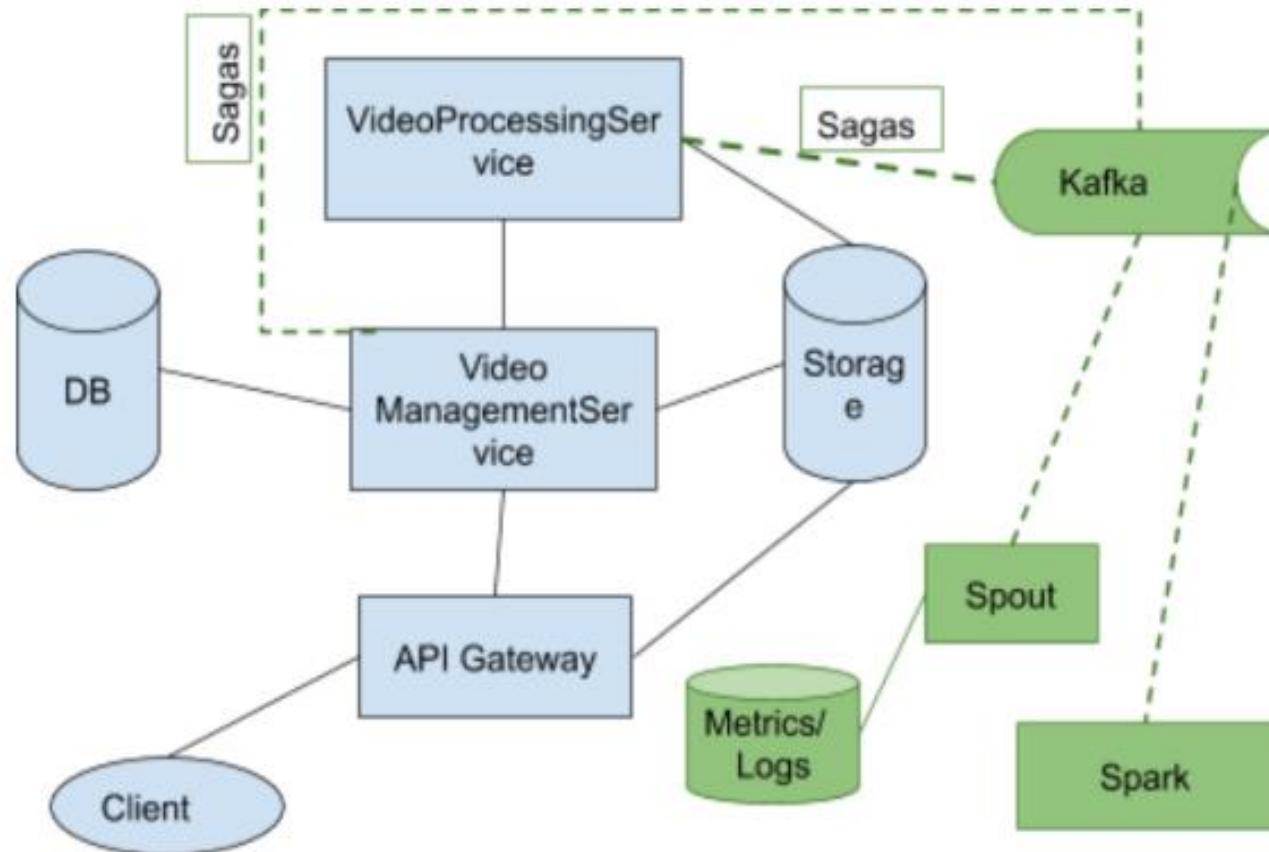


Progetto Distributed Systems and Big Data

Video Server



Strumenti Utilizzati

Docker

Kubernetes

Spring boot

Spring Cloud
Gateway

Flask

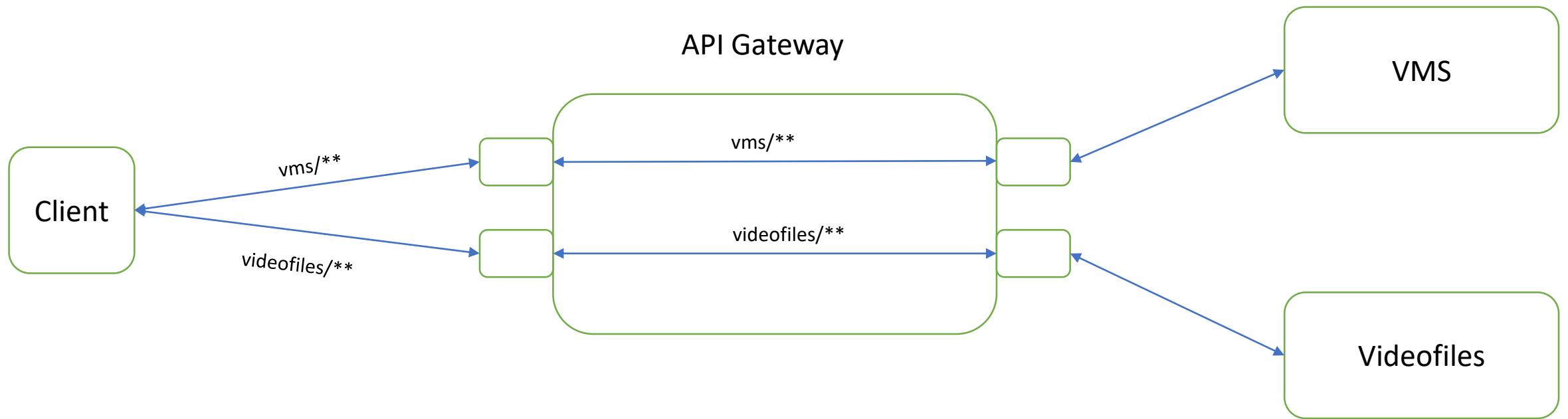
Prometheus

Kafka

Spark

API Gateway

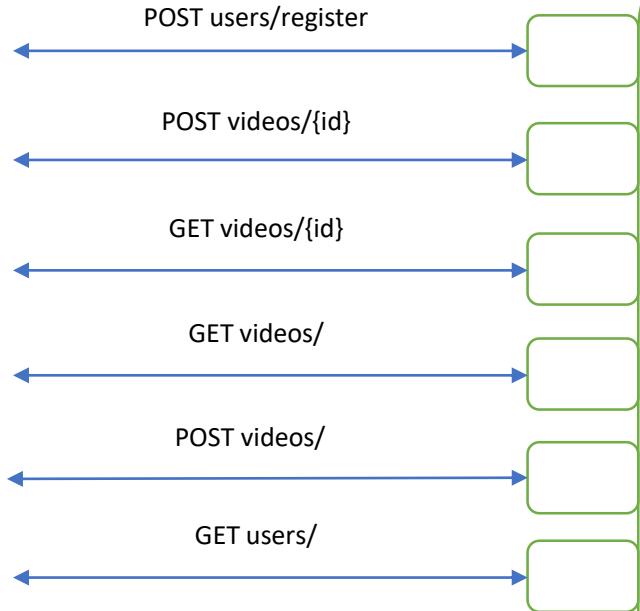
Il modulo API Gateway è stato implementato grazie all'utilizzo di Spring Cloud Gateway. Tutte le chiamate verso il sistema vengono indirizzate a questo modulo, il quale ha il compito di eseguire il routing al servizio corretto.



Video Management Service

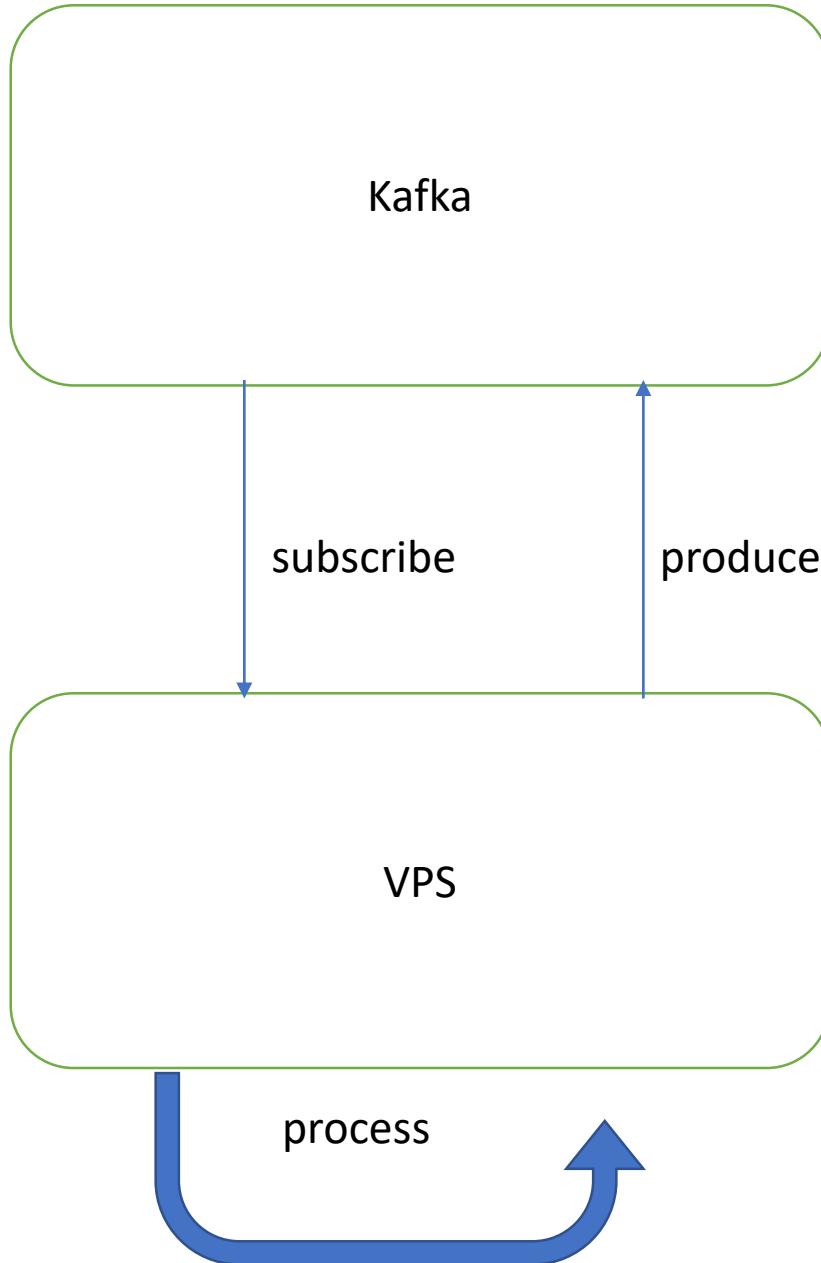
Il video Management Service (vms) è il modulo che contiene la maggior parte della logica di controllo e dei servizi del sistema. Per implementarlo è stato scelto di utilizzare Spring boot, il offre anche soluzione per architetture RESTFULL.

VMS



Video Processing Service

- Per la prima parte si è utilizzato il framework FLASK per riuscire ad implementare una comunicazione HTTP REST con VMS.
- Nella seconda parte Flask è stato rimosso in quanto la comunicazione tra i moduli doveva avvenire in maniera asincrono ed è stato introdotto Kafka.



Utilizzo di Docker

Con l'affermarsi di architetture a microservizi nella quali si necessita l'isolamento delle risorse, i container linux svolgono un ruolo fondamentale. Per questo progetto si è utilizzato docker: piattaforma open source che permette di automatizzare il deployment di applicazioni sfruttando la virtualizzazione tramite container

Prometheus

Prometheus è un sistema open-source per il monitoraggio di sistemi. All'interno del progetto è stato utilizzato per estrapolare le statistiche sulle richieste dei servizi. Per fare ciò Prometheus è stato collegato all'API Gateway e da quest'ultimo estrae tutte le metriche richieste es. Tempi di risposta, numero di richieste ecc...

Spout

Questo modulo ha il compito di estrarre da prometheus le metriche richieste per una successiva elaborazione attraverso Spark. E' stato implementato tramite codice Python. Attraverso le Api esposte da Prometheus Spout estrae le statistiche e le pubblica all'interno di un topic Kafka che verrà consumato da Spark.

Kafka

Kafka in questo progetto è stato utilizzato per implementare la comunicazione asincrona tra il Video Management Service ed il Video Processing Service.

Nello specifico il vms una volta caricato un video in formato .mp4, inserisce all'interno della coda kafka un messaggio contenente l'id del video da elaborare. A questo punto il vps legge dalla coda kafka il messaggio preleva il video da uno storage condiviso, effettua l'elaborazione e una volta conclusa pubblica un messaggio attraverso kafka notificando a vms la fine del processamento.

Apache Spark

Apache Spark è uno strumento di analisi di big data. Nel progetto è stato utilizzato per elaborare le metriche ed estrarre i dati relativi ai tempi di risposta medi e numero di richieste al secondo.

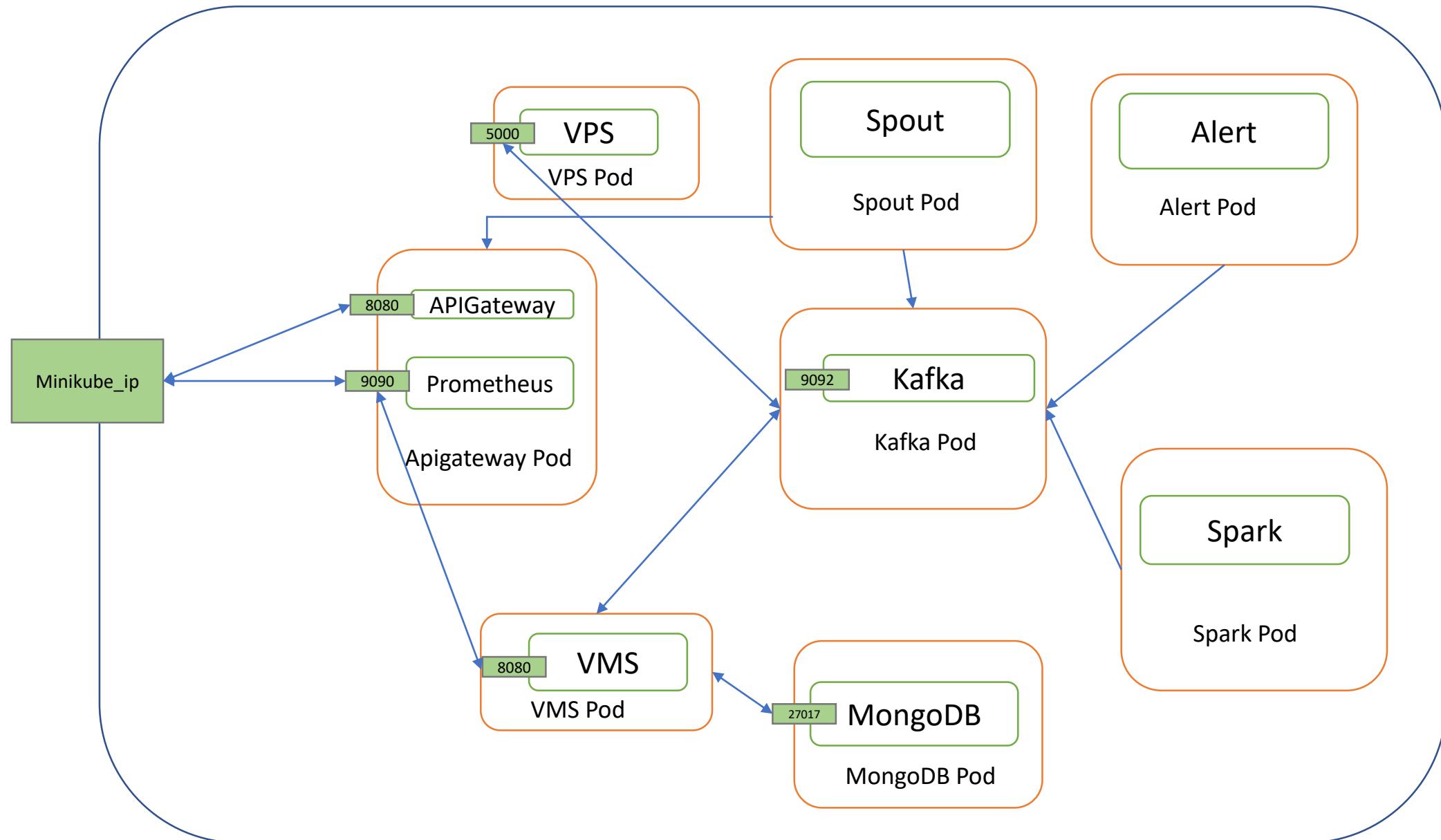
Se le analisi evincono una violazione dei criteri espressi nella consegna, il modulo Spark invia un messaggio ad una cosa Kafka che viene consumato da un componente denominato AlertManagement che visualizza il messaggio.

Migrazione su Kubernetes

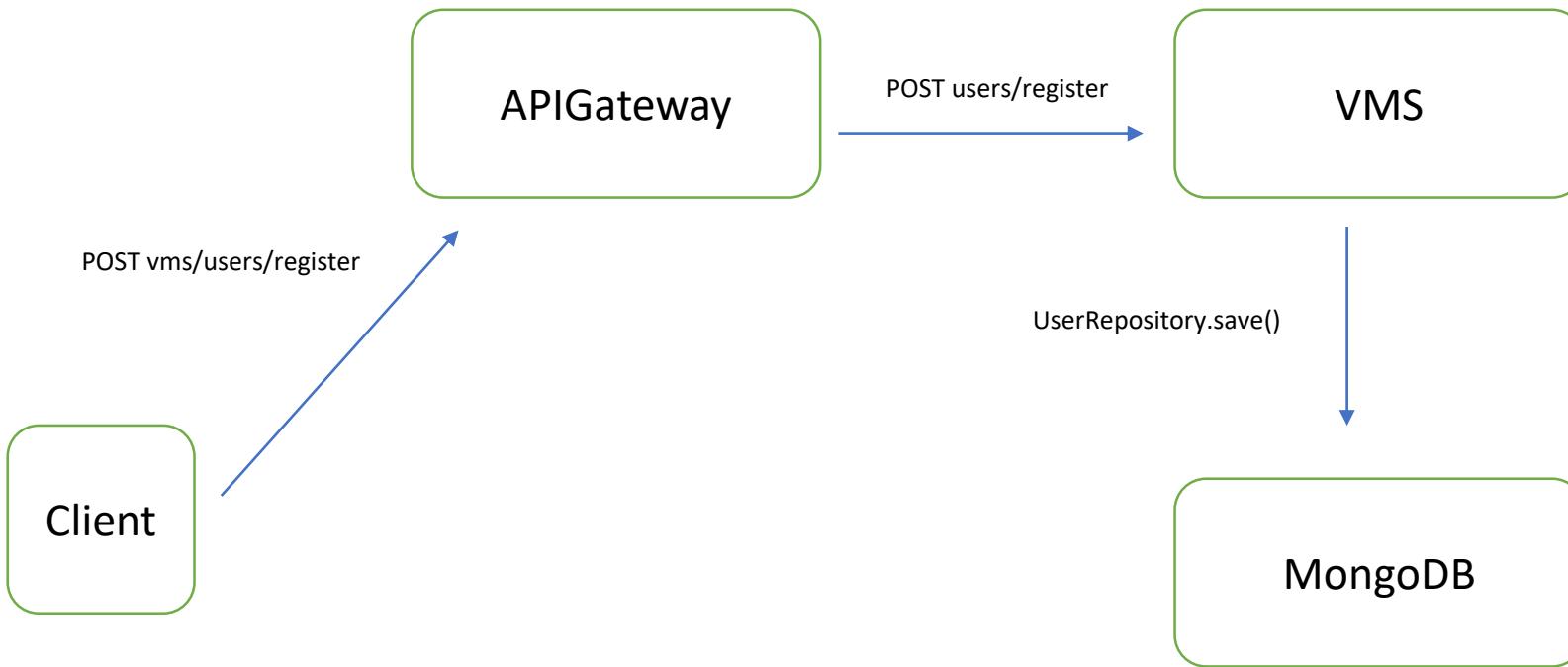
Nonostante docker risulti essere uno strumento che si adatta perfettamente alle esigenze di virtualizzazione e isolamento dei servizi durante le fasi di testing e development; per quanto concerne il deploy si necessita di uno strumento di virutalizzazione più sofisticato quale Kubernetes.

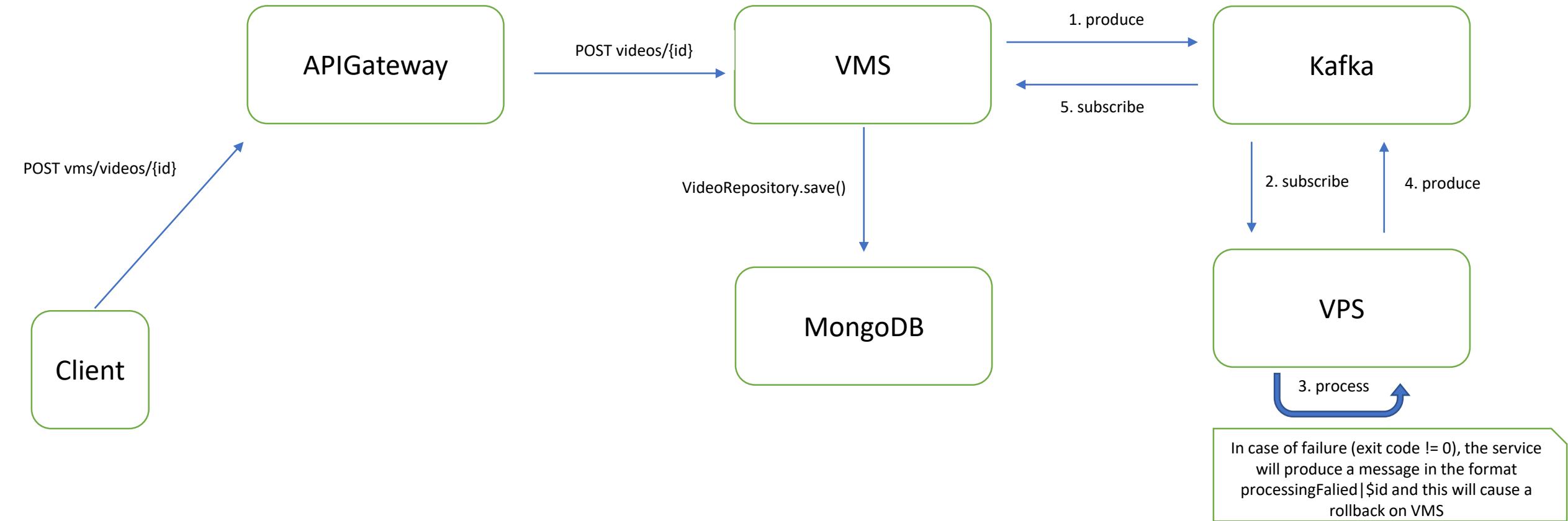
Kubernetes è un framework che si occupa di orchestrare container docker dando la possibilità di scalare i servizi in maniera automatica, tener traccia dello stato dei vari container ed inoltre fornisce dei sistemi di load balancing.

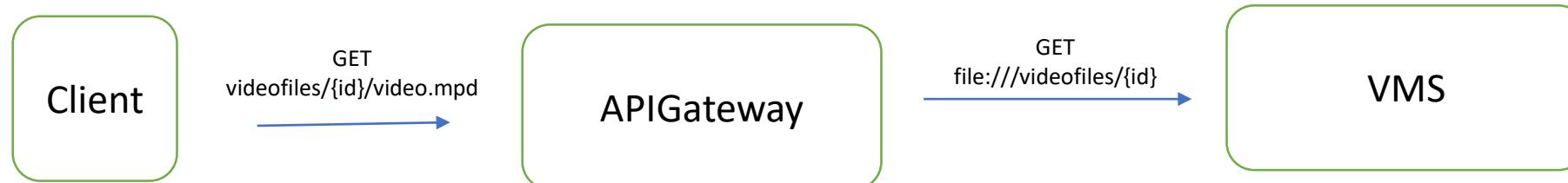
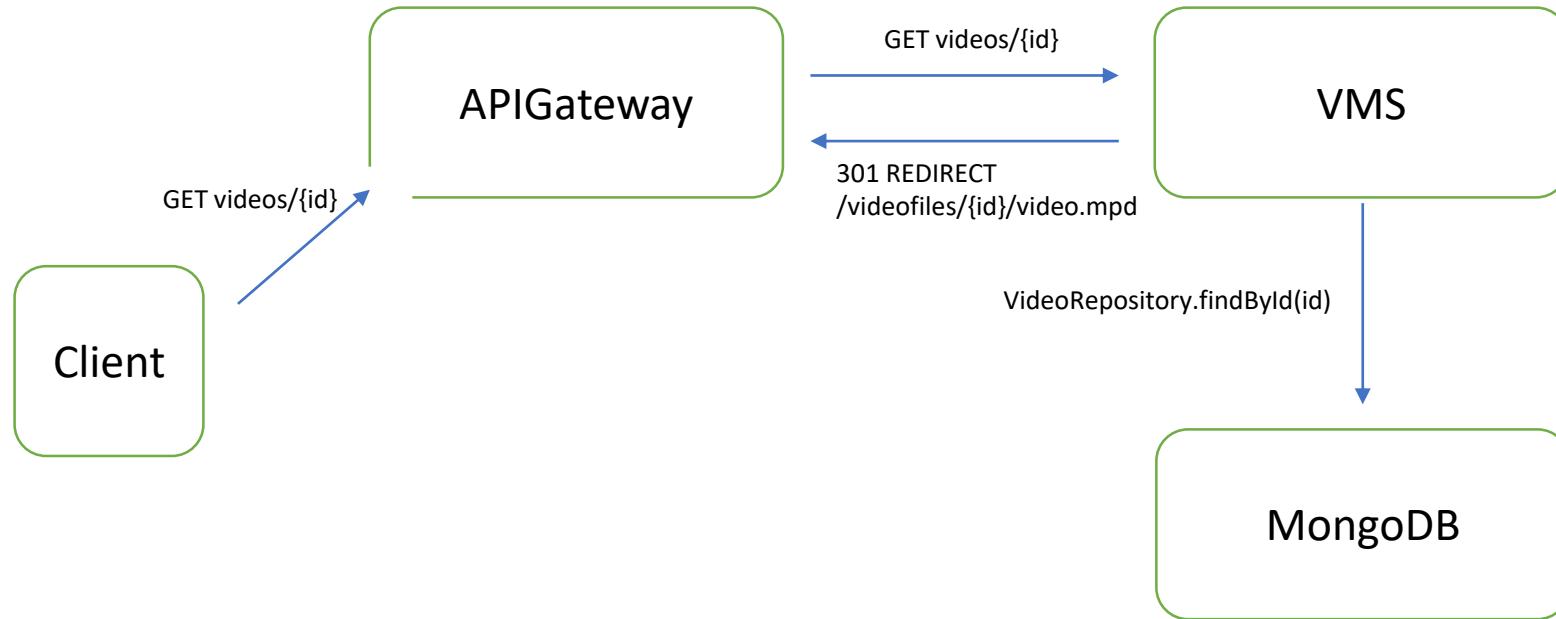
Kubernetes



Flussi







All'interno della cartella Relazione è presente un PDF con la relazione dettagliata.