

Progettare applicazioni a consumo intensivo di dati

Come piattaforme SaaS e API crescono in modo affidabile nel cloud



Scalare concretamente carichi di lavoro a consumo intensivo di dati,

Questo caso d'uso offre una visione pratica e tecnica della scalabilità dei carichi di lavoro che utilizzano molti dati, tra cui:

1

Sfide comuni che emergono con la crescita dei sistemi

2

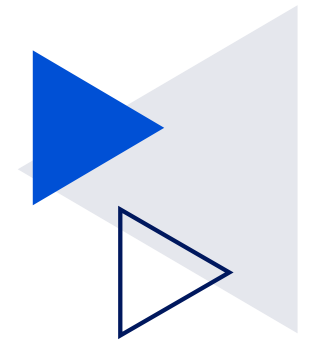
Checklist del design tecnico per guidare le decisioni infrastrutturali

3

Esempi di deploy che mostrano come le organizzazioni mantengono prestazioni, efficienza e controllo

Grazie a queste informazioni, sviluppatori e architetti possono comprendere meglio i modelli e le strategie che consentono alle piattaforme di crescere in modo affidabile gestendo complessità e costi.





1. Sfide tecniche dietro alla scalabilità dei sistemi a consumo intensivo di dati

Man mano che i volumi di dati, il throughput e la concorrenza aumentano, i team si trovano spesso ad affrontare sfide come:



Limiti delle prestazioni dello storage

man mano che i dataset crescono da gigabyte a terabyte e oltre



Previsione difficile dei costi

legati a crescita dello storage, utilizzo di I/O e dati in uscita



Mantenimento della latenza bassa e costante durante i picchi di traffico fra i vari livelli di ingestione, query e API



Aumento del carico operativo dovuto alla gestione di database e pipeline distribuiti



Impossibilità di scalare storage, database e calcolo in modo indipendente, con conseguente inefficienza



Requisiti di residenza dei dati e conformità normativa che aggiungono complessità architetturale



Colli di bottiglia della rete dovuti a traffico est-ovest di dati pesanti tra i servizi

Rispondere a queste sfide in modo efficiente richiede un'infrastruttura in grado di fornire visibilità e controllo sulle modalità in cui i dati vengono archiviati, trasferiti, elaborati e scalati. Questo faciliterà la prevedibilità di prestazioni e costi man mano che le piattaforme crescono.

2. Checklist di progettazione tecnica

Prima di scegliere una piattaforma cloud o adottare un modello architetturale, è importante valutare il comportamento dei carichi di lavoro in produzione. Dataset di grandi dimensioni, I/O elevato e alti volumi di richieste pongono esigenze specifiche in termini storage, rete e calcolo distribuito.

La seguente checklist è pensata per aiutarti a identificare chiaramente le caratteristiche tecniche da considerare nella scelta del provider cloud o dell'architettura più adatta alle tue esigenze professionali.



Storage

- ▶ Modello di storage richiesto: a oggetti, a blocchi, di file o ibrido
- ▶ Requisiti massimi di IOPS
- ▶ Requisiti NVMe per carichi di lavoro a bassa latenza
- ▶ Fattore principale delle prestazioni: throughput vs lettura/scrittura casuale
- ▶ Curva di crescita dei dati (GB/TB/PB)
- ▶ Modello di replica e durabilità: single-zone, multi-AZ o multi-region



Elaborazione e calcolo

- ▶ Profilo del workload: legato alla CPU, legato alla memoria o assistito da GPU
- ▶ Modalità di elaborazione dei dati: batch, streaming o in tempo reale
- ▶ Requisiti di elaborazione distribuita: cluster Spark, Dask, ClickHouse o altri
- ▶ Modello di utilizzo: a carico costante, con picchi di utilizzo o basato su eventi
- ▶ Capacità di containerizzazione e autoscaling



Rete

- ▶ Livello di sensibilità alla latenza
- ▶ Requisiti di banda passante est-ovest richiesti tra servizi o nodi
- ▶ Requisiti di architettura multi-zona o multi-region
- ▶ Requisiti di rete privata
- ▶ Volumi in uscita previsti



Scalabilità

- ▶ Modello di scaling: orizzontale (stateless), verticale (throughput) o ibrido
- ▶ Picchi di traffico o carico previsti: giornalieri, stagionali o legati ad eventi
- ▶ Requisiti di autoscaling, sharding o espansione del cluster
- ▶ Maturità per la containerizzazione



Costi e risorse

- ▶ Prevedibilità dell'utilizzo: stabile vs. molto variabile
- ▶ Principali fattori di costo: storage, calcolo, traffico in uscita o scaling del database
- ▶ Preferenza del modello di spesa: mensile prevedibile vs. pagamento a consumo
- ▶ Impatto dei costi di crescita e replica dell'object storage



Conformità e sicurezza

- ▶ Requisiti di residenza e sovranità dei dati
- ▶ Vincoli normativi sulla localizzazione dei dati
- ▶ Certificazioni necessarie
- ▶ Requisiti di cifratura, isolamento della rete e integrazione IAM
- ▶ SLA, aspettative di clienti e governance in materia di privacy e conformità



Automazione e deploy

- ▶ Requisiti degli strumenti Infrastructure-as-Code
- ▶ Requisiti di integrazione CI/CD
- ▶ Requisiti di velocità di provisioning degli ambienti
- ▶ Necessità di accesso a API e hook di automazione
- ▶ Requisiti di monitoraggio, logging, alert e osservabilità



Database

- ▶ Modello di database: documentale (MongoDB) vs. relazionale (PostgreSQL)
- ▶ Requisiti di lettura/scrittura concorrenti
- ▶ Modelli di query previsti: transazionali vs. analitici
- ▶ Requisiti di alta disponibilità, monitoraggio, backup e failover



Risultati di ingegneria

Piattaforme ben strutturate, progettate con una chiara comprensione di come si evolvono i ei carichi di lavoro esistenti e futuri in ambienti di produzione, permettono agli ingegneri di migliorare affidabilità, prestazioni ed efficienza.

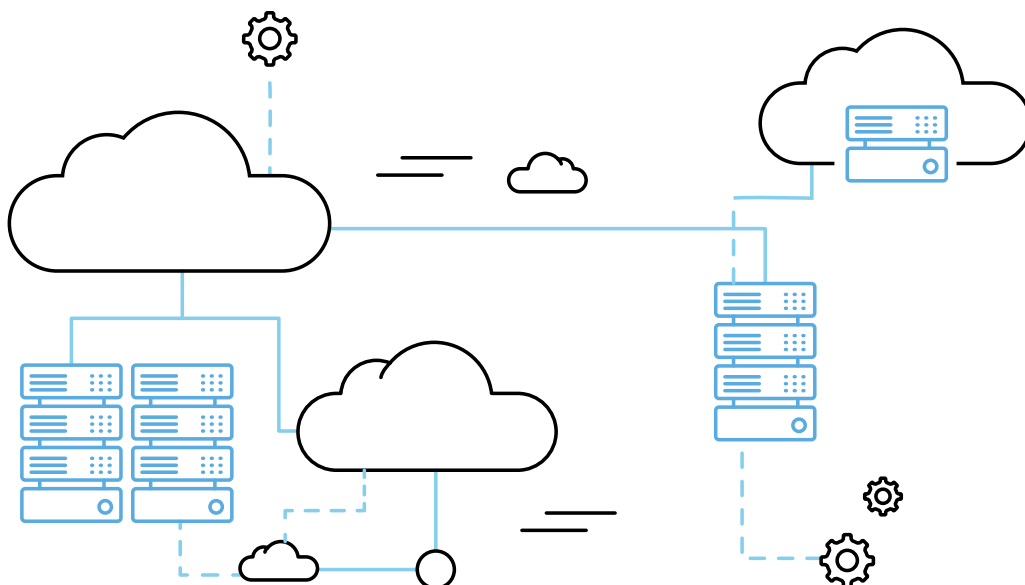
Inoltre, consentono il mantenimento di livelli di servizio coerenti, scalabilità indipendente delle risorse e gestione di servizi e costi.

CARATTERISTICA

- ✓ Throughput stabile di ingestione
- ✓ Tempi di query ed elaborazione più rapidi
- ✓ Scalabilità e storage indipendenti
- ✓ Carico operativo ridotto
- ✓ Costi cloud prevedibili
- ✓ Controllo dei dati e conformità normativa
- ✓ Architetture aperte

VANTAGGI OPERATIVI

- ✓ Meno eventi persi, SLA coerenti
- ✓ Migliore latenza API, reattività analitica
- ✓ Scaling separato per storage, database e risorse di calcolo
- ✓ Meno tempo dedicato a patch, HA, backup
- ✓ Nessun costo inaspettato per traffico in uscita o IOPS
- ✓ Chiarezza di localizzazione e residenza dei dati
- ✓ Evitare il lock-in, portabilità più semplice



Designing for AI workloads

L'IA sta accelerando la crescita di carichi di lavoro a consumo intensivo di dati che combinano alta domanda di calcolo, velocità di trasferimento dello storage e rigorosi requisiti di efficienza. Man mano che questi carichi di lavoro aumentano, i team hanno bisogno di infrastrutture che offrano prestazioni prevedibili, modelli di deploy flessibili e coerenza operativa, senza aggiungere inutile complessità.

Un approccio efficace è adottare uno stack tecnologico standard e ampiamente compatibile con ambienti ibridi.

Piattaforme come quelle di OVHcloud offrono quindi una combinazione di server on premise pronti per il deploy (On-Prem Cloud Platform), server Bare Metal e istanze cloud tradizionali, spesso basate su architetture ad alte prestazioni come AMD EPYC.

In questo modo i team di sviluppo possono selezionare la giusta infrastruttura per ogni applicazione, usufruendo allo stesso tempo di strumenti familiari, comportamento prevedibile ed efficienza su larga scala.

3. Esempi di deploy: come le organizzazioni scalano SaaS e API a consumo intensivo di dati

Le piattaforme SaaS e API ad alto consumo di dati affrontano diverse sfide infrastrutturali legate a traffico, dimensioni dei dataset e tipo di carico di lavoro.

I seguenti esempi mostrano come le aziende in crescita possono mantenere prestazioni prevedibili, scalare in modo efficiente e controllare i costi.

Se la tua sfida è...

Conformità normativa e alta disponibilità multi-AZ

Elaborazione e storage di dati su larga scala

Potrebbe interessarti:

iATROS
MapTiler

iATROS: una piattaforma di e-health sicura e conforme su larga scala

iATROS offre una piattaforma di salute digitale che raccoglie, analizza e gestisce i dati sensibili dei pazienti per centinaia di migliaia di utenti. Per rispondere a esigenti requisiti normativi, come il GDPR e altri standard specifici del settore, e ridurre

la latenza per utenti distribuiti geograficamente, il team ha riprogettato la propria architettura attorno a un'infrastruttura cloud multicluster con alta disponibilità e controlli rigorosi.

SFIDA PRINCIPALE

- ▶ Migrare da un cloud che non rispondeva pienamente ai requisiti europei di protezione dei dati
- ▶ Fornire alta disponibilità e bassa latenza per utenti distribuiti geograficamente
- ▶ Rispondere a rigorosi standard di conformità e governance (GDPR, ISO, settore sanitario)
- ▶ Mantenere un'infrastruttura resiliente tramite fault domain

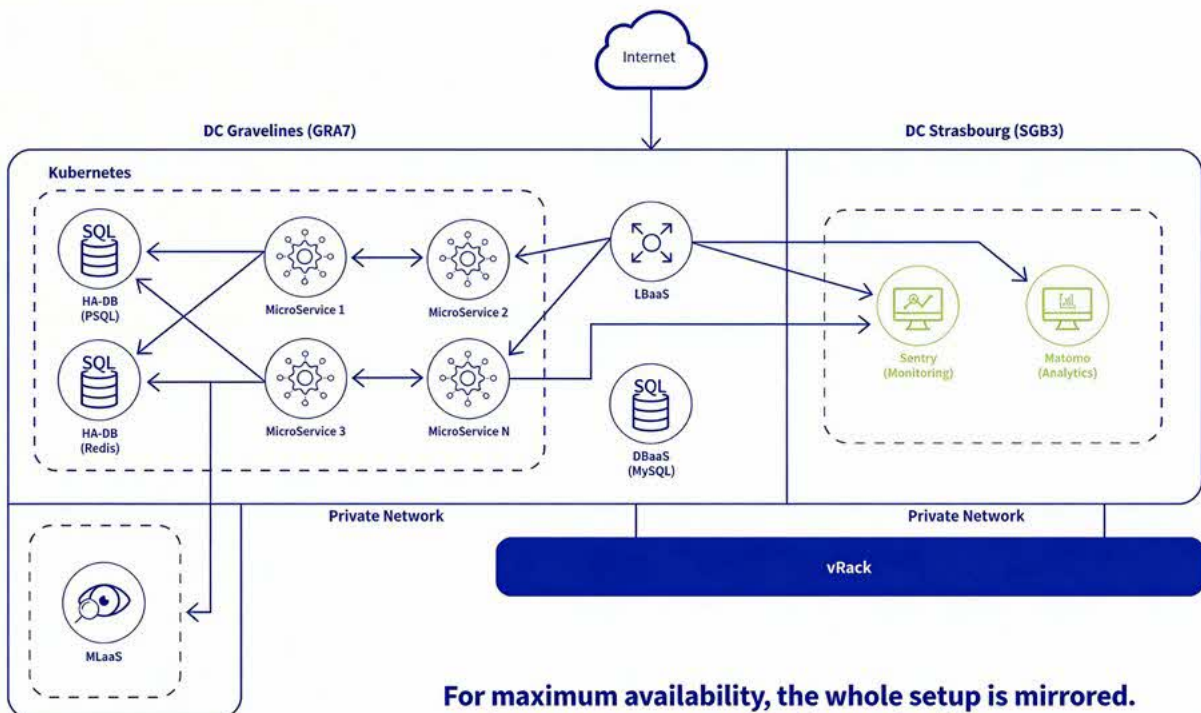
SOLUZIONE

- ▶ Carichi di lavoro migrati nei datacenter OVHcloud localizzati nell'UE per la conformità al GDPR
- ▶ Cluster PostgreSQL ad alta disponibilità eseguiti su più AZ
- ▶ Utilizzo della rete privata vRack per traffico est-ovest a bassa latenza
- ▶ Utilizzo di database gestiti e istanze di calcolo scalabili per prestazioni prevedibili

RISULTATI

- ▶ Riduzione di ~20% delle esigenze di risorse e dei costi rispetto alla configurazione precedente
- ▶ Riduzioni significative della latenza indipendentemente dalla localizzazione degli utenti
- ▶ Hosting dei dati totalmente sicuro e conforme al GDPR con governance solida e modello HA

[Leggi di più](#)



**For maximum availability, the whole setup is mirrored.
Both DC are reverse monitoring each other.**

MapTiler: generazione di mappe satellitari scalabili con istanze cloud illimitate

MapTiler, scale-up geospaziale svizzera, sviluppa e fornisce mappe di base ad alte prestazioni e dati cartografici personalizzati utilizzati da applicazioni in settori come logistica, immobiliare, difesa e turismo. Per restare competitiva con immagini satellitari aggiornate e fornire ogni mese

centinaia di milioni di visualizzazioni di mappe, MapTiler aveva bisogno di un provider cloud in grado di scalare in modo conveniente ed eliminare i colli di bottiglia dell'infrastruttura.



SFIDA PRINCIPALE

- ▶ Gestire esigenze di storage su larga scala e in rapida crescita man mano che i dati satellitari aumentano
- ▶ Supportare in modo affidabile le ~400 milioni di richieste giornaliere per le tessere di mappa
- ▶ Eliminare i costi mensili imprevedibili legati all'instabilità dei volumi di dati
- ▶ Ridurre il tempo di elaborazione da decenni a settimane con capacità di calcolo illimitata



SOLUZIONE

- ▶ Migrazione al Public Cloud di OVHcloud con object storage scalabile compatibile con S3
- ▶ Utilizzo di istanze cloud illimitate per l'elaborazione parallela dei dati
- ▶ Adozione di strutture di costo prevedibili per storage e calcolo
- ▶ Generazione semplificata di mappe satellitari con infrastrutture economicamente vantaggiose

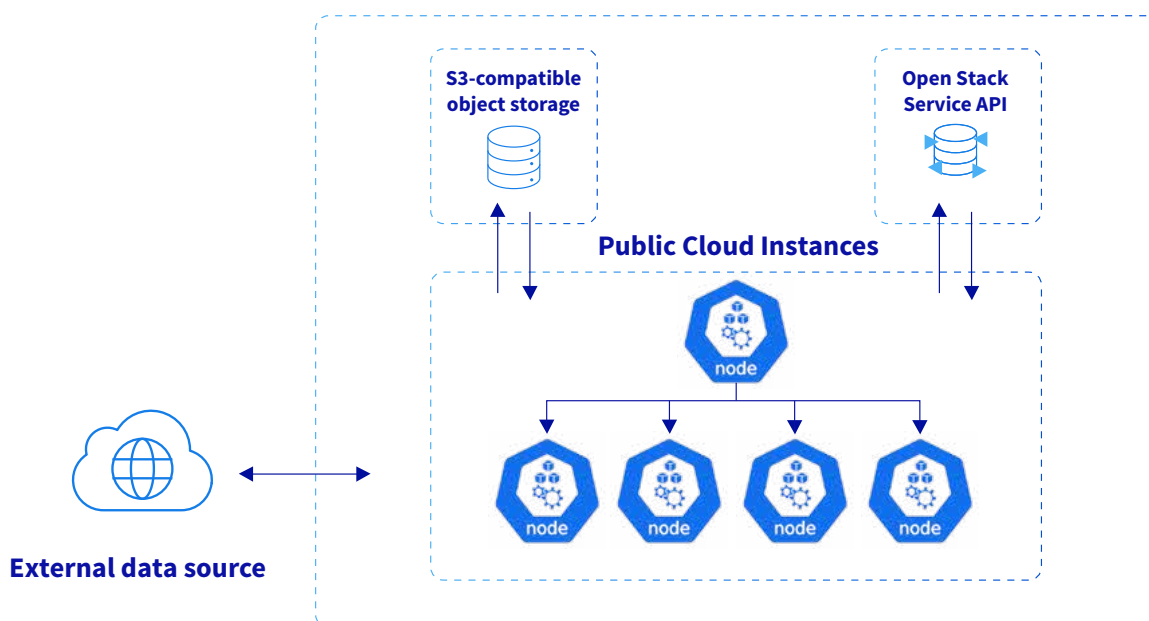


RISULTATI

- ▶ Riduzione dei tempi di elaborazione dei dati satellitari da una stima di oltre 18 anni a poche settimane
- ▶ Costi mensili prevedibili indipendentemente dall'instabilità dei carichi di lavoro dei dati
- ▶ Istanze di calcolo illimitate che prevengono limiti di capacità e accelerano la generazione di mappe

[Leggi di più](#)

MapTiler's cloud infrastructure at OVHcloud



*S3 is a registered trademark of Amazon Technologies, Inc. OVHcloud services are not sponsored or approved by, nor affiliated with Amazon Technologies, Inc. in any way.

Scalare carichi di lavoro a consumo intensivo di dati con la massima affidabilità

As data-intensive workloads grow, operational complexity escalates with data volume, users and services. Without careful design, performance bottlenecks, increasing costs and operational overhead may arise – often when systems are under strain.

Un team competente si assicura che storage, calcolo, database e rete possano scalare in modo indipendente. Quando utilizzati insieme a servizi gestiti per attività come archiviazione, orchestrazione e monitoraggio, gli sviluppatori sono liberi di concentrarsi sulle applicazioni principali mantenendo la flessibilità operativa.

La scalabilità a lungo termine consente ai team di mantenere il controllo anche quando:

- ▶ il volume dei dati aumenta, sia di 10 che di 100 volte
- ▶ i servizi o le pipeline si moltiplicano
- ▶ i malfunzionamenti diventano più frequenti

Un'infrastruttura cloud flessibile permette di ridurre inefficienze e supporta una crescita continua senza compromessi.

Scala SaaS a consumo intensivo di dati e API ad alto volume in massima sicurezza.

Scopri la differenza di OVHcloud.

Vuoi saperne di più?

Organizza una sessione con un Solution Architect:

[Chiedi di essere richiamato](#)

Scopri il cloud scalabile per le aziende in crescita:

[Scopri di più](#)