

Introduzione al Cloud

Parte III

Indice

1	Introduzione	2
1.1	Cloud computing	3
1.2	Data protection	4
2	Caratterizzazione delle sfide di protezione dei dati nel cloud	5
2.1	Le tre dimensioni del problema	5
2.1.1	Proprietà di sicurezza (CIA)	6
2.1.2	Access requirements (funzionalità richieste)	6
2.1.3	Architetture	7
2.1.4	Combinazione delle dimensioni	7
3	Digital Data Market	8

Capitolo 1

Introduzione

Gli avanzamenti tecnologici degli ultimi anni hanno cambiato la nostra società: nella società in cui viviamo oggi la tecnologia è pervasiva (*Internet of Things*), abbiamo tutto quello che avevamo prima però smart; inoltre il cloud viene largamente utilizzato per storage o computing

→ **tutto questo significa flusso di dati da me a fuori.**

Vantaggi:

- + Migliori meccanismi di protezione
- + Resilienza rispetto a malfunzionamenti
- + Miglior prevenzione e risposta (dato che il sistema è smart, può individuare certe anomalie e prevenire un attacco che sta per accadere)

Svantaggi:

- - Maggiore complessità: basta una singola porta per accedere a tutto
- - La stessa velocità che abbiamo in positivo per le funzionalità, ce l'hanno anche gli attaccanti.
- - Incremento di danni e violazioni
- - Perdita di controllo su dati e processi

Due ulteriori problemi:

- - devices vulnerabili ad attacchi esterni
- - devices che contengono dati sensibili che, se attaccati, possono essere portati fuori

Sicurezza... un problema complesso

La sicurezza è un problema complesso poichè richiede una soluzione a molti problemi (protezione delle infrastrutture, dei dati, dei dispositivi, della rete, malware, ...).

Sistema Smart

Ciò che rende un sistema *smart* è la capacità di acquisire, analizzare e processare dati per acquisire conoscenza da iniettare nuovamente nel sistema

→ prevedibilità dell'utente

1.1 Cloud computing

Il cloud permette a organizzazioni e utenti finali di avvalersi di servizi esterni per immagazzinare, processare e accedere ai loro dati.

- + Alta configurabilità
- + Dati e servizi sono sempre disponibili
- + Scalabilità

→ gli utenti perdono il controllo dei loro dati, necessità di nuove soluzioni di sicurezza per **proteggere i dati e processarli in maniera sicura nel cloud**.

Cloud: oggi

Ad oggi i cloud providers applicano misure di protezione solamente da eventuali utenti esterni (protezione rispetto al perimetro). Due scenari possibili:

- **piena fiducia** nel cloud provider in quanto ha pieno accesso ai dati
- proteggiamo i dati anche dai cloud provider ma abbiamo **funzioni limitate**, uso del provider solo come storage

Cloud: nuova visione

Si vogliono adottare soluzioni che offrono garanzie di protezione dando al *data owner* sia pieno controllo dei dati che alta funzionalità su di essi (sia sicurezza verso l'esterno che verso il cloud provider stesso).

- **client-side trust boundary**: solo i comportamenti del client dovrebbero essere considerati fidati

1.2 Data protection

Proteggere i dati non è semplice, bisogna minimizzare l'esposizione ricordandosi:

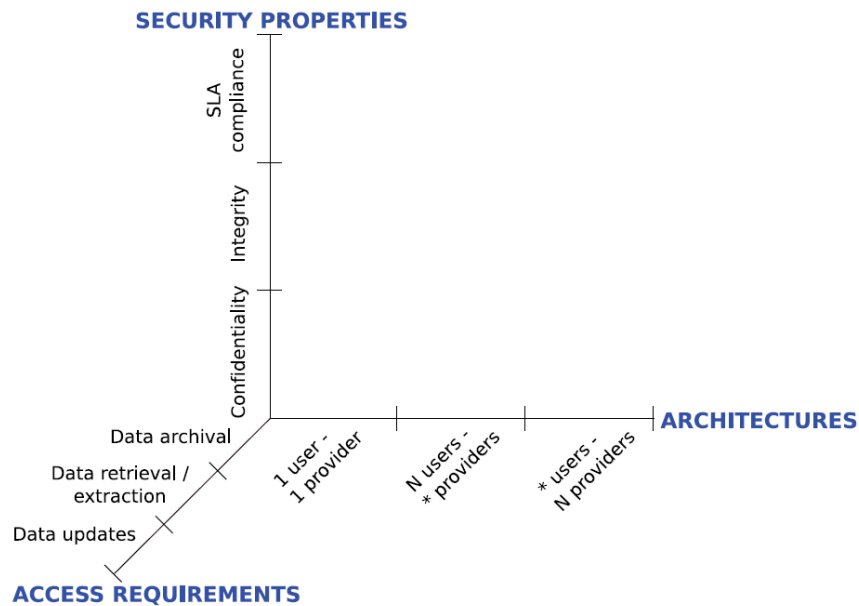
- correlazione tra diverse sorgenti
- esposizione indiretta di informazioni sensibili (ciò che potrebbe portare inferenza)
- non identificabilità \neq anonimato

Capitolo 2

Caratterizzazione delle sfide di protezione dei dati nel cloud

2.1 Le tre dimensioni del problema

Queste sono le tre dimensioni del problema *privacy e protezione nel cloud*. Non c'è una definizione di sicurezza assoluta, dipende sempre da quella richiesta e dal contesto applicativo.



2.1.1 Proprietà di sicurezza (CIA)

- **Confidenzialità:**

- protezione dei dati archiviati nel cloud
- autenticazione fatta non su chi l'utente è ma sulle proprietà che ha (certificazioni); *se devo andare in biblioteca, non serve dire chi sono ma solo che sono uno studente della statale.*
- protezione rispetto alle azioni che può fare un utente (confidenzialità sulla query)

- **Integrità:** per essere integro il dato deve essere corretto, completo e fresco (*up-to-date*)

- integrità rispetto ai dati archiviati nel cloud
- integrità rispetto alla computazione e ai risultati delle queries: è molto più difficile garantirla in computazione, per questo motivo si aggiungono punti di controllo/si pongono domande di cui si sa già la risposta

- **SLA compliance (Server Level Agreement:)** indica l'*availability* nel cloud, quindi non soffrire di negazione del servizio; gli utenti devono poter fare sempre ciò per cui sono autorizzati.

2.1.2 Access requirements (funzionalità richieste)

- **Archivio dati:**

- operazioni di upload/download
- protezione dei dati in storage (devo garantire che i file che ho memorizzato fuori siano protetti); se critto, critto a livello di file.

- **Recupero ed estrazione dei dati:** voglio avere la capacità di eseguire delle query

- supporto per un livello di granularità più fine
- protezione della computazione della query (confidenzialità della query e integrità del risultato)

- **Data update:** In questo contesto ho dei dati dinamici:

- devo supportare operazioni di insert/update (granularità più fine)
- protezione della confidenzialità delle azioni, poichè potrebbero essere osservati i cambiamenti

2.1.3 Architetture

- **1 user - 1 provider** (caso più semplice)
 - protezione dei dati in storage
 - granularità a livello di recupero dati
 - privacy e integrità delle query
- **n users - * providers:**
 - autorizzazione e controllo dell'accesso
 - gestione di scritture multiple
- *** users - n providers:** caso in cui ho diverse sorgenti dati che devono fare computazione insieme

2.1.4 Combinazione delle dimensioni

Ogni combinazione delle istanze delle tre dimensioni, identifica nuovi problemi; le proprietà di sicurezza da garantire dipendono dai requisiti di accesso e sulle assunzioni di fiducia (*trust assumption*) sui provider.

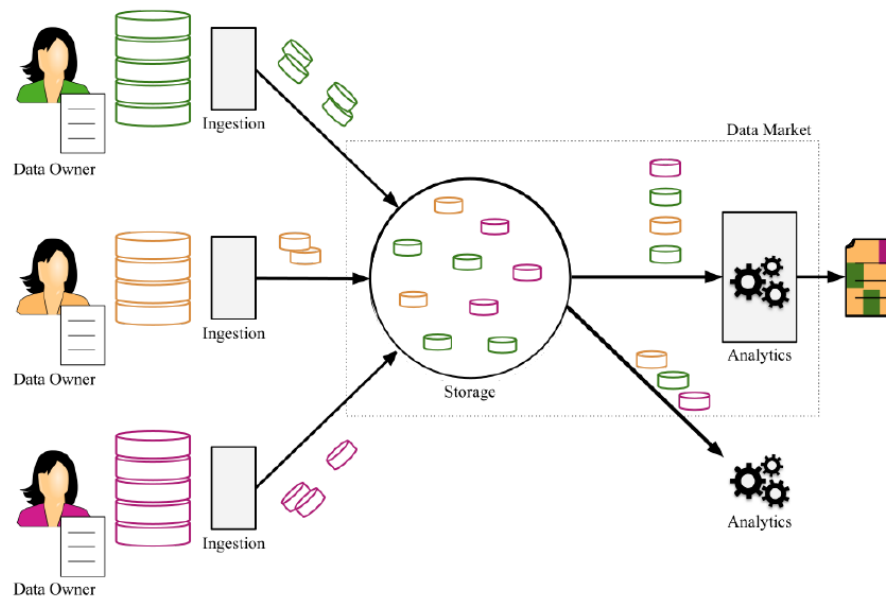
I providers possono essere:

- ***honest-but-curious***: trusted rispetto all'affidabilità del servizio ma non trusted rispetto alla confidenzialità (sbircia nei dati)
- ***lazy***: in questo caso anche problema di integrità, provider che non svolge correttamente il suo compito (es: ritorna query incomplete per risparmiare risorse)
- ***malicious***: maggiore problema di integrità, dall'altra parte c'è qualcuno (non solamente il provider stesso) che fa uno sforzo per compromettere i dati

Capitolo 3

Digital Data Market

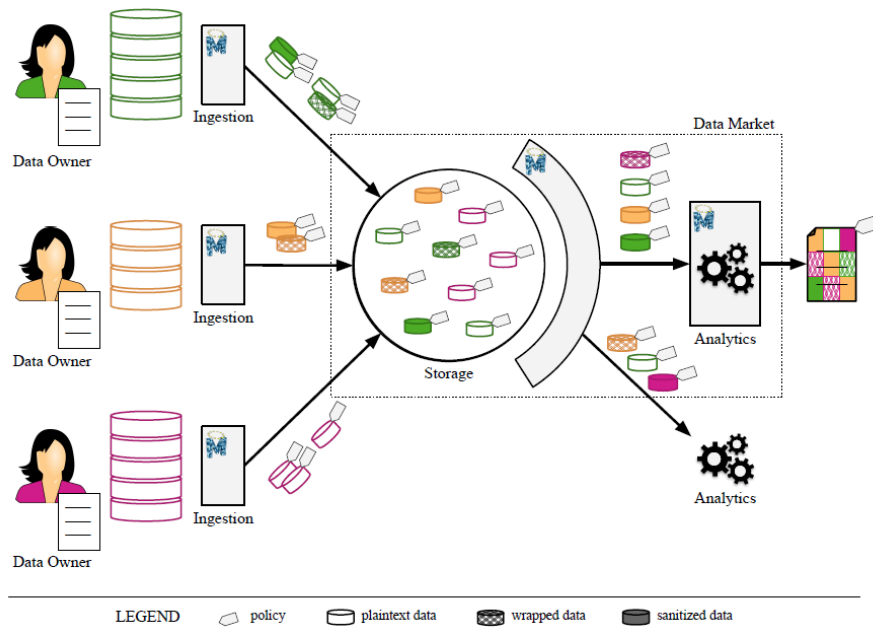
Vi sono diversi data owner che contribuiscono a uno spazio dati comune, per poi fare analisi. Il problema è che i dati possono essere sensibili, bisogna capire il livello di protezione.



Dimensioni del problema:

- **Comprensione dei requisiti:** la politica è difficile da controllare.
Sticky policy: la politica dovrebbe essere attaccata ai dati (non c'è certezza)

- **Tecnologie di applicazione (enforcing):**
 - *Data wrapping*: livello crittografico, può essere invertito
 - *Sanitizzazione*: dati vengono sanitizzati, non può essere invertito
- **Fasi di applicazione: *ingestion, storage, analytics***



- Nella fase di **ingestion** prendiamo i dati;
- in fase di **storage** li archiviamo con tutte le loro politiche, sanitizzati e/o wrappati
- e poi nell'ultima fase li **analizziamo** → dovrebbero restare così come sono stati ceduti.