# Sicurezza dei sistemi informatici

## Sicurezza dei database

### 1. Introduzione

In informatica, il termine database indica un archivio strutturato in modo tale da consentire la gestione dei dati stessi da parte di applicazioni software. Il database è un insieme di informazioni, di dati che vengono suddivisi per argomenti in ordine logico (tabelle) e poi tali argomenti vengono suddivisi per categorie (campi).
La parola “database” viene spesso usata come abbreviazione dell'espressione **Database Management System (DBMS),** che si riferisce ad una vasta categoria di sistemi software che consentono la creazione e la manipolazione efficiente di database da parte di più utenti. Infatti attraverso i DBMS l'utente riesce ad interagire con il database, ma sarà l'amministratore del database a decidere con quali permessi ogni utente può accedere (lettura,scrittura o entrambi). L'amministratore, è anche colui che definisce le regole secondo cui i dati vengono organizzati all'interno del database.

Le basi di dati, oltre ai dati veri e propri, deve contenere anche le informazioni sulle loro rappresentazioni e sulle relazioni che li legano. Spesso, una base di dati contiene le seguenti informazioni:

- Strutture dati che velocizzano le operazioni frequenti;
- Collegamenti con dati esterni, cioè riferimenti a file locali o remoti non facenti parte del database;
- Informazioni di sicurezza, che autorizzano solo alcuni profili utente ad eseguire alcune operazioni su alcuni tipi di dati;
- Programmi che vengono eseguiti, automaticamente o su richiesta di utenti autorizzati, per eseguire elaborazioni sui dati.

### 2. Sicurezza dei database

Le informazioni sono uno dei beni più preziosi della società odierna ed i database rappresentano, in senso virtuale, il luogo nel quale tali informazioni sono custodite in attesa di essere utilizzate per le finalità più disparate.
I moderni **RDBMS** (Relational DataBase Management System, ossia le banche dati o database) costituiscono il fondamento di ogni sistema di business e di altre attività di natura differente; tuttavia essi non vengono tradizionalmente considerati alla stessa stregua dei sistemi operativi e delle periferiche di rete almeno per quanto concerne l'adozione degli standard di sicurezza.
Due sono le ragioni fondamentali che rendono tali sistemi uno dei target più appetibili per gli hackers: la prima è sicuramente riconducibile alla presenza di dati che molto spesso assumono il carattere di estrema confidenzialità (numeri di carte di

credito, dati finanziari, strategici, ecc...) mentre la seconda è data dal fatto che la compromissione di un server di database determina, non di rado, a fronte di una difficoltà di penetrazione talvolta irrisoria, la possibilità da parte dell'attaccante di acquisire il completo controllo della macchina se non addirittura della intera infrastruttura di rete.

### 2.1 La gestione dei rischi

La protezione del database è l'attuazione della sicurezza informatica nel campo delle basi di dati. Il termine “sicurezza” viene spesso sostituito dal termine autorizzazione; si parla quindi di modelli di autorizzazione, gestione delle autorizzazioni, regole di autorizzazione. Qualunque attività diretta ad implementare e rafforzare la sicurezza presuppone il rispetto di questi principi:

1. la definizione di pratiche e procedure di sicurezza chiare e di facile attuazione;
2. la predisposizione di livelli di sicurezza multipli;
3. l'applicazione costante del principio “dei privilegi minori”;

## 3. Requisiti di protezione

Per una base di dati, i requisii di protezione rispetto agli attacchi sono:

- Autenticazione degli utenti;
- Protezione da inferenza;
- Integrità delle basi di dati: uso di tecniche di backup e recovery del sistema operativo e del DBMS;
- Integrità semantica: uso di vincoli semantici e di software del tipo concurrency manager del DBMS;
- Accountability e auditing;
- Verificabilità;
- Identificazione, protezione e gestione dei dati sensibili;
- Protezione multilivello, adatta per database altamente sensitive, i cui dati sono etichettai ne gli utenti identificati e autorizzati mediante clearence;
- Sconfinamento: si tratta di confinare i programmi all'esecuzione entro precisi domini di esecuzione in modo che gli eventuali danni siano limitati a quel dominio.

## 3.1 Autenticazione degli utenti

L' accountability è la capacità di un sistema di identificare un singolo utente, di determinare le azioni e il comportamento all'interno del sistema stesso. Per far questo è supportato dall'audit delle tracce e dal sistema di autenticazione (login).

L'accountability è un aspetto del controllo di accesso e si basa sulla concezione che gli individui siano responsabili delle loro azioni all'interno del sistema.

Tale aspetto è supportato dall'audit delle tracce degli eventi registrati, che può essere usato per il rilevamento di intrusioni e per il rilevamento di eventi passati.

L'auditing è mirato ad accertare la validità e l'affidabilità di un'informazione, ed è anche una verifica del sistema di controllo interno.

Un sistema di autenticazione attraverso login, l'utente fornisce un nome utente ed una password, che vengono validati dal sistema. Questo scambio avviene all'interno di un canale cifrato, per cui non è a rischio di intercettazioni.

## 3.2 Politiche di autorizzazione

Le principali politiche di autorizzazione per una base di dati si possono elencare così:

- Privilegio minimo (need-to-known): ogni utente possiede i minimi privilegi sulle risorse;
- Privilegio massimo: questa politica fornisce a tutti i soggetti la massima visibilità ed il massimo accesso alle risorse;
- Amministrazione centralizzata: esiste un unico amministratore del sistema che ha tutti i privilegi sulle risorse e che può concedere temporanei privilegi su alcune risorse e in seguito rimuoverli;
- Amministrazione decentralizzata (ownership): ogni utente è proprietario delle proprie risorse e può concedere i privilegi temporanemente ad altri soggetti, e successivamente revocarli;
- Amministrazione cooperativa: è necessaria l'autorizzazione da parte di più soggetti per ottenere un privilegio di accesso;
- Amministrazione gerarchica: esistono gerarchie di amministratori, ciascuno responsabile di una parte del database, collegati fra loro da privilegi di GRANT/REVOKE di amministrazione;
- Sistemi chiusi/aperti: in un sistema chiuso tutti i privilegi che non sono autorizzati sono negati, mentre in un sistema aperto tali privilegi sono autorizzati;
- Controllo di accesso discrezionale: i soggetti possiedono l'ownership degli

oggetti da loro creati e possono concedere e revocare a loro discrezione alcuni privilegi ad altri soggetti;
- Controllo di accesso mandatario: i soggetti non possono propagare i privilegi di accesso.

## 3.3 Integrità

L'integrità si applica sia ai singoli dati che al database nella sua interezza, quindi l'integrità risulta un problema fondamentale nella progettazione dei sistemi di gestione dei database.

L'integrità del database: gli utenti devono potersi fidare dell'accuratezza dei valori dei dati immessi, per cui gli aggiornamenti devono essere fatti solo attraverso singole autorizzazioni e i dati devono essere protetti sia dall'azione di programmi esterni illegali sia da forze esterne come calamità naturali. L'integrità del database è completa responsabilità del DBMS, del sistema operativo, e del manager di sistema. Una forma di protezione del database, quindi, può essere data da un regolare backup di tutti i file del sistema.

L’integrità degli elementi: in un database ci si riferisce alla loro correttezza e alla loro accuratezza. Gli utenti autorizzati sono responsabili dell'introduzione corretta dei dati, ma essi spesso possono commettere errori durante l’ immissione, infatti il DBMS aiuta gli utenti a trovare l’errore appena immesso e a correggere gli errori già inseriti grazie a dei controlli di campo. Un controllo di campo può essere un valore numerico, una lettera maiuscola, o uno specifico carattere, questo controllo assicura che un valore cada in specifici limiti, prevenendo così errori di distrazione durante l’immissione dei dati.

## 3.4 Verificabilità

Quando si parla di verificabilità del database, ci si riferisce a un record di verifica di tutti gli accessi, questo record aiuta a mantenere l’ integrità del database indicando chi ha modificato determinati dati e quando è avvenuta la modifica. Un altro vantaggio di questo metodo di protezione è il fatto che gli utenti possono accedere in modo incrementale ai dati protetti, ossia solo una serie di accessi sequenziali visualizzati insieme mostrerà i dati. Grazie al DBMS possono essere aggiunte al database nuove categorie di dati senza dover stravolgere il sistema esistente. Il sistema di sicurezza dei dati impedisce agli utenti non autorizzati di visualizzare o aggiornare il database. Mediante l'uso dell’autenticazione dell’utente agli utenti è permesso l'accesso all'intero database o ad un suo sottoinsieme.

## 3.5 Inferenza

Il problema dell'inferenza nei database è legato a possibili utilizzi di database di dati personali per estrarre informazioni sensibili, o rintracciarle altrove negli archivi informatici.

Ogni database può contenere, un insieme di dati più o meno sensibili, cioè quei dati che non dovrebbero essere di oggetto di dominio pubblico. Tuttavia, la maggior parte dei database può contenere dati di diversi gradi di sensibilità. In questo caso, attraverso alcune strategie di attacco ai DBMS, è possibile dedurre dati sensibili per mezzo di dati non sensibili.

Gli attacchi ai DBMS si dividono in :

- attacchi diretti: si cerca di comporre una query così precisa da restituire esattamente un risultato corrispondente di dati non sensibili, dai quali si possono dedurre facilmente i dati sensibili
- attacchi indiretti: vengono rilasciati valori come le somme, i contatori e le medie, dalle quali attraverso calcoli esterni al database, è possibile poi ricavare dati individuali
- somma: a partire da una somma riportata è possibile desumere un risultato sensibile
- contatore: i valori ottenuti da un contatore possono essere combinati con una somma, e portare così a nuovi dati
- mediani: questo attacco necessita di diverse query tali che tutte abbiano un punto di intersezione con le altre
- attacco del segugio: può indurre in errore il DBMS individuando i dati richiesti con l'utilizzo di query integrative che forniscono come risultati dei record.

Non si possono individuare soluzioni definitive al problema dell'inferenza, vi sono però tre metodi per controllarlo:

1. Soppressione dei dati sensibili più scontati.
2. Tracciamento dei dati in possesso dell'utente. Dei ogni utente si tiene conto di ciò che ha cercato e trovato.
3. Mimetizzazione dei dati. Si apportano arrotondamenti casuali, che possono bloccare attacchi statistici che hanno origine da valori esatti.

I primi due metodi possono essere usati per arginare le query accettate o per limitare i dati restituiti ad una query. Il terzo metodo è applicato solo su dati rilasciati.

## 3.6 Protezione multilivello

I sistemi di database sono spesso sistemi multi-utente, quindi l' accesso di più persone nello stesso database deve essere quindi vincolato, in modo di non creare problemi di interferenza. Il DBMS può mantenere l'integrità del database non consentendo a più utenti di modificare lo stesso record contemporaneamente (blocco del record) e può impedire l'immissione di due record duplicati .Un ulteriore problema dell' accesso concorrente è la lettura/scrittura. Ad esempio se un utente sta aggiornando un valore, mentre un altro desidera leggerlo, il lettore potrebbe ricevere dati aggiornati parzialmente, di conseguenza il DBMS blocca le richieste di lettura fino a quando non viene completata la scrittura.

## 3.7 Transazioni

Una transazione è una sequenza di operazioni, che può concludersi positivamente o negativamente, in caso positivo il risultato delle operazioni deve essere permanente, mentre in caso negativo si deve tornare allo stato prima della transazione.

Le transazioni sono implementate da DBMS o da gestori di transazioni, per essere tali, devono godere delle proprietà ACID, particolarmente importanti nei sistemi in cui si possono eseguire più transazioni contemporaneamente.

Un utilizzo tipico delle transazioni è il seguente:

- Prima di eseguire una transazione, si esegue un'istruzione di "inizio transazione"
- Si eseguono le operazioni di interrogazione e modifica dati
- Se si riscontra un'anomalia, si esegue un'istruzione detta di "rollback", per abortire la transazione
- Se si sono eseguite le operazioni senza anomalie, si esegue un istruzione di "commit", per confermare la transazione.

Se il DBMS riscontra internamente qualche anomalia, esegue automaticamente una rollback, se termina bruscamente, quando riviene attivato, esegue la rollback delle transazioni che erano in corso al momento del crash.

Per implementare una transazione, tipicamente si usa un'area d'appoggio del disco fisso in cui vengono copiati i dati originali appena prima di essere modificati. Quando viene eseguita una commit, i dati originali vengono cancellati, mentre quando viene eseguita una rollback, si ricopiano indietro i dati originali copiati.

# 4. Tecniche di protezione dei database

## 4.1 Protocollo di aggiornamento a due fasi

Le due fasi dell'algoritmo sono la fase di richiesta di validazione (fase dell'intento), e la fase di validazione (fase di commit). La tecnologia di aggiornamento a due fasi è usata quando gli aggiornamenti dei dati necessitano di avvenire simultaneamente in database multipli, è una tecnologia fatta per mantenere l'integrità e l'accuratezza dei dati, attraverso bloccaggi sincronizzati di ogni parte di una transazione.

Quando si usa una commit a due fasi, vengono mantenuti dei valori ombra, tali valori vengono elaborati ed archiviati in locale durante la fase dell'intento e viene copiato nel database effettivo durante la fase di commit.

## 4.2 Controllo della concorrenza

La concorrenza può essere:

- Concorrenza di scrittura di uno stesso elemento, con la conseguente perdita di aggiornamento.
- Concorrenza di lettura e scrittura di un elemento, si ha un'incoerenza fra il dato scritto e quello letto nel database.

Questi problemi di concorrenza possono essere risolti con meccanismi di blocco (lock) temporaneo dell'accesso, in cui si blocca un utente che vuole leggere o scrivere uno stesso dato di una tabella nello stesso momento in cui tale tabella viene elaborata da un altro utente.

# 5 Controlli di sicurezza

I controlli per una base di dati, rispetto agli attacchi, sono principalmente i seguenti:

- Controlli di flusso. Si tratta di controllare le sequenze di operazioni del tipo READ X, WRITE Y che avvengono da un oggetto X (supponiamo autorizzato) verso un oggetto Y (supponiamo non autorizzato). Infatti, il valore contenuto in X viene copiato in Y.

- Controlli di inferenza. Con operazioni di assegnamento tipo Y = f(X), in cui l' insieme di dati Y (supponiamo non autorizzato) è ricavato applicando la funzione f all'insieme X (supponiamo autorizzato). Si possono allora verificare tre tipi di inferenza: accesso diretto, dati correlati, dati mancanti.
- Controlli di accesso. Sono i tipi di controllo più diffusi per applicativi e dati e si basano sul controllo dell'identità dei soggetti, della modalità di accesso della richiesta (ad esempio scrittura, lettura), e dell'oggetto cui il soggetto chiede di accedere. Le modalità di accesso vengono anche dette privilegi di accesso e privilegi amministrativi, che permettono ad alcuni soggetti speciali di concedere e revocare privilegi di accesso alle risorse. I privilegi amministrativi vengono realizzati pressoché tutti in sistemi mediante le due operazioni GRANT e REVOKE.