An toàn Cơ sở dữ liệu

Trần Đức Khánh
Bộ môn HTTT – Viện CNTT&TT
ĐH BKHN

Cơ sở dữ liệu

- Tập hợp các dữ liệu có quan hệ được lưu trữ (tập trung hoặc phân tán) để người dùng có thể truy nhập khi cần

Mục tiêu an toàn CSDL

- Bí mật
  - Dữ liệu nhạy cảm
- Toàn vẹn
  - Vật lý
  - Logic
- Sẵn dùng
- Kiểm soát truy nhập
- Xác thực người dùng

An toàn CSDL

- Các mối đe dọa CSDL
  - Cập nhật CSDL
  - Dữ liệu nhạy cảm
  - Suy diễn
- Thiết kế CSDL tin cậy
  - CSDL đa tầng

An toàn CSDL

- Các mối đe dọa CSDL
  - Cập nhật CSDL
  - Dữ liệu nhạy cảm
  - Suy diễn
- Thiết kế CSDL tin cậy
  - CSDL đa tầng

Cập nhật CSDL

- Sự cố hệ thống
  - Hệ thống bị sự cố khi đang cập nhật CSDL
- Tương tranh
  - Nhiều người dùng truy nhập sửa đổi cùng lúc vào cùng dữ liệu

Sự cố hệ thống: ví dụ CSDL văn phòng phẩm

- Kho chứa văn phòng phẩm
  - Giấy, bút,…
- Quản trị kho chịu trách nhiệm đặt mua văn phòng phẩm
- Phòng, ban sử dụng văn phòng phẩm
  - Mỗi văn phòng phẩm có một quota kinh phí văn phòng phẩm cố định

CSDL văn phòng phẩm: một kịch bản cập nhật

- Phòng kế toán yêu cầu 50 hộp ghim giấy
- Giả sử còn 107 hộp ghim giấy trong kho
- Quản trị kho sẽ đặt hàng nếu số lượng hộp ghim giấy nhỏ hơn 100

Các bước thực hiện kịch bản

1. Kiểm tra kho còn đủ 50 hộp hay không. Nếu không, yêu cầu bị từ chối. Kết thúc giao tác
2. Còn đủ số lượng. Làm phép trừ (107-50=57)
3. Trừ quota kinh phí của phòng kế toán
4. Kiểm tra số lượng hộp có dưới 100 hay không (57). Đặt mua thêm
5. Gửi 50 hộp ghim giấy cho phòng kế toán

Sự cố kịch bản

- Điều gì xảy ra nếu có sự cố hệ thống sau các bước 1,2,3,4

Giải pháp phòng chống sự cố hệ thống

- Giải pháp 2 pha
  1. Intent
     1. Thu thập tài nguyên và thông tin, tính toán và chuẩn bị dữ liệu cho pha sau
     2. Đánh dấu kết thúc pha Intent
        - Set COMMIT-FLAG
  2. Commit
     1. Thực hiện cập nhật CSDL với các dữ liệu đã chuẩn bị ở pha trước
     2. Đánh dấu kết thúc pha Commit
        - Unset COMMIT-FLAG

Ví dụ giải pháp 2 pha: cập nhật CSDL văn phòng phẩm

Intent

1. Kiểm tra giá trị COMMIT-FLAG. Nếu khác 0, chờ cho đến khi COMMIT-FLAG bằng 0
2. So sánh số hộp ghim giấy đang có với số yêu cầu; nếu số yêu cầu lớn hơn, kết thúc
3. Tính TCLIPS = ONHAND - REQUISITION
4. Xem BUDGET hiện nay của phòng kế toán. Tính TBUDGET = BUDGET - COST, trong đó COST là giá thành của 50 hộp ghim
5. Kiểm tra TCLIPS < 100 hay không; nếu đúng, TREORDER = TRUE; nếu không TREORDER = FALSE

Ví dụ giải pháp 2 pha: cập nhật CSDL văn phòng phẩm

Commit

1. COMMIT-FLAG = 1
2. Cóp TCLIPS vào CLIPS trong CSDL
3. Cóp TBUDGET vào BUDGET trong CSDL
4. Cóp TREORDER vào REORDER trong CSDL
5. Chuẩn bị giấy báo giao hàng cho phòng kế toán. Ghi chú giao tác hoàn thành vào log
6. COMMIT-FLAG = 0

Tương tranh: ví dụ CSDL vé máy bay

- Văn phòng du lịch A
  - SELECT (SEAT-NO = '11D') ASSIGN 'MOCK,E' TO PASSENGER-NAME
- Văn phòng du lịch B
  - SELECT (SEAT-NO = '11D') ASSIGN 'EHLERS,P' TO PASSENGER-NAME

Giải pháp tương tranh

- Cập nhật CSDL là một thao tác cơ bản
  - Chỉ một thao tác cập nhật thực hiện trên một dữ liệu
  - Hệ quản trị CSDL sẽ bảo vệ dữ liệu đang được cập nhật
  - Khi thao tác cập nhật một dữ liệu kết thúc, các thao tác khác mới có quyền thực hiện trên dữ liệu đó

Dữ liệu nhạy cảm

- Dữ liệu công chúng không nên có
- Loại dữ liệu nhạy cảm
  - Bảng
  - Bản ghi
  - Trường

Dữ liệu nhạy cảm

- Các loại rò gỉ dữ liệu nhạy cảm
  - Dữ liệu chính xác
  - Cận
  - Kết quả âm
  - Tồn tại
  - Giá trị xác xuất

Bảo vệ dữ liệu nhạy cảm

- Hệ quản trị CSDL quản lý truy nhập dữ liệu nhạy cảm bằng kiểm soát truy nhập

Suy diễn

- Suy diễn dữ liệu nhạy cảm từ dữ liệu không nhạy cảm

Suy diễn

- Các loại tấn công suy diễn
  - Trực tiếp (Direct)
  - Gián tiếp (Indirect)
    - Tổng (Sum)
    - Đếm (Count)
    - …

Ví dụ

Name	Sex	Race	Aid	Fines	Drugs	Dorm
Adams	M	C	5000	45.	1	Holmes
Bailey	M	B	0	0.	0	Grey
Chin	F	A	3000	20.	0	West
Dewitt	M	B	1000	35.	3	Grey
Earhart	F	C	2000	95.	1	Holmes
Fein	F	C	1000	15.	0	West
Groff	M	C	4000	0.	3	West
Hill	F	B	5000	10.	2	Holmes
Koch	F	C	0	0.	1	West
Liu	F	A	0	10.	2	Grey
Majors	M	C	2000	0.	2	Grey

Ví dụ: tấn công trực tiếp

List NAME where SEX=M DRUGS=1

- Trả về thông tin liên quan đến Adam
- Hệ CSDL có thể từ chối câu truy vấn này vì quá đặc biệt

List NAME where (SEX=M and DRUGS=1) or (SEX /= M and SEX /= F) or (DORM=AYRES)

- Câu truy vấn phức tạp hơn nhưng kết quả giống như trên

Ví dụ: tấn công gián tiếp

- Dùng tổng (Sum): tổng của kinh phí hỗ trợ theo giới tính và khu vực

	Holmes	Grey	West	Total
M	5000	3000	4000	12000
F	7000	0	4000	11000
Total	12000	3000	8000	23000

Ví dụ: tấn công gián tiếp

- Dùng đếm kết hợp tổng (Count & Sum): số sinh viên theo giới tính và khu vực

	Holmes	Grey	West	Total
M	1	3	1	5
F	2	1	3	6
Total	3	4	4	11

Biện pháp ngăn chặn tấn công suy diễn

- Phân tích câu truy vấn
- Ngụy trang thông tin
- Loại bỏ thông tin nhạy cảm

An toàn CSDL

- Các mối đe dọa CSDL
  - Cập nhật CSDL
  - Dữ liệu nhạy cảm
  - Suy diễn
- Thiết kế CSDL tin cậy
  - CSDL đa tầng

CSDL đa tầng

- Các tầng CSDL tương ứng với mức độ nhạy cảm của dữ liệu
- Các tiếp cận
  - Phân ngăn (Partitioning)
  - Mã hóa (Encryption)
  - Khóa
    - Khóa toàn vẹn (Integrity Lock)
    - Khóa nhạy cảm (Sensitive Lock)
  - Front-end tin cậy (Trusted Front-end)
  - Bộ lọc giao hoán (Commutative Filter)
  - Cửa sổ (Window/View)

Phân ngăn

- CSDL được chia thành các CSDL khác nhau ở mức độ nhạy cảm khác nhau
- Ưu điểm
  - Quản lý an toàn ở nhiều mức khác nhau
- Nhược điểm
  - Dư thừa
  - Không kết hợp dữ liệu ở các mức nhạy cảm khác nhau

Mã hóa

- Mỗi dữ liệu nhạy cảm sẽ được mã hóa bằng một khóa tương ứng
- Ưu điểm
  - Quản lý an toàn ở nhiều mức khác nhau
- Nhược điểm
  - Tốc độ
  - Không gian lưu trữ

Khóa toàn vẹn

- Mục tiêu đảm bảo tính toàn vẹn và giới hạn truy nhập
- Khóa
  - Checksum
    - Tính toán bằng hàm mã hóa hoặc hàm băm
    - Giá trị phụ thuộc vào Data ID + Data + Sensitivity Label

Khóa nhạy cảm

- Mục tiêu che giấu độ nhạy cảm của dữ liệu
- Khóa
  - Mã
    - Tính toán bằng hàm mã hóa hoặc hàm băm
    - Giá trị phụ thuộc vào Data ID + Sensitivity Level

Front-end tin cậy

- Hoạt động giống "Giám sát thẩm quyền"
- Kiểm soát truy nhập CSDL
  1. Xác thực người dùng
  2. Kiểm tra quyền người dùng
  3. Gửi truy vấn cho hệ quản trị CSDL
  4. Nhận kết quả truy vấn
  5. Phân tích độ nhạy cảm của kết quả truy vấn, so sánh với quyền người dùng
  6. Định dạng lại kết quả truy vấn
  7. Gửi kết quả truy vấn cho người dùng

Bộ lọc giao hoán

- Hoạt động giống "Front-end tin cậy"
- Điều khiển truy nhập CSDL
  1. Xác thực người dùng
  2. Kiểm tra quyền người dùng
  3. Định dạng lại truy vấn
  4. Gửi truy vấn cho hệ quản trị CSDL
  5. Nhận kết quả truy vấn
  6. Phân tích độ nhạy cảm dữ liệu của kết quả truy vấn, so sánh với quyền người dùng
  7. Định dạng lại kết quả truy vấn
  8. Gửi kết quả truy vấn cho người dùng

Cửa sổ

- Mục tiêu giới hạn “tầm nhìn” của người dùng theo quyền
  - Quyền đọc, ghi
- Mỗi cửa sổ là một tập con của CSDL
  - Mỗi tập con tương ứng với dữ liệu mà người dùng có quyền sử dụng