

Nội dung

- Điều khiển dữ liệu phân tán
 - Quản lý khung nhìn
 - Bảo mật dữ liệu
 - Kiểm soát toàn vẹn ngữ nghĩa

32

Kiểm soát toàn vẹn ngữ nghĩa

Duy trì **tính nhất quán** của CSDL bằng cách thực thi một tập các ràng buộc được xác định trên CSDL.

- Ràng buộc về cấu trúc
 - Các thuộc tính ngữ nghĩa cơ bản vốn có của một mô hình dữ liệu, ví dụ: ràng buộc khóa duy nhất trong mô hình quan hệ
- Ràng buộc về hành vi
 - Điều chỉnh hành vi ứng dụng, ví dụ: các phụ thuộc trong mô hình quan hệ
- Hai thành phần
 - Đặc tả ràng buộc toàn vẹn
 - Thực thi ràng buộc toàn vẹn

33

Kiểm soát toàn vẹn ngữ nghĩa

- Thủ tục
 - Điều khiển được nhúng trong mỗi chương trình ứng dụng
- Khai báo
 - Xác nhận trong các phép vị từ
 - Dễ dàng định nghĩa các ràng buộc
 - Định nghĩa về tính nhất quán của CSDL rõ ràng
 - Nhưng không hiệu quả khi kiểm tra các xác nhận cho mỗi lần cập nhật
 - Giới hạn không gian tìm kiếm
 - Giảm số lượng truy nhập/xác nhận dữ liệu
 - Chiến lược phòng ngừa
 - Kiểm tra thời điểm biên dịch

34

Ngôn ngữ đặc tả ràng buộc

Các ràng buộc được xác định trước

Xác định các ràng buộc phổ biến hơn của mô hình quan hệ

- Thuộc tính Not-null

`ENO NOT NULL IN EMP`

- Khóa duy nhất

`(ENO, PNO) UNIQUE IN ASG`

- Khóa ngoại

Một khóa trong quan hệ R là khóa ngoại nếu nó là khóa chính của một quan hệ S khác và sự tồn tại của bất kỳ giá trị nào của nó trong R phụ thuộc vào sự tồn tại của cùng một giá trị trong S

`PNO IN ASG REFERENCES PNO IN PROJ`

- Phụ thuộc hàm

`ENO IN EMP DETERMINES ENAME`

35

Ngôn ngữ đặc tả ràng buộc

Các ràng buộc được biên dịch trước

Thể hiện các điều kiện tiên quyết phải được thỏa mãn bởi tất cả các bộ dữ liệu trong một quan hệ đối với một kiểu cập nhật nhất định

(INSERT, DELETE, MODIFY)

NEW – Phạm vi trên các quan hệ mới được chèn vào

OLD – Phạm vi trên các bộ dữ liệu cũ sẽ bị xóa

Định dạng chung:

```
CHECK ON <relation> [WHEN <update type>]
      <qualification>
```

36

Ngôn ngữ đặc tả ràng buộc

Các ràng buộc được biên dịch trước

□ Ràng buộc miền

```
CHECK ON PROJ (BUDGET ≥ 500000 AND BUDGET ≤ 1000000)
```

□ Ràng buộc miền khi xóa

```
CHECK ON PROJ WHEN DELETE (BUDGET = 0)
```

□ Ràng buộc dịch chuyển

```
CHECK ON PROJ (NEW.BUDGET > OLD.BUDGET AND
                NEW.PNO = OLD.PNO)
```

37

Ngôn ngữ đặc tả ràng buộc

Các ràng buộc chung

Các ràng buộc phải luôn đúng. Công thức tính toán quan hệ theo bộ, trong đó tất cả các biến đều được định lượng.

Định dạng chung

```
CHECK ON <variable>:<relation>, (<qualification>)
```

□ Phụ thuộc hàm

```
CHECK ON e1:EMP, e2:EMP
      (e1.ENAME = e2.ENAME IF e1.ENO = e2.ENO)
```

□ Ràng buộc với hàm tổng hợp

```
CHECK ON g:ASG, j:PROJ
      (SUM(g.DUR WHERE g.PNO = j.PNO) < 100 IF
       j.PNAME = "CAD/CAM")
```

38

Thực thi tính toàn vẹn

Hai phương pháp

■ Phát hiện

Thực thi cập nhật $u: D \rightarrow D_u$

Nếu D_u không nhất quán thì

nếu có thể: bù lại $D_u \rightarrow D_u'$

ngược lại

hoàn tác $D_u \rightarrow D$

■ Phòng ngừa

Thực thi $u: D \rightarrow D_u$ khi và chỉ khi D_u nhất quán

□ Xác định chương trình hợp lệ

□ Xác định trạng thái hợp lệ

39

Hiệu chỉnh truy vấn

- Phòng ngừa
- Bổ sung chứng nhận xác nhận vào truy vấn cập nhật
- Chỉ áp dụng cho các công thức tính toán bộ có các biến định lượng phổ biến

```

UPDATE PROJ
SET BUDGET = BUDGET*1.1
WHERE PNAME = "CAD/CAM"

```

↓

```

UPDATE PROJ
SET BUDGET = BUDGET*1.1
WHERE PNAME = "CAD/CAM"
AND NEW.BUDGET ≥ 500000
AND NEW.BUDGET ≤ 1000000

```

40

Các xác nhận được biên dịch

Bộ ba (R, T, C) trong đó

R quan hệ
 T loại cập nhật (chèn, xóa, sửa đổi)
 C xác nhận trên các quan hệ khác

Ví dụ: Xác nhận khóa ngoại

$\forall g \in \text{ASG}, \exists j \in \text{PROJ} : g.\text{PNO} = j.\text{PNO}$

Các xác nhận được biên dịch:

$(\text{ASG}, \text{INSERT}, C1), (\text{PROJ}, \text{DELETE}, C2), (\text{PROJ}, \text{MODIFY}, C3)$

trong đó:

$C1: \forall \text{NEW} \in \text{ASG}^+ \exists j \in \text{PROJ} : \text{NEW.PNO} = j.\text{PNO}$

$C2: \forall g \in \text{ASG}, \forall \text{OLD} \in \text{PROJ}^- : g.\text{PNO} \neq \text{OLD.PNO}$

$C3: \forall g \in \text{ASG}, \forall \text{OLD} \in \text{PROJ}^- \exists \text{NEW} \in \text{PROJ}^+ :$

$g.\text{PNO} \neq \text{OLD.PNO} \text{ OR } \text{OLD.PNO} = \text{NEW.PNO}$

41

Các quan hệ vi phân

Cho quan hệ R và cập nhật u

R^+ chứa các bộ được chèn bởi u

R^- chứa các bộ bị xóa bởi u

Các loại của u

insert R^+ rỗng

delete R^- rỗng

modify $R^+ \cup (R - R^-)$

42

Các quan hệ vi phân

Thuật toán:

Đầu vào: Quan hệ R , cập nhật u , xác nhận được biên dịch C_i

Bước 1: Tạo các quan hệ vi phân R^+ và R^-

Bước 2: Truy xuất các bộ của R^+ và R^- **không** thỏa mãn C_i

Bước 3: Nếu việc truy xuất không thành công thì xác nhận là hợp lệ.

Ví dụ:

u là delete trên J. Enforcing (EMP, DELETE, C2) :

Trích xuất tất cả các bộ của EMP-

vào trong RESULT

trong đó not(C2)

Nếu RESULT = {}, xác nhận là hợp lệ

43

Kiểm soát tính toàn vẹn phân tán

- Các vấn đề:
 - Định nghĩa các ràng buộc
 - Xem xét các mảnh
 - Nơi lưu trữ
 - Nhân bản
 - Không nhân bản: các mảnh
 - Thực thi
 - Chi phí tối thiểu

44

Các loại xác nhận phân tán

- Xác nhận riêng
 - Quan hệ đơn, biến đơn
 - Ràng buộc miền
- Xác nhận hướng tập hợp
 - Quan hệ đơn, đa biến
 - Phụ thuộc hàm
 - Đa quan hệ, đa biến
 - Khóa ngoại
- Xác nhận liên quan đến tổng hợp

45

Kiểm soát tính toàn vẹn phân tán

- Định nghĩa xác nhận
 - Tương tự với các kỹ thuật tập trung
 - Chuyển đổi các xác nhận thành các xác nhận được biên dịch
- Lưu trữ xác nhận
 - Các xác nhận riêng
 - Một quan hệ, chỉ các mảnh
 - Tại mỗi trạm phân mảnh, kiểm tra tính tương thích
 - Nếu tương thích hãy lưu trữ; ngược lại thì từ chối
 - Nếu tất cả các trạm từ chối, thì từ chối toàn cục
 - Xác nhận hướng tập hợp
 - Liên quan đến các phép nối (giữa các mảnh hoặc các quan hệ)
 - Có thể cần phải thực hiện các phép nối để kiểm tra tính tương thích
 - Lưu trữ nếu tương thích

46

Kiểm soát tính toàn vẹn phân tán

- Thực thi xác nhận
 - Nơi thực thi xác nhận phụ thuộc vào
 - Loại xác nhận
 - Loại cập nhật và nơi phát hành cập nhật
 - Xác nhận riêng
 - Khi update = insert
 - Thực thi tại trạm nơi bản cập nhật được phát hành
 - Khi update = qualified (đủ điều kiện)
 - Gửi xác nhận đến tất cả các trạm có liên quan
 - Thực thi chứng nhận để có được R^+ và R^-
 - Mỗi trạm thực thi xác nhận của nó
 - Xác nhận hướng tập hợp
 - Quan hệ đơn
 - Tương tự như các xác nhận riêng với các cập nhật đủ điều kiện
 - Đa quan hệ
 - Di chuyển dữ liệu để thực hiện nối; sau đó gửi kết quả đến trạm chủ truy vấn

47

Kết luận

- Các giải pháp ban đầu được thiết kế cho các hệ thống tập trung đã được mở rộng đáng kể cho các hệ thống phân tán
 - Khung nhìn cụ thể hóa và kiểm soát truy nhập tùy ý dựa trên nhóm
- Kiểm soát tính toàn vẹn ngữ nghĩa ít được chú ý hơn và nhìn chung không được hỗ trợ tốt bởi các hệ quản trị CSDL phân tán
- Kiểm soát dữ liệu đầy đủ phức tạp hơn và tốn kém hơn trong các hệ thống phân tán
 - Định nghĩa và lưu trữ các quy tắc (lựa chọn trạm)
 - Thiết kế các thuật toán thực thi giúp giảm thiểu chi phí truyền thông

48

Tài liệu tham khảo

- M. Tamer Özsu , Patrick Valduriez, “Principles of Distributed Database Systems”, Fourth Edition, Springer, 2020.
<https://link.springer.com/book/10.1007/978-3-030-26253-2>

49