

-Giao tác là gì: là 1 dãy các thao tác đọc ghi trên CSDL, cùng các phép toán: thêm, xóa,sửa có tính nhất quán để giải quyết các tình huống dữ liệu bị mất tính nhất quán khi có nhiều truy xuất đồng thời.

+Để viết giao tác cần bật dịch vụ: MSDTC (Microsoft Distributed Transaction Coordinator)

-Các loại giao tác :Giao tác phẳng & giao tác lồng

+**Giao tác phẳng:** chỉ có 1 lệnh Begin Transaction & 1 lệnh Commit Transaction

+**Giao tác lồng:** có thể nhiều lệnh Begin Transaction & Commit Transaction lồng với nhau

- **Có 4 tính chất giao tác:** nguyên tử, nhất quán, bền vững, biệt lập.

+**Tính nguyên tử:** 1 tập lệnh được xem như là 1 lệnh, thực hiện được hết hoặc không thực hiện lệnh nào (tính chất quan trọng nhất của giao tác).

+**Tính nhất quán:** nói lên tính đúng đắn của dữ liệu, sau khi xử lý thì chuyển từ trạng thái nhất quán này sang trạng thái nhất quán khác (có 3 độ)
[T ko đề lên dữ liệu rác của ~ g/tác khác, T ko đọc dữ liệu rác của ~ g/tác #]

+**Tính bền vững:** sau khi ủy thác thì dữ liệu k bị xóa khỏi database (để khôi phục dữ liệu)

+**Tính biệt lập:** khi 1 giao tác đang thực thi ko lộ ra các kết quả của nó cho ~ giao tác khác đang cùng hoạt động trước khi nó ủy thác.

+ Read Uncommitted

+ Read Committed

+ Repeatable Read

+ Serializable

+ Snapshot

- **Sự khác nhau giữa giao tác phân tán tập trung & giao tác phân tán**

+Giao tác phân tán: phân bố ở nhiều nơi. Khi 2 giao tác chạy ở 2 server nếu 1 giao tác gặp sự cố thì giao tác kia vẫn chạy bình thường

+Giao tác tập trung: tập trung chỉ ở 1 nơi. Khi giao tác gặp sự cố thì tất cả thao tác đều dừng.

Giao tác với giao tác phân tán giống và khác nhau ntn? Giao tác thì thực thi trên môi trường CSDL tập trung (giao tác phẳng, giao tác lồng), còn giao tác phân tán thì thực thi trên môi trường CSDLPT. Giống 4 t/chất giao tác

Tính trong suốt của 1 SP? Điều kiện để biết 1 SP trong suốt?

- SP trong suốt là SP khi chạy trên tất cả các phân mảnh phải cho ra các kết quả như nhau.

Điều kiện để viết 1 SP trong suốt là tên database ở các phân mảnh phải giống nhau, tên link server các server phải giống nhau

Có bao nhiêu mức trong suốt và người lập trình đứng ở mức nào?

4 mức, ng lập trình đứng ở mức 3 nhưng làm cho ng dùng cảm thấy ở mức 1
+ Trong suốt phân mảnh: ko cần chỉ rõ phân mảnh và vị trí cấp phát cho phân mảnh

+ Trong suốt vị trí: cung cấp tên phân mảnh nhưng ko cần chỉ ra vị trí của phân mảnh.

+ Trong suốt ánh xạ cục bộ: cung cấp tên phân mảnh và vị trí cấp phát

+ Không trong suốt

Trong suốt: khi ng sử dụng thực hiện truy vấn thì k có cảm giác dữ liệu đang phân tán hay tập trung

3 quy tắc phân mảnh:

+ Tính đầy đủ : dữ liệu nằm ít nhất ở 1 mảnh

+ Tính tái thiết đc: thỏa tính đầy đủ thì thỏa tính tái thiết, định nghĩa 1 phép toán để tái thiết dữ liệu: ngang là phép hội (union) , dọc là phép kết tự nhiên trên khóa chính.(join-on)

+ Tính tách biệt: dữ liệu chỉ nằm duy nhất ở 1 mảnh (tính chất này có thể vi phạm đc, table nào nhân bản thì vi phạm)

1. Phân mảnh ngang là gì? Phân mảnh dọc là gì? Phân mảnh ngang nguyên thủy là gì? (Khoa) Phân mảnh ngang dẫn xuất là gì?(giảng viên, lớp, sinh viên , điểm)

-Phân mảnh ngang (phép hội union) là chia (phân rã) 1 quan hệ theo các bộ. Mỗi phân mảnh ngang là tập con của quan hệ, có 2 loại phân mảnh ngang:

+Phân mảnh ngang nguyên thủy: 1 q/hệ dựa trên 1 vị từ định nghĩa trên 1 quan hệ đó. Vd: TENKHOA = “VT”

+P/mảnh ngang dẫn xuất: một q/hệ dựa trên 1 vị từ định nghĩa trên 1 q/hệ #.

VD: SINHVIEN.MALOP = LOP.MALOP

Vd: Chương trình QLDSV phân mảnh ngang theo khoa nên:

Khoa MAKH = “CNTT” => Khoa là phân mảnh ngang nguyên thủy.

Lớp MAKH = “CNTT” => Lớp là phân mảnh ngang nguyên thủy.

KHOA.MAKH = LOP.MAKH => Lớp là phân mảnh ngang dẫn xuất.

-Phân mảnh dọc (phép kết tự nhiên có khóa chính): phân mảnh dựa trên khóa chính của 1 quan hệ.

Điều kiện phân mảnh dọc : có khóa chính

Vì sao các phân mảnh dọc cần có khóa chính: Để tái thiết dc dữ liệu.

-Phân mảnh hỗn hợp : vừa ngang vừa dọc

Điều kiện để phân tán được: phải có dịch vụ SQL Server Agent.

12. Khi nào cần phải tạo link server? Khi tạo cần các ràng buộc gì?

-Tạo link server khi database nằm ở 2 server khác nhau , ràng buộc là: tên các server link tới nhau thì phải cùng tên.

Mục đích: cho phép truy cập dữ liệu từ server này sang server khác

-Các tạo: ví dụ tạo link ở server 1 để truy xuất dữ liệu ở server 2

- Đứng ở server 2 tạo login HTKN, đứng ở server 1 tạo LINK

- Số link : $n(n-1)$; n là số server phân mảnh, số tên link : n-1

Cú pháp: TENLINK.TENDATABASE.DBO.TENBANGTRUYXUAT

Các bước phân tán? Khi phân tán cần bật dịch vụ gì?

- Cấu hình các server:

Bước 1: cấu hình distribution

+ Publisher (server gốc) : chứa bản cơ sở dữ liệu gốc

+ Distributor: điều phối các phân mảnh phân tán

+ Subscriber: server chứa cơ sở dữ liệu sau khi phân tán

Bước 2: Tạo publication

- Định nghĩa Publication: 1 container chứa actual (view, sp, table, UDF), chứa lược đồ phân mảnh , thể hiện mối quan hệ dữ liệu giữa các table khi chúng ta phân tán dữ liệu.

- Định nghĩa Subcrition(database): 1 container nhận Publication (chỉ định cái server nào sẽ chứa cái publication, chứa dữ liệu phân tán theo lược đồ phân mảnh)

Khi phân tán cần bật dịch vụ: SQL Server Agent

KHOA: phân mảnh ngang nguyên thủy, còn lại là dẫn xuất ,

Môn học: là nhân bản, vì môn học có thể có ở khoa này cũng có thể có ở khoa khác

(Continuosly : tự động đồng bộ hóa)

+ UDF: hàm do người dùng tự định nghĩa

-Trong 1 server có nhiều user – tên user có thể trùng nhau

-Trong 1 database có nhiều user – tên user ko đc trùng nhau

-Trong 1 server có nhiều login- tên login ko đc trùng nhau

-1 login chỉ thuộc 1 user

1. Thuộc tính của folder RELPDATA là gì? ở đâu? Mục đích của folder này dùng để làm gì?

- Thuộc tính của folder RELDATA là share, có quyền read/write. Mục đích của folder là để chứa các dữ liệu trao đổi trong quá trình update dữ liệu từ các phân mảnh về CSDL gốc, và từ CSDL gốc đến các phân mảnh.

2. Trong lược đồ phân mảnh, số lượng database trong publication và số lượng database trong subscriber như thế nào?

-Số lượng database trong publication nhiều hơn trong subcrition vì trong publication chứa nhiều subcrition.

3. Khi chọn nào mà khoa đó xuất hiện, sự kiện gì làm điều đó? Lệnh nạp dữ liệu từ database về các datatable là gì? (vd khi nhập điểm thì lệnh gì nạp về): Sai đồ men (lệnh truy vấn lại) -----Sử dụng servername-----

4. Trong database cái nào là nhân bản, cái nào là phân hoạch?

-Môn học là nhân bản, còn lại là phân hoạch.

Loginname nằm trong table nào: Table sys.sysuser trong database đó.

5. Tại sao khi gõ login lại biết họ tên?

-Tại vì login và user liên hệ với nhau qua trường sid trong sys.systemlogin và sys.systemuser

Tại sao biết user liên kết với login nào? Username & loginname liên kết với nhau qua trường sid (trên user và login đều có sid) nên từ loginname biết được username từ sid.

Hàm suy từ loginname => username: SUSER.SID (tên login) => sid => username => userid => groupid => rolename.

Tên nhóm quyền nằm trong table nào? Sys.sysuser

Rowguid: là do hệ thống tự sinh ra, dùng để đồng bộ dữ liệu xuống phân mảnh, nó chỉ rõ table nào, cột nào để đưa dữ liệu xuống đúng vị trí đó.

Ưu khuyết điểm của nhân bản: truy xuất nhanh, đứng ở đâu cũng có thể select (chiều) được. Nhưng update chậm vì có quá nhiều bản sao.

Ưu khuyết điểm của phân hoạch: select chậm nhưng insert và update nhanh do chỉ thao tác trên server gốc hoặc server phân mảnh cần truy xuất.

6. Nếu phân tán xong CSDL, muốn thay đổi cấu trúc(cột) của table server gốc thì ta làm thế nào?

-Có thể thay đổi cột bằng cách viết SP dùng lệnh ALTER TABLE ở server gốc sau khi thêm xong thì đẩy xuống phân mảnh con.

7. Chuyển chữ thường thành chữ hoa : dùng lệnh upper

8.Đưa ra 2 cách xử lý truy vấn dữ liệu cùng lúc ở 2 server khác nhau?

-Dùng câu truy vấn song song cho nó chạy song song ở 2 server khác nhau.

9.Dữ liệu sau khi nhập ở các form sẽ đc đẩy về đâu?

-Đẩy về server publicsher sau đó đồng bộ xuống các Subscriber

10. Trong các table cái nào là tách biệt? cái nào là đầy đủ?

-Tất cả các table đều mang tính chất đầy đủ, các table nhân bản thì vị phạm tính tách biệt.

11. Subform là gì?

-Là form có ít nhất 2 bảng và có sự liên kết với nhau.

14. Quy trình thực hiện tối ưu hóa truy vấn trong CSDLPT?

-Chọn trước, chiếu trước, kết sau. Khử phép kết nếu được, khử kết bằng cách thay phép kết bằng phép truy vấn lồng.

+Tối ưu hóa các phép điều kiện bằng phép biến đổi tương đương

+ Chiếu : Select, chọn :from, điều kiện lọc : where, kết : join- on – where (điều kiện kết)

18. Câu lệnh bắt đầu & kết thúc của 1 giao tác là gì?

Begin distributed Transaction >> bắt đầu

Commit transaction >> Kết thúc

Bắt lỗi : Try (commit transaction) catch (Rollback transaction)

-**PGV**: đc quyền xem cả 2 khoa, và chỉ đc: một tài khoản duy nhất ở 1 phân mảnh nhưng đc quyền xem 2 phân mảnh (PGV)

- **Public static SqlConnection conn = new SqlConnection();**

Biến này dùng để chạy các SP và View, trong SQL nó như là cầu nối và chạy xuyên suốt trong chương trình.

-**ConnectionString**: thuộc tính quan trọng nhất vì chỉ khi mở kết nối kiểm tra xem chuỗi đó đúng hay sai

-**Lệnh Fill** : là lệnh tự động tải dữ liệu về

-**DropDownList**: trong cmb chỉ cho phép chọn, k đc ghi

-**Liên kết dữ liệu giữa 2 form**: biến thì nằm ở program, mà khi chạy thì nó nằm ở đăng nhập.

-**SelectedIndexChanged**: biến cố này tự động chạy khi ta chọn 1 giá trị mới trong cmb và khi nó chạy lấy ra servername nằm trong program.

-**DataSet**: là 1 đối tượng có thể chứa nhiều DataTable.

-**DataTable** dùng để lưu trữ 1 tập dữ liệu dạng bảng đc trích ra từ DataSource

-**DataTable khác với Table**: (1sp là 1 dataTable) là 1 đối tượng mà dữ liệu của nó là 1 phần của table.

- **Rowguid** : Tự sinh ra để đồng bộ hóa dữ liệu.

3. Có mấy loại Join: 3 loại

+ INNER JOIN

+ OUTER JOIN : Left Join & Right Join

+ CROSS JOIN: ko có khóa nó gộp lại nhưng vẫn lấy

INNER JOIN trả về các bản ghi có giá trị phù hợp giữa hai bảng

LEFT JOIN mọi bản ghi bảng bên trái trả về kết hợp với bản ghi phù hợp với bên phải nếu có (nếu không có thì nhận NULL)

RIGHT JOIN mọi bản ghi bảng bên phải trả về có kết hợp với giá trị phù hợp nếu có ở bảng trái

OUTER JOIN (full join) mọi bản ghi ở bảng trái và bảng phải kết hợp lại

-Nhân bản là gì? ở server nào cũng có dữ liệu như nhau .Lí do nhân bản? Nhân bản ko hiệu quả với lệnh gì?

- Các hàm lấy dữ liệu: dataSet và show dữ liệu: datagridview

-Lệnh select...union : dùng để lấy dữ liệu từ các server phân mảnh về

Vd: SELECT MALOP,TENLOP FROM DBO.LOP

UNION SELECT MALOP,TENLOP FROM LINK.QLDSV.DBO.LOP

-Đơn giản hóa 1 b/thức đại số quan hệ được t/hiện dựa trên các tiêu chuẩn :

+Tiêu chuẩn 1: Sử dụng phép biến đổi tương đương, phép chọn và chiếu

+Tiêu chuẩn 2:Thực hiện các phép càng xa càng tốt, tức là đẩy các phép chọn và các phép chiếu xuống phía dưới càng xa càng tốt.

+Tiêu chuẩn 3: Khi các phép chọn đc thực hiện sau 1 phép tích thì kết hợp các phép toán này để tạo thành 1 phép kết.

+Tiêu chuẩn 4: Kết hợp chuỗi các phép toán 1 ngôi liên tiếp nhau áp dụng cho 1 quan hệ toán hạng

+Tiêu chuẩn 5: Khi phát hiện các biểu thức con truy vấn, áp dụng các phép biến đổi tương đương để đơn giản hóa biểu thức truy vấn

- Truy vấn (query) là 1 biểu thức đc biểu diễn bằng 1 ngôn ngữ và dùng để xác định 1 phần dữ liệu đc chứa trong CSDL.

- Phân quyền:

Server 1+2: PGV : datareader, writer, securityadmin,access admin

KHOA: datareader, writer

USER: datareader, writer

Server 3: PKTOAN: accessadmin, reader, writer, securityadmin

-db_accessadmin: ng dùng có quyền quản lý tài khoản SQL server đăng nhập

-db_securityadmin: ng dùng có thể chỉnh sửa vai trò (role) và quản lý các bậc quản lý phân quyền khác

+ muốn tạo tài khoản thì dùng accessadmin + securityadmin (cấp quyền)

-db_datareader: ng dùng có thể đọc được toàn bộ dữ liệu.

-db_datawrite: ng dùng có quyền thêm, xóa, sửa dữ liệu trong bảng

-db_owner: toàn bộ ng dùng có quyền full-access

-sysAdmin: toàn bộ ng dùng đều có thể t/hiện các thao tác trên server.

-Có 3 loại server

+Local server: cài máy nào thì chạy máy đó

+Remote server: điều khiển từ xa giữa 2 máy

+ Link server : Liên kết server này với server khác.

-Dùng link cơ chế nối từ sql server đến database của server khác SQL server hay hệ thống khác

-Tạo link thì tạo remotelogin(HTKN) trước

1. DeadLock là gì ? Các xử lý?

- Một cv B gọi A: Nhưng A bị lỗi nên cv B cứ đợi A-> gọi là DeadLock

- Xử lý: thiết lập Timeout , sau bao lâu đó thì trở về trạng thái ban đầu

Vị từ là gì ? Thực chất là phép chọn trên 1 feild trên 1 giá trị const

2. Dữ liệu rác là gì ? Là dữ liệu chưa được ủy thác