

Các tác vụ quản trị hệ thống

Nội dung

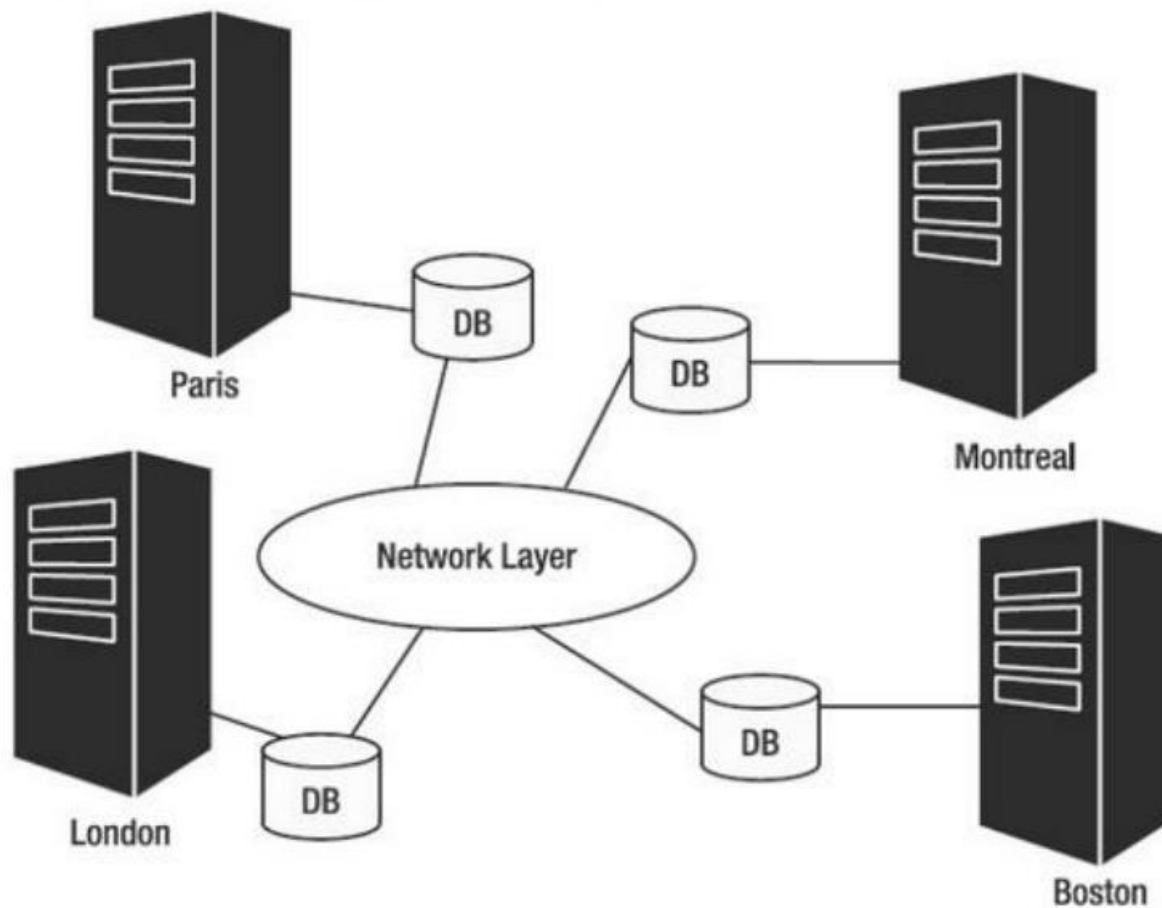
- Phân quyền và bảo mật
- Sao lưu và phục hồi
- Chuyển đổi giữa các loại CSDL
- Kiến trúc nhân bản

Nhân bản

Khái niệm

- Nhân bản dữ liệu (replication) là kỹ thuật để:
 - Tạo ra những bản sao dữ liệu giống hệt nhau
 - Di chuyển các bản sao này đến những vùng khác nhau
 - Đồng bộ hóa dữ liệu một cách tự động để tất cả các bản sao có cùng giá trị dữ liệu
- Nhân bản có thể thực thi giữa những CSDL:
 - Trên cùng một server
 - Những server khác nhau được kết nối bởi mạng LANs, WANs hay Internet

Khái niệm (tiếp)



Ưu điểm của nhân bản dữ liệu

- Dữ liệu được lưu trữ ở nhiều nơi, hiệu quả trong việc có nhiều ứng dụng cùng truy cập, khai thác
- Giảm thiểu xung khắc do số lượng lớn các giao dịch trên mạng
- Tăng tính sẵn sàng: khi server mặc định bị hư hoặc bảo dưỡng, có thể thay thế bằng một server khác
- Có thể kết hợp nhân bản và đồng bộ máy chủ CSDL cho các nền tảng HQT CSDL khác ngoài MSSQL Server (Microsoft OLE DB Provider for ODBC, Microsoft OLE DB Provider for Jet, Microsoft OLE DB for Oracle)

Yêu cầu trong nhân bản

- Trong suốt:

- Các client sẽ không biết đến sự hiện diện của các replica
- Khi client được kết nối vào 1 server để truy xuất dữ liệu thì dữ liệu đó được xem như là duy nhất
- Khi client gửi yêu cầu xử lý đến server thì client chỉ xem như xử lý trên một bản trong khi thực tế nó được xử lý trên nhiều nhân bản (replica) ở mức vật lý

- Nhất quán:

- Khi các client gửi cùng 1 yêu cầu đến các server khác nhau thì quản lý nhân bản phải đảm bảo tất cả các client sẽ nhận về cùng 1 kết quả.

Mục tiêu chính của nhân bản

- SQL Server đưa ra nhiều cơ chế nhân bản để đáp ứng các yêu cầu khác nhau của ứng dụng.
- Mỗi cơ chế cung cấp các khả năng và thuộc tính khác nhau nhằm đạt đến mục tiêu:
 - Sự nhất quán dữ liệu
 - Tính độc lập Site

Nhất quán dữ liệu (Data consistency)

- Có hai cách để đạt được tính nhất quán dữ liệu:
 - Nhất quán giao dịch (Transactional Consistency)
 - Hội tụ dữ liệu (Data convergence)

Nhất quán giao dịch

- Bảo đảm tất cả dữ liệu giống nhau tại mọi site ở bất kì thời điểm nào
- Tất cả các giao dịch thực hiện tại một site duy nhất
- Có 2 loại:
 - Nhất quán lập tức (Immediate Transactional Consistency hay Tight consistency)
 - Nhất quán ngầm (Latent Transactional Consistency hay Loose Consistency)

Nhất quán giao dịch (tiếp)

- **Nhất quán lập tức:**

- Tất cả các site được đảm bảo là luôn thấy cùng một giá trị dữ liệu tại cùng một thời điểm
- Sử dụng 2-phase commit protocol giữa tất cả các site tham gia:
 - Mỗi site phải commit đồng thời mọi thay đổi hoặc không site nào commit những thay đổi
 - Không khả thi khi số lượng site quá lớn

Nhất quán giao dịch (tiếp)

- **Nhất quán ngầm:**

- Các Site tham gia nhất quán ngầm rằng có một sự trì hoãn trong việc phản ánh các giá trị dữ liệu đến các site tham gia, vào lúc này các site không đảm bảo có cùng giá trị dữ liệu.
- Việc sửa đổi các giá trị dữ liệu có thể bị trì hoãn đủ lâu để tất cả các site cùng cập nhật, sau đó tất cả các site sẽ có cùng giá trị dữ liệu.

Hội tụ dữ liệu

- User có thể tự do thao tác trên các site theo các cách khác nhau. Khi các node đồng bộ, tất cả các site sẽ hội tụ về cùng một giá trị.
- Nếu mạng do có sự sửa đổi cùng một dữ liệu tại những site khác nhau thì những sửa đổi này sẽ được giải quyết một cách tự động:
 - Chọn site có độ ưu tiên cao hơn
 - Hay chọn site đưa sửa đổi đến trước

Độc lập site (site autonomy)

- Độc lập site xét đến ảnh hưởng của những thao tác trên một site đến các site khác
- Thường độc lập site càng cao thì tính nhất quán dữ liệu giảm
 - Nhân bản kết hợp (merge replication) có mức độc lập site cao nhất, nhưng không đảm bảo nhất quán dữ liệu
 - Two phase commit có tính nhất quán dữ liệu cao nhưng lại không có tính độc lập site

Kiến trúc nhân bản

Các thành phần chính của nhân bản

- **Publisher:** là server cung cấp dữ liệu để nhân bản đến các server khác.
 - Một Publisher có thể thiết lập nhiều bộ dữ liệu nhân bản (gọi là publication)
- **Subscriber:** là server nhận dữ liệu nhân bản và nhận các tác vụ cập nhật.
 - Quá trình cập nhật dữ liệu ở subscriber không giống như ở Publisher
 - Một Subscriber có thể là một Publisher của các Subscriber khác.

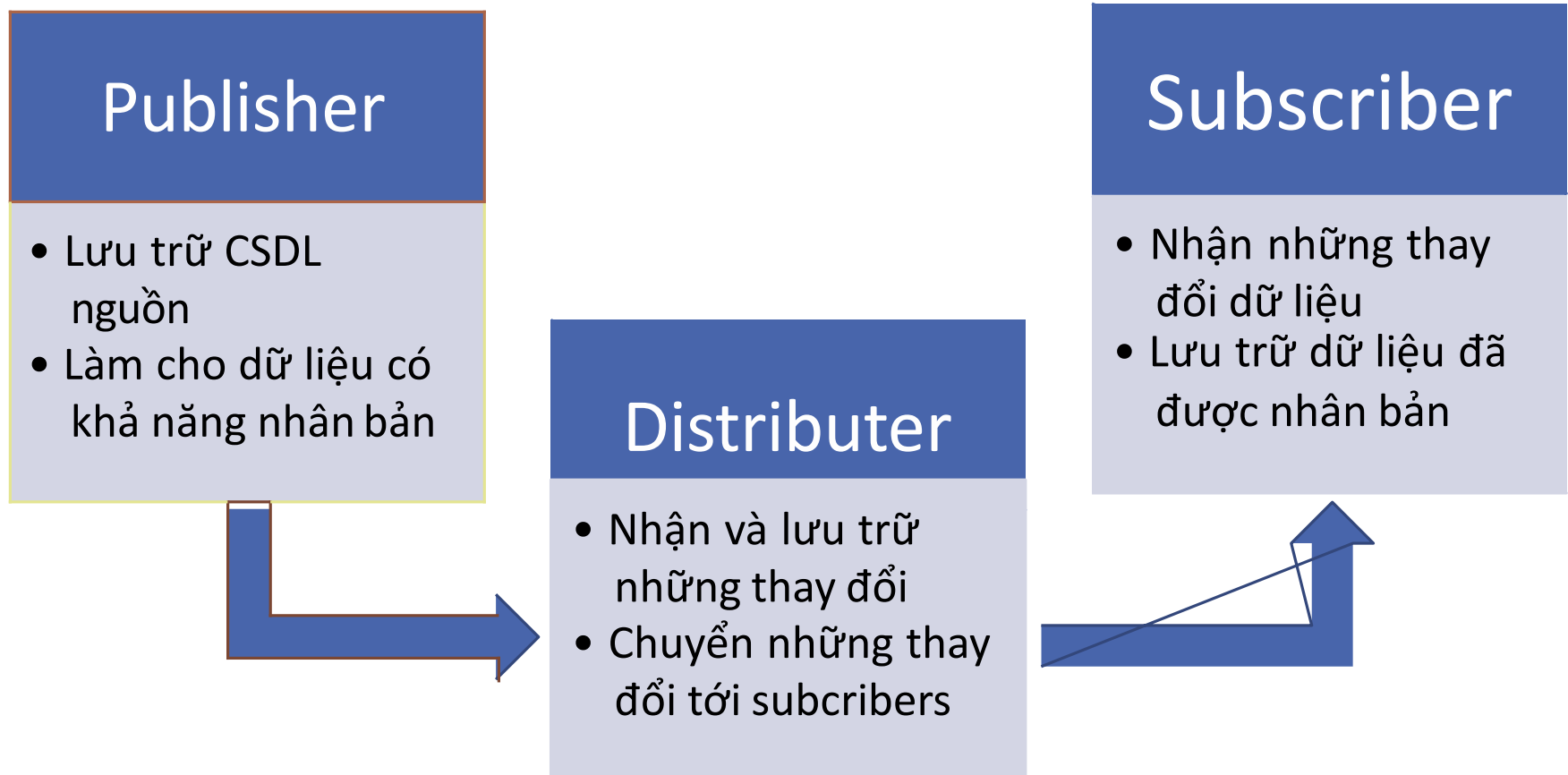
Các thành phần chính của nhân bản (tiếp)

- **Distributer:** là server chứa CSDL phân tán (distribution database) và lưu trữ metadata, history data và transaction.
 - SQL Server sử dụng CSDL phân tán để lưu và chuyển dữ liệu nhân bản từ Publisher đến các Subscriber.
 - Có 2 loại distributor:
 - Local distributor: là server được cấu hình là Publisher và Distributor
 - Remote Distributor: là server tách rời khỏi Publisher và được cấu hình là Distributor

Các thành phần chính của nhân bản (tiếp)

- **Article:** là một bảng, tập dữ liệu hoặc đối tượng của CSDL cấu hình để nhân bản.
 - Một article có thể chỉ bao gồm một vài hàng hay một vài cột của một bảng
- **Publication:** là một tập hợp gồm một hoặc nhiều article.

Các thành phần chính của nhân bản (tiếp)



Chiều di chuyển dữ liệu

- Có 2 kiểu di chuyển dữ liệu
 - Push subscription
 - Pull subscription
- Một publication có thể sử dụng cả hai kiểu di chuyển dữ liệu

Chiều di chuyển dữ liệu (tiếp)

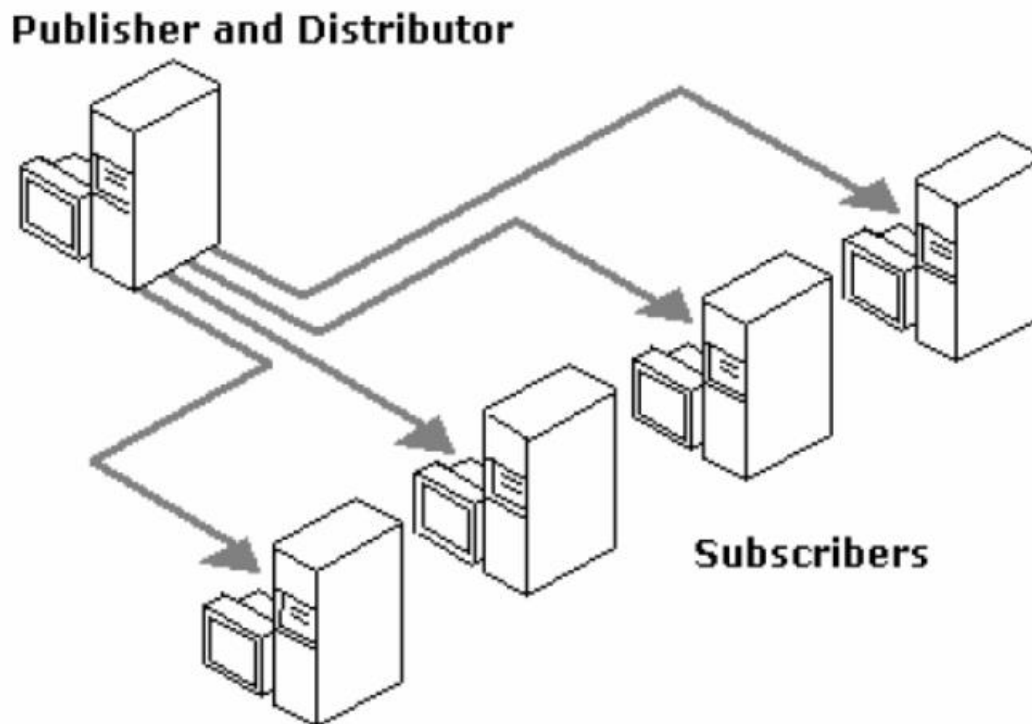
- **Push subscription:** Publisher đẩy những thay đổi đến Subscriber mà không quan tâm Subscriber có cập nhật hay không
 - Pull subscription được sử dụng trong những ứng dụng yêu cầu gửi những thay đổi đến subscriber ngay khi những thay đổi này xảy ra ở publisher
 - Giúp việc quản lý các subscriber đơn giản và tập trung, bảo mật tốt hơn
 - Distributor có thể phải đảm nhận nhiều quá trình phân bổ subscription đến các subscriber cùng một lúc, dễ dẫn đến hiện tượng thắt cổ chai.

Chiều di chuyển dữ liệu (tiếp)

- **Pull subscription:** subscriber kéo những thay đổi tại publisher về theo một khoảng thời gian định kỳ
 - Tốt cho những user độc lập thay đổi bởi vì chúng cho phép user xác định khi nào thì những thay đổi dữ liệu được đồng bộ
 - Pull subscription bảo mật thấp nhưng cho phép số lượng subscriber cao hơn

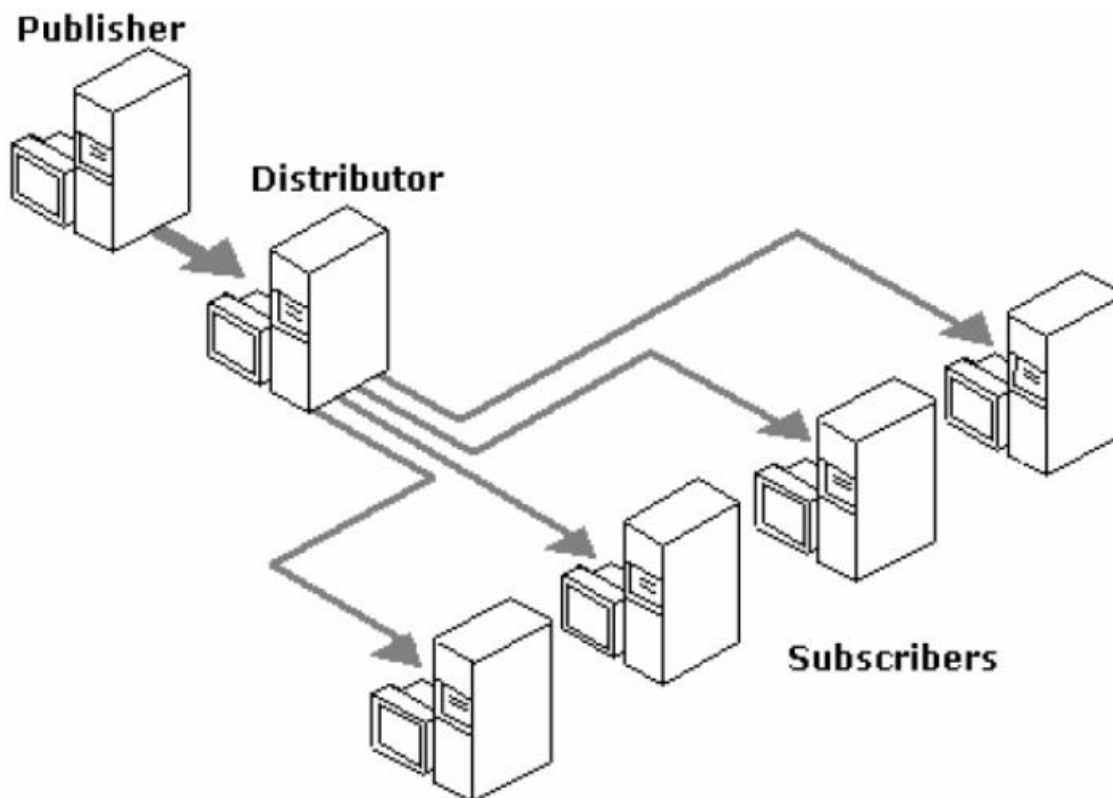
Những mô hình nhân bản dữ liệu cơ bản

- **Central publisher:** là mô hình Publisher và distributor thiết lập trên một máy.



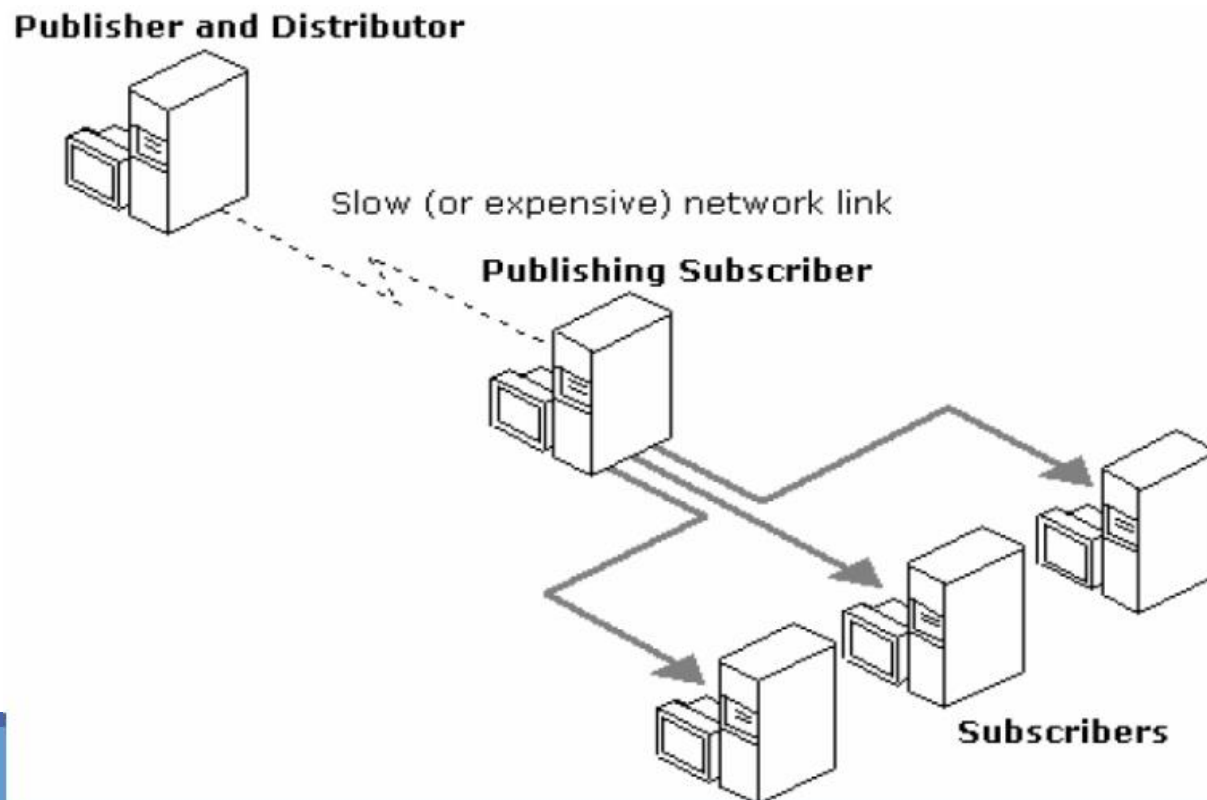
Những mô hình nhân bản dữ liệu cơ bản (tiếp)

- Publisher và Distributor không thiết lập trên một máy.



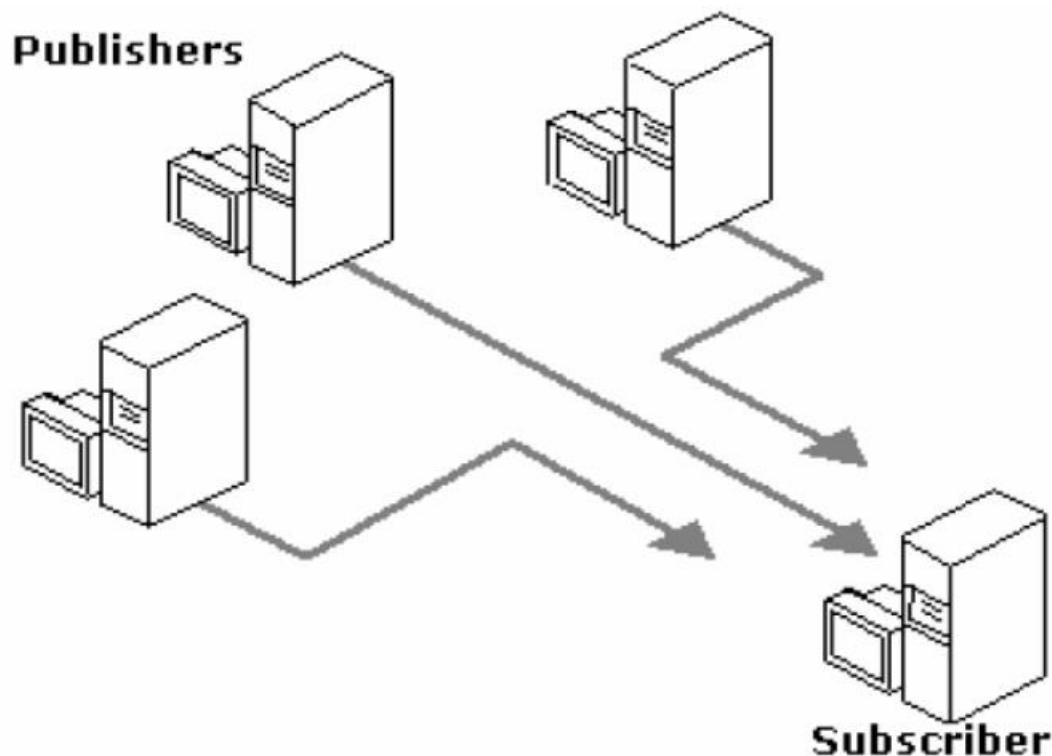
Những mô hình nhân bản dữ liệu cơ bản (tiếp)

- **Republisher:** là mô hình Publisher xuất bản dữ liệu đến Subscriber, sau đó Subscriber được thiết lập là Publisher xuất bản dữ liệu đến các Subscriber khác



Những mô hình nhân bản dữ liệu cơ bản (tiếp)

- **Central Subscriber:** là mô hình Subscriber thiết lập nhận dữ liệu xuất bản từ nhiều Publisher



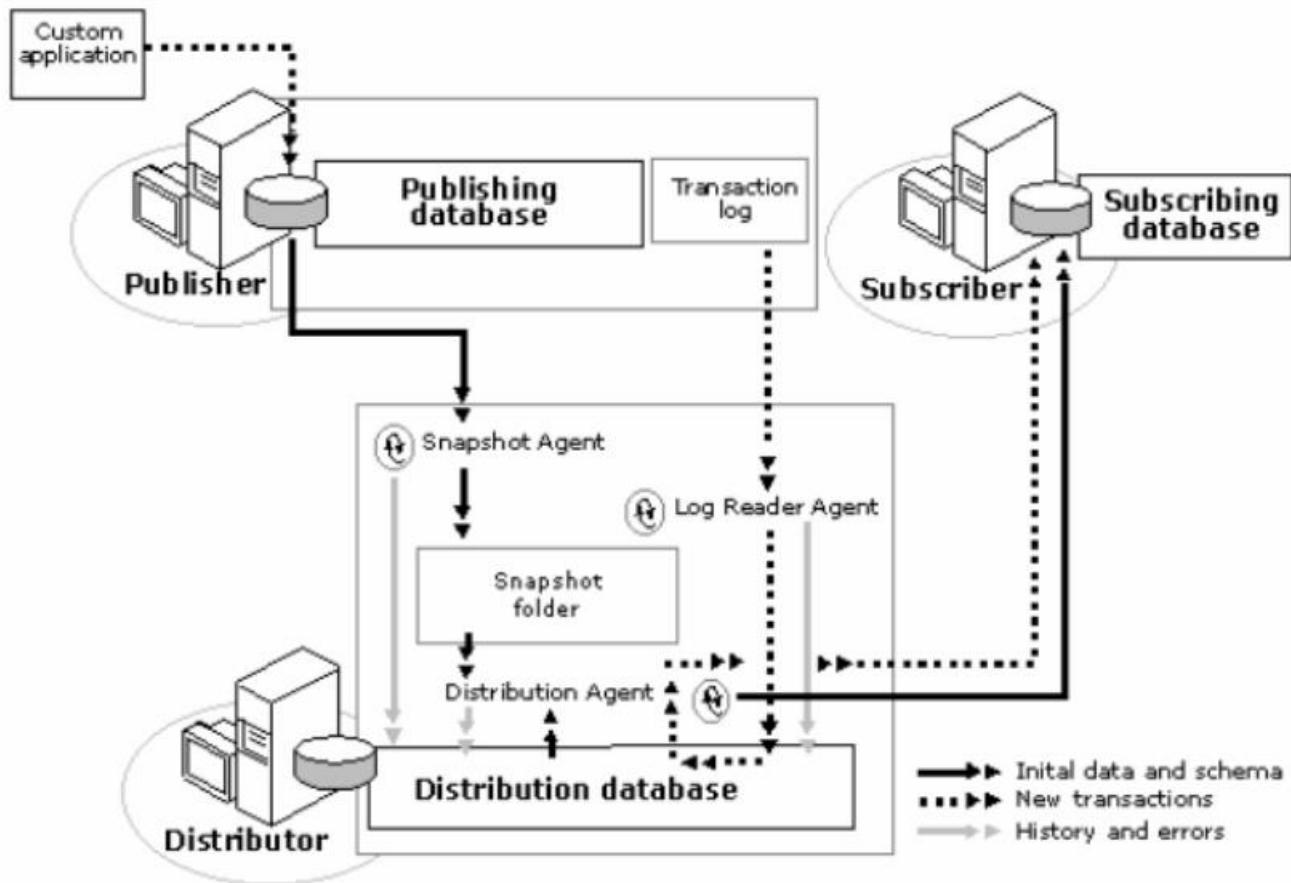
Những kiểu nhân bản dữ liệu

- Có 3 kiểu nhân bản dữ liệu:
 - Snapshot replication
 - Transaction replication
 - Merge replication

Snapshot replication

- Là kiểu nhân bản thực hiện sao chép toàn bộ dữ liệu cần nhân bản (còn gọi là quá trình làm tươi dữ liệu) từ Publisher đến các Subscriber
- Đảm bảo sự nhất quán tiềm ẩn giữa Publisher và Subscriber
- Gửi toàn bộ dữ liệu đến Subscriber thay vì chỉ gửi những thay đổi nên cần hệ thống mạng đủ mạnh để truyền dữ liệu lớn.
- Thường sử dụng khi
 - Dữ liệu thường là tĩnh, ít thay đổi
 - Nhân bản số lượng dữ liệu nhỏ

Snapshot replication (tiếp)



Nhân bản snapshot(Snapshot replication)

Snapshot replication (tiếp)

- Snapshot agent chuẩn bị những snapshot file (*snapshot file là file sao chép lược đồ và dữ liệu của những table phân bố*) chứa lược đồ và dữ liệu của những table phân bố, lưu những file này vào snapshot folder trên Distributor và ghi lại những công việc đồng bộ trong CSDL phân bố (distribution database).
- Distribution agent gửi những snapshot job (tác vụ sao chép dữ liệu) giữ trong bảng dữ liệu phân bố đến Subscriber.

Nhân bản giao dịch (transactional replication)

- Là kiểu nhân bản bắt đầu bằng nhân bản snapshot, sau đó sẽ thực hiện nhân bản các giao dịch dữ liệu theo các sự kiện insert, update, delete trong các table hay những thay đổi liên quan đến việc thực hiện stored procedure, index view.
- Các article trong một publication có thể là một phần hay một table, một hay nhiều stored procedure

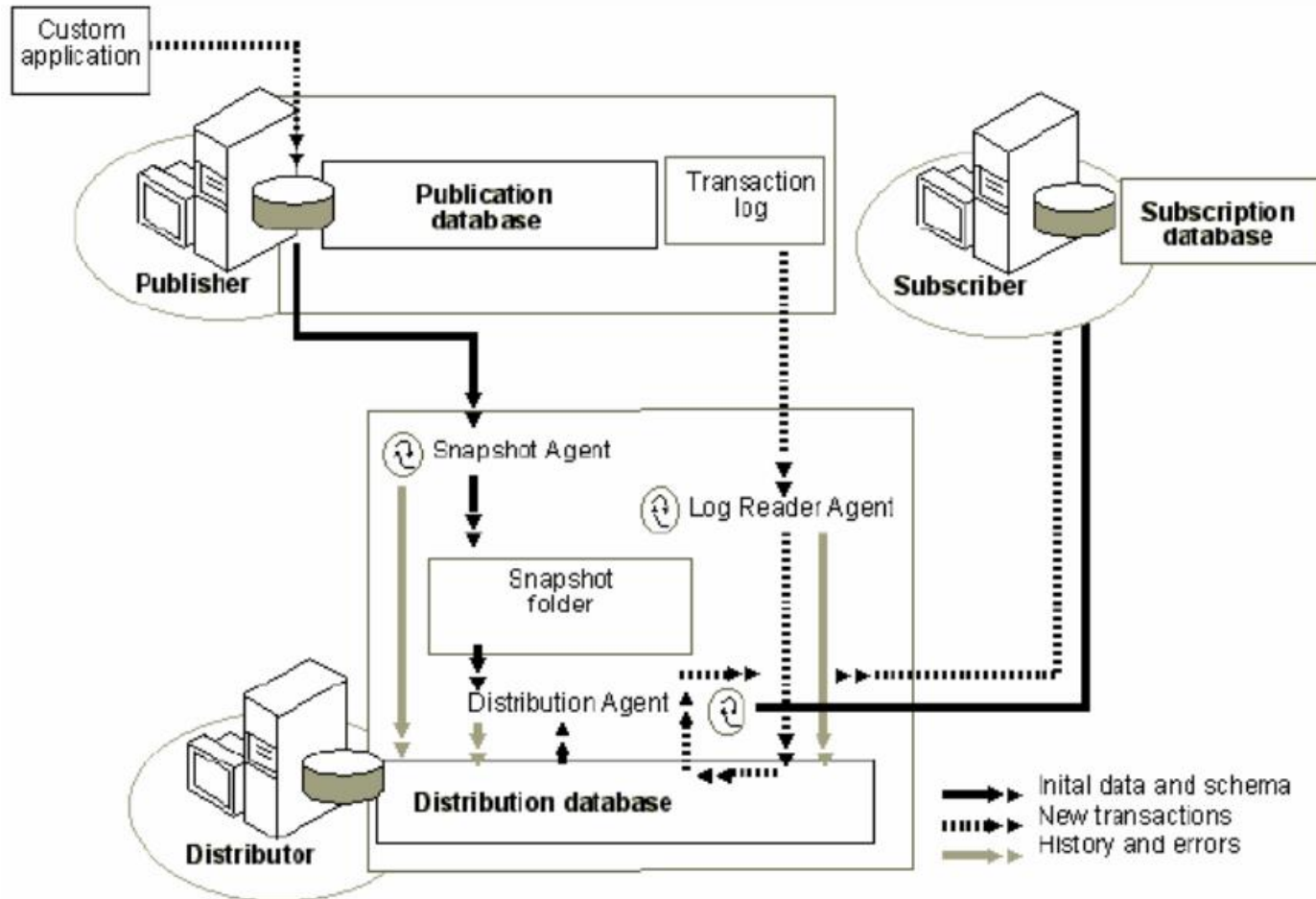
Nhân bản giao dịch (transactional replication) (tiếp)

- Nhân bản giao dịch sử dụng transaction log để giữ những thay đổi được làm trên dữ liệu trong một article.
- SQL Server giám sát những sửa đổi trên dữ liệu và lưu những thay đổi đó lên CSDL phân bố (distribution database).
- Những thay đổi đó sẽ được gửi đến Subscriber và tuân theo một trật tự nhất định.
- Những thay đổi được làm tại Publisher được chảy đến một hay nhiều các Subscriber theo những khoảng thời gian định trước

Nhân bản giao dịch (transactional replication) (tiếp)

- Nhân bản kiểu này được thực hiện khi:
 - Muốn sửa dữ liệu được xuất bản chuyển đến Subscriber, thời gian thực hiện theo giây, hoặc tức thời
 - Cần giao dịch trên toàn bộ hệ thống nhân bản dữ liệu (dữ liệu có thể chuyển đến tất cả các Subscriber hoặc không chuyển đến Subscriber nào)
 - Subscriber thường xuyên kết nối với Publisher

Nhân bản giao dịch (transactional replication) (tiếp)



Nhân bản giao dịch (transactional replication) (tiếp)

- Log Reader agent giám sát transaction log của mỗi CSDL được thiết lập để nhân bản và sao chép những transaction cần nhân bản từ transaction log vào CSDL phân bố (distribution database)
- Distribution agent di chuyển những transaction và những tác vụ khởi tạo snapshot được giữ trong table của CSDL phân bố.

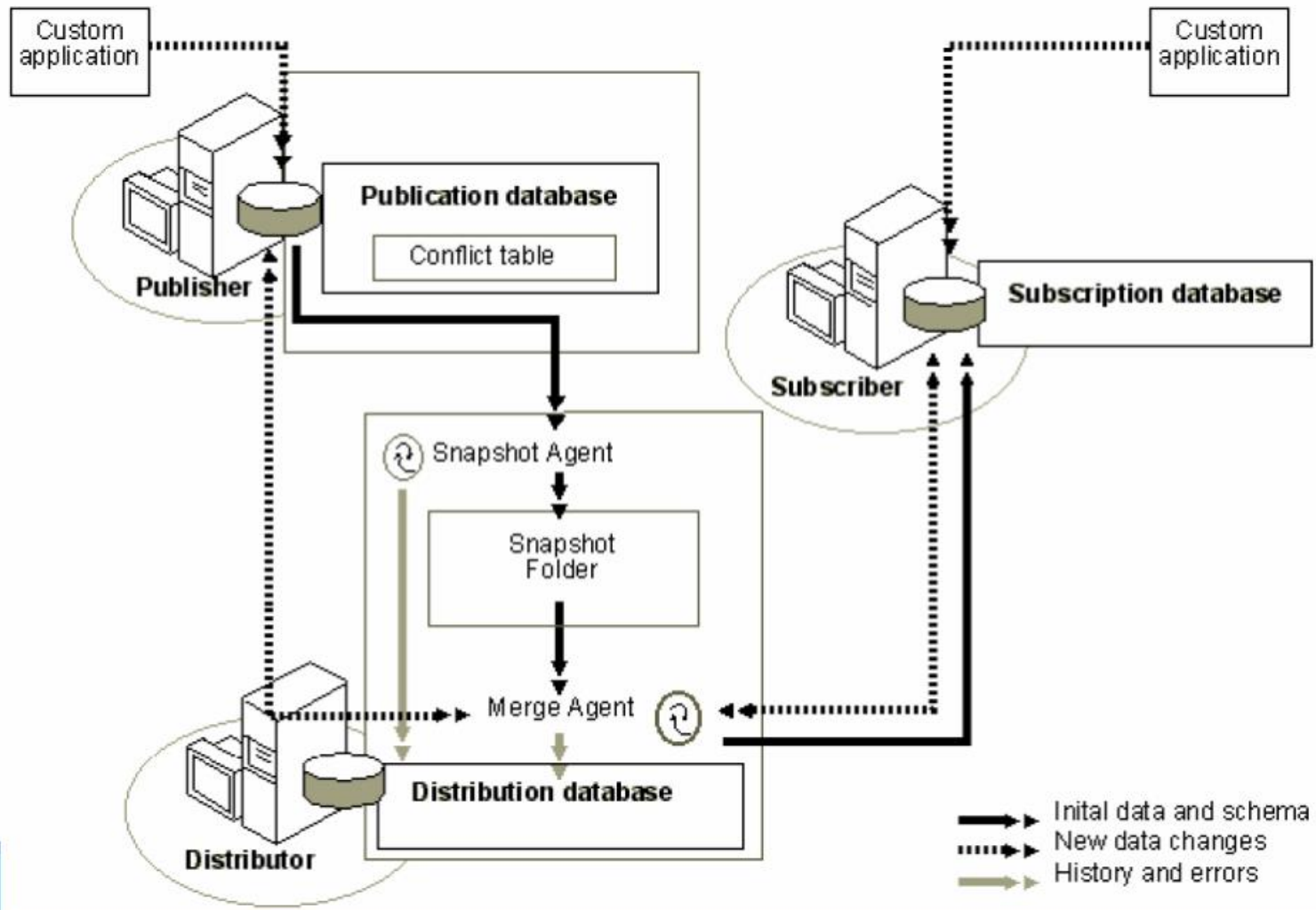
Nhân bản kết hợp (Merge replication)

- Là kiểu nhân bản có tính độc lập site cao nhất. Publisher và Subscriber có thể làm việc hoàn toàn độc lập và sẽ kết nối với nhau theo lịch hoặc theo yêu cầu để hội tụ các kết quả lại.
- Nếu đụng độ xảy ra bởi các site cùng sửa đổi trên cùng một phần tử dữ liệu thì những đụng độ này sẽ được giải quyết tự động:
 - Chọn site có độ ưu tiên cao hơn
 - Hay Site sửa đổi dữ liệu trước
- Các xung đột này có thể được phát hiện và giải quyết theo cấp độ hàng hay cột của bảng dữ liệu.

Nhân bản kết hợp (Merge replication) (tiếp)

- Kiểu nhân bản này thực hiện khi:
 - Nhiều Subscriber có nhu cầu cập nhật dữ liệu và chuyển dữ liệu cập nhật đến Publisher hoặc Subscriber khác
 - Subscriber yêu cầu nhận hoặc chuyển dữ liệu khi offline, đồng bộ dữ liệu với các Subscriber và Publisher sau.

Nhân bản kết hợp (Merge replication) (tiếp)



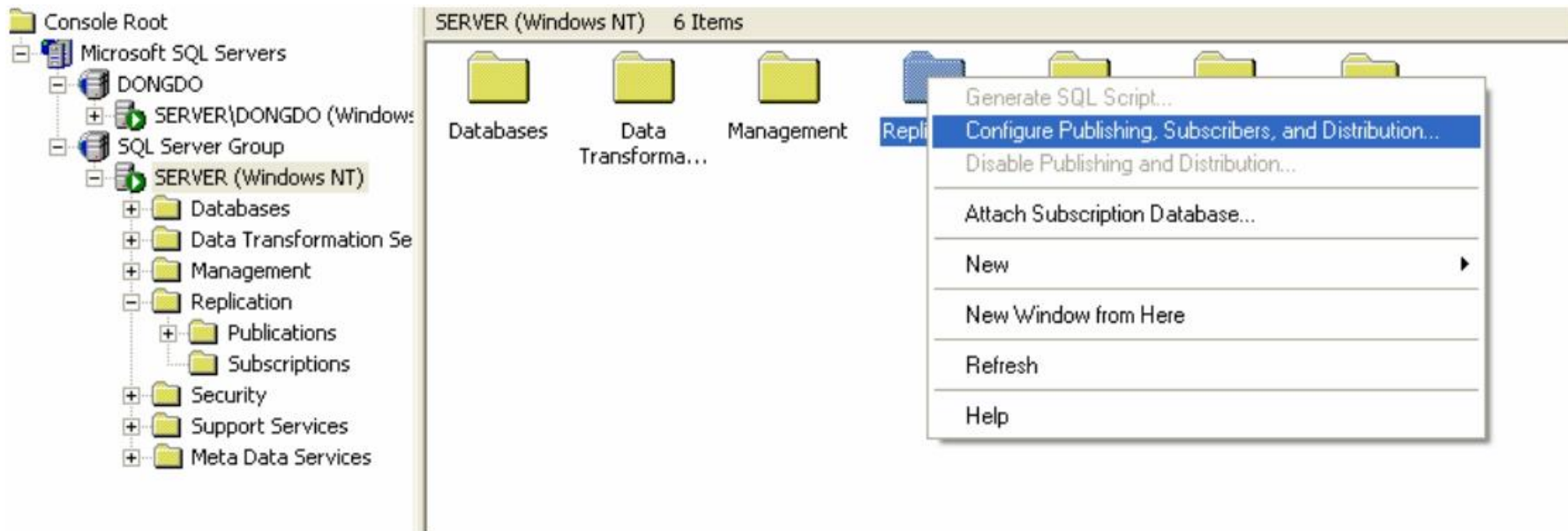
Nhân bản kết hợp (Merge replication) (tiếp)

- Snapshot agent chuẩn bị những snapshot file chứa lược đồ và dữ liệu của những table phân bố, lưu những file này vào snapshot folder trên Distributor và ghi lại những công việc đồng bộ trong Publication.
- Merge agent thực hiện những công việc khởi tạo snapshot được tổ chức trong bảng của publication đến subscriber.
- Merge agent cũng kết hợp những dữ liệu thay đổi xảy ra tại Publisher sau khi initial snapshot được tạo và giải quyết tranh chấp theo những luật mà người dùng đặt ra hay sử dụng bộ giải quyết tranh chấp.

Cấu hình Publisher và Distributor

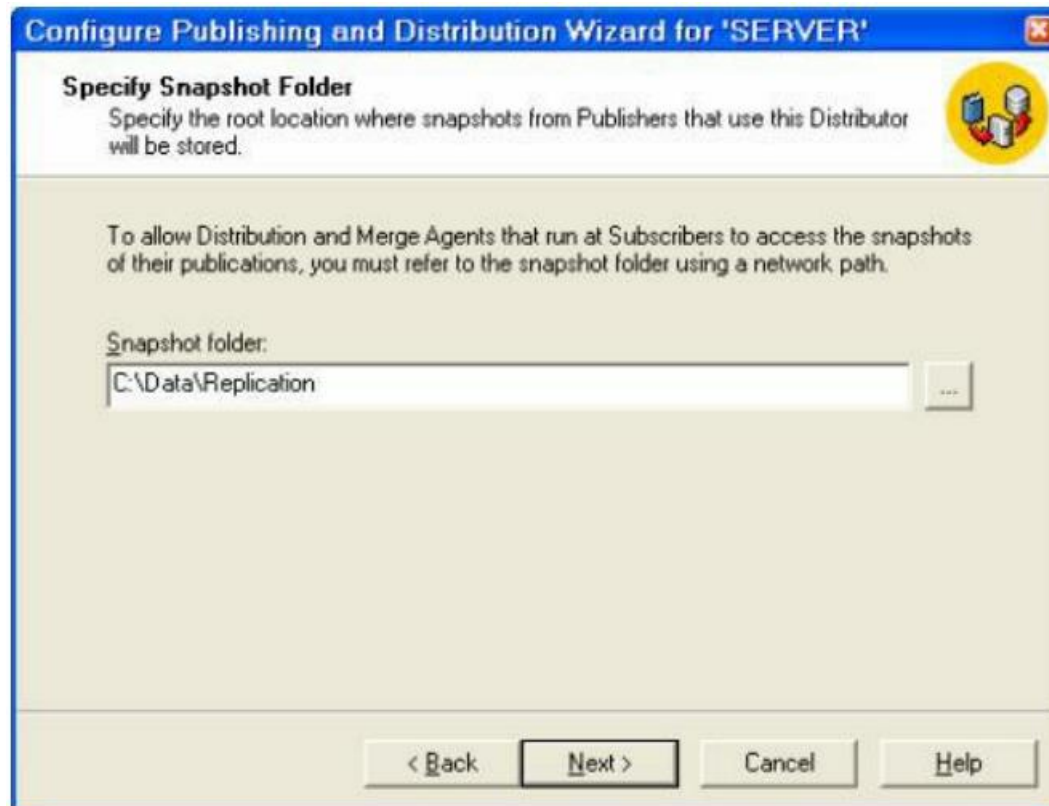
Cấu hình Publisher và Distributor

- Thực hiện chạy dịch vụ SQL Server Agent trong chức năng Service Manager
- Chọn Server cần cấu hình -> Replication
- Nhấn phải chuột -> Configure Publishing Subscription and Distribution...



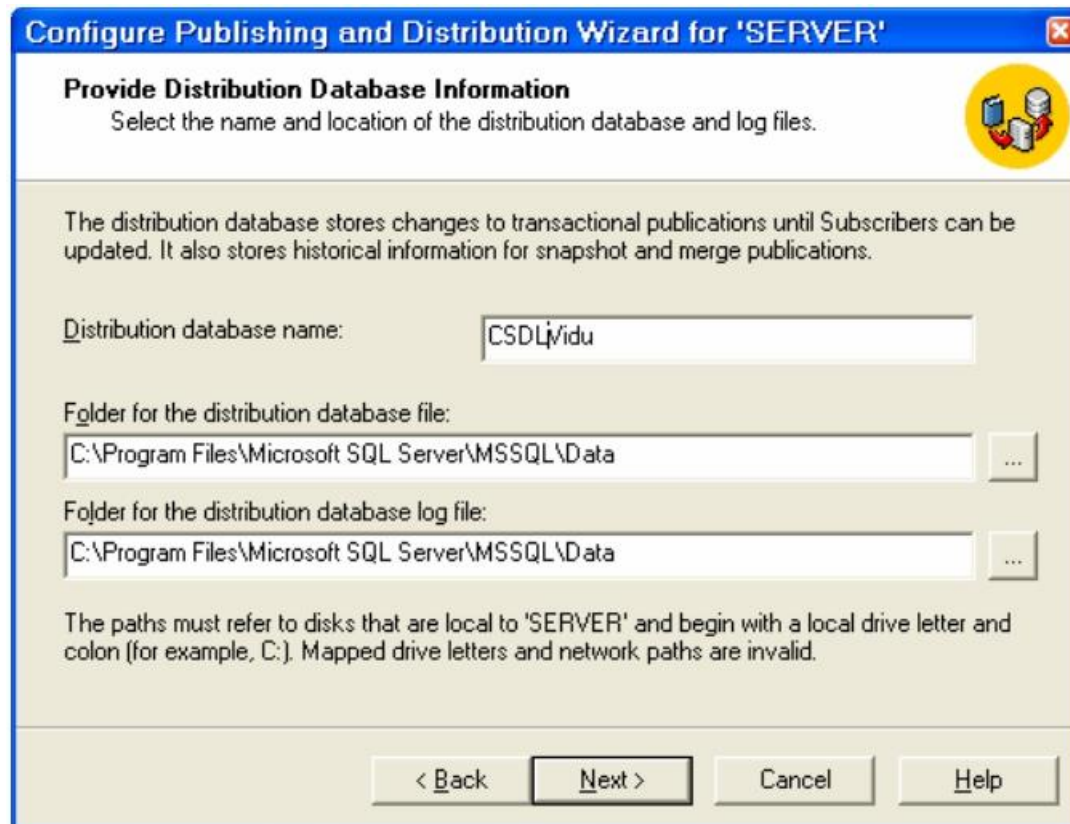
Cấu hình Publisher và Distributor

- Chọn thư mục snapshot



Cấu hình Publisher và Distributor

- Đặt tên CSDL của Distribution



Configure Publishing and Distribution Wizard for 'SERVER'

Provide Distribution Database Information
Select the name and location of the distribution database and log files.

The distribution database stores changes to transactional publications until Subscribers can be updated. It also stores historical information for snapshot and merge publications.

Distribution database name:

Folder for the distribution database file:
 ...

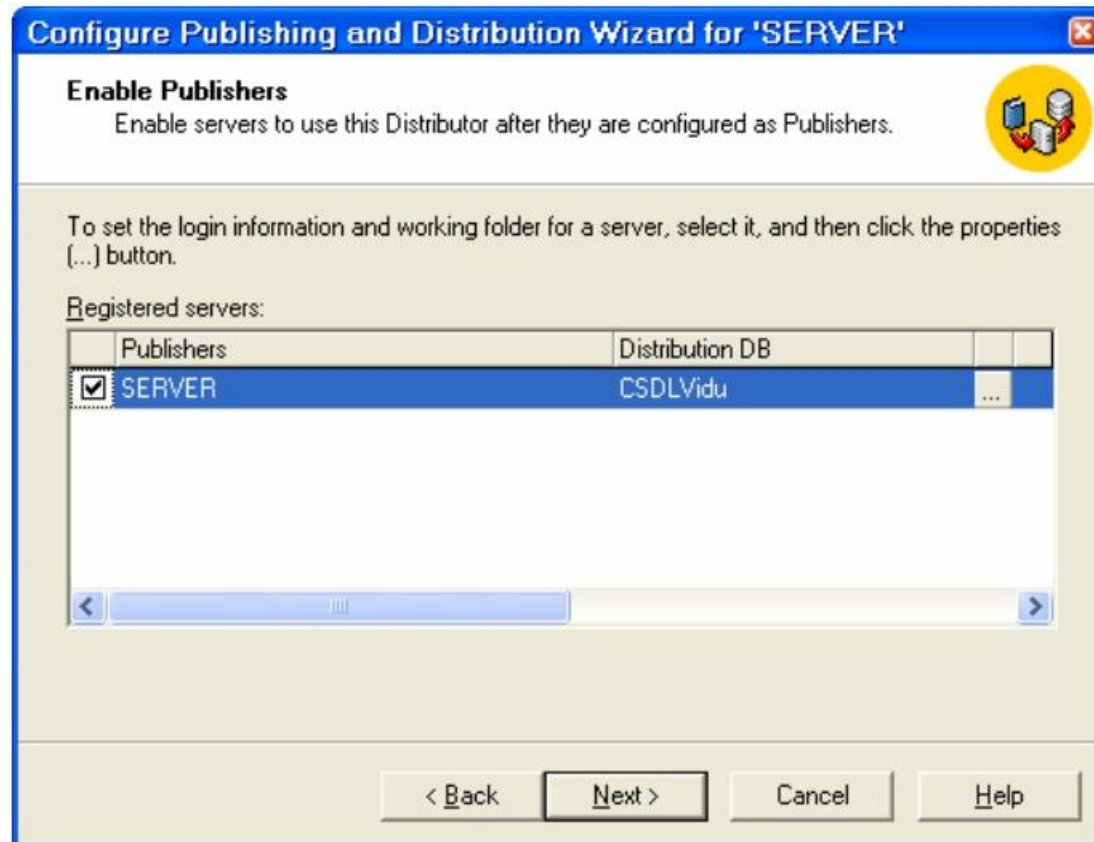
Folder for the distribution database log file:
 ...

The paths must refer to disks that are local to 'SERVER' and begin with a local drive letter and colon (for example, C:). Mapped drive letters and network paths are invalid.

< Back Next > Cancel Help

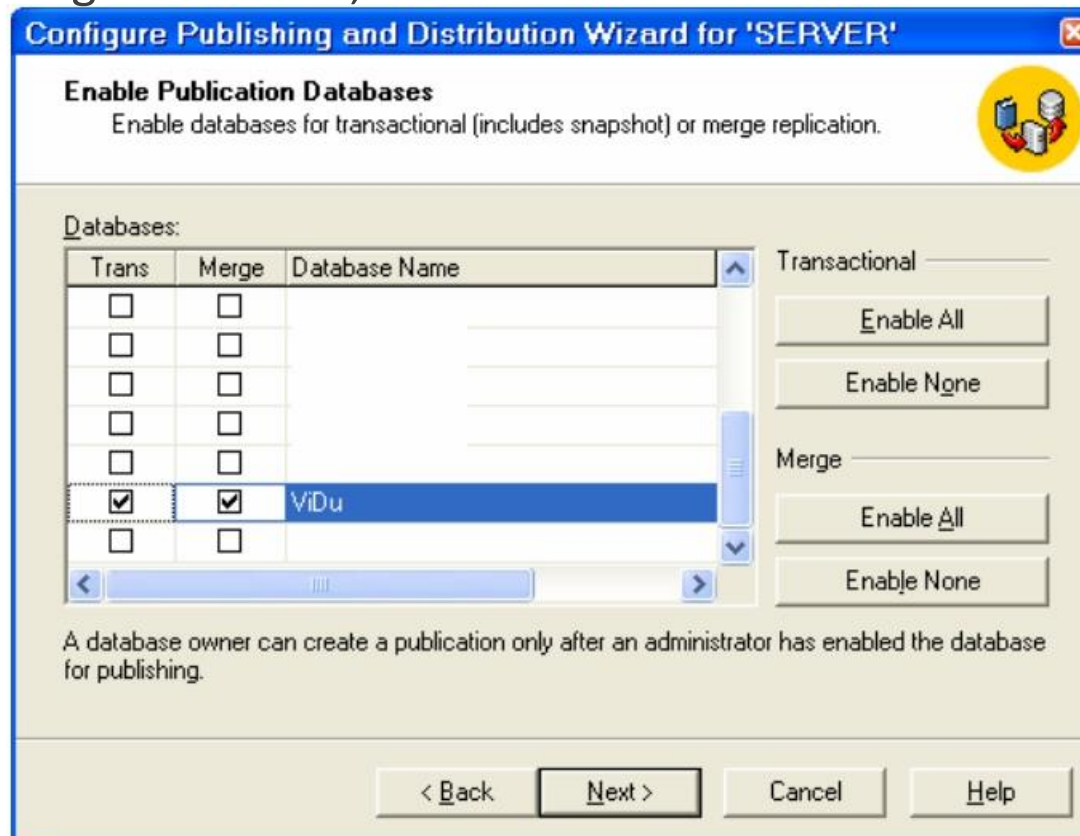
Cấu hình Publisher và Distributor

- Chọn server cấu hình thành Publisher



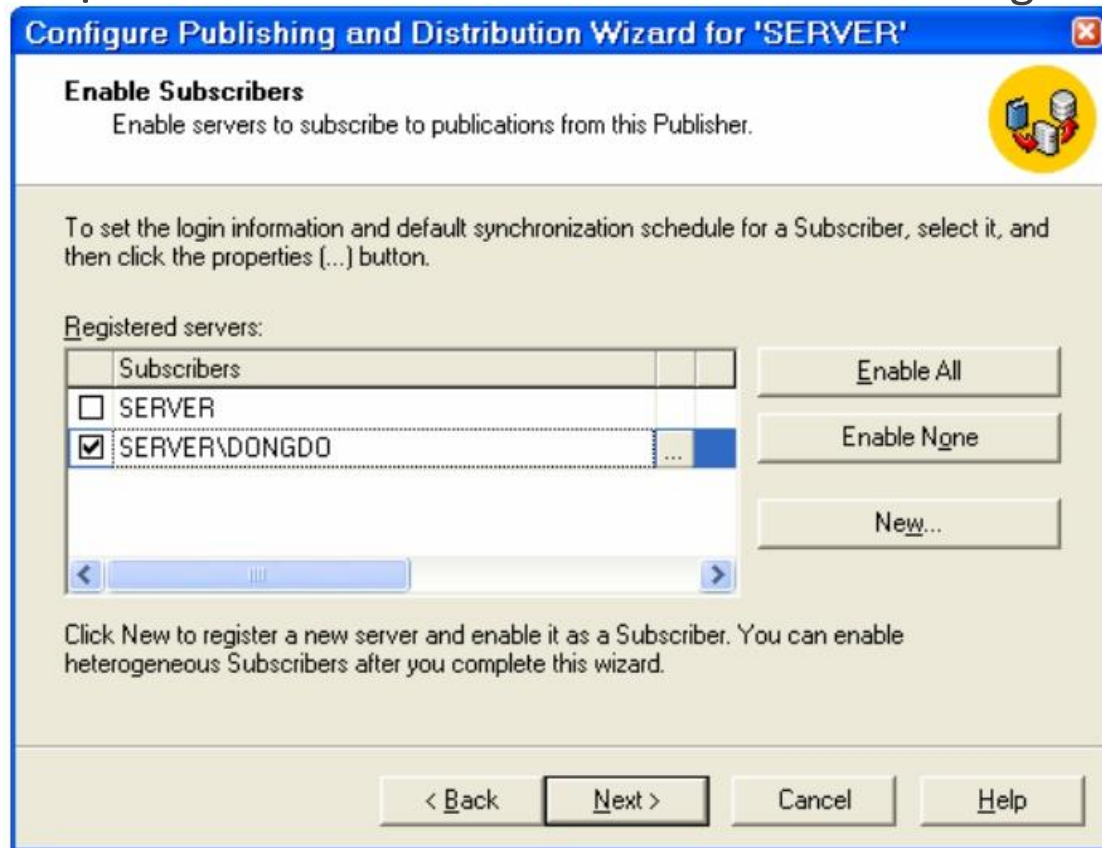
Cấu hình Publisher và Distributor

- Chọn CSDL tham gia nhân bản, kiểu nhân bản



Cấu hình Publisher và Distributor

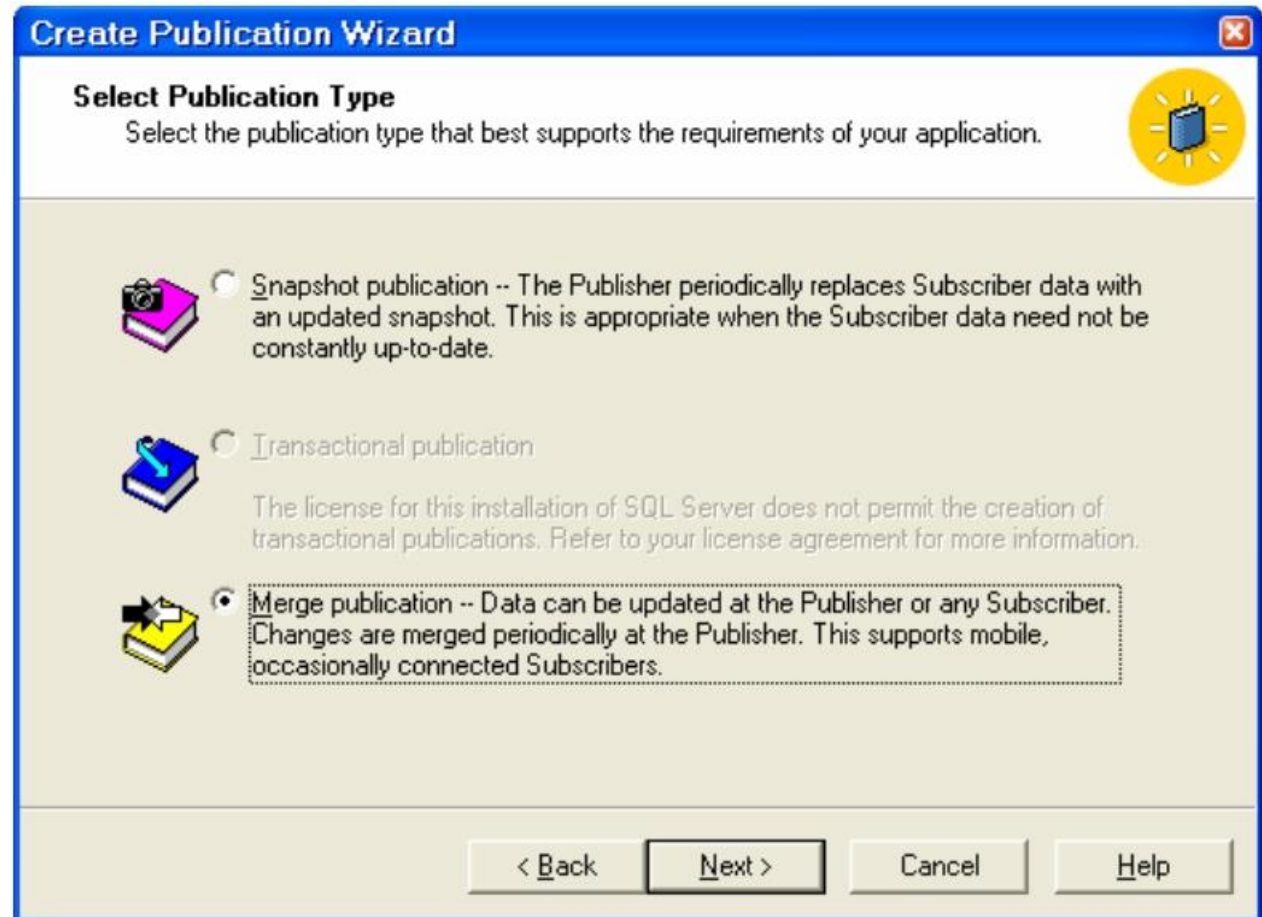
- Chọn Server được cấu hình là Subscriber của Publisher đang thiết lập



Tạo Publication

Tạo Publication

- Chọn Publication trong Replication của Publisher
- Nhấn phải chuột, chọn New Publication
- Chọn CSDL cần xuất bản dữ liệu hoặc đối tượng
- Chọn kiểu nhân bản



Tạo publication

- Chọn phiên bản SQL Server của Subscriber



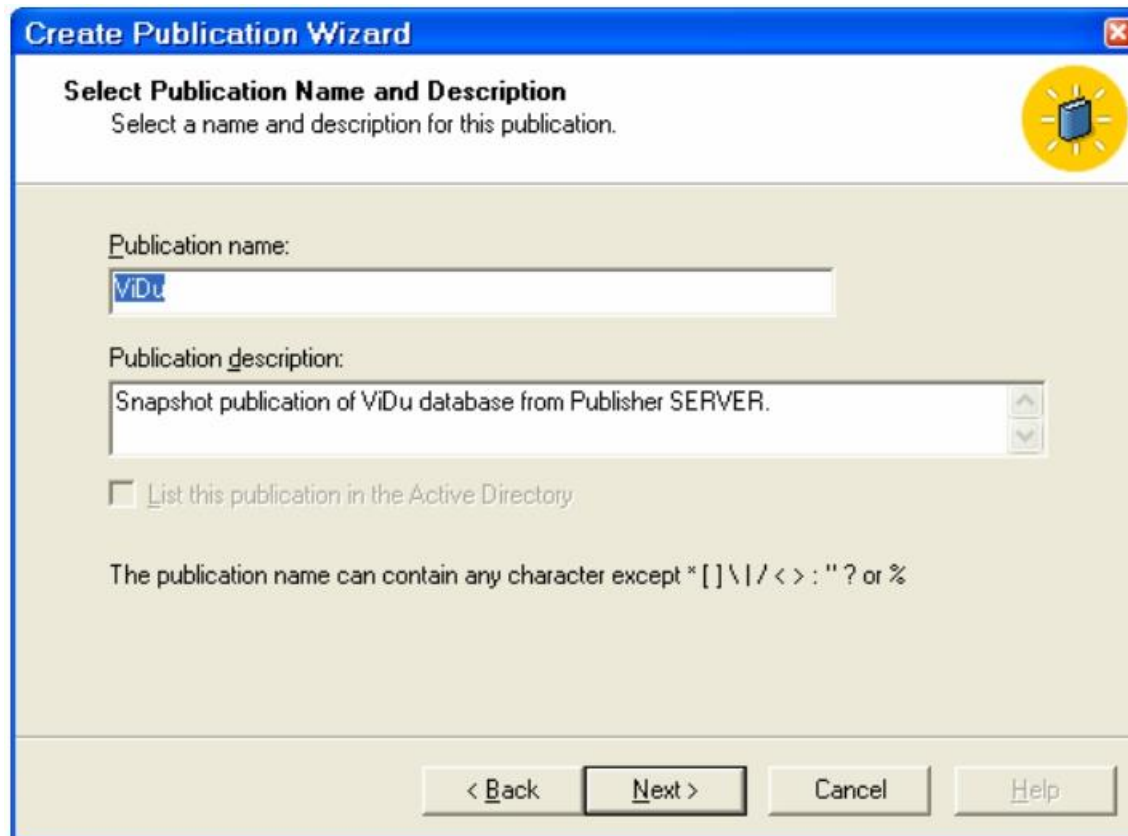
Tạo publication

- Chọn Article tham gia Publication



Tạo publication

- Đặt tên cho Publication



The screenshot shows a Windows-style dialog box titled "Create Publication Wizard". The main heading is "Select Publication Name and Description" with a sub-instruction "Select a name and description for this publication." and a yellow icon of a book with rays. Below this, there are two text input fields. The first is labeled "Publication name:" and contains the text "ViDu". The second is labeled "Publication description:" and contains the text "Snapshot publication of ViDu database from Publisher SERVER.". Below the description field is a checkbox labeled "List this publication in the Active Directory" which is currently unchecked. At the bottom of the dialog, there is a note: "The publication name can contain any character except * [] \ / < > : " ? or %". The bottom of the dialog features four buttons: "< Back", "Next >", "Cancel", and "Help".

Create Publication Wizard

Select Publication Name and Description
Select a name and description for this publication.

Publication name:
ViDu

Publication description:
Snapshot publication of ViDu database from Publisher SERVER.

☐ List this publication in the Active Directory

The publication name can contain any character except * [] \ / < > : " ? or %

< Back Next > Cancel Help

Tạo Push Subscription

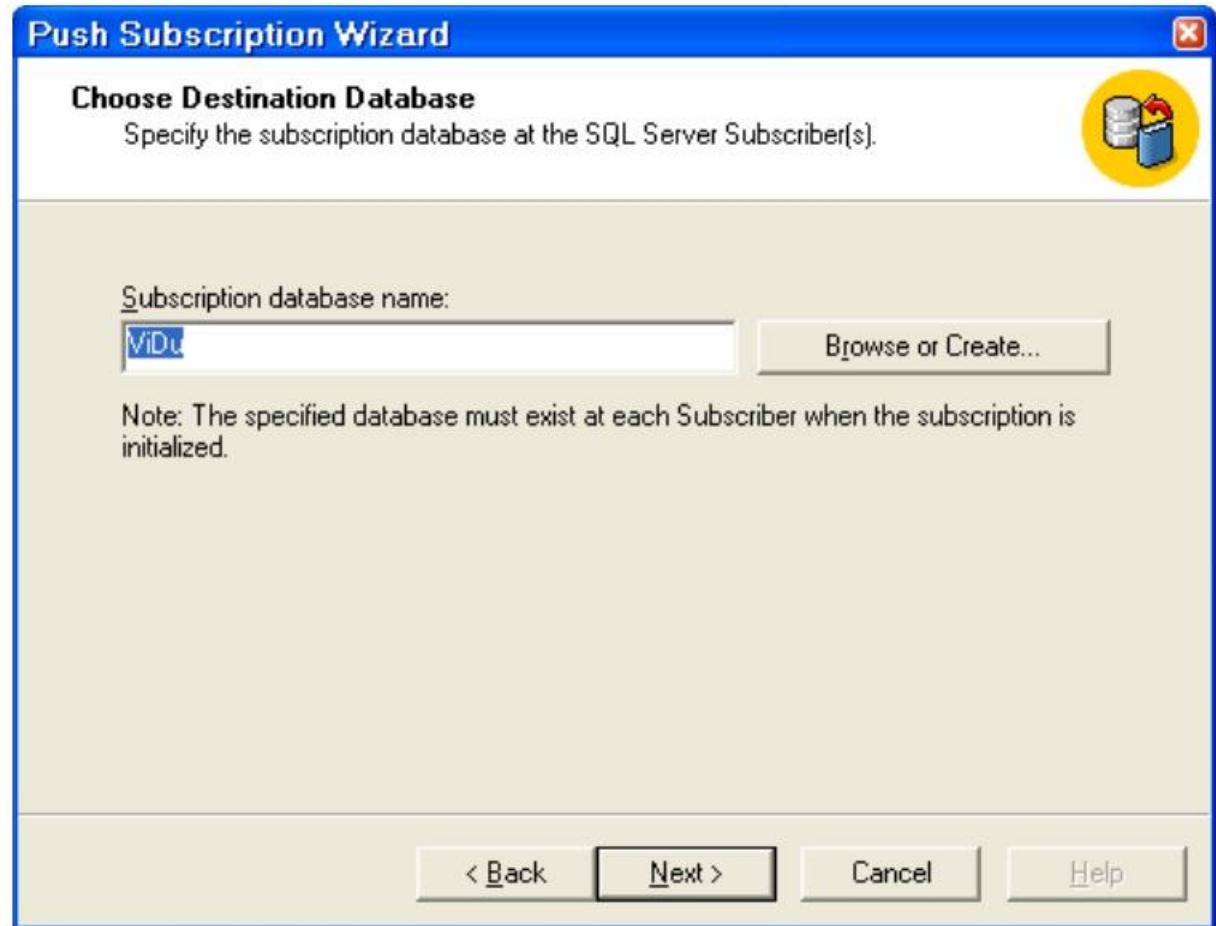
Tạo push Subscription

- Bước này thực hiện thủ tục đẩy từ Publisher lên Subscriber, được thực hiện trên Publisher
- Chọn Publication của Publisher, chọn Push New Subscription
- Chọn Subscriber



Tạo push Subscription

- Chọn CSDL trên Subscriber nếu đã có, nếu chưa, thực hiện chọn chức năng tạo mới



The screenshot shows the 'Push Subscription Wizard' dialog box, specifically the 'Choose Destination Database' step. The title bar reads 'Push Subscription Wizard'. The main heading is 'Choose Destination Database', followed by the instruction 'Specify the subscription database at the SQL Server Subscriber(s)'. There is a yellow circular icon with a database symbol in the top right corner. Below the instruction, there is a text box labeled 'Subscription database name:' containing the text 'VIDU'. To the right of the text box is a button labeled 'Browse or Create...'. Below the text box, there is a note: 'Note: The specified database must exist at each Subscriber when the subscription is initialized.' At the bottom of the dialog, there are four buttons: '< Back', 'Next >', 'Cancel', and 'Help'.

Push Subscription Wizard

Choose Destination Database
Specify the subscription database at the SQL Server Subscriber(s).

Subscription database name:
VIDU

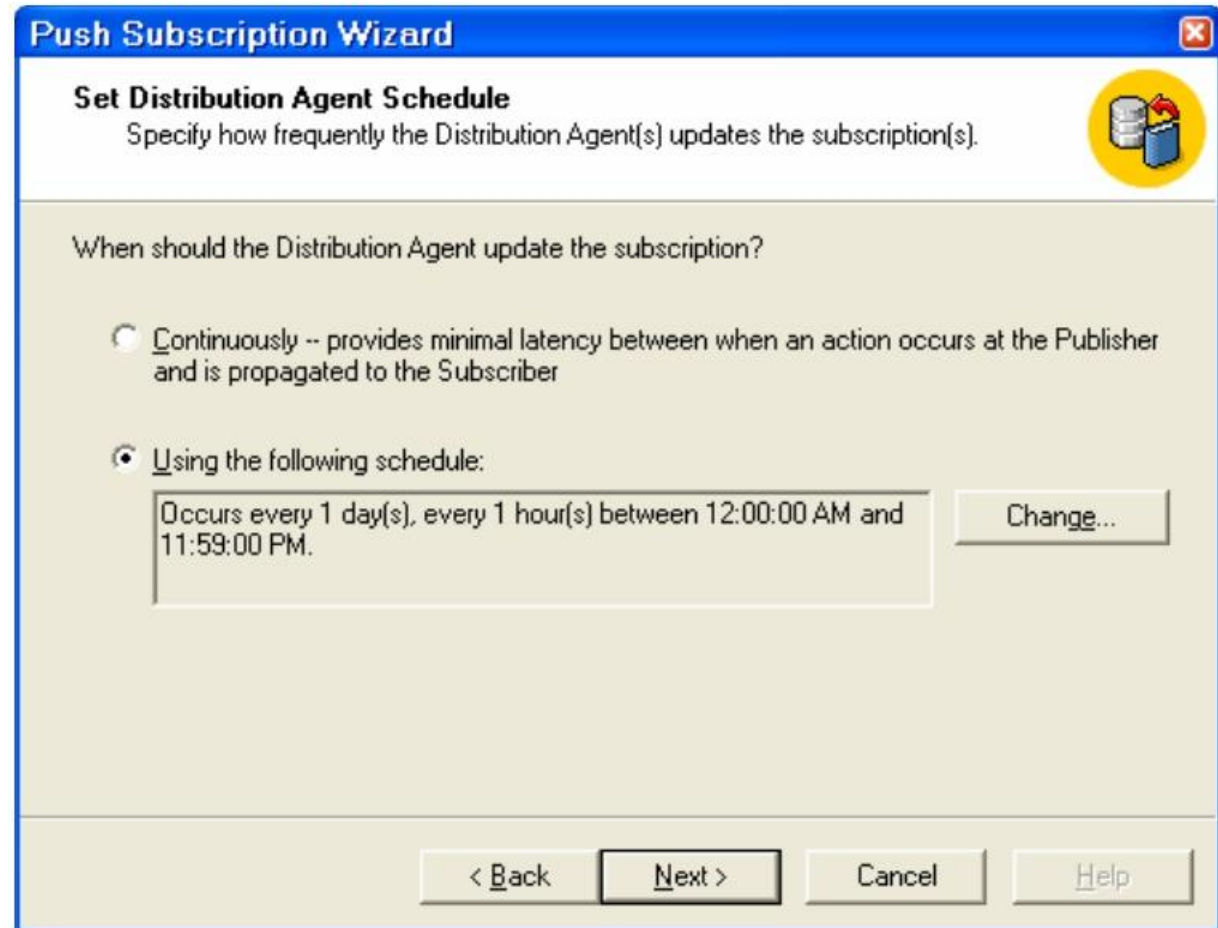
Browse or Create...

Note: The specified database must exist at each Subscriber when the subscription is initialized.

< Back Next > Cancel Help

Tạo push Subscription

- Chọn lịch thực hiện đồng bộ dữ liệu



Push Subscription Wizard

Set Distribution Agent Schedule
Specify how frequently the Distribution Agent(s) updates the subscription(s).

When should the Distribution Agent update the subscription?

☐ Continuously -- provides minimal latency between when an action occurs at the Publisher and is propagated to the Subscriber

☒ Using the following schedule:

Occurs every 1 day(s), every 1 hour(s) between 12:00:00 AM and 11:59:00 PM.

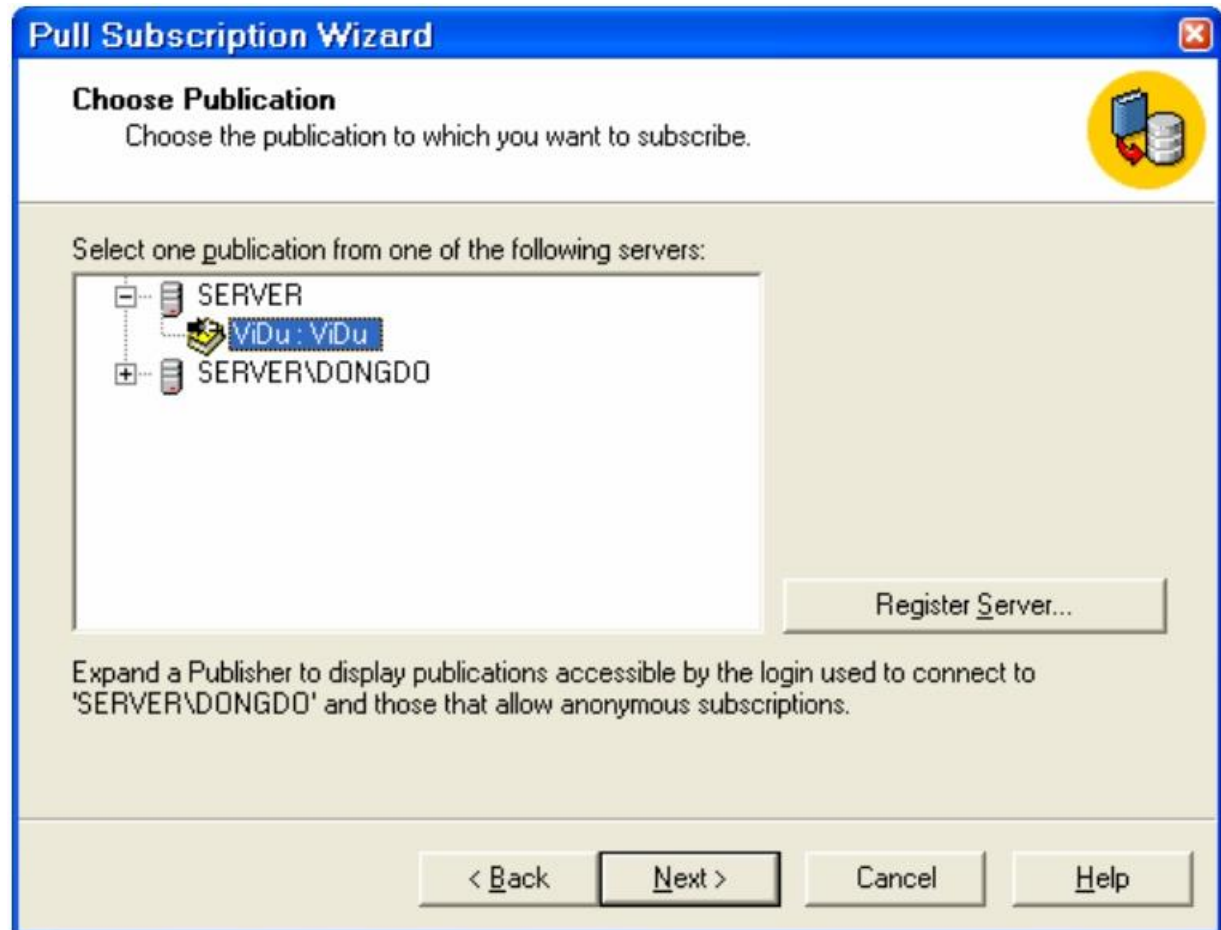
Change...

< Back Next > Cancel Help

Tạo Pull Subscription

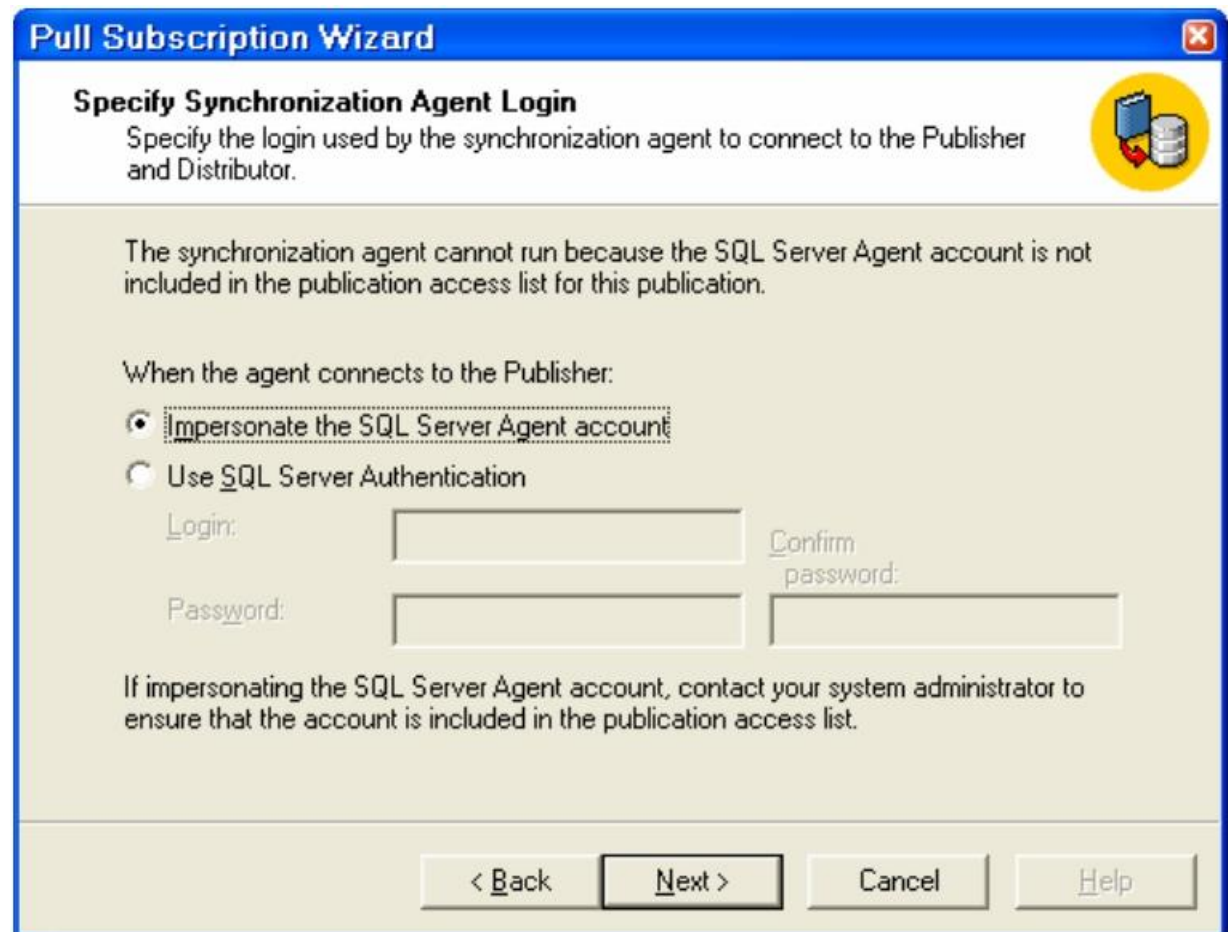
Tạo pull Subscription

- Bước này thực hiện tạo công cụ kéo dữ liệu nhân bản từ Publisher về Subscriber, được thực hiện trên Subscriber
- Chọn Subscription của Subscriber, chọn New Pull Subscription
- Chọn Publication



Tạo pull Subscription

- Chọn Agent tham gia kết nối Publisher



Pull Subscription Wizard

Specify Synchronization Agent Login
Specify the login used by the synchronization agent to connect to the Publisher and Distributor.

The synchronization agent cannot run because the SQL Server Agent account is not included in the publication access list for this publication.

When the agent connects to the Publisher:

☒ Impersonate the SQL Server Agent account

☐ Use SQL Server Authentication

Login: Confirm password:

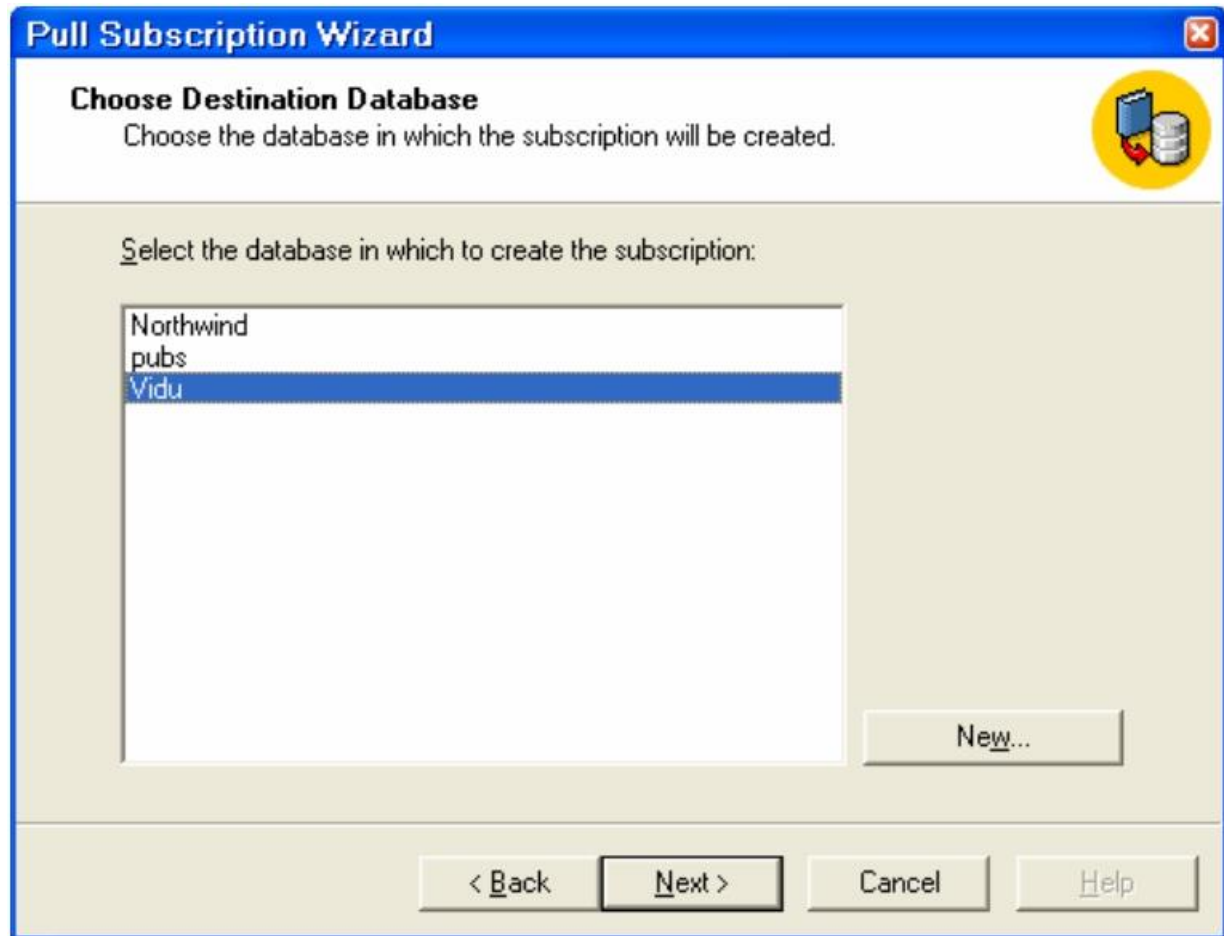
Password:

If impersonating the SQL Server Agent account, contact your system administrator to ensure that the account is included in the publication access list.

< Back Next > Cancel Help

Tạo pull Subscription

- Chọn CSDL đích



Thực hiện đồng bộ dữ liệu

Thực hiện đồng bộ dữ liệu

- Sau khi thiết lập theo các mô hình nhân bản xong, ta có thể thực hiện đồng bộ dữ liệu bằng cách:
 - Thực hiện theo lịch
 - Theo yêu cầu: Chọn Subscription (Push hoặc Pull), nhấn phải chuột chọn Start Synchronizing

