**SQL Server Replication**

1. **Replication**

Replication là một tập các công nghệ sao chép, phân phối dữ liệu và các đối tượng cơ sở dữ liệu từ một cơ sở dữ liệu này đến một cơ sở dữ liệu khác. Sau đó chúng sẽ được đồng bộ với nhau để duy trì tính nhất quán.

1. **Sử dụng Replication**

* Sao chép và phân phối dữ liệu trên nhiều Server khác nhau
* Phân phối bản sao dữ liệu theo lịch trình nhất định.
* Phân phối dữ liệu vừa thay đổi trên nhiều Server khác nhau.
* Cho phép nhiều người dùng và nhiều Server kết hợp dữ liệu khác nhau một cách thống nhất mà không sợ mất dữ liệu.
* Xây dựng CSDL sử dụng cho những ứng dụng trực tuyến hay ngoại tuyến
* Xây dựng ứng dụng Web khi người dùng cần trình bày một số lượng lớn dữ liệu Cho phép copy data đến nhiều server

1. **Mô hình cơ bản của Replication**

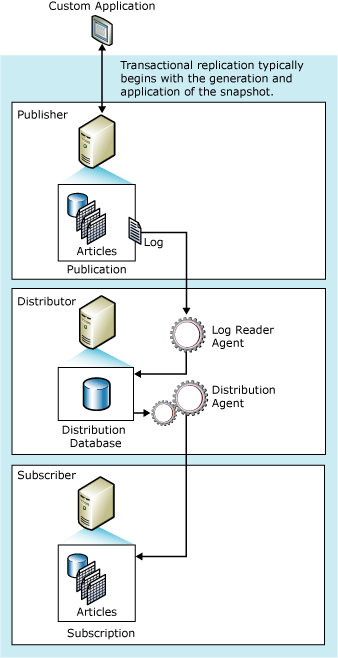
* *Publisher:* Nơi chứa dữ liệu cùng đối tượng cơ sở dữ liệu muốn nhân rộng trên nhiều Server khác nhau.
* *Distributor:* là một server nắm giữ chức năng trung tâm phân phối CSDL và lưu trữ dữ liệu quá khứ cùng với các transaction và meta data.
* *Subcriber:* là những Server dùng để nhận dữ liệu và những Suncribers này đồng ý nhận những dữ liệu mà nó cần chứ không nhất thiết toàn bộ dữ liệu đang có của Publisher.

1. **3 loại replication dùng cho ứng dụng phân tán**

* *Transactional replication*
* *Merge replication*
* *Snapshot replication*

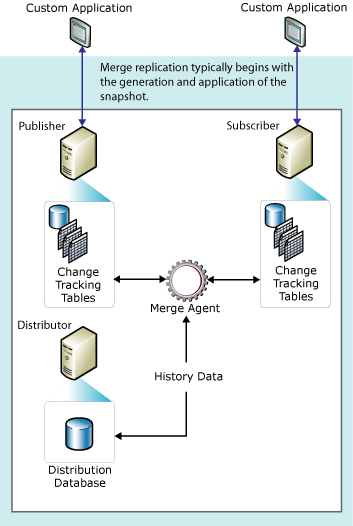
Việc chọn sử dụng loại nào tùy thuộc vào các yếu tố như môi trường nhân bản vật lý, loại và số lượng của dữ liệu được nhân bản.

1. **Replication types**
2. **Transactional replication**



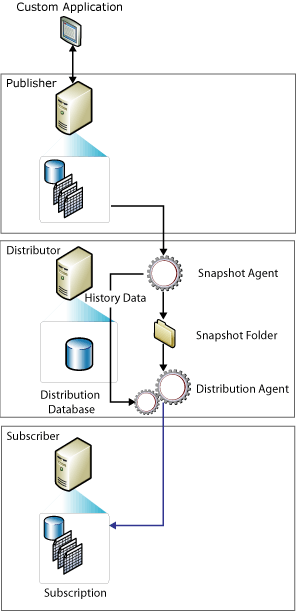
* Bắt đầu với một bản chụp của các đối tượng cơ sở dữ liệu và dữ liệu đã có.
* Sau đó nếu có sự thay đổi dữ liệu và lược đồ tại Publisher thì giao dịch với bên Subscriber cũng sẽ được tiến hành đồng thời. Những dữ liệu thay đổi bên Subscriber đồng nhất với Publisher.
* Transactional replication không cho phép cập nhật tại Subscriber và không truyền trở lại Publisher.
* *Transactional replication được dùng cho môi trường server-to-server, thích hợp với các trường hợp:*
* Khi muốn gia tăng sự thay đổi được truyền tới Subscribers.
* Ứng dụng đòi hỏi độ trễ thấp giữa thời gian thay đổi tại Publisher và thời gian thay đổi tới Subscriber.
* Ứng dụng đòi hỏi truy cập tới dữ liệu trung gian.
* Publisher có khối lượng hoạt động insert, update, delete rất cao.
* Publisher hoặc Subscriber không phải là một cơ sở dữ liệu SQL Server.
* *Transactional replication được thực hiện bởi Snapshot Agent, Log Reader Agent, and Distribution Agent*
* *Snapshot Agent:* chuẩn bị file chụp chứa lược đồ và dữ liệu của các bảng, đối tượng cơ sở dữ liệu được publish. Các file được lưu trữ trong thực mục ảnh chụp và việc đồng bộ các bản ghi trong cơ sở dữ liệu phân phối dựa trên Distribution.
* *Log Reader Agent:* theo dõi nhật ký giao dịch của mỗi cơ sở dữ liệu được cấu hình để Transactional replication và các sao chép giao dịch được đánh dấu cho việc nhân bản từ các bản ghi giao dịch vào cơ sở dữ liệu phân phối, được lưu trữ như một hàng đợi.
* *Distribution Agent:* các bản sao file chụp ban đầu và các giao dịch diễn ra trong các bảng cơ sở dữ liệu phân phối tới Subscriber.

1. **Merge replication**



* Cho phép nhiều server làm việc độc lập (online hay ofline).
* Bắt đầu với một bản chụp của các đối tượng cơ sở dữ liệu và dữ liệu đã có.
* Sau đó nếu có sự thay đổi dữ liệu, lược đồ về sự thay đổi tại Publisher và Subscriber được theo dõi bằng Trigger.
* Subscriber đồng bộ với Publisher khi đã kết nối mạng và trao đổi tất cả các hàng có sự thay đổi giữa Publisher và Subscriber kể từ lần đồng bộ cuối cùng xảy ra.
* *Merge replication được dùng cho môi trường server-to-client, thích hợp với các trường hợp:*
* Nhiều Subscriber có thể cập nhật cùng một dữ liệu tại các thời điểm khác nhau và truyền những thay đổi tới Publisher hoặc các Subscriber khác.
* Subscriber cần nhận dữ liệu, thay đổi ẩn và đồng bộ các thay đổi với Publisher hoặc các Subscriber khác.
* Mỗi Subscriber đòi hỏi một phân vùng dữ liệu khác nhau.
* Xảy ra xung đột khi chúng hoạt động, cần có khả năng phát hiện và giải quyết vấn đề.
* Ứng dụng yêu cầu thay đổi mạng dữ liệu thay vì truy cập tới dữ liệu trung gian.

1. **Snapshot replication**



* Là quá trình thực hiện việc sao chép và phân phối đúng dữ liệu đối với các cơ sở tại một thời điểm.
* Snapshot Replication không yêu cầu kiểm tra sự thay đổi bởi vì dữ liệu thay đổi không được nhân rộng đến Subscriber.
* Các tùy chọn có trong Snapshot Replication cho phép lọc dữ liệu sẽ được nhân rộng và Subscribers có thể làm thay đổi chúng để nhân rộng những thay đổi này với Publisher hay với các Subscribers khác.
* *Snapshot Replication thích hợp với các trường hợp*:
* Dữ liệu thường xuyên thay đổi.
* Chấp nhận bản sao của dữ liệu mà nó có thể hết thời hạn sử dụng với Publisher trong 1 khoảng thời gian.
* Dùng khi bạn cần Replication dữ liệu có dung lượng nhỏ .
* Một khối lượng lớn dữ các thay đổi xảy ra trong một khoảng thời gian ngắn.