**ORACLE REPLICATION**

Môn học IS211.M11.HTCL

Sinh viên thực hiện:

19520814 – Nguyễn Thị Phương Nhi

19521520 – Phan Phạm Quỳnh Hoa

19522304 – Huỳnh Minh Thư

**Phần 1: Tìm hiểu về Oracle Replication**

1. ***Định nghĩa***

* Replication là quá trình sao chép, phân phối dữ liệu và đối tượng CSDL từ một CSDL1 đến CSDL2 sau đó đồng bộ giữa hai CSDL với nhau
* Replication sử dụng công nghệ cơ sở dữ liệu phân tán để chia sẻ dữ liệu giữa nhiều sites, nhưng cơ sở dữ liệu sao chép và cơ sở dữ liệu phân tán không giống nhau.
* Trong cơ sở dữ liệu phân tán, dữ liệu có sẵn ở nhiều vị trí, nhưng một bảng cụ thể chỉ nằm ở một vị trí.

1. ***Ưu và nhược điểm của nhân bản (Replication)***

* Ưu điểm:

+ Lợi ích chủ yếu của nhân bản dữ liệu xoay quanh hiệu suất, nhưng cũng có một lợi ích về đảm bảo độ bền của dữ liệu.

+ Oracle Replication cho phép các tổ chức tạo, đồng bộ hóa và phân phối dữ liệu trên nhiều địa điểm, cho phép chia sẻ dữ liệu với người dùng như đối tác hoặc nhà cung cấp, tổng hợp dữ liệu từ các văn phòng khác nhau

* Nhược điểm:

+ Tiêu tốn nhiều dung lượng vì phải lưu trữ các bản sao của cùng một dữ liệu tại các trang web khác nhau

* Việc sao chép dữ liệu trở nên tốn kém khi các bản sao ở tất cả các trang web khác nhau cần được cập nhật.
* Duy trì tính nhất quán của dữ liệu tại tất cả các sites khác nhau cần các biện pháp phức tạp.

1. ***Đặc trưng của Replication***

* Availability (tính khả dụng): Việc sử dụng nhiều máy chủ với cùng một dữ liệu đảm bảo tính sẵn sàng cao. Nhân bản dữ liệu sao chép dữ liệu sang một số máy, nên việc duy trì quyền truy cập vào dữ liệu sẽ dễ dàng hơn. Nếu một trang web không khả dụng, người dùng có thể tiếp tục truy vấn hoặc thậm chí cập nhật các vị trí còn lại.
* Performance (hiệu suất): Nhân bản cung cấp quyền truy cập cục bộ, nhanh chóng vào dữ liệu được chia sẻ vì nó cân bằng hoạt động trên nhiều trang sites. Một số người dùng có thể truy cập một máy chủ trong khi những người dùng khác truy cập các máy chủ khác, do đó giảm tải ở tất cả các servers .
* Network Load Reduction (giảm tải mạng hay nói cách khác là nâng hiệu suất mạng):
* Replication có thể được sử dụng để phân phối dữ liệu trên nhiều vị trí khu vực. Do đó, các ứng dụng có thể truy cập các máy chủ ở các khu vực khác nhau thay vì truy cập một máy chủ trung tâm. Việc duy trì nhiều bản sao của cùng một dữ liệu thuận tiện cho việc giảm thiểu độ trễ khi truy cập dữ liệu. Ngoài ra, nó có thể giảm chi phí phân tích với giảm tải truy vấn, bằng cách chia nhỏ quá trình xử lý giao dịch trực tuyến (online transaction processing ) và tạo báo cáo trên các hệ thống khác nhau.

1. ***Chức năng của Replication***

* Gồm 2 loại

+ Nhân bản đồng bộ:

* Nhân bản đồng thời dữ liệu sang vùng chính và vùng phụ
* Rất ít khả năng xảy ra lỗi
* Không bị mất dữ liệu sau khi khôi phục nếu có lỗi
* Chi phí nhân bản cao

+ Nhân bản không đồng bộ:

* Thời gian nhân bản đồng thời đến các khu vực chính và phụ có chút chậm trễ
* Dữ liệu được ghi vào mảng chính sau đó nhân bản sang các mảng phụ
* Có thể nhân bản ở khoảng cách xa và vào các khoảng thời gian cụ thể
* Thu thập dữ liệu thay đổi Oracle:

+ CDC đồng bộ

* Thu thập dữ liệu bằng các triển khai Trigger
* Giao dịch không bị đóng trong DB cho đến khi dữ liệu thay đổi chưa được thu thập
* Tại thời điểm thu thập CSDL không thêm chi phí bổ sung mà còn giảm chi phí bằng cách đơn giản hóa việc trích xuất dữ liệu

+ CDC không đồng bộ

* Dữ liệu thay đổi sẽ được ghi lại sau khi giao dịch hoàn tất
* Có ba chế độ Thu thập dữ liệu thay đổi không đồng bộ: Nhật ký nóng, Nhật ký nóng phân tán và tự động hóa.

+ Nhật ký nóng: Trong chế độ này, dữ liệu thay đổi được ghi lại từ các tệp nhật ký làm lại trực tuyến tại nguồn. Có một độ trễ thời gian nhỏ giữa việc cam kết các giao dịch bảng nguồn và sự xuất hiện của dữ liệu thay đổi để thay đổi bảng. Các bảng thay đổi trong chế độ này phải có trong cơ sở dữ liệu nguồn.

+ Nhật ký nóng phân tán: Trong chế độ Nhật ký nóng phân tán, các bảng Nguồn nằm trong cơ sở dữ liệu Nguồn và các bảng thay đổi nằm trong cơ sở dữ liệu dàn. Trong chế độ này, dữ liệu thay đổi cũng được ghi lại thông qua nhật ký trực tuyến. Cơ sở dữ liệu nguồn và cơ sở dữ liệu Staging có thể nằm trên các nền tảng khác nhau và có thể được cài đặt hệ điều hành khác nhau.

+ Nhật ký tự động: Trong chế độ Nhật ký tự động, dữ liệu thay đổi được thu thập từ một tập hợp các tệp nhật ký làm lại được quản lý bởi các dịch vụ vận tải Làm lại. Nó kiểm soát việc chuyển tự động các tệp nhật ký làm lại từ cơ sở dữ liệu nguồn sang cơ sở dữ liệu dàn. Trong chế độ này, dữ liệu thay đổi có thể được ghi lại từ nhật ký trực tuyến cũng như từ nhật ký lưu trữ. Các chế độ này được gọi là Tự động lưu trữ trực tuyến không đồng bộ và Lưu trữ tự động không đồng bộ.

**Phần 2: Hệ thống nhân bản Oracle**

1. ***Các thành phần cơ bản trong Oracle Replication System***

* Replication object*.* Oracle Replication cho phép sao chép các loại đối tượng sau:

+ Tables

+ Indexes

+ Views

+ Packages and Package Bodies

+ Procedures and Functions

+ Triggers

+ Sequences

+ Synonyms

* Replication groups: là một tập hợp các Replication objectcó quan hệ logic với nhau. Trong môi trường nhân bản, Oracle quản lý các Replication object bằng cách sử dụng các replication groups. Các đối tượng trong một nhóm sao chép được quản lý cùng nhau.
* Replication sites: một Replication sites có thể tồn tại ở nhiều Replication sites. Môi trường nhân bản hỗ trợ hai loại sites cơ bản master sites và snapshot sites*.*

+ Một trang có thể vừa là master sites vừa là snapshot sites cùng một lúc.

* Một master site duy trì một bản sao hoàn chỉnh của tất cả các đối tượng trong một replication group.
* Trong khi snapshots trong một snapshot site có thể chứa tất cả hoặc một tập hợp con các bảng dữ liệu trong một nhóm chính (master group).

+ Ví dụ:

* Nếu nhóm chính SCOTT\_MG chứa các bảng EMP và DEPT, tất cả các master sites phải duy trì một bản sao hoàn chỉnh của EMP và DEPT.
* Tuy nhiên, một snapshot site có thể chỉ chứa một snapshot của bảng EMP, trong khi một snapshot site khác có thể chứa snapshot của cả bảng EMP và DEPT.

1. Các loại môi trường nhân bản

* Multimaster replication (hay còn gọi là peer-to-peer hoặc *n*-way replication) (nhân bản đa quản trị viên)

+ Cho phép nhiều sites, hoạt động như các đồng nghiệp ngang hàng, để quản lý các nhóm đối tượng cơ sở dữ liệu được nhân bản.

+ Mỗi trang web trong môi trường multimaster replication là một trang master site.

Diagram

Description automatically generated+ Các ứng dụng có thể cập nhật bất kỳ bảng sao chép nào tại bất kỳ site nào trong cấu hình multimaster.

+ Asynchronous replication (Sao chép không đồng bộ) là cách phổ biến nhất để triển khai multimaster replication.

* Snapshot replication

+ A snapshot có thể là read-only hoặc updateable.

+ Snapshot cung cấp các lợi ích sau:

* Bật quyền truy cập cục bộ, giúp cải thiện thời gian phản hồi và tính khả dụng. Giảm tải các truy vấn từ trang chính, vì người dùng có thể truy vấn
* Snapshot cục bộ thay thế.
* Tăng cường bảo mật dữ liệu bằng cách cho phép bạn chỉ sao chép một tập hợp con đã chọn của tập dữ liệu của bảng chính mục tiêu (target master table).
* Read-only snapshots

+ Mô tả: Trong cấu hình cơ bản, snapshots có thể cung cấp quyền truy cập chỉ đọc( read-only) vào dữ liệu bảng bắt nguồn từ master site.

+ Các ứng dụng có thể truy vấn dữ liệu từ read-only snapshots để tránh truy cập mạng bất kể tính khả dụng của mạng.

+ Tuy nhiên, các ứng dụng trên toàn hệ thống phải truy cập dữ liệu tại trang chủ để thực hiện cập nhật.

Diagram

Description automatically generated

* Updateable snapshots

+ Mô tả: Snapshot có thể cập nhật cho phép người dùng chèn, cập nhật và xóa các hàng của bảng chính đích bằng cách thực hiện các thao tác này trên snapshot.

+ Updateable Snapshot cũng có thể chỉ chứa một tập hợp con dữ liệu trong bảng chính đích.

Diagram

Description automatically generated

* Multimaster and Snapshot Hybrid Configurations

+ Mô tả : multimaster (or *n*-way) replication between two masters có thể hỗ trợ sao chép toàn bảng giữa các cơ sở dữ liệu hỗ trợ hai vùng địa lý.

+ Snapshots có thể được xác định trên masters để sao chép toàn bộ bảng hoặc tập hợp con bảng tới các sites trong mỗi khu vực.

**Diagram

Description automatically generated**

**Phần 3: Hướng dẫn DEMO**

Các bước thực hiện thu thập dữ liệu thay đổi không đồng bộ (Asynchronous Change Data Capture) trên Oracle Database.

* Bước 1: Đặt Cơ sở dữ liệu thành Chế độ FORCE LOGGING

Bước đầu tiên là kích hoạt chế độ FORCE LOGGING trên cơ sở dữ liệu của bạn.

Câu truy vấn ALTER DATABASE FORCE LOGGING;

Text

Description automatically generated

* Bước 2: Cấp quyền Supplemental Logging

Điều này là để cho phép cơ sở dữ liệu đăng nhập trong trường hợp các câu lệnh UPDATE.

Câu truy vấn ALTER DATABASE ADD SUPPLEMENTAL LOG DATA;

Text

Description automatically generated

* Bước 3: Bật **Logging** cho tất cả các cột trong bảng bắt buộc (**required Table)**

Trước tiên, bạn cần xác định các cột mà bạn cần thu thập các thay đổi.

Nếu bạn chọn thu thập các thay đổi từ tất cả các cột, bạn có thể chạy truy vấn sau.

ALTER TABLE DEMO\_CUSTOMER ADD SUPPLEMENTAL LOG DATA (ALL) COLUMNS;

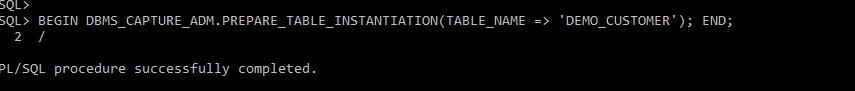
Text

Description automatically generated

* Bước 4:Chuẩn bị Source Tables cho Log Data Capture

Câu truy vấn

BEGIN DBMS\_CAPTURE\_ADM.PREPARE\_TABLE\_INSTANTIATION(TABLE\_NAME => ‘DEMO\_CUSTOMER’); END;



* Bước 5: Tạo một Change Set

Lưu ý: Oracle không hỗ trợ các packages này từ Oracle 12c trở đi.

Để tạo một change set, hãy sử dụng lệnh dưới đây trong Sql Plus command Line. Tạo Change set bằng gói DBMS\_CDC\_PUBLISH

BEGIN  
DBMS\_CDC\_PUBLISH.CREATE\_CHANGE\_SET(  
change\_set\_name => 'DEMO\_CUSTOMER',  
description => 'DEMO\_CDC',  
change\_source\_name => 'HOTLOG\_SOURCE',  
stop\_on\_ddl => 'y',  
begin\_date => sysdate,  
end\_date => sysdate+100);  
END;

Tập hợp thay đổi này sẽ ghi lại các thay đổi bắt đầu từ hôm nay và sẽ ngừng chụp sau 100 ngày. HOTLOG\_SOURCE là một nguồn được xác định trước bởi oracle để thu thập dữ liệu thay đổi không đồng bộ. Text

Description automatically generated

* Bước 6: Tạo Change Table

Tạo Change Table sẽ chứa các thay đổi đối với source tables. Source tables ở đây là bảng ‘Customer’.

BEGIN

DBMS\_CDC\_PUBLISH.CREATE\_CHANGE\_TABLE(

owner => 'masteruser',

change\_table\_name => 'CUSTOMER',

change\_set\_name => 'DEMO\_CUSTOMER',

source\_schema => 'test',

source\_table => 'CUSTOMER',

column\_type\_list => 'ID NUMBER(6),

NAME VARCHAR2(50),

ADDRESS VARCHAR2(500)',

capture\_values => 'both',

rs\_id => 'y',

row\_id => 'n',

user\_id => 'n',

timestamp => 'n',

object\_id => 'n',

source\_colmap => 'n',

target\_colmap => 'y',

options\_string => 'TABLESPACE TS\_USER1');

END;

Text

Description automatically generated

* Bước 7: Enable Change Set

BEGIN

DBMS\_CDC\_SUBSCRIBE.CREATE\_SUBSCRIPTION(

change\_set\_name => 'DEMO\_CUSTOMER',

description => 'Change data for test',

subscription\_name => 'TEST\_CDC\_SUBS');

END;

Text

Description automatically generated

* Bước 8: Tạo một subscription có thể truy cập dữ liệu thay đổi trong Change tables
* Bước 9: Đăng ký subscriber đến source table và các cột trong source table

BEGIN

DBMS\_CDC\_SUBSCRIBE.SUBSCRIBE(

subscription\_name => 'TEST\_CDC\_SUBS',

source\_schema => 'TEST',

source\_table => 'CUSTOMER',

column\_list => 'ID,NAME,ADDRESS',

subscriber\_view => 'SUBS\_VIEW');

END;

* Bước 10: kích hoạt subscription

SQL> BEGIN  
DBMS\_CDC\_SUBSCRIBE.ACTIVATE\_SUBSCRIPTION(  
subscription\_name => 'SUBS\_TEST');  
END;  
/

PL/SQL procedure successfully completed.

* Bước 11:

Khối PL / SQL này sẽ cung cấp cho bạn tập dữ liệu thay đổi tiếp theo.

Bất cứ khi nào dữ liệu thay đổi trong bảng nguồn và được ghi lại trong

Change table sau đó khối PL / SQL này sẽ lấy cho bạn dữ liệu trong subscription window.

BEGIN

DBMS\_CDC\_SUBSCRIBE.EXTEND\_WINDOW(

subscription\_name => 'SUBS\_TEST');

END;

* Bước 12:

Kiểm tra xem các Thay đổi được thực hiện đối với source table có phản ánh trong subscription window hay không.

BEGIN

DBMS\_CDC\_SUBSCRIBE.EXTEND\_WINDOW(

subscription\_name => 'SUBS\_TEST');

END;

PL/SQL procedure successfully completed.

SQL> select  OPERATION$,to\_char(COMMIT\_TIMESTAMP$,’dd-mon-yyyy hh24:mi:ss’),ID,NAME from subs\_view;

Bất kỳ thay đổi nào chúng ta đã thực hiện đối với source table đều được phản ánh trong change table và được liệt kê trong subscription window.

Nếu chúng ta insert dữ liệu trong source table, tất cả các chi tiết bản ghi sẽ được phản ánh trong change table. Nếu chúng ta update, thì cả dữ liệu trước khi cập nhật và dữ liệu sau khi cập nhật sẽ được ghi lại trong Change table và được phản ánh trong subscription window.

TÀI LIỆU THAM KHẢO

Database Replication

<https://docs.oracle.com/cd/A64702_01/doc/server.805/a58227/ch_repli.htm>

Replication Overview

<https://docs.oracle.com/cd/A87860_01/doc/server.817/a76959/repover.htm>

<https://bryteflow.com/oracle-replication-in-real-time-step-by-step/?fbclid=IwAR3nMWIlzKX3GahjospApTzmuKr83mNgKDCjO6CGL3N1XtqhA6dS4gnPkFw>

<https://ecomputernotes.com/database-system/adv-database/data-replication>