HỌC VIỆN NGÂN HÀNG

KHOA CÔNG NGHỆ THÔNG TIN VÀ KINH TẾ SỐ

KHÓA LUẬN TỐT NGHIỆP

NĂM HỌC 2023 – 2024

ĐỀ TÀI:

GIẢI PHÁP DỮ LIỆU DỰ PHÒNG DATA GUARD CHO HỆ THỐNG XẾP HẠNG TÍN DỤNG CÔNG TY TÀI CHÍNH SHBFINANCE

Sinh viên thực hiện:	Lê Hoàng Vũ
Lớp:	K23HTTTA
Khóa học:	2020 – 2024
Mã sinh viên:	23A4040156
Giảng viên hướng dẫn:	ThS. Giang Thị Thu Huyền

Hà Nội, tháng 5 năm 2024

HỌC VIỆN NGÂN HÀNG

KHOA CÔNG NGHỆ THÔNG TIN VÀ KINH TẾ SỐ

KHÓA LUẬN TỐT NGHIỆP

NĂM HỌC 2023 - 2024

ĐỀ TÀI:

GIẢI PHÁP DỮ LIỆU DỰ PHÒNG DATA GUARD CHO HỆ THỐNG XẾP HẠNG TÍN DỤNG CÔNG TY TÀI CHÍNH SHBFINANCE

Sinh viên thực hiện:	Lê Hoàng Vũ
Lớp:	K23HTTTA
Khóa học:	2020 – 2024
Mã sinh viên:	23A4040156
Giảng viên hướng dẫn:	ThS. Giang Thị Thu Huyền

Hà Nội, tháng 5 năm 2024

LỜI CAM ĐOAN

Em xin cam đoan toàn bộ nội dung trong khóa luận tốt nghiệp này là sản phẩm nghiên cứu, tìm hiểu của cá nhân em trong quá trình được học tập và thực tập tại Công ty Công nghệ JProTech.

Trong khóa luận tốt nghiệp của em, bên cạnh những nội dung là kiến thức của bản thân em cũng đã tham khảo một số nguồn tài liệu công khai, uy tín. Tất cả các tài liệu tham khảo trong khóa luận đều được trích dẫn một cách hợp pháp.

Em xin chịu trách nhiệm hoàn toàn và chịu mọi hình thức kỷ luật theo quy định cho lời cam đoan của mình.

Sinh viên thực hiện

LÊ HOÀNG VŨ

LỜI CẢM ƠN

Lời đầu tiên, em xin gửi lời cảm ơn chân thành đến Học viện Ngân hàng, Khoa Công nghệ thông tin và Kinh tế số, các thầy cô thuộc Khoa đã tạo điều kiện để em cũng như các bạn sinh viên đón nhận được kiến thức về chuyên ngành, môi trường học tập ổn định. Tiếp theo là lời cảm ơn ThS. Giang Thị Thu Huyền, giảng viên đã hướng dẫn em khóa luận tốt nghiệp. Trong quá trình nghiên cứu và thực nghiệm, em đã nhận được sự chỉ dẫn sát sao cùng góp ý thực tiễn của cô qua các giai đoạn, để bài viết mang tính khoa học và súc tích, truyền đạt tới người đọc một cách dễ hiểu và bao quát các vấn đề.

Bên cạnh đó, em cũng xin gửi lời cảm ơn đến Công ty TNHH Công nghệ JProTech nói chung và phòng JData nói riêng đã tạo điều kiện cho em có cơ hội được thực tập để học hỏi và tích lũy thêm kiến thức cũng như kinh nghiệm làm việc. Từ lúc được gia nhập vào công ty, anh chị đã tạo cơ hội cho em được tiếp xúc và làm việc với nhiều công việc mới cũng như hỗ trợ về mặt các thủ tục mà trường yêu cầu.

Em cũng xin gửi lời cảm ơn tới gia đình, cũng như bạn bè đã là điểm tựa tinh thần, để em có thể cố gắng hơn mỗi ngày, mong muốn được cống hiến cho gia đình, nhà trường, xã hội những giá trị tốt đẹp.

Với thái độ cầu tiến, chủ động, em đã thực hiện bài báo cáo một cách nghiêm túc, không chỉ về nội dung chuyên môn, mà còn về cách trình bày, thái độ trong quá trình làm việc với giảng viên. Tuy nhiên, em hiểu rằng, luôn luôn có những thiếu sót và sai lầm. Vì vậy, em sẵn sàng tiếp nhận và mong muốn được thầy/cô góp ý, phê bình và sửa đổi.

Em xin kính chúc thầy/cô giữ được nhiệt huyết, đam mê với nghề nhà giáo trân trọng và đáng kính, để không chỉ em cùng các bạn sinh viên khóa 23 - Khoa Công nghệ thông tin và Kinh tế số nói riêng, được tiếp nhận những kiến thức chuyên sâu, mà còn đem lại cho các em sinh viên thế hệ sau, tìm được con đường mình yêu thích trong những ngày tháng của bốn năm đại học.

Em xin chân thành cảm ơn thầy/cô!

MỤC LỤC

DANH MỤC CÁC CHỮ VIẾT TẮT

Chữ viết tắt	Chữ đầy đủ	Diễn giải
ACK	Acknowledgement	Tín hiệu mà tiến trình RFS trả về CSDL chính khi sử dụng AFFIRM/NOAFFIRM
ADG	Active Data Guard	Tính năng trong Data Guard cho phép truy vấn song song với quá trình đồng bộ
ARL	Archived Redo Log	Tệp tin được lưu cất của dữ liệu đồng bộ
ASM	Automatic Storage Management	Tính năng quản lý đĩa, vùng nhớ tự động thay vì File Systems trong Oracle
BMM	Broker Management Model	Mô hình quản lý môi trường Data Guard
CSDL	Cơ sở dữ liệu	Hệ thống lưu trữ dữ liệu. Trong này là CSDL dạng quan hệ
DBWR	Database Writer	Tiến trình ghi dữ liệu từ bộ nhớ xuống đĩa
DGB	Data Guard Broker	Thành phần quản lý môi trường Data Guard
DGMGRL	Data Guard Command Line Interface	Công cụ giao diện dòng lệnh thao tác quản trị môi trường Data Guard
DML	Data Manipulation	Ngôn ngữ thao tác dữ liệu trên CSDL như INSERT, UPDATE
DMON	Data Guard Monitor	Tiến trình trong mô hình Broker thực hiện theo dõi tình trạng môi trường Data Guard
FAL	Fetch Archived Log	Tính năng giúp CSDL dự phòng chủ động trong việc xử lý trễ/thiếu dữ liệu đồng bộ
I/O	Input/Output	Hoạt động đọc/ghi dữ liệu của máy tính
LGWR	Log Writer	Tiến trình ghi dữ liệu đồng bộ từ bộ nhớ xuống đĩa
LNS	LogWriter Network Server	Tên chung cho các tiến trình thực hiện vận chuyển

		dữ liệu đồng bộ
LSP	Logical Standby Process	Tiến trình thực hiện áp dụng thay đổi của CSDL dự phòng dạng lô-gic
MAA	Maximum Availability Architecture	Khung kiến trúc giải pháp sẵn sàng cao của Oracle
MRP	Managed Recovery Process	Tiến trình thực hiện áp dụng thay đổi của CSDL dự phòng dạng vật lý
MTTR	Mean Time To Recovery	Chỉ số thời gian trung bình phục hồi giữa các lần gặp sự cố
NSA	Network Server Async	Tiến trình vận chuyển dữ liệu đồng bộ trong chế độ không đồng bộ
NSS	Network Server Sync	Tiến trình vận chuyển dữ liệu đồng bộ trong chế độ đồng bộ
OCI	Oracle Cloud Infrastructure	Nền tảng hạ tầng điện toán đám mây của Oracle, cung cấp các dịch vụ như IaaS, DaaS, Saas
ODG	Oracle Data Guard	Giải pháp khôi phục sau thảm họa của Oracle
OLTP	Online Transaction Processing	Hệ thống dùng để chuyển xử lý giao dịch, đảm bảo tính ACID của một CSDL
OSB	Oracle Secure Backup	Tính năng giúp bảo mật bản sao lưu
PMON	Process Monitor	Một trong sáu tiến trình quan trọng giúp Instance hoạt động của CSDL Oracle
RAC	Real Application Cluster	Kiến trúc Oracle Database, trong đó một CSDL có thể được sử dụng bởi nhiều Instance (bản thể, multi-Instance)
RFS	Remote File Server	Tiến trình thực hiện nhận dữ liệu đồng bộ của CSDL dự phòng
RMAN	Recovery Manager	Công cụ sao lưu và phục hồi của Oracle
RPO	Recovery Point Object	Chỉ số mục tiêu về lượng dữ liệu phục hồi
RTO	Recovery Time Object	Chỉ số mục tiêu về thời gian phục hồi

TMCP	Thương mại cổ phần	Loại hình công ty kinh doanh theo mô hình cổ phần – vốn góp của các cổ đông
TNHH MTV	Trách nhiệm hữu hạn, một thành viên	Loại hình công ty do một tổ chức/cá nhân làm chủ sở hữu
WTC	World Trade Center	Trung tâm thương mại thế giới tại Mỹ

DANH MỤC BẢNG BIỂU

DANH MỤC HÌNH

LỜI MỞ ĐẦU

1. Lý do chọn đề tài

Dữ liệu là tài sản quý giá, dữ liệu cung cấp "nguyên liệu" cho việc vận hành của doanh nghiệp. Khi cơ sở dữ liệu gặp sự cố, doanh nghiệp không chỉ thiệt hại về dữ liệu, mà còn bị mất đi doanh thu, cơ hội trong khoảng thời gian hệ thống ngừng hoạt động – "downtime".

Tại sự cố ngày 9/11/2001[1], khi tòa nhà Trung tâm thương mại thế giới (WTC) sụp đổ, hơn 800 tổ chức đã mất dữ liệu quan trọng. Trong khi đó, cơ sở dữ liệu của Morgan Stanley – một ngân hàng/công ty dịch vụ tài chính lớn có trụ sở tại WTC, hoạt động bình thường vào ngày hôm sau, nhờ có cơ sở dữ liệu dự phòng được đặt tại một vị trí khác.

Läätsi (2023) cho rằng, các công ty thuộc Fortune Global 500[2] được ước tính thiệt hại 11% tổng doanh thu cho tới năm 2022, xấp xỉ 1.5 nghìn tỷ USD, tăng 8% so với năm 2020 do thời gian hệ thống ngừng hoạt động gây ra. Vấn đề chi phí do thời gian ngừng hoạt động của hệ thống gây ra cũng là một vấn đề được nghiên cứu sâu rộng, nhằm thiết lập kế hoạch, chiến lược tối ưu để giảm thiểu chi phí.

Công ty SHBFinance với sứ mệnh cung cấp sản phẩm tài chính tiêu dùng cho hơn 500.000 khách hàng, cùng 2 triệu hồ sơ vay vốn trong năm 2021. Cơ sở dữ liệu (CSDL) chứa dữ liệu tín dụng cho hồ sơ vay vốn của SHBFinance rất lớn, nhưng chưa đáp ứng được việc hoạt động trở lại kịp thời trong trường hợp gặp sự cố. Đội ngũ quản trị cơ sở dữ liệu phải thực hiện nhiều thủ tục để khôi phục lại từ các bản sao lưu, dẫn đến nhu cầu về triển khai giải pháp khôi phục nhanh và toàn vẹn dữ liệu hơn.

Hiểu rõ được tầm quan trọng của việc lên kế hoạch dự phòng, khôi phục kịp thời, vì vậy, em đã lựa chọn và thực nghiệm triển khai giải pháp mới với tên đề tài "**Giải pháp dữ liệu dự phòng Data Guard cho hệ thống Xếp hạng tín dụng Công ty Tài chính SHB Finance**".

2. Mục tiêu

Với đề tài "Giải pháp dữ liệu dự phòng Data Guard cho hệ thống Xếp hạng tín dụng Công ty Tài chính SHB Finance", bài hướng tới mục tiêu chính sau:

Thứ nhất, nghiên cứu giải pháp khôi phục dự phòng sau thảm họa Oracle Data Guard, kiến trúc và cơ chế hoạt động của các thành phần liên quan.

[1] Sự kiện cuộc tấn công khủng bố hàng loạt của nhóm Hồi giáo cực đoan Al – Qaeda vào nước Mỹ, gây thiệt hại về người và vật chất quy mô lớn.

[2] Bảng xếp hạng 500 công ty/tập đoàn hàng đầu trên thế giới, theo doanh số của tạp chí Fortune.

Thứ hai, lên kế hoạch và xác định các điều kiện tiên quyết trước khi triển khai giải pháp Oracle Data Guard.

Thứ ba, thực nghiệm triển khai giải pháp Oracle Data Guard cho công ty SHBFinance trên nền tảng phù hợp với các tính năng bổ trợ khác nhau.

3. Đối tượng và phạm vi nghiên cứu

Đối tượng: Giải pháp Oracle Data Guard (ODG) trên phiên bản Oracle Database 19c

Phạm vi nghiên cứu: Cơ sở dữ liệu của hệ thống xếp hạng tín dụng nội bộ công ty tài chính SHBFinance

4. Phương pháp nghiên cứu

- Phân tích, tổng hợp: làm rõ các thành phần trong kiến trúc, cơ chế tương tác giữa các thành phần
- So sánh: các chế độ bảo vệ, phương thức đồng bộ hóa dữ liệu thay đổi
- Liệt kê
- Thực nghiệm và đánh giá

5. Bố cục đề tài

Đề tài gồm 3 chương, với nội dung như sau:

Chương 1: Tổng quan về doanh nghiệp và giải pháp Data Guard

Chương 2: Triển khai giải pháp Data Guard cho cơ sở dữ liệu hệ thống xếp hạng tín dụng của SHBFinance

Chương 3: Kết luận

CHƯƠNG 1. TỔNG QUAN VỀ DOANH NGHIỆP VÀ GIẢI PHÁP DATA GUARD

1.1. Khái quát về Công ty Tài chính Ngân hàng TMCP Sài Gòn – Hà Nội SHB Finance

1.1.1. Giới thiệu chung về SHB Finance

Công ty Tài chính Ngân hàng TMCP Sài Gòn – Hà Nội SHB Finance là một công ty hoạt động và cung cấp dịch vụ trong lĩnh vực Tài chính tiêu dùng cho nhóm khách hàng đại chúng. "Công ty Tài Chính TNHH Ngân hàng TMCP Sài Gòn - Hà Nội (SHBFinance), được thành lập ngày 12/12/2016 (Mã số doanh nghiệp: 0107779290) với hình thức pháp lý ban đầu là Công ty Tài Chính TNHH MTV Ngân hàng TMCP Sài Gòn - Hà Nội. Ngày 2/6/2023, Ngân hàng Nhà nước Việt Nam đưa ra Quyết định chuyển đổi Công ty Tài chính TNHH MTV Ngân hàng TMCP Sài Gòn – Hà Nội thành Công ty Tài chính TNHH Ngân hàng TMCP Sài Gòn – Hà Nội." ("Giới Thiệu | SHBFinance," 2024).

- Tên công ty: Công ty Tài Chính TNHH Ngân hàng TMCP Sài Gòn - Hà Nội
- Tên tiếng Anh: SHB Consumer Finance Company Limited
- Tên gọi tắt: SHB Finance
- Địa chỉ: Tầng 6, Gelex Tower, 52 Lê Đại Hành, phường Lê Đại Hành, Quận Hai Bà Trưng, Hà Nội
- Điện thoại: 024 7109 8888
- Website: https://www.shbfinance.com.vn/
- Logo:

Hình 1: Logo thương hiệu công ty SHB Finance

Tiền thân là công ty Tài chính Vinaconex thuộc tập đoàn Viettel, hiện tại công ty hoạt động theo loại hình TNHH từ hai thành viên trở lên. Hiện tại, đang có hai tổ chức sở hữu 50% trên tổng số vốn điều lệ 1000 tỷ mỗi bên là Ngân hàng TMCP Sài Gòn – Hà Nội (SHB) và Ngân hàng TNHH Đại chúng Ayudhya (Krungsri) của Thái Lan.

1.1.2. Thành tựu, mục tiêu, tầm nhìn, sứ mệnh

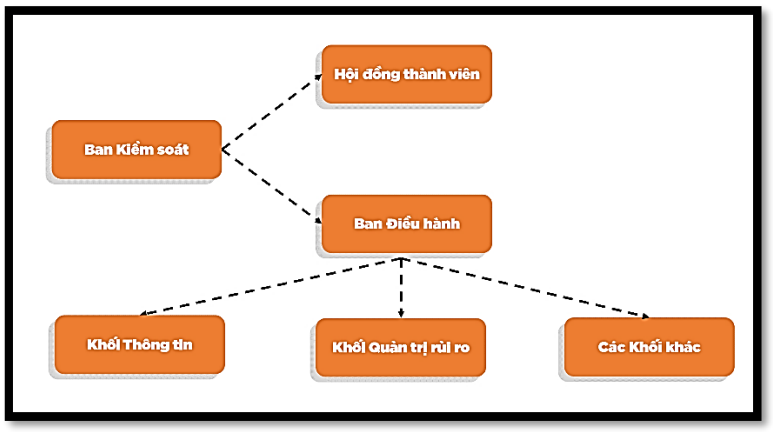
SHB Finance có nhiều điểm nổi bật về việc cung cấp dịch vụ trong lĩnh vực hoạt động là Tài chính tiêu dùng như:

- Top 6 Nhà tuyển dụng được yêu thích trong năm 2018
- Top 1 Nhà tuyển dụng được yêu thích nhất ngành nghề lĩnh vực Finance – Banking năm 2019
- Được Moody's[3] xếp hạng tín nhiệm lần đầu hạng Ba3 năm 2019
- Top 8 công ty tài chính tiêu dùng lớn nhất Việt Nam năm 2023
- Duy trì nợ xấu ở mức ổn định
- Phục vụ cho khoảng 300 nghìn hộ gia đình và 200 nghìn khách hàng

SHBFinance đưa ra tầm nhìn "Trở thành Công ty Tài chính Thuận tiện và Tin cậy với người dân Việt Nam" và sứ mệnh "Cung cấp các giải pháp Tài chính tiêu dùng thông minh, dễ tiếp cận cho mọi người dân Việt". Ngoài ra, SHB Finance cũng sở hữu bộ quy tắc ứng xử độc đáo, dựa trên năm chữ cái viết tắt "SHBFC" gồm: S – Smart, H – Honest, B – Brave, F – Friendly và C – Cooperative.

1.1.3. Cơ cấu tổ chức

Cơ cấu tổ chức của công ty Tài chính SHB Finance gồm: Ban Kiểm soát - giám sát, miễn/bổ nhiệm đối với Hội đồng thành viên và giám sát với Ban Điều hành. Hội đồng thành viên - bầu chọn Ban Điều hành và Ban Điều hành sẽ trực tiếp quản lý các Khối.



Hình 2: Sơ đồ tổ chức của công ty Tài chính SHB Finance

1.2. Mô tả bài toán

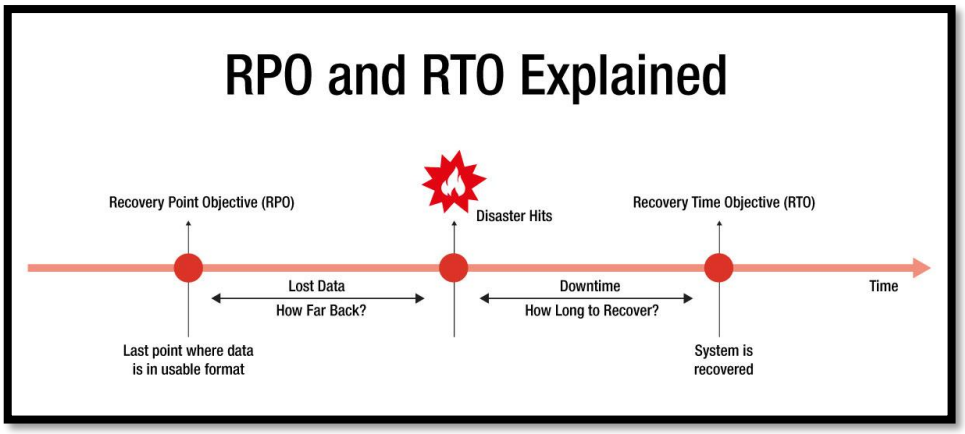
Sao lưu và khôi phục dữ liệu là nhiệm vụ thiết yếu, bắt buộc trong mỗi doanh nghiệp. Công nghệ thông tin ngày càng phát triển, dữ liệu được tạo ra không ngừng, một số công cụ sao lưu không đảm bảo về thời gian khôi phục và tính nhất quán của dữ liệu trong bản sao lưu. Do đó, doanh nghiệp cần ứng dụng giải pháp khác để khắc phục những hạn chế trên, đảm bảo khả năng vận hành ổn định.

[3] Moody's là một cơ quan đánh giá tín nhiệm tín dụng hàng đầu trên thế giới. Cung cấp các báo cáo và xếp hạng tín dụng cho các công ty, quốc gia và các tổ chức khác để đánh giá khả năng trả nợ của họ và độ rủi ro đầu tư.

1.2.1. Thực trạng

Trong quy trình sao lưu và khôi phục dữ liệu, Recovery Time Objective và Recovery Point Objective là hai chỉ số quan trọng, giúp doanh nghiệp xác định được mức độ khôi phục dữ liệu và thời gian khôi phục dữ liệu khi hệ thống gặp sự cố. Đây là nền tảng mà doanh nghiệp sử dụng để quản lý tính liên tục trong kinh doanh – Business Continuity Management (BCM).

- **Recovery Time Objective (RTO):** khoảng thời gian hệ thống được khôi phục, kể từ thời điểm xảy ra sự cố. Thể hiện tốc độ khôi phục dữ liệu.
- **Recovery Point Objective (RPO):** khoảng thời gian tối đa hệ thống chấp nhận mất dữ liệu, kể từ thời điểm sao lưu cuối cùng cho tới lúc xảy ra sự cố. Thể hiện tần suất trong việc sao lưu dữ liệu.



Hình 3: Minh họa khái niệm chỉ số RPO và RTO

Hai chỉ số RTO và RPO có mối liên hệ mật thiết với nhau, RPO càng nhỏ thì RTO càng nhỏ (dữ liệu được sao lưu liên tục, dẫn đến thời gian khôi phục nhanh), nhưng chi phí càng cao do yêu cầu hiệu năng và thiết bị lưu trữ lớn, đòi hỏi người quản trị viên cần có kế hoạch trong việc giám sát và thực hiện sao lưu liên tục.

SHBFinance sử dụng Oracle Database làm cơ sở dữ liệu cho các hệ thống khác ngoài hệ thống lõi, trong đó có hệ thống Xếp hạng tín dụng. Để sao lưu dữ liệu, SHBFinance sử dụng công cụ Recovery Manager (RMAN), thiết lập sao lưu theo lịch cho cơ sở dữ liệu của hệ thống Xếp hạng tín dụng.

Với công cụ RMAN, đội ngũ quản trị có thể sao lưu từ toàn bộ cho đến một phần cơ sở dữ liệu. Tuy nhiên, thời gian khôi phục của RMAN lên tới hàng giờ cho đến hàng ngày, chưa đáp ứng được việc có giá trị RTO nhỏ. Yu & cộng sự (2011) cho rằng, công cụ RMAN có nhược điểm như:

- **Hạn chế về băng thông:** RMAN gặp hạn chế khi trong quá trình sao lưu lượng dữ liệu lớn tới thiết bị lưu trữ không cùng địa điểm với CSDL. Quá trình sao lưu cần rất nhiều tài nguyên cũng như độ tin cậy của hệ thống mạng, gây rủi ro về mặt hiệu suất và việc gián đoạn trong quá trình sao lưu.

- **Hạn chế về thời gian khôi phục:** Trong trường hợp toàn bộ hệ thống trong cùng một trung tâm dữ liệu gặp sự cố, việc khôi phục dữ liệu từ bản sao lưu bằng công cụ RMAN tiêu tốn về mặt thời gian, và lượng dữ liệu mất mát là rất lớn.

1.2.2. Thách thức

Khi hệ thống gặp vấn đề, như sụp đổ hoặc gián đoạn, gây ảnh hưởng đến hoạt động kinh doanh hàng ngày như mất mát dữ liệu, không thể truy cập thông tin quan trọng để đưa ra quyết định kịp thời, gián đoạn trong dịch vụ có thể dẫn đến mất lòng tin từ phía khách hàng, gây tổn hại về uy tín và doanh số của công ty.

Việc phục hồi hệ thống có thể đòi hỏi nhiều thời gian và chi phí, và nếu không có kế hoạch hồi phục hiệu quả có thể làm gia tăng rủi ro, ảnh hưởng đến sức mạnh cạnh tranh của công ty trong ngành.

Hơn hết, trong một môi trường kinh doanh đầy cạnh tranh như ngành tài chính, việc duy trì sự ổn định của hệ thống là một yếu tố quan trọng để bảo vệ uy tín và sự phát triển của doanh nghiệp.

1.2.3. Giá trị mang lại

Để giải quyết thách thức vấn đề tần suất sao lưu và thời gian khôi phục cơ sở dữ liệu hệ thống Xếp hạng tín dụng của SHBFinance, doanh nghiệp cần triển khai giải pháp hiệu quả có khả năng sao lưu liên tục và khôi phục dữ liệu tự động, giảm thiểu thời gian gián đoạn và mất mát dữ liệu. Lợi ích mang lại có thể kể đến như:

- Khôi phục dữ liệu kịp thời: khôi phục dữ liệu nhanh chóng sau sự cố, giảm thời gian gián đoạn và đảm bảo tính liên tục trong hoạt động kinh doanh.
- Tiết kiệm chi phí và thời gian: giảm chi phí và thời gian phục hồi sau sự cố, loại bỏ việc tiêu tốn nguồn lực nhân sự để thực hiện các công việc sao lưu và khôi phục thủ công. Điều này giúp tổ chức tối ưu hóa hiệu suất và tài nguyên, tập trung vào các hoạt động kinh doanh chính.
- Dữ liệu được sao lưu liên tục: dữ liệu được sao lưu với tần suất cao, tự động và sát với thời gian thực của CSDL, đảm bảo tính toàn vẹn của dữ liệu.

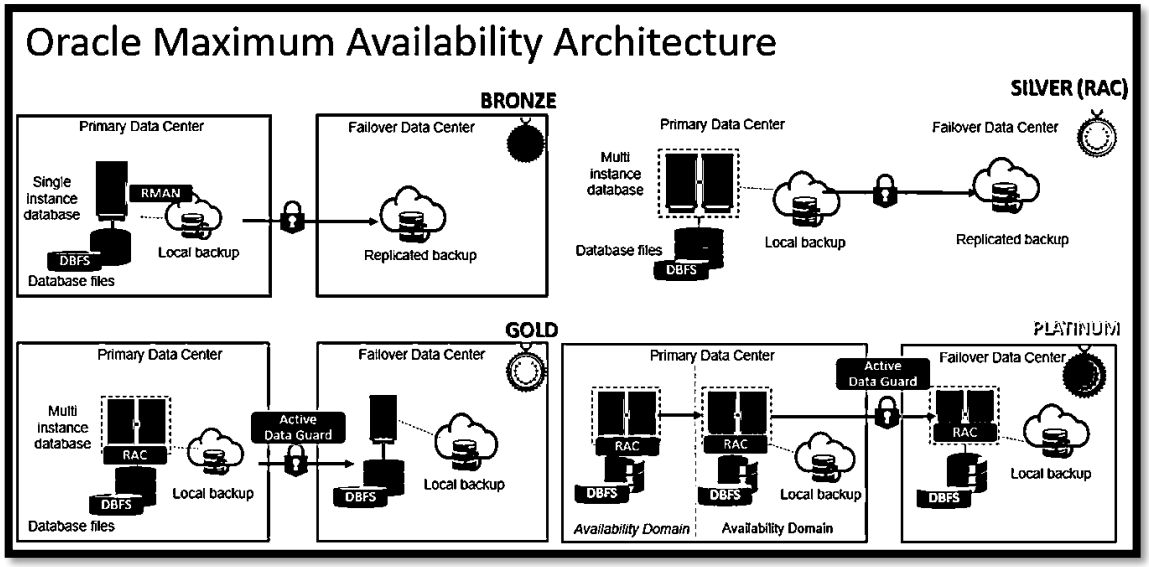
1.2.4. Hướng giải quyết

Giải pháp đảm bảo tính liên tục của Oracle - *Oracle Maximum Availability Architecture (MAA)* được chia ra làm bốn mô hình kiến trúc, phân chia theo hạng với mức độ bảo vệ tăng dần từ trái sang phải: Bronze, Silver, Gold và Plantinum. Mức độ bảo vệ được thể hiện qua các tính năng như bảo vệ dữ liệu, tính sẵn sàng cao, khôi phục sau thảm họa.

Hai chỉ số RTO và RPO được cải thiện thông qua hạng bậc, lượng dữ liệu khôi phục và thời gian khôi phục được đánh giá thông qua ba trường hợp chính: lỗi máy chủ (Server), lỗi thiết bị lưu trữ (Storage) và lỗi toàn bộ trung tâm dữ liệu (Data Center).

Hạng	Công nghệ	RTO/RPO (Server, Storage, Data Center)
Bronze	CSDL dạng Single Instance, công nghệ sao lưu bằng RMAN	RPO: thấp nhất là 0, nhiều nhất là x (tiếng) RTO: thấp nhất là x (phút), nhiều nhất là x (ngày)
Silver	Bronze + CSDL dạng Multi – Instance với công nghệ Real Application Cluster (RAC)	RPO: thấp nhất là 0, nhiều nhất là x (tiếng) RTO: thấp nhất là x (phút), nhiều nhất là x (ngày)
Gold	Silver + CSDL dự phòng với công nghệ Oracle Data Guard	RPO: thấp nhất là 0, nhiều nhất là x (phút) RTO: thấp nhất là x (phút), nhiều nhất là x (phút)
Plantinum	Gold + một số tính năng khác như GoldenGate, Sharding	RPO: thấp nhất là 0, nhiều nhất là x (phút) RTO: thấp nhất là x (phút), nhiều nhất là x (phút)

Bảng 1: Bốn kiến trúc trong giải pháp Oracle MAA



Hình 4: Khung tham chiếu giải pháp Oracle MAA

Cơ sở dữ liệu hệ thống Xếp hạng tín dụng nội bộ của SHBFinance hiện tại đang sử dụng kiến trúc RAC và sao lưu với RMAN (Silver). Với công nghệ RAC, hệ thống

Xếp hạng tín dụng đáp ứng được việc mở rộng hệ thống, cho phép hai hoặc nhiều Oracle Instance sử dụng chung một thiết bị lưu trữ, có khả năng cân bằng tải khi có lượng truy cập lớn từ các chi nhánh SHBFinance. Tuy nhiên, về quá trình sao lưu và khôi phục, công cụ RMAN chưa đáp ứng được thời gian khôi phục nhỏ và tần suất sao lưu lớn.

Vì vậy, để giải quyết bài toán đã nêu, SHBFinance cần nâng cấp mô hình kiến trúc từ Silver lên Gold - ứng dụng được tính năng Oracle Data Guard, hướng tới việc giảm chỉ số RTO và RPO.

Giải pháp Oracle Data Guard có các ưu điểm như sau:

- Sao lưu dữ liệu theo thời gian thực: Oracle Data Guard cho phép sao lưu dữ liệu từ một CSDL chính sang một hoặc nhiều CSDL dự phòng đồng bộ theo thời gian thực, đảm bảo các CSDL dự phòng luôn được cập nhật dữ liệu mới nhất.
- Khôi phục kịp thời: Trong trường hợp sự cố xảy ra ở CSDL chính, Oracle Data Guard tự động chuyển đổi sang một trong các CSDL dự phòng, giúp giảm thiểu thời gian dừng hoạt động hệ thống.
- Phục vụ nhiều mục đích: Oracle Data Guard hỗ trợ nhiều loại CSDL ở chế độ dự phòng gồm Physical, Logical và Snapshot. Trong đó, Snapshot thường được dùng cho môi trường kiểm thử.
- Quản lý tự động: Oracle Data Guard cung cấp các công cụ quản lý tự động để giảm thiểu các tác vụ thủ công của người quản trị, gồm cả việc tự động thực hiện các hoạt động như sao lưu và giám sát.
- Giảm tải cho hệ thống chính: CSDL dự phòng có thể được dùng để truy vấn dữ liệu được đồng bộ, giảm tải cho CSDL chính.

Thông qua các ưu điểm nổi trội của Oracle Data Guard đem lại, đây là giải pháp phù hợp với thực trạng dự phòng của CSDL hệ thống Xếp hạng tín dụng hiện nay. Việc triển khai Oracle Data Guard giúp SHBFinance hướng tới mục tiêu giảm chỉ số RTO và RPO, đảm bảo tính liên tục trong kinh doanh của doanh nghiệp.

1.3. Giới thiệu về Oracle Data Guard

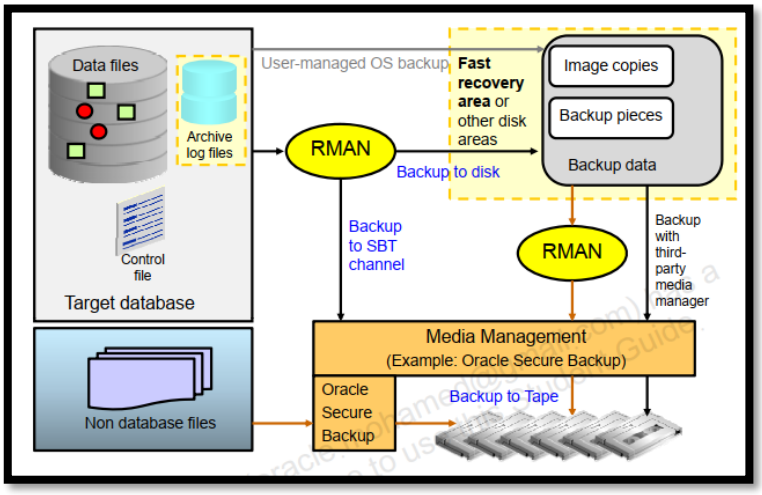
1.3.1. Sao lưu với công cụ Recovery Manager

Recovery Manager (RMAN) là một công cụ quản lý sao lưu và khôi phục dữ liệu của Oracle, được thiết kế để quản lý và bảo vệ dữ liệu trong các cơ sở dữ liệu Oracle. Mục đích cơ bản của RMAN là cung cấp các tính năng và công cụ cho việc thực hiện các tác vụ sao lưu và khôi phục dữ liệu một cách linh hoạt, hiệu quả và an toàn.

Đặc điểm nổi bật chính của RMAN là ngoài việc sao lưu khi CSDL đang tắt (Offline/Consistent/Cold), có thể sao lưu ngay cả lúc CSDL đang hoạt động

(Online/Inconsistent/Hot). Ngoài ra, RMAN cung cấp tính năng sao lưu một phần, gồm chỉ những thay đổi kể từ bản sao lưu trước đó. Tương tự đối với việc khôi phục, công cụ cho phép khôi phục một phần hoặc toàn phần.

Các thành phần mà RMAN sao lưu gồm: Data Files, Control Files, Archived Redo Log, Parameters File. Đối với mức độ block – đơn vị nhỏ nhất trong kiến trúc lưu trữ vật lý, khi sao lưu, RMAN tự động kiểm tra những block rỗng và sẽ bỏ qua block này. Công cụ được tích hợp lên giao diện quản trị là Oracle Enterprise Manager và cũng có thể sử dụng bởi lệnh SQL. Ngoài ra, RMAN được tích hợp thêm thành phần Oracle Secure Backup để mã hóa và sao lưu ra ổ đĩa dạng băng từ (tape) hoặc sao lưu lên đám mây một cách bảo mật, an toàn.



Hình 5: Kết hợp RMAN, Oracle Secure Backup và sao lưu bằng lệnh hệ thống

1.3.2. Khái niệm, kiến trúc của Oracle Data Guard

Khái niệm

Trong các giải pháp phục hồi sau sự cố, Oracle Data Guard (ODG) là một công nghệ được đánh giá cao trong việc đảm bảo tính sẵn sàng và liên tục của CSDL Oracle. ODG được xây dựng và tích hợp trên CSDL Oracle, gồm nhiều mô-đun chức năng như quản lý, giám sát, duy trì một hoặc nhiều CSDL dự phòng nhằm mục đích bảo vệ quy trình vận hành của doanh nghiệp khỏi sự cố.

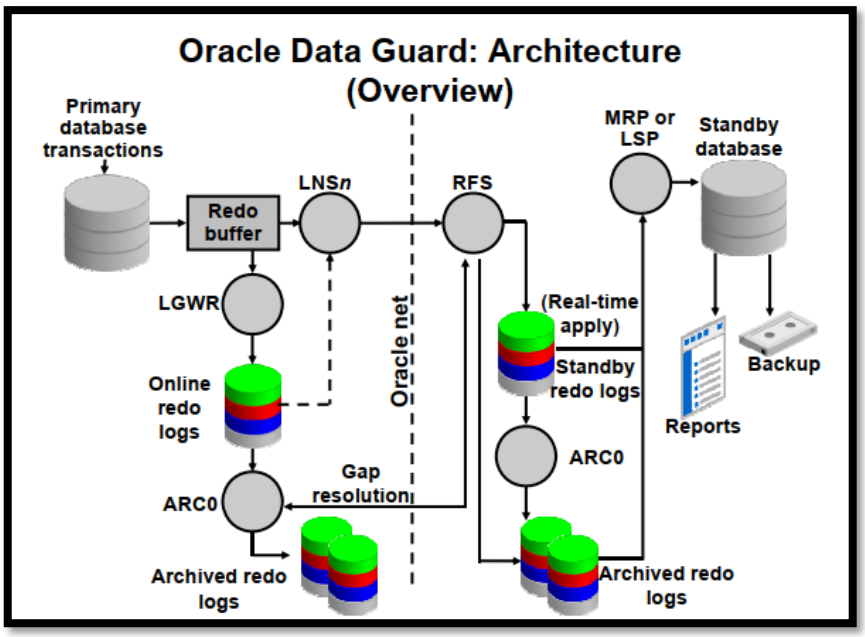
Fuller (2014) cho rằng, CSDL dự phòng duy trì sự ổn định này với vai trò là bản sao và đồng bộ với CSDL chính. CSDL dự phòng có thể đặt cách xa trung tâm dữ liệu của CSDL chính để cải thiện mức an toàn, phòng trường hợp sự cố xảy ra cùng một địa điểm. Khi CSDL chính gặp lỗi, ODG sẽ chuyển đổi vận hành (switch role) từ CSDL chính – đang bị ảnh hưởng sang CSDL dự phòng, CSDL dự phòng sẽ đảm nhận vai trò của CSDL chính.

Kiến trúc

Trong kiến trúc của giải pháp ODG gồm 01 CSDL chính và CSDL dự phòng. Các CSDL liên kết và giao tiếp với nhau thông qua Oracle Network Service - môi

trường mạng của Oracle, các CSDL này có thể đặt cách xa nhau về mặt địa lý. Có nhiều cách triển khai Oracle Data Guard, các CSDL có thể trong cùng một máy chủ và trung tâm dữ liệu, khác máy chủ hoặc khác trung tâm dữ liệu.

Trong giải pháp ODG, CSDL chính gửi những thông tin thay đổi được lưu dưới dạng véc-tơ có trong Online Redo Logs (ORLs) hoặc Redo Buffer Cache (RBC), giữ cho CSDL dự phòng luôn được đồng bộ hóa thông qua việc nhận và áp dụng.



Hình 6: Kiến trúc tổng quan giải pháp Oracle Data Guard

Trong kiến trúc giải pháp ODG sử dụng các tiến trình sao lưu và chuyển đổi vai trò. Môi trường ODG gồm cả tiến trình chung và tiến trình chỉ xuất hiện khi sử dụng giải pháp ODG. Các tiến trình và thành phần của giải pháp ODG gồm:

- Standby Redo Logs: thành phần lưu trữ thông tin thay đổi của CSDL dự phòng.
- Log Writer Network Server (LNSs): tiến trình nhận thông tin thay đổi và chuyển cho CSDL dự phòng.
- Remote File Server (RFS): tiến trình nhận thông tin thay đổi từ LNSs.
- Managed Recovery Process/Logical Standby Process (MRP/LSP): tiến trình áp dụng các thay đổi từ Standby Redo Log trên CSDL dự phòng. Tiến trình MRP cho CSDL dạng Physical và LSP sử dụng cho CSDL dạng Logical.

1.3.3. Loại hình đồng bộ

Giải pháp ODG cung cấp nhiều loại hình đồng bộ, mỗi loại hình có đặc điểm khác nhau, linh hoạt và phù hợp với nhu cầu của doanh nghiệp. Có thể phân loại các loại hình theo hai góc độ: về loại hình CSDL dự phòng và về mức độ bảo vệ trong cơ chế truyền/đồng bộ hóa thông tin thay đổi giữa các CSDL.

Phân loại theo loại hình CSDL dự phòng

- CSDL dự phòng vật lý (Physical):
  - Có cấu trúc File Systems và dữ liệu giống với CSDL chính.
  - Được đồng bộ hóa với CSDL chính thông qua việc áp dụng dữ liệu thay đổi (redo data) từ CSDL chính.
  - Cho phép thực hiện đồng thời tác vụ trả kết quả truy vấn cũng như việc nhận và áp dụng dữ liệu thay đổi vào CSDL. Chỉ mở CSDL ở chế độ chỉ đọc (read only).
- CSDL dự phòng lô-gíc (Logical):
  - Chỉ giống ở cấu trúc lô-gic, bộ nhớ vật lý có thể sử dụng các tính năng khác như Oracle Automatic Storage Management (ASM) để quản lý tập tin khác với File Systems theo mặc định. Ngoài ra, không có view, index giống với CSDL chính.
  - Được đồng bộ hóa với CSDL chính thông qua việc nhận, chuyển hóa redo data thành SQL và thực thi trên CSDL dự phòng để áp dụng thay đổi. Điều này được thực hiện nhờ công cụ phân tích tệp tin lưu trữ logs là LogMiner.
  - Cho phép thực hiện đồng thời các tác vụ trả kết quả truy vấn, áp dụng dữ liệu thay đổi vào CSDL, đặc biệt hơn là cho phép chỉnh sửa đối với các bảng, đối tượng không nằm trong vùng được áp dụng thay đổi. Mở CSDL ở chế độ đọc/ghi (Read/Write)
- CSDL dự phòng Snapshot:
  - Được chuyển từ CSDL dự phòng dạng vật lý.
  - Cho phép thực hiện đọc/ghi trên toàn bộ cơ sở dữ liệu với mục đích kiểm thử.
  - Sẽ không nhận và áp dụng các thông tin thay đổi.
  - Các thay đổi sẽ bị ROLLBACK lại nếu như chuyển về CSDL dự phòng dạng vật lý.

Phân loại theo chế độ bảo vệ

Khi phân loại theo chế độ bảo vệ, cấu hình của các chế độ phụ thuộc vào các đối số được cài đặt (cụ thể là trong tham số *LOG_ARCHIVE_DEST_n*), có 04 đối số chính như sau:

- SYNC: Xác nhận các redo data được gửi sang CSDL dự phòng thành công trước khi giao dịch được đánh dấu là COMMIT, nếu không, hệ thống sẽ dừng hoạt động/tiếp tục tùy thuộc vào chế độ bảo vệ được chọn.
- ASYNC: Không xác nhận việc redo data được nhận bởi CSDL dự phòng, do đó, giao dịch có thể COMMIT ngay lập tức trên CSDL chính.

- AFFIRM: Gửi tín hiệu tín hiệu Acknowledgement (ACK) *sau khi* redo data nhận được đã được ghi vào Standby Redo Logs.
- NOAFFIRM: Khác với AFFIRM ở chỗ sẽ gửi tín hiệu ACK, nhưng *gửi trước* khi được ghi vào Standby Redo Logs.

Ngoài các đối số, CSDL cũng cần phải cấu hình chế độ bảo vệ, đảm bảo CSDL sẽ thực hiện các phương thức bảo vệ dữ liệu khác nhau khi gặp sự cố.:

- Ưu tiên bảo vệ (max. protection):
  - Chế độ bảo vệ mà CSDL chính sẽ đảm bảo rằng không có dữ liệu nào bị sót một cách tuyệt đối trong trường hợp CSDL chính gặp sự cố như thảm họa thiên tai, bị lỗi mạng hoặc bị lỗi với CSDL dự phòng.
  - CSDL chính sẽ dừng hoạt động khi gặp sự cố khiến cho các tiến trình trong việc truyền tải/đồng bộ hóa thông tin thay đổi không thể ghi vào ít nhất một trong các CSDL dự phòng.
  - Cần thiết lập chế độ cho cách truyền redo data với hai đối số: SYNC – đồng bộ hóa, và AFFIRM – xác nhận đã ghi xuống đĩa vật lý tại Standby Redo Logs, với ít nhất một CSDL dự phòng có Standby Redo Logs.
- Ưu tiên tính sẵn sàng (max. availability):
  - Chế độ bảo vệ mà CSDL chính sẽ đảm bảo không có dữ liệu nào bị sót nhưng không hoàn toàn tuyệt đối, vì không tác động tới việc vận hành của CSDL chính trong một ràng buộc về thời gian cho trước
  - Nếu có sự cố, CSDL chính sẽ hoạt động theo cách thức không đồng bộ (ASYNC, hoạt động không đợi xác nhận CSDL dự phòng đã nhận redo data hay chưa) cho đến khi ít nhất một CSDL dự phòng được đồng bộ về mặt thông tin thay đổi và chuyển về SYNC.
  - Cần thiết lập hai đối số: SYNC – đồng bộ hóa và NOAFFIRM (không cần xác nhận đã ghi vào Standby Redo Logs) hoặc AFFIRM cho ít nhất một CSDL dự phòng có chứa Standby Redo Logs.
- Ưu tiên hiệu năng hệ thống (max. performance):
  - Đây là chế độ mặc định của giải pháp ODG. Cung cấp việc bảo vệ dữ liệu thấp hơn hai mức còn lại về tính vẹn toàn, nhưng hiệu năng hệ thống của CSDL chính cao hơn.
  - Giao dịch được xác nhận COMMIT, đồng thời thông tin thay đổi sẽ được lưu xuống tệp tin lưu trữ thông tin thay đổi ngay lập tức mà không cần quá trình xác nhận ACK từ CSDL dự phòng.
  - Thông tin thay đổi (redo data) được truyền tới CSDL dự phòng theo cách không đồng bộ (ASYNC) với những thông tin thay đổi đã được COMMIT.

- Cần cấu hình đối số như sau: ASYNC – không đồng bộ và NOAFFIRM – không xác nhận đã ghi cho CSDL dự phòng đã có Standby Redo Logs.

Bảng so sánh về các chế độ bảo vệ dưới đây sẽ có cái nhìn tổng quan và ngắn gọn hơn. Đây là các kết hợp đối số có ý nghĩa, có một số trường hợp kết hợp đối số khác không được sử dụng như ASYNC/AFFIRM:

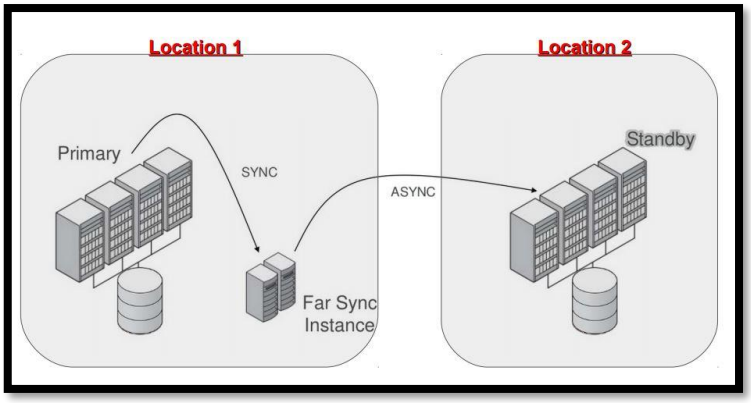
Chế độ	**Rủi ro**	**Chế độ ruyền**	**Nếu tín hiệu ACK không gửi (AFFIRM/NOAFFIRM)**
Ưu tiên bảo vệ	Không mất đồng bộ dữ liệu	SYNC	CSDL chính sẽ treo/dừng hoạt động cho đến khi nhận được tín hiệu ACK
Ưu tiên tính sẵn sàng	Không mất đồng bộ dữ liệu khi CSDL chính gặp sự cố hoặc mất khi cả hai gặp sự cố	SYNC/AFFIRM hoặc SYNC/NOAFFIRM (Fast Sync)	CSDL chính sẽ chờ trong một khoảng thời gian được xác định là NetTimeout, nếu hết, sẽ tiếp tục hoạt động
Ưu tiên hiệu năng	Mất đồng bộ dữ liệu nếu CSDL chính gặp sự cố	ASYNC	CSDL chính không chờ tín hiệu ACK

Bảng 2: Cấu hình đối số phương thức truyền/xác nhận

Đồng bộ trung gian với Far Sync

Tính năng Far Sync xây dựng một Instance làm điểm trung chuyển thông tin thay đổi tới nhiều các CSDL dự phòng khác. Thực tế, Far Sync là một hệ thống có kích thước gọn nhẹ, tiêu thụ ít tài nguyên về lưu trữ cũng như xử lý.

Far Sync giống với CSDL dự phòng thông thường như quản lý các Control File, nhận redo data vào Standby Redo Logs và lưu trữ khi Log Switch xuống Archived Redo Logs. Ngược lại, khác với CSDL dự phòng ở điểm không lưu trữ một số tệp tin không quan trọng như Data Files, không hỗ trợ chuyển đổi vai trò và chỉ thiết lập được tại 02 chế độ bảo vệ: *max. performance* hoặc *max. availability*.



Hình 7: Tính năng Far Sync trong giải pháp Oracle Data Guard

Far Sync không bị giới hạn bởi khoảng cách do cách truyền không đồng bộ, không ảnh hưởng đến hiệu năng của CSDL chính. Hơn hết, Far Sync giúp giảm tải CSDL chính trong việc truyền tải redo data đến hàng loạt CSDL dự phòng. Nếu CSDL chính lỗi, các CSDL dự phòng sẽ lấy các redo data cuối cùng từ Far Sync trong việc chuyển đổi vai trò, đảm bảo tính vẹn toàn. Ngoài ra, việc bổ sung các tính năng như đóng gói, mã hóa dữ liệu được Far Sync đảm nhận nhằm giảm tải cho CSDL chính.

1.3.4. Cơ chế tương tác giữa các thành phần

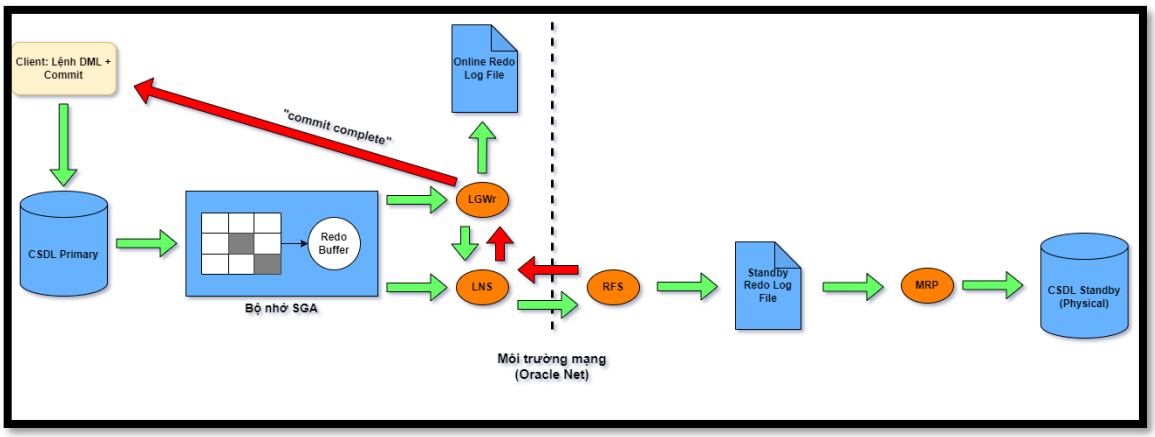
Các thành phần giao tiếp với nhau trong kiến trúc giải pháp Oracle Data Guard thông qua 03 cơ chế chính:

- Cơ chế vận chuyển thông tin thay đổi (Redo Transport Services): gồm các tiến trình ARCn, LGWr, LNSs, RFS phục vụ việc truyền redo data.
- Cơ chế áp dụng thông tin thay đổi (Log Apply Services): gồm các tiến trình MRP, LSP phục vụ việc áp dụng các redo data vào CSDL dự phòng.
- Cơ chế quản lý/chuyển đổi vai trò (Role Management Services): gồm tiến trình DMON và DGR thực hiện việc giám sát và thay đổi vai trò của các CSDL.

Hai chế độ bảo vệ *Ưu tiên hiệu năng* và *Ưu tiên bảo vệ* có những đặc điểm nổi trội khác nhau, đặc biệt hơn so với *Ưu tiên về tính sẵn sàng*. Vì vậy, để hiểu được rõ được sự khác biệt và cơ chế hoạt động của luồng dữ liệu, hình vẽ cùng miêu tả các pha sẽ minh họa cơ chế hoạt động hai chế độ, hoạt động trong loại CSDL dự phòng vật lý.

Với các hình vẽ minh họa, các mũi tên có tông màu nhạt thể hiện luồng truyền redo data từ CSDL chính, tông màu đậm thể hiện luồng trả về tín hiệu ACK từ CSDL dự phòng.

Chế độ bảo vệ *Ưu tiên bảo vệ*



Hình 8: Luồng hoạt động của Oracle Data Guard với chế độ Ưu tiên bảo vệ

Các pha của cơ chế *Ưu tiên bảo vệ* được thể hiện tuần tự với các bước như sau, khi cấu hình SYNC/AFFIRM:

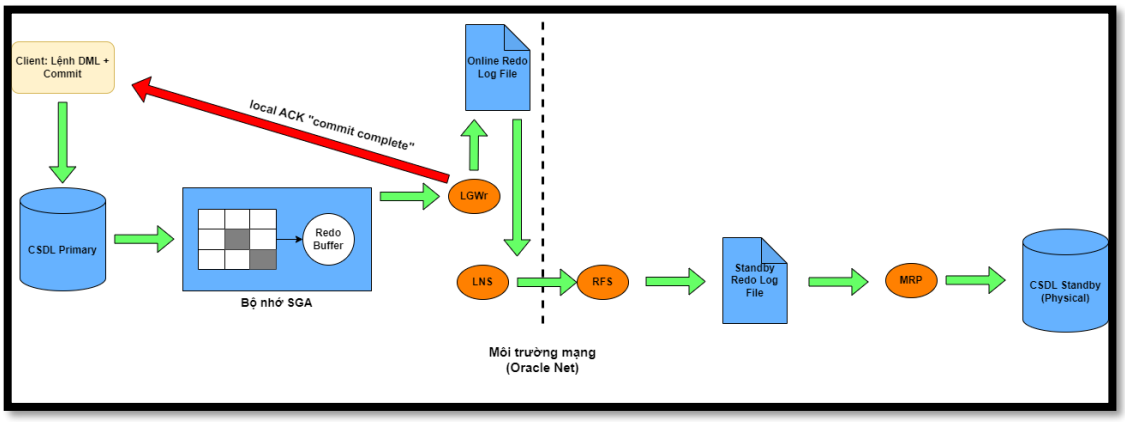
1. Người dùng tạo các giao dịch thông qua các lệnh Data Manipulation Language (DML). Thông tin thay đổi được lưu vào Redo Buffer Cache.
2. Thông tin thay đổi sẽ được tiến trình LGWR đưa và lưu xuống Online Redo Log sau khi COMMIT, nhưng hệ thống chưa báo “commit complete” ngay.
3. Tiến trình LNS sẽ nhận những thông tin thay đổi trong lúc LGWR xuất ra. Những thông tin thay đổi này sẽ được chuyển cho tiến trình RFS thuộc CSDL dự phòng, nhằm thực hiện sao lưu.
4. Sau khi nhận được thông tin thay đổi, tiến trình RFS sẽ ghi những thông tin thay đổi xuống Standby Redo Log Files.
5. Nếu CSDL dự phòng sử dụng tính năng Real-time Apply, thì ngay lập tức, các thông tin thay đổi này sẽ được áp dụng vào dữ liệu lưu trữ vật lý tại CSDL dự phòng với tiến trình khôi phục MRP.
6. Với SYNC/AFFIRM, sau khi dữ liệu đã được áp dụng thành công, tiến trình RFS sẽ phản hồi lại cho tiến trình LNS tín hiệu ACK. Lúc này, hệ thống mới phản hồi người dùng “commit complete”. Tại đây, nếu gặp sự cố về môi trường mạng, khiến cho RFS không thể gửi cho LNS, CSDL chính sẽ đợi tới khi nào nhận được thông tin dẫn tới hệ thống treo

Ngoài ra, khi xảy ra Log Switch trên CSDL chính, sẽ kích hoạt một Trigger giúp CSDL dự phòng cũng thực hiện Log Switch đối với Standby Redo Log Files nhằm đảm bảo tính toàn vẹn.

Đối với RFS, tiến trình này sẽ gửi trực tiếp redo data xuống Archive Log File nếu: Không có Standby Redo Logs (1), Standby Redo Log được cài đặt nhỏ hơn kích thước của Online Redo Logs (2), tất cả Standby Redo Logs đều chưa được lưu trữ

(archived) (3) và nếu RFS thực hiện nhận redo data từ tiến trình ARCn trong cơ chế Gap Resolution (4).

Chế độ *Ưu tiên hiệu năng*



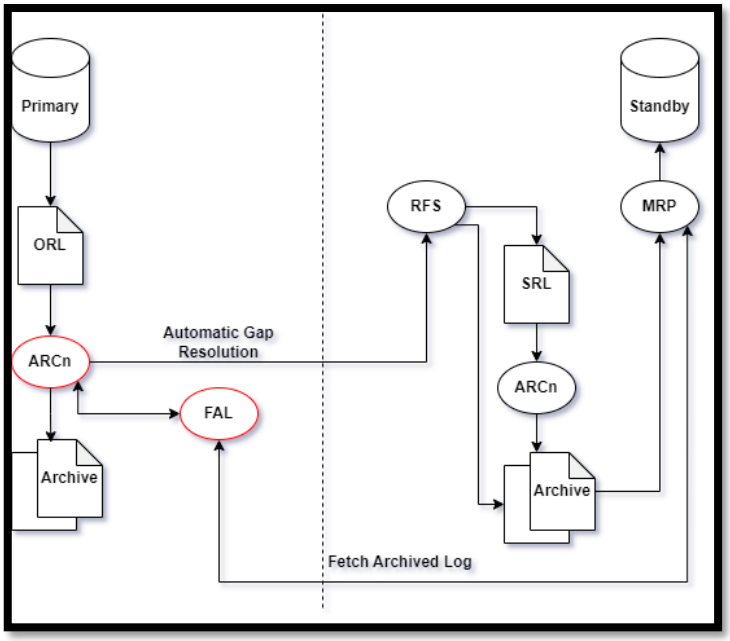
Hình 9: Luồng hoạt động của Oracle Data Guard với chế độ Ưu tiên bảo vệ

Trong chế độ ưu tiên hiệu năng, với cơ chế ASYNC/NOAFFIRM, CSDL chính sẽ không yêu cầu nhận tín hiệu xác nhận nào từ CSDL dự phòng trong việc ghi thành công vào SRLs. Các redo data được LNS tiếp nhận và lấy tại Online Redo Logs nhằm tăng hiệu năng của tiến trình LGWr trong việc ghi redo data.

Khi người dùng gõ lệnh DML và yêu cầu COMMIT, hệ thống sẽ ngay lập tức trả lại tín hiệu COMMIT thành công – "commit complete", do không phải chờ phản hồi từ CSDL dự phòng. Trường hợp này, tín hiệu ACK còn được gọi là "local ACK", có nghĩa là tín hiệu này sẽ không xuất phát từ tiến trình RFS truyền qua môi trường Oracle Net.

Hiện tượng Archive Redo Gap và cơ chế xử lý

Trong môi trường Oracle Data Guard, tình trạng trễ dữ liệu xảy ra khi kết nối giữa các CSDL gặp sự cố hoặc gói tin gửi bị hỏng. Thuật ngữ để miêu tả tình trạng này là "Archive Redo Gap Sequence". Oracle cung cấp 02 cơ chế để xử lý tình trạng này là "Automatic Gap Resolution" và "Fetch Archive Log – FAL".



Hình 10: Minh họa cơ chế xử lý thiếu trong việc truyền thông tin thay đổi

Trong chế độ *Ưu tiên hiệu năng*, do cơ chế ASYNC, các redo data được COMMIT liên tục và không đợi xác nhận từ phía CSDL dự phòng. Khi gặp sự cố mạng, tiến trình LNS của CSDL dự phòng không thể nhận những thông tin thay đổi. Trong một khoảng thời gian như trên sẽ gây ra hiện tượng trễ dữ liệu.

Với cơ chế Automatic Gap Resolution:

- Đây là cơ chế chủ động của CSDL chính. Tiến trình ARCn của CSDL chính sẽ liên tục gửi lệnh ping đến cho tiến trình RFS của CSDL dự phòng để xác định trạng thái khi có kết nối.
- Nếu kết nối mạng khôi phục, ARCn sẽ tiến hành ping kèm truy vấn xác định thông tin về Archive Redo Logs mới nhất. Nếu Log Sequence tại CSDL chính lớn hơn Log Sequence tại CSDL dự phòng, RFS sẽ thông báo lại số lượng Archive Redo Logs còn thiếu, từ đây ARCn sẽ gửi cho CSDL dự phòng theo yêu cầu.
- Ngoài ra, tiến trình LNS cũng sẽ hỗ trợ trong việc cập nhật/gửi đi các redo data mới nhất từ Redo Buffer Cache, Online Redo Logs, giúp CSDL dự phòng nhanh chóng khôi phục lại được trạng thái cập nhật gần nhất với CSDL chính, cho đến khi MRP có thể tiếp tục áp dụng với thời gian thực.

Với cơ chế Fetch Archive Log (FAL):

- Đây là cơ chế chủ động của CSDL dự phòng (dạng vật lý), do hành động "fetch" – yêu cầu/lấy thông tin đang bị thiếu được thực hiện một cách chủ động.
- Sau khi được tiến trình ARCn gửi và RFS nhận vào, CSDL dự phòng sẽ cập nhật trong Control File của nó về tên và địa điểm của các Archive Log. Khi MRP thấy được những thay đổi mới trong Control File, hệ thống sẽ tiến hành áp

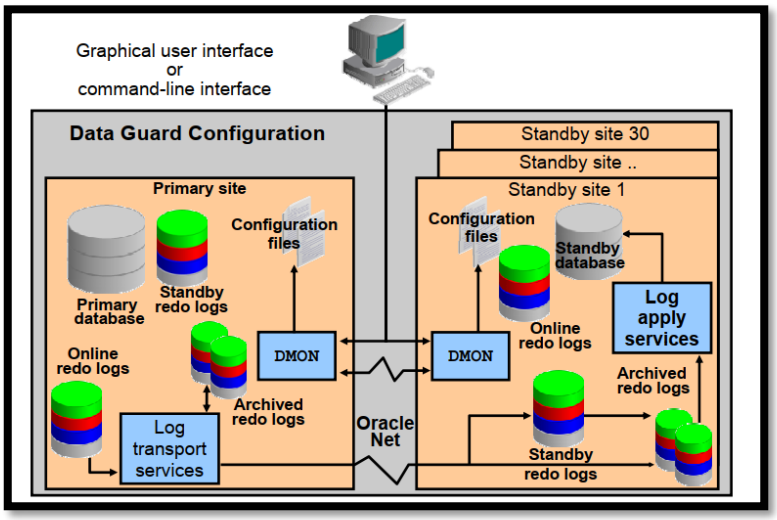
dụng những thay đổi này vào CSDL dự phòng. Nhưng khi MRP thấy thông tin về tệp Archive Log từ CSDL chính lỗi/hỏng/thiếu thì nó sẽ sử dụng cơ chế chủ động thông qua tiến trình FAL, yêu cầu gửi lại Archive Log. Cơ chế này cần thiết lập hai tham số chính trên CSDL dự phòng như sau:

Tham số	Miêu tả
FAL_SERVER	Trỏ tới CSDL nhận yêu cầu và xử lý gửi thông tin thiếu
FAL_CLIENT	Trỏ tới CSDL yêu cầu gửi thông tin thiếu

Bảng 3: Tham số cấu hình cho cơ chế Fetch Archive Log

1.3.5. Oracle Data Guard Broker

Theo Fuller (2014), Data Guard Broker là một tính năng được tích hợp trong Oracle Database Server, dùng để quản trị tập trung các CSDL thuộc môi trường Oracle Data Guard. Các thành phần của Broker gồm: trình điều khiển (thuộc client-side), tiến trình DMON và configuration files (thuộc server-side).



Hình 11: Kiến trúc Oracle Data Guard với tính năng Data Guard Broker

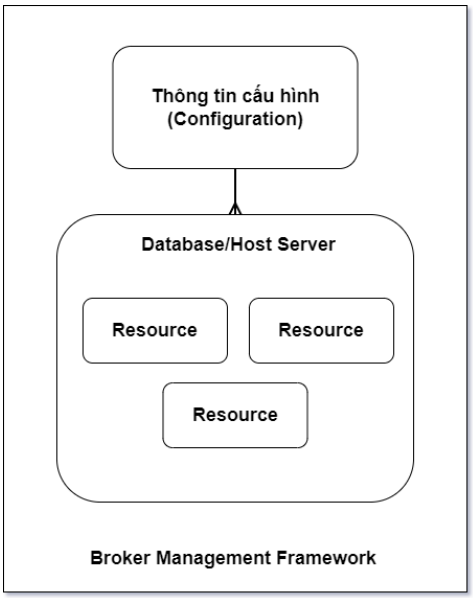
Tính năng Data Guard Broker có các ưu điểm sau khi so sánh với việc không sử dụng Data Guard Broker:

	Sử dụng Broker	Không sử dụng Broker
Quản trị	Quản trị tập trung	Việc quản trị thực hiện riêng lẻ đối với từng CSDL
Tạo CSDL dự phòng	Sử dụng OEMCC để đơn giản hóa và tự động việc tạo	Tạo thủ công (bằng cách sử dụng công cụ RMAN hoặc các công

	CSDL dự phòng (với GUI) (gồm control file, online redo log files, datafiles, và các tệp tin lưu trữ tham số)	cụ khác): - Sao chép các tệp tin của CSDL chính cho CSDL dự phòng - Tạo control file trên CSDL dự phòng - Tạo file tham số trên CSDL dự phòng - Sao chép password file từ CSDL chính sang CSDL dự phòng
Cấu hình và quản lý	Cho phép cấu hình và quản lý nhiều CSDL tập trung và quản lý cấu hình kết nối của các CSDL thông qua một tệp tin duy nhất	- Thiết lập Redo Transport Services và Log Apply Services tại mỗi CSDL - Quản lý CSDL riêng lẻ
Điều khiển	- Tự động thiết lập Redo Transport Services và Log Apply Services - Đơn giản hóa việc chuyển đổi vai trò	- Sử dụng SQL để quản lý - Sử dụng nhiều lệnh hệ thống để quản lý các CSDL cho việc chuyển đổi vai trò cũng như điều khiển các tiến trình
Theo dõi	- Cho phép theo dõi hiệu năng hệ thống, cấu hình và các tham số khác - Cung cấp báo cáo chi tiết về hệ thống	- Chỉ theo dõi cố định vào khoảng thời gian nhất định - Không tập hợp các tham số cần theo dõi cùng một lúc

Bảng 4: So sánh việc sử dụng Broker vào hệ thống

Trong mô hình quản lý của Broker, tập tin cấu hình Configuration chứa thông tin về các CSDL, gồm 01 CSDL chính, tối đa 30 CSDL dự phòng hoặc Far Sync.



Hình 12: Mối quan hệ giữa các thành phần trong mô hình Broker

Resource là đơn vị nhỏ nhất được quản lý bởi Broker, thành phần thể hiện một hoặc nhiều (đối với mô hình Real Application Clusters - RAC) Instance của CSDL. Database hoặc Host Server là tập hợp nhiều Resources, hay chính là hệ thống CSDL chính hoặc dự phòng mà Instance chạy trên chính nó.

Tại server-side, các thành phần của DGB gồm tiến trình DMON và tệp tin thông tin cấu hình. DMON là tiến trình nền, chạy ở mỗi Database Host khi Broker khởi động và được quản lý bởi Broker. Tệp tin thông tin cấu hình chứa các cài đặt về thuộc tính, trạng thái của Database Host.

Tiến trình DMON thực hiện quản lý và sao chép các tệp tin thông tin cấu hình cho mỗi Database Host mà Broker quản lý. Các tiến trình DMON ở mỗi CSDL khác nhau giao tiếp thông qua môi trường mạng Oracle Net để quản lý việc chuyển đổi vai trò CSDL cũng như cung cấp các chỉ số liên quan tới hiệu năng hệ thống.

1.4. Kết luận chương I

Chương I trình bày tổng quan về Công ty Tài chính tiêu dùng SHBFinance. Sau đó, chương đi vào mô tả bài toán của công ty, tập trung vào thực trạng hiện tại, những thách thức mà doanh nghiệp này đang đối diện, giá trị mà doanh nghiệp sẽ đạt được và hướng giải quyết cho những thách thức đó.

Phần tiếp theo của chương giới thiệu về giải pháp Oracle Data Guard. Gồm một số khái niệm quan trọng như kiến trúc, thành phần của Oracle Data Guard, các loại hình đồng bộ dữ liệu, cơ chế tương tác giữa các thành phần và mô hình giám sát hệ thống Data Guard Broker.

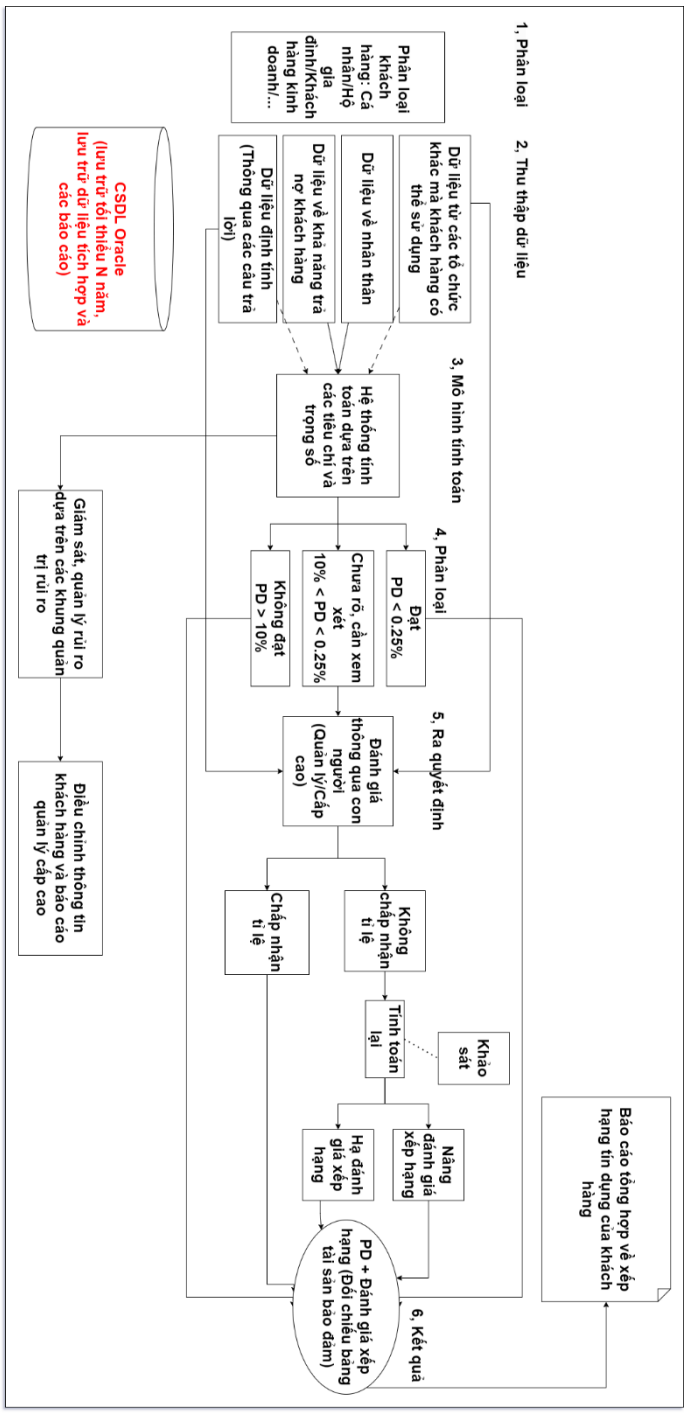
CHƯƠNG 2. TRIỂN KHAI GIẢI PHÁP DATA GUARD CHO CƠ SỞ DỮ LIỆU HỆ THỐNG XẾP HẠNG TÍN DỤNG CỦA SHBFINANCE

2.1. Lên kế hoạch xây dựng giải pháp Data Guard cho SHBFinance

2.1.1. Xác định vấn đề

Theo điều 5, khoản 1, thông tư số 11/2021/TT-NHNN, hệ thống xếp hạng tín dụng nội bộ là hệ thống gồm: "*Các bộ chỉ tiêu tài chính và phi tài chính, các quy trình đánh giá khả năng trả nợ, thanh toán của khách hàng trên cơ sở định tính và định lượng về mặt tài chính, tình hình kinh doanh, quản trị, uy tín của khách hàng (1); Phương pháp đánh giá xếp hạng cho từng nhóm đối tượng khách hàng khác nhau, kể cả các đối tượng bị hạn chế cấp tín dụng và những người có liên quan của đối tượng (2)*" (CỔNG THÔNG TIN ĐIỆN TỬ CHÍNH PHỦ, 2021).

Hệ thống đánh giá tín dụng nội bộ của công ty Tài chính tiêu dùng SHBFinance có quy trình được thể hiện như hình sau:



Hình 13: Tổng quan hệ thống xếp hạng tín dụng nội bộ của SHBFinance

Diễn giải quy trình hoạt động của hệ thống tín dụng nội bộ SHBFinance:

1. Phân loại khách hàng: SHB Finance phân loại khách hàng theo loại khách hàng. Gồm có 3 loại khách hàng chính mà SHB Finance cung cấp dịch vụ: Cá nhân, Hộ gia đình và Khách hàng là người kinh doanh
2. Thu thập dữ liệu: Dữ liệu chính được dùng là dữ liệu về nhân thân, dữ liệu trả nợ của khách hàng, loại dữ liệu này là dữ liệu được dùng để đưa vào mô hình tính toán. Ngoài ra, còn một số loại dữ liệu khác là dữ liệu từ các tổ chức mà khách hàng có thể sử dụng (Internet, viễn thông, bảo hiểm, …) và dữ liệu định

tính (được thể hiện qua một số báo cáo bằng các biểu mẫu đặt câu hỏi dành cho khách hàng như dự định về tương lai, lối sống, …), kiểu loại dữ liệu này thường được dùng làm dữ liệu hỗ trợ cho quản lý ra quyết định khi điểm số chưa chắc chắn. Các dữ liệu này được thu thập và lưu trữ vào CSDL Oracle để thực hiện lưu trữ lâu dài cũng như thiết lập báo cáo.

- Dữ liệu về nhân thân: tuổi, nghề nghiệp, số năm công tác, số người phụ thuộc tài chính, …
- Khả năng trả nợ: dựa trên phương pháp đánh giá uy tín, lịch sử trả nợ, số vòng quay vay nợ, ước lượng khả năng trả nợ dựa trên thông tin thu nhập, …

3. Tính toán: Mô hình thuật toán tính toán điểm tín dụng thông qua các dữ liệu chính và dữ liệu phụ (nếu có), đưa ra tỉ lệ nhất định trong 3 khoảng của xác suất vỡ nợ - Probability of Default (PD). Ví dụ: Trọng số được chia cho các tiêu chí đánh giá có thể kể đến như:

- Lịch sử trả nợ, trọng số 35%
- Dư nợ tại các tổ chức khác, trọng số 30%
- Lịch sử tín dụng (càng dài càng uy tín), trọng số 15%
- Số lần vay nợ mới, trọng số 10%
- Các loại tín dụng sử dụng (loại tín dụng khác nhau sẽ có điểm số khác nhau), trọng số 10%

4. Phân loại: Khi có xác suất, hệ thống sẽ quyết định dựa trên 3 khoảng: Đạt (Rủi ro thấp), Không đạt (Rủi ro cao), Chưa rõ (Rủi ro trung bình). Khoảng Chưa rõ cần sự quyết định của quản lý cấp cao. Lúc này quản lý cấp cao sẽ cần thông tin báo cáo về Dữ liệu của các tổ chức khác cũng như báo cáo về Kết quả biểu mẫu khảo sát định tính của khách hàng để đưa ra quyết định xếp hạng cho khách hàng.
5. Ra quyết định: Nếu cán bộ ra quyết định không đồng ý với tỷ lệ nằm trong khoảng Rủi ro trung bình, sẽ tiếp tục ra quyết định tăng hoặc giảm tỉ lệ dựa trên một biểu mẫu tiếp theo dành cho khách hàng để cán bộ căn cứ và có thể đánh giá tiếp.
6. Kết quả: Xác suất vỡ nợ sẽ được đối chiếu với giá trị của tài sản bảo đảm của khách hàng để đưa ra hạng của khách hàng trong thang điểm xếp hạng tín dụng. Mỗi hệ thống xếp hạng tín dụng nội bộ đều có thang điểm chuẩn khác nhau, dưới đây là hình minh họa về một bảng quy đổi đối chiếu hạng tín dụng nội bộ của khách hàng cá nhân.

Đánh giá xếp loại khách hàng	**AAA**	**AA**	**A**	**BBB**	**BB**	**B**	**CCC**	**CC**	**C**	**D**
Xếp loại rủi ro / **Đánh giá tài sản đảm bảo**	**Rủi ro thấp**			**Rủi ro trung bình**			**Rủi ro cao**			
A (Mạnh)	Xuất sắc			Tốt			Trung bình/Từ chối			
B (Trung bình)	Tốt			Trung bình			Từ chối			
C (Thấp)	Trung bình			Trung bình/Từ chối						

Hình 14: Minh họa bảng quy đổi đối chiếu xếp hạng tín dụng nội bộ

CSDL Oracle được sử dụng để tích hợp thông tin từ nhiều nguồn, chuẩn bị số liệu, dữ liệu cho quy trình đánh giá tín dụng của khách hàng. Dữ liệu được lưu trữ với thời gian lên đến hàng năm.

Có hai trường hợp điển hình (case study) khi ứng dụng giải pháp Oracle Data Guard trong lĩnh vực tài chính, được công bố bởi hãng Oracle bao gồm: Ngân hàng AmTrust và nhà cung cấp hệ thống hỗ trợ vận hành cho trung tâm thanh toán bù trừ - NeuStar. Các yêu cầu chung được đặt ra là chỉ số RPO bằng 0, RTO không lớn hơn 5 – 15 phút và hệ thống tự động sử dụng CSDL dự phòng khi CSDL chính gặp sự cố. Yêu cầu đã nêu được đặt trong bối cảnh khoảng cách của các CSDL lên tới 300 km.

Đối chiếu với CSDL thông tin tín dụng của SHBFinance, là một CSDL quan trọng, tuy nhiên, hệ thống nói trên có một số vấn đề nổi bật như sau:

- Hiện tại, CSDL chính đã sử dụng các bản sao lưu khôi phục ở cả phạm vi nội bộ của trung tâm dữ liệu (local) và ở các nơi lưu trữ khác (điện toán đám mây, băng từ). Tuy nhiên, chưa đáp ứng được việc thời gian khôi phục khi chỉ số RTO lớn, dẫn đến việc khi gặp sự cố, hệ thống cần mất nhiều thời gian hơn gây gián đoạn tới quá trình hoạt động của doanh nghiệp.
- Đối với vấn đề tính toàn vẹn của dữ liệu được thể hiện qua chỉ số RPO, khi hệ thống chính gặp sự cố, việc khôi phục lại dữ liệu có được đầy đủ hay không lại phụ thuộc vào bản sao lưu cuối cùng là bao lâu. Nếu bản sao lưu cuối cùng càng lâu, thì lượng dữ liệu mất càng lớn.
- Khi thực hiện hoạt động truy xuất thông tin trên hệ thống chính với khối lượng lớn sẽ làm giảm tải hiệu năng xử lý thông tin của hệ thống, cần tính toán đến khả năng mở rộng hệ thống để có thể truy xuất thông tin đồng thời

Vì vậy, việc bổ sung, nâng cấp thêm CSDL dự phòng với giải pháp Oracle Data Guard là cần thiết đối với hệ thống xếp hạng tín dụng nội bộ của SHBFinance trong việc đảm bảo quy trình nghiệp vụ được thực hiện không bị gián đoạn bởi sự cố tại trung tâm CSDL chính. Khi đó, CSDL dự phòng do được đồng bộ với CSDL chính, có thể sử dụng CSDL dự phòng để thực hiện các tác vụ như sao lưu dữ liệu, truy vấn dữ

liệu (khi sử dụng Active Data Guard) hoặc để sử dụng làm CSDL thay thế trong trường hợp CSDL chính gặp sự cố hoặc cần bảo trì, nâng cấp.

2.1.2. Xác định cấp độ chuyển đổi dự phòng

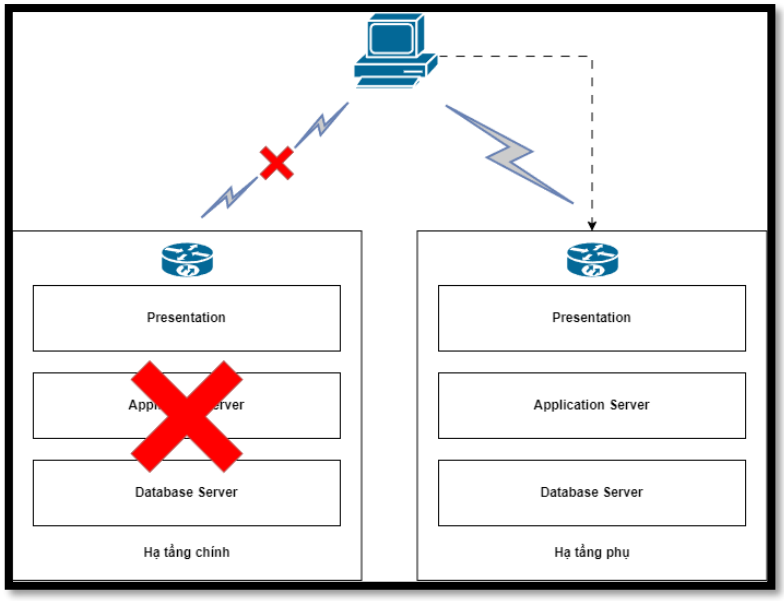
Khi có sự cố, sẽ có trường hợp sau xảy ra: toàn bộ hạ tầng hệ thống chính bị hỏng (gồm cả CSDL), hoặc chỉ máy chủ CSDL bị hỏng. Tùy thuộc vào nhu cầu, chi phí mà doanh nghiệp sẵn sàng bỏ ra cũng như tầm quan trọng của hệ thống ở mức độ nào sẽ có các cách triển khai khác nhau.

Trong hệ thống xếp hạng tín dụng nội bộ của doanh nghiệp được triển khai theo kiến trúc 3 tầng (3-Tiers), gồm các tầng như CSDL, tầng xử lý ứng dụng (Application) và tầng giao diện (Client/Presentation). Tầng giao diện sẽ tương tác trực tiếp với người dùng, các yêu cầu từ lớp này sẽ được gửi tới tầng ứng dụng để xử lý nghiệp vụ, cũng như cung cấp các phương thức bảo mật khi giao tiếp với tầng CSDL. Sau đó, dữ liệu từ tầng CSDL sẽ được trả về lớp giao diện thông qua tầng ứng dụng nếu có.

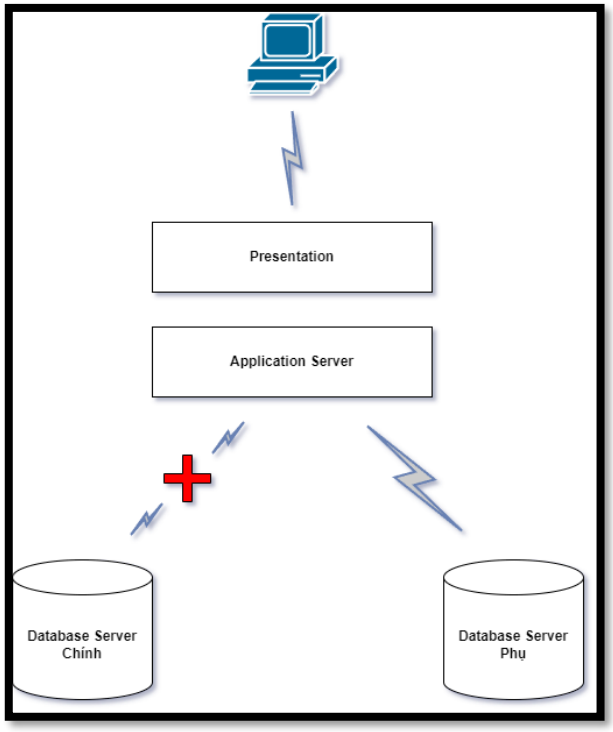
Có cấp độ chuyển đổi dự phòng như sau:

Loại sự cố	Mô tả	Phương thức chuyển đổi
Toàn bộ hạ tầng hệ thống chính gặp sự cố	Hạ tầng hệ thống chính gặp sự cố, gây hỏng/mất kết nối đối với máy chủ chứa lớp ứng dụng với hai lớp giao diện và CSDL. Tất cả máy chủ chứa 3 lớp đều bị ảnh hưởng.	Chuyển đổi toàn bộ hạ tầng hệ thống chính sang hạ tầng hệ thống dự phòng
Chỉ hệ thống máy chủ chứa CSDL chính gặp sự cố	Có thể điều hướng máy chủ ứng dụng kết nối tới CSDL khác để tiếp tục vận hành. Chấp nhận việc người dùng cần mất một khoảng thời gian để kết nối lại, và có thể độ trễ khi truy cập ứng dụng cao hơn.	Phù hợp với mô hình đặt hệ thống máy chủ phân tán, khi hệ thống máy chủ ứng dụng và hệ thống máy chủ giao diện không chung một địa điểm với hệ thống máy chủ CSDL.

Bảng 5: Các trường hợp chuyển đổi



Hình 15: Phương thức chuyển đổi toàn bộ



Hình 16: Phương thức chuyển đổi chỉ hệ thống CSDL

2.1.3. Xác định về đường truyền và đồng bộ dữ liệu

Mô hình Data Guard sử dụng cơ chế đồng bộ redo data thông qua môi trường mạng, từ CSDL chính tới CSDL dự phòng. Vì vậy, cần đảm bảo đường truyền hệ thống của doanh nghiệp SHBFinance có đủ khả năng để truyền dữ liệu đồng bộ, tránh gây ra độ trễ dữ liệu quá lớn.

Độ trễ dữ liệu đồng bộ xảy ra khi CSDL chính không thể truyền hoặc hạn chế trong việc truyền redo data tới CSDL dự phòng. Khi này, các redo data tại CSDL chính được tạo ra liên tục, nhưng CSDL dự phòng có thể không nhận được và có nguy cơ gây ra tình trạng mất hoàn toàn dữ liệu khi hệ thống CSDL chính gặp sự cố mà không thể khôi phục.

Để đánh giá mô hình mạng của hệ thống Data Guard, có hai tiêu chí cần xem xét là độ tin cậy và băng thông mạng. Có một số tiêu chí như:

- Tường lửa và bảo mật đường truyền: sử dụng cơ chế mã hóa, tường lửa có thể làm chậm lưu lượng truyền/nhận dữ liệu thay đổi. Cần phải cân bằng việc bảo mật dữ liệu cũng như việc hạn chế mất mát dữ liệu.
- Sử dụng cơ chế nén thông tin thay đổi nhằm giảm dung lượng lưu trữ làm chậm việc truyền/nhận. CSDL chính sẽ phải thực hiện nén các tệp thông tin thay đổi trước khi gửi và CSDL dự phòng sẽ phải thực hiện giải nén trước khi áp dụng.
- Bảo mật dữ liệu với tính năng mã hóa dữ liệu lưu trữ (Transparent Data Encryption) cũng sẽ làm ảnh hưởng tới tốc độ truyền/nhận trong môi trường Oracle Net. Do dữ liệu truyền được mã hóa và chỉ được giải mã khi đến đích.
- Tối ưu hóa các thông số liên quan tới hệ thống mạng như chỉ số Maximum Transmission Unit (MTU) – kích thước tối đa của một gói tin dữ liệu trong giao thức mạng TCP/IP (hoặc giao thức khác), đo lường số byte tối đa mà một gói tin có thể chứa trước khi gửi qua mạng, giảm thiểu tình trạng phân mảnh gói tin.

2.1.4. Xác định chế độ bảo vệ trong Data Guard

Trong 03 chế độ bảo vệ của Data Guard, cần lựa chọn chế độ để phù hợp với yêu cầu vận hành của doanh nghiệp. Các chế độ có độ ưu tiên khác nhau về hiệu năng, tính sẵn sàng của hệ thống và mức độ mất mát dữ liệu.

Với chế độ Ưu tiên bảo vệ (max. protection):

Chế độ này thực hiện cơ chế chỉ xác nhận một giao dịch đã được COMMIT khi và chỉ khi ít nhất một CSDL dự phòng trả lại tín hiệu ACK cho CSDL chính rằng dữ liệu thay đổi đã được ghi vào CSDL dự phòng. Ngược lại, nếu không có bất kỳ tín hiệu nào trở về, CSDL chính sẽ treo và dừng hoạt động để đảm bảo tính toàn vẹn khi giao dịch chưa được COMMIT ở cả hai CSDL.

Để dự phòng trong trường hợp CSDL dự phòng không hoạt động, doanh nghiệp nên thực hiện triển khai tối thiểu hai CSDL dự phòng trong chế độ Ưu tiên bảo vệ. Với giải pháp này sẽ hạn chế việc CSDL chính rơi vào trạng thái chờ đợi, dẫn đến tự động dừng hoạt động khi ít nhất một trong hai CSDL dự phòng trả lại tín hiệu ACK cho CSDL chính. Chế độ này phù hợp với nhu cầu ưu tiên về tính toàn vẹn dữ liệu hơn là tính sẵn sàng của CSDL.

Với chế độ Ưu tiên tính sẵn sàng (max. availability):

Trong chế độ này, CSDL chính sẽ chờ đến thời gian tối đa được cấu hình trong biến NET_TIMEOUT khi chờ tín hiệu ACK phản hồi lại từ CSDL dự phòng, nếu nhận được tín hiệu, CSDL chính có thể đánh dấu COMMIT và tiếp tục một giao dịch mới. Ngược lại, CSDL chính sẽ hoạt động như chế độ Ưu tiên hiệu năng và liên tục cập nhật trạng thái của CSDL dự phòng.

Để dự phòng trường hợp mất kết nối, doanh nghiệp nên kết hợp thêm tính năng Far Sync với một CSDL trung gian, đứng giữa CSDL chính và CSDL dự phòng để trung chuyển các dữ liệu thay đổi. CSDL Far Sync sẽ có đường truyền tốt hơn để đảm bảo không mất kết nối với CSDL chính, gây ra tình trạng mất dữ liệu. Ngoài ra, cũng luôn phải theo dõi các tiến trình, đường truyền để xử lý các sự cố gây ra độ trễ trong việc đồng bộ. Chế độ này phù hợp với nhu cầu muốn cân bằng về tính toàn vẹn của dữ liệu cũng như tính sẵn sàng của CSDL chính.

Với chế độ Ưu tiên hiệu năng (max. performance):

Chế độ này COMMIT ngay khi có tín hiệu của người dùng, ghi redo data vào Online Redo Log. Hệ thống Data Guard sẽ truyền dữ liệu thay đổi song song tới: Standby Redo Log của CSDL dự phòng, trực tiếp từ Online Redo Logs (đối với đường truyền tốt) (1), tới Archive Redo Log của CSDL chính (2) theo cơ chế không đồng bộ với giao dịch được COMMIT, hạn chế việc mất mát dữ liệu khi CSDL chính xảy ra sự cố.

Chế độ này phù hợp khi không đặt nặng vấn đề về mất đồng bộ dữ liệu và yêu cầu hiệu năng hệ thống chính cần hoạt động với hiệu năng cao.

2.1.5. Xác định yêu cầu phần cứng, phần mềm

Phần cứng:

- Phần cứng của CSDL chính và CSDL dự phòng có thể khác nhau về số lượng bộ xử lý trung tâm (CPU), kích thước bộ nhớ (Memory), và cấu hình lưu trữ (Storage)
- Cho phép hệ thống xử lý cũng như phiên bản cài đặt Oracle Database Software có kích thước đơn vị biểu diễn thông tin khác nhau (32-bit hoặc 64-bit)

Trong trường hợp CSDL chính và CSDL dự phòng cùng một máy chủ (local), cần đảm bảo rằng hệ thống được cấu hình đúng cách để hai CSDL có cùng DB_NAME có thể hoạt động ổn định, không gây ra xung đột.

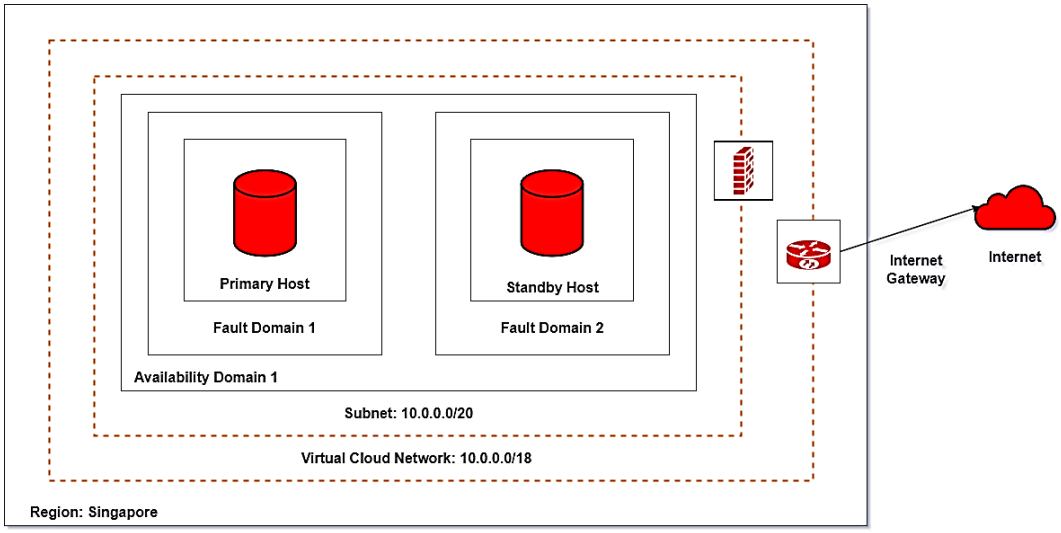
Phần mềm:

- Yêu cầu phiên bản cài đặt cho Oracle Database từ Enterprise Edition trở lên cho cả hệ thống CSDL chính và CSDL dự phòng. Data Guard không hỗ trợ cho Oracle Database Standard Edition.

- Nếu sử dụng công cụ quản lý bộ nhớ Automatic Storage Management (ASM) hoặc Oracle Managed Files (OMF) thì cần sử dụng giống nhau ở cả hai hệ thống CSDL chính và CSDL dự phòng nếu sử dụng CSDL dạng vật lý. Đối với trường hợp kết hợp các phương thức thì cũng tương tự ở cả hai hệ thống.

2.2. Thực nghiệm triển khai giải pháp Data Guard dựa trên RMAN Duplicate và nền tảng điện toán đám mây Oracle Cloud Infrastructure

2.2.1. Kiến trúc tổng quan



Hình 17: Kiến trúc tổng quan thực nghiệm giải pháp Data Guard trên OCI

Region: Là một khu vực địa lý, nơi đặt hạ tầng công nghệ thông tin trải dài trên lãnh thổ địa lý đó, cung cấp nền tảng mạng và tài nguyên cho các ứng dụng, dịch vụ trên điện toán đám mây mà người dùng có thể sử dụng. Mỗi khu vực chứa các hạ tầng công nghệ thông tin này hoàn toàn độc lập về mặt giao tiếp mạng cũng như về vị trí địa lý với các khu vực khác. Thông thường, các ứng dụng sẽ triển khai tại khu vực có lưu lượng sử dụng cao để tăng tốc độ truy cập cho người dùng cuối. Tại các khu vực khác nhau, các dịch vụ có thể triển khai cũng có thể khác nhau do phụ thuộc nhu cầu và hạ tầng công nghệ thông tin, có thể kể đến một số dịch vụ như: Máy ảo, Lưu trữ, Mạng truyền thông, CSDL/Kho dữ liệu, Máy chủ phân giải tên miền (DNS), Bảo mật.

Availability Domain (AD): Là một hoặc nhiều tập hợp hạ tầng công nghệ thông tin trong cùng một khu vực. Mỗi khu vực thường có nhiều nhất 03 AD, mỗi AD được cách ly với nhau về mặt hạ tầng như hệ thống điện, hệ thống làm mát hoặc mạng nội bộ, nên có khả năng chịu lỗi và khó có thể gây ảnh hưởng tới các AD khác. Tất cả các AD trong cùng một khu vực được kết nối với nhau bằng đường truyền có băng thông cao, phù hợp cho việc xây dựng giải pháp dự phòng và sẵn sàng cao.

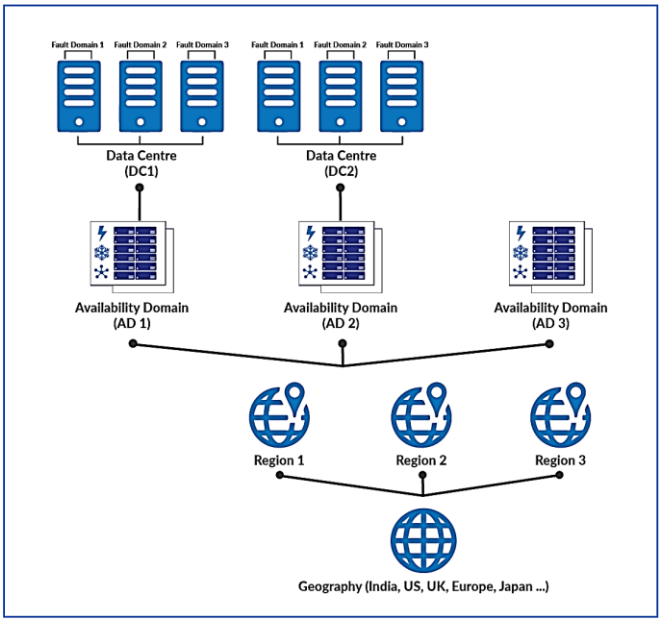
Fault Domain (FD): Là một tập hợp gồm phần cứng hạ tầng trong trong một AD. Mỗi AD có 03 FD. Được thiết kế để bảo vệ khỏi sự cố của các thành phần vật lý

trong cùng một AD khi ứng dụng được triển khai. Nếu một ứng dụng được triển khai gặp lỗi hoặc phần cứng ứng dụng đó cần bảo trì, thì ứng dụng khác được triển khai tại một FD khác sẽ không bị ảnh hưởng về mặt vật lý. Các thành phần vật lý của mỗi FD được tách biệt và đều có nguồn điện dự phòng, phòng chống việc mất điện ảnh hưởng tới các FD khác.

Virtual Cloud Network (VCN) và Subnet: là dịch vụ cung cấp loại mạng ảo, nội bộ giống với mạng truyền thống, có cung cấp tường lửa, các loại hình gateway, được thiết lập ảo hóa dựa trên hạ tầng công nghệ thông tin. Một VCN nằm trong một Region và có thể bao quát hết các đối tượng, dịch vụ/ứng dụng dựa trên tham số CIDR (Classless Inter-Domain Routing) – cấu hình tập hợp các địa chỉ IP có chung tiền tố mạng và số lượng máy đã thiết lập.

Mỗi mạng con – subnet gồm các địa chỉ IP giới hạn bởi số bit mạng và số bit host. Địa chỉ của các subnet không được trùng nhau. Các dịch vụ, ứng dụng trong cùng một VCN sử dụng chung một bảng định tuyến, quy định về bảo mật, tường lửa và máy chủ cấp phát IP động. Subnet có thể đặt ở chế độ nội bộ (private) hoặc công khai (public).

Với mạng công khai, các dịch vụ sẽ được gán một IPv4 công khai bên cạnh IPv4 nội bộ, có thể giao tiếp với Internet. Ngoài ra, Subnet có hai cấp độ về phạm vi. Với subnet cấp Region, dịch vụ thuộc các AD khác nhau có thể giao tiếp nội bộ; với AD thì máy chủ chỉ có thể giao tiếp nội bộ khi trong cùng một AD.



Hình 18: Minh họa phân cấp giữa Region, AD và FD

Trong thực nghiệm, Region khu vực Singapore chỉ có một AD, nhưng cung cấp tới 3 FD tại AD này. Vì vậy, để tận dụng khả năng chịu lỗi trong cùng một trung tâm dữ liệu, các máy chủ CSDL sẽ được đặt tại các FD khác nhau, hạn chế sự cố trong cùng một địa điểm. Cấu hình về mạng lưới và các máy ảo được cài đặt như sau:

Thành phần	Thông tin
VCN	IPv4 CIDR Blocks: 10.0.0.0/8 DNS Hostname: Yes
Route Table	Destination: 0.0.0.0/0 Gateway: Internet Gateway
Security List	*Ingress Rule:* - Source: 0.0.0.0/0, IP Protocol: TCP, Destination Port Range: 22 (SSH) - Source: 0.0.0.0/0, IP Protocol: ICMP (ping) - Source: 0.0.0.0/0, IP Protocol: TCP, Destination Port Range: 1521 (Oracle Database)
Subnet	IPv4 CIDR Blocks: 10.0.0.0/20 Security Lists: Default Subnet Access: Public Subnet Subnet Type: Regional DNS Hostname: Yes

Bảng 6: Thiết lập cấu hình mạng lưới ảo OCI

Máy chủ	Thông tin	Giá trị
Chính	Oracle CPU	1
	Memory (Gigabyte)	6
	Availability Domain	AD-1
	Fault Domain	FD-1
	Subnet	10.0.0.0/20
	Private IP	10.0.12.202
	Public IP	213.35.102.135
	Region	Singapore
	Hostname	source
	Operating System	Oracle Linux 7.9
	Storage (Gigabyte)	50
Dự phòng/Phụ	Oracle CPU	1
	Memory (Gigabyte)	6
	Availability Domain	AD-1
	Fault Domain	FD-2
	Subnet	10.0.0.0/20
	Private IP	10.0.15.63
	Public IP	129.150.61.43
	Region	Singapore

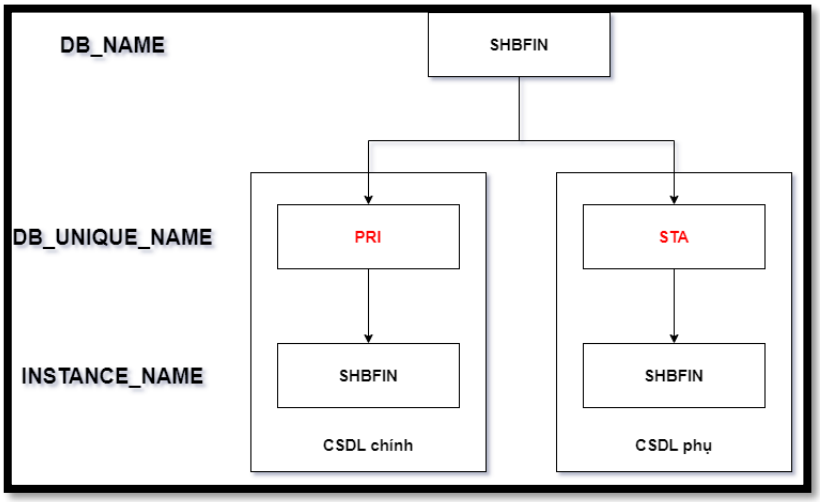
	Hostname	target
	Operating System	Oracle Linux 7.9
	Storage (Gigabyte)	50

Bảng 7: Cấu hình máy chủ ảo cài đặt Oracle Data Guard trên OCI

Để triển khai giải pháp Oracle Data Guard, cả hai máy chủ cần được cài đặt Oracle Database Software phiên bản Enterprise Edition. Ngoài việc cài đặt Software, máy chủ chính cần cài Database. Trong phần thực nghiệm, máy chủ chính đã được cài đặt Database theo kiến trúc Single Instance Database nhằm tối giản việc thực nghiệm, thay vì kiến trúc Real Application Cluster (RAC). Cả hai máy chủ được cài đặt và cấu hình các đường dẫn giống nhau. Các phần tiếp theo sẽ đi sâu về cách triển khai và thiết lập Oracle Data Guard. Lệnh và các cấu hình đầy đủ sẽ được đính kèm theo phụ lục tại cuối bài.

2.2.2. Môi trường Oracle Net và định danh CSDL

Trong môi trường Data Guard, để phân biệt về loại CSDL (dự phòng hoặc chính), cần sử dụng DB_UNIQUE_NAME để hệ thống nhận diện các CSDL này, thay vì sử dụng DB_NAME. Data Guard sẽ sử dụng DB_UNIQUE_NAME để giám sát các CSDL với Broker và thực hiện các phương thức chuyển đổi vai trò.



Hình 19: Minh họa phân cấp tên trong hệ thống CSDL cho SHBFinance

Kiến trúc Data Guard phụ thuộc chủ yếu vào cách cấu hình môi trường Oracle Net làm sao để cả hai hệ thống CSDL chính và phụ đều có thể giao tiếp với nhau. CSDL Oracle sử dụng dịch vụ Listener như một "gateway" trong thiết bị mạng, điều hướng kết nối từ các thiết bị liên lạc với nó tới CSDL để thiết lập phiên làm việc (sessions) dành cho người dùng.

Các CSDL trong môi trường Data Guard cũng sử dụng Listener để có thể kết nối với nhau và truyền tải thông tin, đồng bộ thay đổi dữ liệu, có thể hiểu khi CSDL

chính cần truyền tài thông tin tới CSDL dự phòng thì CSDL chính là đối tượng chủ động, cần tìm Listener để kết nối và chuyển tới.

Thông thường, khi một client kết nối tới máy chủ CSDL, sẽ cần mô tả chuỗi kết nối, gồm: địa chỉ IP hoặc tên host (nếu đã khai báo IP trong tệp hosts), cổng port mở và tên nhận dạng của Instance kết nối tới. Ví dụ như:

```
CONNECT username/password@123.456.789.000:1521/shbfin
```

Các CSDL cũng như cấu hình liên quan trong Data Guard sử dụng một phương thức đơn giản hóa chuỗi kết nối trên bằng việc chứa thông tin chuỗi vào một tên bí danh (alias), tính năng này được gọi là Local Naming Method. Thông tin chuỗi kết nối này sẽ được cấu hình đưa vào bí danh trên mỗi máy khách, mỗi khi kết nối, chỉ cần sử dụng bí danh này là hệ thống có thể biên dịch sang chuỗi truyền thống:

```
CONNECT username/password@pri
# pri chứa thông tin IP, port, instance cần kết nối
```

Bước đầu trong quá trình triển khai kiến trúc Data Guard là cấu hình Listener cho mỗi hệ thống CSDL. Cùng với đó là cấu hình phương thức Local Naming Method để thực hiện việc đơn giản hóa, sử dụng bí danh để các CSDL giao tiếp với nhau. Tại cả hai máy chủ CSDL, thực hiện việc cấu hình Listener có dạng như sau:

```
<listener_name>=
 (DESCRIPTION_LIST =
  (DESCRIPTION =
   (ADDRESS = (PROTOCOL = TCP)(HOST = source)(PORT = 1521))
   (ADDRESS = (PROTOCOL = IPC)(KEY = EXTPROC1521))
  ))
SID_LIST_<listener_name>=
 (SID_LIST =
  (SID_DESC = (GLOBAL_DBNAME = shbfin) (ORACLE_HOME =
/u01/app/oracle/product/19.0.0/dbhome_1) (SID_NAME = shbfin))
 )
```

Trong đó:

<listener_name>: tên của Listener

DESCRIPTION_LIST: chứa danh sách mô tả các kết nối đến mà Listener sẽ xử lý. Tại đây quan tâm đến giao thức TCP dành cho kết nối từ các ứng dụng và giữa các Database với nhau. Với giao thức ICP (Inter-Process Communication), dành cho các ứng dụng cùng trên máy chủ chứa CSDL có thể kết nối nội bộ với nhau.

SID_LIST_<listener_name>: chứa danh sách mô tả các CSDL mà Listener sẽ điều hướng kết nối của người dùng tới CSDL đó.

Cấu hình Local Naming Method cho hai máy chủ chứa CSDL có dạng như sau:

```
<alias_primary_name> =
 (DESCRIPTION = (ADDRESS = (PROTOCOL = TCP)(HOST = db1)(PORT = 1521))
   (CONNECT_DATA = (SERVER = DEDICATED)  (SID = shbfin)))
sta =
 (DESCRIPTION = (ADDRESS = (PROTOCOL = TCP)(HOST = db2)(PORT = 1521))
   (CONNECT_DATA = (SERVER = DEDICATED)  (SID = shbfin)))
```

Trong đó:

<alias_primary_name>: tên bí danh được sử dụng để mô tả kết nối

DESCRIPTION: mô tả kết nối, gồm địa chỉ và dữ liệu kết nối

CONNECT_DATA: chỉ định kết nối sẽ theo phương thức nào và kết nối vào CSDL nào. Trong trường hợp này là phương thức DEDICATED – mỗi client kết nối vào sẽ có một tiến trình nền hỗ trợ riêng với bộ nhớ dành cho việc lưu trữ kết quả của SQL riêng và kết nối vào CSDL với Instance là shbfin.

2.2.3. Cấu hình tham số chung cho hệ thống chính

Bật chế độ FORCE LOGGING: Trong chế độ FORCE LOGGING, mọi thay đổi tại CSDL, cụ thể là trong Redo Buffer Cache đều được ghi xuống thiết bị đĩa cứng tại Online Redo Logs theo cơ chế xoay vòng (ghi đè khi hết) bằng tiến trình LGWr, bởi vậy mà CSDL có thể khôi phục được những thay đổi đã COMMIT sau khi xảy ra sự cố. FORCE LOGGING đảm bảo tính nhất quán của dữ liệu.

Bật chế độ ARCHIVELOG: Khi một Online Redo Logs đầy về mặt dung lượng, chế độ ARCHIVELOG sẽ thực hiện lưu trữ tệp tin này thông qua việc sao chép bằng tiến trình ARCn với điều kiện Online Redo Logs thực hiện cơ chế Log Switch để thực hiện chuyển qua tệp tin khác lưu trữ. Trạng thái của Online Redo Logs (ORLs) thời điểm này sẽ chuyển từ CURRENT qua ACTIVE cho đến khi được lưu trữ thành công, trở về trạng thái INACTIVE.

Tại trạng thái ACTIVE, Online Redo Logs chưa được lưu trữ bằng tiến trình ARCn, tiến trình checkpoint chưa xảy ra và các thông tin thay đổi trong Active Logs đó sẽ được sử dụng cho quá trình Crash Recovery – khôi phục hệ thống sau khi lỗi dừng hoạt động hệ thống đột ngột.

Bật chế độ FLASHBACK: công nghệ Flashback là một phần không thể thiếu trong việc thiết lập chế độ chuyển vai trò tự động – Fast-start Failover trong trường hợp CSDL chính gặp sự cố. Khi thực hiện chuyển đổi failover, CSDL chính (cũ) sẽ rơi

vào trạng thái mất đồng bộ với CSDL chính (là CSDL dự phòng trước đây), trạng thái hiển thị sẽ là “needs Re-instatement”.

Chế độ Flashback giúp CSDL trở về một thời điểm trong quá khứ nhanh chóng, nhờ các dữ liệu được ghi trong flashback logs và được lưu trữ trong phân vùng Fast Recovery Area (FRA). Thông qua cấu hình tham số DB_FLASHBACK_RETENTION _TARGET (phút), CSDL có thể trở về thời điểm trước khi xảy ra lỗi trong khoảng thời gian đã chỉ định với tham số. Flashback loại bỏ việc dựng lại CSDL chính khi thực hiện Failover - gây lỗi và mất đồng bộ.

```
# Truy vấn thông tin cấu hình CSDL
SELECT NAME, DB_UNIQUE_NAME, OPEN_MODE, LOG_MODE,
FLASHBACK_ON, FORCE_LOGGING FROM V$DATABASE;
# Thực hiện bật Force Logging, ArchiveLog và Flashback
SHUTDOWN IMMEDIATE;
STARTUP MOUNT;
SHOW PARAMETER NAME;
ALTER DATABASE ARCHIVELOG;
ALTER DATABASE FORCE LOGGING;
ALTER SYSTEM SET DB_FLASHBACK_RETENTION_TARGET = 60
SCOPE=BOTH;
ALTER DATABASE FLASHBACK ON;
```

Tạo Standby Redo Logs: Standby Redo Logs (SRLs) được dùng khi vai trò của CSDL là dự phòng/phụ, nhận thông tin đồng bộ dữ liệu thay đổi từ CSDL chính. Cần tạo SRLs ở cả hai CSDL chính và phụ, dự phòng trong việc chuyển đổi để chúng có thể nhận và áp dụng thay đổi vào CSDL. Có một số điều kiện bắt buộc khi tạo SRLs như sau: cần tạo nhiều hơn ít nhất 01 groups so với groups của ORLs tại CSDL chính (1), SRLs cần lớn hơn hoặc bằng ORLs của CSDL chính (2). Nếu SRLs được cấu hình sai với các điều kiện đã nêu, tiến trình RFS sẽ ghi vào Archive Redo Log (ARL), mất đi tính năng Real-Time Apply và gây ra hiện tượng trễ.

```
# Kiểm tra dung lượng ORLs theo Megabyte (200MB mỗi ORLs)
select GROUP#,THREAD#,SEQUENCE#,bytes/1024/1024, MEMBERS,STATUS
from v$log;
# Tạo Standby Redo Logs phù hợp
ALTER DATABASE ADD STANDBY LOGFILE THREAD 1 GROUP 4
('/u02/oradata/shbfin/stb_redo04.log') SIZE 200M;
ALTER DATABASE ADD STANDBY LOGFILE THREAD 1 GROUP 5
('/u02/oradata/shbfin/stb_redo05.log') SIZE 200M;
```

```
ALTER DATABASE ADD STANDBY LOGFILE THREAD 1 GROUP 6
('/u02/oradata/shbfin/stb_redo06.log') SIZE 200M;
ALTER DATABASE ADD STANDBY LOGFILE THREAD 1 GROUP 7
('/u02/oradata/shbfin/stb_redo07.log') SIZE 200M;
# Kiểm tra lại các loại Logs hiện tại
SELECT TYPE, MEMBER FROM V$LOGFILE ORDER BY GROUP#;
```

Thiết lập vị trí lưu trữ Redo Log cục bộ: Trong môi trường Data Guard, Redo Transport Services được cài đặt, điều khiển bằng tham số LOG_ARCHIVE_DEST_n. Tham số này cho phép redo data vừa được gửi đồng bộ sang Standby Database, vừa được lưu trữ xuống đĩa. Cụ thể, dạng tổng quát thường dùng của tham số này như sau:

```
LOG_ARCHIVE_DEST_N: [1 | 2 | 3 | … | 31] =
‘LOCATION = path_name | SERVICE = service_name
SYNC | ASYNC
AFFIRM | NOAFFIRM
VALID_FOR = (redo_log_type, database_role)
DB_UNIQUE_NAME = db_unique_name
…’
```

LOG_ARCHIVE_DEST_N là thông tin xác định Redo Transport Services sẽ chuyển redo data xuống cục bộ hay đi sang Standby Redo Logs. Trong trường hợp cấu hình cục bộ, n luôn phải đặt là 1, LOCATION sẽ được đặt giá trị là một đường dẫn của máy chủ cài đặt CSDL chính, SERVICE_NAME dùng khi gửi sang Standby.

Cách truyền phụ thuộc vào kiểu truyền SYNC/ASYNC và AFFIRM /NOAFFIRM, mặc định khi thiết lập AFFIRM thì sẽ thiết lập SYNC. Với tham số VALID_FOR gồm hai đối số đầu vào, khi CSDL có vai trò là *database_role* thì sẽ lưu trữ *redo_log_type* xuống hoặc gửi redo_log_type đi cho CSDL dự phòng. Trong cài đặt cục bộ, thông tin cài đặt sẽ như sau:

```
ALTER SYSTEM SET LOG_ARCHIVE_DEST_1=
'LOCATION=/u02/oradata/shbfin/arch1/
VALID_FOR=(ALL_LOGFILES,ALL_ROLES)
DB_UNIQUE_NAME=pri' scope=spfile;
```

Với cài đặt này, dù CSDL ở vai trò chính hoặc vai trò phụ, các redo data được lưu trữ trong ORLs hoặc SRLs đều được sao chép và cất giữ theo đường dẫn đã cấu hình tại LOCATION.Bảng kết hợp cho tham số VALID_FOR như sau, X là hợp lệ:

Kết hợp	CSDL Chính	CSDL dự phòng (Physical)	CSDL dự phòng (Logical)
ONLINE_LOGFILE, PRIMARY_ROLE	X		
ONLINE_LOGFILE, STANDBY_ROLE			X
ONLINE_LOGFILE, ALL_ROLES	X		X
STANDBY_LOGFILE, STANDBY_ROLES		X	X
STANDBY_LOGFILE, ALL_ROLES		X	X
ALL_LOGFILES, PRIMARY_ROLE	X		
ALL_LOGFILES, STANDBY_ROLE		X	X
ALL_LOGFILES, ALL_ROLES	X	X	X

Bảng 8: Kết hợp đối số trong VALID_FOR

Ngoài ra, còn một số tham số phụ trợ khác cho CSDL chính được cấu hình như sau:

```
# Cấu hình Listener cho Instance
ALTER SYSTEM SET LOCAL_LISTENER='(ADDRESS = (PROTOCOL = TCP)(HOST = source)(PORT = 1521))' SCOPE=SPFILE;
# Cấu hình số lượng tiến trình ARCn
ALTER SYSTEM SET LOG_ARCHIVE_MAX_PROCESSES=30 SCOPE=SPFILE;
# Cấu hình định dạng tên cho Archive Redo Logs
ALTER SYSTEM SET LOG_ARCHIVE_FORMAT='ora_%t_%s_%r.arc' SCOPE=SPFILE;
# Cấu hình dung lượng của Fast Recovery Area
ALTER SYSTEM SET DB_RECOVERY_FILE_DEST_SIZE = 5G SCOPE=SPFILE;
# Cấu hình đường dẫn của Fast Recovery Area
```

```
ALTER SYSTEM SET DB_RECOVERY_FILE_DEST = '/u02/oradata/shbfin/fra/' SCOPE=SPFILE;
# Cấu hình file mật khẩu chỉ sử dụng trong phạm vi máy chủ cài đặt Database
ALTER SYSTEM SET REMOTE_LOGIN_PASSWORDFILE=EXCLUSIVE SCOPE=SPFILE;
# Cấu hình quản lý các tệp tin đồng thời của các CSDL
ALTER SYSTEM SET STANDBY_FILE_MANAGEMENT=AUTO SCOPE=SPFILE;
```

Trong cấu hình phụ trợ này, có thêm cấu hình về Fast Recovery Area (FRA). FRA là một tính năng của Oracle Database giúp quản lý và tự động hóa việc lưu trữ các thành phần quan trọng của cơ sở dữ liệu liên quan đến phục hồi và sao lưu. FRA được sử dụng trong hệ thống Data Guard để nhận những bản sao lưu từ RMAN gồm SRLs, ORLs, Archive Log và Flashback Logs cho công nghệ Flashback.

2.2.4. Tạo hệ thống dự phòng dựa trên RMAN DUPLICATE

Sao chép và gửi file mật khẩu, tham số từ CSDL chính sang CSDL dự phòng: Mọi CSDL trong kiến trúc Data Guard đều cần sử dụng một tệp lưu trữ mật khẩu có chung một mật khẩu giống nhau cho người dùng quản trị SYS. Việc sao chép sang cả tệp tin mật khẩu, tham số đều là việc đảm bảo tính nhất quán dữ liệu trong các tệp.

Khi thực hiện sao chép, hai hệ thống máy chủ đều cần thực hiện sở hữu khóa công khai của bên còn lại để xác thực qua phương thức SSH. Sao chép được dùng bằng công cụ Secure Copy (SCP), sử dụng SSH để mã hóa thông tin truyền đi, khi tệp tin nhận đến sẽ được giải mã bằng khóa bí mật mà hệ thống sở hữu. Cụ thể, việc sao chép từ máy chủ chính sang máy chủ phụ được thực hiện bằng cú pháp sau:

```
scp [other options] [source username@IP]:/[full file name] [destination username@IP]:/[directory]
# source username@IP: thông tin định danh máy chủ chính
# full file name: tệp tin cần gửi
# destination username@IP: thông tin định danh máy chủ dự phòng
# directory: thư mục nhận tệp tin được gửi
# Nếu sử dụng máy chủ chính, không cần đăng nhập từ máy chủ chính
```

Trong hệ thống hiện tại, tệp tin mật khẩu có tên là *orapwshbfin* và tệp tin tham số dạng văn bản thô đã được kết xuất có tên là *initshbfin.ora*, đều nằm ở thư mục *$ORACLE_HOME/dbs*. Dưới đây là câu lệnh để sao chép các tệp tin trên từ máy chủ chính – source sang máy chủ dự phòng – target:

```
scp initshbfin.ora oracle@10.0.15.63:$ORACLE_HOME/dbs
scp orapwshbfin oracle@10.0.15.63:$ORACLE_HOME/dbs
```

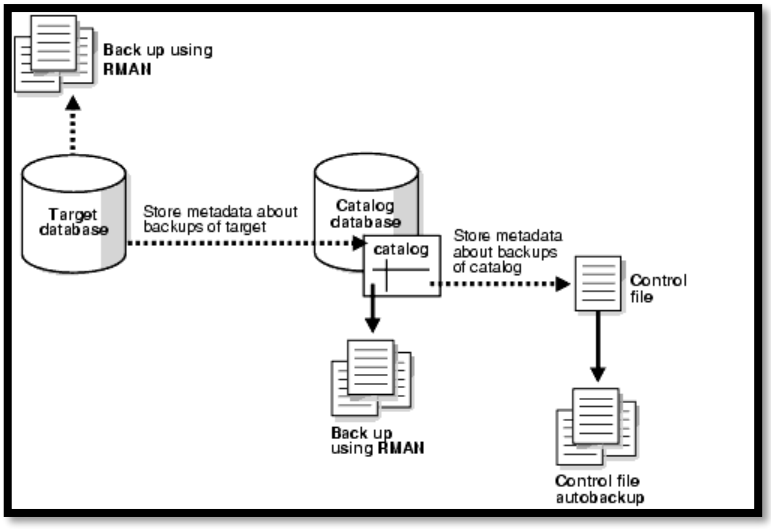
Dựng CSDL dự phòng bằng việc sao lưu dữ liệu từ CSDL chính bằng RMAN DUPLICATE: công cụ RMAN tạo CSDL dự phòng bằng cách nhân bản các tệp tin được được sử dụng bởi CSDL chính, trong khi đó, việc vận hành của CSDL chính vẫn diễn ra bình thường mà không ảnh hưởng. Bằng việc khôi phục các ORLs và ARLs được nhân bản từ CSDL chính, CSDL dự phòng được đồng bộ hóa với CSDL chính.

Ngoài ra, khi thực hiện bằng cách nhân bản – khôi phục, các tham số được cấu hình tại CSDL chính cũng sẽ được đồng bộ sang cho CSDL dự phòng. Tuy nhiên, trong trường hợp này, tập tin tham số đã được sao chép bằng SCP sang trước, rút gọn thời gian trong việc thủ công tạo lại một tệp tham số mới. CSDL được sao chép gọi là TARGET, CSDL nhận và khôi phục bản sao chép gọi là AUXILIARY. Có thể kể đến một số dữ liệu được RMAN sao chép như: các datafiles hệ thống, control files, undo datafiles và tệp tin tham số cấu hình Instance của CSDL (spfile).

```
rman TARGET sys/123@pri AUXILIARY sys/123@sta
DUPLICATE TARGET DATABASE
        FOR STANDBY
        FROM ACTIVE DATABASE
        DORECOVER
        SPFILE
              SET db_unique_name = 'sta' COMMENT 'IS STANDBY'
              SET local_listener =
'(ADDRESS=(PROTOCOl=TCP)(HOST=target)(PORT=1521))' COMMENT 'IS
STANDBY'
              SET log_archive_dest_1 = 'LOCATION=/u02/oradata/shbfin/arch1/
                    VALID_FOR=(ALL_LOGFILES,ALL_ROLES)
                     DB_UNIQUE_NAME=sta' COMMENT 'IS STANBY'
        NOFILENAMECHECK;
```

Ý nghĩa của các cài đặt trong RMAN DUPLICATE như sau:

- *FOR STANDBY:* CSDL dự phòng được tạo ra từ CSDL chính thông qua việc DUPLICATE. Nếu không sử dụng thẻ này, CSDL dự phòng sẽ nhận được mã số Database Identifier (DBID) mới. Mã số DBID này được một CSDL nhỏ tên là Recovery Catalog thực hiện lưu trữ siêu dữ liệu (meta-data) của các CSDL về các thông tin cho RMAN trong quá trình sao lưu khôi phục. DBID của các CSDL khác nhau gây cho hệ thống hiểu rằng đây là các CSDL không cùng thuộc môi trường Data Guard và xung đột. Một số siêu dữ liệu mà Recovery Catalog chứa như sau: cấu trúc CSDL, thông tin về các data files, control files, archived redo logs.



Hình 20: CSDL thực hiện sao lưu thông qua meta-data được RMAN quản lý

- *FROM ACTIVE DATABASE:* RMAN thực hiện nhân bảo các data files trực tiếp từ CSDL chính tới CSDL dự phòng. Khi đó, bắt buộc CSDL chính phải khởi động từ mức MOUNTED trở lên, do tại mức độ này, control files được mở và cung cấp các thông tin về vị trí của các data files cũng như online redo logs. Nếu không sử dụng tùy chọn này, RMAN sẽ thực hiện nhân bản CSDL dựa trên bản sao lưu từ CSDL chính.
- *DORECOVER:* Khi tạo CSDL dự phòng bằng RMAN DUPLICATE, với tùy chọn này, RMAN sẽ dựng lại CSDL thông qua việc khôi phục các Archived Redo Logs, Online Redo Logs. Nếu không sử dụng tùy chọn này, RMAN sau khi nhân bản các tệp tin qua sẽ không thực hiện phục hồi và chỉ mở CSDL ở trạng thái MOUNTED.
- *SPFILE:* thực hiện nhân bản và chỉnh sửa thông tin tệp cấu hình tham số từ CSDL chính. Trong bài này, thông tin cần thay đổi là DB_UNIQUE_NAME và LOCAL_LISTENER.
- *NOFILENAMECHECK:* đường dẫn chứa data files và online redo logs file tại CSDL chính giống với CSDL dự phòng, tùy chọn này bỏ qua việc kiểm tra tên đường dẫn khi thực hiện sao lưu.

2.2.5. Cấu hình môi trường Data Guard

Để các hệ thống có thể truyền và đồng bộ thay đổi, cần cấu hình Log Transport Services và Log Apply Services. Ngoài ra cũng chuẩn bị cài đặt cho việc CSDL dự phòng trở thành CSDL chính trong trường hợp chuyển đổi. Hầu hết, các bước cấu hình môi trường Data Guard cho cả hai CSDL đều giống nhau, duy nhất chỉ có tại CSDL dự phòng, phải bật tiến trình MRPn hay Log Apply Services của CSDL dự phòng dạng vật lý để áp dụng các thay đổi, đồng bộ hóa cho CSDL.

Cấu hình Redo Transport Services đối với CSDL chính: cũng có thể nói đây là việc "*Thiết lập vị trí lưu Redo Log*" nhưng ở phạm vi toàn cục – gửi redo data sang CSDL dự phòng. Phần cấu hình này sẽ dùng tên bí danh – alias được cài đặt trong Local Naming Method thông qua tệp tin tnsname.ora, cho tham số SERVICE thay vì LOCATION với đường dẫn để lưu xuống như thông thường.

```
ALTER SYSTEM SET LOG_ARCHIVE_DEST_2=
'SERVICE=sta ASYNC
VALID_FOR=(ALL_LOGFILES,PRIMARY_ROLE)
DB_UNIQUE_NAME=sta' SCOPE=SPFILE;
```

Khai báo CSDL chính, dự phòng trong môi trường Data Guard: thông qua tham số LOG_ARCHIVE_CONFIG, liệt kê CSDL chính và CSDL dự phòng bằng tham số con DG_CONFIG. Theo mặc định, LOG_ARCHIVE_CONFIG cho phép CSDL chính gửi redo data cho CSDL dự phòng, tuy nhiên cũng có thể cài đặt không cho phép gửi từ CSDL chính hoặc không cho phép nhận từ CSDL dự phòng. Cấu trúc của câu lệnh như sau:

```
LOG_ARCHIVE_CONFIG = {
[SEND | NOSEND] [RECEIVE | NORECEIVE]
[DG_CONFIG] = {remote_db_unique_name 1}
[, … remote_db_unique_name 9] | NODG_CONFIG}
```

- SEND | NOSEND: xác định các redo data từ CSDL chính có được gửi đến CSDL dự phòng không, mặc định sẽ là SEND
- RECEIVE | NORECEIVE: xác định các CSDL dự phòng có nhận redo data từ CSDL chính không, mặc định sẽ là RECEIVE
- DG_CONFIG: xác định danh sách gồm CSDL chính và các CSDL dự phòng được nhận redo data

Cấu hình dành cho hai CSDL thực hiện trong bài sẽ như sau:

```
ALTER SYSTEM SET LOG_ARCHIVE_CONFIG='DG_CONFIG=(pri,sta)';
ALTER SYSTEM SET LOG_ARCHIVE_DEST_STATE_1=ENABLE;
ALTER SYSTEM SET LOG_ARCHIVE_DEST_STATE_2=ENABLE;
```

Cấu hình tiến trình xử lý trễ chủ động FAL cho vai trò CSDL dự phòng: tiến trình FAL có hai thành phần cần cấu hình là FAL Client và FAL Server. Cả hai tiến trình con này đều được sử dụng cho CSDL dự phòng trong việc chủ động xử lý thiếu dữ liệu thay vì bị động với tiến trình ARCn của CSDL chính.

```
# CSDL chính
ALTER SYSTEM SET FAL_CLIENT='pri';
ALTER SYSTEM SET FAL_SERVER='sta';
# CSDL dự phòng
ALTER SYSTEM SET FAL_CLIENT='sta';
ALTER SYSTEM SET FAL_SERVER='pri';
```

Cấu hình chế độ bảo vệ: Chỉ cài đặt phương thức truyền của Redo Transport Services (như ASYNC/NOAFFIRM) là chưa đủ, phương thức truyền không thể đảm bảo được dữ liệu được bảo vệ theo cách nào. Ngoài ra, chế độ bảo vệ cũng cần cấu hình phương thức truyền của Redo Transport Service phù hợp. Ví dụ như đối với Maximum Protection đảm bảo rằng sẽ không có dữ liệu bị mất/lệch bằng cách dừng hoạt động của CSDL chính khi redo data không thể truyền/áp dụng thay đổi vào CSDL dự phòng. Bảng sau đây chỉ ra yêu cầu các thông tin cần để thiết lập chế độ bảo vệ:

Maximum Availability	Maximum Performance	Maximum Protection
AFFIRM/NOAFFIRM	NOAFFIRM	AFFIRM
SYNC	ASYNC	SYNC
DB_UNIQUE_NAME	DB_UNIQUE_NAME	DB_UNIQUE_NAME

Bảng 9: Các thông tin cần để thiết lập chế độ bảo vệ

Thiết lập chế độ ưu tiên hiệu năng sau khi đã thiết lập Redo Transport Services theo phương thức không đồng bộ ASYNC/NOAFFIRM:

```
ALTER DATABASE SET STANDBY DATABASE TO MAXIMIZE
PERFORMANCE;
```

Khởi động Redo Log Apply – tiến trình MRP trên CSDL dự phòng: sau khi cấu hình thành công về môi trường Data Guard, bước cuối cùng là khởi động tiến trình MRP cho CSDL dự phòng dạng vật lý, để CSDL có thể bắt đầu áp dụng các redo data nhận được từ SRLs cũng như Archived Redo Logs.

Theo mặc định, MRP sẽ được tự động bật tính năng Real-Time Apply, hỗ trợ quá trình đồng bộ hóa diễn ra nhanh chóng, sát với CSDL chính thay vì xuất hiện độ trễ và áp dụng từ Archived Redo Logs. Tiến trình MRP sẽ áp dụng redo data từ SRLs sau khi tiến trình RFS hoàn thành việc ghi vào.

Sử dụng Real-Time Apply cùng với việc thiết kế SRLs có số lượng cũng như dung lượng lớn hơn ORLs luôn đảm bảo được việc CSDL dự phòng cập nhật “up-to-date” với CSDL chính vì Log Switch sẽ xảy ra chậm h ơn ít nhất là 1 log file so với ORLs. Nếu không muốn sử dụng Real-Time Apply, cần thêm tùy chọn DELAY cùng với khoảng thời gian giới hạn. Sử dụng thêm tùy chọn DISCONNECT sẽ đưa tiến trình này vào tiến trình chạy nền (background) thay vì theo trực tiếp phiên sử dụng của người dùng (foreground). Ngoài ra, có một số trường hợp redo data sẽ chưa được gửi

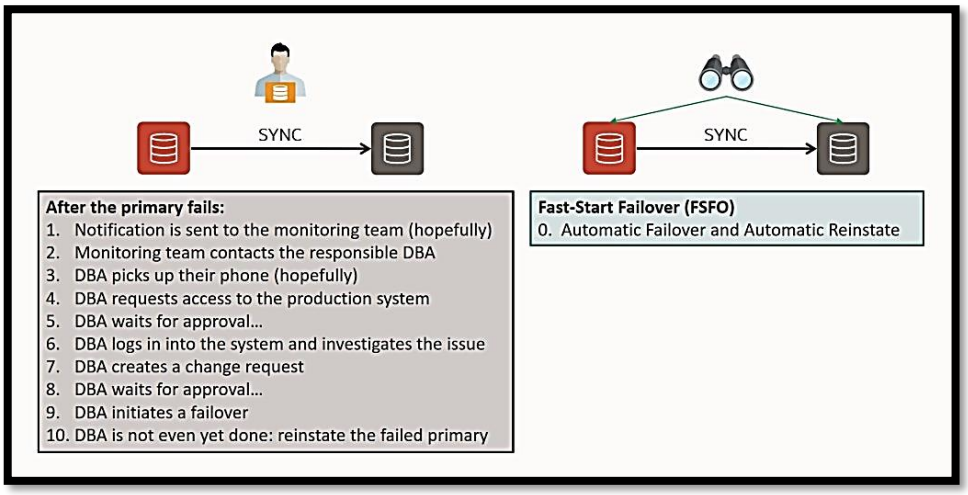
tới CSDL dự phòng nếu chưa xảy ra Log Switch ở CSDL chính, lúc này cần thực hiện thủ công hoặc sử dụng đầy một ORLs.

```
# Sử dụng MRP với Real-Time Apply
ALTER DATABASE RECOVER MANAGED STANDBY DATABASE DISCONNECT;
# Sử dụng Redo Apply với Delay Apply, thông qua Archived Redo Logs
ALTER DATABASE RECOVER MANAGED STANDBY DATABASE USING ARCHIVED LOGFILE DISCONNECT;
# Dừng hoạt động tiến trình Redo Apply
ALTER DATABASE RECOVER MANAGED STANDBY DATABASE CANCEL;
```

2.2.6. Cấu hình Fast-Start Failover với Observer

Mô hình quản lý phân tán Oracle Data Guard Broker cung cấp các chức năng để có thể truy vấn thông tin về CSDL chính và các CSDL dự phòng. Broker tự động hóa được phần lớn các công việc của người quản trị trong môi trường Data Guard như switchover để phục vụ cho việc nâng cấp, bảo trì hệ thống, thay đổi các tham số cấu hình trong môi trường.

Ngoài việc cung cấp thông tin về trạng thái hệ thống, có một tính năng nữa mà mô hình Broker cung cấp là việc đảm bảo tính sẵn sàng của hệ thống CSDL luôn sẵn sàng – tính năng Fast-Start Failover. Tính năng Fast-Start Failover cho phép Broker thực hiện chuyển đổi vai trò tự động của các CSDL. Khi CSDL chính gặp sự cố đột ngột, CSDL dự phòng được chỉ định trước sẽ đảm nhận vai trò thay CSDL chính ngay lập tức (hoặc theo một lượng thời gian được chỉ định) thay vì đợi người quản trị (DBA) thực hiện thủ công, gây mất thời gian do thủ tục hạn chế.

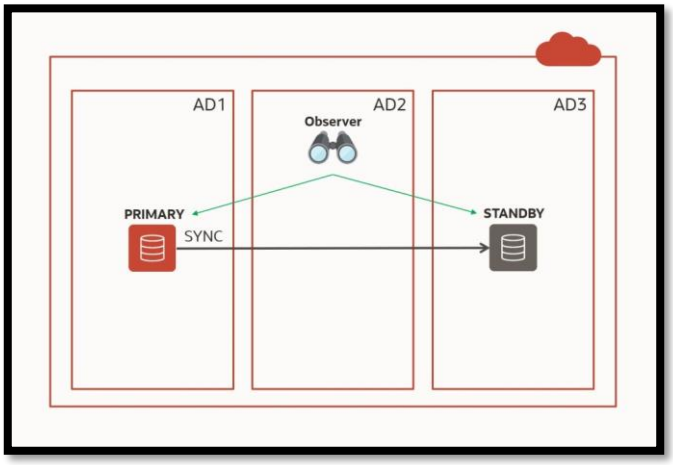


Hình 21: So sánh cơ chế Failover thủ công và tự động bằng Fast-Start Failover

Tính năng Fast-Start Failover (FSFO) dựa vào một cơ chế giám sát khác (dựa trên mô hình Broker Framework) gọi là Observer để thực hiện failover một cách tự

động. Observer có thể coi là một thành phần thứ ba (third party) bên cạnh CSDL chính và CSDL dự phòng trong môi trường Data Guard, là một cơ chế được xây dựng dựa trên nền tảng Broker. Ngoài việc thực hiện nhiệm vụ failover tự động khi CSDL chính gặp sự cố, Observer còn tự động thực hiện việc khôi phục lại CSDL chính (cũ) sau khi gặp sự cố dựa trên công nghệ Flashback. Đây là một thành phần quan trọng, đảm bảo yếu tố dự phòng và chỉ số RTO, RPO.

Theo khuyến nghị của Oracle, Observer cần được cài trên một máy chủ thứ ba ngoài hai máy chủ cài đặt CSDL, dựa trên các thư viện và môi trường hỗ trợ từ Oracle Client. Khi này, bất kể các lỗi xảy ra tại hệ thống của CSDL chính hay dự phòng đều không ảnh hưởng đến máy chủ chạy Observer. Dựa vào mô hình phân tán, hạn chế được lỗi và đảm bảo được việc giám sát, hoạt động luôn chính xác.



Hình 22: Minh họa máy chủ thứ ba chứa Observer trong môi trường Data Guard, dựa trên nền tảng điện toán đám mây OCI

Fast-Start Failover được kích hoạt khi Observer giám sát và nhận tín hiệu từ một số điều kiện sau:

- Máy chủ chứa Observer ổn định, Observer hoạt động
- Observer và CSDL dự phòng mất kết nối với CSDL chính. Nếu chỉ Observer mất kết nối với CSDL chính, thì Observer vẫn mặc định CSDL chính hoạt động thông qua CSDL dự phòng
- Observer vẫn có kết nối với CSDL dự phòng
- Thời gian chờ kết nối lại CSDL chính đã đạt giới hạn
- Các ràng buộc về hệ thống chứa CSDL như data files, các đối tượng, control files gặp lỗi, tiến trình LGWR không thể thực hiện ghi hoặc vùng nhớ chứa Archived Redo Logs bị đầy/không tồn tại
- Instance CSDL gặp lỗi bởi một hoặc nhiều các tiến trình quan trọng sau gặp lỗi: Process Monitor (PMON), System Monitor (SMON), Database Writer (DBWr), Checkpoint (CKPT), Log Writer (LGWr)

- Hoặc kích hoạt thủ công PL/SQL: DBMS_DG.INITIATE_FS_FAILOVER

Cấu hình cho mô hình Broker – Configuration: Thực hiện đặt tham số DG_BROKER_START với giá trị TRUE - khởi động tiến trình nền Oracle Data Guard monitor (DMON) tại mọi CSDL được quản lý bởi Broker. DMON là tiến trình nền nằm ở phía máy chủ (server-side), là thành phần tương tác trực tiếp với CSDL và các tiến trình DMON của CSDL khác để thực hiện giám sát và nhận thông tin.

Các thuộc tính liên quan tới môi trường Data Guard mà Broker Configuration quản lý có mối liên hệ chặt chẽ tới các tham số chung của CSDL. Chính vì điều này, mà các tham số chung đang được quản lý bởi tệp tin Server Parameter File (spfile), là nền tảng để Instance của CSDL khởi động và MOUNT với Database cũng như được quản lý bởi Broker Configuration, có nguy cơ xảy ra xung đột khi thực hiện chỉnh sửa thủ công từ một phía.

Để đảm bảo rằng Broker có thể cập nhật được giá trị của tham số của các tệp tin, người quản trị chỉ được thực hiện cấu hình trực tiếp thông qua Instance (tức là trên spfile, không phải pfile – dạng văn bản có thể đọc). Thông qua tệp tin cấu hình trên, Broker sẽ có cơ chế để xử lý xung đột thuộc tính. Ngoài ra, việc thực hiện chỉnh sửa trên Broker Configuration cũng tương tự, Broker cũng sẽ tự cập nhật ngược lại cho tệp tin cấu hình tham số spfile của Instance.

```
# Thực hiện bật tiến trình DMON ở cả hai loại CSDL
SQL> ALTER SYSTEM SET DG_BROKER_START=TRUE SCOPE=BOTH;
```

Thực hiện xóa thông tin Redo Transport Service qua LOG_ARCHIVE_DEST_n (n >= 2): Với Redo Transport Service, Broker có cơ chế tự nhận biết CSDL chính và CSDL dự phòng, do đó, việc truyền tải sẽ được thực hiện tự động. Việc xóa thông tin cho tham số này đảm bảo cho việc Broker và các tiến trình được cài đặt không bị xung đột với nhau.

```
# Thực hiện trên cả hai loại CSDL
ALTER SYSTEM SET LOG_ARCHIVE_DEST_2='';
```

Thực hiện tạo Broker Configuration cho các CSDL:

```
# Đăng nhập DGMGRL của CSDL chính
DGMGRL sys/123 as sysdba
# Thực hiện tạo Broker Configuration và thêm CSDL chính
create configuration 'DRSHBfinSolution' as primary database is 'pri' connect identifier is pri;
# Thực hiện thêm CSDL dự phòng vào Configuration
add database 'sta' as connect identifier is 'sta' maintained as physical;
```

Trong đó:

‘DRSHBfinSolution’: tên của Broker Configuration

'pri'/'sta': tên thuộc DB_UNIQUE_NAME của CSDL chính và CSDL dự phòng

pri/sta: là giá trị alias đặt trong tnsname.ora, thuộc phương thức phân giải chuỗi kết nối Local Naming Method. Broker sử dụng giá trị này để tương tác với các CSDL khác được cài đặt trong Broker Configuration

MAINTAINED AS PHYSICAL: cung cấp thông tin cho Broker Configuration loại CSDL dự phòng đang dùng

Trong Broker Configuration, có rất nhiều tham số dùng để cấu hình, thường là liên quan đến việc hiển thị trong giám sát, đặt những tiêu chuẩn, giới hạn để cảnh báo sớm và liên quan đến các chuỗi dùng để kết nối. Trong phạm vi thực nghiệm, cấu hình Broker Configuration sẽ không đi quá sâu, mỗi vấn đề lại có các cách cấu hình khác nhau cũng như phân bổ thời gian hợp lý cho các thành phần khác.

```
# Đặt chuỗi kết nối tĩnh sau khi Switchover, để các CSDL có thể tự động kết nối lại
edit database pri set property
staticconnectidentifier='(DESCRIPTION=(ADDRESS=(PROTOCOL=TCP)(PORT=1
521)(HOST=db1))(CONNECT_DATA=(SERVICE_NAME=shbfin)(INSTANCE_NA
ME=shbfin)(SERVER=DEDICATED)))';
edit database sta set property
staticconnectidentifier='(DESCRIPTION=(ADDRESS=(PROTOCOL=TCP)(PORT=1
521)(HOST=db2))(CONNECT_DATA=(SERVICE_NAME=shbfin)(INSTANCE_NA
ME=shbfin)(SERVER=DEDICATED)))';


# Đặt giới hạn để cảnh báo, do trong cùng một subnet nên có thể đặt giá trị 0
# ApplyLagThreshold: giới hạn về việc áp dụng thay đổi bị trễ
# TransportLagThreshold: giới hạn về việc truyền bị trễ
edit database pri set property ApplyLagThreshold=0;
edit database pri set property TransportLagThreshold=0;
edit database sta set property ApplyLagThreshold=0;
edit database sta set property TransportLagThreshold=0;
```

Cấu hình Fast-Start Failover: Để thực hiện sử dụng tính năng Fast-Start Failover, CSDL phải được đặt ở chế độ *Ưu tiên tính sẵn sàng*. Tham số LogXptMode – cấu hình phương thức truyền Redo Transport Services được đặt SYNC hoặc FASTSYNC tùy thuộc vào chế độ bảo vệ hoặc vai trò của CSDL. Tính năng Flashback phải được bật tại CSDL chính để thực hiện việc khôi phục nhanh (re-instate) tự động bởi Observer sau khi Failover bởi sự cố, cũng như việc cấu hình trước CSDL dự phòng nào sẽ đảm nhận vai trò chính.

```
edit database pri set property 'LogXptMode'='sync';
edit database sta set property 'LogXptMode'='sync';
edit configuration set protection mode as maxavailability;
```

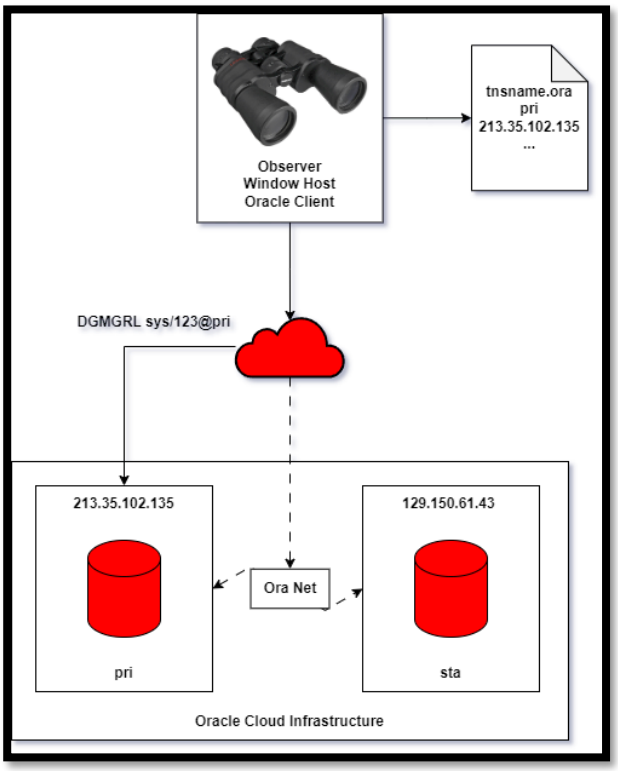
Tính năng Fast-Start Failover có rất nhiều tham số, sau đây là một vài tham số mang tính chiến lược, quyết định tính sẵn sàng của CSDL và vận hành của doanh nghiệp:

- FastStartFailoverThreshold: thời gian giới hạn mà Observer thực hiện để kết nối lại CSDL chính trước khi thực hiện Fast-Start Failover (FSFO). Thời gian này bắt đầu khi Observer bị mất kết nối với CSDL chính. Nếu Observer không thể kết nối lại CSDL chính trong khoảng thời gian kể trên, Observer sẽ kích hoạt FSFO chuyển vai trò sang cho CSDL dự phòng. Mặc định là 30 giây cho tham số này, tuy nhiên, để có cái nhìn khách quan và cấu hình đúng nhất, thì người quản trị nên tham khảo bảng v$FS_OBSERVER_HISTOGRAM để xem dữ liệu thống kê mỗi lần Observer kết nối lại mất bao nhiêu thời gian. Oracle có một số khuyến nghị về lựa chọn thời gian như sau:
    - Với Single-Instance, mạng có độ trễ thấp và tin cậy: 10 – 15 giây
    - Với Single-Instance, mạng diện rộng có độ trễ cao: 30 – 45 giây
    - Với Multi-Instance (RAC): lớn hơn 24 – 40 giây
- FastStartFailoverLagLimit: thời gian giới hạn cho phép có độ trễ trong việc CSDL dự phòng áp dụng redo data so với CSDL chính. Nếu độ trễ có thời gian lớn hơn thời gian đã chỉ định, Fast-Start Failover sẽ không được sử dụng. Tham số này được cài đặt khi CSDL ở chế độ ưu tiên hiệu năng.
- FastStartFailoverAutoReinstate: thực hiện nhiệm vụ khôi phục lại trạng thái của CSDL chính cũ sau khi Fast-Start Failover xảy ra do gặp sự cố. Ngoài ra, Broker cũng không bao giờ tự động khôi phục lại trạng thái của CSDL chính nếu Fast-Start Failover được thực hiện thủ công hoặc được kích hoạt bằng thủ tục DBMS_DG.INITATE_FS_FAILOVER
- FastStartFailoverPmyShutdown: thực hiện dừng hoạt động CSDL chính sau khi Fast-Start Failover xảy ra, thực hiện dừng các hoạt động để chuyển sang cho CSDL dự phòng, không cho phép người dùng thông thường thực hiện truy vấn tại CSDL chính cũ
- CommunicationTimeout: giới hạn thời gian cho phép Broker chờ đợi trước khi ra quyết định cảnh báo mất kết nối giữa CSDL chính và CSDL dự phòng. Với giá trị bằng 0 cho biết các CSDL không bao giờ mất kết nối, mặc định là 180 giây
- ObserverReconnect: quy định chu kỳ mà Observer thiết lập kết nối mới tới CSDL chính. Với giá trị bằng 0, Observer duy trì kết nối với CSDL chính

nhưng không định kỳ thiết lập kết nối mới. Việc thiết lập cũng có lợi trong việc phát hiện kịp thời khi không thể kết nối tới CSDL chính, tuy nhiên gây tốn kém về mặt hiệu suất và chi phí

Cấu hình máy chủ chạy Observer cho Fast-Start Failover: Observer được chạy trên một máy khác, và thường là máy khách với phần mềm Oracle Client (có môi trường giống với Oracle Database Software, nhưng giảm tải các thành phần không cần thiết đối với Client). Tại Oracle Client, sẽ thực hiện khởi động Observer trong giao diện dòng lệnh DGMGRL được kết nối tới bất kỳ CSDL nào, nhưng tiến trình, tệp lưu logs và cấu hình của Observer sẽ được chạy/lưu trên máy mà Observer sử dụng. Observer sẽ dựa vào thông tin mà Broker Configuration cung cấp để giám sát các CSDL.

Với thực nghiệm, Oracle Client được cài đặt trên máy tính Windows, phục vụ khởi động Observer cùng với các tệp tin logs, dữ liệu được lưu trữ tại Windows để giám sát, chuẩn bị cho Fast-Start Failover của hai máy ảo trên nền tảng điện toán đám mây OCI chứa CSDL chính và CSDL dự phòng, đáp ứng được tính phân tán về mặt vật lý giữa các máy chủ. Tại Oracle Client cũng cần thiết lập Local Naming Method thông qua cấu hình tnsnames.ora để Oracle có thể biên dịch chuỗi mô tả kết nối.



Hình 23: Minh họa việc thiết lập Observer tại Windows

```
# Thực hiện kích hoạt Fast-Start Failover trong Broker Configuration trước khi bật Observer trên Window Host
DGMGRL> ENABLE FAST_START FAILOVER;
```

```
# Tại Windows Host, thực hiện việc kích hoạt Observer và cấu hình đường dẫn lưu các tệp tin của Observer thích hợp. Trước khi kích hoạt, đăng nhập vào Configuration tại bất kỳ CSDL nào được quản lý bởi Broker. Nếu không thực hiện cấu hình thư mục, Observer sẽ tự động tạo ở thư mục khác
> DGMGRL sys/123@pri
DGMGRL> START OBSERVER FILE IS D:\Workspace\2023-2024-Ki-I\Do-An-Tot-Nghiep\Bai-Lam\observer\obs.dat LOGFILE IS D:\Workspace\2023-2024-Ki-I\Do-An-Tot-Nghiep\Bai-Lam\observer\log_obs.log;
```

Ngoài ra, tệp ghi trữ log của Observer cũng rất quan trọng để người dùng quản trị có thể theo dõi hành động của Observer thực hiện đối với các CSDL. Tuy nhiên, tệp log này cần mở thủ công mỗi khi người quản trị muốn thực hiện xác định lỗi. Để thuận tiện trong việc theo dõi, trong bài này, tệp log của Observer trên Windows Host được đọc tự động thông qua lập trình Bash scripts, sử dụng môi trường của Git Bash trong thời gian thực mỗi khi dòng dữ liệu mới về thông báo được thêm vào.

```
# Sử dụng Bash để chạy lệnh dưới
#!/bin/bash
# Đường dẫn đến tệp log cần đọc
logfile="D:\Workspace\2023-2024-Ki-I\Do-An-Tot-Nghiep\Bai-Lam\observer\log_obs.log"
# Thực hiện lấy ra số lượng ký tự hiện tại của tệp
lastsize=$(wc -c < "$logfile")
# Thực hiện vòng lặp với điều kiện nếu có dữ liệu mới thì in ra và cập nhật số lượng ký tự mới
while true; do
   currentsize=$(wc -c < "$logfile")
   if ((currentsize > lastsize)); then
        # Lấy ra dòng mới từ vị trí cuối cùng đã đọc trước đó
      newlines=$(tail -c +"$((lastsize + 1))" "$logfile")
        # In ra thông báo
      echo "$newlines"
      lastsize=$currentsize
   fi
        # Thực hiện dùng 1 giây trước khi lặp tiếp
   sleep 1
done
```

2.3. Phân tích sự cố mất ghi dữ liệu trong môi trường Data Guard

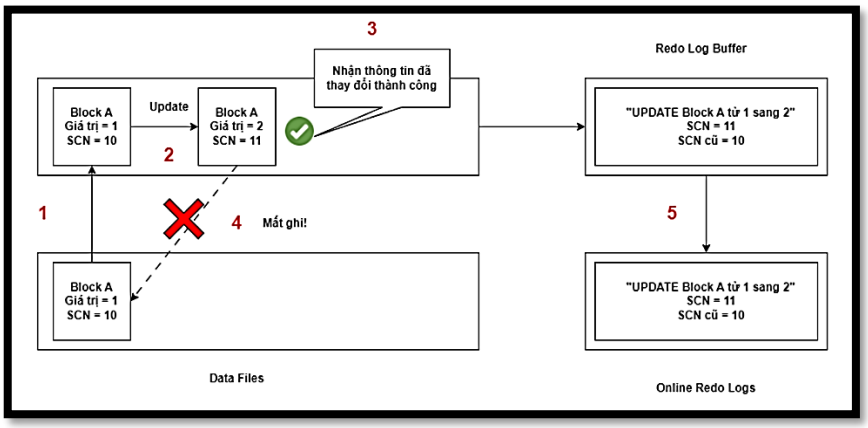
2.3.1. Khái niệm

Klinke (2021) cho rằng, mất ghi dữ liệu – Lost Writes, hay một block dữ liệu vật lý dưới thiết bị lưu trữ bị mất ghi khi các tiến trình thực hiện đọc ghi (I/O) trả về tín hiệu dữ liệu đã được ghi vào block dữ liệu, nhưng thực tế, việc ghi không thực sự xảy ra, dữ liệu tại thiết bị lưu trữ không thay đổi.

Khi dữ liệu được tải lên khu vực Data Buffer Cache của Instance, thay đổi và được COMMIT, cần phải được giải phóng (flushing dirty blocks) – ghi xuống đĩa để dành lại không gian cho các hoạt động đọc ghi khác. Instance nhận được tín hiệu đã được ghi, tuy nhiên, có thể do tiến trình DBWr, bugs trong Oracle, lỗi ổ đĩa hoặc nhiều lý do khác mà dữ liệu cần được ghi lại không thay đổi, và CSDL vẫn tiếp tục hoạt động như chưa hề có vấn đề gì xảy ra.

Lỗi mất ghi có thể xảy ra bất cứ lúc nào bởi những sự cố không thể lường trước. Trong môi trường Data Guard, lỗi mất ghi được dự báo sớm thông qua việc sử dụng tham số DB_LOST_WRITE_PROTECT. Khi này, các redo data được gửi từ CSDL chính sang CSDL dự phòng sẽ thực hiện so sánh thông qua các thông tin gồm: giá trị đọc, giá trị ghi, chỉ số SCN. Khi đặt tham số DB_LOST_WRITE_PROTECT với giá trị là TYPICAL (DB_LOST_WRITE_PROTECT), CSDL dự phòng sẽ thực hiện so sánh trước khi áp dụng thay đổi. Giá trị này cũng yêu cầu redo data chứa thêm thông tin về chỉ số SCN của block khi được đọc từ đĩa. Khi lỗi mất ghi xảy ra, CSDL dự phòng sẽ thông báo lỗi với mã ORA-00756 (lỗi lost writes) thay vì ORA-00600 (lỗi nội bộ), dễ dàng cho người quản trị biết hướng để kịp thời sửa chữa hơn.

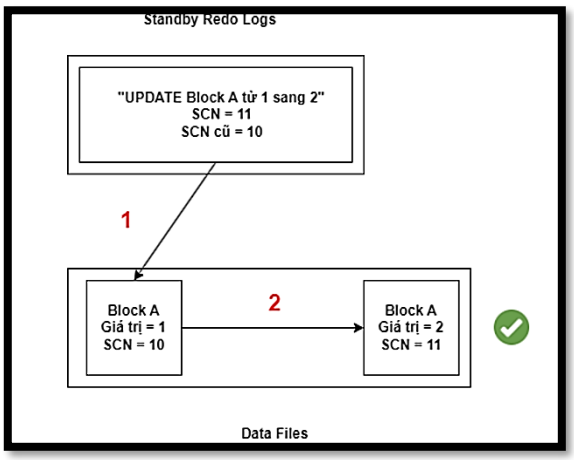
2.3.2. Phát hiện vấn đề mất ghi trong môi trường Data Guard



Hình 24: CSDL chính bị mất ghi

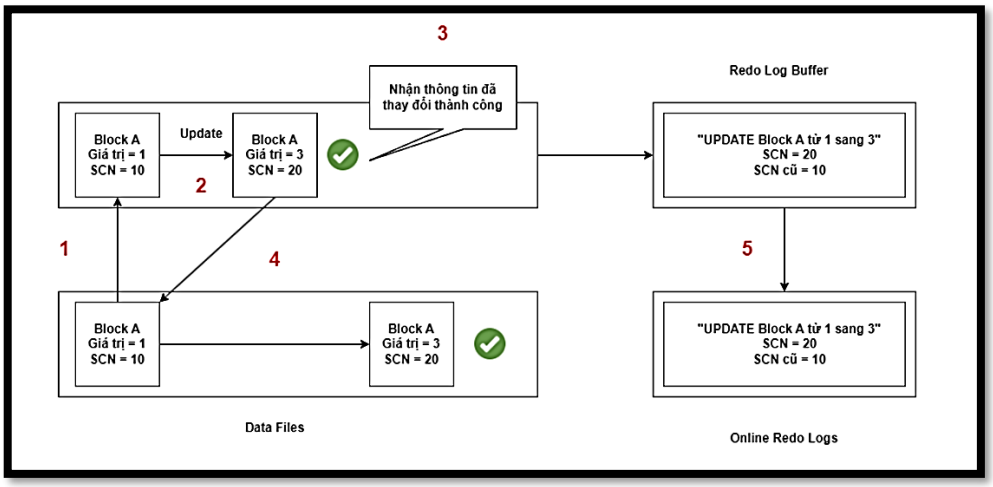
Diễn giải: (1) Block được đưa lên Data Buffer Cache theo yêu cầu của User thông qua Server Process; (2) câu truy vấn thực hiện cập nhật giá trị mới cho block, block trở thành dirty-block; (3) thực hiện Checkpoint, chỉ số SCN tăng từ 10 lên 11,

tiến trình DBWr thực hiện ghi xuống và thông báo thành công; (4) bị lỗi mất ghi, block cũ không được cập nhật và nhất quán với giá trị cũ; (5) đồng thời mọi thay đổi của thông tin mới được lưu vào Redo Buffer Cache, gồm các hành động và chỉ số SCN, từ đây redo data được lưu xuống ORLs và được truyền sang CSDL dự phòng thông qua Redo Transport Services.



Hình 25: Áp dụng thay đổi tại CSDL dự phòng

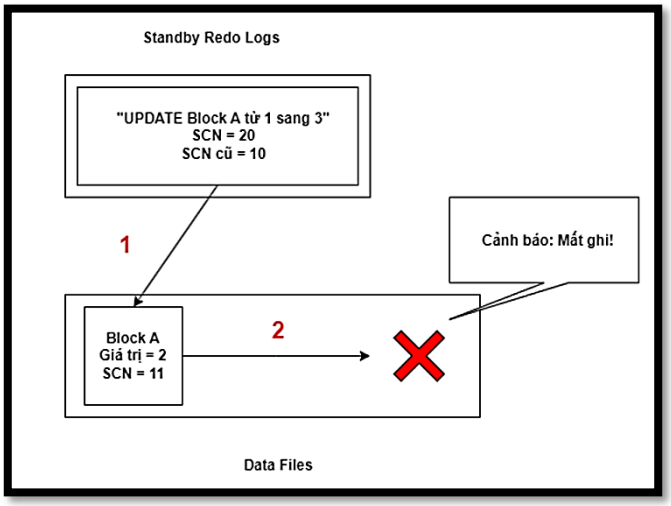
Diễn giải: (1) CSDL dự phòng nhận Block với thông tin mới được tiến trình MRP nhận từ SRLs hoặc ARLs; (2) Tiến trình MRP áp dụng thay đổi thành công dưới đĩa, CSDL dự phòng được cập nhật với thông tin mới. Mặc dù CSDL dữ liệu bị lỗi mất ghi dưới đĩa, tuy nhiên, do vẫn nhận tín hiệu thành công, CSDL chính sẽ "tưởng nhầm" đã ghi thành công thông tin mới và gửi thông tin này cho CSDL dự phòng đồng bộ.



Hình 26: Tiếp tục thay đổi thông tin với block cũ

Diễn giải: (1) block cũ tiếp tục được tải lên Data Buffer Cache để thay đổi dữ liệu theo yêu cầu của người dùng thông qua Server Process; (2) block được thay đổi từ giá trị 1 sang giá trị 3; (3) thực hiện Checkpoint, chỉ số SCN sẽ tăng từ 10 lên 20, hệ thống nhận tín hiệu thành công khi được ghi; (4) đồng thời, tiến trình DBWr sẽ thực

hiện ghi dirty-block xuống đĩa (flushing dirty-block) thành công; (5) Redo Buffer Cache nhận thông tin thay đổi cùng chỉ số SCN mới, lưu vào ORLs và gửi sang CSDL dự phòng.



Hình 27: Áp dụng sau khi xảy ra quá trình mất ghi ở CSDL dự phòng

Diễn giải: (1) Block với SCN 11 và giá trị 2 sẽ được cơ chế Lost Write Protection thực hiện so sánh giữa các block trước kia đã áp dụng và block mới nhận trước khi chuyển cho tiến trình MRP áp dụng thay đổi vào CSDL dự phòng; (2) cơ chế phát hiện SCN trước khi của block lệch với block được ghi ở đĩa (10 và 11). Ngoài ra, giá trị gốc cũng có sự khác biệt khi giá trị cũ là 1 và giá trị trên đĩa là 2. Khi này hệ thống sẽ đưa ra lỗi cụ thể cho CSDL trong tệp tin cảnh báo (alert log file) với mã lỗi ORA-00752 thay vì lỗi mang tính bao quát, không cụ thể như ORA-600.

Khi gặp lỗi ORA-00752, để giải quyết vấn đề này, thông thường (đối với CSDL nhỏ, không quan trọng) sẽ thực hiện chuyển đổi vai trò failover qua CSDL dự phòng với mục đích khiến CSDL dự phòng trở thành CSDL chính và thực hiện dựng lại CSDL chính (cũ) thông qua RMAN DUPLICATE, công nghệ Flashback không được dùng để thực hiện khi không phải Fast-Start Failover xảy ra.

2.3.3. Thực nghiệm cơ chế thông báo lỗi mất ghi

Thực hiện việc tạo data files, bảng và chuẩn bị dữ liệu cho mục đích thực nghiệm cơ chế thông báo lỗi thông qua tham số DB_LOST_WRITE_PROTECTION.

```
# Tạo data files cho một tablespace với mục đích kiểm thử
SQL> CREATE TABLESPACE LOSTWRITE DATAFILE
'/u02/oradata/shbfin/test.dbf' SIZE 100M;

# Kiểm tra lại các data files trong hệ thống, lúc này sẽ trả về kết quả gồm cả TEST.DBF
SQL> SELECT * FROM DBA_DATA_FILES;
```

```
# Tạo một bảng mới trong tablespace đã tạo
SQL> CREATE TABLE LOSTTABLE
   (ID NUMBER,
   PAYLOAD VARCHAR2(100))
   TABLESPACE LOSTWRITE;

# Chuẩn bị dữ liệu, thực hiện đồng bộ sang CSDL dự phòng
SQL> INSERT INTO LOSTTABLE(id,payload) VALUES (1, '2 trieu VND');
SQL> COMMIT;

# Tìm vị trí block của  dòng dữ liệu đã được tạo nằm trong data files, ví dụ, dòng dữ liệu này nằm trong block số 133
SELECT ROWI.D, DBMS_ROWID.ROWID_BLOCK_NUMBER(row.id), a.*
FROM LOSTTABLE a;
```

Sau khi thực hiện chuẩn bị dữ liệu cũng như tìm được vị trí của dữ liệu trong data files với đơn vị là block, tiến hành thử nghiệm sự cố mất ghi thông qua việc cập nhật dữ liệu mới (với trạng thái CSDL đã biết), khôi phục lại data files trước khi cập nhật (tình trạng mất ghi) và thực hiện cập nhật tiếp dữ liệu sau khi đã bị mất ghi.

```
# Sao lưu lại data files trước khi cập nhật, chứa giá trị ‘2 trieu VND’
> dd if=/u02/oradata/shbfin/test.dbf of=cpy_test skip=133 count=1 bs=8192
# Tìm kiếm chuỗi ‘2 trieu VND’ trong data files, đảm bảo rằng đây là tệp chứa dữ liệu đã nhập. Ví dụ, kết quả trả về “Binray file cpy_test.dbf matches”, thì có nghĩa đây là chính là dữ liệu đã nhập
> grep '2 trieu VND' cpy_test.dbf
# Thực hiện cập nhật giá trị mới cho bảng losttable
SQL> UPDATE losttable
   SET payload = '5 trieu VND'
   WHERE id = 1;
   COMMIT;
   ALTER SYSTEM CHECKPOINT;
   ALTER SYSTEM FLUSH BUFFER_CACHE;
# Khôi phục lại data files với giá trị ban đầu (‘2 trieu VND’), khi này, CSDL vừa nhận trạng thái đã ghi, tuy nhiên, giá trị thực sự của data files lại là ‘2 trieu VND’. Như vậy, đã mô phỏng lại sự cố mất ghi thành công
```

```
> dd if=cpy_test of=/u02/oradata/shbfin/test.dbf seek=135 count=1 bs=8192 conv=notrunc
```

Tại CSDL dự phòng lúc này, tiến trình Redo Log Apply – MRP đối với CSDL dự phòng dạng vật lý sẽ bị tắt và không thể thực hiện được việc áp dụng các thay đổi. Tình trạng của tiến trình MRP trong các trường hợp là giống nhau, tuy nhiên, thông báo lỗi lại mang ý nghĩa rất khác đối với giá trị đặt trong tham số DB_ LOST_WRITE_PROTECTION. Các lỗi sẽ được đưa vào tệp tin cảnh báo (alert.log) của CSDL dự phòng, ghi lại mọi hành động và thông báo của tiến trình MRP.

Với giá trị là MANUAL, CSDL dự phòng không chứa thông tin đọc ghi, chỉ có thông tin về chỉ số SCN, so sánh với nhau, nhưng đưa ra mã lỗi không cụ thể là ORA-600.

```
<msg time='2024-04-23T10:26:56.186+07:00' org_id='oracle' comp_id='rdbms'
     type='UNKNOWN' level='16' host_id='db2'
     host_addr='192.168.137.102' pid='2468'>
    …
    ORA-00600: internal error code, arguments: [3020], [5], [135], [20971655], [], [], [], [], [], [], [], []
    ORA-10567: Redo is inconsistent with data block (file# 5, block# 135, file offset is 1105920 bytes)
    ORA-10564: tablespace LOSTWRITE
    ORA-01110: data file 5: &apos;/u02/oradata/shbfin/test.dbf&apos;
    …
</msg>
```

Tuy nhiên, với giá trị là TYPICAL, mọi hành động đọc ghi được ghi theo redo data, CSDL dự phòng có thể thực hiện việc kiểm tra như theo trình bày phần khái niệm. Lúc này, CSDL dự phòng biết được thông tin gây ra hành động mất ghi khi các giá trị trước đó không khớp với nhau, người quản trị sẽ có thông tin rõ ràng hơn để chuẩn bị sửa chữa.

```
…
ORA-00752: recovery detected a lost write of a data block
…
```

2.4. Kết luận chương II

Chương II tập trung vào việc lên kế hoạch và thực nghiệm triển khai giải pháp Data Guard thông qua từng bước cụ thể cho hệ thống xếp hạng tín dụng nội bộ của SHBFinance. Việc thực nghiệm dựa trên Oracle Cloud Infrastructure để mô phỏng sự chống lỗi thông qua việc tách biệt các hệ thống với về mặt vật lý.

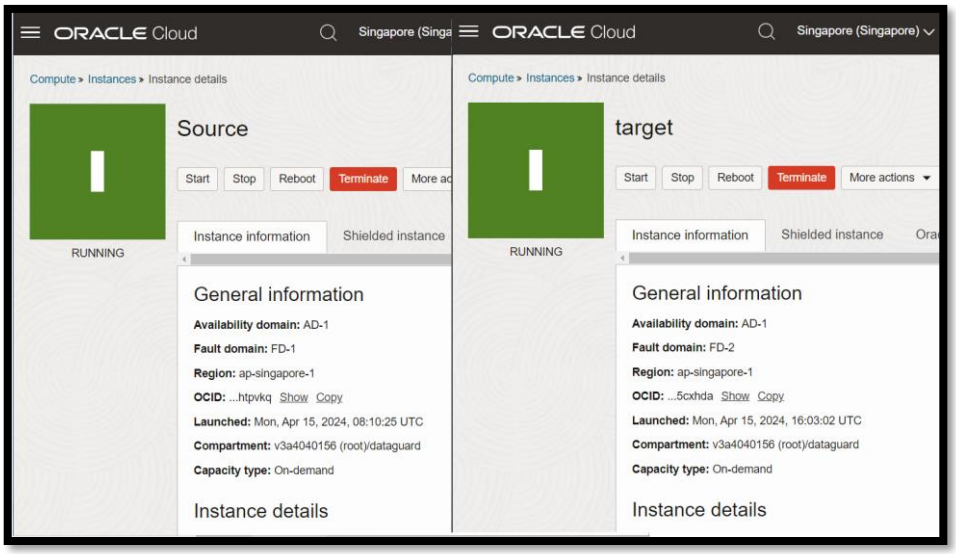
Không chỉ vậy, chương cũng phân tích sự cố mất ghi dữ liệu – "Lost Writes". Đây là sự cố đặc biệt nghiêm trọng khi không có cơ chế thông báo minh bạch nào tới người quản trị, gây mất đồng bộ giữa CSDL chính và CSDL dự phòng khi không được phát hiện kịp thời trong một khoảng thời gian dài. Để giải quyết vấn đề này, trong môi trường Data Guard cung cấp cơ chế thông báo thông qua việc so sánh thông tin gửi/nhận.

CHƯƠNG 3. KẾT LUẬN

3.1. Kết quả đóng góp

Sau khi thực nghiệm triển khai giải pháp Oracle Data Guard cho SHBFinance dựa trên CSDL dự phòng dạng vật lý, ứng dụng nền tảng điện toán đám mây Oracle Cloud Infrastructure, đề tài đạt được một số kết quả đóng góp như sau:

Thứ nhất, triển khai mô hình Data Guard dựa trên nền tảng điện toán đám mây Oracle Cloud Infrastructure, cấu hình hạ tầng vật lý hai máy chủ không cùng nằm trên một thiết bị, đảm bảo về mặt dự phòng, tránh xảy ra lỗi đồng thời trên các CSDL.

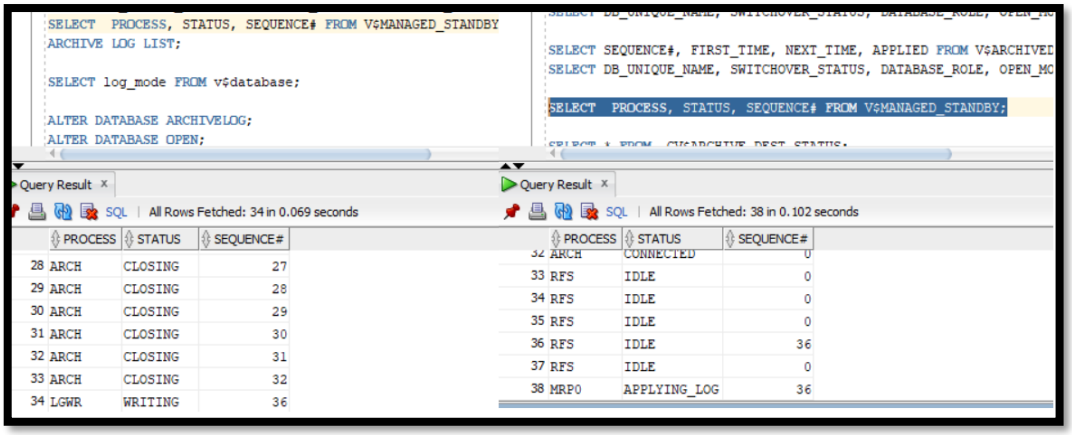


Hình 28: Cấu hình máy chủ chứa CSDL trên OCI

Thứ hai, các tiến trình quan trọng thuộc Log Apply Services và Log Transport Services hoạt động ổn định trong việc truyền và áp dụng thông tin thay đổi từ CSDL chính sang CSDL dự phòng theo chế độ Real-Time Apply, gồm cả các tiến trình xử lý trễ dữ liệu tự động:

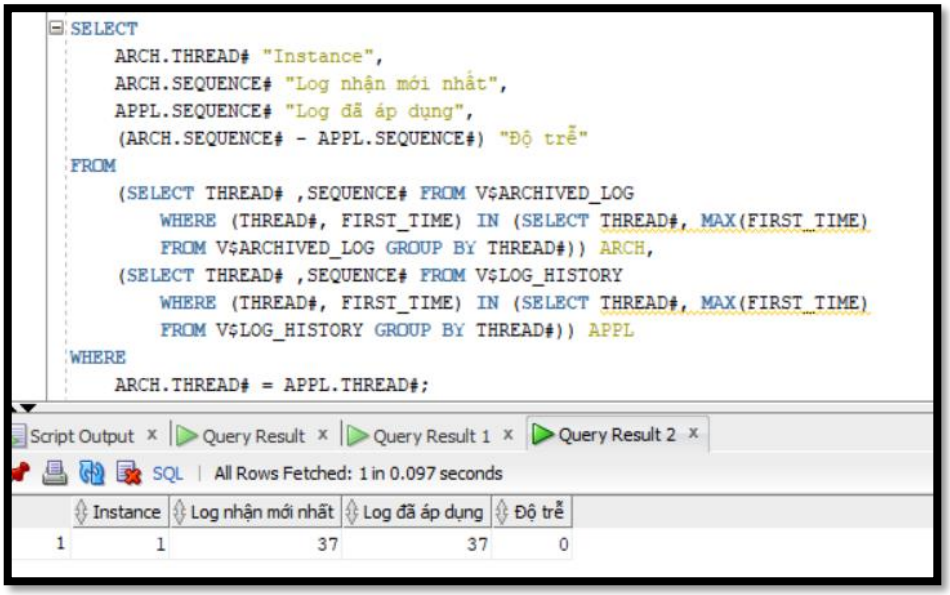
Bảng v$Managed_Standby dùng để truy vấn thông tin các tiến trình thuộc môi trường Data Guard:

- Cột “Process” thể hiện các tiến trình của CSDL chính: LGWR – ghi redo data xuống Online Redo Logs, ARCH – lưu trữ Online Redo Logs thành Archive Redo Logs, RFS – nhận và lưu trữ redo data từ CSDL chính vào Standby Redo Logs và MRP – áp dụng các redo data.
- Cột “Status” cho biết trạng thái hoạt động của tiến trình: CLOSING – hoàn thành việc lưu trữ, WRITING – ghi redo data, IDLE – không hoạt động và APPLYING_LOG – áp dụng thay đổi.
- Cột “SEQUENCE#” hiển thị thông tin thứ tự của Redo Logs đang được sử dụng.



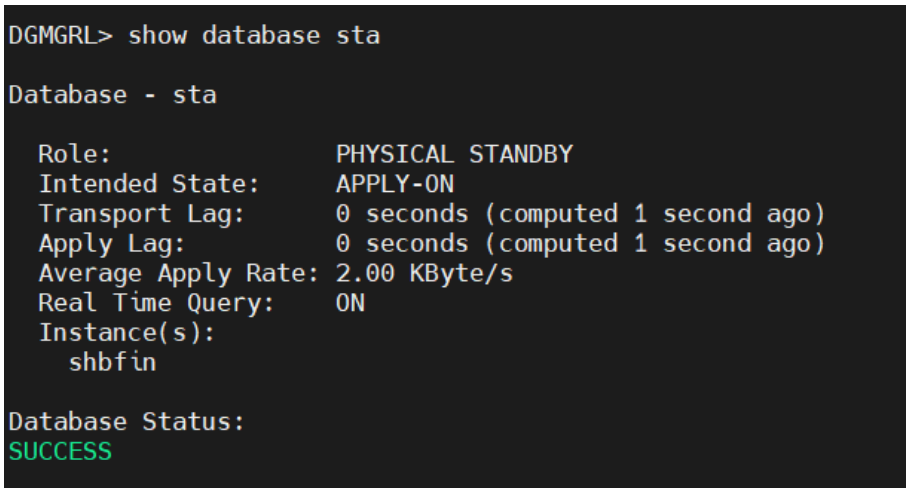
Hình 29: Các tiến trình thuộc hai CSDL trong Data Guard

Bảng v$Archive_Log cho biết thông tin các tệp Archive Redo Log. Để tìm thông tin về các Logs đã áp dụng, sử dụng bảng v$Log_History. Khi kết hợp hai bảng này với nhau, thông qua hiệu số của giá trị Log Sequence lớn nhất thuộc Archive Redo Log đã nhận và Log đã áp dụng, người quản trị sẽ biết được độ trễ hiện tại trong quá trình đồng bộ.

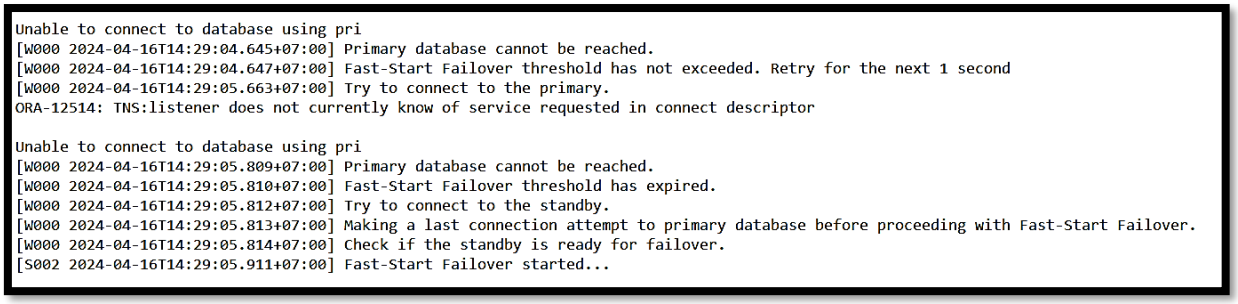


Hình 30: Truy vấn độ trễ đồng bộ

Thứ ba, giám sát thông số, kích hoạt chuyển đổi vai trò cho cả hai trường hợp có kế hoạch và gặp sự cố đột ngột thông qua thành phần giám sát thứ ba – Observer, dựa trên mô hình giám sát Broker Framework. Khi đạt đến điều kiện giới hạn trong thiết lập, Observer ra tín hiệu chuyển đổi trong trường hợp gặp sự cố - Fast-Start Failover tự động, thay vì các thao tác quản trị thủ công và hạn chế về thời gian.



Hình 31: Thông tin được cung cấp bởi Broker



Hình 32: Log của Observer quá trình Fast-Start Failover

Thứ tư, kiểm thử một số trường hợp đồng bộ và xử lý trễ trong việc đồng bộ dữ liệu giữa CSDL chính và CSDL dự phòng trong điều kiện có thể xử lý được thông qua hai cơ chế: Automatic Gap Resolution (CSDL chính chủ động), Fetch Archived Log (CSDL dự phòng chủ động). Các cơ chế xử lý trễ có vai trò quan trọng khi hệ thống gặp các sự cố về đường truyền mạng, gây mất kết nối trong môi trường Oracle Net.

Số thứ tự	Danh mục	Giá trị
1	Tên testcase	Đồng bộ dữ liệu
	Yêu cầu người dùng	Dữ liệu thêm mới tại CSDL chính được cập nhật và đồng bộ hóa tại CSDL dự phòng sau khi được COMMIT
	Dữ liệu kiểm thử	• Bảng test1(c1 INT PRIMARY KEY, c2 CHAR(6)) • Giá trị thêm: (52, ‘rap’), (56, ‘crap’)
	Kết quả mong muốn	Chạy lệnh SELECT * FROM test1 tại CSDL dự phòng sẽ trả về hai dòng giá trị: (52, ‘rap’), (56, ‘crap’)
	Kết quả hệ thống	“2 rows returned” ((52, ‘rap’), (56, ‘crap’))
	Đánh giá	Đạt

2	Tên testcase	Log Switch CSDL chính
	Yêu cầu người dùng	Khi hệ thống thực hiện Log Switch với điều kiện một Redo Log Groups đầy, Redo Logs Groups mới được sử dụng, Log Sequence tăng lên cho Groups tại CSDL chính, đảm bảo Standby Redo Logs bên CSDL dự phòng cũng phải tăng
	Dữ liệu kiểm thử	Cho 03 Redo Log Groups với giá trị Log Sequence và tình trạng như sau: • 13, CURRENT • 13, INACTIVE • 12, ACTIVE Cho 04 Standby Redo Logs với giá trị Log Sequence và tình trạng như sau: • 13, CURRENT • 0, UNASSIGNED • 0, UNASSIGNED • 0, UNASSIGNED
	Kết quả mong muốn	Redo Log Groups: • 13, ACTIVE • 14, CURRENT • 12, INACTIVE Standby Redo Logs: • 14, CURRENT • 0, UNASSIGNED • 0, UNASSIGNED • 0, UNASSIGNED
	Kết quả hệ thống	Tại Redo Log Groups: • 13, ACTIVE • 14, CURRENT • 12, INACTIVE (sai theo yêu cầu)
	Đánh giá	Chưa đạt. Nguyên nhân do sau một khoảng thời gian không sử dụng hoặc log sequence số 12 đã hoàn thành việc lưu trữ (Archived) nên trạng thái đã thay đổi. Tuy nhiên các phần còn lại đúng như yêu cầu.
3	Tên testcase	Xử lý trễ dữ liệu sau khi tiến trình MRP được bật
	Yêu cầu người dùng	Tiến trình áp dụng thay đổi CSDL dự phòng không được bật (Redo Log Apply). Sau khi tiến

		trình MRP được bật, các redo data áp dụng sang
	Dữ liệu kiểm thử	• Bảng test1(c1 INT PRIMARY KEY, c2 CHAR(6)) • Dữ liệu thêm: (61, 'help'), (66, 'trapped')
	Kết quả mong muốn	Kiểm tra không còn trễ với câu lệnh: SELECT * FROM GV$ARCHIVE_DEST_STATUS;
	Kết quả hệ thống	NO GAP
	Đánh giá	Đạt
4	Tên testcase	Xử lý trễ sau sự cố mất kết nối mạng
	Yêu cầu người dùng	Các cơ chế như Automatic Gap Resolution hoặc Fetch Archived Log xử lý trễ do một số sự cố khách quan liên quan tới đường truyền mạng khiến redo data không thể gửi. Khi có kết nối mạng, yêu cầu xử lý trễ hoàn tất. Thực hiện ngắt kết nối và kết nối lại đường truyền mạng thông qua lệnh "ipconfig down" và "ipconfig up"
	Dữ liệu kiểm thử	• Bảng test1(c1 INT PRIMARY KEY, c2 CHAR(6)) • Dữ liệu thêm: (12, 'le'), (14, 'vu')
	Kết quả mong muốn	Kiểm tra không còn trễ với câu lệnh: SELECT * FROM GV$ARCHIVE_DEST_STATUS; Không còn trạng thái "*TNS: Receive timeout occurred*" và "*RESOLVABLE GAP*"
	Kết quả hệ thống	NO GAP
	Đánh giá	Đạt
5	Tên testcase	Thực hiện Switchover cho kế hoạch
	Yêu cầu người dùng	Thực hiện switchover chuyển đổi vai trò thành công cho CSDL chính và CSDL dự phòng
	Dữ liệu kiểm thử	Sử dụng công cụ DGMGRL để theo dõi trạng thái • CSDL chính (pri) với trạng thái: Primary • CSDL dự phòng (sta) với trạng thái: Standby Truy vấn trên CSDL chính và phụ: *SELECT NAME, OPEN_MODE,*

		SWITCHOVER_STATUS, DATABASE_ROLE FROM V$DATABASE; Trả kết quả: • pri | READ WRITE | TO STANDBY | PRIMARY • sta | READ ONLY | NOT ALLOWED | PHYSICAL STANDBY
	Kết quả mong muốn	Sử dụng công cụ DGMGRL để theo dõi trạng thái • CSDL chính (pri) với trạng thái: Standby • CSDL dự phòng (sta) với trạng thái: Primary Truy vấn trên CSDL chính và phụ: *SELECT NAME, OPEN_MODE, SWITCHOVER_STATUS, DATABASE_ROLE FROM V$DATABASE;* Trả kết quả: • sta | READ WRITE | TO STANDBY | PRIMARY • pri | READ ONLY | NOT ALLOWED | PHYSICAL STANDBY
	Kết quả hệ thống	Tương tự kết quả mong muốn
	Đánh giá	Đạt
6	Tên testcase	Thực hiện truy vấn trên CSDL dự phòng
	Yêu cầu người dùng	Nhằm mục đích giảm tải CSDL chính trong tương lai, các truy vấn và tạo báo cáo cần được thực hiện bởi CSDL dự phòng với công nghệ Active Data Guard
	Dữ liệu kiểm thử	Bảng test1 với 2 dòng dữ liệu
	Kết quả mong muốn	Trả về 2 dòng dữ liệu của bảng test1
	Kết quả hệ thống	Trả về 2 dòng dữ liệu của bảng test1
	Đánh giá	Đạt
7	Tên testcase	Thực hiện Failover với Fast-Start Failover khi CSDL chính gặp sự cố
	Yêu cầu người dùng	Observer được đặt ở máy riêng khác biệt với hai máy chủ chứa CSDL. Mô phỏng lại sự cố

		hỏng CSDL chính thông qua việc ngắt hoạt động của tiến trình PMON (quan trọng) tại Instance của CSDL chính. Sau khi đạt một khoảng thời gian nhất định tại CSDL chính, Failover xảy ra tự động và chuyển quyền cho CSDL dự phòng
	Dữ liệu kiểm thử	• CSDL chính: Primary, CSDL dự phòng: Standby • Thực hiện ngắt hoạt động CSDL chính: kill -9 <UID>, với UID của tiến trình pmon trong hệ thống Oracle Linux • CSDL dự phòng được chỉ định Failover: sta
	Kết quả mong muốn	• CSDL chính: sta • CSDL chính (cũ) được thực hiện khôi phục tự động (reinstate) • Thực hiện các tác vụ bình thường của CSDL vận hành trên CSDL chính mới
	Kết quả hệ thống	Đạt các yêu cầu đề ra
	Đánh giá	Đạt
8	Tên testcase	Thực hiện Switchover về CSDL chính cũ, trả lại vai trò sau Failover
	Yêu cầu người dùng	Switchover hoạt động bình thường sau Failover, trả lại đúng vai trò của các CSDL
	Dữ liệu kiểm thử	Hai CSDL sau Failover
	Kết quả mong muốn	Trả về trạng thái cũ bình thường
	Kết quả hệ thống	Trả về trạng thái cũ bình thường
	Đánh giá	Đạt

Bảng 10: Kiểm thử hoạt động Data Guard

Thứ năm, thử nghiệm và chỉ ra sự cố đặc biệt quan trọng trong môi trường Oracle Data Guard – Lost Writes (mất ghi). Hệ thống chỉ ghi lại lỗi chi tiết cho người quản trị rõ ràng khi thực hiện cấu hình tham số DB_LOST_WRITE_ PROTECTION.

3.2. Kết luận

Khóa luận đã trình bày tổng quan về doanh nghiệp SHBFinance, hệ thống xếp hạng tín dụng của SHBFinance và điểm yếu khi công nghệ dự phòng chưa kịp thời đối với hệ thống xếp hạng tín dụng. Nhằm giải quyết vấn đề về tính dự phòng, giải pháp Oracle Data Guard được áp dụng để tạo nên môi trường có tính sẵn sàng tin cậy, đem lại lợi ích về đảm bảo vận hành kinh doanh cũng như tiết kiệm nguồn lực thời gian.

Bài viết cũng đã liệt kê những điểm lưu ý khi triển khai, trình bày chi tiết về từng thành phần trong kiến trúc Oracle Data Guard cũng như cách hoạt động, luồng đi dữ liệu đối với mỗi cách cài đặt khác nhau, giúp người hoạch định lựa chọn cách phù hợp nhất với nghiệp vụ và yêu cầu của tổ chức. Ngoài ra, sự cố nghiêm trọng – Lost Writes được chỉ ra ở mức cơ bản, nhằm cung cấp cho người triển khai có cái nhìn toàn diện hơn

Cuối cùng, việc thực nghiệm triển khai sử dụng nền tảng điện toán đám mây Oracle Cloud Infrastructure đã mô phỏng lại việc đặt hai CSDL có tính khác biệt về mặt vật lý trong cùng một hệ thống, tránh sự cố xảy ra đồng thời trên toàn CSDL, gây ra sự cố cùng một địa điểm.

3.3. Hạn chế

Trong quá trình thực nghiệm, còn một số điểm hạn chế như sau:

- Phạm vi tài liệu của doanh nghiệp SHBFinance cung cấp hạn chế, yêu cầu đảm bảo tính bảo mật của công ty (giữa doanh nghiệp triển khai và doanh nghiệp được triển khai). Chưa phản ánh được hết tính thực tế trong dự án như hạ tầng công nghệ thông tin, quy trình làm việc, thống kê về hiệu năng.
- Hệ thống CSDL trong các doanh nghiệp lớn, Công ty tài chính như SHBFinance đều sử dụng mô hình Multi-Instance hay Real Application Cluster. Trên một máy chủ chứa CSDL, có thể có từ 02 bản thể (Instance) dùng chung một CSDL trở lên nhằm phục vụ mục đích cân bằng tải, đảm bảo tính sẵn sàng cao. Ngoài ra, kiến trúc Container Database cũng được sử dụng để phân chia quyền và dữ liệu theo phòng ban bộ phận. Trong phạm vi của bài khóa luận này chưa thể hiện được các mô hình, kiến trúc trên để đảm bảo thời gian phân chia cho các phần khác, tập trung vào cách hoạt động của Data Guard hơn.
- Trong quá trình quản trị và giám sát CSDL Oracle, công cụ Oracle Enterprise Manager Cloud Control được sử dụng nhằm mục đích tối giản, tập trung trong việc quản lý cũng như cung cấp tính trực quan hóa cho người quản trị. Trong bài viết chỉ có công cụ DGMGRL được sử dụng, cung cấp giao diện dòng lệnh và phải thao tác thủ công nhiều.
- Chưa thể hiện được hết các tính năng cũng như các tham số điều chỉnh do kiến thức của em còn hạn chế về mạng truyền thông, các kiến trúc vật lý về thiết bị lưu trữ. Ngoài ra, em cũng chưa có kinh nghiệm trong việc quản trị môi trường Data Guard để có thể phát hiện những dấu hiệu bất thường khi giám sát qua các tệp tin logs cũng như các chỉ số thể hiện trên Broker Configuration.
- Chưa đưa ra được các cách thức bảo mật trong hạ tầng và trong môi trường Data Guard hay lên lịch cho sao lưu và khôi phục dự phòng bằng CSDL dự phòng. Có các phương thức bảo mật như bảo mật đường truyền trong môi trường Oracle Net, bảo mật thông tin dưới đĩa vật lý, bảo mật xác thực. Ngoài

ra, CSDL dự phòng cũng được sử dụng để tạo bản sao lưu nhằm giảm tải cho CSDL chính khi đang vận hành.

3.4. Hướng phát triển

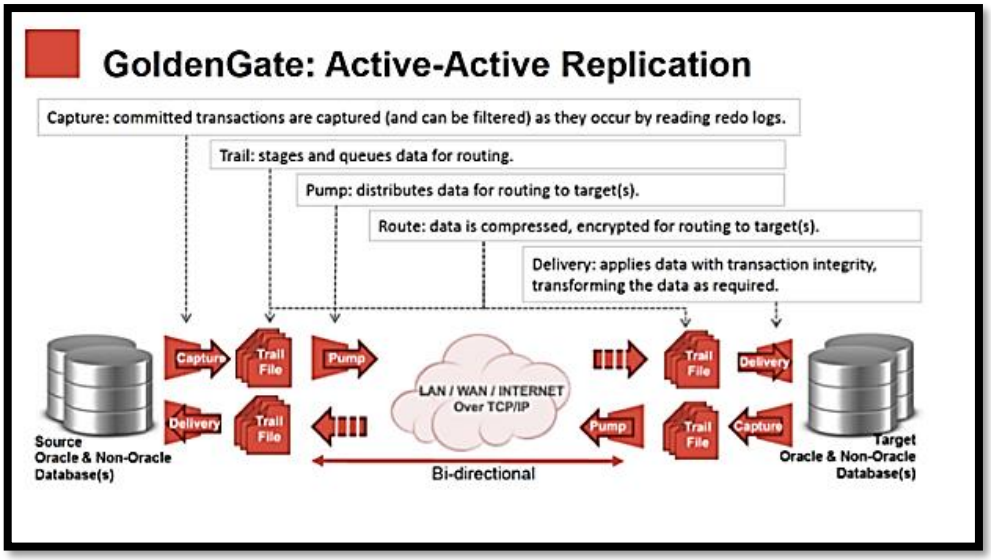
Triển khai giải pháp Oracle Data Guard dành cho hệ thống xếp hạng tín dụng nội bộ của SHBFinance là một trong những cách thức giúp cho doanh nghiệp rút ngắn thời gian về mặt khôi phục cũng như giảm thiểu số lượng dữ liệu mất mát. Hiện nay, Data Guard được xem là giải pháp không thể thiếu cho các doanh nghiệp nếu muốn đạt tới việc khôi phục sau thảm họa, bên cạnh các giải pháp mang tính sẵn sàng cao đối với các doanh nghiệp sử dụng CSDL của Oracle.

Trong tương lai gần, có một số mục tiêu mà em đặt ra để tiếp tục phát triển đề tài này như sau:

- Thực hiện thử nghiệm, triển khai cơ chế hoạt động và các sự cố mang phạm vi rộng hơn đối với toàn hệ thống, dữ liệu lớn, trên các môi trường và kiến trúc khác như Real Application Cluster và Container Database thay vì Single Instance Database.
- Triển khai tính năng DML Redirection, cho phép thực hiện các thao tác chỉnh sửa như INSERT, UPDATE, DELETE đối với CSDL dự phòng, thay vì chỉ có thể SELECT như trước đây đối với Active Data Guard.
- Triển khai mô hình sử dụng CSDL dự phòng dạng Far Sync, thực hiện cơ chế truyền quảng bá (broadcast) tới những CSDL dự phòng khác. Mô hình này giúp CSDL chính có thể hoạt động ổn định hơn, đảm bảo dữ liệu hạn chế mất khi đặt các chế độ bảo vệ Maximum Protection và Maximum Availability.
- Triển khai tích hợp giao diện quản trị Oracle Enterprise Manager Cloud Control trong môi trường Data Guard, đơn giản hóa nhiệm vụ quản trị của người quản trị viên.

3.5. So sánh với giải pháp Oracle Golden Gate

Giải pháp Oracle Data Guard và Oracle GoldenGate đều thuộc chung một loại trong kiến trúc MAA của Oracle là “replication” – nhân bản hệ thống, với mục đích là đồng bộ hóa dữ liệu giữa hai hoặc nhiều hệ thống, phục vụ cho việc dự phòng, di chuyển dữ liệu hoặc xây dựng hệ thống CSDL khác tương tự. Khi tìm hiểu Oracle Data Guard và Oracle GoldenGate, hai giải pháp này thường được xem là có điểm tương đồng, dẫn đến việc nhầm lẫn trong việc lựa chọn giải pháp phục vụ cho việc khôi phục hệ thống sau thảm họa.



Hình 33: Giải pháp Oracle GoldenGate

Trong Oracle Data Guard, các thay đổi được lưu vào redo buffer cache rồi đến redo log. Tùy thuộc vào chế độ cài đặt cho Redo Transport Services như SYNC hoặc ASYNC mà redo data được gửi và lưu xuống đĩa ngay lập tức hoặc có thể trễ hơn tới CSDL dự phòng. Nhờ vậy, các thay đổi tại CSDL chính đều được gửi đến CSDL dự phòng. Khi xảy ra sự cố, quá trình failover diễn ra tự động, kịp thời, việc vận hành được CSDL dự phòng đảm nhận.

Đối với Oracle GoldenGate, CSDL sử dụng Trail – file, lưu lại các thông tin thay đổi của Redo Log, gửi sang cho các CSDL khác, sử dụng kỹ thuật đảo ngược (reverse engineers) thành SQL và áp dụng thay đổi này vào CSDL nhận. Ngoài ra, các dữ liệu được truyền được cấu hình chỉ truyền các thông tin cần thiết mà không truyền toàn bộ dữ liệu thông qua bộ lọc (filters), các hoạt động riêng lẻ (INSERT hoặc UPDATE hoặc DELETE) hoặc bảng/cột cố định.

Oracle GoldenGate được cài đặt riêng đối với mỗi máy chủ nguồn và các máy chủ nhận, cấu hình của các máy chủ cũng có thể khác nhau. Các tiến trình thuộc môi trường GoldenGate cũng khác so với các tiến trình có trong Data Guard, bao gồm: Extract, Pumper, Collector, Replicat, Manager. Cụ thể, các tiến trình Extract, Pumper thuộc máy chủ nguồn, các tiến trình Collector, Replicat thuộc máy chủ nhận, cả hai loại máy chủ đều phải có tiến trình Manager để giám sát các tiến trình trên hoạt động ổn định.

Manager: tiến trình hoạt động ở cả hai máy chủ, các cấu hình môi trường GoldenGate được quản lý và thực hiện bởi tiến trình này. Ví dụ, các Trail – file được trích xuất bởi tiến trình Extract, tiến trình Manager sẽ nhận thông tin này và điều khiển luồng dữ liệu trong môi trường GoldenGate. Ngoài ra, tiến trình này còn quản lý một số nhiệm vụ như:

- Khởi động các tiến trình trong môi trường Oracle GoldenGate

- Duy trì, giám sát các cổng giao tiếp
- Khởi động các tiến trình động
- Quản lý Trail – files
- Thông báo các sự kiện, lỗi và thông báo các ngưỡng giới hạn

Extract: tiến trình hoạt động tại máy chủ nguồn, có nhiệm vụ trích xuất thông tin từ các giao dịch được sinh ra của CSDL, đã được đánh dấu là COMMIT. Thông tin này có thể được trích xuất từ Online Redo Logs hoặc Archived Redo Logs. Quá trình Extract phục vụ cho hai mục đích chính:

- Khởi tạo CSDL nhận: gửi toàn bộ tập dữ liệu tĩnh trực tiếp dưới dạng bảng hoặc các đối tượng khác
- Đồng bộ hóa các thay đổi: trích xuất các thay đổi của lệnh DML và DDL từ hệ thống nguồn tới các hệ thống nhận để thực hiện sao chép và đồng bộ.

Quá trình trích xuất thông tin có thể là từ bảng nguồn, các tệp logs của giao dịch (Redo Logs, Archived Logs, SQL Audit Trails), tùy thuộc vào CSDL nguồn là loại CSDL nào. Ngoài ra, trong môi trường GoldenGate cũng có thể sử dụng mô-đun bên thứ 3 để trích xuất dữ liệu từ bảng nguồn.

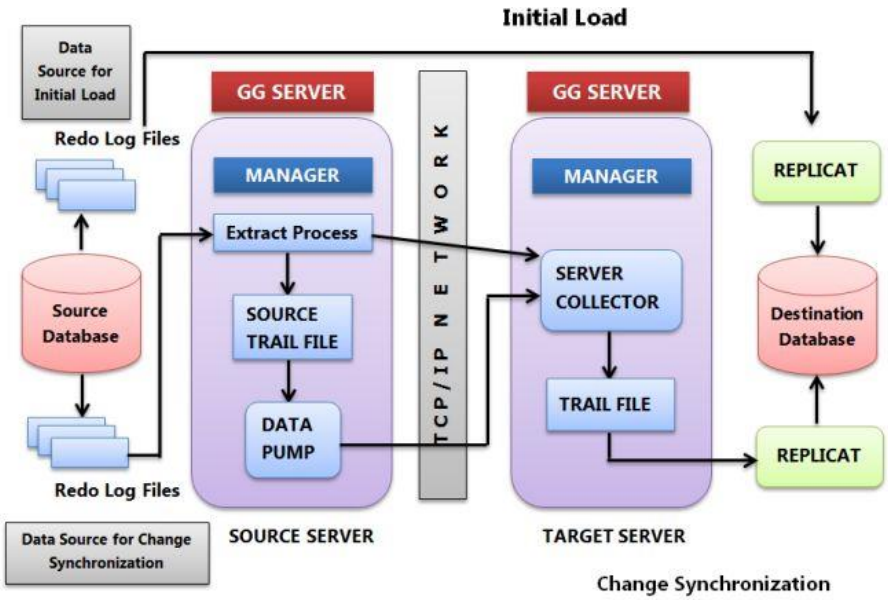
Dữ liệu và các thông tin siêu dữ liệu sẽ được trích xuất thông qua API. Sau khi trích xuất, tiến trình này có thể: gửi tới hệ thống nhận thông qua tiến trình Collector hoặc ghi xuống Local Trail Files tại hệ thống nguồn. Sau khi Extract, các công cụ lọc, chuyển hóa (filter, transformation) cũng có thể sử dụng để lấy ra các dữ liệu phù hợp.

Pumper: tiến trình thực hiện gửi các Trail – files được tạo bởi tiến trình Extract tới hệ thống nhận, thông qua cổng mặc định là 7809 (giao thức là TCP hoặc UDP), dưới sự giám sát của tiến trình Manager. Tiến trình này thực hiện khi Extract không trực tiếp gửi Trail – Files sang hệ thống nhận.

Replicat: tiến trình thực hiện xử lý Trail – files được gửi bởi Pumper. Có thể thực hiện cấu hình xử lý Trail – files theo thời gian thực hoặc sau một khoảng thời gian với độ trễ được cấu hình. Replicat biên dịch các Trail – files thành DDL và DML SQL phù hợp, thực thi những câu lệnh này trên hệ thống nhận để áp dụng và đồng bộ thay đổi với hệ thống gửi. Replicat cũng có thể cấu hình các bộ lọc, chuyển hóa về định dạng phù hợp.

Collector: tiến trình thực hiện tiền xử lý các Trail – files về dạng mà Replicate có thể sử dụng được. Tiến trình này cũng thực hiện cung cấp thông tin về cổng đang mở cho tiến trình Extract, để hai hệ thống có thể giao tiếp với nhau.

Trail – Files (hoặc Extract Files): là tệp tin mà môi trường Oracle GoldenGate sử dụng, phục vụ cho việc tiến trình Extract trích xuất các thông tin về giao dịch từ hệ thống nguồn.



Hình 34: Luồng dữ liệu của Oracle GoldenGate

Bảng so sánh dưới đây sẽ cung cấp thông tin tổng quan hơn để lựa chọn các giải pháp phù hợp:

Số thứ tự	Nội dung so sánh	Oracle GoldenGate	Oracle Data Guard
1	Nền tảng hệ thống	Hỗ trợ nhiều loại CSDL khác nhau	Hỗ trợ chỉ cùng CSDL Oracle
2	Cách thức đồng bộ	Đồng bộ nhiều chiều	Đồng bộ một chiều
3	Loại dữ liệu hỗ trợ	Không hỗ trợ XML và BLOB	Không giới hạn
4	Sao lưu dữ liệu	Chỉ những dữ liệu được nhân bản, giống nhau mới được sao lưu	CSDL chính và CSDL dự phòng chính là bản sao lưu của nhau
5	Giá cả, bản quyền	Active Data Guard trong phiên bản Oracle Database Enterprise Edition, cần mua bản quyền để sở hữu. Data Guard cơ bản không cần mua bản quyền, nhưng thiếu tính năng truy vấn đối với CSDL dự phòng.	Cần mua bản quyền cho tất cả CSDL tham gia vào môi trường GoldenGate. Tuy nhiên, khi mua bản quyền, tính năng Active Data Guard cũng nằm trong gói này.

Bảng 11: So sánh giữa GoldenGate và Data Guard

Như vậy, đối với giải pháp Oracle GoldenGate sẽ phù hợp với các nhu cầu như sau, thay vì dùng giải pháp Data Guard:

- Thực hiện việc đồng bộ dữ liệu giữa một hoặc nhiều bảng tới các CSDL có chế độ đọc ghi
- Thực hiện việc đồng bộ và chuyển hóa dữ liệu tới các bảng
- Đồng bộ thực hiện theo hai chiều, với các cơ chế xử lý xung đột
- Đồng bộ trong hệ thống với các nền tảng CSDL khác nhau (PostgreSQL, SQL Server, Teradata, TimesTen, MySQL, DB2)

TÀI LIỆU THAM KHẢO

[1] *Giới thiệu* | SHBFinance. (2024). Retrieved March 23, 2024, from Shbfinance.com.vn website: https://www.shbfinance.com.vn/ve-chung-toi/gioi-thieu-chung

[2] CỔNG THÔNG TIN ĐIỆN TỬ CHÍNH PHỦ. (2021). *Thông tư số 11/2021/TT-NHNN của Ngân hàng Nhà nước Việt Nam: Quy định về phân loại tài sản có, mức trích, phương pháp trích lập dự phòng rủi ro và việc sử dụng dự phòng để xử lý rủi ro trong hoạt động của tổ chức tín dụng, chi nhánh ngân hàng nước ngoài.* Retrieved April 29, 2024, from Chinhphu.vn website: *https://chinhphu.vn/default.aspx?pageid=27160&docid=203811*

[3] Fuller, M. (2014). *Oracle Database 12c: Data Guard Administration*. Joseph Fernandez, Veena Narasimhan.

[4] K. Keesling, D., & L. Spiller, J. (2014). *Oracle Database 12c: Administration Workshop*. Joseph Fernandez, Veena Narasimhan.

[5] Yu, P., Zhou, N., & Sun, H. (2011). *The application of Oracle Data Guard in the Logistics Distribution Management Platform*. IEEE. https://doi.org/10.1109/iccsnt.2011.6182094

[6] Liu Xiu-ju. (2010). *A brief analysis of the disaster recovery backup technology in Oracle database DataGuard*. IEEE. https://doi.org/10.1109/indusis.2010.5565635

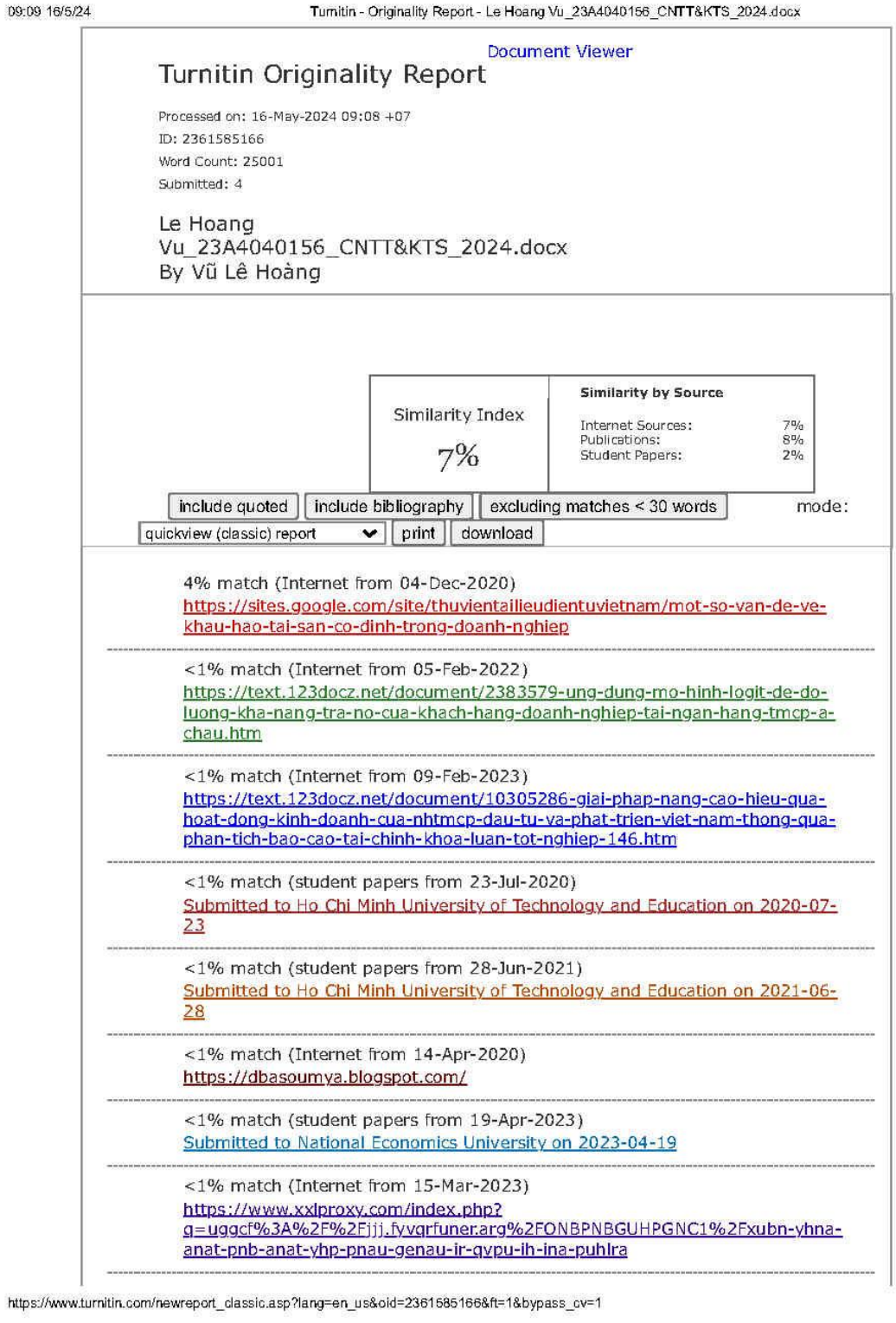
[7] Lääts, M. (2023, July 17). *Cost of Downtime in Manufacturing: Insights & Implications*. Retrieved April 25, 2024, from Evocon website: https://evocon.com/articles/cost-of-downtime-in-manufacturing-insights-implications/

[8] *Oracle Data Guard Best Practices*. (2019). Retrieved April 24, 2024, from Oracle Help Center website: https://docs.oracle.com/en/database/oracle/oracle-database/19/haovw/oracle-data-guard-best-practices.html

[9] Meeks, J., Carpenter, L., & Oracle Corporation . (2006). *Case Study: AmTrust Bank Maximum Availability Architecture – Oracle Database 10g*. Retrieved May 15, 2024, from Oracle.com website: https://www.oracle.com/us/solutions/amtrustprofile-132977.pdf

[10] Carpenter, L., Gupta, M., Meeks, J., & Ray, A. (2008). *Synchronous Data Protection Across 300 Miles NeuStar & Oracle Data Guard*. Retrieved from Oracle.com website: https://www.oracle.com/us/solutions/neustarprofile-133918.pdf

THÔNG TIN KIỂM TRA LIÊM CHÍNH



09:09 16/5/24 Turnitin - Originality Report - Le Hoang Vu_23A4040156_CNTT&KTS_2024.docx

Document Viewer

Turnitin Originality Report

Processed on: 16-May-2024 09:08 +07
ID: 2361585166
Word Count: 25001
Submitted: 4

Le Hoang
Vu_23A4040156_CNTT&KTS_2024.docx
By Vũ Lê Hoàng

Similarity Index	Similarity by Source	
7%	Internet Sources:	7%
	Publications:	8%
	Student Papers:	2%

include quoted | include bibliography | excluding matches < 30 words | mode:
quickview (classic) report | print | download

4% match (Internet from 04-Dec-2020)
https://sites.google.com/site/thuvientailieudientuvietnam/mot-so-van-de-ve-khau-hao-tai-san-co-dinh-trong-doanh-nghiep

<1% match (Internet from 05-Feb-2022)
https://text.123docz.net/document/2383579-ung-dung-mo-hinh-logit-de-do-luong-kha-nang-tra-no-cua-khach-hang-doanh-nghiep-tai-ngan-hang-tmcp-a-chau.htm

<1% match (Internet from 09-Feb-2023)
https://text.123docz.net/document/10305286-giai-phap-nang-cao-hieu-qua-hoat-dong-kinh-doanh-cua-nhtmcp-dau-tu-va-phat-trien-viet-nam-thong-qua-phan-tich-bao-cao-tai-chinh-khoa-luan-tot-nghiep-146.htm

<1% match (student papers from 23-Jul-2020)
Submitted to Ho Chi Minh University of Technology and Education on 2020-07-23

<1% match (student papers from 28-Jun-2021)
Submitted to Ho Chi Minh University of Technology and Education on 2021-06-28

<1% match (Internet from 14-Apr-2020)
https://dbasoumya.blogspot.com/

<1% match (student papers from 19-Apr-2023)
Submitted to National Economics University on 2023-04-19

<1% match (Internet from 15-Mar-2023)
https://www.xxlproxy.com/index.php?q=uggcf%3A%2F%2Fjjj.fyvqrfuner.arg%2FONBPNBGUHPGNC1%2Fxubn-yhna-anat-pnb-anat-yhp-pnau-genau-ir-qvpu-ih-ina-puhlra

https://www.turnitin.com/newreport_classic.asp?lang=en_us&oid=2361585166&ft=1&bypass_cv=1 1/36

NHẬN XÉT CỦA ĐƠN VỊ THỰC TẬP

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM

Độc lập – Tự do – Hạnh phúc

----o0o----

NHẬN XÉT VÀ XÁC NHẬN CỦA ĐƠN VỊ THỰC TẬP

Sinh viên ..... LÊ HOÀNG VŨ .....

Đã hoàn thành quá trình thực tập tại ..... Công ty Công nghệ JProTech .....

Từ ngày 24 / 01 / 2024 đến ngày 29 / 09 / 2024

Trong thời gian thực tập, sinh viên ..Lê Hoàng Vũ..... đã thể hiện được năng lực và hoàn thành công việc được giao ở mức:

☑	**XUẤT SẮC**
☐	**TỐT**
☐	**KHÁ**
☐	**ĐÁP ỨNG YÊU CẦU**
☐	**KHÔNG ĐẠT YÊU CẦU**

Xác nhận của đơn vị thực tập
(Ký và ghi rõ họ tên)

CÔNG TY TNHH CÔNG NGHỆ JPROTECH

TRƯỞNG PHÒNG HCNS
Nguyễn Thị Tohu Phương