**Slide: Lý do chọn đề tài**

Trong lĩnh vực tài chính, thuật ngữ tính liên tục trong kinh doanh – Business Continuity đề cập tới khả năng doanh nghiệp duy trì hoạt động trong và sau sự cố. Các yêu cầu đặt ra không chấp nhận các rủi ro về thời gian và hạn chế mất dữ liệu từ việc khôi phục hệ thống. Vì vậy …

---

Thưa thầy cô, các sự cố ảnh hưởng tới hệ thống CNTT, đều làm gián đoạn quá trình vận hành, kinh doanh của doanh nghiệp, gây ảnh hưởng nặng nề tới doanh thu cũng như dữ liệu bởi thời gian chết trong việc khôi phục lại hệ thống cũng như bản sao lưu không đầy đủ và cập nhật kịp thời.

Để một hệ thống thông tin đảm bảo uy tín về bảo mật, tối thiểu, cần đạt được ISO 27001, trong đó nhắc tới phải đảm bảo yếu tố sẵn sàng…cụ thể là… Chính vì thế, em mong muốn tìm hiểu và triển khai giải pháp phù hợp hơn, thông qua đề tài này trong phạm vi CSDL hệ thống tín dụng của SHBFinance.

**Slide: Xác định vấn đề**

Trong việc lên kế hoạch dự phòng, CSDL này đã triển khai mô hình RAC (một CSDL có nhiều Instance, phục vụ cân bằng tải khi lưu lượng lớn) và lên chiến lược sao lưu. Sao lưu được thực hiện với công cụ RMAN, các bản sao lưu đầy đủ được thực hiện vào Chủ nhật, các bản sao lưu về thay đổi so với bản đầy đủ được thực hiện vào nửa đêm, các ngày trong tuần. Mặc dù, dữ liệu được lên chiến lược và tần suất sao lưu được thực hiện mỗi ngày, nhưng …

--

CSDL lưu trữ các thông tin phục vụ cho mô hình dự đoán khả năng vỡ nợ của khách hàng. Đây là CSDL Oracle với kiến trúc RAC - nhiều Instance trên một CSDL, có thể cân bằng tải khi lượng truy cập lớn từ nhiều chi nhánh. Tuy nhiên, tác vụ sao lưu và khôi phục sử dụng công cụ RMAN, được lên lịch để sao lưu. Qua các bài nghiên cứu uy tín trước đây, được đăng trên website tổ chức IEEE, đã chỉ ra điểm hạn chế của công cụ này là chưa đáp ứng được việc giảm thời gian khôi phục cũng như dữ liệu khôi phục không đầy đủ

**Slide: Kiến thức về Data Guard**

Là giải pháp đảm bảo tính liên tục trong kinh doanh thông qua việc giảm thiểu thời gian ngừng hệ thống bởi các kế hoạch, sự cố làm gián đoạn tới quá trình vận hành.

Cụ thể: Doanh nghiệp cần triển khai và duy trì bản sao của CSDL. Trong khi CSDL chính phục vụ nghiệp vụ của ứng dụng, CSDL bảo sao – dự phòng thực hiện đông bộ hóa thay đổi từ các giao dịch được phát sinh từ CSDL chính. CSDL dự phòng thường được triển khai về hạ tầng vật lý khác với CSDL chính (khác thiết bị lưu trữ, trung tâm dữ liệu, địa lý) nhằm phòng trường hợp lỗi đồng thời. Phần cứng có thể khác nhau, máy chủ CSDL dự phòng thường có cấu hình nhỏ hơn nhằm tiết kiệm chi phí, phần mềm thì phải giống nhau (Enterprise Edition và File Systems).

Khi CSDL chính gặp sự cố, CSDL dự phòng đảm nhận vai trò phục vụ cho ứng dụng thông qua cơ chế chuyển đổi sau sự cố - Failover, ngoài ra còn cơ chế chuyển đổi có kế hoạch gọi là Switchover. Một số tính năng của CSDL dự phòng là: …

**Slide: Công nghệ chính trong Data Guard**

Redo Data: a description of a change made to a single block in the database (include undo segment)

Redo Logs: Each redo log file consists of redo records.

Media Failure: Deletion or corruption of a file needed for database operation -> Media Recovery -> Recover the file by applying redo information

**Công nghệ trong CSDL**

Redo Data: thông tin chứa sự thay đổi của giao dịch. Ví dụ như trên hình: khi thực hiện lệnh update, việc đầu tiên hệ thống sẽ đưa dữ liệu trong datafile dưới đĩa phù hợp với điều kiện, lên bộ nhớ để xử lý (Data Buffer Cache). Dữ liệu thay đổi, cập nhật lại datafiles. Redo data chứa thông tin về giá trị trước đây, hành động, và giá trị sau này như “Giá trị A được cập nhật thành giá trị B”

Redo Logs: Để phục vụ lưu trữ lâu dài, cần đưa redo data từ bộ nhớ xuống đĩa, lưu vào Redo Logs Files

Media Recovery: trong concept recovery và backup của oracle, lỗi media (lỗi thiết bị lưu trữ) – datafile có thể bị hỏng/thiếu dữ liệu -> Media Recovery, thực hiện tận dụng, áp dụng các redo data từ Redo Logs vào để khôi phục datafiles về trạng thái nguyên bản. Không thể hiện trên hình.

**Công nghệ trong môi trường ODG**

Trong môi trường Data Guard, các công nghệ nói trên được sử dụng chính trong việc đồng bộ hóa dữ liệu, cụ thể là: redo data được gửi sang CSDL dự phòng. CSDL dự phòng sử dụng quá trình media recovery để thực hiện khôi phục datafiles từ redo data, giúp cho CSDL dự phòng được đồng bộ với CSDL chính.

SYNC: các giao dịch được xác nhận là COMMIT, KHI redo data được gửi sang CSDL dự phòng đã được ghi xuống Redo Logs File. CSDL dự phòng sẽ gửi lại tín hiệu ACK để xác nhận với CSDL chính.

ASYNC: các giao dịch được COMMIT ngay lập tức, CSDL chính không cần nhận tín hiệu ACK từ CSDL dự phòng.

--

Protection: Sử dụng phương thức gửi SYNC, khi không nhận được tín hiệu ACK, CSDL chính sẽ treo cho tới khi nào nhận được tín hiệu ACK.

Performance: Mặc định, sử dụng ASYNC, nhờ đó, không gây gián đoạn tới quá trình vận hành. Tuy nhiên, lượng dữ liệu mất sau khi xảy ra sự cố lớn hơn so với các chế độ còn lại.

Availability: Tương tự Protection nhưng cài đặt khoảng thời gian giới hạn cho việc treo. Nếu đạt giá trị tới hạn, hoạt động theo chế dộ Perfomance và liên tục gửi lệnh Ping tới CSDL dự phòng để xử lý trễ dữ liệu.

--

Ngoài ra, có thể triển khai thêm một máy chủ giám sát trạng thái các CSDL, thực hiện chuyển đổi tự động khi gặp sự cố với điều kiện ràng buộc.

**Slide Thực nghiệm triển khai**

Sử dụng OCI với các concept hạ tầng vật lý, các cấp độ chống lỗi khác nhau. Nhưng do thời gian có hạn nên em sẽ không trình bày ở đây. Nói một cách tổng quan là: em triển khai trong cùng 1 trung tâm dữ liệu, nhưng các máy chủ khác nhau về tập hợp thiết bị vật lý, khả năng chống lỗi với các sự cố như nguồn điện, xung đột phần mềm sẽ không ảnh hưởng tới máy chủ còn lại.

Availability Domains (AD): tập hợp 1 hoặc nhiều trung tâm dữ liệu độc lập với nhau, nhưng được liên kết bằng hệ thống mạng, chứa các tài nguyên cho điện toán đám mây. Các AD không chia sẻ cùng hạ tầng vật lý -> một AD chết, các AD khác không bị ảnh hưởng. Cấp độ lỗi toàn bộ AD

Trong AD, có nhiều nhóm thiết bị phần cứng độc lập với nhau về mặt logic gọi là Fault Domains, nhằm chống lỗi liên quan đến sự cố phần cứng hoặc xung đột phần mềm trong cùng một AD.

* Vì vậy, em tận dụng concept này của OCI để triển khai hai máy chủ DB nằm ở hai FD khác nhau trong cùng 1 AD