

I. GIẢI PHÁP HỆ THỐNG LƯU TRỮ TẬP TRUNG

1. TIÊU CHUẨN CHUNG

Tiêu chuẩn thiết kế hệ thống lưu trữ

Hệ thống được thiết kế và thi công dựa trên những tiêu chuẩn quốc tế mới nhất, sử dụng công nghệ mới nhất, tiên tiến nhất, đáp ứng tốt nhất cho các ứng dụng và hoạt động hiện tại đồng thời hỗ trợ công nghệ mới trong tương lai.

- a. Được thiết kế theo cấu trúc mở, tuân thủ các tiêu chuẩn Quốc tế về thiết kế, thi công và lắp đặt.
- b. Thiết kế trên cơ sở tổng quát, có tính dự phòng các thiết bị quan trọng, dự phòng các thành phần ngay trong các thiết bị hệ thống, đảm bảo độ tin cậy cho hệ thống, thông tin dữ liệu và cho việc mở rộng sau này.
- c. Sử dụng các vật tư, thiết bị của các hãng sản xuất có uy tín trên thế giới và tại Việt nam.
- d. Hệ thống phải được phân bố một cách tối ưu theo cấu trúc thực tế của công ty, phù hợp với cơ sở hạ tầng sẵn có.
- e. Có khả năng quản trị hệ thống và khắc phục sự cố một cách dễ dàng.
- f. Hệ thống lưu trữ có khả năng quản lý chặt chẽ, thống nhất, bảo đảm cho hoạt động ổn định, giảm thiểu khả năng gây ra lỗi.
- g. Hệ thống có tính dự phòng cao, không xảy ra tình trạng điểm lỗi đơn (SPOF – Single point Of Failure) trên các thành phần.
- h. Giải pháp cho việc sao lưu, phục hồi dữ liệu phải phù hợp với thiết bị lưu trữ về dung lượng, tốc độ. Cơ chế sao lưu linh hoạt, hiệu quả: việc sao lưu, phục hồi dữ liệu có khả năng được thực hiện bằng tay hoặc tự động.
- i. Hệ thống được bảo mật thông qua thiết bị bảo mật phần cứng.
- j. Cho phép mở rộng theo yêu cầu cả về số lượng và tính năng một cách linh hoạt mà vẫn đảm bảo tốc độ hoạt động ổn định cao nhất, bảo vệ chi phí đầu tư ngay trong trường hợp tăng cường các tính năng của hệ thống như : khả năng dự phòng, tính bảo mật,...

2. TIÊU CHÍ THIẾT KẾ

Hệ thống lưu trữ đóng một vai trò hạt nhân vô cùng quan trọng trong một hệ thống. Với vai trò là thành phần trực tiếp lưu trữ dữ liệu tích hợp của tổ chức, thiết kế của hệ thống lưu trữ sẽ ảnh hưởng trực tiếp đến hiệu suất hoạt động, mức độ an toàn dữ liệu cũng như khả năng đáp ứng dịch vụ dữ liệu một cách liên tục của toàn bộ hệ

thống. Chính vì thế khi thiết kế một hệ thống lưu trữ chúng tôi tuân thủ chặt chẽ các tiêu chí sau:

- Khả năng đáp ứng uyển chuyển (*Scalability*): Đây là yêu cầu tối quan trọng của một hệ thống lưu trữ hiện đại, quyết định trực tiếp đến hiệu suất của hệ thống và hiệu quả của sự đầu tư.

- Hệ thống cần có khả năng mở rộng dung lượng dễ dàng, nhanh chóng theo yêu cầu: Điều này cho phép khả năng bắt đầu đầu tư từ một hệ thống nhỏ, và mở rộng dung lượng theo sự phát triển của dữ liệu. Khả năng này sẽ rất có lợi cho một tổ chức, khi mà dung lượng dữ liệu ban đầu ở mức nhỏ và sẽ gia tăng rất nhanh sau đó. Dữ liệu lớn đến đâu đầu tư đến đó, sẽ đảm bảo không có sự lãng phí khi đầu tư trước quá lớn, và cũng loại bỏ rủi ro việc hệ thống lưu trữ không đáp ứng được sự gia tăng đột biến của dữ liệu khi có nhu cầu triển khai nhiều ứng dụng mới.

- Hệ thống lưu trữ phải có khả năng nâng cấp tốc độ truy cập dễ dàng: Để đáp ứng nhu cầu tăng số lượng người dùng truy cập, gia tăng số ứng dụng truy cập và triển khai trên hệ thống sau này, hệ thống SAN được thiết kế để có thể dễ dàng nâng cấp tốc độ truy cập dữ liệu vật lý.

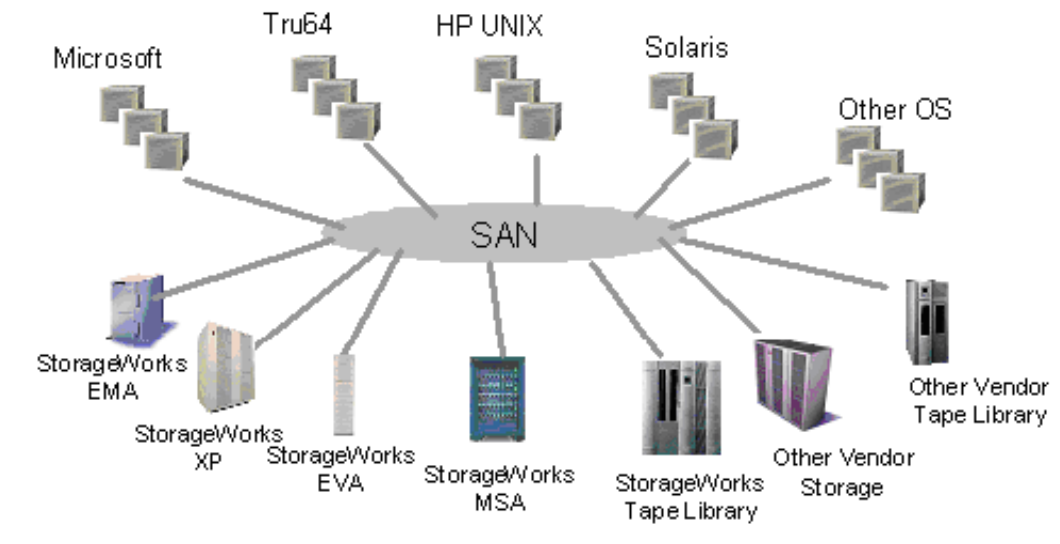
- Hệ thống lưu trữ phải có khả năng dễ dàng cài đặt thêm các máy chủ vào hệ thống: Điều này đảm bảo cho việc triển khai thêm các ứng dụng cần thêm máy chủ sau này được tiến hành một cách dễ dàng. Hệ thống SAN cho phép nâng cấp từ một volume (*phục vụ một máy chủ*) lên tới hàng nghìn volumes (*phục vụ hàng nghìn máy chủ khác nhau*) với chức năng nhiều truy cập đồng thời tới mỗi volume.

- Tính ổn định (*Stability*): Hệ thống lưu trữ SAN được thiết kế để đảm bảo tối đa sự an toàn của dữ liệu, cũng như khả năng sẵn sàng phục vụ liên tục của dữ liệu. Hệ thống cũng cho phép triển khai các chức năng Disaster Recovery cho phép lưu trữ và phục hồi lại dữ liệu từ các thảm họa tồi tệ nhất (*như cháy nổ, động đất ...*)

- Tốc độ (*Speed*): Hệ thống lưu trữ SAN có khả năng giảm thiểu thời gian truy cập cho người sử dụng và các ứng dụng bằng rất nhiều phương án khác nhau: tăng tốc độ các bộ phận cơ học trong hệ thống và đường truyền nội bộ, tăng hiệu suất của các phần mềm điều khiển, và tối ưu hoá việc cấu hình hệ thống.

- Khả năng chia sẻ, dùng chung dữ liệu (*Shareability*): Hệ thống cần được thiết kế cho phép hợp nhất dữ liệu trùng lặp, sao cho số lượng các bản sao vật lý của dữ liệu là tối thiểu, cùng với khả năng cho phép nhiều ứng dụng hay nhiều máy chủ với hệ điều hành khác nhau (*Windows, Linux, Solaris, HP-UX, AIX ..*) có thể truy cập đồng thời vào các dữ liệu này.

- Tính đơn giản (*Simplicity*): Hệ thống lưu trữ SAN được thiết kế để đơn giản hoá tối đa các thao tác quản trị, tích hợp và cấu hình cho phép triển khai hay nâng cấp nhanh chóng, sử dụng dễ dàng, nâng cao độ tin cậy và giảm các chi phí vận hành bảo trì sau này.



Hình 1. Hàng trăm terabytes trong mạng lưu trữ hiệu suất cao multi-platforms

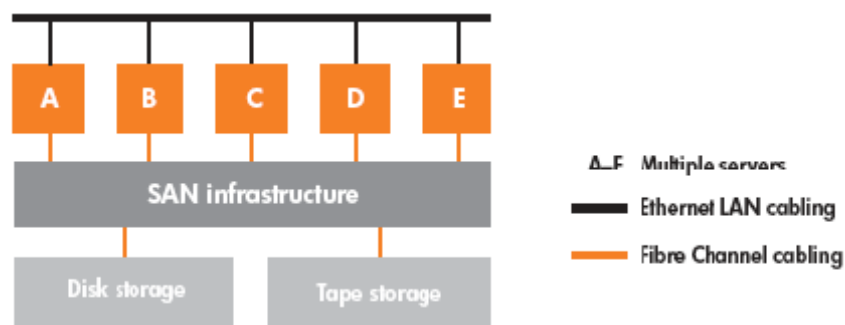
3. GIẢI PHÁP ĐỀ XUẤT

Với các tiêu chí đã nêu, giải pháp dùng hệ thống lưu trữ SAN (*Storage Area Network*) là hợp lý nhất, đây là mô hình hệ thống lưu trữ tập trung dựa trên công nghệ lưu trữ mạng tiên tiến nhất, và là xu hướng phát triển cho mọi hệ thống trung tâm dữ liệu tích hợp trên thế giới hiện nay.

a) Khái niệm hệ thống SAN

SAN là một hệ thống mạng lưu trữ chuyên dụng kết nối nhiều Server và nhiều thiết bị lưu trữ, với mục đích chính là truyền tải dữ liệu giữa hệ thống máy tính và phần tử lưu trữ và giữa các phần tử lưu trữ với nhau.

b) Cấu trúc hệ thống SAN



Hình 2. Cấu trúc cơ bản của hệ thống SAN

c) Các thành phần cơ bản của hệ thống SAN

Server:

- Nhiều server có thể của nhiều nhà cung cấp khác nhau và chạy trên các hệ điều hành khác nhau có thể kết nối đến 1 SAN. Các server này kết nối đến SAN bằng cách sử dụng các card HBA (*Host bus adapter*). Các server có thể kết nối đến SAN bằng 1 hoặc 2 kết nối (*cho cơ chế dự phòng*).

SAN infrastructure:

- SAN infrastructure bao gồm: phần cứng, cáp và phần mềm mà có thể hoạt động trong nó và trong SAN. HBA và Fibre channel switch tạo nên cấu trúc này, cho phép các server và thiết bị lưu trữ kết nối với nhau. Các switch có thể phát hiện ra lỗi hoặc nghẽn kết nối và định tuyến lại dữ liệu một cách thông minh để đến đúng các thiết bị. Khi các switch liên kết với nhau (*cơ chế phân tầng*) thì có thể tăng số kết nối đến SAN, cho phép khả năng thực hiện tốt hơn và có thể tránh được các lỗi kết nối riêng (*kết nối dự phòng*).

Disk storage:

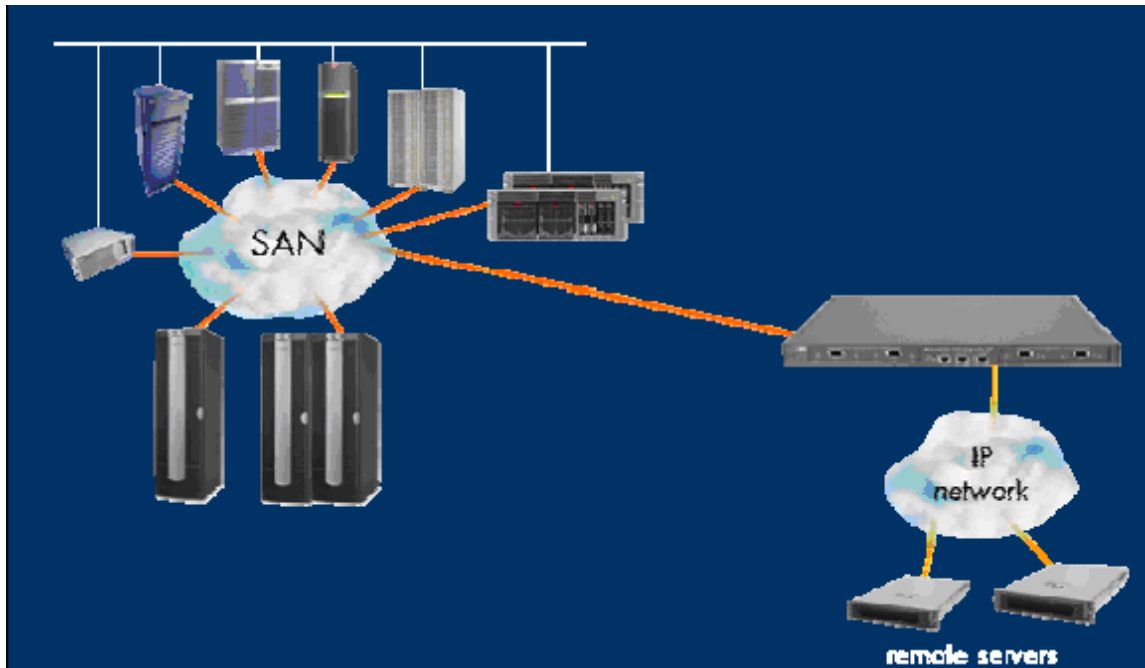
- Disk (*hoặc RAID*) array cung cấp lưu trữ riêng cho các server. Dữ liệu được ghi lên nhiều đĩa sao cho nếu 1 đĩa bị lỗi thì dữ liệu vẫn có thể truy cập từ các đĩa khác. Disk array hỗ trợ khả năng chịu đựng lỗi phần cứng cao hơn. Việc thiết kế dạng modular cho phép tăng dung lượng, thực thi và chia sẻ truy cập cho các server chạy các hệ điều hành khác nhau.

Tape storage:

- Trong SAN thì bất kỳ ổ đĩa nào (*disk array hoặc internal server storage*) có thể kết nối trực tiếp đến 1 tape library bằng Fibre channel. Điều này tạo ra một đường kết nối nhanh và riêng cho backup dữ liệu. Modem SAN là giải pháp để thực hiện cho các server riêng lẻ hoặc ở xa.

Management software:

- Để cho một hệ thống SAN hoạt động thì không thể thiếu thành phần quan trọng này. Nó giúp cho chúng ta có thể cấu hình và tối ưu các thành phần riêng cho cài đặt tốt nhất. Với phần mềm này cho phép chúng ta có thể giám sát toàn bộ mạng, tự động thực hiện các công việc như: backup, replication, mirror, ...



d) Lợi ích của SAN

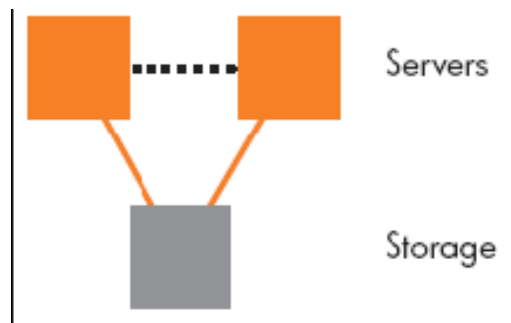
- Lưu trữ tập trung dữ liệu trên một hệ thống đơn nhất.
- Cho phép truy cập dữ liệu với tốc độ rất cao (2Gb/s; 4Gb/s tương lai lên tới 10Gb/s).
- Tính ổn định liên tục của thiết bị rất cao (99,999%).
- Dễ dàng nâng cấp, mở rộng trong tương lai (tăng thêm số lượng máy chủ, hay mở rộng dung lượng lưu trữ lên rất lớn và uyển chuyển): bảo vệ sự đầu tư thông qua khả năng tương thích ngược với các switch thế hệ cũ hơn, và khả năng nâng cấp firmware của thiết bị mà không phải dừng hoạt động hệ thống, đảm bảo mọi dịch vụ đều được duy trì liên tục 100%.
- Có khả năng sao lưu dữ liệu trong nội bộ hệ thống SAN, mà không phải dừng dịch vụ của máy chủ để sao lưu như các hệ thống lưu trữ khác, đồng thời không hề ảnh hưởng băng thông của mạng LAN khi thực hiện các thao tác backup (LAN-free backup).
- Bảo mật tốt: xác thực, xác quyền, điều khiển truy xuất và khả năng quản lý theo vùng tăng thêm mức bảo mật mạng.
- Hỗ trợ nhiều hệ điều hành và môi trường cluster, cho phép thiết kế linh động và bảo vệ sự đầu tư.

- Uyển chuyển trong khoảng cách, kết nối và hiệu suất, với khả năng hỗ trợ nhiều hơn 3000 port trong SAN lắp ghép để cho phép các kết nối cần thiết hiện tại và trong tương lai.

- Hỗ trợ cơ sở hạ tầng đa giao thức gồm FC, iSCSI, và FCIP.

e) Giới thiệu các giải pháp sẵn sàng ứng dụng trong SAN

- Server clustering: là cấu hình 2 hoặc nhiều server kết nối với nhau trong 1 cluster để đạt được một số tính năng sẵn sàng cao.



Hình 3. Server Clustering

- Multipath: kết nối giữa 1 server và 1 network storage bởi 1 vài thành phần: HBA, switch, cables và array controller. Nếu 1 path bị lỗi path còn lại sẽ được thay thế.

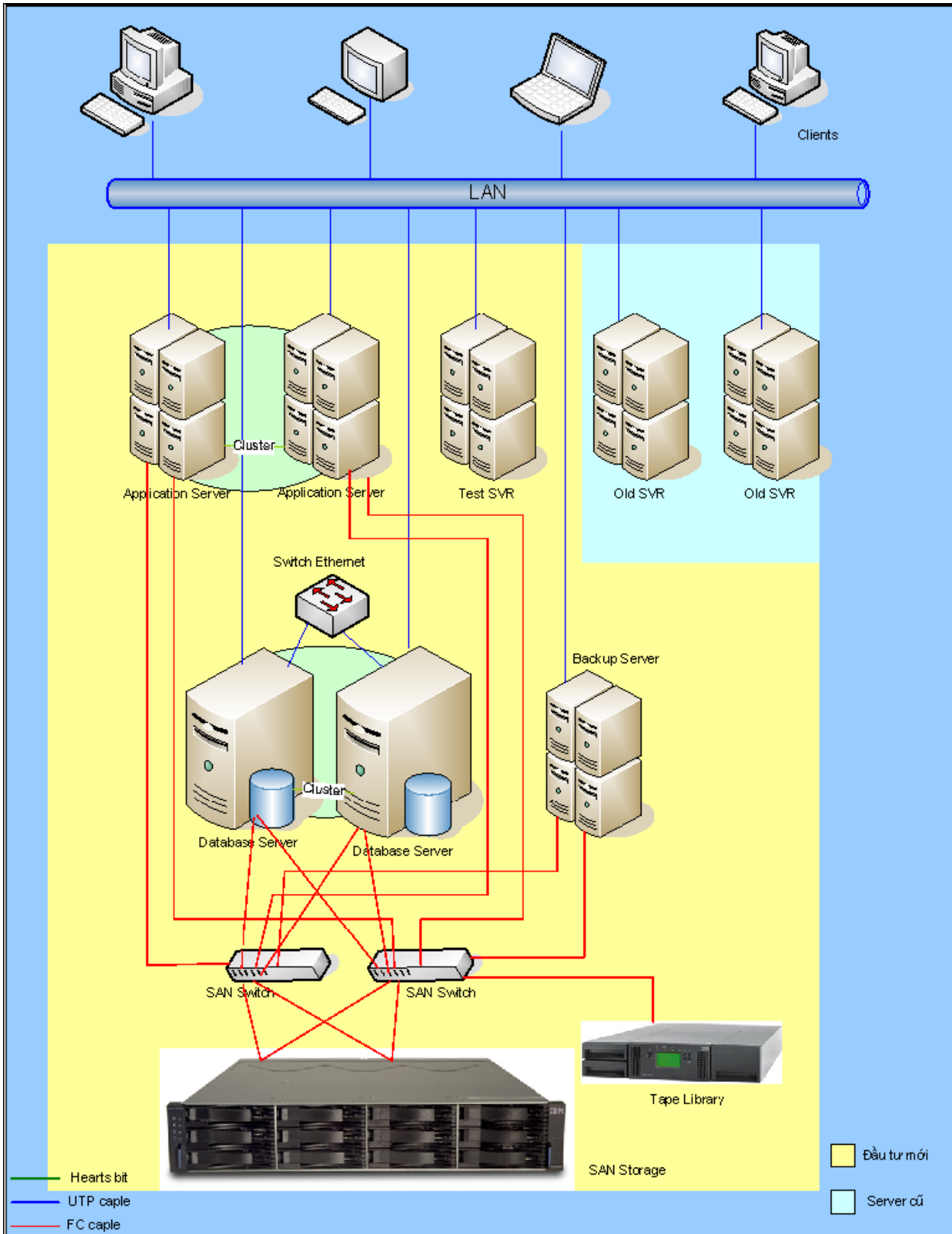


- Storage replication: cho phép duy trì truy cập dữ liệu liên tục trong trường hợp hệ thống lưu trữ gốc bị lỗi bằng cách liên tục nhân bản dữ liệu đến 1 remote secondary array.



Hình 5. Storage Replication

4. GIẢI PHÁP LƯU TRỮ TẬP TRUNG BACKUP TRONG SAN



Hình 6. Mô hình hệ thống SAN cơ chế HA backup trong SAN

a) Ưu điểm:

- Về giải pháp lưu trữ, đáp ứng được yêu cầu sẵn sàng cao, không có điểm đơn lỗi gây ra downtime hệ thống.
- Giải pháp SAN cho phép mở rộng thêm về dung lượng lưu trữ hoặc gắn thêm server dễ dàng.
- Hệ thống được backup ra tape cho phép khôi phục lại hệ thống nhanh chóng trong trường hợp cả hệ thống bị sự cố.
- Tốc độ backup nhanh chóng, không tác động đến băng thông LAN, các giao dịch luôn được đảm bảo không bị tắc nghẽn.
- Phù hợp với môi trường có dung lượng cần backup lớn.

b) Khuyết điểm:

- Chi phí đầu tư cao (*thêm server backup, sử dụng tape library công quang, phần mềm backup trong SAN*)