

# Giới thiệu các giải pháp sao lưu phục hồi dữ liệu

Trần Tuấn Hưng  
HungTT4@fpt.com.vn

## Mục lục

1	Sao lưu phục hồi dữ liệu .....	2
1.1	Sao lưu cục bộ và sao lưu tập trung .....	3
1.2	Sao lưu disk-to-tape và disk-to-disk .....	3
2	Các biện pháp sao lưu tập trung.....	5
2.1	Sao lưu qua mạng LAN .....	5
2.2	Sao lưu qua mạng lưu trữ riêng biệt SAN .....	6
2.3	Sao lưu qua mạng SAN, nhưng giảm thiểu vai trò của máy chủ backup .....	6
3	Phương thức Disk-to-Disk (d2d): bước cải tiến mới trong công nghệ sao lưu.....	7
3.1	Tiền đề dẫn đến sự ra đời của phương thức sao lưu d2d .....	7
3.2	Các cơ chế hoạt động/kích hoạt của sao lưu d2d .....	8
3.3	Ưu điểm của phương thức sao lưu d2d .....	9
4	Lời kết.....	10

## Tóm tắt

Trong quá trình vận hành các hệ thống IT, không thể thiếu chu trình sao lưu phục hồi dữ liệu. Mục đích của chu trình này là sao lưu và lưu trữ một khối lượng lớn dữ liệu bằng các môi trường lưu trữ có giá thành thấp, nhưng vẫn đảm bảo độ toàn vẹn cao cho dữ liệu trong thời gian lâu dài. Dữ liệu đã được sao lưu do đó có thể được phục hồi lại đầy đủ bất cứ khi nào có nhu cầu.

Bài viết giới thiệu và phân tích, đánh giá tính năng kỹ thuật của những biện pháp sao lưu phục hồi dữ liệu đang được áp dụng trong thực tế, đặc biệt là giải pháp sao lưu phục hồi theo xu hướng mới với tên gọi d2d (disk-to-disk).

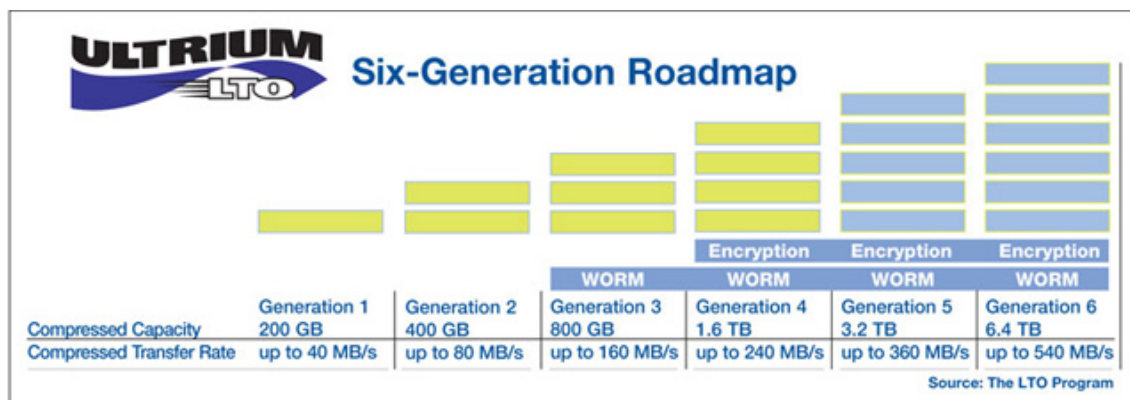
# 1 Sao lưu phục hồi dữ liệu

Hoạt động của hệ thống IT cung cấp các ứng dụng phục vụ hoạt động của bất kỳ một tổ chức, doanh nghiệp nào cũng đều dựa trên cốt lõi là xử lý dữ liệu. Hiển nhiên, theo thời gian, lượng dữ liệu cần lưu trữ và xử lý ngày càng gia tăng. Sự tăng trưởng về dung lượng này làm nảy sinh hai yêu cầu chính. Một là dữ liệu cần phải được quản lý, được lưu trữ theo suốt vòng đời của nó, và chỉ bị hoàn toàn xóa bỏ khi thật sự không còn giá trị sử dụng. Nói cách khác, tính sẵn sàng của dữ liệu cần phải được đảm bảo trong suốt vòng đời sử dụng.

Mặc dù vậy, và đây cũng là yêu cầu thứ hai, không thể và cũng không cần thiết phải lưu trữ tất cả các dữ liệu trên các môi trường lưu trữ có giá thành cao như các ổ đĩa cứng. Để đáp ứng tốc độ tăng trưởng nhanh lên đến 50-70% mỗi năm của dữ liệu theo thống kê thực tế, sẽ là rất tốn kém và không khả thi khi phải liên tục đầu tư thêm các thiết bị lưu trữ dùng các ổ đĩa cứng. Hơn nữa, giá trị sử dụng của dữ liệu nhìn chung giảm dần theo thời gian (ví dụ các file chứa dữ liệu về tình hình kinh doanh 10-20 năm về trước của doanh nghiệp không còn được truy cập đến nhiều nữa tại thời điểm hiện tại), do đó không cần thiết phải lưu trữ những dữ liệu “cao tuổi” như vậy trên các ổ đĩa cứng, một môi trường lưu trữ có giá thành không phải là rẻ. Thay vào đó cần sử dụng một môi trường lưu trữ có mức đầu tư thấp hơn, nhưng lại có độ bền cao, có thể lưu trữ dữ liệu trong thời gian lâu dài qua nhiều năm.

Hai yêu cầu trên dẫn đến sự ra đời của các giải pháp sao lưu dữ liệu truyền thống, sử dụng các thiết bị băng từ. Giá thành đầu tư cho lưu trữ dữ liệu trên băng từ rẻ hơn rất nhiều so với giá thành dùng ổ cứng. Một băng từ thông thường hiện nay (thời điểm đầu năm 2007) với dung lượng 400GByte, có thể chứa được đến 800GByte dữ liệu khi sử dụng các thuật toán nén dữ liệu theo tỷ lệ 2:1. Một băng từ như vậy giá chào bán trên thị trường chỉ vào khoảng 17% giá thành của một ổ đĩa cứng SATA có dung lượng 500GB, 7.2K, bằng khoảng 8% giá thành của một ổ đĩa cứng FC 146,8GB/15K. Có thể thấy qua ví dụ cụ thể trên sự chênh lệch lớn về giá thành đầu tư khi sử dụng đĩa cứng và băng từ để chứa dữ liệu sao lưu. Các băng từ lại có độ bền cao, đảm bảo được sự nguyên vẹn của dữ liệu trong thời gian dài. Thuận lợi nữa là các băng từ có thể được vận chuyển dễ dàng đến một địa điểm vật lý khác với địa điểm của môi trường hoạt động chính. Bảo quản các băng từ tại địa điểm riêng biệt đảm bảo sự toàn vẹn của dữ liệu trong trường hợp có sự cố hoặc thảm họa lớn xảy ra với môi trường hoạt động chính.

Trải qua quá trình phát triển, công nghệ băng từ đã tiến từ những chuẩn DAT (Digital Audio Tape), DLT (Digital Linear Tape), SDLT (Super DLT) đến chuẩn thịnh hành hiện nay là Ultrium LTO (Linear Tape Open). Theo các bước phát triển này, tốc độ truyền tải dữ liệu của các thiết bị băng từ đã ngày càng được nâng cao, và hiện nay đã lên tới 240MB/s với chuẩn mới nhất LTO-4. Không những thế, dòng sản phẩm LTO đã có lộ trình rõ ràng cho những bước phát triển tiếp theo như được hiển thị trong Hình 1.



Hình 1: Lộ trình phát triển của công nghệ băng từ LTO

### 1.1 Sao lưu cục bộ và sao lưu tập trung

Trong mô hình của giải pháp sao lưu truyền thống, bao giờ cũng có sự xuất hiện của thiết bị băng từ. Thiết bị không thể thiếu này là nơi các đầu từ đọc/ghi dữ liệu vào băng từ, đồng thời là nơi chịu trách nhiệm quản lý và xử lý các thao tác liên quan đến băng từ. Các thiết bị như vậy có một hoặc nhiều đầu từ, khay chứa băng từ, cơ chế để đọc/ghi/thay băng từ.

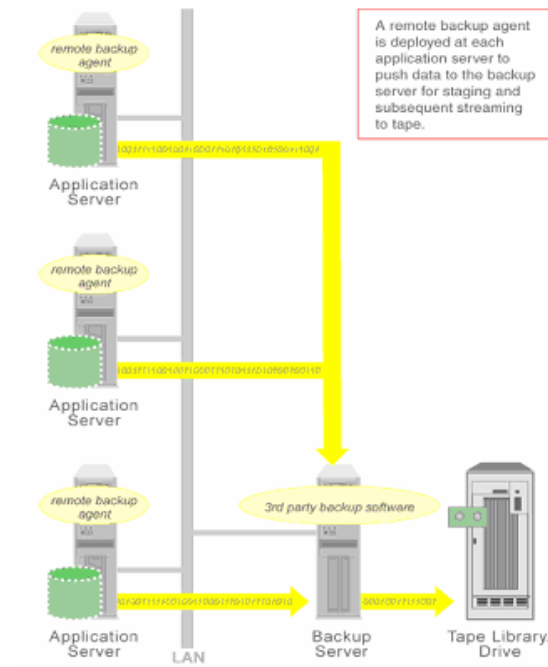
Việc kết nối vật lý cho hoạt động sao lưu về cơ bản có thể thực hiện theo 2 mô hình. Trường hợp thứ nhất là khi đầu từ và băng từ là một thành phần vật lý nội bộ của máy chủ, hoặc thiết bị băng từ là thiết bị riêng, nhưng kết nối vào máy chủ và chỉ phục vụ cho việc sao lưu dữ liệu trên máy chủ đó. Đây là mô hình *sao lưu cục bộ* và được áp dụng phổ biến trong môi trường IT có quy mô nhỏ. Trường hợp thứ hai là khi thiết bị băng từ nối vào một máy chủ, nhưng sẽ cung cấp chức năng sao lưu cho nhiều máy chủ khác nhau. Máy chủ có thiết bị băng từ kết nối trực tiếp đến thường được gọi là máy chủ backup, trên đó cài đặt phần mềm quản trị và điều khiển toàn bộ quá trình sao lưu dữ liệu của các máy chủ khác trong hệ thống. Đây là mô hình *sao lưu dự phòng tập trung*, lưu trữ dữ liệu của nhiều máy chủ về một địa điểm.

Phương pháp sao lưu tập trung phổ biến là sao lưu qua mạng LAN, tận dụng hạ tầng mạng có sẵn để lưu chuyển luồng dữ liệu sao lưu. Trong thời gian gần đây, sự phát triển của công nghệ lưu trữ đã cho ra đời phương pháp sao lưu qua mạng SAN (Storage Area Network). SAN là mạng lưu trữ riêng biệt, có tốc độ cao nhờ vào việc sử dụng giao thức quang. Các mạng SAN đã trở nên khá phổ biến và được triển khai ngày càng rộng rãi trong thực tế. Trong mạng SAN, thiết bị băng từ không nhất thiết phải kết nối trực tiếp đến máy chủ backup, nhưng phải là một thành phần của SAN. Điều này có thể thực hiện được khi thiết bị băng từ kết nối trực tiếp vào thiết bị chuyển mạch của mạng SAN. Sao lưu qua mạng SAN vẫn có sự xuất hiện của máy chủ backup, sử dụng phần mềm chuyên dụng quản lý quá trình sao lưu.

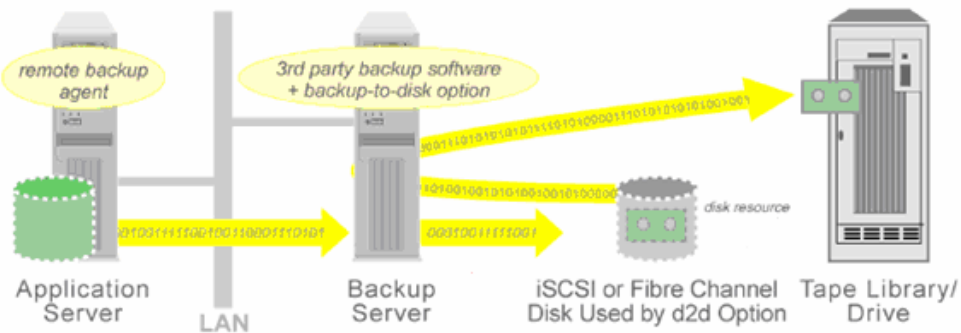
### 1.2 Sao lưu disk-to-tape và disk-to-disk

Như đã trình bày ở trên, phương thức phổ biến và có thể coi là truyền thống trong công nghệ sao lưu phục hồi là sao lưu luồng dữ liệu trực tiếp từ đĩa cứng sang băng từ, được

gọi là *phương thức sao lưu disk-to-tape*. Hiện nay, ngoài phương thức truyền thống đó, nhu cầu sử dụng dữ liệu trong thực tế cùng với sự phát triển của công nghệ lưu trữ và của nguyên lý quản lý vòng đời dữ liệu bắt đầu đẩy mạnh phương thức sao lưu mới theo 2 nấc, từ đĩa cứng sang đĩa cứng và sau đó sang băng từ. Đó là *phương thức sao lưu disk-to-disk* (d2d, hay còn được nhắc đến với thuật ngữ kỹ thuật disk-to-disk-to-tape hoặc tape-cache backup). Hình 2 và Hình 3 minh họa hai phương thức sao lưu dữ liệu disk-to-tape và disk-to-disk theo mô hình tập trung.



**Hình 2:** Mô hình sao lưu tập trung disk-to-tape

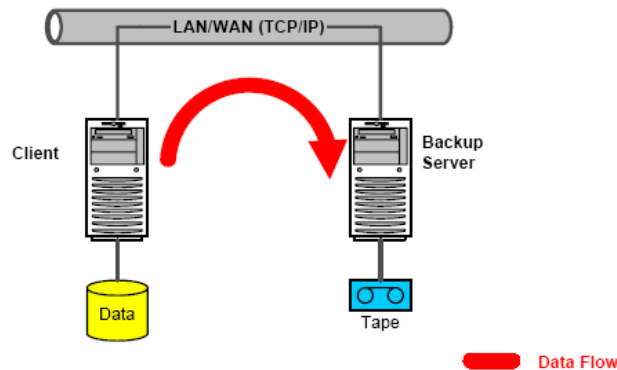


**Hình 3:** Mô hình sao lưu disk-to-disk

## 2 Các biện pháp sao lưu tập trung

### 2.1 Sao lưu qua mạng LAN

Sao lưu qua mạng LAN là khi luồng dữ liệu sao lưu sẽ được chuyển tải qua mạng LAN, tức là dùng chung cơ sở hạ tầng với các luồng dữ liệu khác trong mạng. Mô hình sao lưu là mô hình tập trung với sự có mặt của máy chủ backup có cài đặt phần mềm quản lý quá trình sao lưu. Như có thể thấy trên Hình 4, luồng dữ liệu (mũi tên màu đỏ) sẽ đi từ máy chủ được sao lưu (client), qua máy chủ backup (backup server) về thiết bị băng từ.



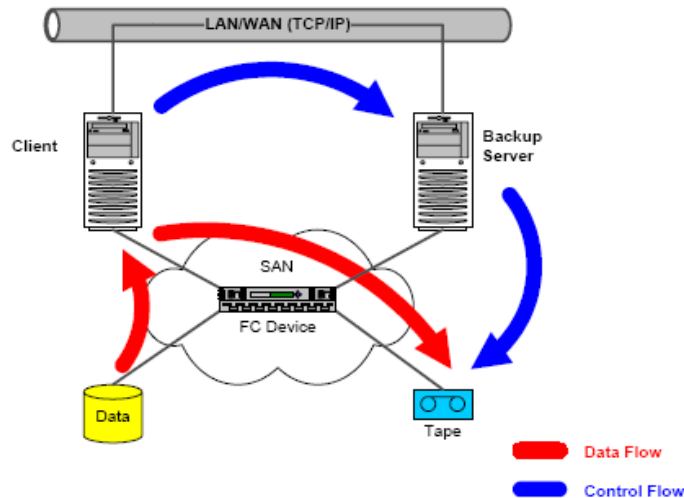
**Hình 4:** Sao lưu qua mạng LAN

Nhược điểm của biện pháp sao lưu qua mạng LAN là sự chia sẻ băng thông trong mạng giữa dữ liệu sao lưu và dữ liệu thông thường. Điều này có thể dẫn đến tắc nghẽn đường truyền (hiệu ứng nút cổ chai) khi dung lượng dữ liệu sao lưu quá lớn, ảnh hưởng đến chất lượng của các ứng dụng thông thường trong mạng. Ví dụ khi dữ liệu của ứng dụng Web bị nghẽn tại đường truyền đang tải dữ liệu sao lưu, thời gian truy cập vào trang Web sẽ trở nên lâu hơn bình thường. Sự bất bình thường, giảm sút về chất lượng dịch vụ này sẽ ảnh hưởng trực tiếp đến cảm nhận của chính người sử dụng. Ngược lại, khi lượng dữ liệu thông thường quá lớn thì cũng có hiệu ứng xấu đến luồng dữ liệu sao lưu. Thời gian cần thiết để sao lưu trọn vẹn lượng dữ liệu cần thiết sẽ kéo dài hơn mức bình thường, ảnh hưởng đến quá trình bảo toàn dữ liệu.

Tuy vậy, ưu điểm của biện pháp sao lưu qua mạng LAN là việc tận dụng nền hạ tầng mạng IP sẵn có, không cần thiết tăng thêm chi phí đầu tư và vận hành hạ tầng IT. Chi phí đầu tư ở đây bao gồm chi phí liên quan đến hạ tầng phần cứng và chi phí bản quyền cho phần mềm quản lý sao lưu tập trung. Khi sử dụng cùng một phần mềm quản lý sao lưu tập trung, đầu tư bản quyền cần thiết cho sao lưu qua mạng LAN là thấp nhất so với đầu tư bản quyền để triển khai các biện pháp sao lưu khác.

## 2.2 Sao lưu qua mạng lưu trữ riêng biệt SAN

Sao lưu qua mạng lưu trữ SAN là khi luồng dữ liệu sao lưu sẽ được chuyển tải qua mạng SAN, tách biệt với cơ sở hạ tầng với các luồng dữ liệu khác trong mạng. Mô hình sao lưu là mô hình tập trung. Như có thể thấy trên Hình 5, chỉ còn luồng dữ liệu điều khiển sao lưu (mũi tên màu xanh) là vẫn sử dụng mạng LAN, còn luồng dữ liệu sao lưu (mũi tên màu đỏ) thì đi qua mạng lưu trữ riêng biệt SAN sử dụng công nghệ truyền dẫn quang. Cần nhấn mạnh rằng về mặt lưu lượng, luồng dữ liệu điều khiển sao lưu là không đáng kể, do đó không còn tồn tại những hiệu ứng bất lợi được liệt kê trong trường hợp sao lưu qua mạng LAN vì sự chia sẻ đường truyền.



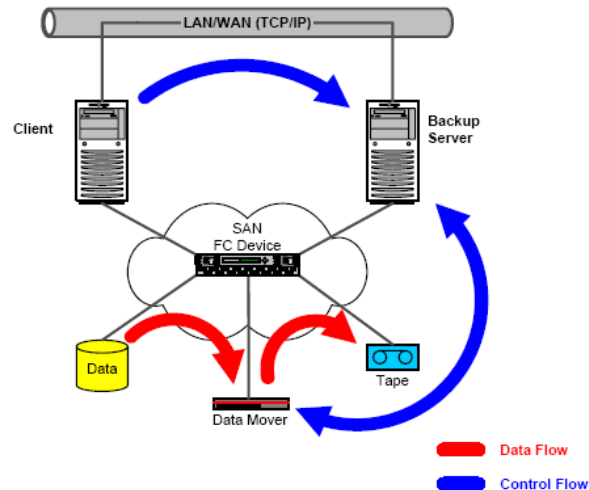
Hình 5: Sao lưu qua mạng SAN

Mạng riêng biệt SAN sử dụng công nghệ quang cho phép tốc độ truyền tải nhanh, lên đến 4Gbps tại thời điểm đầu năm 2007 và sẽ còn được tiếp tục tăng thêm trong tương lai. Ưu điểm nổi trội này giúp quá trình sao lưu qua mạng SAN có thể diễn ra nhanh chóng ngay cả khi lượng dữ liệu cần sao lưu là rất lớn. Tuy vậy, để xây dựng và thiết lập mạng SAN, cần có nguồn đầu tư ban đầu cho các thiết bị, thành phần, phụ kiện chuyên dụng. Mặc dù vậy, tùy vào từng trường hợp cụ thể, xét về lâu dài, nguồn đầu tư ban đầu có thể là xứng đáng dưới thước đo của những tiện ích mà mạng SAN mang lại.

## 2.3 Sao lưu qua mạng SAN, nhưng giảm thiểu vai trò của máy chủ backup

Chú ý kỹ hình vẽ minh họa phương thức sao lưu qua mạng SAN (Hình 5), có thể thấy vai trò của các máy chủ trong cả quy trình sao lưu là rất lớn. Cả luồng dữ liệu điều khiển sao lưu (mũi tên màu xanh), cả luồng dữ liệu sao lưu (mũi tên màu đỏ) đều đi qua máy chủ cần được backup và máy chủ backup (tức là máy chủ có cài phần mềm quản lý sao lưu tập trung). Các máy chủ cần phải xử lý luồng dữ liệu đi qua nó, định tuyến cho luồng dữ liệu đến thiết bị băng từ. Nói cách khác, các máy chủ phải có đủ tài nguyên (CPU, memory, cache...) để xử lý luồng dữ liệu sao lưu. Lượng dữ liệu cần sao lưu càng nhiều

thì yêu cầu về tài nguyên cho các máy chủ backup càng lớn. Hay nói cách khác, khi quá trình sao lưu đang diễn ra, tài nguyên của các máy chủ tham gia vào quá trình này sẽ không còn nhiều cho các ứng dụng khác có nhu cầu sử dụng các máy chủ đó.



**Hình 6:** Sao lưu qua mạng SAN, giảm thiểu vai trò của máy chủ backup

Phương thức sao lưu qua mạng SAN, đồng thời giảm thiểu vai trò của máy chủ backup ra đời xuất phát từ nhu cầu giảm thiểu khối lượng công việc mà các máy chủ tham gia trong quá trình sao lưu phải xử lý. Như có thể thấy trên Hình 6, trong biện pháp này, luồng dữ liệu sao lưu (mũi tên màu đỏ) không còn đi qua các máy chủ nữa. Thay vào đó luồng dữ liệu sẽ được truyền tải trực tiếp từ thiết bị lưu trữ đến thiết bị băng từ thông qua một thiết bị trung gian chuyên dụng trong mạng SAN, gọi là Data Mover. Thiết bị này có thể là một thiết bị vật lý riêng biệt (ví dụ Storage Router của hãng Crossroads hay SAN Data Gateway 2108-G07 của hãng IBM), nhưng cũng có thể được tích hợp sẵn trong các thiết bị chuyên mạch quang (ví dụ dòng SAN switch MDS 9000 của hãng Cisco).

Như đã nhắc đến ở trên, phương thức sao lưu giảm thiểu vai trò của các máy chủ mang lại sự giảm tải đáng kể cho các máy chủ. Như vậy tài nguyên của các máy chủ đó có thể được sử dụng để phục vụ cho các ứng dụng khác. Nhưng bên cạnh đó, việc bắt buộc phải có thiết bị trung gian Data Mover cũng tương đương với việc tăng chi phí đầu tư ban đầu. Do đó cần tùy vào từng trường hợp thiết kế hệ thống sao lưu cụ thể để đánh giá lợi ích tựu chung có được do áp dụng phương thức sao lưu này, và từ đó quyết định có nên đầu tư đưa vào sử dụng hay không.

### **3 Phương thức Disk-to-Disk (d2d): bước cải tiến mới trong công nghệ sao lưu**

#### **3.1 Tiền đề dẫn đến sự ra đời của phương thức sao lưu d2d**

Như đã đề cập đến ở trên, phương thức sao lưu truyền thống disk-to-tape chuyển tải dữ liệu từ các đĩa cứng sang băng từ. Với tốc độ truyền tải cao của mạng SAN hiện nay, có

thể nói cách làm này không bị hạn chế gì về tốc độ truyền dữ liệu. Nhưng nếu chúng ta quan tâm đến quá trình phục hồi, tức là quá trình chuyển dữ liệu ngược lại từ băng từ về các ổ đĩa cứng khi có nhu cầu, thì phương thức sao lưu disk-to-tape lại ẩn chứa nhiều bất cập.

Điểm khác nhau cơ bản giữa truy cập dữ liệu được lưu trữ tại đĩa cứng so với dữ liệu được lưu trữ tại băng từ là cơ chế ngẫu nhiên (random access) cho đĩa cứng và cơ chế thứ tự (sequential access) cho băng từ. Trong cơ chế ngẫu nhiên, thời gian truy cập vào dữ liệu không phụ thuộc vào vị trí vật lý của nó trên đĩa. Trong cơ chế truy cập theo thứ tự, vị trí vật lý của dữ liệu ảnh hưởng trực tiếp đến thời gian truy cập. Băng từ phải được quay đến đúng vị trí dữ liệu đang được lưu trữ, sau đó mới có thể truy cập vào dữ liệu. Như vậy rõ ràng là thời gian để phục hồi dữ liệu từ băng từ sẽ lâu hơn từ đĩa cứng. Đây cũng là nhược điểm lớn nhất của phương pháp sao lưu phục hồi truyền thống sử dụng băng từ.

Phương thức sao lưu disk-to-disk được phát triển để khắc phục thời gian truy nhập dữ liệu trong quá trình phục hồi. Sao lưu theo phương thức disk-to-disk (gọi ngắn gọn là d2d, hay còn được nhắc đến với thuật ngữ kỹ thuật disk-to-disk-to-tape hoặc tape-cache backup) về mặt vật lý là khi dữ liệu trước hết được sao lưu từ đĩa cứng sang đĩa cứng, thay vì sang băng từ như trong phương thức thông thường. Dữ liệu từ đĩa cứng sao lưu, sau khi được lưu trữ một thời gian dài cần thiết, mới được chuyển tiếp (sao chép, hoặc tạo bản clone) sang băng từ vật lý tại thời điểm được kích hoạt tùy theo nhu cầu của người quản trị. Như vậy quá trình phục hồi dữ liệu phần lớn sẽ là chuyển dữ liệu ngược lại từ đĩa cứng sang đĩa cứng, giảm thiểu được thời gian truy cập vào dữ liệu cần được phục hồi, và theo đó là giảm thời gian phục hồi dữ liệu.

### **3.2 Các cơ chế hoạt động/kích hoạt của sao lưu d2d**

Cơ chế host-based là cơ chế đơn giản nhất để thực hiện sao lưu disk-to-disk. Cơ chế này tận dụng tính năng của các phần mềm quản lý sao lưu đang dùng và các ổ đĩa cứng trong hệ thống. Người quản trị chỉ cần kích hoạt tính năng d2d của phần mềm quản lý sao lưu trên máy chủ backup. Qua giao diện điều khiển của phần mềm backup, cần định nghĩa rõ điểm đến cho luồng dữ liệu cần sao lưu là ổ đĩa cứng nào. Thao tác này được thực hiện nhanh và khá đơn giản.

Phương thức tạo bản sao gương (mirroring) từ một thiết bị từ đĩa sang một từ đĩa khác cũng có thể được coi là một hình thức của giải pháp sao lưu d2d. Để làm được điều này, nhiều khi có thể tận dụng các tính năng copy hoặc mirroring sẵn có của thiết bị lưu trữ mà không cần phải sử dụng phần mềm backup chuyên dụng.

Một cơ chế khác là cơ chế appliance-based. Cơ chế hoạt động này dùng một loại thiết bị lưu trữ chuyên dụng trong hệ thống sao lưu, gọi là thư viện băng từ ảo (VTL: Virtual Tape Library). Thiết bị VTL bao gồm phần cứng là các ổ đĩa (như một thiết bị lưu trữ thông thường) và phần mềm điều khiển. Phần mềm này sẽ tạo hoạt động giả lập của VTL như một thiết bị băng từ vật lý nhìn từ quan điểm của hệ thống sao lưu phục hồi. Các máy chủ có dữ liệu cần sao lưu, máy chủ backup, phần mềm quản lý sao lưu trên máy chủ backup đều nhìn nhận VTL như một thiết bị băng từ vật lý (nói cách khác, thiết bị VTL

được ảo hóa thành thiết bị băng từ), và do đó quá trình sao lưu dữ liệu vào VTL cũng sẽ xảy ra giống như vào một thiết bị băng từ vật lý thông thường đối với hệ thống.

VTL có đường kết nối trực tiếp đến thiết bị băng từ vật lý thật, cho phép kích hoạt nấc thứ 2 của giải pháp d2d, tức là sao lưu dữ liệu từ VTL sang băng từ. Với sự có mặt của thiết bị VTL, phần mềm quản lý quá trình sao lưu (tức là phần mềm backup) trước hết sẽ sao lưu dữ liệu của các máy chủ và ứng dụng về VTL. Sau đó, phần mềm backup sẽ kích hoạt việc sao chép dữ liệu từ VTL sang các thiết bị băng từ thật (thời điểm kích hoạt do người quản trị quyết định tùy theo chính sách sao lưu). Mặc dù việc chuyển dữ liệu từ VTL sang thiết bị băng từ vật lý cũng có thể được thực hiện bằng phần mềm quản lý của chính VTL, nhưng trong thực tế các hãng sản xuất không khuyến cáo dùng cách này. Lý do là vì phần mềm backup, khi không trực tiếp quản lý quá trình sao lưu từ VTL sang thiết bị băng từ vật lý, sẽ không quản lý được dữ liệu hiện đang được lưu trữ tại đâu (VTL hay thiết bị băng từ vật lý) dẫn đến việc phục hồi dữ liệu khi có nhu cầu sẽ không thể xảy ra một cách suôn sẻ và nhất quán.

Theo dự đoán của IDC, thị phần trên toàn cầu của VTL sẽ tăng trưởng gấp đôi trong giai đoạn 2006-2011, đạt mức doanh thu 1,4 tỷ đô la Mỹ. Nhận thức được thị trường đầy tiềm năng này, hầu hết các hãng sản xuất thiết bị lưu trữ lớn như IBM, HP, EMC đều đầu tư sản xuất các thiết bị VTL phục vụ cho giải pháp sao lưu d2d. Bảng 1 tổng hợp một số dòng sản phẩm VTL có mặt trên thị trường của các hãng lớn.

Hãng sản xuất	Dòng thiết bị VTL
IBM	- Virtualization Engine TS7500, TS7700
HP	-HP 6000 Virtual Library System -HP 1000i Virtual Library System, -HP 300 Virtual Library System EVA gateway -HP StorageWorks D2D Backup System
NetApp	NearStore VTL 300, 700, 1400
EMC	CLARiiON Disk Library (DL210, series DL4000, DL6000)
SUN	- StorageTek Virtual Tape Library (Series 1000, 2000 và 3000) - Sun StorageTek VTL Plus (Series 1000, 2000 và 3000) - StorageTek Virtual Storage Manager (VSM) System

**Bảng 1:** Một số dòng sản phẩm VTL của các hãng lớn

### 3.3 Ưu điểm của phương thức sao lưu d2d

So với phương thức sao lưu bằng băng từ truyền thống, d2d có những ưu điểm nổi trội như sau:

**Tăng hiệu năng và tốc độ của quá trình sao lưu:** hiệu năng và tốc độ của quá trình sao lưu phục hồi được cải thiện một cách đáng kể do những nhân tố dưới đây:

- **Khả năng mở rộng linh hoạt:** Trong phương pháp sao lưu dùng băng từ thông thường, việc tăng hiệu năng và tốc độ cho quá trình sao lưu chỉ có thể thực hiện được bằng cách tăng số đầu đọc của thiết bị băng từ hoặc thay đầu đọc cũ bằng đầu đọc mới có tốc độ cao hơn. Vấn đề này được giải quyết với sự sử dụng các thiết bị băng từ ảo (VTL) trong giải pháp sao lưu d2d, cho phép tạo ra một cách linh hoạt nhiều đầu đọc ảo.

- **Thông lượng truyền tải tốt và thời gian phục hồi dữ liệu nhanh:** Các ổ đĩa SATA dùng trong giải pháp d2d có tốc độ nhanh. Nguyên lý truy cập ngẫu nhiên (random access) tới dữ liệu của ổ đĩa cứng, thay vì cách truy cập theo thứ tự (sequential access) trong nguyên lý hoạt động của băng từ cho phép dữ liệu được ghi vào (trong quá trình sao lưu) và đọc ra (trong quá trình phục hồi) với tốc độ của đĩa cứng (disk speed), mà không cần tốn thời gian lắp đặt băng từ và quay băng từ đến đúng vị trí lưu trữ dữ liệu cần được phục hồi như trong giải pháp truyền thống disk-to-tape. Điều này dẫn đến tốc độ phục hồi dữ liệu khi sử dụng giải pháp d2d nhanh hơn hẳn so với dùng băng từ thông thường.

**Tăng độ sẵn sàng cho giải pháp:** lỗi phổ biến thường gặp trong quá trình sao lưu dự phòng là các sự cố liên quan đến đầu từ (tape drive) và chính các băng từ. Các lỗi này sẽ không còn xuất hiện trong giải pháp d2d, đơn giản là vì thiết bị băng từ vật lý không còn được dùng cho bước sao lưu đầu tiên (nhắc lại là giải pháp d2d gồm 2 bước: trước hết sao lưu dữ liệu từ đĩa cứng sang đĩa cứng và sau đó sao lưu từ đĩa cứng sang băng từ. Nhưng bước thứ hai xảy ra theo định kỳ và không thường xuyên, do đó các lỗi về thiết bị băng từ sẽ ít gặp hơn). Ngoài ra vì có áp dụng các phương pháp RAID cho các ổ đĩa SATA, nơi chứa dữ liệu sau bước sao lưu thứ nhất của giải pháp d2d, nên giải pháp này có tính bảo vệ cao cho dữ liệu.

**Chi phí hợp lý, tạo mức thu hồi vốn đầu tư (Return of Investment) xứng đáng:** Công nghệ lưu trữ đã và đang đi qua những bước phát triển nhanh chóng. Nhiều công nghệ đĩa cứng mới được thử nghiệm và đưa vào sản xuất hàng loạt trên thị trường. Trong đó các ổ đĩa cứng SATA (dùng trong các thiết bị tủ đĩa để chứa dữ liệu sau bước sao lưu thứ nhất của giải pháp d2d) với dung lượng ngày càng lớn (tối đa 750GB tại thời điểm bài viết) đã không còn quá đắt. Điều này cho phép sử dụng các ổ đĩa cứng làm môi trường chứa các dữ liệu được sao lưu với chi phí đầu tư hợp lý. Bên cạnh đó, với việc mang lại hiệu năng hoạt động cao hơn hẳn so với băng từ vật lý cho quá trình phục hồi dữ liệu, đầu tư xây dựng hệ thống sao lưu d2d là lựa chọn hoàn toàn xứng đáng.

## 4 LỜI KẾT

Có thể nhận thấy lộ trình phát triển của quá trình sao lưu phục hồi dữ liệu trong các hệ thống IT là tiến tới những giải pháp có môi trường lưu trữ dữ liệu có độ an toàn cao (tương đương hoặc tốt hơn khi dùng băng từ), tốc độ phục hồi dữ liệu nhanh (tương đương hoặc tốt hơn khi phục hồi dữ liệu từ đĩa cứng) và giá thành đầu tư vận hành ngày càng giảm. Trong lộ trình như vậy, hiện nay giải pháp sao lưu phục hồi truyền thống từ đĩa cứng sang băng từ (disk-to-tape) vẫn đang được triển khai rộng rãi. Bên cạnh đó, giải pháp sao lưu 2 nấc (disk-to-disk), từ đĩa cứng sang đĩa cứng, sau đó từ đĩa cứng sang băng từ cũng bắt đầu được đẩy mạnh và áp dụng rộng rãi. Với sự phát triển nhanh chóng của lĩnh vực công nghệ lưu trữ, chắc chắn sẽ còn nhiều bước tiến công nghệ được phát minh và ứng dụng, giúp cho các giải pháp sao lưu phục hồi dữ liệu ngày càng hoàn thiện hơn.