Tìm hiểu về RAID

RAID được định nghĩa như thế nào? Trước hết RAID là viết tắt của Redundant Array of Inexpensive Disks (Hệ thống đĩa dự phòng). Đây là hệ thống hoạt động bằng cách kết nối một dãy các ổ cứng có chi phí thấp lại với nhau để hình thành một thiết bị nhớ đơn có dung lượng lớn hỗ trợ hiệu quả cao và đáng tin cậy hơn so với các giải pháp trước đây. RAID được sử dụng và triển khai thành phương pháp lưu trữ trong doanh nghiệp và các máy chủ, nhưng trong 5 năm sau đó RAID đã trở nên phổ biến đối với mọi người dùng.

Lợi thế của RAID

Có 3 lý do chính để áp dụng RAID:

+ Dự phòng

+ Hiệu quả cao

+ Giá thành thấp

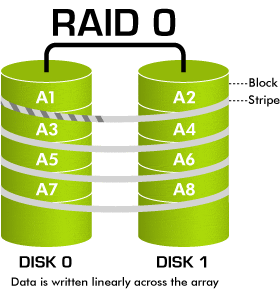
Sự dự phòng là nhân tố quan trọng nhất trong quá trình phát triển RAID cho môi trường máy chủ. Dự phòng cho phép sao lưu dữ liệu bộ nhớ khi gặp sự cố. Nếu một ổ cứng trong dãy bị trục trặc thì nó có thể hoán đổi sang ổ cứng khác mà không cần tắt cả hệ thống hoặc có thể sử dụng ổ cứng dự phòng. Phương pháp dự phòng phụ thuộc vào phiên bản RAID được sử dụng.

Khi áp dụng các phiên bản RAID mạnh bạn có thể thấy rõ hiệu quả tăng cao của nó. Hiệu quả cũng tùy thuộc vào số lượng ổ cứng được liên kết với nhau và các mạch điều khiển.

Tất cả các nhà quản lý những tập đoàn CNTT đều muốn giảm giá thành. Khi chuẩn RAID ra đời, giá thành là một vấn đề chủ chốt. Mục tiêu của các dãy RAID là cung cấp bộ nhớ tốt hơn cho hệ thống so với việc sử dụng riêng biệt các ổ đĩa có dung lượng lớn.

Về phân loại thì có khả nhiều loại Raid được sử dụng như Raid 0, Raid 1, Raid 3, Raid 4, Raid 5, Raid 10 .v.v… Trong bài viết tìm hiểu về Raid ngày hôm nay Vinahost xin được giới thiệu bốn loại Raid chính phổ biến và thường được các khách hàng của Vinahost sử dụng là Raid 0, Raid 1, Raid 10 và Raid 5.

1. RAID 0:



Raid 0 là loại Raid khá phổ biến và được nhiều người sử dụng hiện nay do có khả năng nâng cao hiệu suất tốc độc đọc ghi trao đổi dữ liệu của ổ cứng. Để tiến hành setup Raid 0 thì server cần tối thiểu 2 ổ đĩa (Disk 0, Disk 1).

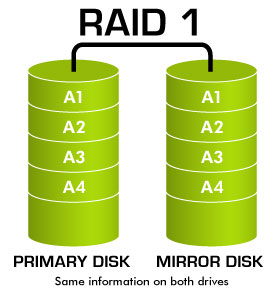
Raid 0 sẽ lưu trữ như sau. Giả sử bạn có 1 file A dung lượng 100MB. Khi tiến hành lưu trữ thay vì file A sẽ được lưu vào 1 ổ cứng duy nhất, Raid 0 sẽ giúp lưu vào 2 ổ đĩa disk 0, disk 1 mỗi ổ 50MB (Striping) giúp giảm thời gian đọc ghi xuống 1 nửa so với lý thuyết .

Ưu điểm: Tốc độ đọc ghi nhanh (gấp đôi bình thường theo lý thuyết).

Nhược điểm: tiềm ẩn rủi ro về dữ liệu. Lý do dữ liệu được chia đôi lưu trên 2 ổ đĩa.Trường hợp 1 trong 2 ổ đĩa bị hỏng thì nguy cơ mất dữ liệu rất cao. Về ổ cứng yêu cầu phải 2 ổ cùng dung lượng, nếu 2 ổ khác dung lượng thì lấy ổ thấp nhất.

Đối tượng sử dụng: Thích hợp với những dịch vụ cần lưu trữ và truy xuất với tốc độ cao. Chẳng hạn như dịch vụ video streaming, chạy cơ sở dữ liệu... (đối với hình thức lưu trữ này Vinahost khuyến cáo khách hàng nên có hình thức sao lưu backup của mình.)

2. RAID 1:



Raid 1 là loại Raid cơ bản được sử dụng khá nhiều hiện nay do khả năng đạt an toàn về dữ liệu. để tiến hành setup Raid 1 thì cũng giống như Raid 0, server cần tối thiểu 2 ổ cứng để lưu trữ.

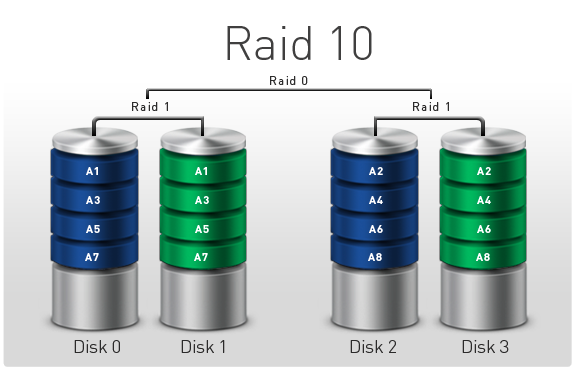
Không giống như Raid 0, Raid 1 đảm bảo an toàn hơn về dữ liệu do dữ liệu được ghi vào 2 ổ giống hệt nhau (Mirroring).

Ưu điểm: An toàn về dữ liệu, trường hợp 1 trong 2 ổ đĩa bị hỏng thì dữ liệu vẫn có khả năng đáp ứng dịch vụ.

Nhược điểm: Hiệu suất không cao, Nâng cao chi phí (giả sử khách hàng sử dụng 2 ổ cứng 500GB. Khi sử dụng Raid 1 thì dung lượng lưu trữ có thể sử dụng chỉ được 500GB). Về ổ cứng yêu cầu phải 2 ổ cùng dung lượng, nếu 2 ổ khác dung lượng thì lấy ổ thấp nhất.

Đối tượng sử dụng: Các dịch vụ lưu trữ, các website vừa và nhỏ không yêu cầu quá cao về tốc độ đọc ghi (in/out) của ổ cứng. Các đối tượng yêu cầu sự an toàn về dữ liệu như các dịch vụ kế toán,lưu trữ thông tin khách hàng, bất động sản v.v…

3. RAID 10:



Raid 10 là sự kết hợp giữa 2 loại raid phổ biến và Raid 1 và Raid 0. Để setup Raid 10 khách hàng cần sử dụng tối thiểu 4 ổ cứng (Disk 0, Disk 1, Disk 2, Disk 3).

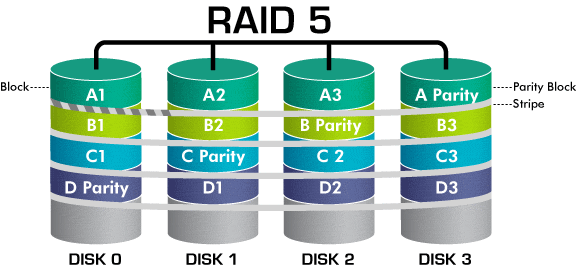
Đối với Raid 10 dữ liệu sẽ được lưu đồng thời vào 4 ổ cứng. 2 ổ dạng Striping (Raid 0) và 2 ổ (Mirroring) Raid 1.

Ưu điểm: Đây là 1 hình thức lưu trữ nhanh nhẹn và an toàn, vừa nâng cao hiệu suất mà lại đảm bảo dữ liệu không bị thất thoát khi 1 trong số 4 ổ cứng bị hỏng.

Nhược điểm: Chi phí cao. Đối với Raid 10 dung lượng sẵn sàng sử dụng chỉ bằng ½ dung lượng của 4 ổ. (giống như raid 1).

Đối tượng sử dụng: Raid 10 thích hợp với tất cả các đối tượng sử dụng (từ những yêu cầu về hiệu suất đến việc đảm bảo an toàn dữ liệu). Về ổ cứng yêu cầu phải 4 ổ cùng dung lượng, nếu 4 ổ khác dung lượng thì lấy ổ thấp nhất.

4. RAID 5:



Raid 5 cũng là một loại Raid được phổ biến khá rộng rãi. Nguyên tắc cơ bản của Raid 5 cũng gần giống với 2 loại raid lưu trữ truyền thống là Raid 1 và Raid 0. Tức là cũng có tách ra lưu trữ các ổ cứng riêng biệt và vẫn có phương án dự phòng khi có sự cố phát sinh đối với 1 ổ cứng bất kì trong cụm.

Để setup Raid 5 ta cần tối thiểu 3 ổ cứng. Theo như hình minh họa phương án lưu trữ của Raid 5 như sau. Giả sử có 1 file A thì khi lưu trữ sẽ tách ra 3 phần A1, A2, A3. Ba phần nãy sẽ tương ứng lưu trên ổ đĩa Disk 0, Disk 1, Disk 2, còn ổ đĩa Disk 3 sẽ giữ bản sao lưu backup của 3 phần này. Tương tự các file sau cũng vậy và tùy theo tiến trình thực hiện mà bản sao lưu có thể được lưu ở bất kì 1 trong những ổ trong cụm Raid.

Ưu điểm: Nâng cao hiệu suất, an toàn dữ liệu, tiết kiệm chi phí hơn so với hình thức lưu trữ Raid 10.

Nhược điểm: Chi phí phát sinh thêm 1 ổ so với hình thức lưu trữ thông thường. (tổng dung lượng ổ cứng sau cùng sẽ bằng tổng dung lượng đĩa sử dụng trừ đi 1 ổ. Giả sử bạn có 4 ổ 500GB thì dung lượng sử dụng sau cùng khi triển khai Raid 5 bạn chỉ còn 1500GB).

Đối tượng sử dụng: Tất cả những website, dịch vụ, ứng dụng có số lượng truy cập và yêu cầu tài nguyên từ nhỏ đến vừa và lớn.

5. JBOD

JBOD (Just a Bunch Of Disks) thực tế không phải là một dạng RAID chính thống, nhưng lại có một số đặc điểm liên quan tới RAID và được đa số các thiết bị điều khiển RAID hỗ trợ. JBOD cho phép bạn gắn bao nhiêu ổ đĩa tùy thích vào bộ điều khiển RAID của mình (dĩ nhiên là trong giới hạn cổng cho phép). Sau đó chúng sẽ được “tổng hợp” lại thành một đĩa cứng lớn hơn cho hệ thống sử dụng. Ví dụ bạn cắm vào đó các ổ 10GB, 20GB, 30GB thì thông qua bộ điều khiển RAID có hỗ trợ JBOD, máy tính sẽ nhận ra một ổ đĩa 60GB. Tuy nhiên, lưu ý là JBOD không hề đem lại bất cứ một giá trị phụ trội nào khác: không cải thiện về hiệu năng, không mang lại giải pháp an toàn dữ liệu, chỉ là kết nối và tổng hợp dung lượng mà thôi.

Raid cứng, raid mềm

1. Raid cứng

RAID cứng thiết lập mảng đĩa cho hệ điều hành sẵn trước khi cài đặt hệ điều hành

RAID cứng: khi hệ điều hành sử dụng không tốn tài nguyên cho việc quản lý đĩa... liên quan đến RAID

RAID cứng chỉ hỗ trợ một định dạng ổ cứng hay khi thiết lập yêu cầu phần cứng khắt khe hơn và không thực hiện được với các ổ cứng bình thường như ATA

Ưu điểm: Độ ổn định cao, có ưu điểm là đẩy cao tốc độ của ổ cứng và có thể thay thế khi card raid hư vì nó không tích vào vào mainboard mà nó được cắm vào cổng PCI Express, nhưng vẫn đảm bảo được data khi chỉ hư card raid mà ổ cứng chứa dữ liệu vẫn còng nguyên không bị hư hại, bên cạnh đó còn có trình điều khiển và phần mềm quản lý, và phục hồi raid.

Nhược điểm: Giá thành cao, không sử dụng được cho các máy tính phổ thông (ví dụ: 1 pc bình thường)

2. Raid mềm

RAID mềm cài đặt hệ điều hành rồi thiết lập RAID

RAID mềm: Phụ thuộc vào hệ điều hành nên bị tốn một phần tài nguyên hệ thống cho việc quản lý RAID

RAID mềm thực hiện được với nhiều loại ổ cứng khác nhau, dung lượng khác nhau... Và trên một đĩa vật lý có thể có những partition với RAID 0, có những partition với RAID 1, RAID 5.

Ưu điểm: Sử dụng cho các máy tính sử dụng các ổ cứng phổ thông nhưng vẫn nâng cao được hiệu năng cho máy tính.

Nhược điểm: Raid mềm chỉ là tập hợp con các tính năng của raid cứng, và nói về độ hư hại của raid mềm là rất lớn không có khả năng phuc hồi, vì raid mềm tích hợp trên mainboard, khi hư là hư main.

**Phân loại RAID**

Xây dựng hệ thống RAID cho Server là một trong những công việc rất quan trọng trong việc quản trị các hệ thống mạng máy tính. Tùy vào quy mô và ứng dụng của mỗi Server mà ta xây dựng hệ thống Raid cho phù hợp và ổn định dựa trên các yếu tố tốc độ, an toàn dữ liệu, tính linh hoạt trong thay thế bảo trì....

Để xây dựng được một hệ thống RAID hoàn chỉnh, việc đầu tiên, ta phải phân biệt được các loại RAID.

Trong bài viết trước Viện đã giải thích thế nào là RAID. Trong bài viết này sẽ đi kỹ hơn vào phân loại của RAID.

RAID được chia thành 2 loại chính : RAID cứng và RAID mềm.

RAID cứng được hiểu RAID dùng phần cứng để điều khiển và ta có thể can thiệp vào giao diện phần mềm của card RAID để tạo các mảng RAID mà không phải dùng bất kỳ phần mềm nào như Windows , Linux ,......

RAID Onboard cũng là một loại RAID cứng nếu nó dùng chip điều khiển RAID (RAID CPU) được tích hợp sẵn trên mainboard, đây là một tùy chọn rẻ tiền kiểu add- in . RAID onboard cũng có những khả năng như card RAID, tuy nhiên, nếu muốn dùng nhiều tính năng nâng cao, mở rộng, khai thác được mọi ưu điểm của công nghệ RAID cung cấp chia mảng cực lớn hàng trăm TB, vận hành tốt hơn thì những card RAID chuyên dụng là sự lựa chọn hàng đầu.

Trong một vài trường hợ, những card RAID không sử đụng card mà dùng Cache memory trên hệ thống, vẫn được gọi là RAID cứng.

Nhìn chung một RAID cứng phải hội tụ các yếu tố sau :

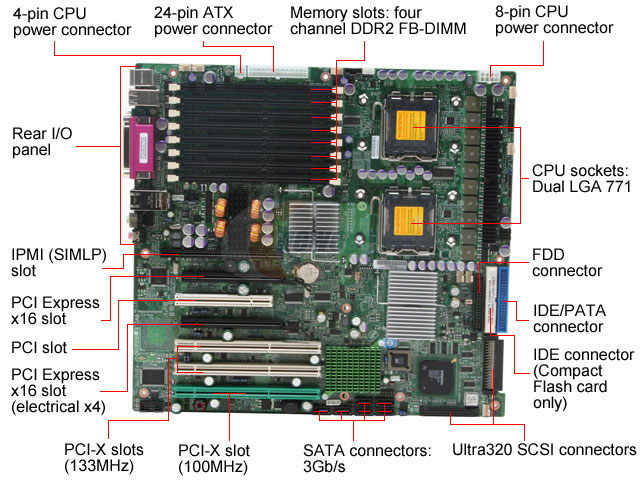
- On-board I/O cho bộ vi xử lý (CPU).

- On-board XOR off-load engine.

- Cache memory.

RAID Mềm: là các trường hợp còn lại của RAID cứng, để cho dễ hình dung "nếu Card RAID đó hoặc mainboard không có CPU riêng để xử lý RAID thì nó là RAID mềm".

Giải thích cho câu nói trên là chúng ta có những Card RAID gắn rời, hoặc hỗ trợ trên mainboard nhưng lấy CPU và RAM của PC/Server đó để điều khiển hệ thống thì vẫn gọi là RAID mềm.



**Phân loại Raid**

Hiện tại RAID có 3 loại chính như sau: Software RAID (RAID mềm), Fake RAID hoặc HostRAID, và Hardware RAID (RAID cứng).

**1. Software RAID**

Sử dụng hệ điều hành Windows hoặc Linux để tiến hành cài đặt RAID. Trên Windows có thể sử dụng Windows Based Raid - Nằm trong mục Disk Management. Trên Linux sử dụng mdadm utility. Nó thực hiện tất cả các lệnh I / O và các thuật toán toán học RAID chuyên sâu trực tiếp trên các CPU của máy chủ lưu trữ.

Ưu điểm: Không mất tiền mua card RAID, chạy sẵn ứng dụng có sẵn trong Windows và Linux

Nhược điểm: Sử dụng CPU và RAM để xử lý luồng dữ liệu, ảnh hưởng đến tốc độ truy xuất và xử lý của hệ thống. Dễ mất dữ liệu khi hệ điều hành bị lỗi !

Đối tượng sử dụng: Khách hàng sử dụng cho gia đình, văn phòng nhỏ hoặc lưu trữ dữ liệu ở mức độ bình thường nhất.

**2. Fake RAID / Host RAID**

Những mainboard có tính năng RAID trong Bios, thường thì bạn phải kích hoạt chức năng này lên và sẽ có ngay cấu hình RAID ngoài BIOS, các ổ cứng được thiết lập được cắm từ các cổng SATA hoặc cổng lưu trữ khác (M.2 tùy vào mainboard có hỗ trợ hay không) đây là hình thức hardware Raid. Sau khi chúng ta cài hệ điều hành lên thì việc quản lý RAID này sẽ hoàn toàn phụ thuộc vào hệ điều hành đây là hình thức software RAID. Nó sử dụng firmware nhận diện/ đánh giá ổ cứng trước khi HĐH được khởi động. Và sau khi hệ điều hành khởi động xong (lấy quyền kiểm soát lại từ Bios) thì khi đó nó giao quyền điều khiển RAID cho OS. Vì vậy nó có tên gọi là Fake RAID hay còn gọi là Host RAID. Hình thức này vẫn sử dụng CPU giống như Software RAID.

Fake RAID được gắn với các giải pháp phần cứng, nhưng không có bộ nhớ cache DRAM trên bo mạch chủ mà cho phép bộ nhớ cache ghi lại dữ liệu trên bộ nhớ RAM của hệ thống. Đây cũng là ưu điểm và nhược điểm của giải pháp này. Trường hợp máy chủ bị mất điện đột ngột, dữ liệu lưu trữ trong bộ nhớ chính sẽ bị mất. Do đó những máy này buộc phải có giải pháp điện backup qua UPS hoặc máy phát điện tức thì.



Ưu điểm: Không mất tiền mua card RAID, Không lo hệ điều hành lỗi là mất dữ liệu vì những config này đã được lưu trong Firmware của mainboard.

Nhược điểm: Mainboard lỗi là mất dữ liệu config RAID phải tìm một mainboard tương tự. Mất điện hệ thống cũng dẫn đến việc bảo tồn toàn vẹn dữ liệu, mất an toàn dữ liệu.

Đối tượng sử dụng: Khách hàng sử dụng cho gia đình, tăng tốc cho một số ứng dụng nhờ tính năng RAID như: Ứng dụng làm việc, Game... Các máy chủ nhỏ cần yêu cầu RAID backup dữ liệu (Raid 1, Raid 0+1...)

3. Hardware RAID

Thành phần của Hardware RAID như một máy tính thu nhỏ với Bios riêng, giao diện quản lý riêng. Thành phần bên trong bao gồm chip xử lý (CPU), RAM (RAM ECC), nguồn cấp điện (Pin lưu trữ), Cổng kết nối SATA, SAS..., Chân kết nối giao tiếp với các cổng PCI / PCI-E / PCI-X.

Thường dưới dạng card Add-in được cắm trên các khe PCI / PCI-E / PCI-X. Các card Raid Controller ngày nay thường sử dụng khe cắm PCI-E, sử dụng băng thông tốc độ cao này để truyền dữ liệu cao nhất. Các card Raid Controller này giảm tải hệ thống máy chủ trong một số hoặc tất cả các lệnh I / O, dành các hoạt động tính toán RAID cho một hoặc nhiều bộ vi xử lý thứ cấp mà nó có. Góp phần giảm phụ thuộc từ CPU, nâng tối đa hiệu năng của hệ thống.

Ngoài việc cung cấp những lợi ích chịu lỗi của một RAID thông thường , bộ điều khiển hardware RAID còn thực hiện các chức năng kết nối tương tự như bộ điều khiển trên máy chủ tiêu chuẩn. Và cũng bởi nhờ nó có riêng cho mình tài nguyên (CPU, memory,...) , nên chúng thường cung cấp hiệu suất cao nhất cho tất cả các loại RAID. Hardware RAID cũng cung cấp tính năng chịu lỗi mạnh mẽ hơn đa dạng hơn software RAID. Ví dụ như RAID 0/1/5/6/10/50/60.



Tùy vào lựa chọn của khách hàng mà có giải pháp RAID dành riêng. Ví dụ như các cấp độ RAID 1 hoặc RAID 10 hoạt động khá tốt khi không cần 1 Raid Controller đắt tiền, nhưng với giải pháp RAID 5 và 6 nên sử dụng card RAID có hỗ trợ Pin.

Ưu điểm: Tối ưu tốc độ truyền dữ liệu, tính ổn định và toàn vẹn dữ liệu cao nhất, giảm tải cho hệ thống xử lý của máy tính.

Nhược điểm: Giá thành cao

Đối tượng sử dụng: Doanh nghiệp vừa và nhỏ, các doanh nghiệp lớn có dữ liệu lớn và cần mức độ an toàn dữ liệu cũng như tốc độ tối đa