**Tính sẵn sàng (khả dụng ) cao là gì?**

**Mục lục**

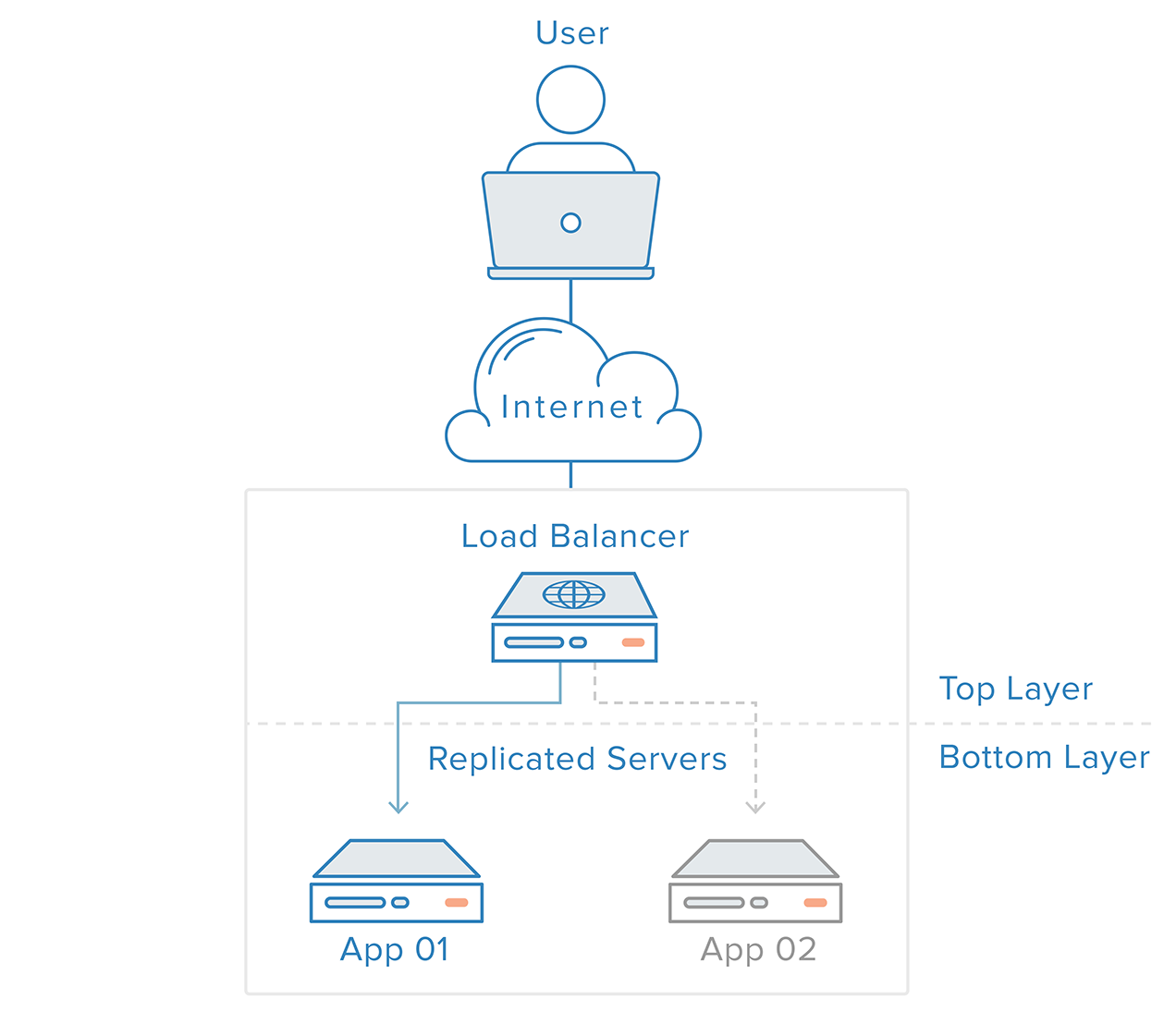
* [1. Giới thiệu](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#introduce)
* [2. Tính sẵn sàng là gì?](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#whatis)
* [3. Định lượng tính sẵn sàng](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#measuring)
* [4. Tính sẵn sàng cao hoạt động như thế nào?](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#how-work)
* [5. Tính sẵn sàng cao quan trọng khi nào?](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#important)
* [6. Điều gì tạo nên một hệ thống sẵn sàng cao?](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#what-make)
* [7. Những thành phần cần thiết nào cần cho một hệ thống đảm bảo tính sẵn sàng cao?](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#what-need)
* [8. Phần mềm nào có thể được sử dụng để cấu hình tính sẵn sàng cao cho hệ thống?](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#what-pro)
* [9. Tổng kết](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#conclusion)
* [Các nội dung khác](https://github.com/dule07/thuctap012017/blob/master/TVBO/docs/Posts_Translated/what-is-ha.md#content-others)

**Nội dung**

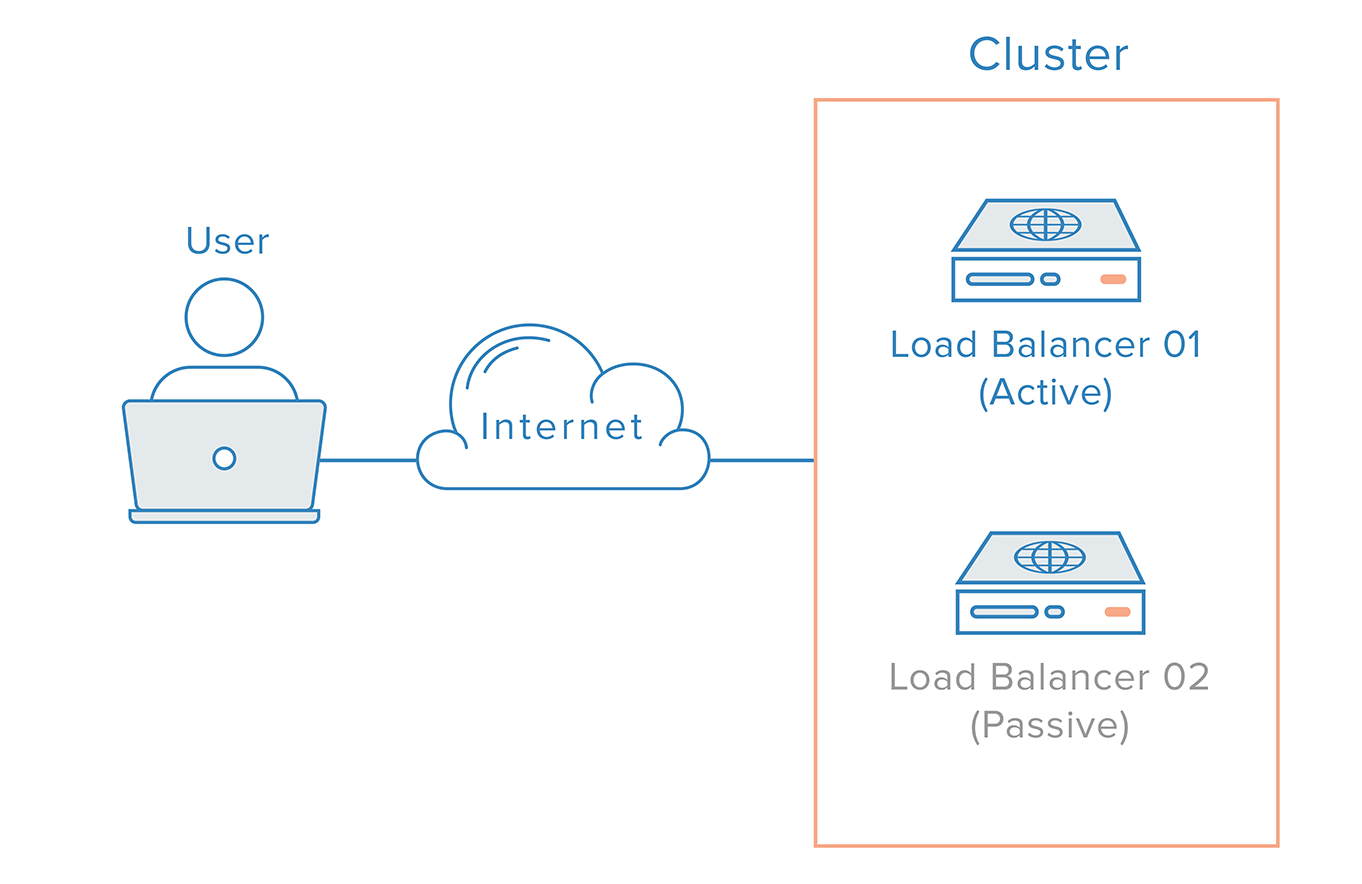
* **1. Giới thiệu**
  + Với nhu cầu tăng cho việc thiết kế các cơ sở hạ tầng đáng tin cậy với hiệu năng cao để phục vụ cho các hệ thống quan trọng, các thuật ngữ về khả năng mở rộng (Scalability ) và mức độ khả dụng cao hay tính sẵn sàng cao (High Availability ) không thế không phổ biến hơn bao giờ hết. Trong việc sử lý tải hệ thống tăng đang là một vấn đề quan tâm chung, việc giảm thời gian thời gian ngừng hoạt động (downtime) và loại bỏ trường hợp "single points of failure" ([SPOF](https://www.google.com/search?q=what+is+single+points&ie=utf-8&oe=utf-8)) cũng đang dần trở lên hết sức quan trọng. Tính sẵn sàng cao được xem là chất lượng cho việc thiết kế cơ sở hạ tầng theo quy mô nhằm giải quyết các điều cần cân nhắc đã và đang tồn tại trong các hệ thống.
  + Trong bài viết này, chúng ta sẽ cùng thảo luận những gì chính xác về tính sẵn sàng cao và cách mà nó có thể làm để cải thiện mức độ độ tin cậy trong cơ sở hạ tầng hệ thống của bạn.
* **2. Tính sẵn sàng là gì?**
  + Trong tính toán, cụm từ sẵn sàng (Availability ) được sử dụng để mô tả cho khoảng thời khi mà một dịch vụ có thể sử dụng được cũng giống như khoảng thời gian yêu cầu của hệ thống để có thể đáp ứng đối với những yêu cầu được tạo ra bởi người dùng.
  + Tính sẵn sàng cao (High Availability ) là chất lượng của một hệ thống hoặc một bộ phận đảm bảo mức độ hoạt động cao trong khoảng thời gian nhất định.
* **3. Định lượng tính sẵn sàng**
  + Tính sẵn sàng thường được biểu diễn bằng chỉ số phần trăm (%) cho biết thời gian hoạt động dự kiến cho một hệ thống cụ thể hay một bộ phận của hệ thống trong một khoảng thời gian cụ thể là bao nhiêu. Nếu giá trị là 100% điều này có nghĩa là hệ thống luôn luôn sẵn sàng và không bao giờ xảy ra tình trạng ngừng hoạt động. Ví dụ: Một hệ thống có tính tính sẵn sàng là 99% trong một năm. Điều này có nghĩa là hệ thống có thể có thời gian ngừng hoạt động là 3,65 ngày.
  + Các giá trị này được tính dựa trên một số yếu tố bao gồm cả thời gian bảo trì theo kế hoạch và không theo lịch cũng như thời gian để phục hồi một lỗi hệ thống có thể xảy ra.
* **4. Tính sẵn sàng cao hoạt động như thế nào?**
  + Tính năng sẵn sàng cao hoạt động như một cô chế để phản hồi lại việc thất bại của một hạ tầng hệ thống. Cách thức mà nó hoạt động khá đơn giản nhưng thường đòi hỏi một số phần mềm và cấu hình chuyên biệt.
* **5. Tính sẵn sàng cao quan trọng khi nào?**
  + Khi các thiết lập hệ thống hoạt động một cách mạnh mẽ, giảm thiểu thời gian ngừng hoạt động và thời gian gián đoạn dịch vụ thường được xem là những ưu tiên cao. Bất kể hệ thống và phần mềm của bạn có đáng tin cậy hay không, các vẫn đề có thể xảy ra có thể làm gián đoạn các ứng dụng hoặc các máy chủ của bạn.
  + Triển khai tính sẵn sàng cao cho cơ sở hạ tầng của bạn là một chiến lược hữu ích để giảm tác động của các loại sự cố. Các hệ thống sẵn sàng cao có thể phục hồi các hệ thống hoặc các thành phần của hệ thống bị hư hỏng một cách tự động.
* **6. Điều gì tạo nên một hệ thống sẵn sàng cao?**
  + Một trong những mục tiêu của tính sẵn sàng cao cho hệ thống là để loại bỏ những [SPOF](https://www.google.com/search?q=what+is+single+points&ie=utf-8&oe=utf-8) trong cơ sở hạ tầng của hệ thống. Một SPOF chính là một thành phần trong các công nghệ mà bạn sử dụng có thể gây ra gián đoạn dịch vụ nếu nó trở nên không có sẵn (dừng hoạt động ). Như vậy, bất kỳ thành phần nào là cần thiết cho các chức năng của các ứng dụng của bạn mà không có dự phòng được xem là một SPOF.
  + Để loại bỏ các SPOF, mỗi lớp của ngăn xếp công nghệ phải được chuẩn bị cho việc dự phòng. Ví dụ, bạn có một cơ sở hạ tầng gồm 2 máy chủ web giống hệt nhau, đằng sau được dự phòng là một cân bằng tải (Load Balancer). Lưu lượng đến từ các clients sẽ được phân bố đều đến các máy chủ web. Nhưng nếu một trong các máy chủ web bị dừng hoạt động, cân bằng tải sẽ thực hiện chuyển hướng các lưu lượng một cách trực tiếp đến máy chủ web còn lại.
    - Lớp máy chủ web trong ví dụ này không phải là một SPOF bởi vì:
      * Có các thành phần dự phòng cho máy chủ web được đặt ra.
      * Cơ chế ở phía trên của lớp này (bộ cân bằng tải) có thể phát hiện các hỏng hóc trong các thành phần và thích nghi với hành vi của nó để phục hồi kịp thời

nhưng chuyện gì sẽ xảy ra khi mà bộ cân bằng tải dừng hoạt động?

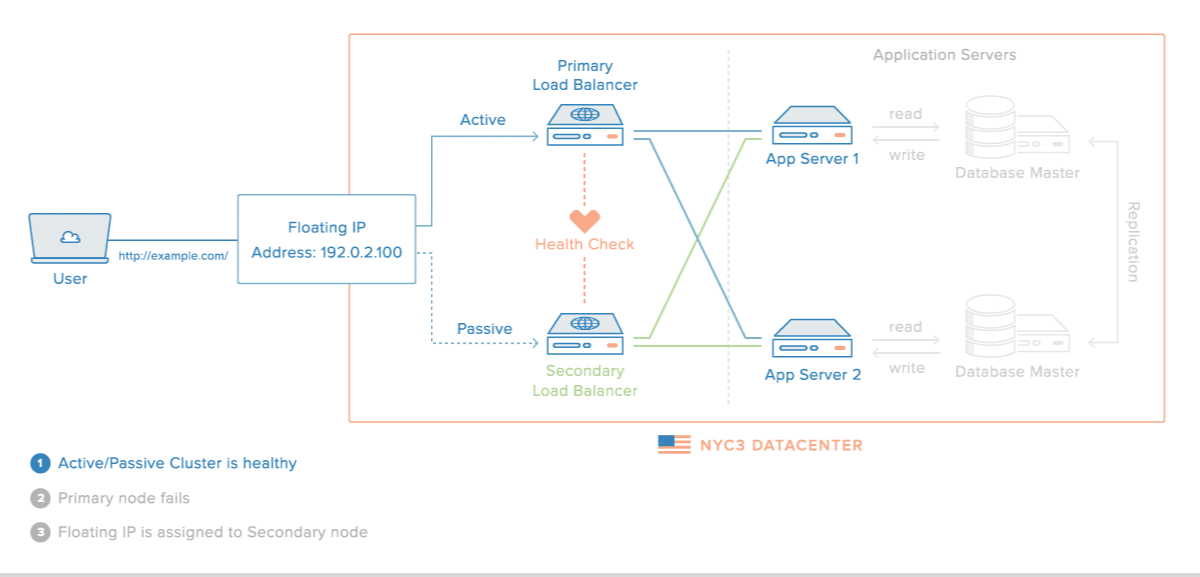
* + - Điều này không phải là không phổ biến trong thực tế, lớp cân bằng tải vốn dĩ bản thân nó là một SPOF vì nó không có dự phòng. Loại bỏ SPOF này có thể là một thách thức. Tuy nhiên bạn có thể dễ dàng cấu hình bổ sung để thiết lập dự phòng cho cân bằng tải, nhưng nó lại không có máy chủ nào thực hiện phát hiện và khắc phục lỗi.
    - Sự dự phòng không thể đảm bảo cho tính sẵn sàng cao. Phải có một cơ chế để phát hiện lỗi và thực hiện hành động khi một trong các thành phần của ngăn xếp hạ tầng trở lên không khả dụng.
    - Việc phát hiện và phục hồi lỗi hỏng hóc cho các hệ thống dự phòng có thể được thực hiện bằn cách sử dụng phương pháp tiếp cận từ trên xuống dưới. Lớp trên cùng sẽ có trách nhiệm theo dõi các lớp ngay bên dưới. Như trong ví dụ trên, trình cân bằng tải chính là trên cùng và các máy chủ web là lớp dưới so với trình cân bằng tải. Khi mà một máy chủ web không khả dụng, cân bằng tải sẽ khắc phục bằng việc ngừng chuyển hướng các request từ clients đến máy chủ web bị lỗi cụ thể đó.

[](https://camo.githubusercontent.com/9d661b4de1f5863bb1793ae9fa5a08db25bc4826/68747470733a2f2f6173736574732e6469676974616c6f6365616e2e636f6d2f61727469636c65732f686967682d617661696c6162696c6974792f4469616772616d5f322e706e67)

* + - Cách tiếp cận này có xu hướng đơn giản hơn nhưng nó có những hạn chế: sẽ có một điểm trong cơ sở hạ tầng của bạn, nơi mà ở lớp trên cùng không tồn tại hoặc ngoài tầm liên lạc giữa các máy chủ - đó là trường hợp với lớp cân bằng tải. Việc tạo ra một dịch vụ phát hiện sự hỏng hóc cho cân bằng tải tại một máy chủ bên ngoài hạ tầng đơn giản sẽ tạo ra một SPOF.
    - Với trường hợp như vậy, việc tiếp cận một cách phân tán là cần thiết. Nhiều máy chủ dự phòng phải được kết nối với nhau tại cùng một nơi và mỗi máy chủ nên có khả năng phát hiện và khắc phục các lỗi.

[](https://camo.githubusercontent.com/5412bb2acc4278eca66b063c8005fa05386bb12e/68747470733a2f2f6173736574732e6469676974616c6f6365616e2e636f6d2f61727469636c65732f686967682d617661696c6162696c6974792f4469616772616d5f312e706e67)

* + - Đối với trường hợp cân bằng tải, tuy nhiên sẽ có một sự phức tạp thêm bởi cách làm việc của nameservers. Việc phục hồi từ một cân bằng tải lỗi thường có nghĩa là chuyển giao lại toàn bộ chức năng mà cân bằng tải lỗi này đang thực hiện sang cho một cân bằng tải dự phòng khác để nó tiếp tục đảm nhiệm, điều này có nghĩa rằng ta phải thay đổi DNS để thực hiện trỏ tên miền tới địa chỉ IP của cân bằng tải dự phòng. Việc thay đổi này có thể mất một khoảng thời gian để bộ cân bằng tải dự phòng xuất hiện trên Internet, điều này có thể dẫn đến hệ thống ngừng hoạt động một cách nghiêm trọng.
    - Một giải pháp có thể sử dụng lúc này đó là sử dụng [cân bằng tải với DNS áp dụng thuật toán round-robin](https://www.digitalocean.com/community/tutorials/how-to-configure-dns-round-robin-load-balancing-for-high-availability). Tuy nhiên, cách tiếp cận này là không đáng tin cậy vì nó để lại việc chuyển đổi dự phòng ứng dụng phía clients.
    - Giải pháp mạnh mẽ hơn và đáng tin cậy hơn là sử dụng các hệ thống cho phép sửa địa chỉ IP một cách linh hoạt giống như ta sử dụng các địa chỉ IP động ([floating IPs](https://www.digitalocean.com/community/tutorials/how-to-use-floating-ips-on-digitalocean)). Việc ánh xạ lại địa chỉ IP sẽ loại bỏ các vấn đề về quảng bá và caching vốn có trong các thay đổi của DNS bằng cách cung cấp một địa chỉ IP tĩnh mà có thể dễ dàng sửa đổi lại khi cần thiết. Tên miền vẫn có thể được liên kết với cùng 1 địa chỉ IP, trong khi địa chỉ IP này có thể linh hoạt di chuyển hay được sử dụng giữa các máy chủ.
    - Đây là cách mà cơ sở hạ tần có tính sẵn sàng cao sử dụng floating IP có thể miêu tả như sau:

[](https://camo.githubusercontent.com/a9a4d0ec6bfdb681b3826696005ad3f6b89c0f31/68747470733a2f2f6173736574732e6469676974616c6f6365616e2e636f6d2f61727469636c65732f686967685f617661696c6162696c6974792f68612d6469616772616d2d616e696d617465642e676966)

* **7. Những thành phần cần thiết nào cần cho một hệ thống đảm bảo tính sẵn sàng cao?**
  + Có một số thành phần phải được xem xét cẩn thận để đảm bảo cho việc xây dựng một hệ thống cơ sở hạ tầng đảm bảo tính sẵn sàng cao trong thực tế. Cao hơn hẳn so với việc thực hiện triển khai các phần mềm. Tính sẵn sàng cao phụ thuộc vào các yếu tố sau:
    - Môi trường: Nếu tất cả các máy chủ của bạn cùng nằm trong cùng một khu vực địa lý, điều kiện môi trường như động đất hoặc lũ lụt có thể làm mất toàn bộ dữ liệu hệ thống của bạn. Việc có các máy chủ dự phòng tại các trung tâm dữ liệu khác nhau và các khu vực địa lý khác nhau sẽ làm tăng độ tin cậy.
    - Phần cứng: Các máy chủ đảm nhiệm chức năng cung cấp tính sẵn sàng cao phải có khả năng phục hồi mạnh mẽ phòng cho việc mất điện hay hỏng hóc về phần cứng bao gồm cả về ổ đĩa cứng cũng như các giao diện mạng (network interfaces )
    - Phần mềm: Toàn bộ các phần mềm, bao gồm cả hệ điều hành và ứng dụng của chính nó phải được chuẩn bị để xử lý sự cố không mong muốn mà có thể sẽ yêu cầu nó phải khởi động lại.
    - Dữ liệu: Mất dữ liệu và sự không thống nhất về dữ liệu giữa các máy chủ có thể được gây ra bởi một số yếu tố, và nó không ngoại trừ việc ổ đĩa lưu trữ dữ liệu bị hỏng hóc. Các hệ thống sẵn sàng cao phải đảm bảo cho tính an toàn của dữ liệu trong trường hợp xảy ra hỏng hóc.
    - Mạng internet: Mất mạng là một vấn đề điển hình xảy ra cho một hệ thống đảm bảo tính sẵn sàng cao. Điều quan trọng ở đây là cần một chiến lược mạng dự phòng để sử dụng nó khi xảy ra các sự cố về mạng.
* **8. Phần mềm nào có thể được sử dụng để cấu hình tính sẵn sàng cao cho hệ thống?**
  + Mỗi tầng của một hệ thống đảm bảo tính sẵn sàng cao sẽ có nhu cầu khác nhau về các phần mềm và cấu hình đi kèm. Tuy nhiên, ở mức độ ứng dụng, cân bằng tải đại diện cho một phần cần thiết của phần mềm để tạo ra bất kỳ thiết lập có tính sẵn sàng cao nào.
  + [HAProxy](https://www.digitalocean.com/community/tutorials/an-introduction-to-haproxy-and-load-balancing-concepts) (High Availability Proxy ) là sự lựa chọn phổ biến cho các loại máy chủ khác nhau, bao gồm cả các máy chủ đảm nhiệm chức năng làm cơ sở dữ liệu ([database servers](https://www.digitalocean.com/community/tutorials/how-to-use-haproxy-to-set-up-mysql-load-balancing--3))
  + Đi lên dần các tầng của một hệ thống, điều quan trọng là phải triển khai các giải pháp dự phòng đáng tin cậy cho các điểm truy cập ứng dụng, thường thì là sử dụng cân bằng tải. Để loại bỏ SPOF này, như đã đề cập ở trên, chúng ta cần phải thực hiện nhóm các cân bằng tải lại vào với nhau và sử dụng một floating IP cho các máy chủ cân bằng tải. Corosync và Pacemaker là những lựa chọn phổ biến cho việc tạo ra một thiết lập như vậy. Chúng đều có thể triển khai trên cả hai hệ điều hành máy chủ [Ubuntu](https://www.digitalocean.com/community/tutorials/how-to-create-a-high-availability-setup-with-corosync-pacemaker-and-floating-ips-on-ubuntu-14-04) và [CentOS](https://www.digitalocean.com/community/tutorials/how-to-create-a-high-availability-setup-with-pacemaker-corosync-and-floating-ips-on-centos-7)
* **9. Tổng kết**
  + Tính sắn sàng cao là một phần quan trọng trong kỹ thuật của độ tin cậy, tập trung vào việc đảm bảo rằng một hệ thống hoặc một bộ phận trong hệ thống có mức hoạt động cao và ổn định trong một khoảng thời gian nhất định. Thoạt nhìn, việc thực hiện nó có vẻ khá phức tạp. Tuy nhiên nó có thể mang lại những lợi ích to lớn cho các hệ thống đòi hỏi về độ tin cậy cao