**Chương I. Giới thiệu đề tài**

Công nghệ mạng và các hệ thống máy chủ ngày nay đã có những bước phát triển vượt bậc. Nhờ đó đáp ứng được các ứng dụng thương mại rộng lớn như các giao dịch tài chính, hệ thống cơ sở dữ liệu, hệ thống web server, file server, các ứng dụng truyền media . Các ứng dụng này có tần suất sử dụng cao, yêu cầu hoạt động liên tục. Do vậy, hệ thống mạng phải có khả năng mở rộng tối ưu để đáp ứng được một lượng lớn các yêu cầu ngày càng tăng của người dùng mà không gây ra bất kỳ một độ trễ không mong muốn nào.

Chúng ta đều biết các máy chủ là trái tim của mạng máy tính nếu máy chủ mạng hỏng , hoạt động của hệ thống sẽ bị ngưng trệ. Điều đáng tiếc là dù các hãng sản xuất đã cố gắng làm mọi cách để nâng cao chất lượng của thiết bị , nhưng những hỏng hóc đối với các thiết bị mạng nói chung và các máy chủ nói riêng là điều không thể tránh khỏi. Do vậy , vấn đề đặt ra là cần có một giải pháp để đảm bảo cho hệ thống vẫn hoạt động tốt ngay cả khi có sự cố xảy ra đối với máy chủ mạng. Việc lựa chọn một máy chủ đơn lẻ có cấu hình cực mạnh để đáp ứng nhu cầu này sẽ kéo theo chi phí đầu tư rất lớn và không giải quyết được các vấn đề đặt ra của các tổ chức. Giải pháp hiệu quả được đưa ra là sử dụng một nhóm máy chủ cùng thực hiện một chức năng dưới sự điều khiển của một công cụ phân phối tải - Giải pháp cân bằng tải. Có rất nhiều hãng đưa ra giải pháp cân bằng tải như Cisco , Coyote Point , Sun Microsystems… với rất nhiều tính năng phong phú. Tuy nhiên , về cơ bản, nguyên tắc cân bằng tải điển hình là RRDNS ( Round Robin DNS). Với giải pháp này , nếu một máy chủ trong nhóm bị lỗi , RRDNS sẽ vẫn tiếp tục gửi tải cho máy chủ đó cho đến khi người quản trị mạng phát hiện ra lỗi và tách máy chủ này ra khỏi danh sách địa chỉ DNS. Điều này sẽ gây ra sự đứt quãng dịch vụ. Sau những phát triển , từ các thuật toán cân bằng tải tĩnh như Round Robin, Weighted Round Robin đến các thuật toán cân bằng tải động như Least Connection , Weighted Least Connection, Optimized Weighted Round Robin và Optimized Weighted Least Connection, kỹ thuật cân bằng tải hiện nay nhờ sự kết hợp các thuật toán trên ngày càng trở nên hoàn thiện mặc dù nhược điểm vốn có như tạo điểm lỗi đơn và vấn đề nút cổ chai do sử dụng bộ điều phối tập trung (centralized disatcher ) vẫn còn. Ngoài khả năng áp dụng với Web server , kỹ thuật này còn có thể áp dụng với các hệ thống server ứng dụng khác. SLB (Cân bằng tải máy chủ ) không chỉ làm nhiệm vụ phân phối tải cho các server mà còn cung cấp cơ chế đảm bảo hệ thống server luôn khả dụng trước các client. SLB không có yêu cầu đặc biệt gì về phần cứng, bất cứ máy tính nào hợp chuẩn đều có thể được sử dụng làm server. Chi phí triển khai nhờ đó giảm đáng kể. Kiến trúc phần mềm phân tán của SLB cho phép cung cấp hiệu năng và tính khả dụng của kỹ thuật này ở mức cao nhất.

Với lý do trên, chúng tôi đã lựa chọn đề tài: “Cân bằng tải cho web server ”. Đây là một vấn đề khá quan trọng trong việc thiết kế các hệ thống mạng hiện nay, các trung tâm tính toán hiệu năng cao và các trung tâm dữ liệu phục vụ cho các ứng dụng lớn.

Đề tài sử dụng Nginx làm web server đồng thời đóng vai trò cân bằng tải và sử dụng Nginx với Keepalived để xây dựng hệ thống dự phòng.

Đề tài này sẽ giúp chúng ta biết được các quá trình cần thiết để có thể cân bằng tải cho một hệ thống mạng , hiểu sâu về khái niệm cũng nh biết cách cài đặt và cấu hình cân bằng tải cho web server sử dụng Nginx và Keepalived

Để hiểu rõ vấn đề chúng tôi đã xây dựng bố cục đề tài gồm có 7 chương :

Chương I . Giới thiệu đề tài

Chương II . Tổng quan về cân bằng tải

Chương III . Tổng quan về Nginx

Chương IV . Hệ thống dự phòng cho web server với Keepalived

Chương V . Cài đặt và cấu hình hệ thống

Chương VI. Demo

Chương VII. Kết luận