

# 리눅스 클러스터링 웹 서버의 고가용성에 대한 연구

## A study on high availability of the linux clustering web server

°박 지 현, 이 상 문, 홍 태 화, 김 학 배

연세대학교 전기·전자공학과(Tel : 82-2-2132-2778 ; Fax : 82-2-362-2780 ; E-mail : jihyune@yonsei.ac.kr)

**Abstract** : As more and more critical commercial applications move on the Internet, providing highly available servers becomes increasingly important. One of the advantages of a clustered system is that it has hardware and software redundancy. High availability can be provided by detecting node or daemon failure and reconfiguring the system appropriately so that the workload can be taken over by the remaining nodes in the cluster. This paper presents how to provide the guaranteeing high availability of clustering web server. The load balancer becomes a single failure point of the whole system. In order to prevent the failure of the load balancer, we setup a backup server using heartbeat, fake, mon, and checkpointing fault-tolerance method. For high availability of file servers in the cluster, we setup coda file system. Coda is a advanced network fault-tolerance distributed file system.

**Keywords** : high availability, linux clustering, fault-tolerant, checkpoint, coda

### 1. 서론

최근 웹을 통한 다양한 서비스로 인터넷 사용자 수가 급격히 증가하고 있으며, 통신기기, 가전/멀티미디어기기 등의 여러 분야에서 인터넷을 이용한 응용프로그램의 개발이 확대되고 있는 추세이다. 또한, 초고속통신망 등의 인터넷관련 기본 인프라 구축되면서 다양한 인터넷 응용 프로그램들이 웹기반으로 통합화되고 기존의 파일서버 및 프린터 서버뿐만 아니라 웹서버와 애플리케이션 서버, 멀티미디어 서비스를 위한 VOD/AOD 서버, 인터넷 기반 사업용 서버 등으로 서버의 영역이 점차 확대되어 나가고 있으며 그 사용 환경이 크게 개선되고 있다. 따라서, 실제로 웹 서비스를 제공하는 서버단에서는 인터넷 사용의 증가로 인해 막대한 양의 트래픽(traffic)이 발생하여 웹 서비스 전반에 걸친 성능의 저하가 우려된다. 즉, 웹 서버의 접속용량 및 작업처리 속도가 서비스의 질 향상에 큰 영향을 줄 수가 있으며, 특히 business-critical 한 경우에는 웹서버의 안정성 및 가용성 문제가 갈수록 가장 중요한 요소로 대두되고 있다[1,8,9].

일반적으로 기존의 단일 웹서버에서는 고가·고성능의 시스템을 구축하여 웹서버에서 발생될 수 있는 이러한 문제들을 해결하고 있으나, 급격한 접속수에 따른 서버의 확장성 및 유연성 등이 부족하다. 이를 위해 최근에는 리눅스에 의해서 일반 PC를 클러스터링하여 클러스터 가상 웹서버를 구축하고 있는 사례가 크게 증가하고 있다. 이러한 방법은 사용자측에서는 마치 하나의 서버에만 접속한 것처럼 투명성(transparency)을 제공하며 서버측에 대해서는 확장성을 보장할 수 있다[1]. 확장성을 목적으로 하는 이러한 클러스터 웹서버의 구조는 증가하는 부하를 모두 처리할 수 있어야 한다. 따라서 부하를 공유하는 독립된 서버들로 구성된 분산구조가 확장성 웹서버를 위해 보다 적절한 방안이 된다[2]. 클러스터링 웹서버의 고가용성을 위해 각 요소의 고장 발생에 대해서 신속한 고

장 진단 및 검출 그리고 적절한 복구에 대한 고려가 필수적이다.

이러한 리눅스 기반의 클러스터 환경에서 동작하는 웹 서버나 대용량의 멀티미디어 서버, 자료 검색 서버 등은 대용량의 파일 입출력의 실시간 처리뿐만 아니라 서비스에 대한 고가용성을 만족시킬 수 있어야 하고, 각 서버는 자신의 로컬 디스크뿐만 아니라 네트워크로 연결된 하나의 가상 서버내의 다른 컴퓨터들에 저장되어 있는 데이터들에 대해서도 효과적으로 대처 할 수 있어야 하기 때문에 이러한 클러스터 환경에 적합한 분산 병렬 파일 시스템이 필수적이라 할 수 있다[6,7].

본 연구에서는 이러한 클러스터링 가상웹서버를 구현하고, 이를 위한 고가용성 및 고장포용기법 그리고 여기에 적용된 고장 진단 기법에 대한 방안을 제시한다. 또한, 진보된 네트워크 파일 시스템인 Coda 분산 파일 시스템에 대해서 그 기본 개념과 동작 원리 및 Coda 시스템이 리눅스 클러스터 환경에서 어떠한 방식으로 고가용성을 보장하는지에 대해서 알아보고, Coda 시스템을 실제 리눅스 클러스터 웹 서버에 적용한 결과에 대해서 살펴보도록 하겠다.

### 2. 클러스터 웹 서버 구조

단일 서버와는 달리 고확장성 및 고가용성을 위해 독립적인 여러 서버들을 그룹화하여 마치 하나의 서버처럼 운영하는 것을 클러스터(cluster)라고 한다[3]. 본 연구에서는 리눅스를 기반으로 하여 그림 1과 같은 클러스터링 웹서버를 구현한다. 클러스터 시스템은 다양한 구조로 구현 될 수 있으나, 일반적으로 부하를 분산해주는 로드밸런서(load balancer)와 그에 딸린 여러 대의 실서버(real server)들로 구성될 수 있다[4,8]. 비록 많은 실서버들로 이루어져 있어도, 인터넷을 통해 웹 서버에 접속하는 사용자에게는 단일 서버 또는 단일 서버이미지를 갖게 하며, 사용자는 클러스터를 하나의 서버인 것처럼 간주하게 된다. 이러한 클러스터링 기술은 저가의 시스템을 이용하여 고가용성 및 고성능의 시스템을 구현하기 위해 널리 사용되고 있다.

본 연구는 2000년 정보통신연구진흥원 산업기술개발사업의 지원 하에 진행되었습니다.