제 1 장 서 론

인터넷 사용 환경이 보편화되면서 최근 웹을 통하여 다양한 서비스를 제공받고자 하는 인터넷 사용자 수가 급격히 증가하고 있으며, 통신기기 뿐만 아니라, 가전 / 멀티미디어 기기 등의 여러 분야에서도 인터넷을 이용한 응용프로그램의 개발이 확대되고 있는 추세이다. 또한, 초고속통신망 등을 통한 인터넷관련 기본 인프라가 폭넓게 확장되어 다양한 인터넷 응용 프로그램들이 웹 기반으로 통합화되어가고 있다. 따라서 이러한 각종 인터넷 응용 서비스들을 제공하기 위해서 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 있다. 이렇게 인터넷 사용인구가 증가하고 그에 따른 대량의 인터넷 서비스 제공을 위해 실제로 웹을 통한 인터넷 서비스를 제공하는 서버단에서는 접속자 증가와 해당 컨텐츠 사용 증가로 인해 막대한 양의 트래픽(traffic)이 발생하여 웹서비스 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. 실제로, 특정 사이트에 사용자가 폭주하여 서비스의 병목현상이 발생하여 웹서버의 처리 불가로 사이트가 다운되는 경우를 자주 접할 수 있다. Contingency Planning Research에서 제공하는 연구에 의하면, 평균적인 시스템의 중단으로 기업체들은 몇 백만 달러의 손해를 입는다고 나와 있다[1]. 따라서 웹 서버의 접속용량 및 작업처리 속도의 향상은 전체 서비스의 질을 높이고 사용자를 만족시키는데 아주 중요한 의미를 가지게 된다.

일반적으로 기존의 단일 웹서버에서는 고가·고성능의 시스템을 구축하여 웹서버에서 발생될 수 있는 이러한 문제들을 해결하고 있으나, 급격한 접속 사용자 수 증가에 따른 서버의 확장성 및 유연성 등이 부족하다. 이를 위해 최근 리눅스에 의해서 일반 PC를 클러스터링 하여 클러스터 가상 웹서버를 구축하고 있는 사례가 크게 증가하고 있다[2]. 이러한 가상서버는 외부에서 접속을 시도하는 가상

IP만을 사용자에게 공개하고 실제로는 부하분산서버가 실제서버들에게 부하를 분산하여 사용자에게 서비스를 제공하는 방식으로 사용자측에서는 마치 하나의 서버에만 접속한 것처럼 투명성(transparency)을 제공하며 서버측에 대해서는 고가용성과 신뢰성 및 확장성을 보장할 수 있다. 또한, 클러스터링 웹서버의 고가용성을 위해 각각의 서버에서 발생한 수 있는 고장에 대해서 신속한 진단 및 검출 그리고 적절한 복구 및 무정지 시스템 구축에 대한 고려가 필수적이다. 또한 이러한 클러스터 환경에서 동작하는 웹 서버는 대용량의 파일 입출력의 실시간 처리에 대한 고가용성을 만족시킬 수 있어야 하고, 각 서버는 공유자원의 이용을 위해서 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[3].

본 논문에서는 리눅스 기반의 클러스터링 가상 서버를 구축 방법뿐만 아니라 클러스터링 가상 서버의 고가용성을 위한 MON, Heartbeat, Fake, Coda 의 구축 방법에 대하여 설명한다. MON은 클러스터를 구성하는 실제서버들의 상태를 모니터링하기 위해 사용되고, Heartbeat 와 Fake 는 부하분산서버의 SPOF (Single Point of Failure) 발생에 대비하여 사용된다. Coda는 진보된 네트워크 고장 포용 파일 시스템으로서 Replication 서버와 클라이언트에서의 캐시 정책을 통하여 클러스터링 가상서버의 고가용성을 향상시킨다. 그리고 이러한 고가용성 기법들을 실제 여러 대의 리눅스 머신을 이용하여 구축한 클러스터링 가상서버에 적용한 결과에 대해서도 살펴보도록 하겠다.