

RPC 기반 동적 가중치를 적용한 웹 클러스터 시스템의

부하분산에 관한 연구

김수진, *최용귀, **윤천균
호남대학교, *Jaulinux, **호남대학교

naxunicorn@naver.com, *elgar@jaulinux.com, **chqyoun@honam.ac.kr

A study on load distribution of the web clustering system based on RPC dynamic weight

Sujin Kim, *Younggui Choi, **Chunkyun Youn
Honam University, *Jaulinux, **Honam University

요 약

인터넷의 급속한 성장과 인터넷이 일상 생활에서 차지하는 중요도는 점점 커지고 있다. 그에 따라 인터넷의 정보 전송량도 극적으로 증가하고 있으며, 과중한 데이터의 양은 네트워크의 병목현상, 데이터 처리량에 따른 시스템의 과부하 등의 문제점들을 야기시켰다. 이러한 문제점들을 극복하기 위한 방안으로 확장성과 높은 가용성을 가진 서버를 이용한 클러스터 시스템에 대한 요구가 증가하고 있다.

본 논문은 클러스터링의 부하분산 알고리즘 중 기존의 고정 가중치에 의한 부하분산 방법을 개선한 RPC(Remote Procedure Call)기반 동적 가중치를 적용하여 효율적인 부하분산 클러스터링 방법론을 제안하고, prototype 을 구현하여 성능을 실험하였다. 클러스터 시스템에서 일정 시간 동안 누적된 평균 real 서버 상태를 참조하는 기존의 동적 가중치 적용 사례와는 달리, 본 논문에서는 RPC 라이브러리를 이용해 실시간으로 real 서버의 상태를 측정하여 보다 정확한 각 서버의 상태정보를 이용하여 특정한 서버에 편중됨이 없는 효율적인 부하분산이 가능하도록 하였다.

실험 결과, 기존의 고정 가중치 방법보다 real 서버들간 향상된 부하분산 효과가 있었으며, 보다 빠른 응답 시간을 통해 웹 클러스터 시스템의 성능 향상을 확인할 수 있었다.

I. 서 론

인터넷의 폭발적인 성장과 인터넷이 우리 생활에서 차지하는 중요도가 커짐에 따라 인터넷의 정보 전송량도 매년 100%를 넘는 증가가 이루어지고 있다[1]. 최근 그래픽 처리와 CGI(Common Gateway Interface) 및 관련 응용 프로그램의 사용이 증가하면서 웹 서버 시스템에 미치는 부하는 기하급수적으로 늘어나고 있다. 이러한 상황에서 온라인상의 웹 사이트들은 사용자에게 보다 빠르고 안정적인 서비스를 제공할 수 있도록 웹 서버 시스템에 대한 높은 가용성과 확장성을 요구하고 있다 [2]. 이러한 해결책으로 일반 PC 나 저가형 워크스테이션들을 고속 네트워크로 연동하여 고성능 컴퓨팅을 제공하는 클러스터링 시스템에 대한 관심이 급증하고 있다[3].

클러스터란 여러 대의 시스템을 네트워크를 통해 연결하여 하나의 단일 시스템처럼 동작하도록 제작한 시스템을 말한다. 클러스터 시스템은 그 사용 목적에 따라, 순수한 과학 계산을 위한 시스템과 부하분산을 목적으로 하는 시스템으로 구분될 수 있다. 그 중에서 부하분산 클러스터는 단일 서버를 사용할 때 발생할 수 있는 부하를 다른 서버로 분산시켜 웹 서버의 과부하를 해결할 수 있는 방법이다.

클러스터 시스템은 슈퍼컴퓨터나 병렬컴퓨터보다 확장성 측면에서 우수하고 단위 시스템의 성능 향상과 분

산/병렬처리 기술의 발달로 고속 컴퓨팅 및 멀티미디어 데이터 처리를 효과적으로 수행할 수 있다[4]. 이러한 장점을 이용하여 웹 서버를 클러스터 형태로 구성하였을 때 기존의 웹 서버가 가져야 할 요구 조건인 확장성, 가용성, 경제적인 효율성을 만족시킨다.

본 연구에서는 RPC 기반 동적 가중치를 적용한 웹 클러스터 기법을 통해서 실시간으로 서버의 자원상태를 측정하여 특정한 서버에 편중됨이 없는 효율적인 부하분산 방안을 제안한다.

본 논문의 구성은 다음과 같다. 제 2 장에서는 기존의 클러스터링 방법론에 관한 관련연구를 소개하며, 제 3 장에서는 본 논문에서 제안하는 RPC 기반 동적 가중치를 적용한 웹 클러스터 시스템의 부하분산 방법론에 대해 기술한다. 제 4 장에서는 제안한 방식에 대한 실험 및 성능을 평가하고 문제점을 제기한다. 마지막으로 제 5 장에서는 결론을 맺고 추후 연구 방향을 소개한다.

II. 기존의 부하분산 클러스터링 기법

웹 클러스터 시스템의 부하분산을 담당하는 핵심적인 기능으로 다음과 같은 부하분산 알고리즘이 있다.

Round-Robin(RR) 알고리즘은 real 서버의 접속수나 응답시간을 고려하지 않고 모두 동일하다는 가정하에 순서적으로 작업 서버를 선택하는 방식이다. 스케줄링에서 소요되는 처리 시간을 효과적으로 단축시킬 수 있