

웹서버 성능 향상을 위한 로드 밸런서 성능 연구

장동원^{*}, 조평동^{*}

^{*}한국전자통신연구원

Load Balancer Performance Testing and Analysis for Improving Web Server Performance

Dong-won, Jang, Pyung-dong, Cho

^{*}Electronics and Telecommunications Research Institute

E-mail : dwjang@etri.re.kr

요 약

본 고에서는 웹 서버 팜(farm)의 성능을 향상시키기 위해서 사용되는 로드 밸런서(load balancer)의 성능을 측정하고 분석하였다. 로드 밸런서는 입력되는 트래픽을 서버 팜의 각 서버에 고르게 분산시켜서 클라이언트의 접속 요구를 최대한 수용해서 전체적인 서비스 품질을 향상시킨다. 본 고에서는 로드 밸런서가 초당 최대 수용할 수 있는 접속 요구를 측정하고 이때 로드 밸런싱(load balancing)이 서버 팜의 각 서버로 고르게 이루어졌는지 측정한다.

ABSTRACT

In this paper, we tested and analysed the performance of the load balancer which advances the performance of web server farm. Load balancer distributes the traffic that is inputting from internet and accepts client's connection requests to its maximum capacity. Then we test the maximum connection requests per second that load balancer processed and the load balancing that established with the load balancer.

키워드

로드 밸런서, 인터넷, 라우터, 게이트웨이, 로드 밸런싱, 알고리즘

1. 서 론

인터넷의 급속한 성장에 따라서 트래픽이 많아지고 복잡한 프로토콜 및 서비스들을 지원할 수 있는 요구도 증가되었다.

또한 인터넷이 중요한 정보 교환 수단으로 사용되게 됨으로써 가용도(availability) 및 이상 장애(fault tolerance)에 발생이 중요한 문제로 부상되었다. 이러한 문제를 해소하기 위한 시도는 동일한 데이터를 갖는 서버를 여러 개 두어서 운영하는 것이었다. 그러나 이 방법은 서버 수가 늘러나면 비용이 증가될 뿐만 아니라 관리가 복잡해지고 특정 서버에 과부가 걸리면 전체에 이상 장애가 과급되는 등 만족한 해결 방법은 되

지 않는다. 이러한 문제를 지능적으로 처리할 수 있는 방법이 로드 밸런싱 기능이다.

로드 밸런서는 효율적으로 트래픽을 분산시킨다. 또한 웹 서버 팜(farm)을 감시하고 최적의 성능과 가용도를 얻기 위한 라우팅 결정을 수행한다.

로드 밸런서는 이와 같이 많은 트래픽을 처리하기 위해서 더 많은 하드웨어 장비를 늘리지 않고도 서버를 클러스터링(clustering)하거나 팜을 만들어 효율적으로 트래픽을 관리해서 네트워크의 서비스 품질을 높여 준다.

본 고에서는 Fast Ethernet 포트 2개를 지원하는 전용 하드웨어 형태의 Layer 4 스위칭 기

능을 수행하는 로드 밸런서를 대상으로 하였다.

로드 밸런서의 성능을 시험하기 위해서 Spirent사의 SmartBits 2000에 ML7710카드 14장을 장착하고 SmartTCP v1.1을 사용하여 maximum session capacity 및 maximum TCP connection setup rate을 측정하였다.

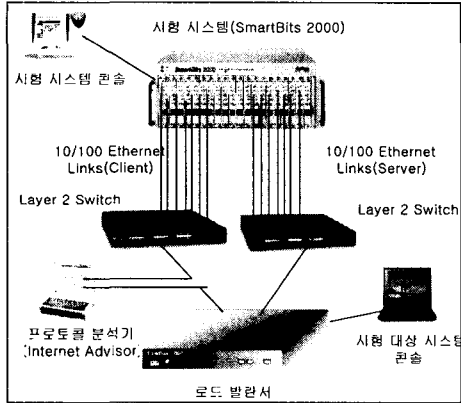


그림 1. 성능 시험 구성도
Fig. 1 Test Configuration

본 고에서는 로드 밸런서의 성능을 평가하기 위해서 초당 몇 개의 TCP connection setup을 지원하고, 초당 TCP session을 몇 개나 처리하며, 최대 session capacity는 얼마인지를 측정한다. 또한 최대 connection setup rate 및 다양한 트래픽 환경에서도 안정적으로 동작하고 로드 밸런싱 성능 저하가 없는지를 분석하였다.

II. 본 론

본 고에서는 로드 밸런서의 로드 밸런싱 성능을 분석하기 위하여 maximum TCP connection-setup rate 및 maximum session capacity 성능을 측정한다.

TCP connection-setup rate를 시험하기 위하여 시험기의 클라이언트 포트를 통하여 TCP request rate를 증가시키며, 이에 대한 로드 밸런서의 최대 connection setup rate와 서버 포트간의 로드 밸런싱 상태를 측정한다.

본 고에서는 로드 밸런서에 대한 사전 throughput 시험을 통하여 성능의 범위를 정하고, connection rate를 2800 개/sec부터 차례로 증가시키며 각 시험에 대해 112,000개씩

connection request를 발생시켰다.

Session capacity 시험에서는 connection setup rate 성능시험에서 얻은 최고 connection request rate를 중심으로 8,400개/sec에서 14,000개/sec까지 초당 connection request를 변화시키며 최대 session capacity를 측정하였다.

III. 시험 환경

로드 밸런서의 성능 측정을 위해서 그림 1과 같이 실제 환경과 똑같은 테스트베드를 구축한다.

시험 시스템은 클라이언트와 서버 역할을 한다. 클라이언트측은 모두 7개의 포트가 구성되며 각 포트별로 TCP 트래픽을 생성한다. 또한 서버측은 7개의 포트가 구성되며 각 포트가 1개의 서버 역할을 한다.

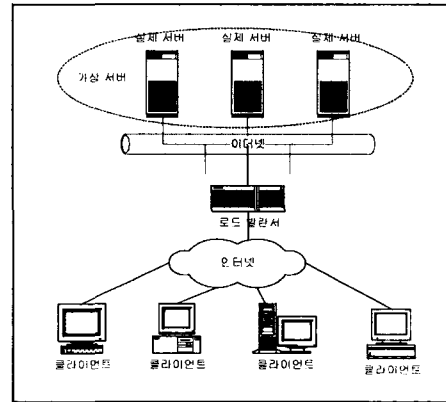


그림 2. 로드 밸런싱 네트워크
Fig. 2 Load Balancing Network

IV. 시험 방법

Connection setup 시험은 로드 밸런서가 TCP 접속을 수행할 때의 성능을 측정한다(그림 3). 특히 로드 밸런서가 접속 요구를 서버간에 얼마나 잘 밸런싱하면서 TCP 세션의 접속 성립 단계를 수행하는지 분석할 수 있다. 이 시험은 시험 시스템의 클라이언트측 7개 포트를 통해서 각 trial당 모두 112,000개의 TCP connection request가 생성된다. 이때 초당 2,800개에서 21,000개까지 1,400개씩 증가시키면서 14번의 trial을 수행한다.

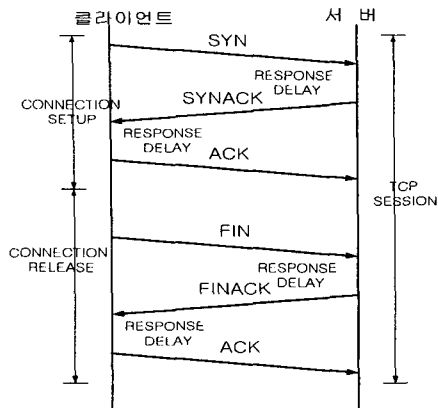


그림 3. TCP 메시지 교환 절차
Fig. 3 TCP Message Exchange Procedure

세션 rate 시험은 로드 밸런서가 TCP 접속 및 해제할 수행할 때의 성능을 측정한다(그림 3). 그러나 세션 rate 시험은 로드 밸런서가 TCP 세션의 접속 성립 단계 뿐 만 아니라 접속 해제를 포함한 전체 TCP 세션을 수행하면서 로드 밸런싱을 얼마나 잘 수행하는지를 분석한다.

세션 rate 시험은 세션 rate이 최고인 점을 중심으로 접속 setup이나 teardown을 적정 범위에서 증가시키면서 로드 밸런서의 성능을 측정한다. 일반적으로 세션 rate 최고점은 connection setup 최고점의 약 1/2이다. 그러므로 이 점을 중심으로 해서 세션 rate 시험을 수행한다.

V. 시험 결과

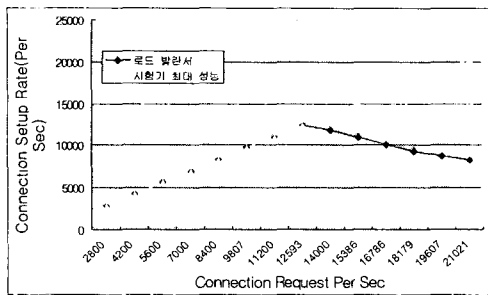


그림 4. Connection Setup 측정 결과
Fig. 4 Connection Setup Test Result

로드 밸런서에 대한 connection setup 시험 결과는 그림 4과 그림 5와 같다. Connection request rate을 증가시키면서 로드 밸런서의

connection setup rate를 측정하는 이 시험에서 최대 초당 12,384개의 connection을 설정하였다. 이 시험과 함께 로드 밸런서가 round-robin 방식의 로드 밸런싱 알고리즘을 사용하여 7개의 서버에 트래픽을 고르게 분산시키는 것을 관찰하였다.

그림 4는 로드 밸런서가 초당 입력된 12,600개의 TCP connection request 트래픽을 성능 열화 없이 처리하고 있음을 보여 준다. 그러나 초당 12,600개의 TCP 트래픽을 초과하게 되면 성능이 점차 떨어짐을 알 수 있다.

그림 5는 connection setup 시험을 수행하면서 로드 밸런서가 7개의 각 서버로 트래픽을 분산시킨 결과를 보여 준다.

본 시험에서는 round-robin 알고리즘을 사용한 결과이다. 로드 밸런서는 거의 균일하게 7개의 각 서버로 트래픽을 분산시키고 있음을 알 수 있다.

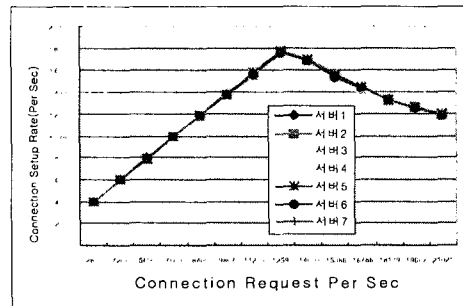


그림 5. 로드 밸런싱 측정 결과(Connection Setup)
Fig. 5 Load Balancing Test Result

로드 밸런싱 알고리즘은 로드 밸런서의 connection setup 성능에 영향을 줄 수 있다. 즉 알고리즘 자체가 로드 밸런서에 영향을 줄 수 있다.

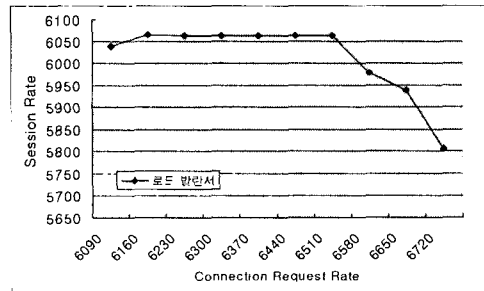


그림 6. Session Rate 측정 결과
Fig. 6 Session Rate Test Result

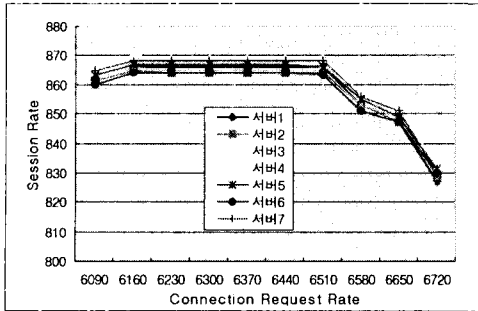


그림 7. 로드 발란싱 측정 결과(Session Rate)
Fig. 7 Load Balancing Test Result

그림 6과 그림 7은 session rate 시험 결과이다. 최대 session rate는 최대 connection setup rate의 약 1/2에 해당하는 6,065개/sec로 측정되었으며(그림 6) 이는 session rate이 호성립 단계 뿐 만 아니라 호 해제 단계까지 포함하기 때문이다. 이 경우에도 역시 로드 발란서는 성능 최고치까지 안정적인 로드 발란싱을 유지하고 있음을 알 수 있다(그림 7).

VI. 결 론

본 고에서는 로드 발란서의 connection setup 및 session rate 성능 측정 및 로드 발란싱 성능 측정 및 분석에 대해서 기술하였다.

로드 발란싱은 하드웨어적인 방법과 소프트웨어적인 방법으로 매우 다양하게 이루어질 수 있다. 본 고에서는 일종의 라우터인 하드웨어적인 방법으로 구성된 로드 발란서를 대상으로 하였다. 로드 발란서는 외부 네트워크와 내부 서버 팜 사이에 연결되는 장비로 클라이언트와 서버간에 트래픽을 효율적으로 전달하고 분배한다.

로드 발란싱은 알고리즘에 따라서 약간의 차이를 나타낸다. 본 고에서는 정적 알고리즘인 round-robin 알고리즘을 수행하는 로드 발란서를 측정하였다. 또한 로드 발란싱 알고리즘은 입력되는 트래픽이 클라이언트 별로 매우 불규칙하면 로드 발란싱에 영향을 받는다. 본 고에서는 각 클라이언트에서 고르게 트래픽을 발생 시켜서 측정을 수행하였다.

본 고에서 측정된 connection setup 및 session rate 성능 및 로드 발란싱 성능은 로드

발란서간의 절대적인 비교보다는 상대적인 비교에 활용되어야 한다. 왜냐하면 하드웨어적인 장점과 소프트웨어적인 장점을 모두 갖는 로드 발란서 등 여러 가지 형태로 설계될 수 있기 때문이다. 그러므로 본 고에서 측정된 성능값은 로드 발란싱을 위한 비용에 대한 상대적인 성능을 비교하는데 사용되어야 한다.

참고문헌

- [1] Netcom Systems, Smart TCP User Guide, V2.01, April, 2001.
- [2] Zhiruo Cao, Zheng Wang, and Ellen Zegura, "Performance of Hashing-Based Schemes for Internet Load Balancing", Tech. Rep. GIT-CC-99-14, Georgia Tech, 1999.
- [3] Wensong Zhang, Shiyao Jin, Quanyuan, Wu, "Creating Linux Virtual Servers", National Laboratory for Parallel & Distributed Processing, LinuxExpo'99, May, 1999.