

웹서버 클러스터 시스템의 성능 요소 분석

박진원^{1†}

Analysis on the Performance Elements of Web Server Cluster Systems

Jin-Won Park

ABSTRACT

This paper is on the research result for analyzing the performance of GLORY(GLOBAL Resource management sYstem) used for Web Server Cluster system, which was developed at ETRI(Electronic and Telecommunication Research Institute). The paper includes the definition of Web Server Cluster System, the characteristics of the system, user oriented system performance, current performance enhancement methods, computer simulation model for GLORY and its experimental results for the performance of GLORY. GLORY is composed of 2048~1,000,000 units of PCs, and is used for Internet servers. From the results of the simulation experiments, we notice that GLORY has enough capacity to fully serve the appropriate level of Internet services. Also, the results show that Web server service time is longer than that for network transmission time but requires more DNS than expected, and that 100Mbps LAN is good enough for directly connecting Internet to the Web servers while not affecting the total system performance.

Key words : Web Server Cluster System, Internet Server, Performance Evaluation

요약

본 논문은 국내 연구기관에서 개발한 웹서버 클러스터 시스템인 GLORY의 성능을 평가하기 위한 시뮬레이션 모델을 구축하고 이를 통하여 GLORY를 구성하고 있는 각 자원들의 활용률을 중심으로 성능을 평가한 것이다. 이를 위하여 웹서버 클러스터 시스템에 대한 정의, 사용자 측면과 운영자 측면에서 본 시스템 성능, 기존의 시스템 성능 향상 방안 등에 대해 살펴보았다. 또한, 웹서버 클러스터 시스템에 대한 컴퓨터 시뮬레이션 모델링, 시뮬레이션 실험 결과에 대해 기술되어 있다. 본 논문의 분석 대상인 GLORY는 2,048~1,000,000대의 PC로 구성되는 대용량 인터넷 서버시스템이다. 컴퓨터 시뮬레이션 수행 결과, GLORY는 적절한 수준의 인터넷 서비스를 충분히 수용할 수 있는 능력을 보이고 있으며, 웹서버 서비스 시간이 네트워크 전송 시간보다 많이 소요되고 있고, DNS가 예상보다 많이 소요될 것으로 보인다. 웹서버와 직접 연결되어 있는 네트워크는 100Mbps 수준의 LAN을 구축하여도 전체 시스템 성능에 큰 영향을 미치지 않을 것으로 보인다.

주요어 : 웹서버 클러스터 시스템, 인터넷 서버, 성능 평가

1. 웹서버 클러스터 시스템

웹서버 클러스터 시스템(Web Server Cluster System)

*이 논문은 2009학년도 홍익대학교 학술진흥연구비에 의하여 지원되었음.

2010년 7월 9일 접수, 2010년 9월 7일 채택

¹⁾ 홍익대학교 게임학부 게임소프트웨어전공

주 저자 : 박진원

교신저자 : 박진원

E-mail: jinon@hongik.ac.kr

은 인터넷의 폭발적인 성장으로 기존의 소규모 인터넷 서버들의 처리 용량이 한계에 부딪히자 저가의 서버들을 클러스터 형태로 구성하여 인터넷 서버의 확장성, 성능, 신뢰성을 향상시키는 방안으로 제시되어 사용되고 있는 대규모 인터넷 서버이다⁵⁾. 일반적으로 웹서버 클러스터 시스템은 수 천대에서 수 만대에 이르는 PC급 서버들을 한 장소에 네트워크로 연결하여 하나의 대용량 웹서버 시스템을 구축하고 있다.

예를 들면 네이버의 경우 대략 80,000대의 서버로 시

시스템을 구축한 것으로 알려져 있으며 Google의 경우 DB 구축을 위해서만 80,000대의 서버를 사용하고 있어 전체 시스템은 이보다 훨씬 많은 수의 서버들을 사용하고 있는 것으로 추측된다^[1,4]. 웹서버 클러스터 시스템의 속성을 파악하기 위하여 이와 유사한 기능을 제공하는 것으로 생각할 수 있는 그리드 컴퓨팅(Grid Computing)과 비교해 보면 표 1과 같다. 표 1을 보면 웹서버 클러스터 시스템은 예전의 그리드 컴퓨팅과 용도, 기능 측면에서 매우 다른 시스템으로 그리드 컴퓨팅이 계산 능력 제고에 초점을 맞춘 것이라면 웹서버 클러스터 시스템은 자료 검색 기능 제고에 초점을 맞춘 시스템으로 이해될 수 있다. 그리고 그리드 컴퓨팅이 작업부하 분산을 위해 구축되는 것인데 비해 웹서버 클러스터 시스템은 데이터베이스 분산을 위해 구축된 것으로 볼 수 있다.

웹서버 클러스터 시스템의 구성 단위는 일반 성능을 갖는 PC급 컴퓨터로써 운영체제를 탑재하고 메모리와 하드디스크를 장착한 시스템이다. 따라서 웹서버 클러스터 시스템의 단위 구성 요소인 서버들은 독립적인 계산, 검색, 저장 기능을 수행할 수 있다. 웹서버 클러스터 시스템은 이들 단위 PC를 수 천대에서 수 만대까지 한 장소(하나의 빌딩 수준)에 LAN으로 연결하여 대용량 DB 저장 시스템을 구축하는 것이 주 목적이다.

본 논문은 웹서버 클러스터 시스템으로 개발된 GLORY (GLObal Resource management sYstem)의 성능을 평가하기 위해 컴퓨터 시뮬레이션 모델을 구축하고, 이를 통하여 GLORY를 구성하고 있는 각 자원들의 활용률을 중심으로 성능을 평가할 것이다. 이를 위하여 본 논문에는 웹서버 클러스터 시스템에 대한 정의, 시스템 특징, 사용자 측면과 운영자 측면에서 본 시스템 성능, 기존의 시스템 성능 평가 방안 검토, GLORY에 대한 컴퓨터 시뮬레이션 모델, 상세한 모델링 내용, 시뮬레이션 결과 분석에 관한 내용이 기술되어 있다.

표 1. 웹서버 클러스터 시스템과 그리드 컴퓨팅 비교

분류	웹서버 클러스터	그리드 컴퓨팅
용도	검색 중심	계산 중심
네트워크	Open Network	Closed Network
운영체제	Windows/Linux 기반	Linux 기반
주요기능	DB 분산, IDC 기능	작업 부하 분산, Supercomputin 기능
구성단위	독립적인 단위 컴퓨터	특정 기능용 부품

본 논문의 분석 대상인 GLORY는 한국전자통신연구원에서 설계 개발한 웹서버 클러스터 시스템으로, 적게는 2,048대에서 많게는 1,000,000대에 이르기까지의 PC로 구성되는 인터넷 서버용 시스템이다^[2]. 대규모 웹서버 시스템으로 활용될 클러스터 시스템은 상세한 성능 평가 작업이 아직은 충분히 이루어지지 않은 상황에서 성능평가 방안을 제시한다는 점에서 의미가 있는 것으로 보인다.

2. GLORY

GLORY는 웹서버 클러스터 시스템의 일종으로 GLORY의 규격과 모습을 살펴보면 그림 1과 같다. 그림 1은 1만 대 규모의 단일 데이터센터를 클러스터로 구성하는 모습을 나타내고 있다. 이러한 단일 데이터센터는 100개까지 확장 가능하여 최대 100만대의 클러스터 시스템으로 확장 가능하다.

GLORY의 단일 데이터센터내 클러스터 컴퓨팅 플랫폼 규격은, 랙(Rack) 당 16 노드, 24 ports GbE(down link 1Gb, up link 1Gb*2)로 구성되고, 서브넷 당 16 랙, 40 ports GbE(down link 1Gb, up link 1Gb*4)로 연결되며, 10 GbE Core Router 당 8개 서브넷 연결(down link 1Gb, up link 10Gb)*5 이 가능하다. 다중 데이터 센터용 클러스터 컴퓨팅 플랫폼 규격은 데이터 센터들간 전용선으로 100개까지 연동 가능하다. GLORY는 대규모 저가 노드와 공개 소프트웨어를 기반으로 글로벌 인터넷 서비스를 지원하는 솔루션을 제공한다.

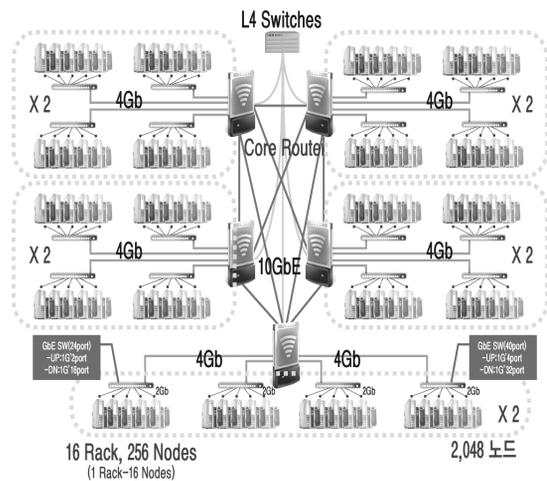


그림 1. 단일 데이터센터내 클러스터 H/W 구성도

표 2. Load Balancing Strategy 비교

Approach	특징	장점	단점
Client (1)	Client 측면 Distributed	No Server Overhead	Limited Applicability
DNS (2)	Web Server 측면 Centralized	No Bottleneck	Partial Control
Dispatcher (3)	Web Server 측면 Centralized	Fine Grained Load Balancing	Dispatcher Bottleneck Packet Rewriting Overhead
Server (4)	Web Server 측면 Distributed	Fine Grained Load Balancing Distributed Control	Latency Increase Packet Rewriting Overhead

3. 웹서버 클러스터 시스템의 성능 요소

웹서버 클러스터 시스템의 성능은 두 가지 측면에서 살펴 볼 수 있다. 시스템 운영자 측면에서는 시스템 응답 시간과 더불어 시스템의 균등 활용률이 중요한 성능 지수가 된다. 균등 활용률은 네트워크로 연결되어 있는 웹서버 클러스터 시스템의 각 구성 요소들이 균등하게 활용됨으로써 시스템 운영의 신뢰성과 안정성을 도모하고 효율적인 시스템 운영을 위해 제시되는 성능 지표이다. 시스템 사용자 입장에서는 단연 시스템 응답 시간이 최대 관심사이다. 사용자가 원하는 작업을 얼마나 빨리 웹서버 클러스터 시스템이 제공할 수 있는가를 나타내는 시스템 응답 시간은 사용자들이 가장 중요하게 생각하는 시스템 성능 지표이다^[4].

웹서버 클러스터 시스템의 서비스 품질은 하드웨어 측면에서 다음 두 가지 요소에 좌우된다. 첫째, 네트워크의 자료 전송 속도이다. 이는 인터넷 링크의 대역폭에 따라 결정되는 요소이다. 둘째, 하드웨어 구성 요소의 성능이다. 즉, CPU, RAM 및 I/O 성능 등 각 서버 구성 요소의 성능에 따라 결정되는 요소이다. 최근의 컴퓨터 시장은 하드웨어 가격이 상대적으로 저렴해짐에 따라 하드웨어 측면에서의 시스템 성능 향상 방안은 단순히 추가적인 하드웨어 구성 요소를 투입하는 방향으로 전개되고 있다.

웹서버 클러스터 시스템 서비스 품질을 소프트웨어, 혹은 운영 측면에서 향상시키려는 노력이 계속되고 있다. 소프트웨어 측면에서 시도되고 있는 성능 향상 방안은 Load Balancing Strategy, DNS(Domain Name Server) 사용의 최소화, Core 사용, Replica 사용(Google, Hadoop 등)이 시도되고 있다. 다음에는 이들 각각의 방법에 대해 살펴보자.

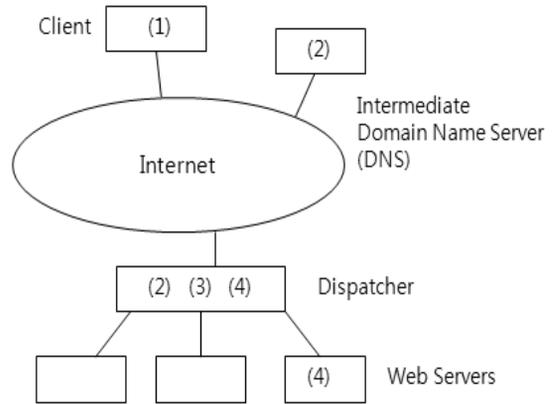


그림 2. Load Balancing 방식에 대한 도식화

3.1 Load Balancing Strategy

Load Balancing Strategy는 엄밀히 말하면 서비스 응답시간 측면에서의 시스템 성능 향상 방안이라기보다는 시스템 운영자 측면에서 전체 웹서버 클러스터 시스템의 균형적인 활용을 도모하기 위한 방안으로 사용되는 것이다. Cardellini^[6]는 클러스터 형태의 웹서버 시스템에서 서버들 간의 Load Balancing을 위한 방안을 Client Based Approach, DNS Based Approach, Dispatcher Based Approach^[8], Server Based Approach 등 4가지로 분류하고 이들의 특징, 장단점을 분석하였다. 표 2는 이들의 내용을 간단히 비교해 본 것이다.

Load Balancing 방식을 표 2의 괄호 안의 번호에 맞춰 표현하면 그림 2와 같다.

3.2 DNS 사용의 최소화

DNS based Approach는 클라이언트와 서버의 상태를 고려하여 URL을 Real IP로 번역하는 과정에서 TTL(Time To Live)이라는 변수를 설정하고, 이를 활용하여 웹서버

클러스터 외부의 Intermediate Name Server를 Real IP 번역 과정에 활용하는 방식이다. 외부 DNS를 활용함으로써 내부 DNS에 걸리는 작업부하를 최소화하는 효과를 기대할 수 있다. TTL 값을 짧게 설정하면 외부 DNS의 역할이 감소하고 내부 DNS의 작업부하가 증가하지만 TTL 값을 크게 설정하면 외부 DNS의 역할 증대와 작업 부하 증대 효과가 나타난다. 이는, TTL 값이 너무 작으면 내부 DNS의 작업부하가 많아서 HTTP 처리 업무가 많아지고, 반대로 TTL 값이 너무 크면 외부 DNS가 웹서버 클러스터 시스템의 자료를 오랫동안 유지하고 있어야 하므로 전체 인터넷에 작업부하가 많이 걸리기 때문이다. TTL 값은 실제 운영시 1시간 정도로 설정하는 것이 적절한 것으로 알려져 있다.

3.3 Core 사용

일반적으로 사용자 서비스 요청의 80~90%가 지역 서버에 의해 처리되는 것으로 알려져 있다. 이는 대부분의 서비스가 어느 특정 응용, 혹은 내용(Contents)에 국한되어 있음을 의미한다. 여기서 Core라는 개념은 지역 서버에서 빈번하게 요구되는 작은 크기의 파일을 의미한다. Core를 사용한다는 것은, 모든 서버에 특정 파일을 모두 유지하게 하여 사용자의 서비스 요구를 처리하게 하는 것을 말한다. 한 연구 결과에 의하면⁹⁾ 일정 기간동안 발생한 총 접근 파일(2GB)의 5%(100MB)를 Core로 지정하고 이를 RAM에 저장하여 접근 시간을 단축시키는 방식을 활용하여 16개의 Node로 구성되어 있는 웹서버 클러스터 시스템을 대상으로 실험한 결과, Core 개념을 사용하지 않고 자료를 분산해서 RAM에 저장한 경우보다 처리율 기준으로 32% 정도 우수한 성능을 보였다고 한다. 또한, 모든 서버가 모든 자료를 디스크에 저장한 경우보다는 무려 250% 이상 우수한 성능을 보인 실험 결과를 발표하였다. 자료를 분산해서 RAM에 저장한 경우에는 TCP handoff overhead가 15/16, 즉 94%의 확률로 발생했고, 모든 자료를 Disk에 저장한 경우에는 Disk 접근 속도가 RAM 접근 속도에 비해 10배 정도 느리기 때문에 TCP handoff overhead가 발생하지 않는 이점에도 불구하고 낮은 성능을 보인 것으로 판단하고 있다.

3.4 Replica 사용

Replica란 복제본을 의미한다. 웹서버 클러스터의 성능 향상을 위하여 Replica를 사용한다는 의미는 여러 웹서버에 주요 데이터의 복수 복사본을 유지한다는 것을 말한다. Google 파일 시스템이나 Hadoop 분산파일 시스템

에서 Replica 개념을 사용하고 있다. Google은 3개의 Replica를, Hadoop은 2~3개의 Replica를 유지하고 있는 것으로 알려졌다. Replica를 이용하면 전체적으로 네트워크의 부하를 감소시켜 준다. 그리고 Replica는 고장 감내 기능을 유지하기 위한 방안으로도 활용할 수 있다. 그러나 복수 데이터들 간의 데이터 일치성을 유지하기 위한 오버헤드와 쓰기 작업시 작업부하를 증가시키는 단점을 갖고 있다.

4. GLORY에 대한 컴퓨터 시뮬레이션

GLORY는 최대 100만 대까지의 웹서버를 연결하여 사용할 수 있도록 설계된 시스템이다. 그러나 실제로 100만 대의 웹서버를 연결하여 사용하는 시스템은 당분간 흔한 경우가 아닐 것으로 예상하여 본 시뮬레이션에서는 최대 구성 가능 시스템보다 훨씬 규모를 축소하여 모델링하고자 한다.

본 시뮬레이션의 대상 시스템은 2,048대로 구성된 웹서버 클러스터 시스템으로서의 GLORY이다. 그림 3은 시뮬레이션 모델링을 위하여 2,048대의 웹서버를 논리적으로 재구성한 것이다. 이는 물리적인 구조와 다소 다를 수 있으나 동작 모형을 구축하기 위하여 재구성한 것이다. 2,048대는 256대로 구성된 기본 노드가 8개 연결되어 있는 구조로 GLORY의 최소 구성 단위로 볼 수 있다. 이보다 더 큰 규모를 대상으로 할 경우 네트워크의 단계가 한 더 추가될 뿐 기본적인 동작 모형은 동일하다고 판단하여 2,048대로 구성된 시스템을 시뮬레이션 대상으로 설정하였다. GLORY를 대상으로 구축한 시뮬레이션 모델은 Arena로 작성되었다. 2,048대의 개별 웹서버들은 Resource로 정의되었고 DNS, Dispatcher, 네트워크 등도 Resource

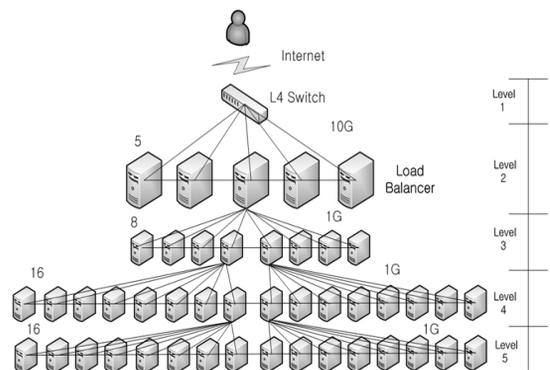


그림 3. 시뮬레이션을 위한 GLORY의 논리적 구조도

로 정의되었다. 특히 네트워크는 라우팅 단계별로 연결되는 라인 수에 맞춰 Resource 개수를 정의하였다. 사용자들의 웹서버 클러스터 사용 요구는 Page View, Query, 동영상 서비스 등 3가지 형태에 따라 별도의 엔티티로 정의하였다.

그림 3의 구조는 실제로 GLORY가 운용될 상황과 다소 차이가 있을 것으로 예상된다. GLORY는 10,000대를 기준으로 1,000,000대까지 확장 가능한 구조로 설계되어 있고, 인증, DB, 트랜잭션 등 특정 서비스에 특화된 서버들로 분리 운영될 것이므로, GLORY의 전체 구조 역시 모든 서버가 동일한 기능과 네트워크상의 위치를 갖는 그림 3과는 달리 서비스별, 기능별, 지역별로 조정된 구조를 가질 것으로 예상되기 때문이다. 그러나, 본 시뮬레이션 모델이 구축되고 있는 시점에서 GLORY의 구체적인 운용 구조를 정의하기 어려우므로 여기서는 단순한 구조, 즉 모든 서버가 동일한 기능과 네트워크 상의 위치를 갖는 것으로 가정하고 시뮬레이션을 수행하였다.

4.1 GLORY 동작 개요

2,048대의 웹서버로 구성된 GLORY를 대상으로 시뮬레이션 모델을 구축하는 과정에서 각 서버는, Core 개념을 도입하여 사용자 파일 서비스 요구의 95%는 Local에서 처리하고 5%만이 redirecting을 통해 다른 웹서버에 의해 처리된다고 가정했다. 또한, 네트워크는 Tree 형태로 구성되어 있으며 DNS는 인터넷 Gate에 1대 혹은 그 이상 설치되어 있다고 설정했다. 작업 부하는 실제 상황의 1/20 정도 수준에서 시작하여 점차 증가시키는 방향으로 실험을 진행하였고, 궁극적으로 추출하고자 하는 시스템 성능 지수는 네트워크 사용률과 서버 사용률, 각각 상이한 서비스 요구 패턴에 대한 응답 시간 등으로 설정하였다.

4.2 작업 부하 분석

웹서버 클러스터 시스템의 실제 작업부하는 직접 기업이 발표한 자료를 사용하는 것이 타당하나, 시뮬레이션을 수행할만큼 세부적인 자료를 구하기가 쉽지 않다. 따라서 본 연구에서는 언론에 공개된 자료와 논문이나 문헌에 발표된 자료를 조합하여 해석하는 방식을 택했다.

작업 부하와 관련하여 수집된 1차 자료는 국내에서 운영되는 대표적인 포털 업체인 Naver를 중심으로 Google의 자료를 일부 활용한 것으로 구체적인 내용은 다음과 같다^[3].

- 방문자 수: 1일 2,200만명(Naver, 2007년) 기준
- Page View: 1일 10억회, 도착시간 간격은 평균 0.864

ms(지수분포) 가정

- Query: 1일 1.1억회, 도착시간 간격은 평균 7.854ms(지수분포) 가정
- MPC(동영상 보기): 크기 16~80MB, 1일 200만건 기준, 도착시간 간격은 평균 432ms(지수분포) 가정
- UPLD(동영상 업로드): 크기 16~80MB, 1일 20,000건(Google U-tube), 도착시간 간격은 평균 2,160ms(지수분포) 가정

4.3 컴퓨터 시뮬레이션 모델을 위한 시스템 파라미터

요구되는 서비스 유형별 작업부하 외에 컴퓨터 시뮬레이션 모델을 구축하기 위해 필요한 GLORY 시스템 자체의 파라미터들은 다음과 같다.

4.3.1 DNS 작업 처리 시간

내부 DNS가 Real IP Address를 변환하는데 소요되는 TCP Communication Overhead setup time은 75ms 소요되는 것으로 설정하였다. 이는 Olshefski^[11]에 근거하여 산출한 것이다. 그리고 최초 방문자를 위한 인증시간(Authentication time)은 20ms로 설정하였다. 이는 Chung^[10]에 따른 것이다.

4.3.2 네트워크 전송 시간

10Gbps Ethernet을 통한 자료 전송 시간은 네트워크의 유효 전송률을 감안하여 조정된 전송률로 계산하였다. 즉, 10Gbps의 용량을 갖는 네트워크는 실제로 4.8~6.0 Gbps 정도의 유효 전송률을 보인다는 연구 결과^[12]에 따라 각 데이터 크기에 따른 전송 시간을 계산하였다.

4.3.3 HTTP 처리 시간

HTTP를 처리하기 위한 시간은 300ms^[11]로 설정하였다.

4.3.4 요구 파일 검색을 위한 시간

사용자가 원하는 자료를 검색하는데 소요되는 시간은 웹서버 클러스터 시스템의 하드웨어 구조, 파일 시스템, DB 검색 및 저장구조에 따라 차이가 있을 것으로 예상된다. 본 연구에서는 Zhu^[13]를 인용하여 평균 400ms가 소요되는 것으로 설정했다. 그러나 좀 더 현실을 반영하기 위해 자료 검색 시간 분포를 삼각분포[200ms, 400ms, 600ms]에 따르는 것으로 가정하였다.

4.3.5 메모리 접근 시간

메모리 접근 시간은 Cardellini^[7]에 기술된 바에 따라,

1MB 데이터를 메모리에서 읽을 때 10ms, 디스크에서 읽을 때 100ms가 소요되며, 쓰기에는 이의 2배 시간이 소요된다는 사실을 근거로 계산하였다. 그리고 메모리에서 직접 데이터를 가져 올 확률을 95%로 가정하고 디스크에서 데이터를 가져 올 확률을 5%로 가정하여 계산한 결과, 14.5KB의 경우 $14.5 * 0.002048 = 0.0297\text{ms}$ 로 설정하였고 16.125KB 데이터를 읽어 오는 시간은 $(16.125 * 10 / 1,000) * 0.95 + (16.625 * 100 / 1,000) * 0.05 = 0.2338\text{ms}$ 로 설정하였다. 그러나 동영상 자료를 읽는 경우(MPC)나 쓰는 경우(UPLD)에는 메모리 읽기와 데이터 전송이 최초의 1.5 KB를 읽은 다음에 파이프라이닝 된다는 가정에서 최초의 1.5KB(read 단위) read 시간, 혹은 write 시간을 감안한 후 별도의 메모리 접근 시간은 산정하지 않았다. 읽기는 $1.5 * 10 / 1,000 = 0.015 \text{ ms}$, 쓰기는 $1.5 * 20 / 1,000 = 0.030\text{ms}$ 로 각각 계산되었다.

4.3.6 MPC 및 UPLD를 위한 파일 크기

동영상 자료의 크기는 일정하지 않다. 그러나 포털에서는 대개 동영상 파일 크기에 제한을 둔다는 전제 아래 동영상 데이터는 16MB~80MB 크기의 일양분포를 따른다고 가정하였다. 이는 인터넷에서 통용되는 동영상 파일 포맷을 기준으로 대략 1분~5분에 해당하는 데이터 크기를 설정한 것이다.

4.4 Arena를 이용한 GLORY 컴퓨터 시뮬레이션 모델

본 GLORY 대상 컴퓨터 시뮬레이션 모델은 Arena를 이용하여 구축되었다. 구체적인 모델의 모습은 여기에 서술하지 않는다. 본 시뮬레이션은 사용자로부터 서비스 요구가 시작되어 사용자에게 최종적으로 결과가 전송되는 전 과정을 다룬 것이 아니다. 사용자의 인터넷 사용 행태, 인터넷의 구조 및 특성 등을 감안해야 하는 어려운 점이 있고 본 시뮬레이션의 목적이 웹서버 클러스터 시스템 자체의 성능을 분석하는 것이므로 모델링의 범위를 웹서버 클러스터 시스템 자체에 국한시켰다. 즉, 인터넷을 통해 입력된 사용자 요구 내용이 웹서버 클러스터 시스템에 도착한 시점부터 웹서버 클러스터 시스템이 서비스를 마치고 인터넷에 그 결과를 띄워 보내는 순간까지를 분석 대상으로 하였다.

5. 시뮬레이션 결과 분석

2,048개의 웹서버로 구성된 GLORY를 대상으로 Arena

를 이용하여 구축된 모델을 활용하여 입력 파라미터를 변화시키면서 반복 실험을 수행하여, 네트워크 사용률, 웹서버 사용률, 각 서비스 유형별 응답시간 등의 실험 결과를 분석한다. 그러나 시스템 파라미터(네트워크 속도, 서버 처리 속도, GLORY 구조 등)는 변화시키지 않았다. 이는 실제 시스템이 구축되어 운영되기 전에는 정확한 시스템 파라미터를 설정하는 것이 어렵고, 시스템 성능을 입력 파라미터에 대해 상대적인 수치를 관찰하는 것이 본 실험의 목표이기 때문이다.

5.1 실험 입력 파라미터

반복 실험은 Page View 서비스 요청 빈도(평균 서비스 요청 발생 시간 간격), Query 서비스 요청 빈도, MPC 서비스 요청 빈도 및 UPLD 서비스 요청 빈도 등 4개의 입력 파라미터를 변화시키면서 수행하는 것이다. 다음 표 3은 입력 파라미터 조합을 나타내고 있다. 표 3에 서술된 숫자는 ms 단위로 각 서비스들에 대한 평균 서비스 도착시간 간격을 의미한다. 이들은 Naver의 실제 서비스를 기준으로 1일 24시간 서비스를 가정하여 산출한 것인데, 2,048대 웹서버 클러스터 시스템 규모를 고려하여 기본적으로 실제 서비스의 1/20 정도의 작업 부하를 고려하였고 이후에는 Peak time을 고려하여 이보다 더 무거운 작업부하를 부과하였다.

표 3에서 보는 바와 같이 Page View와 Query는 실제 서비스의 1/20, 1/10 또는 1/5에 해당하는 작업부하를 부과하였으나 MPC와 UPLD는 실제 작업부하와 같은 다소

표 3. 실험 입력 파라미터*(ms, 개)

Case	Page View	Query	MPC	UPLD	DNS 개수
1	0.864	7.854	432.0	2,160.0	1
2	0.864	7.854	432.0	2,160.0	2
3	0.864	7.854	432.0	2,160.0	3
4	0.864	7.854	432.0	2,160.0	4
5	0.864	7.854	216.0	2,160.0	3
6	0.864	7.854	432.0	1,080.0	3
7	0.864	7.854	216.0	1,080.0	3
8	0.432	3.927	432.0	2,160.0	3
9	0.432	3.927	216.0	1,080.0	4
10	0.432	3.927	216.0	1,080.0	5

* 1일 10억회(Page View)의 1/20, 1.1억회(Query)의 1/20, 2,000,000회(MPC), 400,000(UPLD)회 방문 기준

높은 수준의 작업부하를 부과하였고 경우에 따라서는 더 무거운 작업부하도 부과하였다. 이는 예비 실험을 수행해 본 결과 MPC와 UPLD의 작업부하가 너무 가벼워서 전체 시스템의 성능에 큰 영향을 미치지 못하는 것으로 나타났다. 그러나 실제 상황에서도 이와 같은 현상이 일어나고 있는지 다소 의문이 있다. 이는 MPC 관련 자료 수집 과정에서 오류가 있을 수도 있고 실제 상황에서도 이와 같은 현상이 나타나고 있을 수도 있는 것으로 보인다.

5.2 실험 결과 분석

실험은 총 30,000ms 동안 반복하여 실시하였고 실험 결과는 네트워크 사용률, 웹서버 사용률, 각 서비스 유형별 응답시간 등을 중심으로 수집하였다. 표 3에 제시된 실험 입력 파라미터에 따라 수집된 실험 결과를 서비스 종류에 따른 성능 지수(대기 시간)를 중심으로 표현하면 표 4와 같고, 자원 중심으로(자원 활용률) 표현하면 표 5와 같다.

표 4를 살펴보면 Page View는 DNS 수에 따라 대기 시간이 크게 변하는 것을 알 수 있고 기본 수준의 서비스 요구 빈도(Cases 1~7)에 맞추기 위해서는 적어도 3개 이상의 DNS를 운용해야 하는 것으로 나타났다. Page View 요구가 2배로 증가하는 경우(Cases 8, 9, 10) DNS가 3, 4개일때 다시 대기 시간이 크게 증가하여 DNS를 5개로 확장시켜야 대기 시간이 다시 짧아지는 것을 알 수 있다.

현재 수준의 작업 부하에서 Query 서비스는 대기 시간이 크게 변하지 않는 것을 알 수 있다. 이는 Page View에 비해 Query는 상대적으로 작업 부하(서비스 요구 빈도)가 매우 낮은 수준이기 때문인 것으로 판단된다. 동영상 서비스(MPC)나 업로딩(UPLD) 서비스는 Page View 서비스의 작업부하가 비교적 낮은 수준일 때는 대기시간이 비교적 안정적으로 짧은 것으로 나타나고 있으나 Page View 서비스 요구 빈도수가 2배로 증가하는 경우와 MPC, UPLD 자체 서비스 요구 빈도가 2배로 증가하는 경우에는 대기 시간이 크게 증가함을 알 수 있다.

표 5를 살펴보면 DNS의 활용률이 전체적으로 매우 높음을 알 수 있다. 대체로 네트워크의 활용률보다 웹서버나 DNS의 서버 활용률이 더 높게 나타나고 있다. 이는 GLORY가 10Gbps와 1Gbps 네트워크를 사용함으로써 전송 시간을 획기적으로 감소시켰기 때문이다. 네트워크 사용률 측면에서 10Gbps LAN의 활용률이 높게 나타나고 있는데 이는 2,048대의 웹서버를 모두 연결해야 하는 부담과 동영상 자료의 전송에 상당한 시간을 할애해야 하

표 4. 서비스 중심 실험 결과(대기 시간, ms)

Case	Page View	Query	MPC	UPLD	DNS 개수
1	78.6355	0.3203	4.5781	5.4815	1
2	22.9827	0.3627	7.3031	9.2825	2
3	1.3024	0.3634	6.5810	7.7831	3
4	0.4798	0.3467	6.3624	5.6696	4
5	1.3210	0.8054	13.9445	14.1994	3
6	1.1374	0.3453	6.1738	5.7356	3
7	1.6561	0.9501	16.4445	17.9710	3
8	60.8629	0.3322	5.6725	4.0969	3
9	16.8604	1.0145	18.7566	19.1757	4
10	2.0766	0.9504	16.7934	21.3189	5

표 5. 자원 중심 실험 결과(활용률)

Case	DNS	10Gbps (L4-Core)	1Gbps (Core-Root)	1Gbps (Root-Local)	WS
1	0.9989	0.2088	0.0907	0.0113	0.2278
2	0.9963	0.2285	0.0983	0.0123	0.2300
3	0.7526	0.2277	0.0979	0.0175	0.2286
4	0.5486	0.2242	0.0966	0.0173	0.2306
5	0.7476	0.3945	0.1605	0.0200	0.2317
6	0.7577	0.2232	0.0963	0.0120	0.2307
7	0.7554	0.4426	0.1785	0.0223	0.2292
8	0.9996	0.2414	0.1156	0.0145	0.4563
9	0.9994	0.4762	0.2035	0.0254	0.4599
10	0.8570	0.4550	0.1961	0.0245	0.4648

기 때문인 것으로 보인다.

표 5의 마지막 경우, 즉 Page View의 서비스 요구 빈도가 Case 1에 비해 2배, Query의 서비스 요구 빈도도 2배, 동영상 서비스 및 동영상 Uploading 서비스 요구까지 모두 2배의 작업 부하가 걸려있는 경우 DNS를 5대 사용하더라도 DNS의 사용률이 85.7%에 달하고 10Gbps 네트워크의 활용률이 45.5%, Web Server 활용률이 46.5%에 이르는 등 2,048대의 웹서버 클러스터 시스템으로 감당하기에는 한계 상황에 도달해 있다고 판단된다.

6. 결론 및 향후 연구 방향

본 논문은 일반적인 웹서버 클러스터 시스템의 성능 요소를 분석하고 국내 연구기관에서 개발한 GLORY

(GLObal Resource management sYstem)의 성능을 자원 활용률을 중심으로 분석한 것이다. 본 논문에는 웹서버 클러스터 시스템에 대한 정의, 시스템 특징, 사용자 측면과 운영자 측면에서 본 시스템 성능, 기존의 Web Server Cluster 시스템 성능 평가 방안 개요, GLORY에 대한 컴퓨터 시뮬레이션 모델, 시뮬레이션 결과 분석에 관한 내용을 포함하고 있다.

2,048대로 구성되는 최소 규격의 GLORY를 대상으로 실제 운용되는 상황을 가정하여 Page View, Query, 동영상 서비스 및 동영상 업로딩 서비스 등 4가지 형태의 서비스를 수행하는 시나리오를 설정하였다. GLORY가 운용되는 상황을 최대한 반영하여 GLORY의 성능을 평가하기 위한 컴퓨터 시뮬레이션 모델을 구축하고 실험을 수행하였다.

컴퓨터 시뮬레이션 수행 결과, GLORY는 적절한 수준의 인터넷 서비스를 충분히 수용할 수 있는 능력을 보이고 있다. 실험 결과, 일반적으로 웹서버 서비스 시간이 네트워크 전송 시간보다 많이 소요되고 있고, DNS가 예상보다 많이 소요될 것으로 보이며 웹서버와 직접 연결되어 있는 네트워크는 100Mbps 수준의 LAN을 구축하여도 전체 GLORY 성능에 큰 영향을 미치지 않을 것으로 보이기도 한다. 물론 이는 세밀한 시뮬레이션 실험이 추가로 수행되어야 의미있는 결론을 내릴 수 있을 것이다. 또한, 향후 2,048대의 웹서버를 DNS, 인증 서버, DB 서버, 트랜잭션 서버, 동영상 서버 등 특정한 용도에 맞춰서 재구성하고 이를 바탕으로 세밀한 시뮬레이션 연구가 수행되어 효율적인 GLORY 운용 방안을 설정하는 노력을 기울여야 할 것이다.

참 고 문 헌

1. 강병준, 류현정, 구글 vs 네이버, 전자신문사, 2008.

2. 김명준, 김준, 이미영, 김학영, 김성운, 김영균, 남궁한, "GLORY: 대규모 저가 노드 기반 글로벌 인터넷 서비스 솔루션," 한국정보처리학회 논문지, 14(3), pp. 387-394, 2007년 6월.
3. 조선일보, "인터넷 제국 NHN의 빛과 그림자", 2007년 11월 16일자, 2007년 11월 27일자.
4. S.E. Arnold, Google Version 2.0, Infonortics Ltd., 2007.
5. M. Banikazemi, V. Moorthy and D.K. Panda, "Efficient collective communication on heterogeneous networks of workstation," National Science Foundation, U.S.A., June 2000.
6. V. Cardellini, et. al., "Dynamic load balancing on web server systems," IEEE Internet Computing, May/June 1999.
7. V. Cardellini, et. al., "Enhancing a web server cluster with quality of service mechanism," Proc. of the IEEE Int'l Performance, Computing, and Communications Conference, Phoenix, U.S.A., 2002.
8. D. Caromel, et. al., "Asynchronous peer-to-peer web services and firewalls," IWJPCD, Denver, U.S.A., 2005.
9. L. Cherkasova and M. Karlsson, "Scalable web server cluster design with workload-aware request distribution strategy WARD," HP Workshop, 2001.
10. Y.H. Chung, D.S. Moon, T.H. Kim and J.W. Park, "Workload dispatch planning for real time fingerprint authentication on a sensor-client-server model," PDCAT 2004, LNCS 3320, pp. 833-838, 2004.
11. D. Olshefski and J. Nieh, "Understanding the management of client perceived response time," SIGMetrics/Performance '06, Saint Malo, France, 2006.
12. W.D. Smith, TPC W: Benchmarking an e-commerce solution, www.tpc.org/tpcw.
13. H. Zhu, et. al., "Demand-driven service differentiation in cluster-based network servers," Proc. of IEEE INFOCOM, Anchorage, U.S.A., 2001.



박진원 (jinon1@hongik.ac.kr)

1975 서울대학교 산업공학과 학사
 1982 미국 오하이오주립대 산업시스템공학과 석사
 1987 미국 오하이오주립대 산업시스템공학과 박사
 1988~1999 한국전자통신연구원 책임연구원
 2000~현재 홍익대학교 게임학부 교수

관심분야 : 모델링&시뮬레이션, 컴퓨터시스템 성능평가, 공학교육