

## 로그자료 분석을 통한 인터넷 서버의 용량 분석

김수진\* · 윤복식\*\* · 이용주\*\*\* · 강금석\*\*

### Capacity Analysis of Internet Servers Based on Log-Data Analysis

Su Jin Kim\* · Bok Sik Yoon\*\* · Yong Joo Lee\*\*\* · Jin Xi Jiang\*\*

#### ■ Abstract ■

Due to the rapid increase in the internet traffic volume, ISPs are faced with the definite need of the expansion of server capacity. In order to provide prompt services for customers and still prevent excessive facility cost, it is critical to determine the optimum level of internet server capacity. The purpose of this paper is to provide a simple but effective strategy on the expansion of server capacity according to the increase in internet traffic. We model an internet server as an M/G/m/m queueing system and derive an efficient method to compute the loss probability which, in turn, is used as a basis to determine proper server capacity. The process of estimating the traffic parameter values at each server based on log data analysis is also given. All the procedures are numerically demonstrated through the process of analyzing actual log data collected from a game company.

Keyword : Internet Server Capacity, M/G/m/m, Log Data Analysis

### 1. 서 론

인터넷의 등장으로 정보화 사회로의 전환이 가

속화되면서 인터넷 사용자 뿐만 아니라 인터넷 서비스 업체들도 날로 증가하고 있다. 이러한 속에서 인터넷 서비스 업체들은 수익기반을 다수의 고객

논문접수일 : 2001년 5월 14일    논문게재확정일 : 2001년 12월 24일

\* EK Communication

\*\* 홍익대학교 기초과학과 응용수학 전공

\*\*\* 이화여자대학교 경영학과

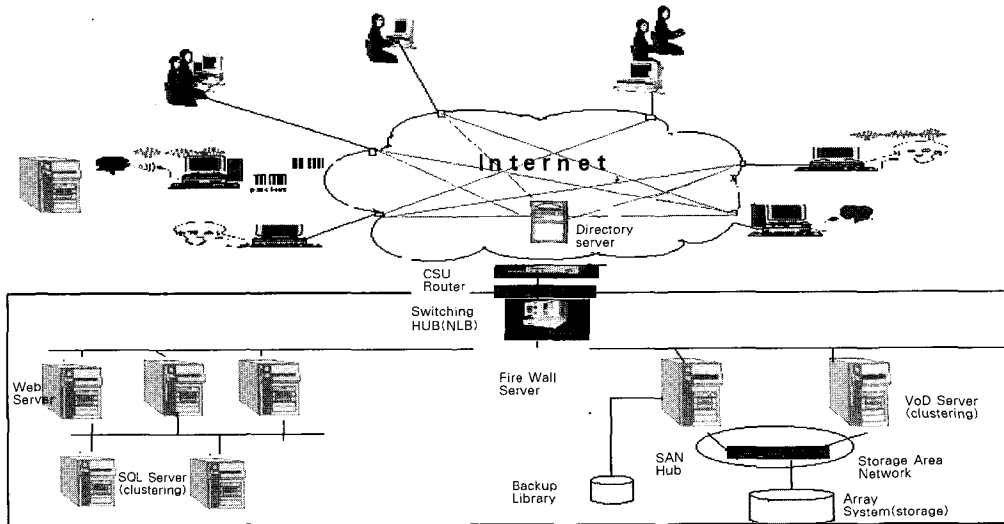
유치로 인한 접속률을 높이는 것에 있다고 보고 더 많은 사용자를 확보하기 위해 다양한 콘텐츠 개발을 통하여 서비스 영역을 확대해 나가고 있다. 이와 함께 인터넷 서비스 업체들은 고객들이 늘어나면서 고객의 지연에 따른 불편사항을 해결해야 하는 또 다른 상황에도 직면해 있다. 인터넷 이용 시의 접속지연은 서비스에 대한 불만으로 이어져 고객들의 이탈을 불러오고, 심리적인 측면에서는 인터넷 서비스 운영기업의 이미지에 타격을 입히게 된다. 서비스 불만으로 인한 「실망이탈고객」을 다시 유인하는데는 적지 않은 비용과 시간이 소모되므로 성장기에 있는 인터넷 서비스 사업에 커다란 장애 요소가 될 수 있다.

국내 조사결과에 따르면 국내 인터넷 이용자수는 2000년 5월 1,534만 명, 6월 1,575만 명, 7월 1,603만 명으로 월 평균 약 37만 명씩 증가하여 2000년 8월말 현재 1,640만 명으로 추산되고 있으며(한국인터넷정보센터, <http://www.nic.or.kr>), 해외보고서에 따르면 1999년 5월 250만이던 host의 수가 7월에는 310만으로, 9월에는 340만으로 증가하였고, 이러한 추세로 증가한다면 2001년까지 대략 1,000만개의 웹 호스트(web host)가 있을 것으로 예상되지만 다수의 웹사이트가 존재함에도 불구하고 웹 트래픽(web traffic)의 80%는 모든 사이트 중 0.5%에만 집중된다고 한다(Alexa Internet Trends Report, <http://www.alexa.com>). 인터넷 이용자 및 서비스 공급자 양측의 급격한 증가가 진행되고 있는데, 이 상황에서 인터넷 기업의 생존 전략은 고객욕구에 부응하는 다양성, 신뢰성, 적합성을 갖춘 콘텐츠 개발과 함께, 폭증하는 인터넷 트래픽에 대비한 인터넷 서버의 증설을 비롯한 고품질의 인터넷 환경의 구축에 초점이 맞추어 져야 할 것이다. 물론 기업 경영의 측면에서 과도한 시설투자는 잠식 투자자본(sunk investment)이 되어 이로 인한 자금의 가용성 및 회전율(turn over)저하가 발생하므로 이를 최소화하고

미래 수익성을 살릴 수 있는 적절한 투자범위의 결정이 요구된다고 볼 수 있다. 그러나 간과할 수 없는 다른 한 측면은 소비고객의 증가에 대응하는 적절한 서버의 증설로 소비자 인식(consumer reputation)의 양(+)적 향상을 위한 투자를 함으로써 QoS(quality of service)를 만족시키는 서비스를 제공하여 기업 가치창출의 경쟁력을 갖춰야 한다는 점이다.

본 연구의 목적은 인터넷 서비스 업체에서 사용자 증가에 따라 합리적으로 서버 증설을 결정하는 단순하면서도 효과적인 방법을 제시하는 것이다. 인터넷 성능에 관한 연구는 망의 입력 트래픽의 측정 및 적절한 모델링에 관한 이론적인 연구가 많은데(Heyman,1998, Harris et al.,2000 등), 인터넷 서버의 성능에 관한 실용적인 연구는 의외로 많이 발표되지 않고 있다. 본 연구에서는 동질의 인터넷 입력 트래픽을 가정하고 서버에서 관측할 수 있는 로그자료만을 사용하여 트래픽 대비 서버 용량의 적정성을 판단하고 적절한 증설시점을 찾아내는 방법을 제시한다. 결국 일정수준 이상의 서비스 품질을 유지하기 위한 서버구축 방안이 주로 연구되는데, 필요 이상의 복잡성을 회피하기 위해 본 연구에서는 사용자 단말기에서 망까지의 개별적 연결의 차이에 의한 지연 문제는 제외시키고, 사이트 운영자 측면에서 서버에서의 처리시간(서버내부의 처리지연 포함)만을 고려한다.

본 서론에 이어 2절에서는 인터넷 서비스의 일반적인 진행과정을 소개하고 3절에서는 인터넷 서버를 m서버 손실시스템으로 모형화하여 손실함수를 계산하는 방법을 제시하고 4절에서는 실제의 인터넷 기업의 로그자료를 분석하여 트래픽 파라미터를 구하는 사례를 상세히 설명하고 5절에서 대상 기업의 세 종류 인터넷 서버들에 대한 용량 분석 사례를 제시한 후 6절에서 결론을 맺는다.



<그림 1> 일반적인 인터넷 서비스 과정

## 2. 인터넷 서비스의 일반적 과정

### 2.1 일반적인 인터넷 서비스 과정

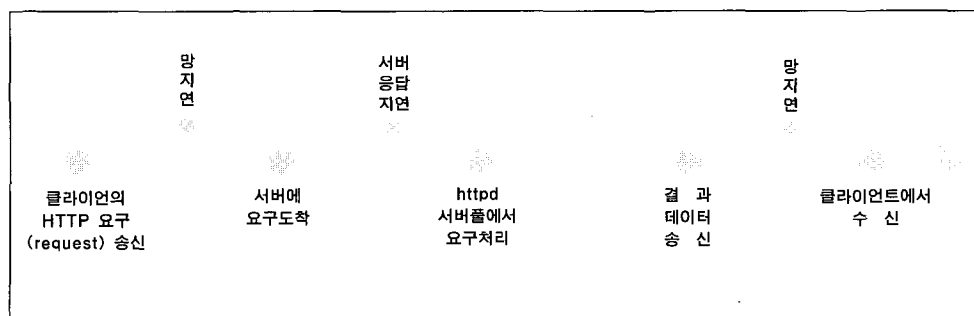
일반적으로 고객이 인터넷 망을 통해 시스템에 접속하여 서버에 의해 원하는 서비스를 받게 되는 과정은 <그림 1>과 같다.

<그림 1>을 보면 고객은 단말 서브시스템(user agent)에 의해 원하는 검색내용을 입력하게 되고, 이러한 고객요구(request)는 통신 서브 네트워크를 통해 정보를 처리해주는 서브 시스템으로 입력된다. 라우터에 의해 스위칭 허브로 입력되어 기업

서브 시스템으로 입력되는 고객요구들은 방화벽 (fire wall)이라 불리는 보안 서버를 거쳐 각 웹서버로 할당된다. 고객이 웹화면에서 발생시키는 사건에 따라 요구는 DB서버나 VoD서버 등으로 할당된다.

### 2.2 서버 내에서의 지연

웹서비스의 경우 클라이언트/서버 구조에 기반을 두고 진행된다. 클라이언트의 응용프로그램이 서버 호스트를 찾아 호스트의 HTTP(hyper text transfer protocol) 서버와 TCP 접속을 한 후 프로



<그림 2> 웹서비스에서의 지연

토콜 HTTP에 따라 한 개 또는 여러 개의 다중매체 문서를 요구하게 된다. 서버는 요구된 데이터의 전송을 완료한 후 연결을 끊게 된다.

이 과정을 구체적으로 나타내면 <그림 2>와 같은 데 로그자료의 분석에 의해 웹서버에서의 처리 시간은 쉽게 계산될 수 있으나 웹서버 이전, 이후의 망 관련 지연은 망에 별도의 측정장치를 설치하여야 한다(Dilley et al.,1998). 본 연구에서는 웹서버의 성능에 초점을 맞추어 로그자료에 의한 분석을 주로 행한다. 본 연구의 분석대상이 되는 <그림 2>에서의 처리시간(httpd 서버풀에서 요구처리)은 서버내부에서의 지연을 포함하게 된다.

### 3. 인터넷 서버의 대기행렬 모형

#### 3.1 M/G/m/m 모델링

인터넷 서비스는 보통 여러 대의 서버가 병렬로 연결되어 동시에 많은 수의 고객을 처리하도록 구성되어 있고 또한 하나의 서버 내에서도 프로세스 공유를 통해 많은 고객의 동시 접속이 가능하도록 설계되어 있다. 인터넷 서버의 동시접속 용량은 서비스에 동원되는 망, CPU, RAM, 하드디스크 등의

하드웨어와 소프트웨어 자원에 의해 결정되는데 보통 서버의 운영자가 적정 서비스율이 유지될 수 있는 한도 내에서 동시접속자 수(m)를 제한하게 된다. 인터넷 서버를 모델링 할 때 서버의 수를 하나로 보고 1 서버 프로세스 공유(process sharing) 모형(Yaschkov, 1987)으로 접근할 수도 있다. 하지만 컴퓨터의 용량 한도 내에서는 접속자의 증가에 따른 인터넷 서비스 점유시간의 변화가 거의 없다고 관찰되므로 m서버 손실 시스템으로 모델링 하는 것이 더 현실에 근접할 것이라고 판단된다. 실제로 4장에서 언급할 A사의 기업포털 웹서버의 사례에서 얻은 <표 1>에서 보듯이 고객의 접속이 많을 때와 적을 때의 고객 별 평균 서버점유시간이 큰 차이가 없음을 알 수 있다. 예를 들어 최빈시간대(4장의 <그림 4> 참조)인 오후 4시대의 경우 59.99초 정도인데 한산한 시간대(4장의 <그림 4> 참조)인 오전 5시대의 경우 60.1초로 오히려 약간 커짐을 관찰할 수 있다. 또한 <표 1>에서 서비스 시간(세션시간)은 평균 58초로 계산할 수 있고 시간대별로 최저 51.67초에서 최대 63.7초까지 분포되어 있음을 볼 수 있는데 이들의 모두 평균에서 표준편차(=3.26) 2배의 범위 내에 들어 있어 평균

<표 1> 기업 웹서버에서의 시간대별 체재시간

(단위 : 초)

시 간 간 격	체재시간	시 간 간 격	체재시간	시 간 간 격	체재시간
00:00:00 ~ 00:59:59	56.76	08:00:00 ~ 08:59:59	57.70	16:00:00 ~ 16:59:59	59.99
01:00:00 ~ 01:59:59	63.94	09:00:00 ~ 09:59:59	54.94	17:00:00 ~ 17:59:59	63.70
02:00:00 ~ 02:59:59	61.58	10:00:00 ~ 10:59:59	56.78	18:00:00 ~ 18:59:59	62.11
03:00:00 ~ 03:59:59	61.80	11:00:00 ~ 11:59:59	58.56	19:00:00 ~ 19:59:59	58.22
04:00:00 ~ 04:59:59	60.55	12:00:00 ~ 12:59:59	57.81	20:00:00 ~ 20:59:59	56.14
05:00:00 ~ 05:59:59	60.10	13:00:00 ~ 13:59:59	56.84	21:00:00 ~ 21:59:59	57.58
06:00:00 ~ 06:59:59	51.67	14:00:00 ~ 14:59:59	57.58	22:00:00 ~ 22:59:59	53.06
07:00:00 ~ 07:59:59	54.74	15:00:00 ~ 15:59:59	56.50	23:00:00 ~ 23:59:59	53.34
총 합					1392.00
평 균					58.00
분 산					10.60
표준편차					3.26

의 차이가 없다는 결론을 내릴 수 있다. 이로부터 서버용량 한도 내에서는 고객수의 과다에 무관하게 평균 서비스 시간이 일정하다는 가정을 할 수 있게 되어 프로세스 공유 모형보다는 서버 수가 매우 많은 병렬서버 시스템이 더 현실에 근접한 것으로 보인다. 또한 인터넷 서비스의 고객 도착과정은 포아송 과정으로 간주하는데 무리가 없으므로 본 연구에서는 개별 인터넷 서버를 M/G/m/m으로 가정한다. 실제로 인터넷 서비스 시간은 지수분포가 아닌 heavy-tail 분포로 많이 간주되고 있다(Crovella et al.,2000, Harris et al.,2000, Menascé et al.,1998 등).

### 3.2 손실확률의 계산

M/G/m/m 시스템의 경우 고객이 용량초과로 인해 시스템에서 기각될 확률인 고객손실확률은 Erlang의 손실공식에 의해 M/M/m/m 시스템과 같게 된다. 도착률  $\lambda$ , 단일서버의 서비스율  $\mu$  인 M/M/m/m 시스템에서의 고객손실확률은

$$P_m = \frac{(\lambda/\mu)^m/m!}{\sum_{n=0}^m (\lambda/\mu)^n/n!} \quad (1)$$

에 의해 직접 구할 수도 있으나 본 연구에서 같이 서버의 용량을 의미하는  $m$ 이 클 때는 수치오차로 인하여 정확한 값의 계산이 불가능하다. 따라서 본 연구에서는 다음과 같은 관계식(Wolff, 1989)을 사용한다.

$$B(m, a) = \frac{aB(m-1, a)}{aB(m-1, a) + m} \quad (2)$$

$$B(0, a) = 1 \quad (3)$$

여기서  $a = \frac{\lambda}{\mu}$  이며,  $B(m, a)$ 는 서버수가  $m$  이고, 제공로드가  $a$  일 때의 고객손실확률이다.

식 (2) 는 서버수  $m$  인 시스템에서  $m-1$ 명의

고객이 있을 확률은  $m-1$ 개의 서버만을 고려한 시스템이 가득 찬 사건과 여기에 서버가 하나 첨가된 시스템은 가득 차지 않는 사건이 동시에 일어날 확률과 같다는 관찰에서 다음 식

$$P_{m-1} = B(m-1, a)(1 - B(m, a)) \quad (4)$$

을 얻을 수 있고, 또한 균형방정식

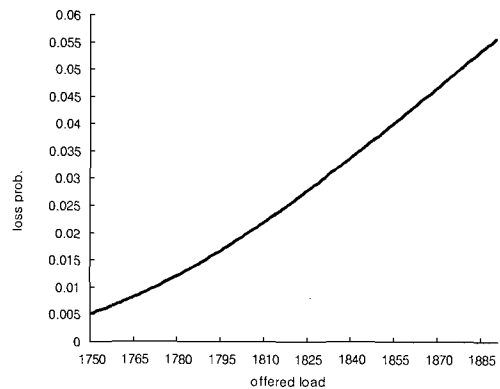
$$\lambda P_{m-1} = m\mu P_m$$

에 의해 다음 식

$$P_{m-1} = \frac{m\mu}{\lambda} P_m = \frac{m}{a} B(m, a) \quad (5)$$

을 얻어 이 두 식의 관계로부터 유도될 수 있다.

식 (2)를 이용하여 수치계산을 하여 4장의 사례 분석에서 사용하기 위해  $m=1800$  일 때의 제공부하  $a$ 에 대한 고객손실확률의 변화를 그래프로 그리면 <그림 3>과 같다.



<그림 3> M/G/1800/1800 시스템의 손실확률

## 4. 로그자료의 분석

### 4.1 A사의 서버의 구성

본 연구에서는 A사의 기업사이트를 담당하는 웹서버에 대해 분석을 한다. 인터넷 포털 사이트

운영업체 A사의 경우 현재(2000년 11월 2일 기준) 가입자 수는 629만 명이며, 동시 접속자수는 14만 명, 1일 평균 페이지 뷰는 120만 명이며, 최빈시(peak time)에는 18만에서 20만 명이 방문한다. 대상기업 A사는 인터넷 상에서 회원제 서비스를 하는 업체로 개인회원 및 기업회원으로 구분되는 회원풀을 운영하고 있으며, 인터넷 솔루션, 게임포탈 및 정보 서비스를 제공하는 업체이다. A사는 웹 서비스를 위해 내부에 현재 353대의 서버를 보유하고 있으며, H사가 운영하는 인터넷 정보센터(IDC)에서 221대의 서버를 위탁관리하고 있다. 353대의 서버 구성은 기업 및 자사 제품 소개를 담당하는 웹 서버 2대, 검색 서버 245대, 게임 서버 106대, 협력업체 서버 1대로 되어있다. A업체 server의 CPU이용률(utilization)은 0~100%까지 부하가 적은 시스템에서는 평균 30%정도의 이용률을, 부하가 큰 시스템에서는 70%의 이용률 분포를 보인다.

인터넷 서버를 서비스 특성 별로 웹서버, DB서버, VoD서버, 게임서버 등으로 분류할 수 있는데 담당하는 서비스에 따라 서비스의 비트 율과 서비스 점유시간의 차이가 있게 된다. A사의 경우 기업 웹서버, 게임포탈 사이트의 서버, 특정 게임의 서버들로 구성되어 있는데, 고객이 기업 웹서버를 통해 기업정보를 살펴본 후 게임에 접속을 원하면 게임포탈로 연결이 되며 게임포탈에서 특정 게임을 선택하여 게임서버로 이동하게 된다. 물론 직접 게임서버로 연결할 수도 있다.

본 연구에서는 A사에서 사용하고 있는 3가지 인터넷 서버, 즉 기업의 홈페이지를 담당하는 기업 웹서버와 다양한 게임을 소개하고 원하는 게임으로 연결해 주는 게임포탈 서버, 또한 실제로 게임을 담당하는 게임서버를 분석하고, 효율적인 증축 방안을 연구하고자 한다. 이 기업의 서버들은 모두  $m=1800$ 으로 설정되어 있다.

인터넷 서버의 분석은 주로 서버의 로그자료의 분석을 통해 이루어진다. 본 연구에서는 Web

Trends의 출력결과 중에서 단위 시간별 session 수, session 별 서비스 시간, session 별 hit 수, 그리고 page view의 수 및 view되는 시간 등을 이용해서 M/G/m/m 모형의 파라미터를 추정하게 된다.

#### 4.2. 기업 웹서버의 분석

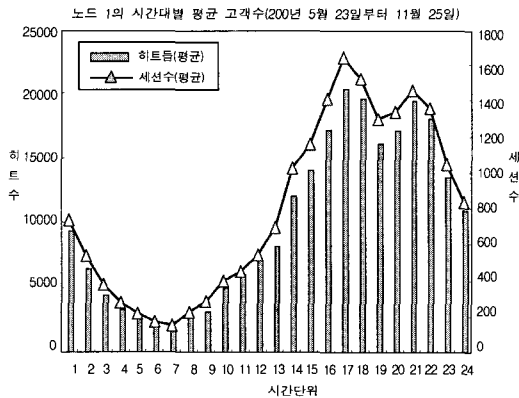
<그림 4>와 <그림 5>는 A업체의 5월 23일에서 11월 25일까지의 로그데이터를 Web Trends로 분석한 후 출력 데이터를 정리한 결과인데 시간대별 평균 고객수 분포와 대역폭의 변동을 볼 수 있다. 이들을 비교하면 대역폭이 고객수(세션수)에 대체적으로 비례하는 것을 관찰할 수 있다. 이것은 웹 서버에서 고객들이 주로 요구하는 서비스들이 http 서비스이므로 예상할 수 있는 결과라고 볼 수 있다. 따라서 입력 고객의 도착률을 (트래픽 비트율이 아닌) 세션수로 파악하기로 한다. 본 연구의 실제 분석에서는 최근의 1달간의 추세만을 반영하기 위해 10월 자료를 위주로 분석한다.

<그림 4>에서 보면 고객 도착의 최빈시간(peak time)은 16:00:00에서 16:59:59 사이임을 알 수 있는데 본 연구에서는 일별 평균 자료와 최빈시간의 자료로 나누어 분석해 본다. 2000년 10월 2일부터 10월 28일까지를 대상으로 하였을 때 노드 1에서의 일별 평균 히트수는 10019(hit/hour)였으며, 최빈시간대에는 20337(hit/hour)를 기록하였다. 평균 시간대 세션 수는 평균 796(session/hour)이었으며, 최빈시간대에는 1639(session/hour)를 기록하였다. 세션과 뷰(view)수의 로그 자료를 통해 노드 1에서 하나의 세션당 히트수는 12(hit/session)이다. 이 자료로부터 초당 고객 도착률은 평균 0.22(세션/초) 최빈시간의 도착률은 0.46(세션/초)로 얻을 수 있다(<표 2> 참조).

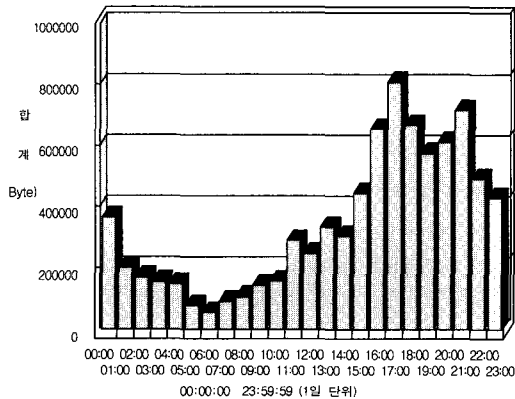
평균 서비스 처리시간(세션시간)은 3장의 <표 1>로부터 58초로 얻을 수 있다.

<표 2> 기업 웹서버의 단위 시간당 고객 도착률

단 위 시 간	고객도착률 (세션/초)	단 위 시 간	고객도착률 (세션/초)	단위시간	고객도착률 (세션/초)
00:00:00 ~ 00:59:59	0.20	08:00:00 ~ 08:59:59	0.08	16:00:00 ~ 16:59:59	<b>0.46</b>
01:00:00 ~ 01:59:59	0.15	09:00:00 ~ 09:59:59	0.11	17:00:00 ~ 17:59:59	0.42
02:00:00 ~ 02:59:59	0.10	10:00:00 ~ 10:59:59	0.12	18:00:00 ~ 18:59:59	0.36
03:00:00 ~ 03:59:59	0.08	11:00:00 ~ 11:59:59	0.15	19:00:00 ~ 19:59:59	0.37
04:00:00 ~ 04:59:59	0.06	12:00:00 ~ 12:59:59	0.19	20:00:00 ~ 20:59:59	0.41
05:00:00 ~ 05:59:59	0.05	13:00:00 ~ 13:59:59	0.29	21:00:00 ~ 21:59:59	0.38
06:00:00 ~ 06:59:59	0.04	14:00:00 ~ 14:59:59	0.32	22:00:00 ~ 22:59:59	0.29
07:00:00 ~ 07:59:59	0.06	15:00:00 ~ 15:59:59	0.39	23:00:00 ~ 23:59:59	0.23
<b>총 합</b>					5.31
<b>평 균</b>					<b>0.22</b>
<b>분 산</b>					0.02
<b>표준편차</b>					0.14



<그림 4> 웹서버의 시간대별 고객 도착수

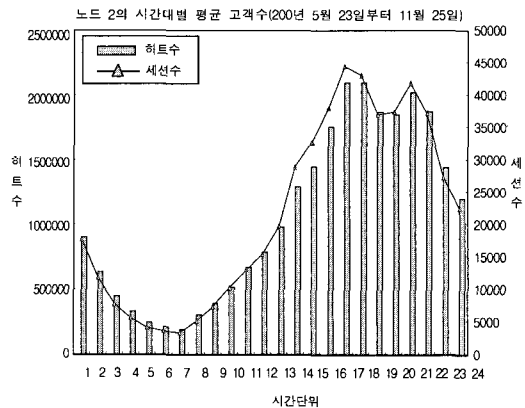


<그림 5> 웹서버의 시간대별 대역폭

### 4.3 게임포털서버의 분석

<그림 6>은 게임포털서버의 경우 2000년 5월 28일부터 2000년 11월 25일까지 시간대별 평균 고객수분포를 나타낸 그래프이다.

게임포털 서버의 10월 2일부터 10월 25일까지 시간대별 시스템로그를 분석하여 구한 도착률은 다음<표 3>과 같다. <표 3>에서 볼 수 있듯이 게임포털 서버의 경우 최빈시간대의 도착률은 4.88(세션수/초)이며, 평균시간대의 도착률은 2.30(세션수/초)이다.



<그림 6> 게임포털 서버의 단위 시간당 고객 도착 수

〈표 3〉 게임포털 서버의 단위 시간당 고객 도착률

단 위 시 간	고객도착률 (세션/초)	단 위 시 간	고객도착률 (세션/단위)	단 위 시 간	고객도착률 (세션/단위)
00:00:00 ~ 00:59:59	1.82	08:00:00 ~ 08:59:59	0.70	16:00:00 ~ 16:59:59	<b>4.88</b>
01:00:00 ~ 01:59:59	1.21	09:00:00 ~ 09:59:59	1.08	17:00:00 ~ 17:59:59	4.71
02:00:00 ~ 02:59:59	0.75	10:00:00 ~ 10:59:59	1.33	18:00:00 ~ 18:59:59	3.95
03:00:00 ~ 03:59:59	0.53	11:00:00 ~ 11:59:59	1.61	19:00:00 ~ 19:59:59	4.10
04:00:00 ~ 04:59:59	0.39	12:00:00 ~ 12:59:59	2.14	20:00:00 ~ 20:59:59	4.62
05:00:00 ~ 05:59:59	0.32	13:00:00 ~ 13:59:59	3.40	21:00:00 ~ 21:59:59	3.87
06:00:00 ~ 06:59:59	0.29	14:00:00 ~ 14:59:59	3.84	22:00:00 ~ 22:59:59	2.73
07:00:00 ~ 07:59:59	0.47	15:00:00 ~ 15:59:59	4.25	23:00:00 ~ 23:59:59	2.20
<b>총 합</b>					55.19
<b>평 균</b>					<b>2.30</b>
<b>분 산</b>					2.67
<b>표준편차</b>					1.63

〈표 4〉 게임포털 서버의 시간대별 체제시간

시간 간격	체제시간	시간 간격	체제시간	시간 간격	체제시간
00:00:00 ~ 00:59:59	275.00	08:00:00 ~ 08:59:59	279.55	16:00:00 ~ 16:59:59	290.64
01:00:00 ~ 01:59:59	309.78	09:00:00 ~ 09:59:59	266.19	17:00:00 ~ 17:59:59	308.61
02:00:00 ~ 02:59:59	298.35	10:00:00 ~ 10:59:59	275.11	18:00:00 ~ 18:59:59	300.90
03:00:00 ~ 03:59:59	299.42	11:00:00 ~ 11:59:59	283.71	19:00:00 ~ 19:59:59	282.07
04:00:00 ~ 04:59:59	293.34	12:00:00 ~ 12:59:59	280.06	20:00:00 ~ 20:59:59	271.97
05:00:00 ~ 05:59:59	291.15	13:00:00 ~ 13:59:59	275.39	21:00:00 ~ 21:59:59	278.98
06:00:00 ~ 06:59:59	250.31	14:00:00 ~ 14:59:59	278.94	22:00:00 ~ 22:59:59	257.08
07:00:00 ~ 07:59:59	265.20	15:00:00 ~ 15:59:59	273.74	23:00:00 ~ 23:59:59	258.44
<b>총 합</b>					6743.92
<b>평 균</b>					281.00
<b>분 산</b>					248.86
<b>표준편차</b>					15.78

게임포털 서버의 경우 평균 체제시간은 281초이며, 각 시간대별 체제시간은 <표 4>와 같다.

평균 도착률을 0.4216(세션/초)로, 평균 체제시간은 1696초로 계산하여 사용한다.

#### 4.4. 게임 서버의 분석

게임 서버의 경우에도 위와 유사하게 분석할 수 있다. 단지 이 경우 시간대별 자료가 주어지지 않아 하루 전체의 평균량을 계산하여 서버 한 대 당

#### 5. 각 서버의 용량 분석

4.2절의 추정치를 근거로, A사의 웹서버의 경우 제공로드  $a$ 는 최빈시간의 경우  $\lambda=0.46$ ,  $\mu=1/58$ 에서 26.4일량, 평균적으로는  $\lambda=0.22$ 에서



12.83 일량으로 얻어진다.

제공로드가 각 서버에서 수용 가능한 최대량인  $m=1800$ 과 동일한 1800일량일 때, 3.2절의 방법으로 계산하면 손실 확률은 0.01857이 되며, 이때의 1초당 도착률은  $\lambda = a\mu$ 에서 31.04로 얻어진다. 서버에 도착하는 제공로드가 1877일량일 때 고객 손실률은 0.05이 되며, 서버에 도착하는 제공로드가 1772일량일 때 고객 손실률은 0.01이 된다. 따라서 A기업이 고객 손실률을 0.05 이하로 유지하기를 원한다면 초당 도착률이 32.36(시간당 116,503)을 넘으면 서버를 증설해야하며, 고객 손실률을 0.01 이하로 유지하기를 원한다면 초당 도착률이 30.55(시간당 109,986)을 넘는 시점에서 서버의 증설을 고려해야 할 것이다.

한편 현재의 서버이용률( $\rho$ )는

$$\rho = \frac{\lambda}{m\mu} = \frac{a}{m}$$

를 통해 계산하면 최빈시간에 0.015 정도이므로 서버의 용량대비 이용률이 매우 낮음을 알 수 있다. 따라서 웹서버의 경우 서버의 용량은 낭비되고 있으며, 현 시스템에서 서버의 증설은 고려할 필요성이 없다고 볼 수 있다.

게임포털 서버의 경우에는 4.3절로부터 작업(session)별 평균 서비스 시간은 281초가 되므로  $\mu = \frac{1}{281}$ 이며, 제공로드는 최빈시간의 경우 1371.55 일량, 평균적으로 646.15일량이다. 서버 한 대당 동시접속자의 처리 용량이 1800세션이므로 현재의 서버의 이용률(= 트래픽 밀도(traffic inten-

sity))은 최빈 시간에  $\rho \cong 0.76$ 정도이며, 평균  $\rho \cong 0.36$ 이므로 서버는 안정적인 운용을 유지하고 있음을 알 수 있다. 손실확률을 0.05 이하로 유지하기 원하면 제공로드가 1877일량 이하가 되어야 하는데(<그림 2> 참조) 이때의 도착률은 초당  $\lambda \cong 6.68$ 이다. 손실확률을 0.01 이하로 유지하려면 제공로드가 1772일량 이하가 되어야 하는데 이때의 도착률은  $\lambda \cong 6.31$ 가 된다. 최빈시간에 현재의 도착률이 4.88이므로 아직까지는 서버가 용량 한계에 도달하지 않았다고 볼 수 있다.

게임서버의 경우에는 다수의 서버가 구동되고 있는데 총 고객 도착률로부터 각 서버의 평균 도착률을 구하여 0.4216(세션/초)를 얻고, 서비스율을

$$\frac{1}{1696}$$

로 구할 수 있다. 이로부터 각 서버당 제공로드는 초당 715 일량으로 얻어지고 서버의 초당 이용률은  $\rho \cong 0.3972$  정도를 얻어진다. 평균적으로는 과부하가 아니나 최빈시간의 부하를 측정된 후에 정확한 판정을 내려야 할 것이다. 손실확률이 0.05에 도달할 때의 도착률은 초당  $\lambda \cong 1.17$ 이 되고, 손실확률이 0.01일 때의 도착률은  $\lambda \cong 1.05$ 이 되므로 최빈시간의 도착률이 이 값에 도달할 때 게임 서버의 용량증설을 고려해야 할 것이다.

아래 <표 5>에 각 서버 별로 손실확률에 따른 고객도착률을 정리해 놓았다.

참고로 Little의 법칙에 의해 시스템의 평균 고객수를 구하면 현 상태에서 기업 웹서버의 평균 고객수는 약 13명(최빈시간대에는 27명)이고, 게임포털 서버의 경우 평균 고객수는 약 646명(최빈시간

<표 5> 손실확률별 고객 도착률

손실확률	제공로드	웹서버	게임포털서버	게임서버
0.01	1772	30.5517	6.3060	1.0448
0.02	1804	31.1034	6.4199	1.0637
0.03	1830	31.5517	6.5125	1.0790
0.04	1854	31.9655	6.5979	1.0932
0.05	1877	32.3621	6.6797	1.1067

대 약 1371명), 하나의 게임 서버의 경우 평균 고객수는 약 715명이다.

## 6. 결 론

본 연구에서는 웹서버의 로그자료를 분석하여 고객손실 확률을 바탕으로 인터넷 서버용량의 적정성 여부를 판정하는 간편한 방법이 제시되었다. 고객의 요구가 기각되면 고객은 서비스를 받기 위해 재시도를 하며 기다려야 하므로 결국은 지연시간이 이 방법에 반영된다고 볼 수 있다. 본 연구의 접근 방법을 적용하면 고객의 도착률 증가를 관찰하다가 서버의 적절한 증설시점을 파악하는 하나의 기준을 마련할 수 있는데, 실제적인 증설시점 및 증설 용량의 결정을 위해서는 고객수의 변화 예측은 물론 서버의 증설비용과 고객손실에 따른 고객의 손실비용 간의 이득(trade-off) 관계를 분석하여 보다 정밀한 분석이 뒤따라야 할 것이다.

본 연구는 기업내의 다양한 서버간의 이동에 따른 지연을 포함하지 못했으나 트래픽 자료가 보다 상세히 얻어질 수 있으면 대기망 모델을 통해 보다 정밀한 분석이 가능할 것이다. 이를 위해서는 서버 내의 요구신호이동 상황을 보다 상세히 기록하고 분석할 수 있는 유틸리티가 먼저 설치되어야 하는데 현 시점에서는 미비되어 추후 연구과제로 남긴다. 또한 경영적인 측면에서 인터넷 서버 자원에 대한 비용분석이 행해진다면 보다 현실적인 서버 관리 정책이 가능할 수 있으므로(Menascé, 2000), 향후 다양한 비용자료가 얻어지면 이러한 관점의 연구가 필요할 것이다.

## 참 고 문 헌

- [1] M. Crovella, C. Lindemann, M. Reiser, Internet performance modelling : the state of the art at the turn of the century, *Performance Evaluation*, 42, 2000, pp.91-108.
- [2] Dilley, J., R. Friedrich, T. Jin, J. Rolia, Web server performance measurement and modelling techniques, *Performance Evaluation*, 33, 1998, pp.5-26.
- [3] Harris, C.M., P.H. Brill, M.J. Fischer, Internet-type queues with power-tailed interarrival times and computational methods for their analysis, *INFORMS Journal on Computing*, 12, 2000, pp.261-271.
- [4] Heyman D.P., Some issues in performance modeling of data teletraffic, *Performance Evaluation*, 34, 1998, pp.227-247,
- [5] Menascé D.A., V.A.F. Almeida, R. Fonseca, M.A. Mendes, Business-oriented resource management policies for e-commerce servers, *Performance Evaluation*, 42, 2000, pp. 223-239.
- [6] Nabe, M.Murata, H.Miyahara, Analysis and modeling of WWW traffic for capacity dimensioning of internet access lines, *Performance Evaluation*, 34, 1998, pp.249-271.
- [7] Wolff, R.W., Stochastic Modeling and the Theory of Queues, Prentice-Hall, Englewood Cliffs, 1989.
- [8] Yashkov, S.F., Process-sharing queues : some progress in analysis, *QUESTA*, 2, 1987, pp.1-17.