

웹 및 시스템 로그 분석 기반 공간정보 오픈플랫폼 서비스 사용 현황 분석

Service Status Analysis About the Spatial Information Open Platform based on the Analysis of Web Server Log and System Log

장한솔* · 홍성훈** · 김민수*** · 장인성****

Han Sol Jang · Seong Hun Hong · Min Soo Kim · In Sung Jang

요약 공간정보 오픈플랫폼 서비스인 브이월드가 2012년 시범 서비스를 시작한 후부터 지금까지 많은 국민들의 관심으로 매년 사용자가 폭발적으로 증가하고 있다. 증가하는 사용자의 요구에 따라 앞으로 추가될 서비스 및 사용자의 이용 환경 개선 등의 지표로 활용하기 위해 구체적인 서비스 사용 현황을 알 필요가 있다. 하지만 현재의 시스템에는 실시간 모니터링 시스템이 구축되어 있지 않아 브이월드가 제공하는 2D/3D/Portal 서비스별 하드웨어 자원 사용량과 실 사용자들의 이용 패턴 등의 구체적인 서비스 사용 현황을 파악하는데 어려움이 있다. 이에 본 논문에서는 웹 로그를 바탕으로 2D/3D/Portal 서비스별 시스템 사용 현황 및 시스템 로그를 바탕으로 자원(CPU, Memory) 사용 현황 분석을 통해 2015년도 서비스 사용 현황을 분석하고 2014년도 대비 분석 결과를 비교하여, 현 시스템의 문제점 및 해결 방안을 제시한다.

키워드 : 공간정보 오픈플랫폼, 웹서비스 사용 현황, 웹 로그, 시스템 사용률

Abstract Since the V-World, the Spatial Information Open Platform service, has started in 2012, a lot of people have increased explosively every year with their interest. It is necessary to know the specific service status in order to serve as indicators of the improvement of user's environment and the service to be added in the future based on the user's increasing need. However, there is difficulty to figure out more specific service status, such as the usage of hardware resources for 2D / 3D / Portal services and the actual user usage patterns, because the current system does not have the real-time monitoring system. Therefore, in this paper, through the analysis of the usage of system resources for 2D / 3D / Portal services based on web server log and the usage of hardware resources such as CPU, Memory based on system log, we analyze the usage of service in 2015 and compare with the results of the 2014, to present problems of the current system and the solutions about the problems.

Keywords : Spatial Information Open Platform, Web Service Status, Web Log, System Resource Useable

1. 서 론

현재 공간정보 오픈플랫폼(이하 오픈플랫폼) 서비스는 국내외에서 매우 활성화 되어 있다. 세계적으로 Google Map, Bing Map, Openstreet Map 등이 널리 이용되고 있으며 국내에서는 Naver Map, Daum Map, Olleh Map, VWorld Map 등이 활발히 이용되고 있다. 특별히 VWorld Map(이하 브이월드)은 국토교통부에서

처음으로 국가가 보유하고 있는 공간정보를 국민이 편리하게 사용할 수 있도록 하는 개방형 오픈플랫폼 서비스로 2012년에 시범 서비스가 시작된 이래로 다양한 서비스가 추가되고 있다[18]. 전국 단위의 영상 지도, 3차원 공간데이터, 지적도, 토지이용현황, 용도 지역지구도 등 다양한 공간정보를 제공함으로써 많은 국민들로부터 관심을 받아 사용량이 급격히 늘고 있다. 늘어나는 사용량에 맞추어, 앞으로 추가될 서비스

† This research was supported by a grant(13 도시건축 A02) from Spatial Information Open Platform Infra Technology Development Research Project funded by Ministry of Land, Infrastructure and Transport government.

* Han Sol Jang, Researcher, IT Research Laboratory, ICTWAY. hsjang@ictway.co.kr

** Seong Hun Hong, Deputy Manager, Technical Dept., SNAInformation. hongsh@snainfo.com

*** Min Soo Kim, Associate Professor, Dept. of IT Convergence, Daejeon University. minsoo@dju.ac.kr (Corresponding Author)

**** In Sung Jang, Senior Researcher, Spatial Information Technology Research Team, ETRI. e4dol2@etri.re.kr

및 사용자의 이용 환경 개선 등의 지표로 활용하기 위해 구체적인 서비스 사용 현황을 알 필요가 있다. 하지만 아직 실시간 모니터링 시스템이 구축되어 있지 않아 실시간으로 서비스 사용 현황을 파악할 수 있는 방법이 없다.

이에 따라 본 논문에서는 2014년에 수행된 연구[11]로부터 약 1년이 지난 지금, 동일한 방법으로 서비스 사용 현황을 파악하여 현재의 서비스 사용 현황과 그동안 어떠한 변화가 있었는지 비교 분석하였다. 2014년에 수행된 연구와 마찬가지로 웹 로그를 분석하여 현재 시스템의 데이터 전송량, 네트워크 대역폭, 접속 수, 접속 경로에 대한 현황을 확인 및 비교하였다. 또한, 본 논문에서는 추가적으로 Portal 서비스, 2D 서비스, 3D 서비스별로 시스템 로그 분석을 통해 각각의 하드웨어 자원별 CPU와 Memory 평균 사용률을 분석하였다. 이에 향후 브이월드 시스템의 자원을 효율적으로 사용할 수 있는 방안에 대한 연구가 필요하다고 판단하였다.

본 논문의 구성은 다음과 같다. 2장의 관련 연구에서는 웹 로그 분석을 통해 얻어질 수 있는 정보에 대해 살펴볼 것이다. 3장에서는 현재 브이월드 시스템의 구성 및 웹 로그를 분석한 결과를 제시하며, 4장에서는 시스템 로그를 분석한 결과를 제시할 것이다. 끝으로 5장에서는 본 논문의 결론 및 향후 연구 방향을 제시하고자 한다.

2. 관련 연구

2.1 웹 로그 분석

Kim et al.[8]은 공개 서비스 웹 서버와 비공개 인터넷 웹 서버의 사용자 접속 로그를 이용하여 웹 서버의 부하 패턴을 분석하였다. Jeong and Lee[4]는 웹트렌즈를 이용하여 웹 서버에 저장된 로그로부터 히트 수, 페이지별 평균 사용 시간, 웹사이트 평균 사용 시간 등의 정보를 도출하여 가장 노출횟수가 많은 콘텐츠, 일일 접속자 변화량, 상위 검색 키워드 등의 정보를 분석하였다. Kwon and Zhang[10]은 웹로그 익스퍼트를 활용하여 쇼핑몰 웹사이트의 방문 데이터를 분석하였다. 히트수, 방문자수, 방문자 IP 등을 통해 접속 유휴 시간 감소, 방문자에 대한 정보 분석, 웹사이트 재구축 전략 평가 등의 결과를 도출하였다. Kang et al.[7]은 한국자료분석학회 홈페이지의 한 달 동안 수집된 웹 로그파일에 대해 방문자, 트래픽, 콘텐츠 분석을 통하여 홈페이지의 개선점을 제시하였다. Choi

et al.[2]은 2009년도 정시 모집 기간에 대학교 홈페이지에 방문하는 신입생들이 가장 많이 방문한 시기나 페이지, 지역별 방문 순위와 같은 행태를 바탕으로 차후년도의 입시전략을 기획하기 위해 웹 로그 데이터를 수집 및 분석하였다. Oh et al.[16]은 웹 사이트 방문자의 행동을 실시간으로 수집 및 분석하여 사이트의 전환율을 높이기 위한 실시간 온라인 마케팅 시스템(ROMS¹⁾)을 제안하였다. Hwang and Yoo[3]는 웹 로그 분석을 통해 방문자의 Session을 파악하고 방문자가 읽은 뉴스 기사간의 연관 규칙을 확인하여 최종적으로 규칙의 의미와 온라인 뉴스 사이트에서 고려해야 하는 의미를 제시하였다. Kim and Lim[9]은 특정 웹 사이트에 대해 약 60일간의 로그 데이터를 수집하고 분석하여 방문자 형태를 분석하고, 웹 접근성 및 품질 평가를 Google Analytics 활용하여 시행하였다. Lee[12]도 웹 이용자의 행태를 분석하고 다양한 정보요구를 파악하기 위한 수단으로 Google Analytics를 도입하는 것을 제안하고, 실제 아카이브 웹사이트에 적용한 사례를 소개하였다.

이러한 연구 결과들을 통해 앞으로 추가될 서비스 및 사용자의 이용 환경 개선 등의 지표로 활용하기 위한 목적으로 방문자 수, 접속 경로와 같은 구체적인 서비스 사용 현황을 파악하는데 웹 로그 분석 방법의 의미가 있을 것이라고 판단하였다.

2.2 시스템 로그 분석

Mendoza[13]는 AIX 기반 서버에서 SAS 환경을 구축하였을 때, 병목현상 방지를 위한 모니터링을 위해 vmstat이나 iostat과 같은 명령어 대신 NMON을 사용하는 것을 추천하였다. 구체적으로 Sharp[17]는 제안된 MDCF²⁾와 logger의 성능을 평가하기 위해 NMON을 사용하여 CPU, Memory 및 I/O 사용량을 측정하였다. Chawla[1]는 제안된 ECM³⁾ 시스템의 성능을 평가하기 위해 IBM Rational Performance Tester를 사용하여 서버의 응답 시간 등을 측정하였고, NMON을 사용하여 CPU, Memory 및 I/O 사용량을 측정하였다. Jung and Choi[5,6]는 데이터 중복 제거 파일 시스템을 제안하면서 제안 시스템에 대한 성능 분석 방법으로 NMON을 사용하여 CPU 점유율을 측정하였다.

이러한 연구 결과들을 통하여 Linux 및 AIX 기반의 브이월드 시스템의 하드웨어 자원에 대한 성능 분석

1) ROMS: Realtime Online Marketing System[16]
2) MDCF: Medical Device Coordination Framework[17]
3) ECM: Enterprise Content Management[1]

을 위해 NMON을 활용하는 것이 무리가 없다고 판단하였다.

3. 브이월드 시스템 및 웹 로그 분석

현재 브이월드 시스템은 Figure 1과 같이 웹 서비스를 제공하기 위해 WEB-WAS-DB로 구성된 3-tier 형태로 구성되어 있고, 하드웨어 사양은 Table 1과 같다 (2015. 03. 27 기준).

네트워크 장비는 모두 이중화 구성을 통해 장애에 대비하고 있으며, 사용자 분산을 통한 성능 안정화를 위해 L4 Switch의 로드 밸런싱 기능을 사용하고 있다.

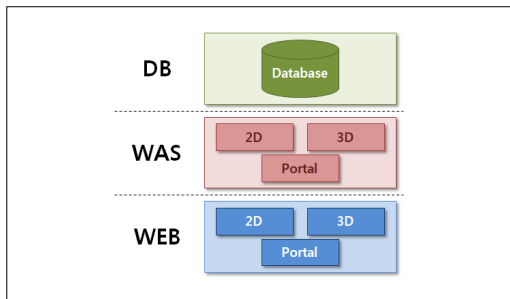


Figure 1. 3-Tier Structure of Open Platform

Table 1. Hardware Specification of Open Platform

Type	Features	Quantity
WEB Server	Portal WEB Server	6
	2D WEB Server	4
	3D WEB Server	6
WAS	Portal WAS	2
	2D WAS	3
	3D WAS	6
DB Server	DB Server	2
Etc. Server	Link Server	1
	Backup Server	1
Switch	L2 Switch	2
	L3 Switch	2
	L4 Switch	2
	L7 Switch	2
Security	WEB Firewall	2
	Firewall	2
	SSL VPN	1
Storage	Storage	1
	SAN Switch	2

특히 2014년 하반기에는 사용자 폭주로 인한 네트워크 서비스 지연 및 불가 상황을 방지하는 기능(Max Session Limit)을 지원하는 L7 Switch 도입을 통해 사용자 폭주에도 허용 한계 내의 Session 사용자는 원활하게 브이월드 서비스를 이용할 수 있도록 네트워크 시스템을 개선하였다.

브이월드 서비스에서 제공하는 다양한 서비스들을 Portal 서비스, 2D 서비스, 3D 서비스로 구분하여 제공하고 있다. 각각의 서비스는 다음과 같은 정보를 제공하고 있다[11].

- Portal 서비스: 검색, 개발자 지원 서비스, 사용자 참여 서비스 등에 해당하는 정보를 제공한다.
- 2D 서비스: 2차원 형태의 지형도 서비스, 다양한 주제도, 부동산 정보, 건축물 정보 등에 해당하는 정보를 제공한다.
- 3D 서비스: 3차원 형태의 기본 지형도 서비스, 3D 건물 콘텐츠, 고해상도 텍스처 등에 해당하는 정보를 제공한다.

본 논문에서는 브이월드 시스템의 위 3가지 서비스를 구성하고 있는 서버 자원별 평소 사용 현황을 알아보고자 하였다. 이를 위해 웹 로그 분석을 실시하였고, 2014년 연구 결과와 비교 분석하였다.

웹 로그 분석은 Figure 2와 같이 ① 웹 서버로부터 로그 파일을 수집하는 단계; ② 이를 DB 테이블 형태로 가공하는 단계; ③ 마지막으로 생성된 DB 테이블에서 SQL을 이용하여 필요한 정보를 산출하여 가시화하는 단계로 구성된다. 이러한 로그 분석을 통해서 얻을 수 있는 정보는 다음과 같다[11].

- 데이터 전송량: 서버에서 전송된 데이터 전송량을 1시간 단위로 분석함. 해당 웹 서버에서 응답한 총 전송량(단위: Byte)
- 네트워크 대역폭: 서버에서 전송된 네트워크 대역폭을 5분 단위로 분석함. 해당 웹 서버에서 응답 처리에

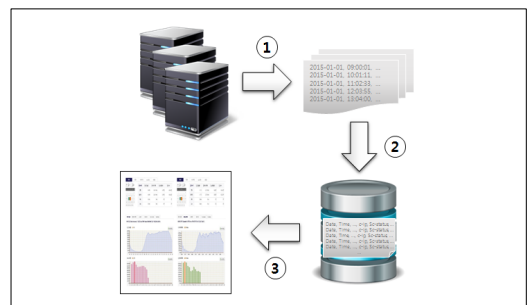


Figure 2. Procedure for Analysis of WEB Server Log

Table 2. Measurement of Analysis Data (2014 & 2015)

Day	Data Transfer (GB)	Network Bandwidth (Mbps)	Unique Visitor	Peak Hit	Day	Data Transfer (GB)	Network Bandwidth (Mbps)	Unique Visitor	Peak Hit
2014-05-20	341.89	104	116,295	35,239,384	2015-02-01	269.37	65	67,756	11,678,261
2014-05-21	343.44	85	117,756	34,363,772	2015-02-02	671.09	143	123,786	29,972,160
2014-05-22	336.13	107	116,669	32,883,771	2015-02-03	791.23	202	132,953	34,337,137
2014-05-23	303.7	75	112,508	30,288,192	2015-02-04	975.78	331	143,849	42,129,649
2014-05-24	166.64	42	78,562	16,460,755	2015-02-05	1,013.99	249	146,329	46,040,482
2014-05-25	165.07	87	68,586	15,647,693	2015-02-06	878.92	190	148,185	35,437,811
2014-05-26	327.59	92	109,648	32,540,163	2015-02-07	555.21	203	114,564	30,901,132
2014-05-27	359.83	99	117,466	35,717,849	2015-02-08	446.95	92	95,780	23,559,476
2014-05-28	330.27	108	78,176	33,216,055	2015-02-09	915.63	205	161,395	43,570,511
2014-05-29	319.04	103	76,094	33,288,330	2015-02-10	931.82	230	150,424	44,917,943
2014-05-30	310.99	104	83,984	32,503,853	2015-02-11	945.15	220	153,106	40,804,165
2014-05-31	153.74	38	51,448	15,745,453	2015-02-12	932.29	232	155,729	38,639,139
2014-06-01	137.82	41	64,255	14,001,265	2015-02-13	865.02	190	154,785	37,582,262
2014-06-02	324.06	97	129,162	33,113,140	2015-02-14	502.10	105	112,553	20,079,366

사용된 네트워크의 대역폭 사용량(단위: bps)

- 순 방문자 수: 1시간 단위로 중복 IP를 제거하며, 서비스를 사용하기 위해 방문한 사용자의 총 수를 의미함(단위: user)
- 접속 수: 모든 콘텐츠의 Hit 수를 시간 단위로 분석하여 서비스 사용 횟수를 집계 함(단위: 건수)
- 접속 경로: Hit, 데이터 전송량을 기준으로 콘텐츠에 접속한 경로를 수집하여 집계 함

Table 2는 2014년과 2015년에 수행한 연구에서 얻은 웹 로그의 데이터 전송량, 네트워크 대역폭, 순 방문자 수, 접속 수에 대한 일별 사용 현황을 나타내며, 각 항목에 대한 자세한 분석은 다음과 같다.

3.1 데이터 전송량(Data Transfer)

사용자의 요청에 의해 서버에서 전송된 데이터 전송량은 Figure 3과 같이 2014년 일일 평균 280.02GB에서 763.90GB로 483.88GB만큼 증가하였다. 이는 2014년 대비 173%의 증가율을 보인 셈이다. 또한, 주단위로 분석해보면 주말에는 전송량이 감소하고 주중에는 전송량이 증가하는 패턴을 보이고 있음을 알 수 있다.

Figure 4와 같이 일일 데이터 전송량을 시간대 별로 분석해보면 00시부터 07시까지 데이터 사용량이 감소

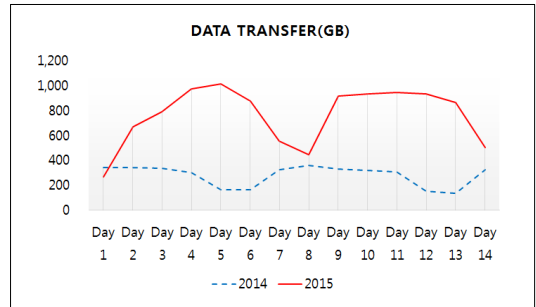


Figure 3. Avg. Data Transfer (2014 & 2015)

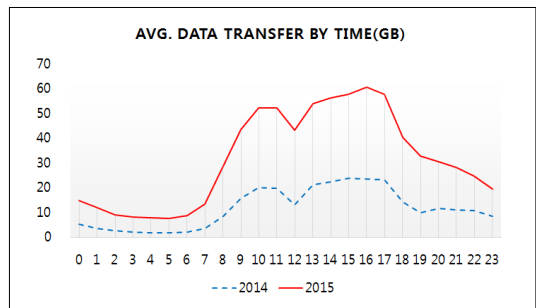


Figure 4. Avg. Data Transfer by Time (2014 & 2015)

하고 08시부터 18시까지 데이터 사용량이 증가하며 18시 이후부터는 데이터 사용량이 줄어드는 패턴을 보이는 것을 알 수 있다. 이는 평균 업무 시간대에 데

이더 사용량이 많은 것으로 확인된다. 하루 중 가장 많이 사용하는 시간대는 16~17시 사이로 집계되며 이때의 데이터 전송량은 60.56GB로써 2014년 동시간대와 비교하면 36.99GB만큼 증가한 것을 확인할 수 있다.

3.2 네트워크 대역폭(Network Bandwidth)

서버에서 전송되는 데이터의 네트워크 대역폭은 Figure 5와 같이 2014년 일일 평균 84Mbps에서 190Mbps로 106Mbps만큼 증가하였다. 이는 2014년 대비 126%의 증가율을 보인 셈이다. 또한, 데이터 전송량과 마찬가지로 주말에는 사용량이 감소하고 주중에는 사용량이 증가하는 패턴을 보이고 있음을 알 수 있다.

Figure 6과 같이 일일 네트워크 대역폭을 시간대별로 분석해보면 현재 브이월드 사용자의 서비스 집중 시간대인 평균 업무 시간대에서 1.2Gbps 이상의 사용량을 보이는 것으로 확인되었다. 향후 네트워크 사용량의 증가 및 사용자의 접속 폭주 현상에 대비할 수 있는 여유 네트워크 대역폭이 필요하다고 판단된다.

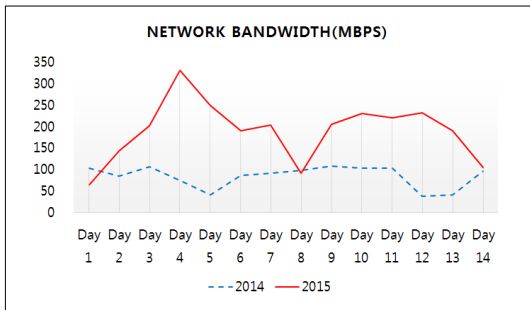


Figure 5. Avg. Network Bandwidth (2014 & 2015)

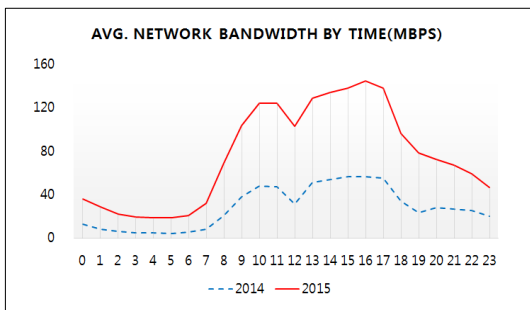


Figure 6. Avg. Network Bandwidth by Time (2014 & 2015)

3.3 순 방문자 수(Unique Visitor)

중복 IP를 제거하여 브이월드 서비스에 방문하는 순

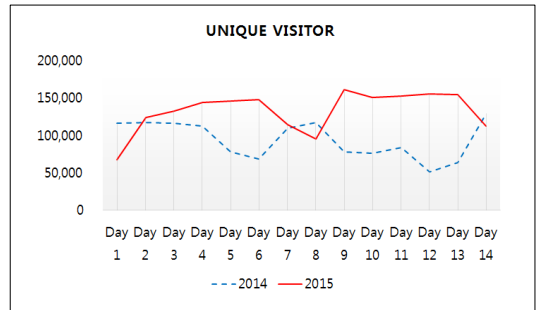


Figure 7. Avg. Unique Visitor (2014 & 2015)

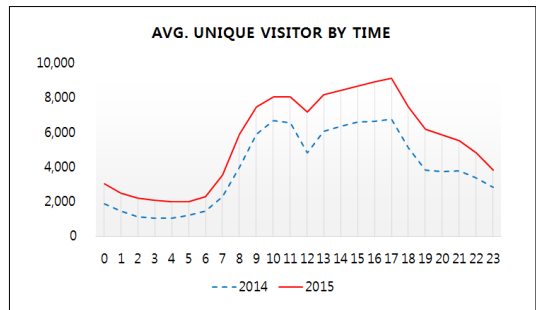


Figure 8. Avg. Unique Visitor by Time (2014 & 2015)

방문자 수는 Figure 7과 같이 2014년 일일 평균 94,329명에서 132,942명으로 38,613명만큼 증가하였다. 이는 2014년 대비 41%의 증가율을 보인 셈이다.

Figure 8과 같이 일일 순 방문자 수를 시간대별로 분석해보면 일일 평균 약 1.5배의 방문자가 방문하였고, 2014년과 유사한 패턴으로 평균 업무 시간대에 방문자가 집중되어 있는 것을 확인하였다.

3.4 접속 수(Peak Hit)

웹 로그에는 직접 사용자가 브이월드 서비스에 접속하여 서비스를 요청하는 것 외에 정보통신산업진흥원의 GeoWeb 3D Platform, 국립환경과학원의 물 환경 정보시스템, 광운대학교의 캠퍼스투어 등과 같이 외부에서 Open API를 통하여 서비스를 요청하는 것도 포함된다[19]. 이에 따른 총 접속 수는 Figure 9와 같이 2014년 일일 평균 28,214,977건에서 34,260,678건으로 6,045,701만큼 증가하였다. 이는 2014년 대비 21%의 증가율을 보인 셈이다. 또한, 주 단위로 분석해보면 2014년과 유사한 패턴을 보이고 있음을 알 수 있다.

Figure 10과 같이 일일 접속 수를 시간대별로 분석해보면 평균 업무 시간대인 08시부터 서서히 접속 수

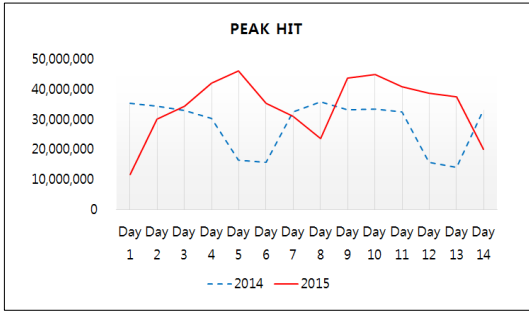


Figure 9. Avg. Hit Count (2014 & 2015)

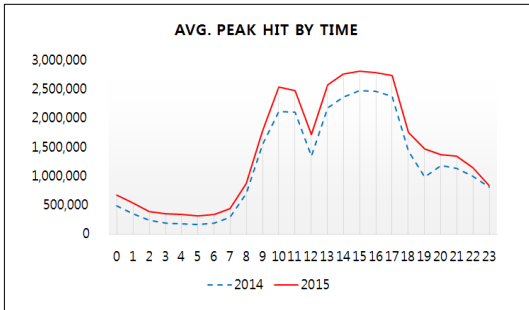


Figure 10. Avg. Hit Count by Time (2014 & 2015)

가 지속적으로 증가하다가 18시부터 감소하는 패턴을 보였다. 2014년 데이터와 비교해보면 접속 수는 증가하였으나, 사용자의 시간대별 사용패턴은 변하지 않음을 확인하였다.

3.5 접속 경로(Connection Path)

브이월드 서비스에 접속하는 경로를 파악하기 위해서 접속 수와 데이터 전송량을 기준으로 3천 건의 경로를 수집하여 상위 5가지 접속 경로에 대해 비교 분석하였다. Figure 11, 12에서 나타나듯이 접속 경로의 대부분을 차지하는 경로는 사용자가 직접 브이월드 사이트에 접속하는 `http://vworld.kr`이며, 그 외는 외부 사이트에서 Open API 형태로 제공하는 서비스에 접속한 경로이다.

특히 2015년 추가된 국토교통부에서 제공하는 “교통길잡이” 어플리케이션[14]이 상위 5가지 접속 경로에 포함되어 있는 것을 알 수 있었다. 이 서비스는 고속도로와 국도의 실시간 교통 정보, 공사 및 사고 정보 등을 확인 가능하도록 만들어졌다. 도로의 교통 상황을 파악하고 목적지까지 빠르게 가고자 하는 사용자들의 필요와 맞물려 신규 서비스임에도 사용자들이 많이 이용하는 서비스임을 확인할 수 있었다. 향후에

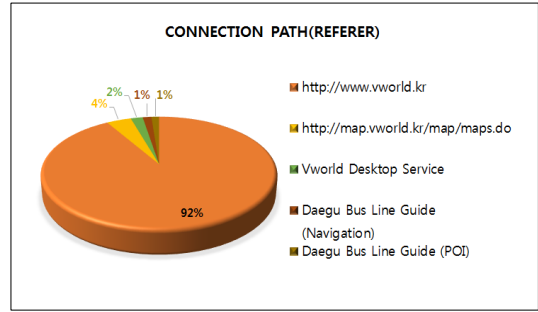


Figure 11. Top 5 of Connection Path (2014)

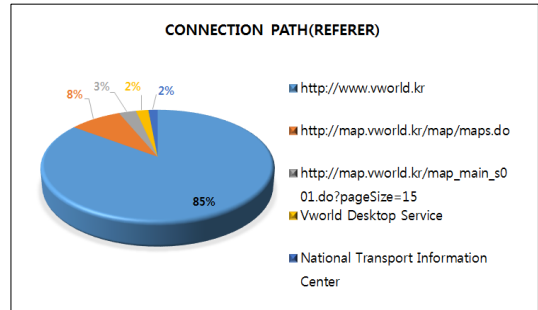


Figure 12. Top 5 of Connection Path (2015)

도 이와 유사한 서비스가 추가된다면 많은 사용자들의 관심과 활용이 예상된다.

4. 시스템 로그 분석

본 논문에서는 2014년에 수행한 연구와 달리 추가적으로 NMON[15]을 이용한 시스템 로그 분석을 통해 브이월드 서비스에서 제공하는 3가지 서비스별 하드웨어 자원의 사용 현황을 파악하였다. NMON으로 수집된 로그를 NMON Analyser를 이용하여 도출된 엑셀 형태의 결과를 정리하여 분석하였다. 분석 대상 범위는 WEB/WAS/DB 서버 29대 중에서 서버 교체 중인 2대의 3D WEB 서버를 제외하고 나머지 27대를 대상으로 수행하였다. 이러한 로그 분석을 통해서 얻을 수 있는 정보는 다음과 같다.

- CPU 사용률: 정해진 시간을 기준으로 모든 물리적 스레드의 평균 사용률을 보여줌(단위: %)
- Memory 사용률: Memory 사용률은 크게 System Memory, User Memory, File System Cache Memory 사용률을 보여줌(단위: %)

4) NMON: Nigel's performance Monitor for Linux[15]

Table 3. Avg. System Utilization of Resources

Service Name		CPU (%)	MEM (%)
WEB	Portal	0.74	32.72
	2D	1.08	6.67
	3D	0.13	13.39
WAS	Portal	10.98	69.64
	2D	5.21	46.26
	3D	2.40	75.74
DB		8.54	60.39

3가지 서비스별 하드웨어 자원에 대한 CPU와 Memory의 평균 사용률은 요약하면 Table 3과 같다. 미미한 CPU 사용률을 알아보기 쉽도록 다음에 나오는 CPU 사용률에 대한 그래프들의 Y축을 최대 20%로 제한하였다.

4.1 Portal WEB/WAS

4.1.1 CPU

Figure 13과 같이 Portal WEB 서버는 평균 CPU 사용률이 약 1% 정도로 미미한 사용률을 보이는 반면, Portal WAS는 08시부터 사용률이 점차 증가하다 18시 이후부터는 줄어드는 패턴을 보이는 것을 알 수

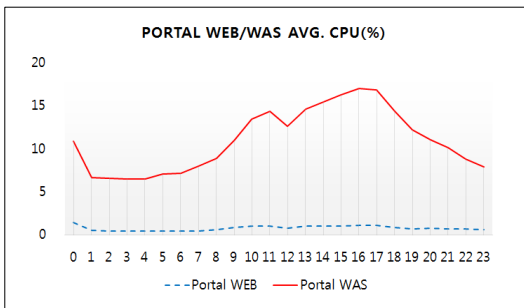


Figure 13. Avg. Portal-WEB/WAS CPU(%)

있듯이 평균 업무 시간대에 Portal 서비스의 사용량이 많은 것으로 확인하였다. 하루 중 가장 많이 사용하는 시간대인 16~17시 사이의 평균 CPU 사용률은 약 17% 정도로 안정적인 서비스가 가능한 상태로 확인하였다.

4.1.2 Memory

Figure 14와 같이 Portal WEB 서버는 평균 Memory 사용률이 약 35%를 유지하는 안정적인 상태로 나타났으며, Portal WAS는 평균 Memory 사용률이 약 70%를 유지하는 안정적인 상태로 나타났으나 추후 서비스 추가로 인해 Memory 부족 현상을 미연에 방지하고자 Memory 증설을 고려하는 것이 좋다고 판단하였다.

4.2 2D WEB/WAS

4.2.1 CPU

Figure 15와 같이 2D WEB 서버는 평균 CPU 사용률이 약 1% 정도로 미미한 사용률을 보이는 반면, 2D WAS는 08시부터 20시까지 사용률이 증가하는 현상을 볼 수 있다. 이는 브이월드 서비스가 제공하는 Open API 등의 외부 링크 서비스 중 많은 수가 공공 서비스 임을 고려한다면 의미 있는 분석 결과라고 할 수 있다.

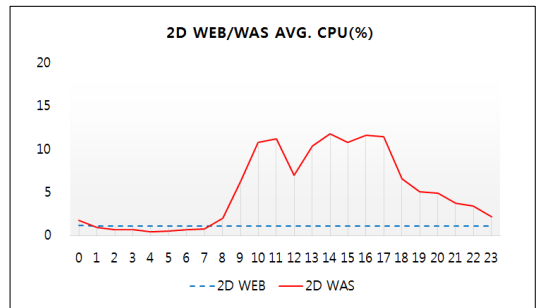


Figure 15. Avg. 2D-WEB/WAS CPU(%)

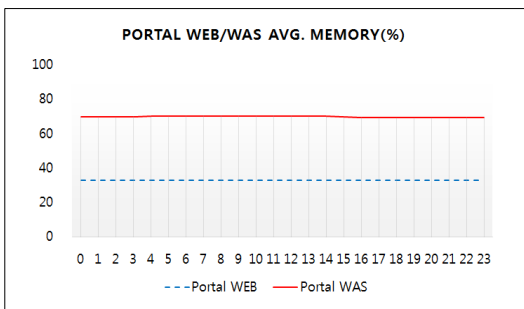


Figure 14. Avg. Portal-WEB/WAS Memory(%)

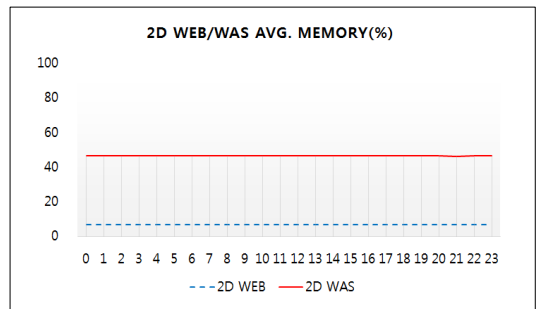


Figure 16. Avg. 2D-WEB/WAS Memory(%)

4.2.2 Memory

Figure 16과 같이 2D WEB 서버는 평균 Memory 사용률이 약 7%를 유지하는 매우 안정적인 상태로 나타났으며, 2D WAS는 평균 Memory 사용률이 약 50%를 유지하는 안정적인 상태로 나타났다. 2D WEB 서버에 비해 2D WAS의 평균 Memory 사용률이 높게 나타나는 것은 지적도, 산사태 위험지도 등과 같은 사용자의 요청에 따라 실시간으로 2D WAS에서 지도 이미지 생성 작업을 수행하기 때문으로 판단된다.

4.3 3D WEB/WAS

4.3.1 CPU

Figure 17과 같이 3D WEB 서버는 평균 CPU 사용률이 약 1% 정도로 미미한 사용률을 보이는 반면, 3D WAS는 사용률이 높진 않지만 2D 서비스와 마찬가지로 평균 업무 시간대인 08시부터 18시까지 사용률이 집중 분포되어 있는 것을 볼 수 있다.

4.3.2 Memory

Figure 18과 같이 3D WEB 서버는 평균 Memory 사용률이 약 15%를 유지하는 안정적인 상태로 나타났으며, 3D WAS는 평균 Memory 사용률이 약 75%를

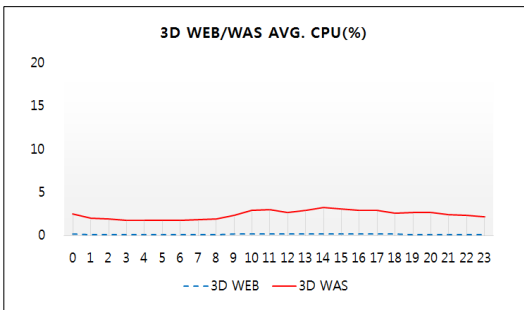


Figure 17. Avg. 3D-WEB/WAS CPU(%)

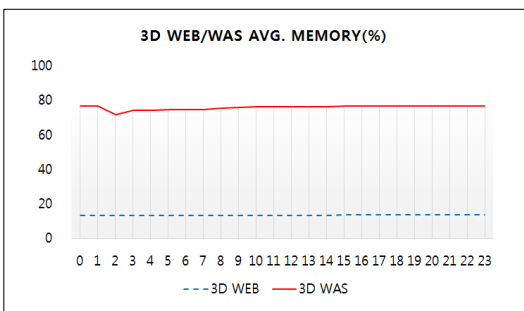


Figure 18. Avg. 3D-WEB/WAS Memory(%)

유지하는 안정적인 상태로 나타났다.

3D WEB 서버에 비해 3D WAS의 평균 Memory 사용률이 높게 나타나는 것은 브이월드 서비스의 특성상 2D 배경지도와 3D 서비스가 DB에 저장된 데이터를 검색하여 서비스하는 방식이 아니라, 자주 변하지 않는 지도 데이터들을 파일 형태로 저장하고 서비스하기 때문이다. 따라서 3D 서비스를 위한 데이터뿐만 아니라 2D 서비스의 배경 지도 데이터를 3D WAS에 함께 두는 구조가 되었고, 사용자들이 이러한 데이터들을 자주 이용하기 때문에 당연히 3D WAS에 잦은 호출이 발생하고 그에 따라 Memory 사용률이 높아질 수밖에 없다. 언론 홍보 등과 같이 특별한 이슈가 있어 평소보다 많은 사용자의 유입이나, 추후 서비스 추가로 인해 Memory 부족 현상을 미연에 방지하고자 Memory 증설을 고려하는 것이 좋다고 판단하였다.

4.4 DB

4.4.1 CPU

Figure 19와 같이 DB 서버는 평균 CPU 사용률이 높진 않지만 Portal 서비스와 유사한 패턴으로 평균 업무 시간대인 08시부터 18시까지 사용률이 집중 분포되어 있는 것을 볼 수 있다. 주로 이 시간대에 많은 사용자들이 Open API 등의 외부 링크 서비스들을 많이 활용하는 것으로 판단할 수 있다. 이러한 서비스들은 DB 서버에서 사용자가 원하는 특정 지역의 좌표나 부동산 실거래가 등과 같은 속성 정보를 검색하는 기능을 제공하기 때문이다.

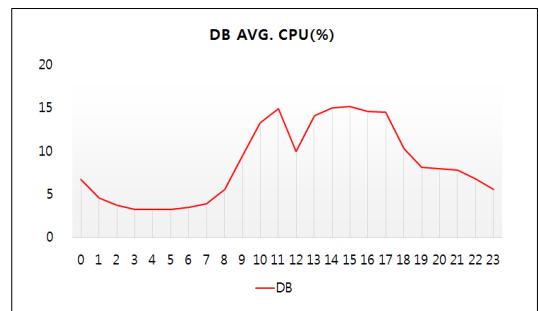


Figure 19. Avg. DB CPU(%)

4.4.2 Memory

Figure 20과 같이 DB 서버는 평균 Memory 사용률이 약 60%를 유지하는 안정적인 상태로 나타났다.

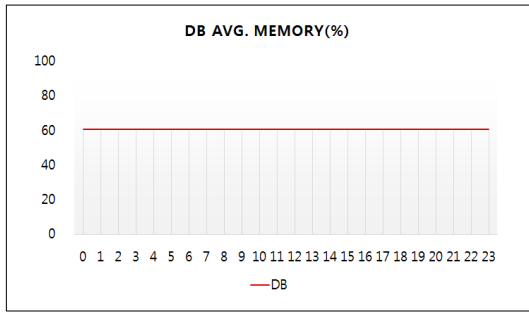


Figure 20. Avg. DB Memory(%)

5. 결론 및 향후연구

본 논문에서는 웹 및 시스템 로그 분석을 통해 브이월드 서비스의 사용 현황을 파악하였다. 현재의 브이월드 시스템은 실시간 모니터링 시스템의 부재로 인해 본 논문에서 사용한 웹 및 시스템 로그 분석 방법은 실시간 서비스 사용 현황을 파악하기에는 부족한 점이 있다. 하지만 2014년에 수행한 연구 결과 대비 2015년 현재 브이월드 서비스별 사용 현황을 비교함으로써 브이월드 시스템의 향후 하드웨어 증설 및 자원 재조정에 필요한 적절한 방향성을 제시할 수 있는 지표를 얻기에는 충분하다고 본다. 웹 로그 분석 결과 2014년 대비 데이터 전송량은 173%, 네트워크 대역폭은 126%, 순 방문자 수는 41%, 접속 수는 21% 증가한 것을 확인하였다. 이는 브이월드 서비스의 사용자가 지속적으로 계속 증가하고 있음을 나타낸다고 할 수 있다. 또한, 브이월드 서비스의 콘텐츠를 Open API 형태로 이용하고 있는 국토교통부의 “교통길잡이” 모바일 어플리케이션 이용자 또한 뚜렷하게 증가한 것을 볼 수 있다. 이를 통해 2016년에는 더욱 많은 사용자가 브이월드 서비스를 이용 할 것으로 예측 가능하다.

시스템 로그 분석 결과 높은 사용률을 나타내는 WAS에 반해 WEB 서버의 사용률은 낮게 나타나는 것을 볼 수 있었다. 이를 극복하기 위하여 향후 브이월드 시스템의 자원을 효율적으로 사용할 수 있는 서버 가상화와 같은 아키텍처 조정이나 물리적인 자원 재조정 방안 등에 대한 연구가 필요하다고 판단된다. 뿐만 아니라 2D 서비스와 3D 서비스의 데이터가 혼재되어 있는 상황이어서 두 서비스의 정확한 데이터 구분이 어려운 점이 있다. 이에 각 서비스별로 데이터를 분리 적용한 아키텍처로 구현 할 필요가 있다고 판단하였다. 향후 연구에서는 Google Analytics와 같이 실시간으로 모니터링 할 수 있는 시스템을 통해 브이월드

서비스의 사용 현황을 파악할 필요가 있다고 판단된다.

References

- [1] Chawla, G. 2014, Optimizing the Resource utilization of Enterprise Content management workloads through measured performance baselines and dynamic topology adaptation, Master’s Thesis, University of Stuttgart, Stuttgart, Germany.
- [2] Choi, S. B; Kang, C. W; Cho, J. S. 2009, Behavior analysis of entrance applicants using web log data, Journal of the Korean Data Analysis Society 20(3):493-504.
- [3] Hwang, H. S; Yoo, K. D. 2013, Mining Association Rules from the Web Access Log of an Online News website, Journal of the Korea Industrial Information System Society 18(2):47-57.
- [4] Jeong, S. K; Lee, C. W. 2003, Web log Data Analysis Apply to Web Contents Analysis Result data, Proceedings of the Korean Information Science Society Conference, 30(1):579-581.
- [5] Jung, S. O; Choi, H. 2014, Performance analysis of the virtualized data deduplication file system, Paper presented at the Korea Computer Congress 2014, Busan, June 25-27.
- [6] Jung, S. O; Choi, H. 2014, Performance Analysis of Open Source Based Distributed Deduplication File System, KIISE Transactions on Computing Practices, 20(12):623-631.
- [7] Kang, C. W; Kim, K. K; Kim, T. W. 2009, Weblog Analysis of The Korean Data Analysis Society’s Website, Journal of the Korean Data Analysis Society 11(1):361-370.
- [8] Kim, K. H; Park, N. S; Choi, E. H; Ahn, G. S. 1998, Analysis of Web Server Load Patterns from Web Log File, Proceedings of the Korean Information Science Society Conference, 25(1):134-136.
- [9] Kim, D. L; Lim, Y. H. 2014, Research in the Direction of Improvement of the Web Site Utilizing Google Analytics, Journal of the Korean Society of Cartoon and Animation Studies 36(1):553-572.
- [10] Kwon, Y. J; Zhang, H. W. 2009, Web Log Analysis Model, Paper presented at the Conference of Korea Society Industrial Information System

2009, Daegu, May 8.

- [11] Lee, S. H; Cho, T. H; Kim, M. S. 2014, A Study on the Service Status of the Spatial Open Platform based on the Analysis of Web Server User Log: 2014.5.20.~2014.6.2. Log Data, Journal of Korea Spatial Information Society 22(4):67-76.
- [12] Lee, H. E. 2015, A study of user behavior of archive using web analytics, Master's Thesis, Myongji University, Seoul, South Korea.
- [13] Mendoza, A. 2008, Using nmon to Monitor SAS Applications on AIX Servers, Paper presented at the SAS Global Forum 2008, Texas, March 16-19.
- [14] National Transport Information Center, Accessed April 3. <http://m.its.go.kr:8089/map/v-world.do>.
- [15] NMON, Accessed March 24. <http://nmon.sourceforge.net/pmwiki.php>.
- [16] Oh, J. H; Kim, J. H; Kim, J. W. 2011, A Study on the Development of Realtime Online Marketing System Using Web Log Analytics, Journal of Society for e-Business Studies 16(3):249-261.
- [17] Sharp, J. 2014, A Logging Service as a Universal Subscriber, Master's Thesis, Kansas State University, Manhattan, Kansas, United States.
- [18] VWorld, Accessed March 26. <http://www.vworld.kr>.
- [19] VWorld: case of application, Accessed March 26. http://www.vworld.kr/po_apiappl_s001.do.

Received : 2015.04.08

Revised : 2015.06.11

Accepted : 2015.06.18