

웹 사용자 로그 분석 기반 공간정보 오픈플랫폼 서비스 사용현황 연구: 2014.5.20.~2014.6.2. 수집자료 대상 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

이승한* · 조태현** · 김민수***

Seung Han Lee · Tae Hyun Cho · Min Soo Kim

요약 최근 IT, 모바일 기술의 발전으로 공간정보는 국민생활과 국가경제의 인프라 역할을 수행하고 있으며, SNS·소셜커머스를 비롯한 많은 어플리케이션에서 공간정보를 활용하여 다양한 서비스를 제공하고 있다. 이와 관련하여 국가 공간정보 서비스를 안정적으로 제공할 수 있는 공간정보 오픈플랫폼 서비스(VWorld)의 등장에 사람들의 관심이 크게 증가하고 있다. 그러나 현재 공간정보 오픈플랫폼 시스템은 실시간 모니터링 시스템이 도입되지 않은 상태로 오픈플랫폼의 실시간 부하 및 사용 현황을 파악하는데 많은 어려움이 있다. 이에 본 연구에서는 오픈플랫폼 시스템의 웹 사용자 로그정보를 이용하여 데이터 전송량, 네트워크 점유율, 서비스 방문자 수, Hit 건수, 접속 경로, 주요 이용 콘텐츠와 관련된 사용현황을 분석하고 그 결과를 제시하고자 한다. 본 연구에서 제시한 결과가 오픈플랫폼 서비스의 모든 사용 현황을 파악하기에는 부족한 면이 있지만, 해마다 변화하는 오픈플랫폼의 사용현황을 파악하기 위한 기본 자료로, 또한 오픈플랫폼 시스템 확장에 있어서의 기초자료로서 활용이 가능할 것으로 기대된다.

키워드 : 공간정보, 오픈플랫폼, 브이월드, 웹서비스, 사용현황, 웹로그

Abstract Recently, through the development of IT and mobile technology, spatial information plays a role of infrastructure of the people life and the national economy. Many kinds of applications including SNS and social commerce is to leverage the spatial information for their services. In the case of domestic, spatial open platform that can provide national spatial data infrastructure services in a stable manner has been released. And many people have been interested to the open platform services. However, the open platform currently has many difficulties to analyze its service status and load in real time, because it does not hold a real-time monitoring system. Therefore, we propose a method that can analyze the real-time service status of the open platform using the analysis of the web server log information. In particular, we propose the results of the analysis as follows: amount of data transferred, network bandwidth, number of visitors, hit count, contents usage, and connection path. We think the results presented in this study is insufficient to understand the perfect service status of the open platform. However, it is expected to be utilized as the basic data for understanding of the service status and for system expansion of the open platform, every year.

Keywords : Spatial Information, Open Platform, VWorld, Web Service, Service Status, Web Log

1. 서론

고해상 3차원 지도를 기반으로 다양한 국가공간정보를 전 국민에게 웹 기반으로 제공하고 오픈 API(Application Programming Interface)를 통해 누구나 손쉽게 3차

원 공간정보를 활용할 수 있도록 하는 개방형 서비스 모델의 공간정보 오픈플랫폼(이하 오픈플랫폼) 서비스가 국토교통부 주도하에 출시되었다[19]. 현재 오픈플랫폼은 브이월드(VWorld)라는 이름으로 2012년 1월 시범 서비스가 시작된 이후, 지금까지 다양한 기능과

† 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

* Seung Han Lee. Senior Researcher. IT Research Laboratory. ICTWAY. shlee@ictway.co.kr (Corresponding Author)

** Tae Hyoun Cho. Senior Researcher. IT Research Laboratory. ICTWAY. thcho@ictway.co.kr

*** Min Soo Kim. Principal Researcher. Spatial Information Technology Research Team. ETRI. minsoo@etri.re.kr

기술이 추가되고 있으며 사람들의 관심도 급격히 높아지고 있다. 오픈플랫폼 서비스는 그 동안, 2012년 2월 G20 핵 안보 정상회의 경호업무 서비스를 시작으로 여수엑스포 영상지도서비스, 전국 토지정보서비스, 북한 주요지역 영상지도서비스, 6대 광역시 3차원 지도서비스, 독도 및 남극 3차원 지도서비스, 산사태 위험지도 서비스 등 사회적 이슈를 반영한 다양한 서비스와 행정정보를 활용한 서비스를 제공함으로써 온라인 접속 건수의 증가뿐만 아니라, 분야별 활용도도 크게 증가하고 있다.

이와 같이 오픈플랫폼 서비스는 이용 건수 및 활용 증가에 따라 오픈플랫폼 시스템의 용량 증대, 신규 서비스 제공을 위한 기술 개발 및 도입, 또한 사용자 확장에 따른 네트워크 및 보안 이슈 등과 같은 다양한 기술지원이 추가적으로 필요하게 되었다. 이에 국토교통부는 오픈플랫폼의 지속적인 고도화를 위해 중장기 로드맵을 수립하고, 이에 따라 매년 오픈플랫폼의 서비스 추가 및 확장, 데이터 유통 확대, 사용자 참여 확대, 시스템 용량 증설 등의 업무를 추진하고 있다. 특히, 오픈플랫폼 시스템의 경우 해마다 하드웨어 용량 증설이 이루어지고 있으며 현재는 최초 시범 시스템 대비 WEB 서버의 경우 8배, WAS(Web Application Server)의 경우 2배, 기타 DB(Data Base) 및 네트워크 용량의 경우 2배 이상 증가된 상황에 이르고 있다.

그러나, 이러한 지속적인 오픈플랫폼 시스템의 확장에도 불구하고 신규 서비스 오픈에 따른 언론 홍보 등의 새로운 이슈에 따라 사용자 접속이 폭주하게 되고, 이에 따라 사용자에게 원활한 서비스를 제공하는데 문제가 빈번히 발생하였다. 예를 들어, 2012년 9월에 북한지역의 3차원 위성영상 신규 서비스를 홍보하면서 접속이 폭주함에 따라 문제가 발생하였으며 그 긴급 조치로 전산서버를 증설하고 시스템 운영 요원의 비상근무 체계를 가동하였다[10]. 2013년 9월에는 북한 주요지역 고해상도 영상서비스 제공에 따라 9월 29일부터 9월 30일까지 약 40만 명이 일시적으로 접속하여 서비스가 원활하지 못하였으며, 이를 위하여 사용자 접속속도, 사용자 요구 처리속도 및 3차원 데이터를 포함한 모든 데이터 처리속도를 2배 이상 보강하고 사용자 폭주에 대비한 통신장비를 보강하는 조치를 급하게 취하였다[11].

웹 서비스 제공 시스템에서 종종 발생하는 접속 폭주로 인한 서비스 지연 문제는 해당 시스템에 큰 영향을 미치게 되고 복구와 보완 작업이 이루어지기 전까지 정상 서비스가 불가능하게 된다. 특히, 오픈플랫폼의 경우 민간과 공공기관에서 필요한 데이터와 서비

스를 연계하여 사용하기 때문에 오픈플랫폼 시스템에 장애가 발생할 경우, 그 영향이 오픈플랫폼에서 제공하는 웹 서비스의 단순 지연뿐만 아니라 오픈플랫폼을 활용하는 다른 모든 서비스에 영향을 주게 되므로 매우 심각한 문제를 야기 시키게 된다. 이에 웹 서비스 제공 시스템은 용량 증설을 대비하기 위하여 기존 서비스 사용 현황을 근거로, 추가로 제공해야 할 서비스 용량을 미리 산정할 수 있어야 한다. 이를 위해서는 현 시스템의 수용 가능한 용량과 서비스 제공에 따른 점유 용량 등을 정확히 파악할 수 있어야 한다. 예를 들어, 시스템이 수용할 수 있는 동시 접속자 수 대비 현재 접속자 수, 또는 지원 가능한 네트워크 대역폭 대비 일일 네트워크 대역폭 사용량 등을 파악할 수 있어야 한다.

그러나, 오픈플랫폼의 경우 기 구축된 시스템 구성의 현황 파악은 가능하지만, 실시간으로 운영 상태를 모니터링 할 수 있는 시스템이 구축되어있지 않아, 실시간 사용 현황 분석을 위해 필요한 정보를 파악하는데 많은 어려움을 가지고 있다. 이에 본 논문에서는 웹 사용자 로그정보를 이용하여 현재 오픈플랫폼 시스템의 사용 현황을 명확히 파악하고 분석하여 현 시스템의 미래사용 현황을 예측할 수 있는 자료를 제시하고자 한다. 구체적으로 데이터 사용량, 네트워크 대역폭, 서비스 방문자 수, 서비스 Hit 수, 콘텐츠 사용량, 접속 경로에 대한 분석 결과를 제시하고자 한다.

본 논문의 구성은 다음과 같다. 2장의 관련 연구에서는 오픈플랫폼과 웹 기반의 사용 현황 파악 및 분석에 대한 방법에 대해 살펴볼 것이다. 3장에서는 본 논문에서 제안하는 오픈플랫폼 시스템과 현황 파악을 위한 웹 로그 수집 및 분석 방안에 대해 살펴보고, 4장에서는 오픈플랫폼 시스템 현황에 대한 분석 및 평가 결과를 제시할 것이다. 끝으로 5장에서는 본 논문의 결론 및 향후 연구 방향을 제시하고자 한다.

2. 관련 연구

현재 오픈플랫폼 서비스는 국내·외에서 매우 활성화 되어 있다. 세계적으로 Google Map, Bing Map, OpenStreetMap 서비스가 널리 이용되고 있으며, 국내에서도 다음맵, 네이버맵, 올레맵, 브이월드맵 등이 활발히 이용되고 있다. 이러한 오픈플랫폼은 웹 GIS의 플랫폼 형태로 2차원 지도서비스, 3차원 지도서비스, 항공사진, 위성영상 등의 다양한 정보와 길 찾기, 버스·지하철 노선 검색 등의 분석 서비스뿐만 아니라, 오픈 API를 제공하여 지도 관련 정보를 사용자들이

다른 응용에서 자유롭게 이용할 수 있도록 다양한 서비스를 제공하고 있다[4,7].

오픈플랫폼을 지원하는 기반 기술인 웹 시스템 기술도 빠른 속도로 발전하고 있으며, 이러한 웹 시스템은 클러스터 시스템(Clustering System) 기술, CDN(Contents Delivery Networking) 기술, Cloud 기술 등의 도입으로 점점 대형화 되고 있다[8,14,16,17]. 또한 웹 시스템을 운영, 관리하기 위한 실시간 모니터링 기술 또한 크게 발전하고 있다. 실시간 웹 모니터링 시스템은 웹 서버나 클러스터링 장비를 단순히 감시하는 기능뿐만 아니라 시스템에서 지원하는 서비스에 대해 실시간 분석을 수행할 수 있는 모니터링 기능도 포함하고 있다. 최근에는 개인 블로그 및 SNS 등의 활성화로 인하여 소규모 개인 웹사이트에 대한 실시간 모니터링 기능 서비스를 제공하는 사이트도 증가하고 있는 추세이다[5].

현재 웹 모니터링 시스템 구축과 관련하여 웹에 산재한 다양한 정보 제공에 대한 분석 과정에서 해당 서비스 페이지가 호출 될 때 분석에 필요한 정보를 전달하기 위하여 해당 페이지에 간단한 자바스크립트 형태의 소스코드를 삽입하여 정보를 수집하는 방법이 있으며, 또한 로봇 형태의 에이전트를 추가하여 정보 검색, 필터링, 통합, 추출 기능을 제공하도록 하는 웹 변경 탐지를 위한 에이전트를 추가하는 방법 등의 기술이 연구되고 있다[6,15].

오픈플랫폼 시스템은 대용량 고해상도 지도 정보 서비스를 제공하고 2차원 지도, 3차원 지형, 3차원 건물 등의 다양한 콘텐츠를 제공하는 시스템으로 현재 클러스터링 기술 기반으로 여러 대의 WEB 서버와 WAS로 구성되어 있다. 이와 같이 대용량의 다양한 정보를 서비스해야 하는 웹 시스템은 반드시 실시간 모니터링 시스템을 통해서 사용 현황을 파악함으로써 운영, 관리 및 유지보수를 효율적으로 수행할 수 있어야 한다[3]. 그러나 현재 오픈플랫폼은 이와 같은 모니터링 시스템을 전혀 지원하지 않고 있으며, 더구나 2012년 시범 서비스 시작 이래로 시스템 사용 현황 분석을 위한 정보 및 분석 결과가 전무한 상태이다. 이에 본 연구에서는 웹 사용자 로그 분석 결과를 이용하여 오픈플랫폼 시스템의 명확한 사용 현황을 파악하고자 한다.

3. 오픈플랫폼 시스템 및 사용현황 분석 방법

본 장에서는 오픈플랫폼 성능 분석을 위해 오픈플랫폼 시스템의 구성 및 서비스 환경에 대하여 소개하

고 웹 사용자에게 대한 정보 분석을 위하여 필요한 WEB 서버의 로그 분석 방법에 대하여 제시하고자 한다.

3.1 오픈플랫폼 시스템 구성

3.1.1 오픈플랫폼 시스템 구성

현재 오픈플랫폼의 시스템 구성은 Figure 1과 같다. Figure 1에서 오픈플랫폼의 장비 구성은 네트워크 장비, WEB 서버, WAS, 기타 서버(연계, 백업, 로그 등을 담당), SAN(Storage Area Network) Storage, DB 시스템으로 구성된다.

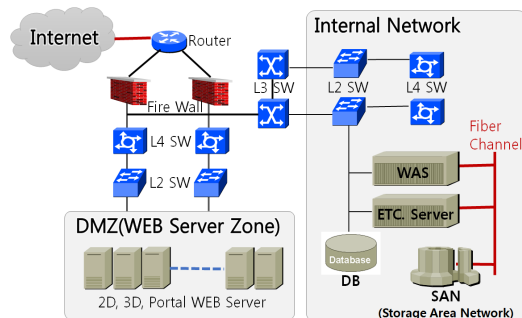


Figure 1. System Architecture of Open Platform

Table 1. Hardware Specification of Open Platform

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

네트워크 장비는 모두 이중화 구성을 통해 장애에 대비하고 있으며 L4 Switch 구성으로 부하 분산(Load Balancing) 기능을 사용하여 WEB 서버 접속자에 대한 분산 기능을 수행하고 있다. WEB 서버는 서비스 기능 별로 오픈플랫폼 포털, 2D 지도, 3D 지도 기능으로 분류되어 있으며 각 기능별로 WAS와 연동을 통해 서비스를 제공하고 있다. 또한 WEB 서버는 기능별로 L4 Switch를 통해 클러스터링 형태로 구성되어 있으며 WAS 또한 각 기능별로 WEB과 L4 Switch를 통해 Load Balancing 될 수 있도록 구성되어 있다. 오픈플랫폼에서 제공하는 서비스 중에서 통합 지도 서비스는 대용량 콘텐츠가 필요하므로 SAN 형태로 구성된 대용량 Storage를 통해 안정적인 콘텐츠 제공을 하고 있다. DB 시스템은 고가용성(HA:High Availability) 구성을 통해 이중화 및 장애에 대응하도록 구성되어 있으며 사용자 정보 및 각종 콘텐츠에 해당하는 정보를 저장하고 있다. Table 1은 이러한 하드웨어 구성에 대한 상세 정보를 보여준다[20].

3.1.2 오픈플랫폼 웹 서비스의 3 계층 구조

오픈플랫폼에서 제공하는 서비스는 웹 기반 서비스로 클라이언트는 PC의 웹 브라우저 또는 모바일 앱에서 서비스를 이용할 수 있다. 이러한 웹 서비스를 제공하기 위하여 오픈플랫폼 시스템은 Figure 2에서와 같이 WEB 계층, 응용 계층(WAS), DB 계층이 물리적으로 구성된 3 계층(3-tier) 형태로 구성되어 있다. 이러한 3 계층 구조에서 각 계층별 역할은 다음과 같이 정의되어 있다[18].

- **프리젠테이션 계층(WEB 계층):** 오픈플랫폼 웹 서비스와 사용자 인터페이스로 구현되며 브라우저를 이용한 HTML, XML 등에 대한 웹 서비스를 지원하는 계층

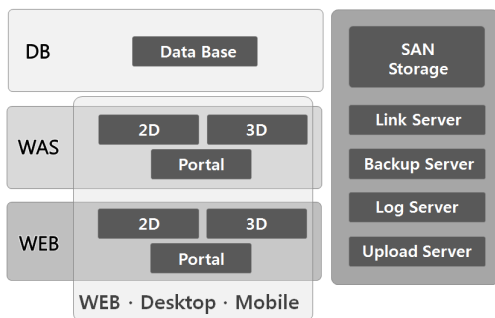


Figure 2. 3-Tier Structure of Open Platform

- **어플리케이션 계층(WAS 계층):** 오픈플랫폼 비즈니스(서비스) 로직을 구현하기 위한 어플리케이션을 지원하는 계층으로 웹 서비스에서는 WAS가 그 역할을 지원함
- **데이터 계층(DB 계층):** 데이터베이스 서버와 메인 프레임으로 구성되고 오픈플랫폼 서비스에 필요한 데이터, 콘텐츠 등 정보를 저장 및 가공할 수 있는 계층

또한, Figure 2에서 보는 바와 같이 각각의 WEB 서버와 WAS는 오픈플랫폼에서 제공하는 통합지도 서비스, 개발자 지원 서비스, 데이터 제공 서비스, 사용자 참여 서비스 등의 다양한 서비스를 Portal, 2D, 3D로 구분하여 제공하고 있다.

- **Portal 서비스:** 개발자 지원 서비스, 데이터 제공 서비스, 사용자 참여 서비스 등의 다양한 정보를 제공하는 서비스로, 기본 2D 배경지도, 3D 지형도 서비스 외에 다양한 검색, 3D 건물 콘텐츠, 다양한 주제도, 부동산 정보, 건축물 정보 등에 해당하는 정보를 제공
- **2D 서비스:** 2차원 형태의 지형도 서비스 외에 다양한 검색, 다양한 주제도, 부동산 정보, 건축물 정보 등에 해당하는 정보를 제공
- **3D 서비스:** 3차원 형태의 기본 지형도 서비스 외에 다양한 검색, 3D 건물 콘텐츠, 고해상도 텍스처 등에 해당하는 정보를 제공

3.2 오픈플랫폼 웹 사용자 정보 분석

본 연구에서는 제안된 오픈플랫폼 서비스의 사용 현황을 분석하기 위하여 WEB 서버의 로그 분석 방법을 사용하였다. 이는 현재 오픈플랫폼 시스템이 실시간으로 서비스 사용 통계를 집계할 수 있는 모니터링 시스템을 보유하고 있지 않기 때문이다. 예를 들면, Google Analytics와 같은 웹 서비스 모니터링 기능을 사용하기 위해서는 오픈플랫폼의 응용프로그램에 Java Script 구문을 삽입하여야 하는데 이를 위해서는 상용 동작 중인 오픈플랫폼의 응용프로그램에 추가 개발하는 과정이 필요하다. 그래서 기존 시스템의 변경 없이 서비스 사용현황을 파악할 수 있는 로그 분석 방법을 선택하였다.

WEB 서버 로그 분석을 통하여 사용자에 대한 통계를 집계하는 절차는 Figure 3과 같이 WEB 서버를 통해 텍스트 형태의 로그 파일 정보를 수집하는 단계,

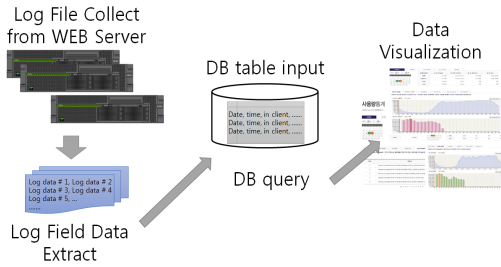


Figure 3. Procedure for Analysis of WEB Server Log

Table 2. Type and Description of WEB Server Log Field

Field	Description	Example
Date	Date of client request	2014-03-11
Time	Time of client request	10:00:03
s-computer name	WEB Server name. Separate server name for redundant server	WIN-SMV7T CHPPAL
Cs-uri-stem	Client request page or file URL - This means contents	/default.asp
c-ip	Client IP address of request	211.58.184.39
Cs (referer)	Refer page URL. Requested to the previous page in the client	-
Cs-host	The host server to process the request. (Domain information)	center01.gw.p ubedge.com
Sc-status	HTTP return code for the request	302, 404, 200
Sc-bytes	The total response for a request for a client network traffic	645
Time-taken	The total elapsed time means the time it takes for the request client	202

이를 DB 테이블 형태로 가공하는 단계, 그리고 생성된 DB 테이블에서 DB Query를 이용하여 필요한 결과를 산출하고 가시화 하여 통계 자료로 사용하는 단계로 구성된다.

오픈플랫폼은 현재 16대의 WEB 서버가 운용 중인 상태로 각 WEB 서버의 로그 기록을 통합 분석함으로써 본 연구에서 제시하고자 하는 서비스 사용자 현황을 보다 정확하게 집계할 수 있다. 현재 로그는 각 WEB 서버에서 파일 단위로 저장되며 해당 로그에서

필요한 데이터는 Table 2와 같이 WEB 서버 로그 필드를 참조하여 Configuration 값을 조정함으로써 WEB 서버로 인입하는 클라이언트 요청들에 대하여 필요한 로그 정보를 기록할 수 있다. 또한 오픈플랫폼 WEB 서버 16대에서 각각의 로그 파일을 수집하여 분석을 수행하여야 한다.

로그 파일의 분석은 텍스트 파일 형태로 생성된 로그 파일의 텍스트 정보를 각각의 필드에 맞게 분석을 수행하기 위하여 미리 구성된 DB 테이블에 입력하고 그 분석 결과를 엑셀 또는 그래프 형태로 가공하여 표출한다.

이러한 로그 파일 수집을 통해 얻을 수 있는 정보는 데이터 전송량, 네트워크 대역폭, 서비스 방문자 수, 접속 수(Hit 수), 접속 경로 등이며 이에 대한 상세한 정의는 다음과 같다.

- **데이터 전송량:** 서버에서 전송된 데이터 전송량을 1시간 단위로 분석함. 해당 웹 서버에서 응답한 총 전송량을 의미함 (단위: Byte)
- **네트워크 대역폭:** 서버에서 전송된 네트워크 대역폭을 5분 단위로 분석함. 해당 웹 서버에서 응답 처리에 사용된 네트워크의 대역폭 사용량을 의미함 (단위: bps)
- **서비스 방문자 수:** 1시간 단위로 중복 IP를 제거하며, 서비스를 사용하기 위해 방문한 사용자의 총 수를 의미함 (단위: user)
- **접속 수(Hit):** 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
- **콘텐츠 사용량 집계:** Hit를 통해 가장 많이 사용되는 콘텐츠를 검색하여 집계함

4. 오픈플랫폼 사용현황 분석 및 평가

본 장에서는 오픈플랫폼 시스템의 웹 서버를 통하여 수집된 웹 사용자 로그를 분석하여 서비스 사용현황의 통계 결과 및 분석 결과를 제시한다. 본 연구를 수행하기 위한 로그 분석 기간은 2014년 5월 20일 자정부터 2014년 6월 2일 자정까지이며, Portal WEB 서버 6대, 2D WEB 서버 4대, 3D WEB 서버 6대에서 일 단위로 생성된 로그 파일을 수집하여 그 결과를 분석 하였다.

특히, 분석 결과 주말의 사용량이 주중에 비해 약 50% 수준이었으며, 1일 기준으로 평균 업무 시간대인

Table 3. Measurement of Analysis Data

Date	Data Transfer(GB)	Network Bandwidth(Mbps)	User Count	Hit Count	Data Transfer/User (Byte)	Data Transfer/Hit (Byte)
2014-05-20 TUE	341.89	104	116,295	35,239,384	2,939,884	9,702
2014-05-21 WED	343.44	85	117,756	34,363,772	2,916,531	9,994
2014-05-22 THU	336.13	107	116,669	32,883,771	2,881,041	10,222
2014-05-23 FRI	303.70	75	112,508	30,288,192	2,699,399	10,027
2014-05-24 SAT	166.64	42	78,562	16,460,755	2,121,113	10,123
2014-05-25 SUN	165.07	87	68,586	15,647,693	2,406,806	10,549
2014-05-26 MON	327.59	92	109,648	32,540,163	2,987,682	10,067
2014-05-27 TUE	359.83	99	117,466	35,717,849	3,063,311	10,074
2014-05-28 WED	330.27	108	78,176	33,216,055	4,224,715	9,943
2014-05-29 THU	319.04	103	76,094	33,288,330	4,192,668	9,584
2014-05-30 FRI	310.99	104	83,984	32,503,853	3,702,990	9,568
2014-05-31 SAT	153.74	38	51,448	15,745,453	2,988,215	9,764
2014-06-01 SUN	137.82	41	64,255	14,001,265	2,144,946	9,844
2014-06-02 MON	324.06	97	129,162	33,113,140	2,508,935	9,786

9시에서 18시까지의 사용량이 높은 것으로 분석 되었다. 이는 오픈플랫폼에서 제공하는 서비스가 단순한 통합 지도 제공 서비스 이외에 개발자 지원 서비스, 공공 데이터 제공 서비스 등이 있기 때문에 이를 이용하는 사용자들이 그들의 업무와 연관된 서비스를 주로 업무 시간 내에 이용을 하였던 것으로 판단된다.

Table 3은 분석된 결과를 토대로 작성한 일별 사용 현황에 관한 통합 수치자료이며 각 항목에 대한 분석은 다음과 같다.

4.1 데이터 전송량

우선 클라이언트의 요청에 의해 서버에서 전송된 데이터 전송량은 Figure 4에서 보듯이 일일 누적 주중 평균 330GB, 주말 평균 150GB 정도로 나타나고 있다. 일일 방문자 수로 본 데이터 전송량을 나누어보니, 사용자 1명당 하루 평균 2MB에서 4MB 정도의 데이터를 사용하였음을 알 수 있었다. 또한 일일 Hit 수를 나누어보니, 한 번의 서비스 요청(Request)에 대해 평균 10KB 정도의 데이터가 전송되었음을 알 수 있었다.

하루 동안의 데이터 전송량을 세부적으로 분석해보면 1시부터 8시사이의 데이터 전송량이 가장 낮으며 9시부터 증가하여 15시에서 17시 사이에 가장 높은 데이터 전송량을 보이고, 19시 이후 다시 낮아짐을 볼 수 있다(Figure 5). 이는 예측된 바와 같이 사용자들이 오픈플랫폼 서비스를 주로 평균 업무시간 대에 많이 사용을 하고 있음을 의미한다. 시간별 최대 데이터 전

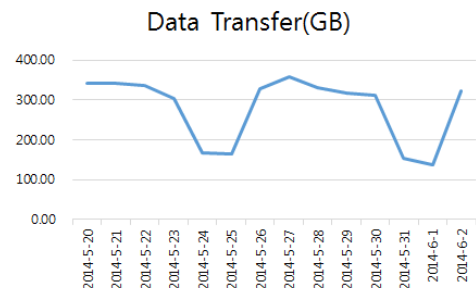


Figure 4. Daily Amount of Data Transfer

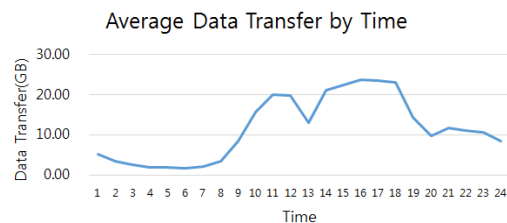


Figure 5. Hourly Average Data Transfer Rate

송량은 5월 22일 목요일 17시부터 18시 사이로 동시간대 평균값인 23.11GB의 150% 정도 높은 값인 34.6GB로 나타난다.

4.2 네트워크 대역폭

서버에서 전송되는 데이터의 네트워크 점유율을 의미하는 네트워크 대역폭은 Figure 6에서 보듯이

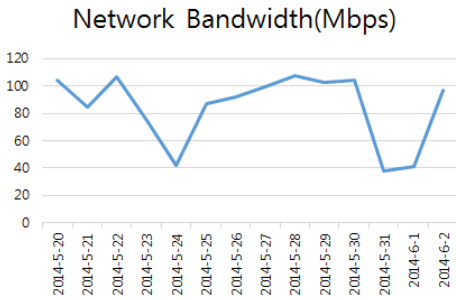


Figure 6. Daily Network Bandwidth

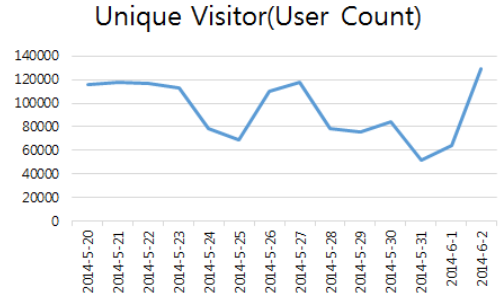


Figure 8. Daily User Count

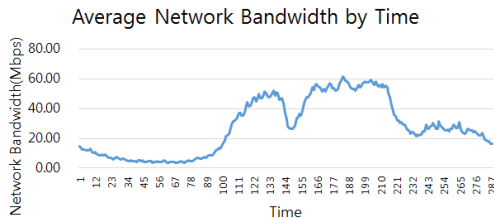


Figure 7. Hourly Average Network Bandwidth

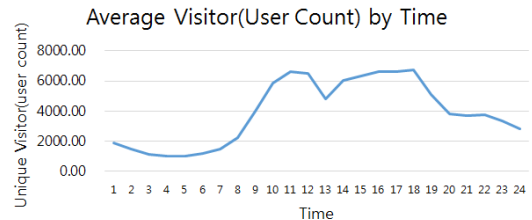


Figure 9. Hourly Average User Count

주중에는 평균 80Mbps~110Mbps 정도로 주말에는 평균 50Mbps 정도로 나타나고 있다.

하루 동안의 시간대별 평균 네트워크 대역폭은 Figure 5의 시간대별 평균 데이터 전송량 그래프와 유사한 패턴을 보이고 있다. 이는 동일 네트워크 대역폭을 가진 시스템에서 전송되는 데이터에 비례해서 전송 데이터가 네트워크를 점유하기 때문이다. 구체적으로 9시부터 증가된 네트워크 대역폭은 10시에서 17시 사이에 55Mbps 정도의 대역폭을 보이며 19시 이후에는 감소함을 볼 수 있다(Figure 7). 네트워크 대역폭에서 특히 주목해야 할 사항은 시스템에서 지원할 수 있는 최대 네트워크 대역폭 대비 일일 최고 네트워크 대역폭 사용량이다. 현재 오픈플랫폼 시스템의 일일 최대 네트워크 대역폭은 1Gbps로 분석 결과 데이터에 의하면 일별 최고 대역폭이 107.6Mbps(5월 28일 수요일 15시)로 아직은 평소의 안정적인 상황을 고려할 때 네트워크 용량은 여유가 있다고 판단된다.

4.3 서비스 방문자 수

오픈플랫폼 서비스에 접속하는 방문자 수로서, 이는 오픈플랫폼 서비스 사이트에 직접 접속하여 서비스를 이용하는 사용자 이외에, 오픈플랫폼에서 제공하는 오픈 API 등의 기능들을 사용하여 구축된 웹 응용 시스템에서 접속한 사용자도 집계에 포함한다. 예를

들면 오픈플랫폼 오픈 API 서비스 등을 이용하고 있는 ‘대구광역시 버스 노선 안내’, ‘산림청 산사태 정보 시스템’, ‘북한 정보 포털’ 등의 웹 사이트에서 접속한 사용자가 이에 해당된다. Figure 8을 보면 오픈플랫폼의 순 방문자 수가 일일 누적 기준으로 주중에는 평균 11만 명 수준이고, 주말에는 평균 6만 5천 명 수준임을 알 수 있다. 본 연구에서는 로그 정보를 기반으로 방문자 수를 도출하는데 있어서 한 시간 단위로 오픈플랫폼 시스템에 접속한 클라이언트의 IP 중에서 중복 IP를 제거 한 후에 IP를 카운트하여 방문자 수를 계산하는 방법을 이용하고 있다. 그러므로 한 시간 이상 접속한 사용자는 일 2회 이상 방문한 것으로 산정될 수 있기 때문에 분석 결과로 주어진 방문자 수가 실제 방문자 수보다 많을 수도 있음을 고려할 필요가 있다.

Figure 9는 하루 동안 시간대별 평균 방문자 수를 보여주는데, 방문자 수가 9시부터 증가하여 18시까지 6천명 이상을 유지하다가 19시 이후부터는 서서히 감소하고 있음을 보여주고 있다. 특히 새벽 시간인 3시부터 6시까지도 1천명 이상의 사용자가 꾸준히 유지되고 있다는 점은 흥미로운 결과라고 할 수 있겠다. 본 연구 수행 동안 측정 일 중 시간당 최대 방문자 수는 10,413명(6월 2일 월요일 17시)이었다. 만약 사용자 당 평균 접속시간을 10분으로 가정한다면, 본 정보를 이용하여 평일 시간당 방문자 수 6천명 기준으로

평균 동시접속자 수가 1천명(=6,000/60x10) 수준임을, 시간당 최대 방문자 수 1만 명 기준으로는 1천 7백 명 수준임을 예측할 수 있다. 현재 오픈플랫폼의 WEB 서버, WAS, 기타 서버 환경을 고려할 때 1천 7백명 수준의 동시접속자에게 서비스를 제공하는데 아직은 무리가 없다고 판단된다.

4.4 접속 수(Hit)

오픈플랫폼 서비스에 접속하는 Request에 대한 Hit 수로서 오픈플랫폼 서비스 사이트에 직접 접속하여 서비스를 사용하는 것 이외에 외부 연결에 의하여 Hit 된 경우도 포함한다. Figure 10에서 보듯이 주중에는 평균 3천만 건 이상의 Hit, 주말에는 평균 1천 5백만

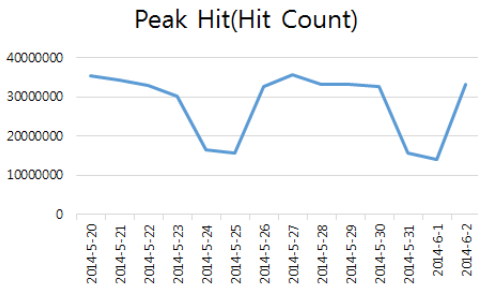


Figure 10. Daily Hit Count

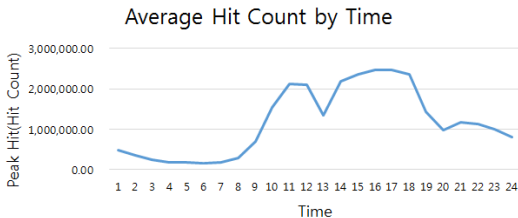


Figure 11. Hourly Average Hit Count

Connection Path(Referer)

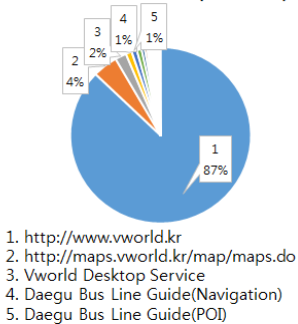


Figure 12. Percentage of Connection Path

건의 Hit처럼 방문자 수에 비하여 예상했던 것보다 상당히 많은 Hit가 발생함을 알 수 있다. 이는 오픈플랫폼 서비스에서 통합지도 서비스의 지도 화면을 표시하기 위해서는 상당히 많은 수의 클라이언트 Request가 요청되어야하기 때문에 Hit수가 실제 서비스 방문자 수에 비하여 크게 증가하게 되는 것이다.

Figure 11에서 보듯이 하루 동안의 시간대별 Hit 수는 Figure 9의 시간대별 방문자 수와 매우 유사한 패턴을 보여주고 있음을 알 수 있다. Hit 수가 9시부터 증가하여 18시까지 매시간 평균 2백만 건 이상을 유지하다가 19시 이후부터는 서서히 감소함을 알 수 있다.

4.5 접속 경로

오픈플랫폼 서비스에 접근하는 경로를 파악하기 위해 Hit 수와 대역폭(Traffic)을 기준으로 3천 건의 경로를 수집하여 분석하였다. Figure 12에서 보는 바와 같이, 그 중 Hit 수를 기준으로 전체의 95%에 해당하는 상위 5위 이내의 결과를 분석해 보면, 오픈플랫폼 서비스 사이트로 직접 접속하는 경우가 상위 1~3위에 해당하여 대부분을 차지하고 있음을 알 수 있다. 나머지는 다른 사이트에서 제공하는 서비스에서 오픈 API를 통해서 연결되거나 링크 형태로 연결된 것으로, 대구광역시 버스 노선 안내 서비스의 노선검색 서비스와 출발지 도착지 POI(Point Of Interest) 검색 서비스가 꾸준히 상위 4~5위를 차지하고 있음을 알 수 있다. 이외에도 산림청의 산사태 정보 시스템, 통일부의 북한정보포털 지도 서비스, biz-gis.com 서비스, 안전행정부의 생활안전지도 서비스, 필지별 산림 정보 시스템, 취약노인 지원 시스템, 공항 소음 포털 서비스 등과 같이 다양한 공공 및 민간 기관을 통한 접속 경로도 확인할 수 있었다[1,2,9,13].

4.6 콘텐츠 사용 순위

오픈플랫폼 서비스에서 제공되는 콘텐츠의 사용 빈도를 측정하고자 Hit 수를 기준으로 3천 건의 콘텐츠를 수집하여 분석하고자 하였으나, 통합지도 서비스의 경우 지도 페이지 하나에 다수의 클라이언트 Hit가 발생하므로 집계된 Hit 수 분석을 통하여 클라이언트가 정확히 어떤 서비스를 요청하였는지 확인하기에는 어려움이 있었다. 그럼에도 불구하고, 분석결과를 보면 대부분의 서비스가 통합지도 서비스 요청에 해당되는 콘텐츠로서 지도 정보 관련 콘텐츠 요청이 90% 이상임을 예측할 수 있었다.

5. 결 론

공간정보에 대한 사람들의 관심이 크게 증가하면서 오픈플랫폼에 대한 방문자 수의 급증과 함께 오픈플랫폼 시스템이 중단되는 사태가 수차례 발생하게 되었다. 더욱 큰 문제는 이러한 사태를 해결할 수 있는 실시간 모니터링 시스템도 도입되지 않은 상태이다. 이에 본 연구에서는 오픈플랫폼 시스템의 WEB 서버 16대의 웹 사용자 로그정보를 이용하여 데이터 전송량, 네트워크 점유율, 서비스 방문자 수, Hit 건수, 접속 경로, 주요 이용 콘텐츠에 대한 분석 결과를 제시하였다. 평일 기준 하루 10만 명 이상의 방문자 수와 300GB 이상의 데이터 전송량을 보여주는 분석 결과에 따르면, 현재 오픈플랫폼 서비스에 대한 관심이 크게 증대되었음을 수치적으로 알 수 있었다. 아울러, 대구 버스정보시스템을 포함한 다양한 공공 및 민간 기관을 통한 접속 및 활용도 확인할 수 있었다.

현재 본 연구에서 제시한 결과로 오픈플랫폼 서비스 사용 현황을 모두 파악하기에는 아직 부족한 점이 많다. 그러나 본 연구결과는 해마다 변화하는 오픈플랫폼의 사용 현황을 파악하기 위한 근거 자료로, 또한 오픈플랫폼 시스템 확장에 있어서의 기초자료로 활용이 가능할 것으로 판단된다. 향후 연구 방향으로서는 서비스별 사용현황과 같은 세부 서비스 및 주요 사용 콘텐츠에 대한 분석을 통해 실제 오픈플랫폼 서비스를 사용하는 사용자별 현황을 파악함으로써 서비스별 활용도를 판단할 수 있는 결과 도출이 필요하다고 판단된다. 또한, 오픈플랫폼 활용을 위해 제공하는 오픈 API 키는 2014년 3월까지 1,996건이 발급되었으나 발급 후 사용하지 않는 사용자도 있을 수 있으므로 실제로 사용하고 있는 오픈 API 사용에 대한 현황 파악이 필요하다고 판단된다[12].

References

- [1] Biz-GIS Corporation, Accessed July 25. <http://www.biz-gis.com/>.
- [2] Daegu Metropolitan City Bus Line Guide, Accessed July 25. <http://businfo.daegu.go.kr/>.
- [3] Ha, E. O; Kim, Y. H. 2009, Design and Implementation of Web-based Monitoring System for an EAI Environment, Journal of Society for e-Business Studies, 14(3):1-13.
- [4] Jun, S. H; Doh, K. T. 2013, Design and Implementation of Web GIS Server Using Node.js, Journal of KSIS, 21(3):45-53.
- [5] Kim, D. J; Jeong, S. J; Kwon, Y. H; Kim, J. S; Kim, J. K. 2003, Design and implementation of a web based IDC oriented monitoring and management system, The Journal of the Institute of Information and Telecommunication, 10(1):37-43.
- [6] Kim, D. Y; Kim, J. T. 2009, Efficient Design of Web Searching Robot Engine Using Distributed Processing Method with Javascript Function, KIMCIS Journal, 13(12):2595-2602.
- [7] Kim, E. H. 2009, The Strategies of Technology Development for Geospatial Web Platform, The Journal of Geographic Information System Association of Korea, 17(2):171-181.
- [8] Kim, I. K; Shin, J. C. 2013, A Study of Advanced N-Tier Model for Improving Maintainability of Web Applications in Cloud Environment, A study of Information Technology Architecture, 10(4):537-553.
- [9] Landslide Information System, Korea Forest Service, Accessed July 25. <http://sansatai.forest.go.kr/>.
- [10] MOLIT(Ministry of Land, Infrastructure and Transport) News, 2012, North Korea 3D satellite image map services expanding the congestion, Accessed September 27. http://www.molit.go.kr/USR/NEWS/m_72/dtl.jsp?id=95071032.
- [11] MOLIT(Ministry of Land, Infrastructure and Transport) News, 2013, Vworld service delay with congestion, Accessed September 29. http://www.molit.go.kr/USR/NEWS/m_72/dtl.jsp?id=95072769.
- [12] MOLIT(Ministry of Land, Infrastructure and Transport) News, 2014, Reference for Joongang Sunday news of Vworld, Accessed April 14. http://www.molit.go.kr/USR/NEWS/m_72/dtl.jsp?id=95073878.
- [13] North-Korea Information Portal, Accessed July 25. <http://nkinfo.unikorea.go.kr/>.
- [14] Oh, K. s. 2013, Changes in the structure of internet trends and implications: Specially CDN, Information and Communication Policy, KISDI, 25(11):1-24.
- [15] Park, C. J; Cho, J. K. 2007, A Study on the Implementation and Practical Application of Web

- Monitoring System, The Journal of Information Systems, 16(2):35-49.
- [16] Park, J. W. 2010, Analysis on the Performance Elements of Web Server Cluster Systems, The Korea Society for Simulation Journal, 19(3):91-98.
- [17] Park, J. W. 2013, Operational Scheme for Large Scale Web Server Cluster Systems, The Korea Society for Simulation Journal, 22(3):71-79.
- [18] TTAK.KO-10.0292, 2008, A Guideline for Hardware Sizing of Information Systems, TTA.
- [19] Vworld: Spatial Information Open Platform, 2014, MOLIT(Ministry of Land, Infrastructure and Transport), Accessed July 29. <http://www.vworld.kr>.
- [20] Vworld: Spatial Open Platform Hardware system, 2014 March.

논문접수 : 2014.7.31
수 정 일 : 2014.8.20
심사완료 : 2014.8.22