2015 IEEE 17th International Conference on High Performance Computing and Communications (HPCC), 2015 IEEE 7th International Symposium on Cyberspace Safety and Security (CSS), and 2015 IEEE 12th International Conf on Embedded Software and Systems (ICESS)

Survey of Big Data-as-a-Service Type

Yunkon Kim, Yong-Hyun Kim, Ga-Won Lee, Eui-Nam Huh
Department of Computer Engineering
Kyung Hee University
Yongin-si, South Korea
{ykkim, kyh0104, gawon, johnhuh}@khu.ac.kr

***Abstract*—Big Data is used to find new value and brings us several benefits, which we didn't know before. Various analytics has been studied in Big Data area for benefits. Moreover, to reduce analysis time and to support real-time service to user, distributed processing can be an alternative solution. Big Data also requires high performance resource for distributed analysis, for this reason, Big Data and Cloud Computing seem to be naturally combined-called Cloud-based Big Data. However, Cloud-based Big Data had no criteria to evaluation. Also, it is hard to decide which Cloud-based Big data is well designed and how much resource should be provided to provide qualitied Cloud-based Big data service. Hence, in this paper, we surveyed enterprises and their product and then, deduct criteria to classify and evaluate Cloud-based Big Data service.**

***Index Terms*—Big Data, Cloud Computing, Big Data-as-a-Service (BDaaS), Service Type.**

I. Introduction

ICT(Information and Communications Technology) development has increased portability of device and size of digital media. Many types of device have made human's life safe and convenient. Devices are newly released in short period and are generating data every time. Small amount of data, which devices generate, comes together and becomes large amount of data. It means the devices such as PC, mobile phone, and sensor device in IoT (Internet of Things), create huge amount of data every day. For example, globally 500 million Tweets are sent per day and 80% of Twitter active users are on mobile (Twitter Inc.). New 30 billion contents would be shared on Facebook in a month. Information in the world was forecasted to be increased 1.8 zettabyte every year and be multiplied fifty times until 2020 (IDC, 2012). The more data are generated, the more complex diversity of data become. Current technique is focused on development to handle large amount and high diversity data. Scientists and engineers call it Big Data. Big Data can be a core of digital information age and a source of new value.

Big Data concept spreads out to whole area of human activity. Big Data is applied IT as well as human's life, hence it is considered as a key factor to decide competitiveness. Big Data as economical input can create new business value. With proper method of analysis, Big Data brings high quality life for human, becomes innovation itself, and becomes source of new value.

Big Data is not just big and complex data. When big and complex data is analyzed, it can be Big Data in the true sense of the term. Not only data is stored on the storage, but also all data can be important when it is analyzed. Therefore, analytics is important to Big Data. Expression, "Garage-in, Knowledge-out" maybe represent importance of analysis in Big Data. Analysis for expectation, advanced analysis in banking, social media analysis has been researching to extract knowledge or service. Because huge amount of data is analyzed with complex algorithm, distributed processing is required to reduce analysis time.

Although Big Data has lots of possibility in point of mining new information, it has problem of resource allocation and lack of resource because it has to process huge amount of data. One of resource problem is caused that size of Big Data or complexity of Big Data is hard to be measured. For storing and analyzing huge amount of data, many storages and computing resource are required. Expandability and economic feasibility of Cloud Computing is proper structure for Big Data. In other words, Big Data can be considered as a key source of service and Cloud Computing can be a part of platform and infrastructure. Cloud-based Big Data will solve that resources aren't elastic and support Big Data system to analyze unstructured, large data which is created in real-time because resource can scale up and down automatically. These role allocations between Big Data and Cloud Computing have an effect of reducing cost and usability of management. To analyze Big Data, servers which can handle distributed, parallel processing are necessary. After analyzing Big Data using Cloud Computing infrastructure, users just pay their usage. That can be cheaper than what they deploy their own infrastructure. Thus, combining Big Data and Cloud Computing system has advantages, such as distributed processing, extensible structure, and economical competitiveness.

Despite of benefits, an existing problem is that there are no criteria to evaluate Cloud-based Big Data. Each enterprise has different criteria for Cloud-based Big Data services. Some enterprises use their own infrastructure while others use part of or whole Cloud resources. Some organizations just have Big Data analysis technique while others just have distributed processing infrastructure. However, Big Data service products in each enterprises looks similar and hard to compare. And also, Big Data platform is necessary IT environment to extract knowledge by analyzing various data from different sources and provide intelligence services. Therefore, classification of Big Data service is important for users.

In this paper, we surveyed many Big Data enterprises and their products. In section II, we investigated Cloud-based Big

978-1-4799-8937-9/15 $31.00 © 2015 IEEE
DOI 10.1109/HPCC-CSS-ICESS.2015.279

IEEE computer society

Data platform and service. We classified enterprises based on survey to organize common requisites. We evaluated enterprises with common requisites. We categorized evaluation result as a type. At the conclusion, we deducted relation between Big Data-as-a-Service and type that we categorized.

II. Related Work

A. Big Data

The term "Big Data" was first introduced to the computing world by Roger Magoulas from O'Reilly media in 2005, in order to define a great amount of data that traditional data management techniques cannot manage and process due to the complexity and size of this data. Big Data has five characteristics called 5V's. 1) Volume refers to the quantity of data gathered by a company,2) Velocity refers to the time in which Big Data can be processed, 3) Variety refers to the type of data that Big Data can comprise, 4) Value refers to the important feature of the data which is defined by the added-value that the collected data can bring to the intended process, activity or predictive analysis/hypothesis, and 5) Veracity refers to the degree in which a leader trusts information in order to make a decision[1][2][3][4][9].

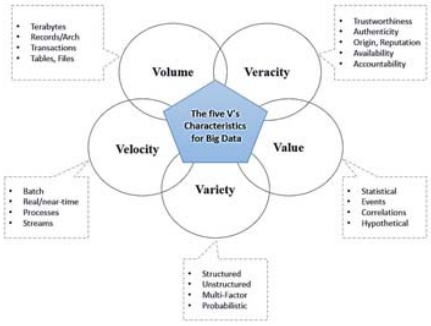

Figure 1. Five Vs Big Data Characteristics[1]

Challenges in Big Data area include Analytics, visualization, storage, information privacy, and so on. Storage management is necessary due to importance of distributed processingto analyze Big Data. Recently, most Big Data systems incorporate Hadoop-based architectures and are quickly becoming the center of the enterprises technology stack for data management[1][3][5].

B. Cloud Computing

Cloud computing is a paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand. In rec. ITUT Y.3500 | ISO/IEC 17788, there is a role which is a set of cloud computing activities that serve a common purpose. Roles has been defined Cloud service customer (CSC), cloud service provider (CSP), and cloud service partner (CSN). Cloud service categories include Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), Software-as-a-Service(SaaS), Network-as-a-Service (NaaS)[3][6][7][8]. These service levels are not new, but technology has evolved to make the consumption patterns look different from how they looked in the past[3]. Five essential characteristics of Cloud Computing are on-demand self-service, borad network access, resource pooling, rapid elasticity, and measured service[7].

- ***On-demand self-service***: A consumer can access different services as needed automatically without service provider's intervention.
- ***Broad network access***: All services are available over the network and are also accessible through standard protocols.
- ***Resource pooling***: The resources are pooled to serve the users at a single physical location and/or at different physical location according to the optimality conditions.
- ***Rapid elasticity***: The resources can be provisioned without service provider intervention and can be quickly scale in and scale out according to the user needs.
- ***Measured service***: A metering capability is deployed in cloud system in order to charge users.

C. Big Data-as-a-Service

Big Data-as-a-Service encapsulates various big data storage, management, and analytics techniques into services and provides common big data related services to users via programmable APIs, which greatly enhances efficiency, reduce cost and enables seamless integration[10]. Data can be analyzed for a lot of purposes, such as enhancing system performance, guiding decision making, assessing risk, trimming costs, lifting sales, and so on[11]. It is necessary to have a single point(infrastructure) which provides common functionality of big data management, and flexible enough to handle different types of big data and big data analysis tasks[12].

Scale-out infrastructure provides necessary computing and storage capacity for big data. Infrastructure for Big Data have to combine with storage designs[13]. A Big Data platform allows users to access, analyze and build analytic application on top of large data sets[14]. Big Data analytics is the process of examing large amounts of data of various types to uncover hidden patterns, unknown correlations and other useful information[15]. There are too few skilled big data practitioners available for every organizations. Therefore, the organizations turn to Big Data Analytics Software-as-a-Service[16].

Issues in related researches can be devided into two big categories; big data management and scalable computing resource for Big Data analysis. Academic researchs about these issues are not extended yet. Therefore, Big Data combined Cloud Computing can be one of key solution to solve these issues easily. Enterprises also focus on Cloud-based Big Data service or Big Data-as-a-Service regarding these manners. Criteria to classify Big Data-as-a-Service is required due to different criteria of each enterprises. Furthermore, the criteria can also help Big Data-as-a-Service customers to choose suitable Big Data services.

III. Service Types of BigData-as-a-Service

In this paper, we investigated Cloud-based Big Data enterprises and deducted two common factors. One factor is to

classify Big Data platform or solution and the other factor is to evaluate requirement satisfaction of enterprises about Cloud-based Big Data. At the beginning of this section, we showed current state of affairs of Cloud-based Big Data platform and service. And then we arranged an evaluation criteria based on Big Data requisites and Cloud Computing requisites. Next, we classified enterprises as some groups by evaluation criteria. Classified groups were used to make service types. Lastly, we defined concept of Big Data-as-a-Service and deducted relation between Big Data-as-a-Service and type that we categorized.

A. Current state of affairs of Cloud-based BigData Platform and Solution[10, 11]

Big Data platform enterprises from Big Data Landscape were candidate for analysis. Big Data Landscape Version 3.0 categorized as Application group, Analytics group, Infrastructure group, Infrastructure /Analytics group, Open Source group and Data Sources group.

TABLE I. GROUPS FROM BIG DATA LANDSCAPE

Groups	Detail
Application	To provide service using Big Data
Analytics	To provide analysis technique such as Big Data search and analysis and Machine Learning
Infrastructure	To provide Hadoop-based platform, NoSQL Database, Storage, Security, Monitoring, and soon
Cross Infra./ Analytics	Combined group, Infrastructure and Analytics
Open Source	Framework and open source techniques
Data Source	Base group for the other groups

We categorized Big Data platform enterprises into 4 groups: Application, Analytics, Infrastructure, Cross Infrastructure/Analytics. Because open source and data source were possible to make ambiguousness, we excluded these two groups from criteria.

Total 37 enterprises providing Cloud-based Big Data service were analyzed. 10 enterprises were in Application group, 14 enterprises were in Analytics group, 9 enterprises were in Infrastructure group, and 4 enterprises were in Cross Infrastructure/Analytics group. Following our survey, trend of Big Data might be analytics and infrastructure area. To make application or service, companies tried to analyze data and needed infrastructure to support analysis. As our survey, sub-part of analytics and infrastructure is like following sentences. Social and business analytics, real-time analytics, advanced analytics such as expecting analysis, and specified analytics such as banking and music were located in various areas which requires Big Data analytics. Infrastructure part were researched and developed in storage, Massive Parallel Processing (MMP) Database, In-memory analytics, data management, and so on. As mentioned, details of survey is in Table 2.

TABLE II. DETAIL OF CLASSIFIED ENTERPRISES

Groups	Company	Details
Application	Clickfox	Providing customized information based on customer's experience and location
	Sailthru	Customized solution service based web site browsing habit
	Bloomreach	Social data analysis solution for retail institutions
	KNEWTON	e-learning solution using machine learning and statistical analysis
	NextBigSound	Social analysis service about music
	Bloomberg	Banking market analysis and monitoring solution
	Numberfire	Sport statistics analysis and monitoring solution
	Yieldex	Trace user and personalized advertisement solution
	Bluekai	Customized advertisement solution by data analysis
	Hortonworks	Integrated platform for analysis on distributed processing
Analytics	SKYTREE	Machine learning and advanced analysis solution
	Revolution Analytics	Big Data analysis solution
	SPSS	Statistics analysis solution
	Bluefin Labs	Social TV analysis solution
	Dataminr	Tweeter monitoring and analysis solution
	SimpleReach	Social analysis solution
	Recorded Future	Analysis solution for expectation
	Place IQ	Location-based customer analysis solution
	Centrifuge	Visual network analysis
	Dell	Big Data search and analysis platform
	Ayasdi	Data analysis and multi-tenant cloud platform
	metaLayer	Big Data visualization, monitoring and analysis
	Tableau	Browser-based analysis and collaboration solution
	Altair	Intelligent analysis solution
Infra.	Cloudera	Batch processing, interactive SQL, and searching and analysis solution
	10gen	Management solution for monitoring and backup MongoDB
	DATASTAX	NoSQL-based Big Data management solution
	TERADATA	Big Data analysis and Hadoop solution
	EMC^2	MPP architecture solution and appliance
	MarkLogic	Real-time interactive analysis solution and distributed store system
	Paradigm4	In-Database advanced analysis solution
	OceanSync	Cloud management and processing platform for Hadoop
	MapR	Advanced Big Data analysis platform
Cross Infra. /Analytics	Amazon	Marketing solution using analysis and NoSQL
	HP	Real-time MPP and in-memory analysis
	IBM	Stream data analysis and MPP data warehouse
	Oracle	Advanced analysis integration platform for in-database based on NoSQL

According to the details of classified enterprises, enterprises of Application group mainly carried out general data analysis, visualization, and semantic web. Enterprises of Analytics group has a goal to analyze and process data faster than previous analyze approach. They mostly use advanced analysis technique such as expecting analysis and scientific analysis by data scientist. Enterprises of Infrastructure group could support the other groups by providing NoSQL, MPP and Map-reducing technique. A Goal of Infrastructure group enterprise is to provide faster data build and more stable data management by using non-relational database, parallel processing, and in-memory technique.

B. Method of deducting requisites

To evaluate Cloud-based Big Data Service in the enterprise classification, we matched five Vs of Big Data area with five characteristics of Cloud Computing area. And then, we applyed

classification result to the matched one. Correlation of Big Data and Cloud Computing in Big Data perspective was arranged as follows: Rapid Analysis, Data Collection, Scale up and down, Distributed Store, Processing Support, and SLA. At Cloud Computing perspective, Correlation was arranged as Big Data service customer, and Big Data service provider. In sequence, requsities were selected from the five Vs of Big Data and five characteristics of Cloud Computing by enterprises classification result. Through this classification, we accumulate points for requisites by requirement in enterprise's solution. Consequently, Volume, Velocity, Variety became requisite of Big Data and all five characteristics of Cloud Computing became requisite as shown as Figure 2. The reason Veracity and Value didn't become requisite was that we couldn't measure 2Vs by opened information of enterprises.

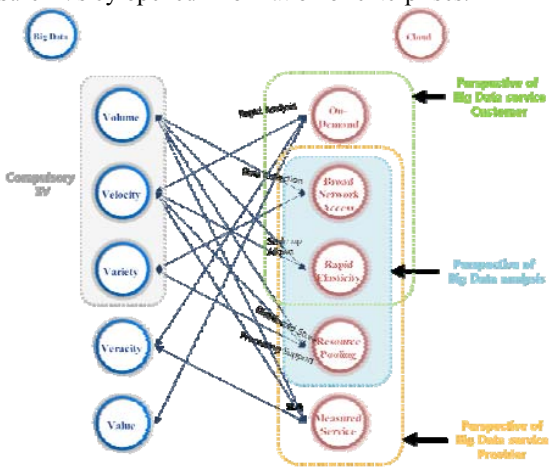

Figure 2. Correlation Between Big Data and Cloud Computing

C. Definition of Criteria for Cloud-based Big Data service Enterprise

There are various Cloud-based Big Data services in the area of IaaS, PaaS, and SaaS. However, it is hard to classify Cloud-based Big Data services because lack of evaluation criteria. Moreover, each enterprise has different criteria to provide Cloud-based Big Data service. Thus, customers, who are using Cloud-based Big Data, and even Cloud-based Big Data service providers are hard to know whether the services are veritably Cloud-based Big Data service or not and how much resources, tools, methods, and other elements should be provided for Cloud–based Big Data Service. Therefore, we analyzed Cloud-based Big Data enterprises by using three Big Data requisites and five Cloud requisites. 3 Big Data requisites are Volume, Velocity, and Variety. 5 Cloud requisites are Rapid Elasticity, Broad Network Access, Resource Pooling, On Demand Self-service, and Measured Service. We valued each requisite as a point from 0 to 3. A criteria of point is as follows:

TABLE III. BIG DATA AREA CRITERIA

Point	Volume	Velocity	Variety
3	Over PB	Real-time Analysis	Structured, semi-structured, unstructured
2	TB	Near real-time analysis	Structured, semi-structured
1	GB	Simplicity analysis	Structured
0	none	none	none

TABLE IV. CLOUD AREA CRITERIA

Point	Rapid Elasticity	Broad Network Access	Resource Pooling	On Demand Self-service	Measured Service
3	Automated scale-up/down	Linkage various mobile device	Unlimited resource use	Any-time, Any-where and Various Device Use	Possible to measure SLA
2	Scale-up/down as request	Linkage inter/ external network	Limited resource use	Any-time, Any-where or Various Device Use	-
1	Use Third-party cloud storage	Linkage internal network	-	-	-
0	Limited scale-up/down	Impossible to access network	Impossible to share resource	Use under manager's permission	Not be measured

D. Analysis relation between the criteria and provided information of solution

We selected ten enterprises which had commercialized solution. Table 5 shows relation analysis points and scores based on the provided information from the enterprises. We use anonymous name for the security of reasons regarding name of the enterprises. To analyze relation based on similarity from the above table we visualized as Figure 3.

TABLE V. RELATION ANALYSIS POINT

Name	Classification	Function	Volume	Velocity	Variety	Rapid Elasticity	Broad Network Access	Resource Pooling	On Demand Self-,	Measured Service
A	IaaS, PaaS, SaaS	Elastic MapReduce, DynamoDB, Redshift, S3, Kinesis	3	3	3	3	3	3	3	3
B	PaaS	Hadoop/HDFS, Autonomy IDOL, Vertica, Enterprise Security, Apps	3	3	3	3	1	2	3	3
C	IaaS, PaaS, SaaS	Management, Analysis, Application	3	3	3	3	3	3	3	3
D	IaaS, PaaS, SaaS	Query, Analysis, DataSet Sharing	3	3	3	3	3	3	3	3
E	IaaS, PaaS, SaaS	Hadoop Cluster, Visualization, Analysis, Unstructured Data Processing, Scale Up	3	3	3	3	3	3	3	3
F	PaaS	Hadoop Cluster, Analysis	3	3	3	3	1	2	3	3
G	PaaS	Scale Up, Pricing, Analysis	3	3	3	3	1	2	3	3
H	PaaS	Collection, Indexing, Analysis, Hadoop, NoSQL	2	2	3	3	1	2	3	0
I	PaaS	AWS, Hive	3	2	3	3	3	3	3	3
J	PaaS	Hadoop	3	1	3	2	0	2	2	0

It describes some characteristics and similarities between ten enterprises. We try to classify enterprises as service types according to characteristics in Table 6. As shown from enterprise A, B, C and D, we called Type I, the enterprises are satisfy all eight criteria regarding big data and cloud. This pattern has storage which can store huge amount of Big Data, resource which can process Big Data rapidly, network which can communicate with various data source. Type II satisfy seven criteria except velocity such as enterprise I which might not be able to support real-time analysis. Type III, then, satisfied six criteria has weak point at resource pooling for cloud and rapid elasticity scored 2 and zero point respectively, it discovered from enterprises B, F and G. This platform might provide limited storage and might not support scale-up/down. Lastly, enterprise H and J just meet five criteria has week point at cloud technologies, when it comes to considering user's request, volume, and various data. It is Service Type IV. Type V is out of cloud-based big data service. These classified service types and characteristics are appeared to similar shape of chart regardless shape size.

TABLE VI. SERVICE TYPES

Service Type	Details
I	This type maintain balance between Cloud and Big Data part. In Cloud part, providers have their own infrastructure. Resources can be managed automatically to scale-up/down for distributed processing. In Big Data part, providers have huge volume and various type of data. High data processing speed can be supported by distributed processing. This type can provide on demand self-service.
II	This type satisfied more Cloud part than Big Data part. Resources, which providers have, can be managed automatically. However, they have weak point in Big Data. It is possible for them to have small volume of data, not fully enough algorithm, and not various type of data. This type can provide on demand self-service.
III	This type satisfied more Big Data part than Cloud Part. They have advanced Big Data analysis technique, Large volume of data. However, They rent Cloud infrastructure or have own infrastructure. It is hard to manage infrastructure to scale-up/down. This type can provide on demand self-service.
IV	This type partially satisfied both Big Data part and Cloud part. Providers can give on demand self-service by in limited resource use and specific Big Data analysis technology.
V	This type is out of Cloud-based Big Data service.

E. Big Data-as-a-Service

According to this survey, we infer some requirements. First, huge amount of data including structured, semi-structured, and unstructured data are needed. Second, BDaaS might be able to various type of data. Third, Cloud resources might be able to scale-up/down is more helpful to analyze various types of data. Fourth, real-time or near real-time analysis is possible. Fifth, resouces can be scaled-up and down as autometically by usage while analyzing and processing. Finally, BDaaS customer just pays as amount of thier usage.

Big Data can make new information from huge data sourceHigh perfirmance and discributed computing environment are neccessary. Then cloud computing can manage resources including storage automatically. Customer can save cost setting up their own infratructure because customers just pay for their usage. Customer will spend network traveling time. For this reason, combining big data with cloud computing will make synergy. On cloud environment, analysis speed will be increased, thus network traveling time is relatively smaller.

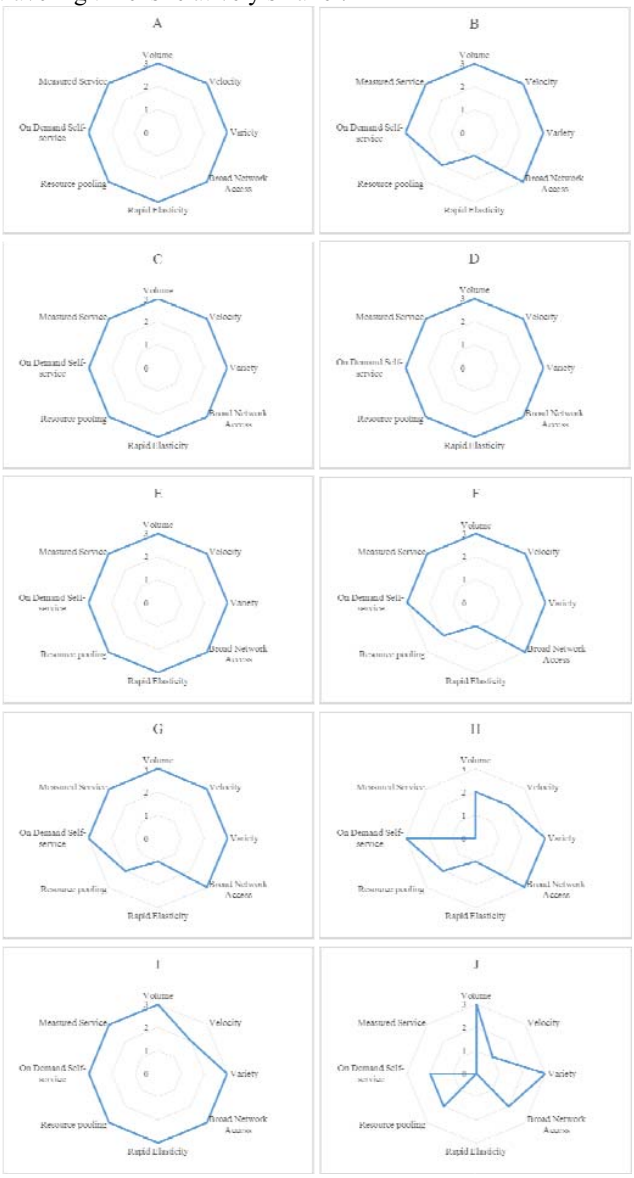

Figure 3. Chart of relation analysis point

Relation between BDaaS and the service types is that BDaaS should be included in Service Type I, II strictly as shown in Figure 3. The reason is that rapid elasticity is important factor for BDaaS. Loosely, service type III is optional.

Finally, we drew a concept in Figure 4 based on our survey. Cloud might support storage and computing resources. Data is managed and analyzed on the Cloud. As a result of analysis, various service for user will be made. This environment might

be called Big Data-as-a-Service(BDaaS). Data analysis and management of Big Data are able to become better by Elasticity of Cloud.

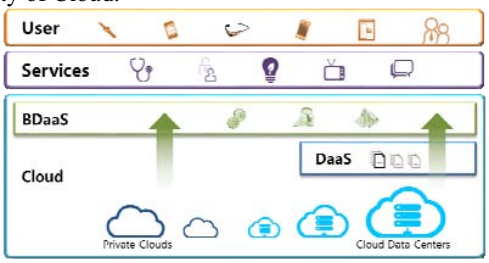

Figure 4. Big Data-as-a-Service Concept

IV. Conclusion and Future Work

As importance of Big Data has been incresed, Cloud-based Big Data Services were provided from verious enterprises. However, these BDaaS was hard to evaluate due to lack of Cloud and Big Data co-relation criteria. Therefore, each enterprises had different criteria so that these problem makes customers confused and customers and even providers hard to know which services were veritable Cloud-based Big Data service. Furthermore, it is hard to make a decision that how much services such as resource provided to be a good service. For this reason, we surveyed and analyzed BDaaS enterprise services to make criteria for BDaaS evaluation. Through enterprise services analysis, we investigated what kind of services were provided and organized the investigation result.

After select requisites in Big Data and Cloud, we point each requisite from 0 to 3. Total eight requisites were selected and used to evalute solution of enterprises. By evaluation, we found some patterns and each pattern became a service type as a criteria. Service type I satisfied eqaully between Big Data and Cloud, service type II satisfied more Cloud than Big Data, type III satisfied more Big Data than Cloud, type IV partially satisfied both Big Data and Cloud and type V was out of Cloud-based Big Data service, we concluded

Defined BDaaS in this paper is a service model which is meaningful information from large, complex dataset using Cloud resources. BDaaS can be classified and evalutated by proposed criteria service type. We suggest that Big Data should be included in Service type I and II strictly. Loosely, service type II can be opional. Lastly, We expect that this criteria can be used to evaluate Big Data-as-a-Service.

Our next step is to expand the point system which was used to evaluate Cloud-based Big Data solution to include not only general cases but also abnormal cases. Furthermore, we will suggest standardized BDaaS framework and following requirement, function will define. Finally, our goal is to implement BDaaS.

Acknowledgment

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (R0166-15-1011, Development of Interoperable Standards for Big Data System) and (R-20150223-000247, Cloud Storage Brokering Technology for Data-Centric Computing Standardization) (Corresponding Author: Eui-Nam Huh)

References

[1] HIBA JASIM HADI, AMMAR HAMEED SHNAIN, SARAH HADISHAHEED, AZIZAHBT HAJI AHMAD, “BIG DATA AND FIVE V’S CHARACTERISTICS”, IRF International Conference, 2014. 11.

[2] Xindong Wu, Xingquan Zhu, Gong-Qing Wu, Wei Ding, “Data Mining with Big Data”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 26, NO. 1, 2014. 01.

[3] Eli Collins, “Bringing Big Data Systems to the Cloud”, IEEE Cloud Computing, 2014.

[4] Krishnaprasad Thirunarayan and Amit Sheth, “Semantics-Empowered Approaches to Big Data Processing for Physical-Cyber-Social Applications”, Association for the Advancement of Artificial Intelligence (www.aaai.org), 2013.

[5] Zaheer Khan, Ashiq Anjum, Kamran Soomro and Muhammad Atif Tahir, “Towards cloud based big data analytics for smart future cities”, Journal of Cloud Computing: Advances, Systems and Applications, 2015.

[6] ISO/IEC 17789, “Information Technology -- Cloud Computing -- Reference Architecture”, ISO/IEC, 2014. 10.

[7] Mohammad Sajid, Zahid Raza, “Cloud Computing: Issues & Challenges”, International Conference on Cloud, Big Data and Trust, 2013. 11.

[8] Lihong Jiang, Li Da Xu, Hongming Cai, Zuhai Jiang, Fenglin Bu, and Boyi Xu, “An IoT-Oriented Data Storage Framework in Cloud Computing Platform”, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 10, 2014. 5.

[9] Rasu B. Shrestha, “Big data and cloud computing”, Applied Radiology, 2014. 03.

[10] Zibin Zheng, Jieming Zhu, and Michael R. Lyu, “Service-generated Big Data and Big Data-as-a-Service: An Overview”, IEEE International Congress on Big Data, 2013.

[11] S. Lohr, “The age of big data,” New York Times, vol. 11, 2012.

[12] “ Challenges and opportunities with big data,” leading researchers across the United States, Tech. Rep., 2011.

[13] “Big data-as-a-service: A market and technology perspective,” EMC Solution Group, Tech. Rep., 2012.

[14] J. Horey, E. Begoli, R. Gunasekaran, S.-H. Lim, and J. Nutaro, “Big data platforms as a service: challenges and approach,” in Proceedings of the 4th USENIX conference on Hot Topics in Cloud Ccomputing, ser. HotCloud’12, 2012, pp. 16–16.

[15] M. Rouse, “Definition of big data analytics,” “http://searchbusinessanalytics.techtarget.com/definition/big-dataanalytics”, January 2012.

[16] “ Why big data analytics as a service?” “http://www.analyticsasaservice.org/why-big-data-analytics-as-aservice/”, August 2012.

[17] Matt Turck, “The State Of Big Data in 2014: a Chart”, FirstMark Capital, 2014.05.

[18] Doug Henschen, “16 TOP Big Data Analytics Platforms”, InformationWeek, 2014.01.