Journal of Knowledge Information Technology and Systems(JKITS), Vol. 15, No. 3, pp. 347-355, June 2020
https://doi.org/10.34163/jkits.2020.15.3.004

Journal of Knowledge Information Technology and Systems

ISSN 1975-7700

http://www.kkits.or.kr

A Study on Adaptive Smart Platform for Intelligent Software in Big Data Environment

Jeong-Sig Kim[1], Jinhong Kim[*2]

[1]*Department of Computer and Mobile Convergence, Gyeonggi University of Science and Technology*
[2]*Department of Computer Engineering, Paichai University*

ABSTRACT

Since smart platform convergence became a main issue of 4th industrial revolution, intelligent software-centric industry has carried out a policy from Internet of Things/Everything, Big Data, Artificial Intelligent, and Deep Learning. Now a day, this highest super technology provide user's convenience, awareness, adaptation, and reactivity from intelligent software of IT. From their user-oriented service, we could make a great new leap forward to quality of software product. In addition, various convergence system of super national level is being made to stimulate both economy and industry. Nevertheless, it must be considered that the scale of these infrastructures is very large and that the conventional methods of inspecting infrastructures are very fast and time-consuming. These conventional methods are also dangerous for inspection team since the inspectors need to move or even climb on massive infrastructures to inspect places and areas that are difficult to reach. Besides, most conventional inspection methods are visual and the approaches applied are manual. Because of these problems mentioned above, adaptive smart platform and intelligent software have become important issues for users-aspect, industry-side, and Nation-oriented as time goes on. Accordingly, in this research paper, we propose adaptive smart platform that intelligent software have analyze, process, and refinement by big data. This platform is designed based on convergent software framework, such as Hadoop, HDFS and so on architectures utilizes our platform analyze massive data on their field by streamed data processing.

KEYWORDS: Smart Platform Convergence, 4th Industrial Revolution, Intelligent Software, Internet of Things/Everything, Big Data, Artificial Intelligent, Deep Learning, Highest Super Technology, Hadoop, HDFS

ARTICLE INFO: Received 23 May 2020, Revised 1 June 2020, Accepted 11 June 2020.

[*]Corresponding author is with the Department of Computer Engineering, Paichai University, Baejae-ro, Seo-Gu, Daejeon, 35345, Korea
E-mail address: jinhkm@pcu.ac.kr

1. Introduction

This research paper is about to basic concepts with adaptive smart platform for intelligent software relating to big data [1]. It a well-defining attempt to integrate that make up big data, what metrics clarify the size and each characteristics of big data, and what methods and technologies consist in harness of using intelligent software for adaptive smart platform. Big data are the main concept to appeared in new technological developments [2]. In addition, big data-driven intelligent software is read to acceptance of their concept rapidly, and also, adaptive smart platform advance in high technology have increasingly complex and large-scale computer and communication systems [3]. Such as, Internet of things has exponentially increased the scale and complexity of existing their systems that adaptation is thus an imperative property for Internet of Everything. After all, main concept of this research paper is to bring 4th industrial revolution by big data environment, that is charge of commercial efforts by large intelligent software and smart platform. So to speak, big data focuses on 'Prediction', 'Adaptation', 'Well-defined Structure', 'Analysis and Visualization'.

This paper is organized as follows. In Section 2 details the Adaptive Smart Platform. Then Intelligent Software in Section 3. Also, Section 4 presents Big Data-based Adaptive Smart Platform. Finally, Section 5 concludes the paper by briefly summarizing the main points.

2. Adaptive Smart Platform

The state of infrastructures and their technology represent the level of development in each platforms [4]. They remain and occupy an advanced position among the most developed technological and scientific project/research, it is important to maintain the state of their infrastructures in high quality condition/ consequency to improve their performance, sustainability, and reliability [5]. Smart Platform importance of reliable infrastructures in user's life, as well as the enormous costs associated with rebuilding infrastructures, require the deployment and implementation of predictive maintenance programs based on intelligent software techniques, accompanied by periodic and corrective maintenance where needed [6-8]. Nevertheless, it must be considered that the scale of these infrastructures is very large and that the conventional methods of inspecting infrastructures are very fast and time-consuming. These conventional methods are also dangerous for inspection team since the inspectors need to move or even climb on massive infrastructures to inspect places and areas that are difficult to reach [9-11]. Besides, most conventional inspection methods are visual and the approaches applied are manual [12]. Because of these problems mentioned above, adaptive smart platform and intelligent software have become important issues for users-aspect, industry-side, and Nation-oriented as time goes on. Eventually, we describe the 'A', 'S', 'P'. 1) *A*; *A*daptation mean that platform learn by itself what users and software is

operating as well as to operate in the designed environment without changing their system. 2) *S*; *S*mart means software and hardware with knowledge and intelligence. In addition, smart is a digital device that allows users to expand and reconfigure functions. *P*; *P*latform is a technology that enables product development or a process that supports current or future development. And also, this is a common execution environment that provides various functions. Therefore, adaptive smart platform is adaptation technologies that are used as based upon with intelligent SW (applications), processes or basic hardware-oriented, as well as use specific machine learning-based language with cooperative different platform.

3. Intelligent Software

Intelligent software is an autonomous process that performs tasks on behalf of the user for a specific purpose. This is a system that does not exist independently and is part of an environment or operates within it. The environment here refers to an operating system, a network, and so on [13]. Intelligent software has a knowledge-based and reasoning function, and attempts to solve problems through information exchange and communication with users, or resources. Intelligent software has the ability to recognize changes in the environment by itself, take action to respond, and learn based on experience, and they don't only perform a given task passively, but has an active attitude to pursue its purpose with its own purpose [14]. Intelligent Software operates inside an adaptive smart platform and provides the following features [15]. 1) Autonomy. It operates by judging by itself without direct interference from people or other objects, and has some kind of control over their behavior or internal state [16]. 2) Social Ability. It can interact with user and other software component using a communication language [17]. 3) Reactivity. It recognizes the real world, users, such as graphical user interface, a set of different software component, and the Internet-like environment and responds appropriately in time to the changes that take place within them [18]. 4) Proactivity. This is not simply acting in response to the environment, but taking initiative and acting goal-oriented [19]. 5) Temporal continuity. This is not a process that simply processes a given input and displays the result and exits, but rather a continuous daemon-like process that runs in the front and pauses for a while on the back [20]. 6) Goal-driven. It performs complex high-level tasks. The work is divided into smaller detailed tasks, and the processing order is determined and processed [21-23].

4. Big Data-based Adaptive Smart Platform

Big Data-based Adaptive Smart Platform(ASP) provide highest technology platform to helped solve a various type data problems. Such problems involve organizing miscellaneous, high volume, and potentially messy data sources

to refine timely and actionable. This systems enabled new solutions to established data problems, such as search platform, social networks-based, and cloud computing and data-driven service providers. Recently, Big Data-based Adaptive Smart Platform also contribute to first line, business decisions making of industrial technology, including stock, finance, manufacturing, media, and so on. Big data platform perform data collection/indexing /retrieval, data pattern matching and abnormally detection, formal natural language processing for domain identification and semantic analysis, information extraction from abnormal data (image, audio, video data, or a variety of other data). We are shown characteristics of big data-based adaptive smart platform include as bellows:

1) Storage capacity; It collects various information from the big data environment.

2) Distributed components; It processes software component in massive connected by social network environment.

3) Data Classification; It classifies the variety of data from big data layers.

Eventually, Big platform technology is high-3V (Volume, Velocity, Variety) information, as well as it is important to enhance the capability of their systems. For big data-based adaptive smart platform, our approach provide three conceptual components: Data access patterns, Computation patterns, adaptation patterns, and we shown by <Figure 1>.

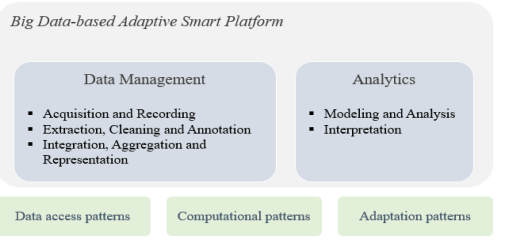

Figure 1. Concept of Big Data-based Adaptive Smart Platform

From the point of view of big data analytics, To take advantage of the multi processing that Hadoop provides, it needs to express data query with a MapReduce such as 'text analytics − social network feeds, emails, messenger, blogs, facebook, and instagram so on'. MapReduce process by breaking the processing with map phase and reduce phase, and each phase has key-value pairs of which specifies two functions; map function and reduce function. Map function is only a data preparation phase, setting up their data, and also a good place of to drop bad records. <Figure 2> shows MapReduce logical data flow.

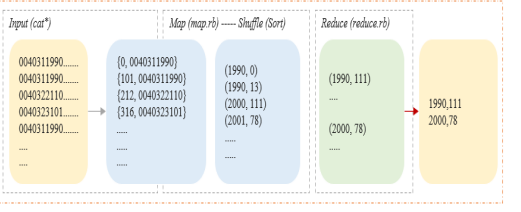

Figure 2. MapReduce Logical Data Flow
(This map function is represented by the Mapper,
that define abstract map( ))

As we mentioned above, MapReduce Job is a unit of operate that client-side wants to be performed. This consists of the input data (cat*), MapRduce, and configuration information. Hadoop

runs the job by dividing into tasks (*map tasks and reduce tasks*). There are two types of nodes that control the job execution with both jobtraker and tasktracker. Jobtraker run on the big data system by reports to keeps a record of the overall progress of each job using tasktracker. After all, Hadoop divides the input to MapReduce job into input splits, it creates one map task for each input splits, which runs the user defined map function for each record in their split. Splits is meant to the time taken to process, there is small compared to the time to process the whole input with by logical data flow in <Figure 2>. In order to most jobs, a good spilt size tends to the size of an HDFS block, although this could be changed for the cluster or specified when each file is created. So, we called data locality optimization, Hadoop is running the map task on a node or doing their best the map. Even though, HDFS block don't use valuable cluster bandwidth, but input split of map task is trying to run other map tasks, so the job scheduler will find a free map slot on a node in the same blocks. we are illustrated in <Figure 3>

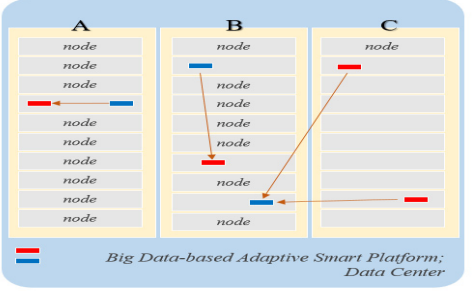

Figure 3. Hadoop using MapReduce of Data Center
(Red Block is map task, and Blue Block is HDFS block. Left A is Data-local, Center B is rack-local, and Right C is map tasks.)

When we have proposed data flow by each nodes for the general case, multiple reduce tasks is shown in <Figure 4>, and this diagram makes it clear how each data flow between map and reduce tasks. Futhermore, shuffle in <Figure 2> is more complicated that this diagram suggests, and tuning it can have a job execution time.

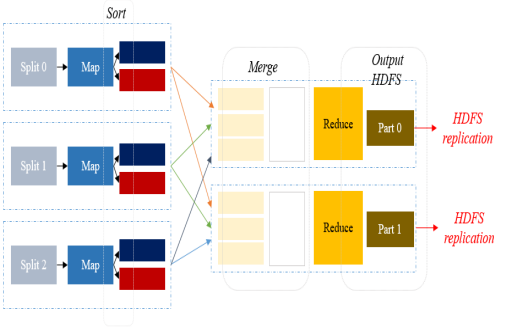

Figure 4. MapReduce Data Flow with Multiple Reduce Tasks

Specifying a combine function is defined using the MapReduce class for this big data-based adaptive smart platform. It is implementation as the MapReduce Job function using Java as bellows. This Job object forms the specification of the job and gives control over how the job is run. When we run this job on a Hadoop cluster, we will package the code which Hadoop will distribute around the cluster (JAR file). In other words, Job's setJarByClass() method use to locate the relevant JAR file by looking for the this file containing this class. Having constructed a Job object, we specify the input and output paths. After all, input path is specified by calling the static addInputPath() method on FileInputFormat, and it can be a single file, a directory, or a file pattern. Output path is also specified by the static

setOutputPath() method on FileOutputFormat, that it means a directory where the output files from the reducer functions are written by Figure 3 and 4.

```
public class MaxtextWithCombiner {
 public static void main(String[] args) ... {
  if (args.length !=2) {
  system.err.println("usage: MaxtextWithCombiner <input
  path>" + "<output path>");
  system.exit(-1);
  }
  Job job = new Job ();
  job.setJarByClass(MaxtextWithCombiner.class);
  job.setJobName("Max Text")

  FileIputFormat.addInputPath(job, new path(args[0]));
  FileOutputFormat.addInputPath(job, new path(args[0]));

  Job.setMapperClass (.....);
  Job.setCombinerClass (.....);
  Job.setReducerClass (.....);
  } ,..... }
```

5. Conclusions

In this paper, we have proposed adaptive smart platform for intelligent software in big data environment. The main idea of this research paper is to how we could have a concept big data. Other point a view, 3V of big data primary focus has been on analytics to obtain valid/valuable from adaptive smart platform. We also deal with Hadoop, HDFS, and their map flow. After all, big data platform, so we generally called 'Adaptive platform', 'intelligent platform', or 'smart platform', makes the case for new paradigm of their environment, as well as they have statistical methods manage for larger data sets. Technological advances of computations have enable data-driven, data classification, data optimization by using data mining, machine learning, deep learning for big data in timely manner. Especially, predictive analytics comprise a variety of big data techniques that predictive future based on historical/current data among their analytics. In addition, this analytics primarily based on statistical methods, so called Python and R, are developing new statistical methods for big data. Accordingly, our conceptual in adaptive smart platform derive on statistical significance the challenges of computational efficiency as following characteristics; Heterogeneity, Noise Accumulation, Spurious Correlation, Explanatory Data Variables, Predictor. Accordingly, from now to near future, we try to approach to integrate real data of industrial platform, in which the use of a informal data resource discovery, and manage the adaptation of smart-*, intelligent-*, and super connectivity world.

References

[1] M. Zolkepli, F. Dong, and K. Hirota, *Automatic switching of clustering methods based on fuzzy inference in bibliographic big data retrieval system*, International Journal of Fuzzy Logic and Intelligent Systems, Korean Institute of Intelligent Systems, Vol. 14, Issue 4. pp. 256-267, 2014.

[2] Y-J. Kim, and K-M. Lee, *Big numeric data classification using grid-based bayesian*

inference in the MapReduce framework, International Journal of Fuzzy Logic and Intelligent Systems, Korean Institute of Intelligent Systems, Vol. 14, Issue 4. pp. 313-321, 2014.

[3] J-H. Kim, and S-Ch. Kim, *Toward hybrid model for architecture-oriented semantic schema of self-adaptive system*, GPC 2013, LNCS 7861, Springer-Verlag, pp. 832-837, 2013.

[4] J-H, Kim, and E-S. Lee, *Semantic web recommender system based personalization service for user XQuery pattern.* In: Deng, X., Ye, Y. (eds.) WINE 2005. LNCS, vol. 3828, pp. 848–857. Springer, Heidelberg, 2015.

[5] J-D. Lee, and J-H. Lee, *Constructing efficient regional hazardous weather prediction models through big data analysis*, International Journal of Fuzzy Logic and Intelligent Systems, Korean Institute of Intelligent Systems, Vol. 16, Issue 1. pp. 1-12, 2016.

[6] J-S. Kim, and J-H. Kim, *A study on workflow processes for self-adaptive software system approaches*, Journal of Knowledge Information Technology and Systems(JKITS), KKITS, Vol.14, No.3, pp. 267-274, June 2019.

[7] S-H. Kim, D-S. Son, M-H. Park, and H-S. Hwang, *Developing a big data analytic model and a platform for particulate matter prediction: A case study*, International Journal of Fuzzy Logic and Intelligent Systems, Korean Institute of Intelligent Systems, Vol. 19, Issue 4. pp. 242-249, 2019.

[8] M. Nie, Z. Xiong, R. Zhong, W. Deng, and G. Yang, *Career choice prediction based on campus big data—mining the potential behavior of college students*, Applied Sciences, MDPI, Vol. 10, Issue 8, 14, Apr, 2020.

[9] J. Wu, K. Ota, M. Dong, J. Li, and H. Wang, *Big data analysis-based security situational awareness for smart grid*, IEEE Transactions on Big Data, IEEE, Vol. 4, Issue 3, pp. 408-417, 2018.

[10] P. O'Donovan, K. Leahy, K. Bruton, and D. T. J. O'Sullivan, *An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities*, Journal of Big Data, SpingerOpen Journal, Vol. 2 No. 25, 2015.

[11] R. Singh, and P. J. Kaur, *Analyzing performance of apache tez and MapReduce with hadoop multinode cluster on amazon cloud*, Journal of Big Data, SpingerOpen Journal, Vol. 3 No. 19, 2016.

[12] A. Gandomi, M. Reshadi, A. Movaghar, and A. Khademzadeh, *HybSMRP: A hybrid scheduling algorithm in hadoop MapReduce framework*, Journal of Big Data, SpingerOpen Journal, Vol. 6 No. 106, 2019.

[13] S-Y. Heo, *Implement of big data collecting system based on distributed messaging(BDCS-DM) using the rasberry pi*, Journal of Knowledge Information Technology and Systems(JKITS), KKITS, Vol. 14, No. 3, pp. 257-266, Jun. 2019.

[14] K-H. Kim, and Y-J. Kim, *A study on the enterprise's big data utilization*, Journal of Knowledge Information Technology and Systems(JKITS), KKITS, Vol. 14, No. 5, pp. 445-453, Oct. 2019.

[15] I. Yaqoob, I. A. T, Hashem, A, Gani, S. Mokhtar, E. Ahmed, N. B. Anuar, and A. V. Vasilakos, *Big data: From beginning to future*, International Journal of Information Management, Elsevier, Vol. 36, No. 6, pp. 1231-1247, Dec. 2016.

[16] A. Siddiqa, I. A. T. Hashem, I. Yaqoob, M. Marjani, S. Shamshirband, A. Gani, and F. Nasaruddin, *A survey of big data management: Taxonomy and state-of-the-art*, Journal of Network and Computer Applications, Elsevier, Vol. 71, pp. 151-166, Aug. 2016.

[17] Z. Sun, K. Strang, and R. Li, *Big data with ten big characteristics*, ICBDR 2018: Proceedings of the 2nd International Conference on Big Data Research, ACM Digital Library, pp. 56-61, Oct. 2018.

[18] M. Ptiček, and B. Vrdoljak, *Big data and new data warehousing approaches*, ICCBDC 2017: Proceedings of the 2017 International Conference on Cloud and Big Data, ACM Digital Library, pp.6-10, Sep. 2017.

[19] K. Demertzis, K. Rantos, and G. Drosatos, *A dynamic intelligent policies analysis mechanism for personal data processing in the IoT ecosystem*, Big Data Cognitive. Computing, MDPI Journal, Vol. 4 No. 9, Apr. 2020.

[20] M. Shen, X. Tang, L. Zhu, X. Du, and M. Guizani, *Privacy-preserving support vector machine training over blockchain-based encrypted IoT data in smart cities,* IEEE Internet Things Journal, pp. 7702-7712, Jun. 2019.

[21] G. Napoles, M. Leon, I. Grau, and K. Vanhoof, *Fuzzy cognitive maps tool for scenario analysis and pattern classification,* In Proceedings of the 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), Boston, MA, USA, Vol. 6 No. 8, pp. 644-651, Nov. 2017.

[22] V. J. Hodge, S. O'Keefe, and J. Austin, *Hadoop neural network for parallel and distributed feature selection,* Neural Network. Vol. 7 No. 8, pp. 24-35, 2016.

[23] H. S. Munawar, S. Qayyum, F. Ullah, and S. Sepasgozar, *Big data and its applications in smart real estate and the disaster management life cycle: A systematic analysis*, Big Data Cognitive. Computing, MDPI Journal, Vol. 4 No. 2, Mar. 2020.

빅데이터 환경에서의 지능형 SW를 위한 적응형 스마트 플랫폼에 관한 연구

김정식[1], 김진홍[2]

[1]*경기과학기술대학교 컴퓨터모바일융합학과, 조교수*

[2]*배재대학교 컴퓨터공학과, 조교수*

요 약

스마트 플랫폼 컨버전스는 4차 산업 혁명의 주요 이슈가 된 이래로, 지능형 소프트웨어 중심 산업은 사물 인터넷, 빅 데이터, 인공 지능 및 딥 러닝에서 정책을 수행되었다. 오늘날, 이들의 최고 슈퍼 기술은 지능형 IT 소프트웨어에서 사용자의 편의성, 인식, 적응 및 반응성을 제공한다. 사용자 중심의 서비스를 통해, 우리는 소프트웨어 제품의 품질로 새로운 도약할 수 있다. 또한 경제와 산업 모두를 자극하기 위해 초국가 차원의 다양한 융합 시스템이 만들어지고 있다. 그럼에도 불구하고, 이러한 대규모의 인프라를 위해 빠른 적용 방법을 요구하지만, 많은 시간이 걸리는 문제점이 있다. 또한, 인프라 확산을 위한 소프트웨어 접근 방법에 있어서도 문제점이 따른다. 게다가, 기존의 방법들의 시각화 측면에 있어 수동적으로 수행된다. 이러한 문제점으로 인해 적응형 스마트 플랫폼과 지능형 소프트웨어는 사용자 측면, 산업 측면 및 국가

지향의 중요한 문제점으로 자리잡고 있다. 따라서, 본 논문에서는 지능형 소프트웨어가 빅데이터 분석, 처리 및 개선을 위한 적응형 스마트 플랫폼을 제안한다. 이 플랫폼은 Hadoop, HDFS와 같은 수렴형 소프트웨어 프레임워크를 기반으로 설계되었으며, 아키텍처는 스트리밍된 데이터 처리를 통해 현장에서 방대한 데이터를 분석하기 위함이다.

Acknowledgments

This work was supported by the research grant of Pai Chai University in 2020.

Jeong-Sig Kim he is he is Assistant Professor of Department of Computer and Mobile Convergence, at the Gyeonggi University of Science and Technology, Gyeonggi-do, and Korea. He respectively received his Ph.D. degrees in Electronic, Electrical and Computer Engineering from Sungkyunkwan University, Korea, in 2007. His research interests include smart vehicular network, smart platform, software engineering, wireless sensor networks, scalable reliable communication protocols, mobile computing, network security protocols, proxy caching systems.

E-mail address: arius70@gtec.ac.kr

Jin Hong Kim, he is he is Assistant Professor of Department of Information Computer Engineering at the Pai Chai University, Daejeon, and Korea. He respectively received his Ph.D. degrees in Electronic, Electrical and Computer from Sungkyunkwan University, Korea, in 2006. Dr. Kim was Information System Research Center at the University of New South Wales in Australia. Moreover, He also served or currently serving as a reviewer and Technical Program Committee for many important Journals, Conferences, Symposiums, Workshop in Big Data area. His research interests include Smart Vehicular Network, Smart Platform, Machine Learning, Artificial Intelligent, Intelligent Software, Intelligent Agent System, and Big Data. He is a life member of the KKITS.

E-mail address: jinhkm@pcu.ac.kr