Data Center Infrastructure Management Systemfor Efficient Utilization of Computing Resources

HeeJune Han, Heejun Yoon, and Sang Oh Park

Abstract—A data center is a physical or virtual central repository for storing, managing, and disseminating data and information related to a particular business, as well as providing computing resources for analyzing stored data. As science and technology of modern society develops, the size of experiment increases, and as the size and amount of data generated through experiments increase exponentially, the computing resources for analyzing this increase are also increasing. As the size of the analytical data increases, the effort and research for efficiently managing the computing resources have been continuously carried out. We also increase the number of experiments we support and the computing resources we need to manage as we grow. Therefore, this paper describes and proposes programs and structures for managing and monitoring increasing computing resources each year.

Research Keywords—Data Center, DCIM, Infrastructure, Resources, Management

1 Introduction

As you can see from the name GSDC (Global Science Experimental Data hub Center) of KISTI(Korea Institute of Science and Technology Information), we support various experiments from around the world such as ALICE[1] (Large Ion Collider Experiment), CMS[2] (Compact Muon Solenoid), LIGO[3] (Laser Interferometer Gravitational wave Observatory) etc. Because we are supporting not only all the world's experiments but also domestic experiments, the service environment, structure and requirements are different according to each experiments. In addition, resources for maintaining and providing services are increasing as the size of research, size of generated data, and input / output amount of data for analysis are large. Therefore, there is a lot of difficulty in constructing and managing the equipment in size and configuration changes every year.

- *HeeJune Han is with the Global Science experimental Data hub Center, Korea Institute of Science and Technology Information, Daejeon, 34141, Korea E-mail: hjhan@kisti.re.kr.*
- *Heejun Yoon is with Global Science experimental Data hub Center, Korea Institute of Science and Technology Information, Daejeon, 34141, Korea E-mail: k2@kisti.re.kr.*
- *Sang Oh Park (corresponding author) is with the School of Computer Science and Engineering, Chung-Ang University, 84, Heukseok-ro, Dong-jak-gu, Seoul, Republic of Korea, 06974 E-mail: sopark@cau.ac.kr*

In this paper, we introduce a system for managing physical infrastructure and computing resources such as CPU, memory and storage of KISTI-GSDC. The system is expected to solve the underprovisioned rack problem in the datacenter that Gartner analyst David Capuccio addressed at Gartner Symposium / ITxpo 2016[4]. The composition of this paper is as follows. Section 2 describes the concept and structure of the system, Section 3 presents the results and lastly Section 4 presents conclusions and future work.

2 System Concept and Architecture

In this section, we discuss the concept and structure of the system to manage the infrastructure of the data center. The front-end of our system is structured as shown in Fig.1.

Front-end systems include IPA[5], which manages user accounts, access to user groups and DNS, and Foreman[6], which helps system administrators provision and manage servers. And RackTable[7], which is a system that manages physical servers that GSDC services. The existing system, RackTables, had difficulty retrieving information about computing resources such as CPU, memory and storage. To solve these problems, we added the ability to query computing resources such as CPU, memory and storage based on the data stored in the database. The results can be viewed

Fig. 1. Configuration of Frontend System.

through a web browser such as chrome, firefox and internet explorer. The structure of the system is shown in Fig.2. First, it receives information about the device based on the data stored in the database. Second, we use information received to calculate the CPU core, memory information, and disk information in use. Finally, the retrieved result is displayed to the user through a web browser. The database uses mysql, which stores the hardware CPU, memory, and storage information for the device. The computing resources for the experiments being serviced based on the data stored in the database are checked and the results can be confirmed by the users through the php-based web page.

3 Result

In this section, we run the system and review its performance. As a result of the system performance, we were able to check the CPU, storage, etc. for each service we support. And you can see how much of your total computing resources are being used Fig.3 shows an example of the results of running the program.

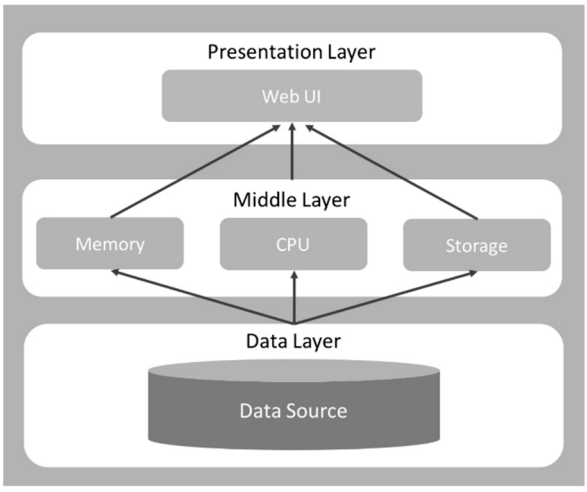
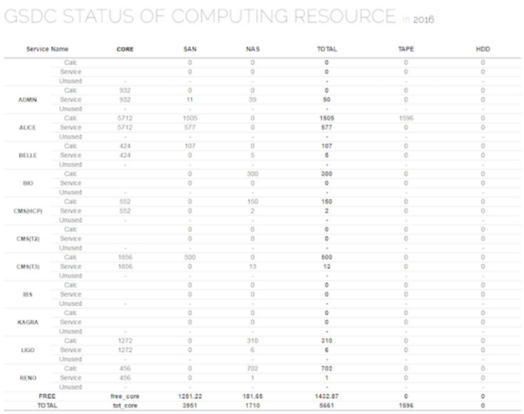


Fig. 2. System Architecture

4 Conclusions and Future Work

In this paper, we propose a system that enables system administrators to easily manage and monitor physical infrastructure and computing resources in the data center, and discussed the concept and system structure of the system. As system administrators have developed to make, infrastructure easier and easier to manage, we hope that this system will make it easier for system administrators to manage infrastructure and computing resources. Future work will be to create a system that can identify and monitor more detailed information such as temperature and load for all physical equipment supporting the service. It is not only replacing existing systems, but also being applied to existing and new systems without any problems.


GSDC STATUS OF COMPUTING RESOURCE in 2016

Service Name		CORE	SAN	NAS	TOTAL	TAPE	HDD
	CAK		0	0	0	0	0
	Service		0	0	0	0	0
	Unused		-	-	-	-	-
	CAK	932	0	0	0	0	0
ADMN	Service	932	11	39	50	0	0
	Unused	-	-	-	-	-	-
	CAK	5712	1505	0	1505	1596	0
ALICE	Service	5712	577	0	577	0	0
	Unused	-	-	-	-	-	-
	CAK	424	107	0	107	0	0
BELLE	Service	424	0	5	5	0	0
	Unused	-	-	-	-	-	-
	CAK		0	300	300	0	0
BIO	Service		0	0	0	0	0
	Unused		-	-	-	-	-
	CAK	552	0	150	150	0	0
CMS(HCP)	Service	552	0	2	2	0	0
	Unused	-	-	-	-	-	-
	CAK		0	0	0	0	0
CMS(T2)	Service		0	0	0	0	0
	Unused		-	-	-	-	-
	CAK	1856	500	0	500	0	0
CMS(T3)	Service	1856	0	13	13	0	0
	Unused	-	-	-	-	-	-
	CAK		0	0	0	0	0
IRS	Service		0	0	0	0	0
	Unused	-	-	-	-	-	-
	CAK		0	0	0	0	0
KAGRA	Service		0	0	0	0	0
	Unused		-	-	-	-	-
	CAK	1272	0	310	310	0	0
LIGO	Service	1272	0	6	6	0	0
	Unused	-	-	-	-	-	-
	CAK	456	0	702	702	0	0
RENO	Service	456	0	1	1	0	0
	Unused	-	-	-	-	-	-
FREE		free_core	1251.22	191.66	1442.87	0	0
TOTAL		tot_core	3951	1710	5661	1596	0

Fig. 3. Result of Computing Resource Status.

Acknowledgment

This work was supported by the program of the Construction and Operation for Large-scale Science Data Center, 2017, funded by the KISTI and by the program of the Global hub for Experiment Data of Basic Science 2017, funded by the NRF (N-16-NM-CR01).

References

[1] The ALICE Collaboration, “The ALICE experiment at the CERN LHC”, Journal of Instrumentation, 3, S08002, 2008, doi: 10.1088/1748-0221/3/08/S08002.

[2] Collaboration, C. M. S., and R. Adolphi. "The CMS experiment at the CERN LHC." JInst 3.08, 2008, doi: 10.1088/1748-0221/3/08/S08004.

[3] Althouse, William E., and Michael E. Zucker. "LIGO: The laser interferometer gravitational-wave observatory." Science, vol. 256, no. 5055, pp.325-333 1992.
[4] http://www.gartner.com/smarterwithgartner/top-10-technology-trends-impacting-infrastructure-operations/.
[5] IPA, https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/htmlsingle/Identity_Management_Guide/.
[6] Foreman, https://www.theforeman.org.
[7] RackTables, http://racktables.org.

HeeJune Han received a M.S from the Digital Informatics and Convergence at Chungbuk National University in 2016. He is currently a Researcher in the department of Global Science experiment Data hub Center at KISTI (Korea Institute of Science and Technology Information). His research interests are infrastructure management such as servers and storages of data center.

Heejun Yoon received the M.S from the Computer Engineering at Chungnam National University, KR in 1997. He has been serving as a Senior Researcher at Korea Institute of Science and Technology Information since 2000. His research interests include Large data processing, Parallel Computing, Cyber-Infra system, Computer Education.

Sang Oh Park received the B.S., M.S., and Ph.D. degrees from the School of Computer Science and Engineering at Chung-Ang University, in 2005, 2007, and 2010, respectively. He has been serving as an Assistant Professor of the School of Computer Science and Engineering at Chung-Ang University since 2017. He served as a Senior Researcher of Global Science experimental Data hub Center at Korea Institute of Science and Technology Information from 2012 to 2017 and a Research Professor at the School of Computer Science and Engineering. His research interests include high-performance computing, big data system, tape storage system, embedded system, cyber physical system, home network, smart factory and linux system.