Data Science Journal, Volume 12, 10 April 2013

A SCIENCE CLOUD FOR DATA INTENSIVE SCIENCES

Ken T Murata[1*], Shinichi Watari[1], Tsutomu Nagatsuma[1], Manabu Kunitake[1], Hidenobu Watanabe[1], Kazunori Yamamoto[1], Yasufumi Kubota[1], Hisao Kato[1], Takuya Tsugawa[1], Kentaro Ukawa[1,2], Kazuya Muranaga[1,2], Eizen Kimura[3], Osamu Tatebe[4], Keiichiro Fukazawa[5], and Yasuhiro Murayama[6]

[1]Space Weather and Environment Informatics Lab, National Institute of Information and Communications Technology (NICT), 4-2-1 Nukui-kita, Koganei, Tokyo 184-8795, Japan
**Email:* ken.murata@nict.go.jp
[2]Systems Engineering Consultants Co. Ltd. 4-10-1, Yoga, Setagaya, Tokyo 158-0097, Japan
Email: kentaro.ukawa@nict.go.jp
[3]Dept. Medical Informatics of Medical School of Ehime Univ., Situkawa, Toon City, Ehime, 791-0295, Japan
Email:ekimura@m.ehime-u.ac.jp
[4]Dept.Computer Science, Univ. of Tsukuba, Tenoudai 1-1-1, Tsukuba Science City, Ibaraki 305-8573, Japan
Email: tatebe@cs.tsukuba.ac.jp
[5]Research Institute for Information Technology, Kyushu University, 6-10-1 Hakozaki, Higashi-Ku, Fukuoka, 812-8581, Japan
Email: fukazawa@cc.kyushu-u.ac.jp
[6]Integrated Science Data System Research Lab, NICT, 4-2-1 Nukui-kita, Koganei, Tokyo 184-8795, Japan
Email: murayama@nict.go.jp

ABSTRACT

It is often discussed that the fourth methodology for science research is "informatics". The first methodology is a theoretic approach, the second one is observation and/or experiment, and the third one is computer simulation. Informatics is a new methodology for data intensive science, which is a new concept based on the fact that most scientific data are digitalized and the amount of data is huge. The facilities to support informatics are cloud systems. Herein we propose a cloud system especially designed for science. The basic concepts, design, resources, implementation, and applications of the NICT science cloud are discussed.

Keywords: Science cloud, Informatics, Observation, Experiment, Computer simulation, Data intensive science, Data-oriented science, Large-scale storage, Super computer

1 INTRODUCTION

During these past 50 years, along with the appearance and development of high-speed computers (and super-computers), numerical simulation has been considered to be the "third methodology" for science, following the theoretical (first) and experimental and/or observational (second) approaches (Figure 1). The variety of data yielded by the second approach has been getting larger and larger. This is due to progress in the technologies of experiments and observations. The amount of the data generated by the third methodology has also been getting larger and larger because of the tremendous development in and programming techniques of super computers.

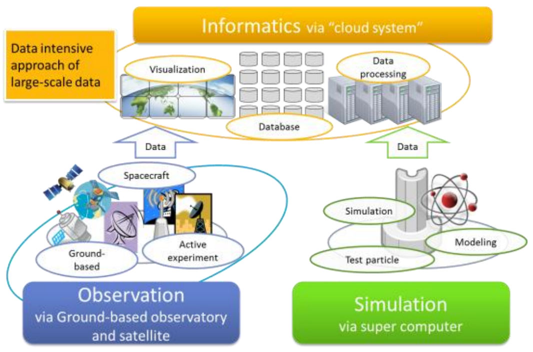

Figure 1. Informatics and the Cloud System

Most of the data files created by both experiments/observations and numerical simulations are saved in digital formats and analyzed on computers. The researchers (domain experts) are interested in not only how to make experiments and/or observations or perform numerical simulations but also in what information (new findings) can be extracted from the data. However, the data do not usually tell anything about the science. This is implicitly hidden in the data. Researchers have to extract information to find new science from the data files. This is the basic concept of data intensive (data oriented) science.

As the scale of experiments and/or observations and numerical simulations gets larger, new techniques and facilities are required to extract information from such large numbers of data files. This technique is called "informatics", which is the fourth methodology for new science.

Any methodologies must work on their facilities. For example, the space environment is observed via spacecraft, and numerical simulations are performed on super-computers. The facility of informatics, which deals with large-scale data, is a computational cloud system for science.

This paper proposes a cloud system for informatics that has been developed at NICT (National Institute of Information and Communications Technology), Japan. The NICT science cloud, "OneSpaceNet (OSN)", is the first open cloud system for the scientist who wishes to carry out informatics on his/her own science.

2 NICT SCIENCE CLOUD (ONESPACENET)

2.1 Overview of the Science Cloud

As discussed above, as the size of data files in any type of science gets larger, we need a new paradigm to analyze the data: that is informatics, the fourth methodology for science as shown in Table 1. The facility to support the fourth methodology is the science cloud (Microsoft, 2009).

The science cloud is not for simple uses. Many functions are expected from the science cloud, such as data standardization, data collection and crawling, a large and distributed data storage system, security and reliability, databases and meta-databases, data stewardship, long-term data preservation, data rescue and preservation, data mining, parallel processing, data publication and provision, the semantic web, 3D and 4D visualization, out-reach and in-reach, and capacity building.

Table 1. Four methodologies and their facilities

	methodology	facility
First	theory	human being
Second	experiment/observation	e.g., spacecraft
Third	numerical simulation	super computer
Fourth	informatics	cloud system

Figure 2 is a schematic picture of the NICT science cloud. It should be noted that there are two types of observation data, and both are stored in the storage system in the science cloud. One type is from archive sites outside of the cloud: this is data to be downloaded through the Internet to the cloud. The other is data coming from the equipment directly connected to the science cloud. These are often called sensor clouds. One of the great advantages of the scientific cloud over other legacy systems is its integrated function. A large-scale disk area is provided to users but not necessarily for data file storage. For instance, cluster systems with parallel data processing are also mounted in the NICT science cloud. Since each node is responsible for both a data file node and a data processing node, users don't have to copy (or move) large size data files to their data processing system sites. Note that it usually takes more than one week to copy data files with 10TB over the Internet.

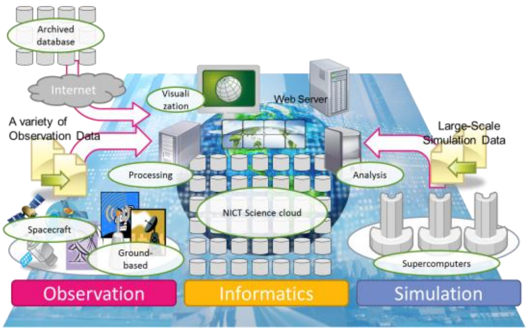

Figure 2. A basic concept of NICT science cloud

2.2 Implementation of the NICT Science Cloud

One of the definitions of a cloud system is its multi-functionality. It has to satisfy a variety of requests from users. Science clouds must be, in general, more functional than commercial clouds. The providers of commercial cloud provide simple services, which mainly work on virtual machines. Science clouds have to provide two different types of services: test-beds for developers of cloud computing (science/technology for the cloud) and facilities to perform high-level science (the cloud for science/technology). This suggests that the science cloud has to be equipped with many resources to satisfy both the science/technology of the cloud and the cloud of the science/technology. The science cloud must be designed and implemented for such a variety of intensive data processing studies.

Figure 3 is a schematic picture of the current implementation of the NICT science cloud. The resources in the figure were installed by the end of 2011 and will be developed step by step in the following years. The cloud system is composed of several clusters of computational resources deployed over the JGN-x network. The JGN (Japan Gigabit Network) is a wide-area network with 10Gbps or more, covering all Japan from Hokkaido to Okinawa. Most of the access points (APs) are located at research institutions or universities.

What should be paid attention to in the NICT science cloud is that the network over JGN-x is an L2 (layer-2) network. Wide-area network systems are often constructed with L2 because of its easy maintenance and security. The L2 network also has an advantage over L3 in terms of its routing-less data transfer.

Because the NICT science cloud is a wide-area (domestically distributed) cloud, data transfer speed inside the cloud is important to performances of the cloud. To avoid long-distance routing inside the cloud through routers widely distributed over JGN-X (over Japan), the L2 network is preferable to the L3 network. Another reason is security: an L2 network is a closed network; thus the traffic between nodes in the cloud stays inside the cloud.

The most important resource in Figure 3 is a set of large-scale storages. In the NICT science cloud, we deployed a set of distributed file storages over the JGN-x. These computing resources are discussed in Section 2.3.

In Figure 3 four super-computers are connected to the science cloud. The virtual super-computing environment also plays important roles in the science cloud because it usually yields a large size of data to be processed and visualized. For such post-processing, a parallel computing environment must be in use. Here, note that these parallel computers are either cluster type or hetero-type.

Large-size displays are also necessary for the use of the science cloud. Because the spatial sizes of numerical simulations get larger (Murata et al., 2007), high-resolution displays are required to preview visualized data without data compressions (with full resolutions). This will be discussed in Section 3.4.

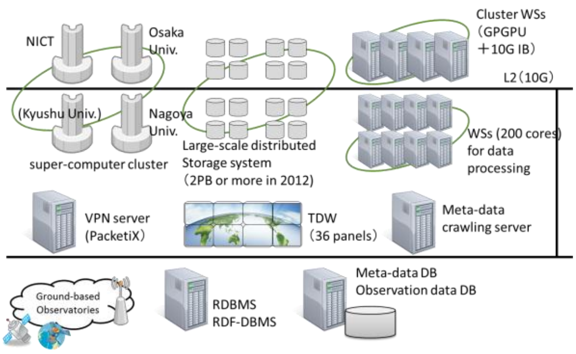

Figure 3: Implementation of NICT science cloud

2.3 Distributed Storage System

As discussed above, one of the most important resources in science clouds is their data storage systems. Because we need to store all of the digital data from both observations/experiments and simulations, it is crucial to have data storage with sufficient size and high performance.

In the NICT science cloud, we constructed a distributed data storage system named “OSN cloud storage”, equipped over the JGN-x. Because we have many access points throughout the JGN-x, we were able to use the data storage at the data centers (DCs) of the JGN-x. To construct the distributed data storage system, we adopted the Gfarm (Grid Datafarm), which is a middleware for such wide-area distributed data storage systems (Mikami et al, 2011; Kobayashi et al., 2011). The newest version of Gfarm is 2.5. We used 2.4.2 because it is the most stable version of Gfarm.

The deployment of the Gfarm DCs at present is displayed in Figure 4. There are several advantages to these widely distributed storage systems that use Gfarm. The most important function of Gfarm for OSN cloud storage is data file redundancy. Once a user drops (saves) a data file on the OSN cloud storage, the system automatically make replications of the file that are saved at different nodes of the system. When another user accesses the data file, the system provides him/her with the most accessible data file (closest to the user). Thus the latency between Hokkaido and Okinawa, a distance of more than 3000km, is negligible. Using this system, a user in Hokkaido accesses the file instance located in the Hokkaido DC, and the user in Okinawa accesses that in the Okinawa DC.

Another important function based on the redundancy of the OSN cloud storage is its backup free service. In Figure 4, two replications of each data file are automatically created (a total of 3 files on the Gfarm system) as discussed above. Assume that one of the data file instances is broken and lost. In this case, the system detects this loss and makes another replication of the file as soon as possible (see Figure 4). Therefore, the number of copies of the file in the system remains three. This suggests that neither backup nor restoring is necessary. It usually takes a few days to a one week to restore a large size, 10TB for example, of backup data files. This often stops the researchers' work. The concept of the BCP (Business Continuity Plan) or BCM (Business Continuity Management) should be applied not only for business but for scientific work. To obtain good research results, such continuity is expected in a research environment.

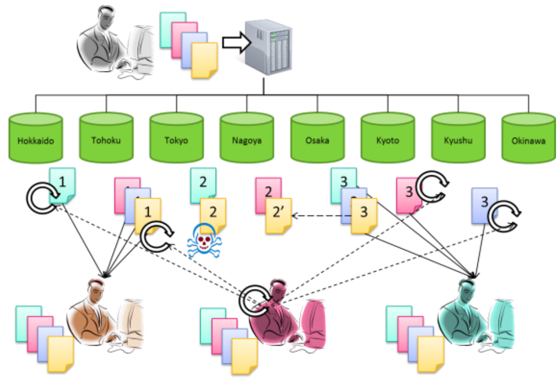

Figure 4: Distributed Storage in the NICT science cloud (OSN cloud storage) at data centers (DCs). DCs at Hokkaido, Tohoku, Kyoto, and Kyushu are under contemplation.

The availability of the distributed storage system is also important from the viewpoint of cost performance. The price of data storage (Hard Disk Drive: HDD) is getting lower, suggesting that scalable storage additions are more reasonable than equipping large-size storage. Figure 5 is a time-dependent graph of the OSN cloud storage size since October, 2009. Note that the total storage size changes frequently. This means that we have often added or removed small size file system nodes (with about 50TB) on the OSN cloud storage system without terminating service. In the OSN cloud storage system, one file node with 50TB cost about $5,000 (US dollars); this means the cost for 1TB is about $100. We don't need a large budget to develop the storage system.

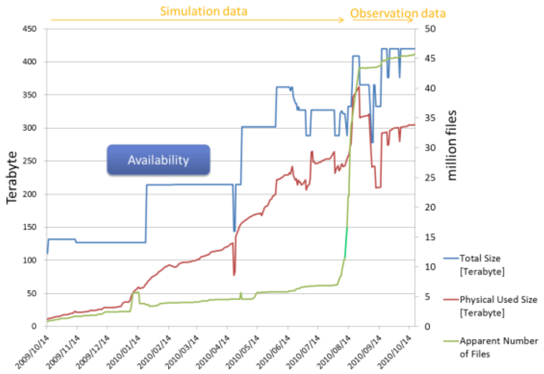

Figure 5: Availability of OSN cloud storage

3 APPLICATIONS OF THE NICT SCIENCE CLOUD

3.1 Overview

As discussed above, the science cloud must be applied for a variety of data intensive research work. In this section, we discuss a few good examples of the NICT science cloud usage. Because it is not quite one year since the opening of the science cloud, we have not yet archived any outstanding results on the cloud. Most of them are, thus, initial reports. However, these initial results are valuable because they have been derived with the support and use of the science cloud.

3.2 Large-scale visualization

One effective target of the use of the science cloud is visualization of numerical simulations. As the development of super-computers has increased, the size of simulation data has gotten larger. The recent trend of the spatial size of numerical simulation is 1000^3 (Giga) as the main memory size increases. Herein we consider a computer simulation via the Global MHD simulation code (Fukazawa et al., 2006). The 1000^3 grid number corresponds to about 10 to 100GB data size as 10 to 100 components are allocated on each grid.

In most time-dependent computer simulations, such as Global MHD simulations, we have discarded most of the numerical data with sampling in time. For example, the time resolution of the present simulation is 0.5 sec., but the usual sampling time of visualization is 1 min. This suggests that 99.2% of the total simulation data are not used.

The major reason to discard most of the data is post-processing. The data size for all of the data is almost 15TB for 2 hours of simulation with 0.5 sec. time resolution. Figure 6 shows the comparison of post-processing times between the legacy method and parallel visualization via the NICT science cloud. Even if one has large-scale storage for this 15TB data, it is unreasonable to analyze (visualize) the data with the legacy method. The left-hand side of the panel in Figure 6 shows that it takes 18 days to simply read this 15 TB data. Visualization with 1 core (1 CPU) takes 16 days, which is also too long.

We are now developing a new parallelization method to visualize such large-scale simulation data using 14 hetero-machines in the NICT science cloud, as shown in the right-hand side panel in Figure 6. Theoretical examination estimates that the data read (I/O) time will be as short as 30 min., and visualization will be parallelized so that it will take only 4 hours.

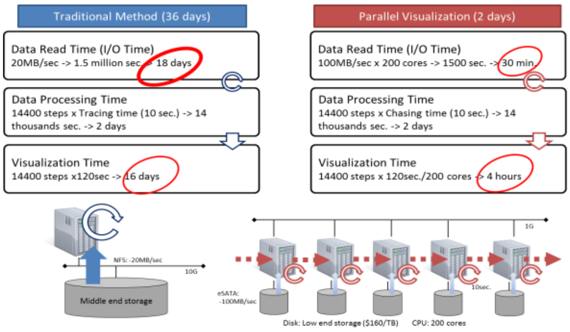

Figure 6: Comparison of post-processing times between legacy method and parallel visualization via NICT science cloud (14 cores).

We have developed a proto-type of the parallel visualization for Global MHD simulations. The performance will be reported in other papers, but it took less than one day and obtained some new visualization results as shown in Figure 7.

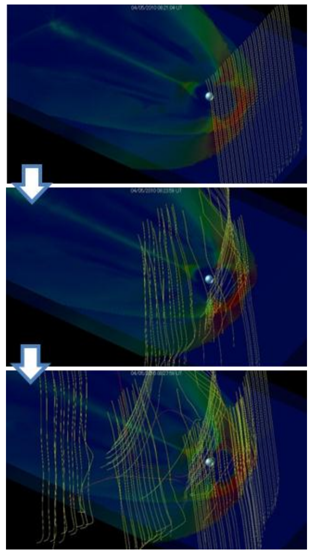

Figure 7: High time resolution visualization of Global MHD simulation

3.3 Data Collection (Crawling)

In Figure 2, we discussed that observation and/or experiment data are transferred to the cloud storage. However, apart from the simulation data that comes from super-computers directly connected to the cloud as shown in Figure 2, observation data are usually stored and managed at institutions out of the cloud. We need to independently collect such public data from these institutions through the Internet. For data processing, especially long-term data processing, on-demand data collection systems often do not work because the data file download time takes a longer time than the processing time. To avoid these issues, an automatic data collection (crawling) system is crucial.

Figure 8 shows an example of the number of daily data of permanent GPS receivers collected for research studies in NICT. The number of data files to be collected is increasing, and we currently need to collect more than 5000 data files a day via FTP/HTTP servers of more than 20 domestic and world-wide institutions. The policies or ways to obtain these data from each institution often change without any information given to users. This means that manual data file collection is now very difficult.

We have developed an automatic data collection (crawling) system that works in the NICT science cloud. The system has two functions: collection of data file information (metadata), called NICTY, and data file crawling based on metadata, called DLA (DownLoad Agent) (Ishikura et al., 2006). Figure 9 shows a procedure to collect metadata and data files using NICTY and DLA. These systems are already in use in the NICT science cloud, as shown in Figure 10. Metadata records number more than 9 million, and crawled data files number more than 5 million so far.

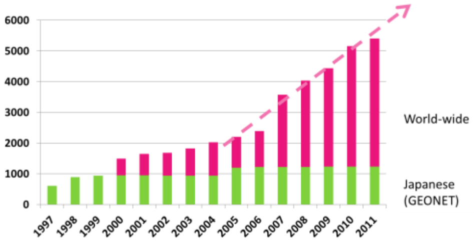

Figure 8: The annual trend of the number of GSP receivers (world-wide since 2000 and domestic (Japan) since 1997).

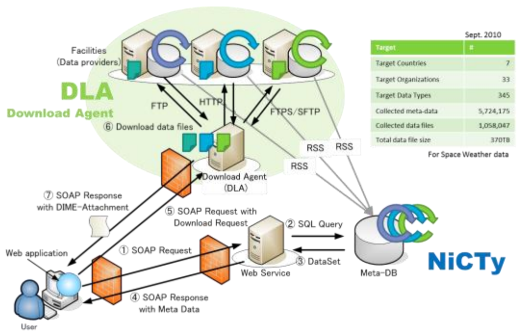

Figure 9: An automatic data crawling system on the NICT science cloud: NICYT and DLA (DownLoad Agent)

3.4 High Resolution Display

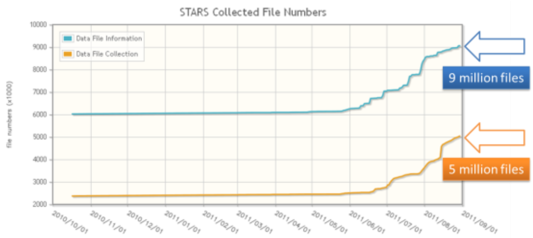

Figure 10: Automatic data crawling system on the NICT science cloud: NICYT and DLA (DownLoad Agent)

As the size of computer simulation becomes bigger, the spatial resolution of visualized data gets larger. To preview or analyze such large-scale visualized data, we need a high-resolution display that directly refers to the OSN cloud storage. (Otherwise, large scale data must be transferred to the display before previewing. This takes a long time and leads to poor usability.) The Tiled Display Wall (TDW) is one possible solution in the near future for the issue of high-resolution visualization. Figure 11 is a picture of a TDW in the NICT main hall. The TDW is composed of 25 panels, and each panel has SXGA resolutions. Here we should note that for the best performance of the TDW we need to make use of the science cloud. The typical resolution of TDWs is more than 10 HD (HD is high-vision with 1024 x 2048 resolution). It requires more than 1Gbps data transfer. This suggests that we need to prepare a 10Gbps network to connect both master server and client servers of TDW. Figure 12 shows a plan to transfer movie data to a TDW using a 10Gbps network and the distributed storage system discussed in Section 2.3. The I/O time will be a bottle-neck for the data transfer. Note that the nominal I/O speed of an eSATA disk is as low as 100MB, which is slower than 10Gbps. The authors have a plan to make a parallel data transfer from Gfarm storage nodes to a TDW with a speed as high as 10Gbps using a memory mapping technique.

Figure 11: A Tiled Display Wall at NICT main hall

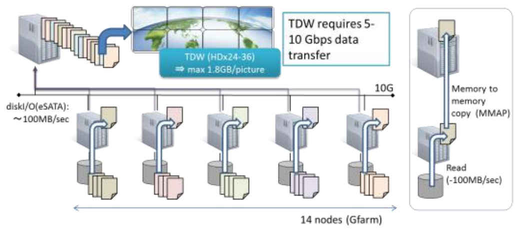

Figure 12: High-speed data transfer to TDW

4 CONCLUSION

The first workshop of the science cloud was held in February 2010 in Chicago, IL, USA (Grossman et al., 2010). This implies that the concept of the application of a cloud system to science is rather new, and nobody has ever succeeded in the construction of a real science cloud. We need more trial-and-error to understand when our science cloud works and when it does not work.

Microsoft is one of the companies interested in science cloud systems: they seem to apply their own cloud service, Microsoft Azure (Microsoft, 2010), to science research, as was presented at the workshop above. However, one success on one science project does not necessarily mean the cloud is successful. The science cloud is an environment of a variety of sciences, and an environment in which any researcher may make his/her own customizations. We often describe a cloud system as having three layers: SaaS (Software as a Service), PaaS (Platform as a Service), and IaaS (Infrastructure as a Service).

In Figure 13 we compare the NICT science cloud, "OneSpaceNet" (for space weather works), with other famous cloud services. So far, the Google cloud services are SaaS and PaaS. Users mainly make use of the Google cloud service through the Web. Another cloud service via Amazon is PaaS and IaaS based. It provides storage and computational resources but has no application services.

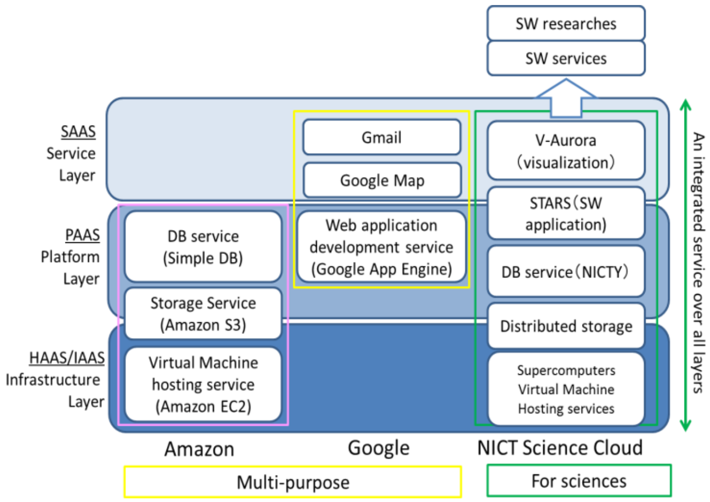

Figure 13. Comparison of the NICT science cloud for space weather research works with other famous cloud systems

Murata et al. (2005) have attempted to construct a software system to merge different types of data from both spacecraft observations and numerical simulations. Their trial was so challenging that both data types were simultaneously previewed on 4D (3D in space and time) space. However, the system was not completed because the inside system was too complicated and data size was too large even though all of the software design was based on object-oriented methodology (Murata et al., 2001). To develop such a large-scale and multi-functional system, we need a computational environment on which to construct the system. The concept of the science cloud is suitable for that. The cloud system generally provides a variety of functions required for science work.

In the present study, we propose a multi-functional science cloud and presented several studies based on this cloud system. The system is still in development, but several research works have started. Many issues are left for more practical uses of the system for research works. New research findings, which are not able to be obtained without the science cloud, are expected in the near future.

5 ACKNOWLEDGEMENTS

The present work was done by using resources of the OneSpaceNet (the NICT science cloud) and JGN-X (Japan Gigabit Network).

6 REFERENCES

Fukazawa, K., Ogino, T., & Walker, R.J. (2006) The Configuration and Dynamics of the Jovian Magnetosphere. *J. Geophys. Res. 111*, A10207, doi:10.1029/2006JA011874.

Grossman, R., Gu, Y., Sabala, M., Slides, P., Mambretti, J., Szalay, A., & White, K. (2010) An Overview of the Open Science Data Cloud. *The Science Cloud 2010*, Illinois, USA. Retrieved from the World Wide Web, February 19, 2013: http://datasys.cs.iit.edu/events/ScienceCloud2010/p01.pdf

Ishikura, S., Kimura, E., Murata, K.T. Kubo, T., & Shinohara, I. (2006) Automatic meta-data collection of STP observation data. *AGU Fall Meeting*, San Francisco, CA, USA.

Kobayashi, K., Mikami, S., Kimura, H., & Tatebe, A. (2011) The Gfarm File System on Compute Clouds. *Proceedings of 1st International Workshop on Data Intensive Computing in the Clouds (DataCloud 2011)*.

Microsoft (2009) The Fourth Paradigm: Data-Intensive Scientific Discovery. Retrieved from the World Wide Web, February 19, 2013: http://research.microsoft.com/en-us/collaboration/fourthparadigm/

Microsoft (2010) Windows Azure Platform. Retrieved from the World Wide Web, February 19, 2013: http://www.microsoft.com/azure/default.mspx

Mikami, S., Ohta, O., & Tatebe, O. (2011) Using the Gfarm File System as a POSIX compatible storage platform for Hadoop MapReduce applications. *Proceedings of 12th IEEE/ACM International Conference on Grid Computing (Grid 2011)*.

Murata, K.T., Fujita, S., & Tanaka, T. (2007) 3D Visualizations and Analyses of Magnetic Flux Rope. *AOGS 2007 4th Annual meeting*, Bangkok, Thailand.

Murata, K. T., Yahara, H., & Toyota, K. (2001) Software design via object-oriented methodology & network database for solar-terrestrial observation data. *Database System* 123-5, pp 31-36.

Murata, K.T. Yamamoto, K., Matsuoka, D., Matsumoto, H., Okada, M., Mukai, T., Sigwarth, J.B., Fujita, S., Tanaka, T., Yumoto, K., Ogino, T., Shiokawa, K., Tsyganenko, N.A., Green, J.L., & Nagai,T. (2005) Development of the Virtual Earth's Magnetosphere System (VEMS). *Advances in Polar Upper Atmosphere Research 19*, pp135-151.

(Article history: Available online 23 March 2013)