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[12], Kazuya Muranaga[12], 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 researches is "informatics"; the first methodology is theoretic approach, the second one is observation and/or experiment, and the third one is computer simulation. The informatics is expected as a new methodology for data intensive science, which is a new concept based on the fact that most of the scientific data are digitalized and the amount of the data are huge. The facilities to support informatics are cloud systems. Herein we propose a cloud system especially designed for science. The basic concepts, design, resource, implementation and applications of 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 50 years, along with appearance and development of high-speed computers (and super-computers), numerical simulation is considered to be a "third methodology" for science, following theoretical (first) and experimental and/or observational (second) approaches (Figure 1). The variety of data yielded by the second approaches has been getting more and more. It is due to the progress of technologies of experiments and observations. The amount of the data generated by the third methodologies has been getting larger and larger. It is because of tremendous development and programming techniques of super computers.

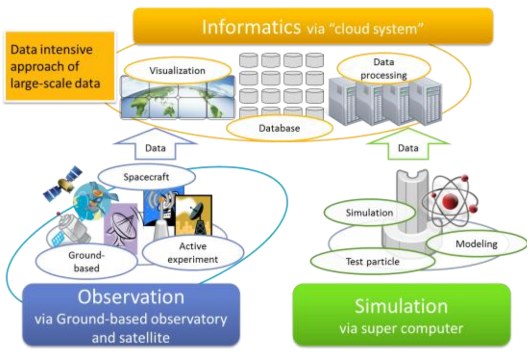

Figure 1. Informatics and Cloud System

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 what information (new findings) to extract from the data. However, data does not usually tell anything about the science; sciences are implicitly hidden in the data. Researchers have to extract information to find new sciences from the data files. This is a basic concept of data intensive (data oriented) science.

As the scales of experiments and/or observations and numerical simulations get larger, new techniques and facilities are required to extract information from a large amount of data files. The technique is called as "informatics" as a fourth methodology for new sciences.

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

This paper is to propose a cloud system for informatics, which has been developed at NICT (National Institute of Information and Communications Technology), Japan. The NICT science cloud, we named as "OneSpaceNet (OSN)", is the first open cloud system for scientist who is going to carry out their informatics for their own science.

2 NICT SCIENCE CLOUD (ONESPACENET)

2.1 Overview of Science Cloud

As is discussed above, as the size of data files in any types of science gets larger, we need a new paradigm to analyze the data: that is informatics, a 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 to the science cloud; such as data standardization, data collection and crawling, large and distributed data storage system, security and reliability, database and meta-database, data stewardship, long-term data preservation, data rescue and preservation, data mining, parallel processing, data publication and provision, 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. Both types of data from observation and simulation are stored in the storage system in the science cloud. It should be noted that there are two types of data in observation. One is from archive site out of the cloud: this is a data to be downloaded through the Internet to the cloud. The other one is data from the equipment directly connected to the science cloud. They are often called as sensor clouds. One of the great advantages of the scientific cloud to other legacy systems is its integrated function. A large-scale disk area is provided with the users, but not necessarily for the 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 data file node and data processing node, uses 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.

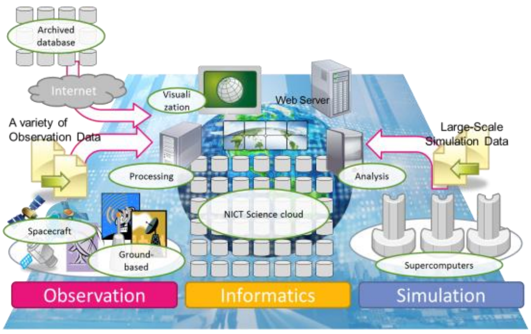

Figure 2. A basic concept of NICT science cloud

2.2 Implementation of NICT Science Cloud

One of the definitions of 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 types of different services: test-beds for developers of the cloud computing (science/technology for cloud) and facilities to perform high-level science (cloud for science/technology). It suggests that the science cloud has to be equipped with many resources to satisfy both science/technology of cloud and cloud of 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 implementation of the NICT science cloud so far. The resources in the Figure are installed by the end of 2011, thus 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 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 to L3 in terms of the routing-less data transfer.

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

The most important resource in Figure 3 is a set of large-scale storages. In the NICT science cloud, we deploy 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. Virtual super-computing environment also plays important roles in the science cloud, since they usually yield a large size of data to be processed and visualized. For such post-processing, 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. Since the spatial sizes of numerical simulations get larger (Murata et al., 2007), high-resolution display are required to preview visualized data without data compressions (with full resolutions). This will be discussed in Section 3.4.

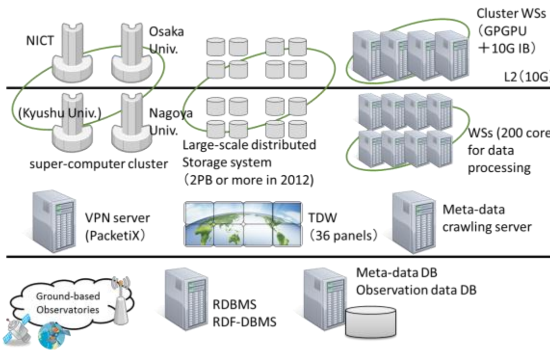

Figure 3: Implementation of NICT science cloud

2.3 Distributed Storage System

As discussed above, one of the most important resources in science clouds is data storage system. Since we need to store all of the digital data from both observations/experiments and simulations, the size and high performance of the data storage are crucial.

In the NICT science cloud, we construct a distributed data storage system named "OSN cloud storage", equipped over the JGN-x. Since we have many access points over the JGN-x, we are able to deploy data storages at the data centers (DCs) of the JGN-x. To construct the distributed data storage system, we adopt the Gfarm (Grid Datafarm), which is a middleware for such wide-area distributed data storage system (Mikami et al, 2011; Kobayashi et al., 2011). The newest version of the Gfarm is 2.5, and we use 2.4.2 since it is most stable version of the Gfarm.

The deployment of the Gfarm DCs so far is displayed in Figure 4. There are several advantages of these widely distributed storage system using the Gfarm. Most important function of the Gfarm for the 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 saved at different nodes of the system. When another user accesses the data file, the system provides him/her with most accessible data file (closest to the user). The latency between Hokkaido and Okinawa, the distance between them is more than 3000km, is negligible. Using this system, a user at Hokkaido accesses the file instance located in Hokkaido DC and the user at Okinawa accesses that in Okinawa DC.

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

The availability of the distributed storage system is also important from a viewpoint of cost performance. Since the price of data storage (Hard Disk Drive: HDD) is getting lower. This suggests that scalable storage addition is more reasonable than equipping large-size storage at once. 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 dollar); this means the cost for 1TB is about $100. We don't have to get large budget to develop the storage system.

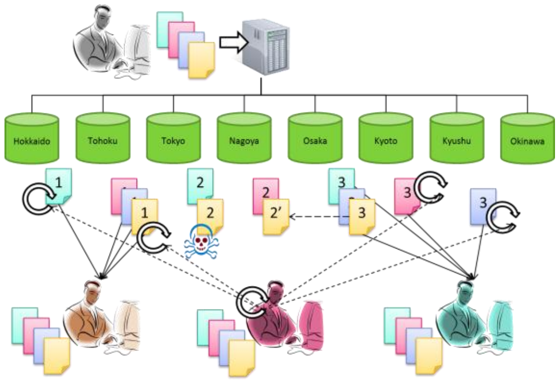

Figure 4: Distributed Storage in NICT science cloud (OSN cloud storage) at data centers (DCs). The DCs at Hokkaido, Tohoku, Kyoto and Kyushu are under contemplation.

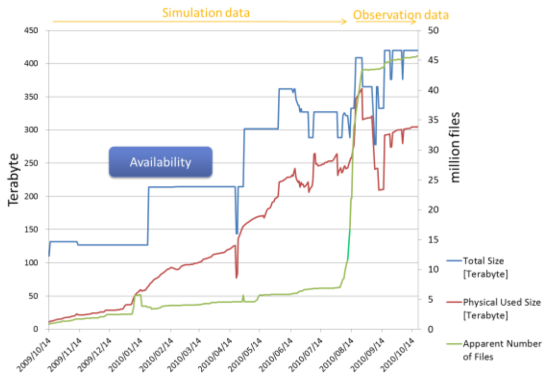

Figure 5: Availability of OSN cloud storage

3 APPLICATIONS OF THE NICT SCIENCE CLOUD

3.1 Overview

As discussed above, science cloud must be applied for a variety of data intensive research works. In this section, we discuss few examples of good use of the NICT science cloud. Since it is almost one year after opening of the science cloud, we have not archived any outstanding results on the cloud. Most of them are, thus, initial reports. However, these initial results are valuable since they are not derived without supports or uses of the science cloud.

3.2 Large-scale Visualization

One of the effective targets of use of the science cloud is visualization of numerical simulations. As the development of super-computers, the size of simulation data is getting 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 Global MHD simulation code (Fukazawa et al., 2006). The 1000^3 grid number corresponds to about 10 to 100GB data size since 10 to 100 components are allocated on each grid.

In most of the time-dependent computer simulations like 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 usual sampling time of visualization is 1 min. It suggests that 99.2% of total simulation data are not used.

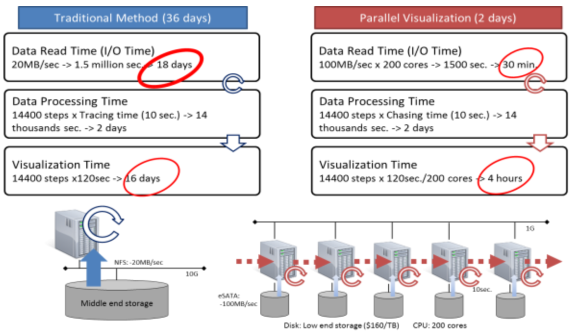

Figure 6: Comparison of post-processing times between legacy method and parallel visualization via NICT science cloud (14 cores).

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 simulation with 0.5 sec. time resolution. Figure 6 shows comparison of post-processing times between legacy method and parallel visualization via NICT science cloud. Even if one has a large-scale storage for this 15TB data, it is unreasonable to analyze (visualize) the data with 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 not enough, too.

We are now developing a new parallelization method to visualize such a 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.

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 within one day, and obtained some new visualization results as shown in Figure 7.

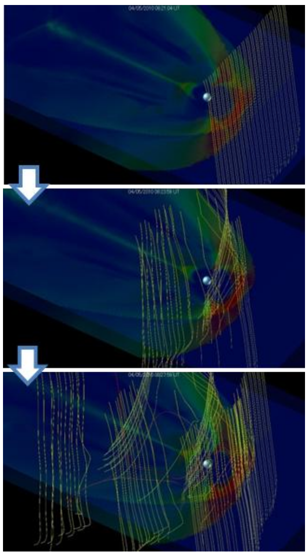

Figure 7: High time resolution visualization of Global MHD simulation

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 which 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 the institutions through the Internet. For data processing, especially long-term data processing, on-demand data collection system often does not work since the data file download time takes longer time than processing time. To avoid these issues, 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 need to collect more than 5000 data files a day via FTP/HTTP servers of more than 20 domestic and world-wide institutions now. The policy or way to provide these data from each institution often change without any information to users. It means that manual data file collection is now very difficult.

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

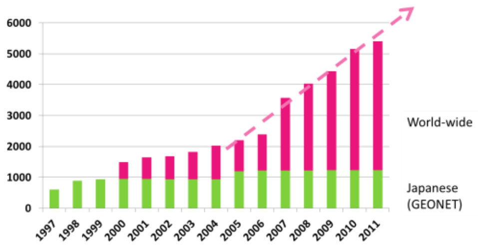

Figure 8: The annual trend of the number of GSP receivers (world-wide since 2000 and

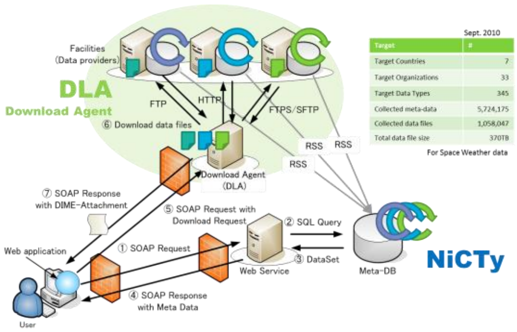

Figure 9: An automatic data crawling system on NICT science cloud: NICYT and DLA (DownLoad Agent).

3.4 High Resolution Display

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 which directly refers to the OSN cloud storage. (Otherwise, large scale data must be transferred to the display before previewing; it takes long time and leads to poor usability.) Tiled Display Wall (TDW) is one of the possible solutions in the near future for the issue of the high-resolution visualization. Figure 11 is a picture of a TDW at NICT main hall. The TDW is composed of 25 panels, and each of them has SXGA resolutions. Here we should note that for the best performance of the TDW we need to make use of 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 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 10Gbps network and distributed storage system discussed in Section 2.3. The I/O time will be a bottle-neck of the data transfer. Note that nominal I/O speed of 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 as high speed as 10Gbps using memory mapping technique.

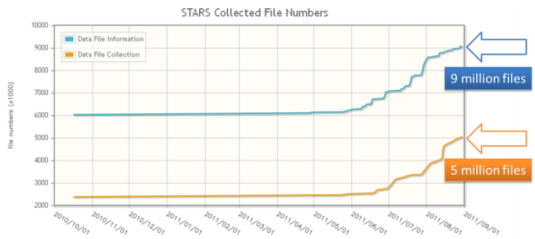

Figure 10: Automatic data crawling system on NICT science cloud: NICYT and DLA (DownLoad

Figure 11: A Tiled Display Wall at NICT main hall.

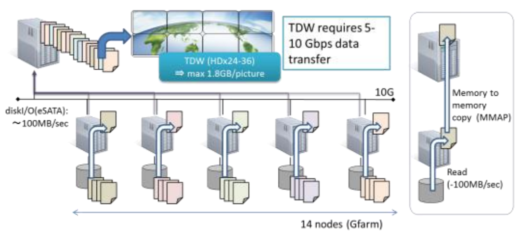

Figure 12: High-speed data transfer to TDW.

4 CONCLUSION

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

Microsoft is one of the companies who is interested in science cloud systems: they seem to apply their own cloud service, Microsoft Azure (Microsoft, 2010), to science researches as presented at the workshop above. However, one success on one science project does not necessary mean the success of the cloud. The science cloud is an environment of a variety of sciences, and an environment on which any researchers make their own customizations. We often describe a cloud system with 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 named as “OneSpaceNet” (for space weather works) with other famous cloud services. So far, Google cloud services are as SaaS and PaaS. Users mainly make use of the Google cloud service through Web. Another cloud service via Amazon is PaaS and IaaS based. They provide storage and computational resources, but no application services.

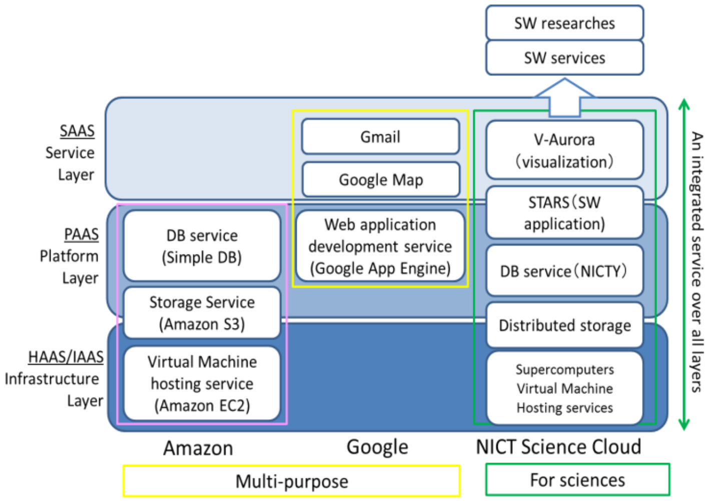

Figure 13. Comparison of the NICT science cloud for space weather research works with other famous cloud systems.

Murata et al. (2005) has ever 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 are simultaneously previewed on 4D (3D in space and time) space. However, the system was not complete 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 we 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 works.

In the present study, we propose a multi-functional science cloud and presented several studies based on the cloud system. The system is still on 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 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

Microsoft (2009), The Fourth Paradigm: Data-Intensive Scientific Discovery, http://research.microsoft.com/en-us/collaboration/fourthparadigm/

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, database, 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*, No.19, pp.135-151.

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.

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.

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

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

Microsoft (2010) Windows Azure Platform, 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)*.

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)*