

Security approaches of a distributed storage system in the NICT Science Cloud

WATANABE, Hidenobu^{1*} ; SUZUKI, Yutaka² ; MURANAGA, Kazuya² ; UKAWA, Kentaro² ; MURATA, Ken T.¹

¹National Institute of Information and Communications Technology, ²Systems Engineering Consultants Co., LTD.

Open data and global data citation of a scientific data have been discussed in international projects of scientific research field in recent years. Nowadays, internet services of them are gradually being provided. In these circumstances that a scientific researcher can manipulate large amounts of scientific data, the Informatics for Earth and Space Sciences fields is expecting the data-intensive science which provides new scientific knowledge or realizes multidisciplinary research by cloud computing technology.

On the other hand, there are not much success cases of the data-intensive science because there are not many free cloud services which provides a storage with large capacity, high available and high security in addition to high performance computers at the disposal of users. Also, many scientific researchers are much concerned about security of cloud services according to the results of the questionnaire survey by a promotion project of the Ministry of Education, Culture, Sports, Science and Technology.

National Institute of Information and Communications Technology (NICT) established about 500 CPU and 3 petabyte-scale cloud systems (NICT Science Cloud) for the data-intensive science. The NICT Science Cloud has data centers in 5 regions (Tokyo, Nagoya, Kyoto, Osaka and Okinawa) of Japan and provides a distributed computing environment with the Gfarm file system on a 10Gbps Layer 2 network (JGN-X). Scientific researchers can use it anytime without charge.

We introduce security approaches of a distributed storage system with the Gfarm in the NICT Science Cloud. In particular, we explain an application and a system to verify integrity, authenticity and traceability of stored data. Also, we discuss about a contribution of the approaches to the data-intensive science.

Keywords: data-intensive science, cloud service, distributed storage system, security, open data