§32. "Cloud" Distributed Storage toward Unified Data Access Platform for Fusion Researches

Nakanishi, H., Ohsuna, M., Kojima, M., Nonomura, M.,
Imazu, S. (Pretech Corp.),
LABCOM Group

The data acquisition and management system of LHD has been operating since 1998. At the beginning of the LHD experiment, it started with less than 10 CAMAC acquisitions. In the 6th campaign, however, it went over 30 even though they were still all CAMACs. From the 7th campaign the long-pulse trials were started in earnest, and thereafter many real-time (RT) acquisition digitizers have been installed. In the 13th period, the total number of plasma diagnostics went up to 85. More than half of the total DAQs use the RT-capable digitizers, such as Yokogawa WE7000 and National Instruments PXI.

Fig. 1 shows some peaks, each of which means a series of long- pulse trials. The highest peak of 90 GB is the world record of acquired data amount in a single shot of fusion experiment. In short-pulse experiments, 10.6 GB raw data are acquired for every shot recently. As for the total amount of the cumulated data, it continues increasing exponentially to be above 0.25 peta-bytes.

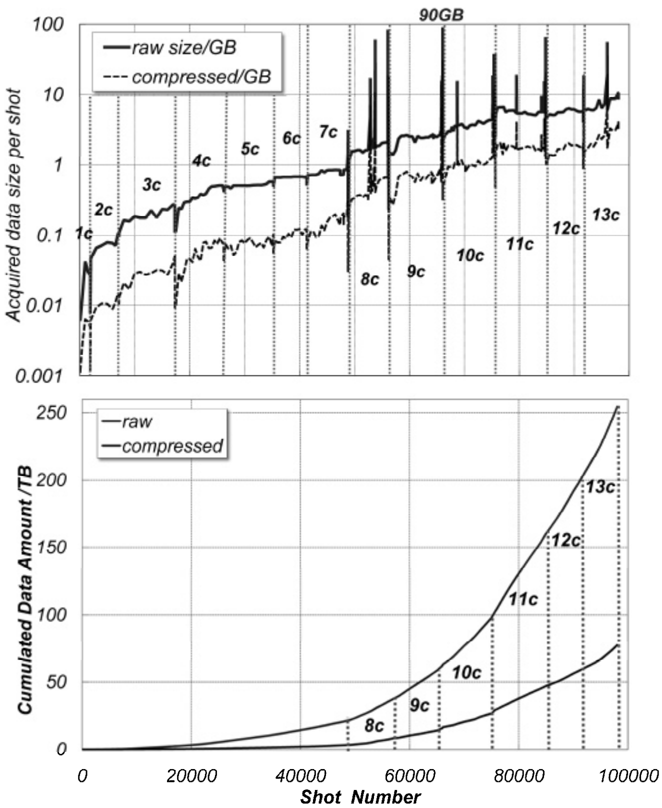

Fig. 1. Growth of acquired raw data per shot (top) and the cumulated amount (bottom) in LHD. It has grown 100 times bigger for the past 10 years. LHD's daily number of shots is almost constant near 170 or 180.

"Cloud" Data Storage in LHD

The LHD plasma diagnostics have already 85 data acquisition (DAQ) nodes in parallel. All of them work concurrently with the experimental sequences; however, disorders of some nodes never affect the whole DAQ processing and the failed nodes can be replaced even in the experimental sequence. One of the biggest advantages in "cloud" computing is considered this hot-swapping ability of each node.

The previous data storage system consisted with a symmetrically mirrored cluster of 8 hard-drive arrays (RAIDs) which connected to each other through 4 Gbps FibreChannel storage area network (FC-SAN)[1]. As it replicated the newly appeared data through the software-based batch-processing method, there existed a single point of failure (SPOF) in primary writable RAID.

In order to advance both of the system reliability and the performance scalability, our storage system has shifted to the cloud-based one, named *IznaStor* (Fig. 2).

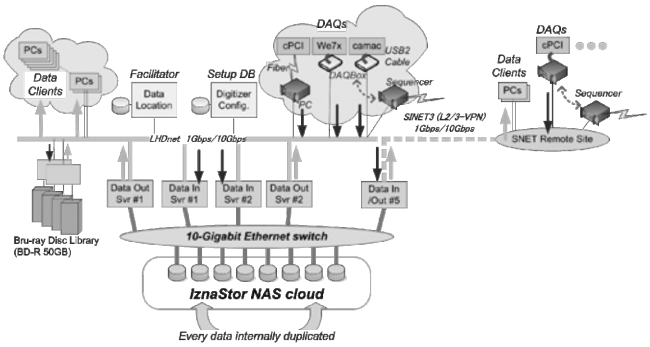

Fig. 2. "Cloud" structure of many DASs and the distributed storage "*IznaStor*", with extensions of SNET remote sites.

IznaStor has a completely symmetric structure that consists of multiple network-attached RAID storage (NAS). The interconnect is also upgraded from 4 Gbps FC to 10 Gbps Ethernet. Other advantages never provided by the prior system are as follows:

1. Automatic file replication will be done from the primary node to another one. Free capacities are balanced among all the participant nodes.
2. The primary disk to write will be selected at random.
3. SPOF-less symmetric architecture enables us to increase or decrease nodes on demand or disorder.
4. NAS RAID is more cost-effective than FC ones.

Toward the Unified Data Access Platform

"Cloud" data storage has a good possibility in parallel and redundant I/O processing that provides better scalability and reliability at once. The structure shown in Fig. 2, however, does not fully enable peer-to-peer data storing because they still use some I/O gateways. It is owing to the compatibility for working application software. When we completed to modify the related software suitable for *IznaStor*, it could provide much better I/O performances even for the data analyzing supercomputer or "cloud" PCs.

1) Nakanishi, H. *et al.*: NIFS Annual Report (2009) 159.