

“The solution provided by Hitachi Data Systems ensures that we will face no barriers to growth. ... Their storage arrays are equipped with functions for virtualization of resources, which add an extra improvement as they facilitate reconfiguration, division into logically independent parts, and data migration.”

*Michał Białoskórski
Administrator in the Supercomputer
Department
CI TASK in Gdansk*



Polish Supercomputer Center CI TASK Stores More Than 1.7PB of Data on Hitachi Storage With the Lustre File System

CI TASK, one of Poland's largest supercomputer centers, offers massive processing capabilities that enable scientists throughout the country to conduct research projects, resulting in massive data growth. To provide a solution that enabled current processes and ensured no barriers to future growth, CI TASK turned to Hitachi Data Systems.

Known as Data Processing Centre of the Tri-City Academic Computer Network, CI TASK gathers and provides data processing resources to researchers. These resources include supercomputers that offer enormous processing capabilities, systems for storage of large data volumes, and dedicated software. Over the 20 years since it was established, CI TASK's computers have modelled the wings of Polish and foreign aircraft, supported research on new materials, and helped conduct virtual tests of medicinal substances, among myriad other uses. They have also carried out various simulations and computing in areas such as chemistry, physics, fluid mechanics, weather prediction models, mechanics, astrophysics, and so forth.

Computers of extremely large processing capacity, or supercomputers, have become a key driver of progress in science. Processing centers like CI TASK enable Polish scientists to conduct research projects that would otherwise have to wait several years for their implementation within a laboratory environment. Costs are also considerably lower when computer models are deployed to analyze complex problems.

“Science means continuous experiments, usually very expensive ones. Our center enables research from throughout Poland to enhance the boundaries of knowledge at a considerably lower cost, as subsequent trials and errors are incomparably cheaper in the virtual environment than in the real world,” says Rafał Tylman, head of the Supercomputer Department of CI TASK in Gdansk.



CI TASK

INDUSTRY

IT Services: Supercomputing for Science and Education

SOLUTION

Enterprise Platform, Modular Platform

HARDWARE

Hitachi Unified Storage VM (3)
Hitachi Unified Storage 150 (1)

SERVICES

Provided by Hitachi TrueNorth Platinum Partner Wasko S.A.

Benefits at a Glance

- Higher performance of the Lustre cluster file system.
- Flexible management of resources, scalability of storage into petabytes.
- Higher availability of the supercomputer to scientists.

The Challenge: Capacity Irrespective of Disk Failures

CI TASK supercomputer's processing offers high capacity. However, the expectations of research usually go beyond what can be delivered. In 2008, when CI TASK was starting the Galera supercomputer, its processing capacity (about 50 TFLOPS) was ranked 46th in the famous TOP500 ranking of the fastest supercomputers in the world.

Galera is a cluster of servers that contain a total of 1,344 four-core Intel Xeon processors, that is, 5,376 processing cores, operating with the frequency of 2.33 GHz. In 2012, the Galera cluster was enlarged with an additional 5,520 cores and, as a result, the Galera Plus cluster emerged with its processing capacity exceeding 100 TFLOPS.

In order to ensure that such a processing capacity is used effectively, 2 things are required: a flexible approach to management of a supercomputer's tasks and high-performance data storage. It seems rather obvious, but in practice this is a huge challenge. The current capacity of the Galera Plus disk storage system is over 1.7PB.

"For us, data storage flexibility means an ability to divide available disk resources in any proportions among various applications running in parallel on the supercomputer. This effect can be achieved only when all the resources are covered by the same namespace within one global file system. The supercomputers we used before required management of multiple file systems, which was both cumbersome and time consuming. Therefore, when we were designing Galera we decided to try the Lustre cluster file system, which allows for scaling of one namespace over multiple disk systems. It soon turned out that it was an excellent choice," recalls Michał Białoskórski, an administrator in the Supercomputer Department of CI TASK in Gdansk.

CI TASK decided to allocate the majority of the funds available for the creation of the Galera supercomputer to the processing layer. The data storage system was considered of secondary importance; however, this approach revealed its weaknesses very soon. A relatively low capacity of storage system controllers coupled with a large number of high-capacity disks resulted in difficulties; it

“ Processing tasks that involve modelling of physical, chemical or engineering phenomena, especially in real time, require high interactivity. In such applications, efficient communications between servers and mass storage devices is a must. In addition, there are many areas of analysis where one 'run' of processing is not enough to accomplish the targeted objective. And with Hitachi storage arrays, the Lustre file system can at last demonstrate its true performance, so valued by supercomputer experts globally. ”

*Michał Białoskórski
Administrator in the Supercomputer
Department
CI TASK in Gdansk*

took a lot of time to recover a RAID group after a disk failure, leading to a reduced capacity of the whole environment. When planning the purchase of a new data storage environment for supercomputers a few years later, CI TASK decided to invest in a solution that would guarantee the efficient operation of the data storage system even if a few disks were lost at the same time.

"The mass storage environment for the Galera Plus supercomputer operates as a temporary data warehouse of 'raw' data received as a result of processing. Consequently, the failure of a disk is mainly a performance-related problem for us. When a storage system starts automatically rebuilding the logical volume, the performance of the entire disk storage system, and as a result the whole supercomputer, goes down. Unlike in database systems, losing some data as a result of a disk failure is practically negligible. Any missing data can be forecast or, in the worst-case scenario, part of the processing can be simply rerun. We store final data safely in a separate environment," says Tylman.

The Solution: Lustre File System on Hitachi Unified Storage

During the tender it was Wasko that presented a solution that was able to both meet the performance-related expectations of CI TASK and ensure the highest usable storage capacity. The recommended architecture was

based on Hitachi storage and servers with Intel Xeon processors, serving as the platform for the Lustre cluster file system. At the test stage, the solution offered by Wasko achieved a stable data recording performance of about 15GB per second, also while rebuilding a damaged disk.

"Our solution consisted of 3 enterprise storage systems, Hitachi Unified Storage VM (HUS VM) for data storage, and 1 HUS 150 midrange system for storing index data and responding to queries. The platform created for the Lustre cluster included four 2-processor servers for reading and writing data and 2 similar servers with solid-state disks and a larger RAM for management of Lustre metadata. All the servers run Linux," says Krzysztof Małek, strategic client manager at Wasko S.A.

The architecture of the environment consists of: (1) one Galera Plus cluster that includes 864 servers and is connected internally using an InfiniBand network of the fat tree typology; and (2) six servers with the Lustre file system that are connected to both InfiniBand network and storage area network (SAN) via Fibre Channel 8 GB/sec cards. The SAN consists of two 24-port Fibre Channel switches, to which disk storage systems are cross-connected. Making the Lustre file system work with a Fibre Channel SAN was a big challenge in the project, as Lustre usually does not use Fibre Channel-based SAN networks, but rather Ethernet/iSCSI-based networks.

**Hitachi
Unified
Storage VM**

[LEARN MORE](#)

"A Fibre Channel-based SAN network assumes fixed distribution of resources among its servers, while Lustre requires that all servers should have access to all resources. The performance of Lustre results, among other things, from the fact that data is distributed among as many disks as possible. It was not an easy task, but thanks to our experience gained on other HPC projects for large clients we were able to skillfully resolve that problem," says Krzysztof Małek.

A second parameter that was significant to CI TASK, apart from the processing capacity, was the usable capacity of the disk storage system. Wasiko provided storage systems with licenses for the storage management software without any restrictions to the number and capacity of disks, which was of great importance to CI TASK. A total of 564 SAS-2 disks, 3TB each, were installed on the storage systems. After defining the RAID groups, the usable capacity of the storage of the Galera supercomputer is slightly above 1.7PB.

The Benefits: Efficiency and New Opportunities

At present, CI TASK has a very efficient supercomputer environment. In addition to the huge processing capability, resulting from the possibility to divide calculations between over 10,000 processing cores, the Galera Plus supercomputer has gained vast capacity in the area of data recording and retrieving with its mass storage. "We have managed to build an efficient processing environment. With over 10,000 cores and a nearly 2PB data repository in a single and consistent environment, we can help researchers resolve even the most complex scientific problems," says Mściślaw Nakonieczny, director of CI TASK in Gdansk.

"Processing tasks that involve modelling of physical, chemical or engineering phenomena, especially in real time, require high interactivity. In such applications, efficient communications between servers and mass storage devices is a must. In addition, there are many areas of analysis where one 'run'

Innovation is the engine of change, and information is its fuel. Innovate intelligently to lead your market, grow your company, and change the world. Manage your information with Hitachi Data Systems.

INNOVATE
WITH INFORMATION™
www.HDS.com/innovate

of processing is not enough to accomplish the targeted objective. And with Hitachi storage arrays, the Lustre file system can at last demonstrate its true performance, so valued by supercomputer experts globally," emphasizes Białoskórski.

By implementing the Hitachi platform, CI TASK has increased the availability of the Galera Plus to scientists, as the high capacity of the disk storage system enables efficient execution of multiple processing tasks in parallel. The storage systems can be configured so that a failure of one or even several disks does not significantly impact the performance of the environment as a whole. "For researchers, a supercomputer is a virtual lab. When one trial fails to give unambiguous results they will try again. It sometimes happens that a team wants us to repeat processing with modified function parameters or input," explains Tylman.

With the solution implemented by CI TASK, the environment is freely scalable in multiple dimensions. The license for the storage system software has no restrictions regarding the number of disks or controllers, or the volume of data processed in the

environment. And, Hitachi storage systems can manage other data storage systems via virtualization, which Hitachi has been developing, testing and implementing in large systems for several years.

"We assume that with the expectations of clients steadily rising our processing and storage resources will be growing accordingly, as has been the case constantly throughout the 20 years of CI TASK's history," says Białoskórski. "The solution provided by Hitachi Data Systems ensures that architecture or technology will not be barriers to our growth. Hitachi storage systems provide resource virtualization, facilitating storage reconfiguration and partitioning of the systems into logically independent parts as well as easy data migration," he explains.

"The ultimate goal for the Lustre file system implemented on the Hitachi storage arrays is to support a new processing cluster with performance over 10 times higher (1.2 PFLOPS) than the supercomputers CI TASK uses today. The procedure to buy a new supercomputer is pending now," adds Nakonieczny.

CI TASK INNOVATES WITH INFORMATION

- Higher performance of the disk storage system for supercomputers enables individuals who use the processing resources of CI TASK to start simulation and modelling applications in real time. As a result, CI TASK is potentially a more attractive partner for the industry.
- Improved capacity of the disk storage system expands the efficiency of multiphase processing, where results of one stage are used as input at a subsequent stage. This enhancement supports new opportunities to study extremely complex natural phenomena.

© Hitachi Data Systems



Corporate Headquarters

2845 Lafayette Street
Santa Clara, CA 95050-2639 USA
www.HDS.com community.HDS.com

Regional Contact Information

Americas: +1 408 970 1000 or info@hds.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hds.com
Asia Pacific: +852 3189 7900 or hds.marketing.apac@hds.com

© Hitachi Data Systems Corporation 2014. All rights reserved. HITACHI is a trademark or registered trademark of Hitachi, Ltd. All other trademarks, service marks, and company names are properties of their respective owners.

SS-525-A DG August 2014