HITACHI
Inspire the Next

> “Impressive Hitachi NAS Platform capabilities ensure enormous data transmission capacity and enormous data processing power for CAS’ image data.”

Liu Shi Bin
Manager of Satellite Data Image Project
Chinese Academy of Sciences

Chinese Academy of Sciences (CAS)

INDUSTRY	Education: Research
SOLUTIONS	File and Content Services, Modular Platform **Hardware** — Hitachi NAS Platform 3100, powered by BlueArc®; Hitachi Adaptable Modular Storage 2300 **Software** — Hitachi NAS Platform software and Hitachi Data Migrator software, powered by BlueArc®

Robust Hitachi NAS Platform Answers High Performance Storage Needs of Chinese Academy of Sciences

To meet the increasing demands of its research activities, the Chinese Academy of Sciences (CAS) needed to establish a large scale storage system for recording images from its Remote Sensing Satellite. Hitachi Data Systems provided the robust, scalable and reliable architecture CAS required with the Hitachi NAS Platform, powered by BlueArc®.

About CAS and the Remote Sensing Satellite Ground Station

The CAS China Remote-Sensing Satellite Ground Station (RSGS) was founded through a Sino-American Scientific-Technical Co-operation Memorandum signed by Mr. Deng Xiaoping. This high tech scientific operation and research unit is under the direct management of the Chinese Academy of Sciences.

The RSGS was inaugurated and put into operation in December 1986. Its principal tasks are to receive, process, archive and distribute earth resources data from various remote-sensing satellites, and also to use that data in scientific research. To keep pace with evolving user demands, the RSGS has introduced a number of new initiatives over the years, such as strengthening its data service, product distribution, technical consultation and quality assurance, as well as the continuous development of new services.

Two Major Challenges

CAS identified two important objectives the new system would have to achieve.

Firstly, it had to be powerful enough to provide more than 40TB storage capacity with massive scalability to expand to petabyte (PB) capacity. Secondly, the storage platform had to guarantee the efficiency and reliability of data sharing even when dealing with huge volumes of satellite data and images.

The CAS RSGS provides remote-sensing images to more than 600 local and foreign organizations. This satellite data is widely used in numerous applications, including China's land resources survey, national forest resources survey, ecological studies, urban development studies, desertification monitoring, crop-yield assessments, disaster monitoring and assessments, geological exploration and topographic mapping. Indeed, the CAS RSGS plays a vital role in enhancing China's high tech industrialization and provides valuable scientific support for ongoing economic development.

Tackling the Rapid Growth of Unstructured Data

Due to a continuous increase in the volume of satellite images, the CAS RSGS needed to build a massive unstructured data storage system with petabyte-level capacity. This new system would be relied upon not only to provide multitier data management, but also to enable comprehensive data migration solutions. It would likewise have to be advanced, extremely reliable and highly manageable in order to allow the efficient sharing of unstructured data.

"Hitachi NAS Platform and Adaptable Modular Storage 2300 provide a total storage solution to meet our requirements. Hitachi NAS Platform provides robust architecture, unprecedented reliability and vigorous scalability, which enables us to not only effectively manage large scale image files, but also achieve tiered storage and data migration in a convenient and hassle-free way. With Hitachi, we now enjoy superb performance and efficiency for large scale file sharing and processing."

Liu Shi Bin
Manager of Satellite Data Image Project
Chinese Academy of Sciences

CAS had a number of other decisive requirements for the design of its new storage system. Critical issues included the creation of a high performance, large scale, document storage platform; vastly scalable storage capacity; easy upgradability to petabyte-level capacity to manage the rapid growth of unstructured data; and seamless interoperability between the storage system and all I/O servers to ensure uninterrupted data transmission.

A Powerful Storage Platform for Large Scale Image Data Processing

As a mature technology, Hitachi NAS Platform is a proven solution for processing massive-scale data transmissions. With hybrid storage capabilities that combine a hardware storage utility and Network Data Management Protocol (NDMP) backup using FPGA (field programmable gate array) chips, Hitachi NAS Platform provides unsurpassed performance and reliability for concurrent file serving activities. As CAS frequently and concurrently receives a high volume of image data, it was vital that the new system be highly flexible and scalable to support unexpected online expansion.

Hitachi NAS Platform addressed this need by providing high performance, unmatched reliability and superb availability. Other key advantages include:

Leading-edge Architecture

Following years of continuous improvement, Hitachi NAS Platform has developed into a mature technology that can comprehensively solve common NAS bottleneck issues. The wide-ranging functionality of Hitachi NAS Platform now uses many leading technologies to deliver a best-of-breed operational platform:

- FPGA technology enhances system performance and reliability.
- A hardware-based file sharing system provides high speed data transmission of 128TB to 256TB for a single file.
- Seamless integration with SAN enables effortless and dynamic expansion.
- Support for multiple network protocols includes CIFS, NFS, iSCSI and NDMP.
- Support for single and multiple-node clustering operations (up to eight nodes) is provided, and all clusters support active-active failover function.

Traditionally, a NAS platform needed a parallel file system to facilitate temporary file sharing. Since the launch of Hitachi NAS Platform, organizations can now enjoy a more advanced storage solution that quickly and efficiently processes even large scale file sharing operations.

Advanced Functionalities

Hitachi has enriched the traditional NAS with a number of functions designed to meet the specific needs of document management, sharing and storing, realizing storage virtualization, managing large scale file systems, protecting data, delivering virtualized services and more.

Hitachi NAS Platform virtualization services now include virtual server, virtual volumes and virtual storage. With virtualization, storage administrators can effectively allocate different types of storage for different applications. For instance, virtual volumes minimize the complexity of physical disks and flexibly meet the capacity demands of users. The virtual server also enables resource partitioning management and provides an auto-switch function in the event of overloading or system failure. Other benefits include:

- **Multitier storage management.** Data can easily be migrated from one tier of storage to another according to predefined document specifications. For user convenience, the document path is also automatically directed to the new storage tier, providing a high degree of transparency.
- **Snapshot.** Snapshot is a major data protection strategy. Hitachi NAS Platform can support up to 1,024 snapshots for each document.
- **Modular design.** Users can flexibly change function modules based on their specific needs, thus maximizing the overall IT investment.
- **Heightened capacity.** Hitachi NAS Platform supports 64 virtual servers and up to 256TB of memory for each file system. This enables each file system to support up to four million documents and directories, with a maximum storage capacity of 2PB.
- **Integration with storage management software.** Hitachi NAS Platform seamlessly integrates with the Hitachi storage management package to enable highly efficient and highly scalable storage solutions.

Data Processing Performance

In an authoritative third-party assessment, Hitachi NAS Platform was found to offer the market's highest performance. Indeed, its unique architecture provides unparalleled data processing power:

- Multiple FPGA-based hardware architecture ensures up to 800Mb/sec data throughput for a single node. A pair of clustering nodes can scale up to 193,000 IOPS SpecSFS processing power.
- Hitachi NAS Platform can simultaneously expand from a single node to eight nodes. This delivers linear improvements in performance that enable continuous expansion with exceptional reliability.

Until a few years ago, some users of parallel file systems faced performance and data processing bottlenecks. For example, when the data processing of a parallel file system exceeded 30TB, the system would begin to sprawl and even fail. Likewise, when the hit rates reached 100 or above, the performance would suddenly diminish. In contrast, Hitachi NAS Platform provides unrivaled power, enabling it to drive a single file system of up to 128TB capacity, or expand a single node to between 2PB

and 4PB. These impressive capabilities ensure enormous data transmission capacity and enormous data processing power for CAS' image data.

After intensive evaluation, CAS chose the high performance Hitachi NAS Platform 3100 and the Hitachi Adaptable Modular Storage 2300, a 1TB SATA disk storage system, and consolidated them into a virtual storage pool. Centrally managed by NAS Platform 3100, the file sharing space for the application servers' queries is allocated dynamically with NFS or CIFS protocols. CAS also deployed automatic data migration to support multiple file systems with Hitachi NAS Platform software and Hitachi Data Migrator software, powered by CommVault®.

This solution ensures that all I/O servers can process data simultaneously and reliably.

Solution Topology

The solution uses the high performance Hitachi NAS Platform 3100 as a core system for file sharing storage. Overall, the system comprises two key technologies:

- Hitachi NAS Platform 3100 master control node unit (NAS master control equipment)
- Hitachi Adaptable Modular Storage 2300 storage unit with 50TB of disks installed for NAS file sharing storage, eight disks forming a RAID-5 array group and over 40TB of usable capacity in total

By consolidating the NAS Platform 3100 and Adaptable Modular Storage 2300 storage system, CAS created a virtual storage pool for file sharing management and storage. NAS Platform 3100 works as a file sharing management center, using NFS and CIFS protocols to support UNIX- or Microsoft® Windows®-based servers for image-document sharing. The Adaptable Modular Storage 2300 is also configured with multiple Fibre Channel ports, enabling business applications to share data through the Fibre Channel interface.

A Super-capacity Unstructured Data Storage Platform

In view of CAS' unique requirements for unstructured data and large scale image-file sharing and management, Hitachi provided a high performance NAS solution that provides massive capacity to meet every need and application. Data migration and multitier storage management functions now enable users to flexibly migrate data from one tier of storage to another according to predefined document specifications. By automatically directing the path to the new storage tier, the system ensures absolute migration transparency. CAS can also analyze the performance and cost factors of each storage platform to work out a best-fit storage strategy.

Hitachi Data Systems Corporation

Corporate Headquarters
750 Central Expressway
Santa Clara, California 95050-2627 USA
www.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