

Postprint of the NSRT Observation Data Storage System at Xinjiang Astronomical Observatory

Authors: Zhang Hailong, Zhu Yan, Nie Jun, Yuan Jianping, Wu Gang, Liu Jun, Wang Jie, Wang Wanqiong, Ye Xinchun, Tohtinur, Zhang Meng

Date: 2018-06-21T13:05:54+00:00

Abstract

The Nanshan 26m Radio Telescope of Xinjiang Astronomical Observatory (abbreviated as NSRT) has amassed extensive scientific data through years of observations. To address the online storage and backup challenges of NSRT astronomical observation data, a remote, off-site disaster recovery backup system was established. Independent storage systems were deployed at both the Xinjiang Astronomical Observatory headquarters and the Nanshan Observation Station, enabling real-time remote, off-site disaster recovery data backup between the two systems. The storage system was implemented based on the object-based storage technology Lustre, and comprehensive testing was conducted on its read and write performance. The constructed disaster recovery backup system has effectively resolved the online storage and data security issues for both observation data and derived data of the Xinjiang Astronomical Observatory.

Full Text

ChinaXiv Cooperative Journal: Xinjiang Astronomical Observatory NSRT Observation Data Storage System

Zhang Hailong^{1,2}, Zhu Yan^{1,3}, Nie Jun^{1,2}, Yuan Jianping¹, Wu Gang¹, Liu Jun¹, Wang Jie¹, Wang Wanqiong¹, Ye Xinchun¹, Tohtinur¹, Zhang Meng^{1,3}

¹Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi 830011, China

²Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Nanjing 210008, China

³University of Chinese Academy of Sciences, Beijing 100049, China

Abstract

After years of observation, the Nanshan 26-meter Radio Telescope (NSRT) of Xinjiang Astronomical Observatory has accumulated a vast amount of scientific data. To address the online storage and backup challenges for NSRT astronomical observation data, a remote, off-site disaster recovery backup system was established. Independently operational storage systems were deployed at both the Xinjiang Astronomical Observatory headquarters and the Nanshan observation station, enabling real-time remote and off-site disaster recovery backup between the two storage clusters. The storage system was implemented based on the object-based storage technology Lustre, and comprehensive read/write performance tests were conducted. The disaster recovery backup system effectively resolves the online storage and data security issues for both observational and derived data at Xinjiang Astronomical Observatory.

Keywords: Observational Data; Storage; Backup; Virtual Observatory; Data Safety

1. NSRT Data Overview

The Murchison Widefield Array (MWA), a Square Kilometre Array precursor telescope consisting of 2,048 low-frequency array elements, archives approximately 400 MB of correlated data per second. This data is first stored online at the MRO observatory's data storage facility, then transmitted in real-time via a 10 Gbps dedicated line to the Pawsey Data Center located 700 km away from MRO. Additionally, Pawsey Data Center backs up data as needed via 1 Gbps lines to MIT (USA), VUW (New Zealand), and RRI (India). The National Astronomical Observatories Data Center is currently China's largest astronomical database, encompassing data from various observational facilities under the National Astronomical Observatories as well as from other stations, with partial backups at the Chinese Academy of Sciences Network Center and Alibaba Cloud. The Purple Mountain Observatory's publicly accessible databases include millimeter-wave radio astronomy, solar radio spectrum, near-Earth object telescope, and solar spectrum databases, all of which provide online access and have established corresponding backup systems. Other astronomical research institutions such as the Strasbourg Astronomical Data Center, European Southern Observatory Science Data Center, CSIRO ATNF Data Archive, Shanghai Astronomical Observatory, and Yunnan Astronomical Observatory have each constructed their own data management systems.

The Xinjiang 25-meter radio telescope, completed in December 1993, was upgraded to a 26-meter aperture and renamed the Nanshan Radio Telescope (NSRT). NSRT undertakes important astronomical observation tasks for international cooperation and major domestic research programs, and is a formal member of four international cooperation organizations: the European VLBI Network (EVN), International VLBI Service for Geodesy and Astrometry (IVS), Russian Low Frequency VLBI Network (LFVN), and East Asian

VLBI Network. It participates in 11 international cooperation programs and supports national key projects including the National Climbing Program, major science engineering projects, lunar exploration, Mars exploration, National Natural Science Foundation programs, and key basic research projects of the Chinese Academy of Sciences, as well as numerous single-dish international cooperative astronomical observation tasks.

NSRT has conducted multiple research programs including pulsar observations, molecular line surveys, IDV monitoring, and supports projects such as the Galactic plane magnetic field survey and Jupiter studies. The telescope's effective observation time has repeatedly set new domestic records for similar radio telescopes, playing an active role in astronomical research both domestically and internationally. With the explosive growth of observational data, permanently preserving these valuable astronomical data, effectively managing online storage, and efficiently implementing remote off-site disaster recovery backup have become critical challenges for the operation of the Xinjiang Astronomical Observatory's 26-meter radio telescope.

From January 2000 to June 2002, the NSRT pulsar observation system utilized a dual-channel room-temperature receiver with 320 MHz bandwidth centered at 1540 MHz, employing a $2\text{S} \times 128 \times 2.5$ MHz analog filter bank for dedispersion, producing data in "Timer" format. In late 2002, a cryogenic receiver system was commissioned, improving antenna sensitivity to 0.5 mJy. In January 2010, the Digital Filter Bank (DFB) system was deployed, offering higher time resolution and enabling routine observations of approximately 280 pulsars, including 10 millisecond pulsars. The DFB system records data in "Psrfit" format, readable and analyzable by the "psrchive" software package. Over a decade of observations has yielded more than 56,000 valid raw data records, with total raw and processed data volume approaching 20 TB.

NSRT has performed molecular line observations of OH, H₂CO, NH₃, H₂O, and other species. Since 2010, the digital dedispersion system has produced raw data in RPFits format. Calibrated molecular line data enables estimation of interstellar medium and molecular cloud physical and chemical properties, with approximately 5 TB of archived data currently available.

Since 2004, NSRT's 6 cm continuum observation system has monitored extragalactic radio source flux densities, including large-sample rapid variability surveys of northern blazars and long-term radio flux monitoring of Fermi AGN. The continuum system terminal, developed by the Max Planck Institute for Radio Astronomy, operates at 4800 MHz center frequency with 600 MHz bandwidth. Raw data in FITS format requires calibration for pointing, atmospheric opacity, gain, and time-dependent effects before conversion to absolute flux densities for scientific analysis. Years of observations have obtained approximately 250,000 valid raw data records for ~800 radio sources, totaling about 10 TB.

2. Data Storage and Backup System Design

The NSRT data storage and backup system design is illustrated in [Figure 1: see original paper], comprising three components: observation data acquisition and online archiving, off-site backup of raw observational data, and data release and sharing.

The data acquisition and online archiving component is implemented at the Nanshan observation station, consisting of the 26-meter telescope, receiver system, DFB system, temporary storage system, and archiving system. Using pulsar observations as an example, data collection, processing, period calculation and folding, disk recording, dedispersion display, schedule querying, graphical output, and antenna control are performed in real-time by the telescope's acquisition program. Typical sampling rates are 1 ms, with observation durations of 2-16 minutes depending on pulsar flux density at the observed frequency. Preprocessed data are written to temporary storage servers and permanently archived after scientific validation.

Raw observation data off-site backup (the Xinjiang Astronomical Observatory headquarters is located approximately 100 km north of the Nanshan observation station) is achieved via a dedicated 300 Mbps line between the stations. Data are synchronized daily from Nanshan's Permanent Online Archive to the headquarters' Long-term Online Archive starting at midnight, currently using the Next-Generation Archive System (NGAS) for raw data transmission. Users can log into the Xinjiang Observatory's Taurus high-performance computing system to download and process data. Taurus connects to the Long-term Online Archive via 56 Gbps InfiniBand switches. Processed data can be archived and released as needed. GAVO (German Astrophysical Virtual Observatory) servers primarily handle data release and processed data storage. Extracted metadata from raw observations are imported into corresponding databases to prepare for data release. Data storage, Taurus, and GAVO servers interoperate via NFS. Database data are backed up at both the Xinjiang Astronomical Observatory headquarters and Nanshan observation station. A 100 Mbps MSTP (Multi-Service Transfer Platform) dedicated line has been established between the Qitai observation station and headquarters. To meet initial construction and multi-instrument data acquisition needs, a 20 TB storage system has been deployed at Qitai, which can also serve database backup requirements.

The third component consists of data backup systems and release platforms located at Nanshan and Qitai stations. Both backup systems implement remote off-site disaster recovery for critical headquarters data via dedicated networks. The data release system, based on the Xinjiang Astronomical Observatory data center, enables Virtual Observatory-standard data release, efficient data retrieval, and access.

3. Storage System Implementation

3.1 Storage Technology Introduction

Storage systems can be classified by server type into closed-system storage (primarily for mainframes) and open-system storage (based on Windows, UNIX, or Linux servers). Open-system storage divides into internal storage and external storage. External storage connects via Direct-Attached Storage (DAS) or Fabric-Attached Storage (FAS). FAS further subdivides into Network-Attached Storage (NAS) and Storage Area Network (SAN), as shown in [Figure 2: see original paper].

DAS, the current predominant architecture, directly attaches storage to servers, relying on the host operating system for I/O and storage management. Data backup and recovery consume server CPU and system I/O resources, with data flowing back through the host to storage, typically occupying 20–30% of server resources. DAS performance depends entirely on the connected server.

NAS, or Network-Attached Storage, adds storage devices to single computers or high-performance computing systems via standard network topologies. NAS provides file-level storage, enabling rapid capacity expansion with plug-and-play support and multi-platform compatibility for Unix/Windows LAN environments. However, NAS consumes significant bandwidth during backup operations, as the network must handle both normal user traffic and storage disk I/O requests.

SAN, or Storage Area Network, constructs dedicated storage networks by connecting storage arrays to servers via Fibre Channel switches, supporting both SCSI and IP protocols on the same physical channel and allowing any server to connect to any storage array. Fibre Channel SAN (FCSAN) provides higher bandwidth through fiber interfaces supporting distances over 10 km, facilitating physically separated backup storage in different machine rooms.

Object-Based Storage (OBS) separates data path (read/write) from control path (metadata). Object Storage Targets (OST) enable each device to autonomously manage data distribution with built-in intelligence. OBS architecture comprises objects, object storage devices, metadata servers, and clients. Network bandwidth, I/O throughput, file system capacity, and processing power scale linearly with storage nodes, providing excellent performance and scalability at all levels—nodes, object count, and object space. OBS enables highly parallel access to massive datasets by separating file data from metadata, striping traditional file data across storage objects while preserving file metadata in metadata objects with globally unique identifiers and file attributes. While SAN and NAS represent the two mainstream network storage architectures, OBS is an emerging architecture that combines SAN's high-speed direct access with NAS's distributed data sharing advantages, delivering a high-performance, reliable, cross-platform, and secure data sharing storage architecture.

3.2 Storage System Implementation

Considering the advantages and disadvantages of DAS, NAS, SAN, and OBS technologies, along with Xinjiang Astronomical Observatory's current storage requirements, we implemented the storage system using object-based storage technology. The system is built upon Lustre, a high-performance distributed file system for Linux with open-source code that uses object-based disk storage, with metadata servers providing file system metadata services. As shown in [Figure 2: see original paper], the system employs dual network interconnects: 56 Gb Infiniband for high-speed data exchange between storage servers, and Gigabit Ethernet for management. Two metadata servers (MDS) operate in active-passive mode with real-time data synchronization; when the primary MDS fails, the standby MDS takes over, reducing system failure rates and ensuring reliable metadata access. Three Object Storage Targets (OST) provide 100 TB of usable storage space.

The centralized Lustre storage system connects to two I/O servers also configured in active-passive mode, with one online and one standby. Linux users install Lustre client software to mount and access the 100 TB storage space. As illustrated in [Figure 3: see original paper], the Linux Virtual File System (VFS) implements any file operation in Linux through a unified I/O system, independent of specific file system formats. VFS defines basic conceptual interfaces and data structures supported by all file systems. The Lustre file system provides the abstract interfaces and data structures expected by VFS, maintaining consistency between its file and directory concepts and VFS definitions to enable data transfer. The Logical Object Volume (LOV) aggregates OST information into a single volume, while user read/write operations are implemented through the Object Storage Client (OSC). Upon receiving user requests, OSC queries the metadata server via the Metadata Client (MDC) to locate data positions in OSTs and returns address information, enabling OSC to perform parallel data read/write operations.

3.3 Storage Performance Testing

We conducted comprehensive read/write performance tests using the professional storage benchmarking tool IOZONE on both single-node and multi-node configurations.

Single-Node Performance: Test command: `./iozone -a -g 64G -i 0 -i 1 -i 2 -f /home/iozone -Rb single.xls`. Parameters: automatic mode generating comprehensive test reports, block sizes from 4 KB to 16 MB, maximum test file 64 GB, test node file `/home/iozone`, results output to `single.xls`. Optimal performance was achieved with 8 MB block size and 8-16 GB file sizes, yielding 420 MB/s single-node write speed and 2.2 GB/s single-node read speed.

Multi-Node Performance: Test command: `./iozone -a -g 64G -i 0 -i 1 -i 2 -f /home/nodes -Rb multi.xls`. Parameters: automatic mode, block

sizes 4 KB to 16 MB, maximum file 64 GB, eight nodes specified in `/home/nodes` file, results output to `multi.xls`. Optimal performance occurred with 8192 KB block size and ~65 MB file size, achieving approximately 960 MB/s multi-node write speed and 5.1 GB/s multi-node read speed, as detailed in [Figure 4: see original paper] and [Figure 5: see original paper].

Conclusion

Based on the actual requirements of Xinjiang Astronomical Observatory, we designed and implemented an online storage and backup system for NSRT observation data. Two storage systems were constructed at the Xinjiang Astronomical Observatory headquarters and Nanshan observation station, enabling remote off-site disaster recovery backup. Read/write performance tests demonstrate that both single-node and multi-node speeds currently meet NSRT data management requirements. The object-based storage technology provides excellent performance and scalability. Testing was completed at the Xinjiang Astronomical Observatory data center and Taurus high-performance computing system.

References

- [1] Tingay S J, Goeke R, Bowman J D, et al. The Murchison Widefield Array: the Square Kilometre Array Precursor at low radio frequencies[J]. *Publications of the Astronomical Society of Australia*, 2013, 30(30): 109-121.
- [2] Zhang Hailong, Wang Jie, Wang Wanqiong, et al. Construction and application of the data center in Xinjiang Astronomical Observatory[J]. *Astronomy Research and Technology*, 2017, 14(1): 94-102.
- [3] Wang N, Manchester R N, Zhang J, et al. Pulsar timing at Urumqi Astronomical Observatory: observing system and results[J]. *Monthly Notices of the Royal Astronomical Society*, 2001, 328(3): 855-866.
- [4] Yuan J P, Manchester R N, Wang N, et al. Pulse profiles and timing of PSR J1757–2421[J]. *Monthly Notices of the Royal Astronomical Society*, 2017, 466(1): 1234-1241.
- [5] Hailong Zhang, Markus Demleitner, Na Wang, et al. Data retrieval from Xinjiang Astronomical Observatory's Pulsar Data Archive[J]. *Astronomy Research and Technology*, 2016, 13(4): 473-480.
- [6] Shirley Y L. The critical density and the effective excitation density of commonly observed molecular dense gas tracers[J]. *Publications of the Astronomical Society of the Pacific*, 2015, 127(949): 299-310.
- [7] Liu B R, Liu X, Marchili N, et al. Two-year monitoring of intra-day variability of quasar 1156+295 at 4.8 GHz[J]. *Astronomy & Astrophysics*, 2013, 555(4): 334-345.

[8] Liu X, Mi L G, Liu J, et al. Intra-day variability observations and the VLBI structure analysis of quasar S4 0917+624[J]. *Astronomy & Astrophysics*, 2015, 578: A34-A42.

[9] Zhang Hailong, Ye Xincheng, Li Huijuan, et al. Astronomical data query and release review[J]. *Astronomy Research and Technology*, 2017, 14(2): 212-228.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.