# Postprint of the Taurus High-Performance Computing System at Xinjiang Astronomical Observatory

**Authors:** Hailong Zhang, Ye Xinchen, Nie Jun, Longfei Chen, Tohti Nur, Wang Jie, Cui Chenzhou, Li Changhua, Zhu Yan, Zhang Meng

**Date:** 2018-05-15T00:00:00+00:00

## Abstract

The Taurus high-performance computing system of Xinjiang Astronomical Observatory consists of one login node, 16 compute nodes, 2 I/O nodes, and 100 TB of high-speed storage. Its CPU theoretical double-precision floating-point computing capability is 6.7584 Tflops, with an actual peak performance of 6.289 Tflops measured via the Linpack test, representing 93.06% of the theoretical value; the GPU theoretical computing capability is 18.72 Tflops, with an actual measured peak performance of 14.882 Tflops, corresponding to a computational efficiency of 79.5% of theoretical capability. Both compute and storage nodes of the system utilize 56 Gb Infiniband switched interconnect; the storage system, tested via IOZone, achieves a single-node write speed of 460 MB/s and a multi-node write speed of 800 MB/s. The system has been applied in domains including polyphase filtering and GPU algorithm acceleration for interference cancellation, Monte Carlo simulations, among others.

## Full Text

## Taurus High Performance Computing System of Xinjiang Astronomical Observatory

**Zhang Hailong[1,2], Ye Xinchen[1], Nie Jun[1,2], Chen Longfei[1], Tohtonur[1], Wang Jie[1], Cui Chenzhou[3], Li Changhua[3], Zhu Yan[1] , Zhang Meng[1]**

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

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

[3] National Astronomical Observatories, Chinese Academy of Sciences, Beijing

100012, China
University of Chinese Academy of Sciences, Beijing 100049, China

## Abstract

The Taurus high-performance computing (HPC) system at Xinjiang Astronomical Observatory comprises one login node, sixteen compute nodes, two I/O nodes, and 100TB of high-speed storage. The theoretical double-precision floating-point computational capacity of the CPUs is 6.7584 Tflops. Linpack testing reveals an actual peak performance of 6.289 Tflops, yielding 93.06% of the theoretical capacity. The GPU theoretical computational capacity is 18.72 Tflops, with an actual tested peak of 14.882 Tflops, corresponding to 79.5% efficiency. Both compute and storage nodes are interconnected via 56Gb Infiniband switches. IOZone testing of the storage system demonstrates single-node write speeds of 460MB/s and multi-node write speeds reaching 800MB/s. The system has been successfully applied to GPU-accelerated algorithms for polyphase filtering and interference mitigation, as well as Monte Carlo simulations.

**Keywords:** HPC; Lustre; IOZone

## 1. Taurus High Performance Computing System

Astronomy is a data-intensive science grounded in observation and simulation. Frontier research topics such as dark energy, dark matter, black holes, cosmic origins, celestial formation, and the origin of life require, first and foremost, observational data, followed by high-performance computing systems to model and analyze massive volumes of real or simulated data, thereby unveiling the mysteries of the universe. Rapid advancements in information technology and high-performance computing have made it possible to realize theoretical and experimental studies through computational means, establishing HPC as an essential tool for astronomers [**?**].

The Xinjiang Astronomical Observatory of the Chinese Academy of Sciences has dedicated decades to astronomical observation and theoretical research, focusing on pulsars, star formation, active galactic nuclei, and radio and optical telescope technologies. To enhance the efficiency of data processing and simulation analysis, we have constructed a high-performance computing system tailored to our scientific research needs, named Taurus[1].

In recent years, GPU-based parallel computing has emerged as a research hotspot in high-performance computing, substantially accelerating scientific analysis and simulation applications [**?**]. NVIDIA introduced GPU-accelerated computing in 2007 [**?**], enabling computational-intensive tasks to be offloaded to GPUs while CPUs handle remaining workloads, thereby significantly improving data processing speeds. In astronomy, GPU computing frameworks are particularly well-suited for astronomical image processing, large-scale cosmological simulations, and space object orbit modeling, and have found widespread application in astronomical research.

The computational performance of HPC systems derives from parallel computing across multiple nodes, making data transmission and communication between nodes critical to system design [?]. The establishment of the Taurus HPC system represents a breakthrough for Xinjiang Astronomical Observatory in high-performance computing support. In future work, Taurus will facilitate research in stellar evolution modeling, radio astronomy polyphase filtering, coherent and incoherent dedispersion, numerical simulations, and numerous other scientific endeavors.

[1] http://taurus.xao.ac.cn/

### 1.1 System Topology

Taurus employs a CPU+GPU hybrid architecture, currently consisting of one login node, one management node, sixteen compute nodes, two I/O nodes, and 100TB of high-speed storage. Each compute node is equipped with two 12-core Intel Xeon E5-2692 v2 CPUs operating at 2.20GHz, 64GB of memory, and one Nvidia Tesla K20m GPU. Compute and I/O nodes are interconnected via a 56Gb Infiniband switch for high-speed data transfer and message passing, while Gigabit Ethernet and IPMI networks handle cluster management. The system topology is illustrated in Figure 1 [Figure 1: see original paper].

### 1.2 Computational Performance Testing

Linpack [?] is the most widely used benchmark for testing floating-point performance of high-performance computing systems and serves as the standard test for the TOP500 supercomputer rankings. The benchmark consists of 64-bit double-precision floating-point operations that measure the number of multiply-add calculations per second (flops) a system can perform. Linpack offers three variants: Linpack 100, Linpack 1000, and HPLinpack [?]. The first two are unsuitable for testing parallel computer clusters; therefore, this study employs HPLinpack to evaluate the Taurus system.

**1.2.1 CPU Performance Testing** Taurus comprises sixteen compute nodes with the configuration detailed in Table 1 . The E5-2692 v2 processor can perform eight operations per clock cycle, yielding a theoretical CPU double-precision floating-point performance of $2.2 \times 8 \times 24 \times 16 = 6758.4$ Gflops for the entire system.

HPLinpack solves a dense linear system of order N using Gaussian elimination to evaluate floating-point performance. The HPL.dat configuration and test results for sixteen nodes are presented in Tables 2 and 3 , respectively. The final results demonstrate that Taurus achieves an actual CPU double-precision floating-point performance of 6.289 Tflops, corresponding to 93.06% efficiency (6.289/6.7584).

**1.2.2 GPU Performance Testing** Each Taurus node includes one Nvidia Tesla K20m GPU based on the Kepler architecture, featuring 2496 CUDA cores, a core frequency of 706MHz, and memory bandwidth of 208GB/s. The theoretical double-precision floating-point performance for sixteen GPUs is 16 $\times$ 1.17 $\times$ $10^3$ Gflops = 18.72 Tflops.

Single-node GPU performance was tested using HPLinpack with the configuration shown in Table 4 . Results for various Ns values are presented in Table 5 . Through extensive testing, we derived an empirical formula for determining the optimal Ns value:

$$Ns = \mathrm{ROUND}\left(\frac{\sqrt{M \times N \times 1024 \times 1024 \times R/8}}{128}, 0\right) \times 128$$

where M is the total memory per node, N is the number of nodes, and R is a coefficient. GPU clusters achieve peak performance when R ranges between 0.8 and 0.9. Testing revealed that the sixteen GPU nodes of Taurus reach a maximum computational peak of 14.882 Tflops, yielding an actual efficiency of 79.5% (14.882/18.72).

### 1.3 Storage System Performance Testing

Taurus utilizes the Lustre file system [?] for storage, providing 100TB capacity. The storage architecture consists of one storage node connected to two expansion enclosures via SAS interfaces in a loopback configuration. Lustre is an open-source, object-based parallel file system that delivers reliable, secure, user-friendly, and scalable storage [?]. The architecture is illustrated in Figure 2 [Figure 2: see original paper].

Taurus employs centralized storage management, with all compute nodes mounting the same file system. To address fault tolerance and I/O bottlenecks [?], two independent controller cards on the storage controller connect via fiber to two redundant I/O nodes. If one I/O node fails, the other immediately takes over, ensuring storage stability. Additionally, I/O nodes and servers are linked via 56Gb Infiniband [?], supporting multiple concurrent connection transitions for high-speed data transfer.

We evaluated the file system using the IOZone benchmark, which tests write, read, re-read, re-write, reverse read, stride read, fread, fwrite, random read, pre-read, memory-mapped file I/O, async I/O read, and async I/O write operations [?]. Test files must be smaller than physical memory to avoid affecting results. Taurus achieves multi-node read speeds of approximately 6GB/s. Write performance is shown in Figure 3 [Figure 3: see original paper], reaching a maximum of 650MB/s for 256MB files with 16MB block sizes.

### 1.4 Software Environment

Taurus employs the PBS workload manager for hardware resource management and allocation, enabling users to submit job scripts, query job status, and manage job queues. The module system manages software and library versions, allowing users to easily configure environment variables and switch between Python versions, MPI parallel environments, and programming models. For GPU programming, each compute node supports the CUDA architecture, enabling direct submission of data-intensive jobs to GPUs via PBS.

## 2. Applications of the High Performance Computing System

Since its deployment, over thirty researchers have utilized Taurus for scientific computing.

### 2.1 Monte Carlo Simulation of Chemical Evolution in Molecular Clouds

Astrochemistry models the synthesis of various molecules in molecular clouds, with Monte Carlo simulation being a common approach [**?**]. This stochastic method simulates the occurrence of individual chemical reactions within a reaction network. The astrochemistry group at Xinjiang Astronomical Observatory uses Taurus to simulate chemical evolution in systems with size distributions, modeling a small volume of a molecular cloud containing a dust grain and surrounding gas. Each system contains multiple chemical reactions that can be simulated computationally. When considering multiple dust grains, parallel evolution is required, with subsystems mixed periodically to maintain homogeneity across the larger system. A major drawback of Monte Carlo methods is their computational expense—simulating 200,000 years of evolution requires dozens of days on a standard workstation. Using Taurus, the same simulation completes in approximately seven days. Simulation results for different molecular evolutions are shown in Figure 4 [Figure 4: see original paper].

### 2.2 GPU-Accelerated Multi-phase Filtering and Radio Frequency Interference Mitigation

Radio Frequency Interference (RFI) identification and mitigation present significant technical challenges, particularly the need to rapidly and accurately distinguish RFI from genuine astronomical signals. The massive data volumes from large-aperture telescopes and complex RFI environments demand real-time processing capabilities. Researchers at Xinjiang Astronomical Observatory are developing CUDA-accelerated RFI mitigation algorithms, having successfully implemented an adaptive RFI processing method on Taurus GPUs with promising results, as shown in Figure 5 [Figure 5: see original paper]. Taurus substantially reduces processing time, providing a hardware platform for coherent dedispersion implementation.

### 2.3 Application Process for Taurus HPC System

Taurus operates under an open management policy and accepts applications from researchers in relevant fields. Specific requirements can be sent to zhanghailong@xao.ac.cn. Detailed usage instructions and example programs are available at http://taurus.xao.ac.cn.

## Conclusion

In response to the scientific computing requirements of Xinjiang Astronomical Observatory, we have constructed a sixteen-node high-performance computing system. Testing demonstrates that the system achieves 93% of theoretical CPU performance and 80% of theoretical GPU performance. The Lustre storage system attains optimal write speeds with 16MB block sizes and files larger than 256MB; smaller files degrade overall storage performance. Currently, over thirty users conduct scientific computing on Taurus, achieving promising results in Monte Carlo simulations of molecular cloud chemistry and GPU-accelerated polyphase filter bank algorithms.

## References

[1] Yang Zherui, Gao Na, Liu Liang. Construction and management of large scale astronomical data analysis and multi-dimensional information visualization platform[J]. E-science Technology & Application, 2015, 6(5): 73-83.

[2] Fan Z, Qiu F, Kaufman A, et al. GPU cluster for high performance computing[C]// Proceedings of the ACM/IEEE SC2004 Conference. 2004: 47-47.

[3] Kirk D. NVIDIA CUDA software and GPU parallel computing architecture[C]// Proceedings of the 6th international symposium on Memory management. 2007: 103-104.

[4] Chervenak A, Foster I, Kesselman C, et al. The data grid: towards an architecture for the distributed management and analysis of large scientific datasets[J]. Journal of Network and Computer Applications, 2000, 23(3): 187-200.

[5] Dongarra J J, Luszczek P, Petitet A. The LINPACK benchmark: past, present and future[J]. Concurrency and Computation: Practice and Experience, 2003, 15(9): 803-820.

[6] Barrett R F, Chan T H F, D' Azevedo E F, et al. Complex version of high performance computing LINPACK benchmark (HPL)[J]. Concurrency and Computation: Practice and Experience, 2010, 22(5): 573-587.

[7] Zhao T, March V, Dong S, et al. Evaluation of a performance model of lustre file system[C]// 2010 Fifth Annual ChinaGrid Conference. 2010: 191-196.

[8] Kosta L, Hunter H, George G, et al. Measuring I/O Performance of Lustre and the Temporary File System for Tradespace Applications on HPC Sys-

tems[C]//Proceedings of the SouthEast Conference. 2017: 187-190.

[9] Yildiz O, Dorier M, Ibrahim S, et al. On the root causes of cross-application I/O interference in HPC storage systems[C]// 2016 IEEE International Parallel and Distributed Processing Symposium. 2016: 750-759.

[10] Pfister G F. An introduction to the infiniband architecture[M]// High Performance Mass Storage and Parallel I/O: Technologies and Applications. 2001: 617-632.

[11] Zhang Chunming, Rui Jianwu, He Tingting. An approach for storing and accessing small files on Hadoop[J]. Computer Applications and Software, 2012, 29(11): 95-100.

[12] Lamberts T, Cuppen H M, Ioppolo S, et al. Water formation at low temperatures by surface O  hydrogenation III: Monte Carlo simulation[J]. Physical Chemistry Chemical Physics, 2013, 15(21): 8287-8302.

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