Lessons Learned from Optimizing the Sunway Storage System for Higher Application I/O Performance

Qi Chen, Kang Chen, Zuo-Ning Chen, Wei Xue, Xu Ji and Bin Yang

Chen Q, Chen K, Chen ZN et al. Lessons learned from optimizing the sunway storage system for higher application I/O performance. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 35(1): 47–60 Jan. 2020. DOI 10.1007/s11390-020-9798-5

Research problem

- Problem: It is hard for applications to make full utilization of the peak bandwidth of the storage system in high-performance computers (HPC) because of
  - I/O interferences from different applications.
  - storage resource misallocations due to wrong I/O patterns or performance fault of servers and devices.
  - complex long I/O paths from applications to storage devices.
- Prior researches are focused on some topic of them, but lack of system-level analysis and end-to-end system-level optimization.

kernel contributions

- Present the Sunway storage system and its design considerations leading to this architecture
- Introduce the end-to-end performance monitoring and diagnosis tool Beacon
- Present the limitations of the static I/O forwarding strategy and its solution the application-aware I/O forwarding allocation framework (DRFA)
- Introduce our performance-ware data placement framework that is used to avoid I/O interference and performance anomalies of storage devices.
- Introduce our remote node-local storage stack to shorten I/O path and improve metadata performance of applications with N-N I/O pattern.

conclusions

- Our works addressed I/O interferences and storage resource misallocations in the storage stack that limited application I/O performance.
  - DRFA addressed the I/O interference and resource misallocation at the I/O forwarding layer.
  - The data placement framework efficiently controlled data distribution of parallel I/O processes in the parallel file system.
  - The remote node-local storage system shortened the I/O path from computing nodes to storage devices and break the metadata bottleneck of parallel I/O processes with N-N I/O pattern.
- Our studies and lessons can provide some reference to other storage systems that adopt an I/O forwarding architecture.