



YanRong Technology and Mellanox High Performance Distributed File Storage Joint Solution

Overview

YRCloudFile, the distributed file storage vendor from Beijing YanRong Technology Co., Ltd. (hereinafter referred to as YanRong Technology), fully supports InfiniBand and RoCE high-speed RDMA networks with enterprise features, such as ultra-high performance, high availability, high reliability, horizontal expansion, and flexible configuration. The YRCloudFile metadata cluster uses a fully distributed and scalable architecture to support horizontal capacity and performance expansion, providing standard POSIX interfaces that can support access for thousands of clients, simultaneously.

Mellanox EDR 100Gb/s InfiniBand switches and ConnectX-5 VPI EDR adapters provide single-port 100Gb/s full line rate, and zero packet loss network performance without the complex configuration of port operating modes and link speeds. The InfiniBand network is easy to deploy and manage, dramatically reducing network complexity and providing significant advantages for high-performance storage system solutions that require scalability and reliability.

To make better use of the performance benefits of RDMA network offload architecture, YRCloudFile implemented a lot of software optimization for RDMA in data/task affinity, cache alignment, request encapsulation, and page locking, etc. YRCloudFile is a file system that provides a global namespace, can be deployed on commercial x86 servers, and supports the mainstream LINUX operating system (CentOS/RHEL/SUSE), with a unified file system managing all disks on servers.

RDMA Advantages

The statistics show that when using traditional Ethernet for data transmission, the CPU spends a lot of time on memory copy operations of/for the TCP stack, preventing further reduction of write latency and full use of the high-performance disks. YRCloudFile leverages RDMA to reduce the write latency of distributed storage clusters by 60% and improve IO performance by 400% with the help of RDMA's zero-copy memory features.

By applying the Remote Direct Memory Access (RDMA) technology of the EDR InfiniBand network, I/O tasks can be handled without the CPU, freeing up compute resources to accelerate application performance and increase data transfer efficiency. RDMA enables network adapters to directly access application data, bypassing the kernel, CPU, and protocol stack, and allowing the CPU to work on more valuable tasks while transferring data through the network.

YRCloudFile implements an RDMA-based network with the following hardware options:

- End-to-end InfiniBand network: VPI network adapter + InfiniBand switch
- End-to-end Ethernet RoCE network: Ethernet network adapter + Ethernet switch

PERFORMANCE TEST RESULTS

In the benchmark performance test, four 2U x86 client servers were used to test the YRCloudFile performance. IO requests are initiated from clients through EDR InfiniBand to YRCloudFile storage clusters composed of two x86 servers (with two 3.5TB NVMe disks in each server respectively).

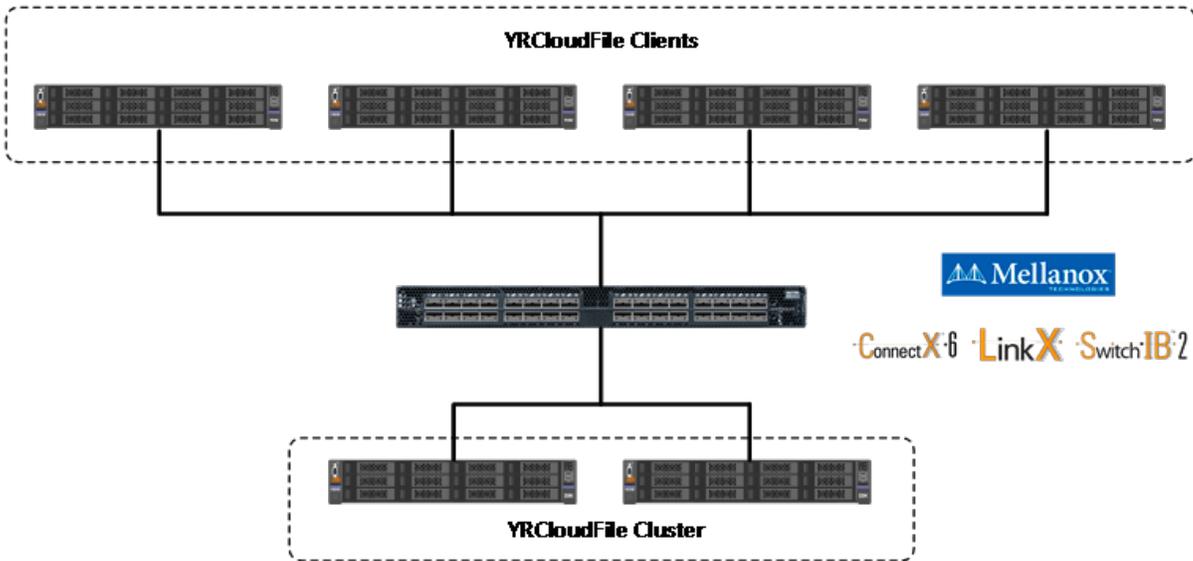


Figure 1: Benchmark testing YRCloudFile performance

In a 4K random read and write test, the performance of a single NVMe disk reached 220,000 IOPS, and two to 4 NVMe disks displayed a linear growth capability. At high IO pressures, the random read latency is approximately 300 microseconds and the write latency is always less than 200 microseconds.

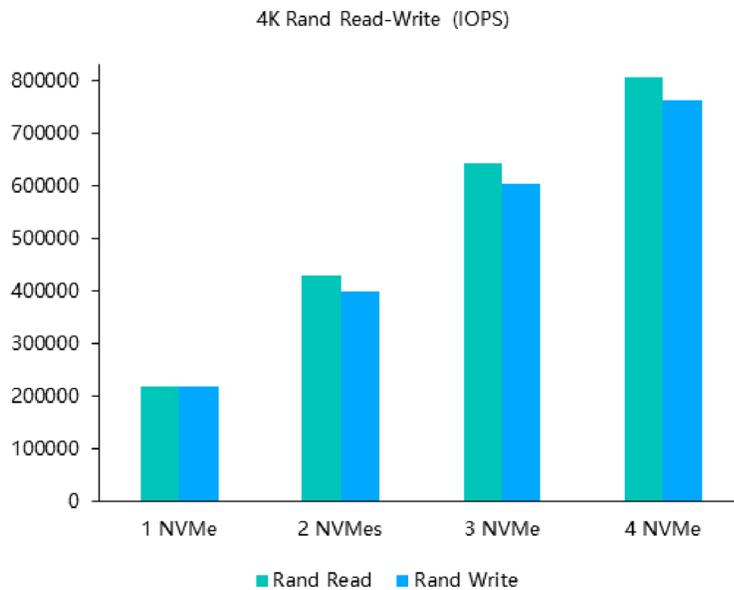


Figure 2: 4K Rand Read-Write (IOPS)

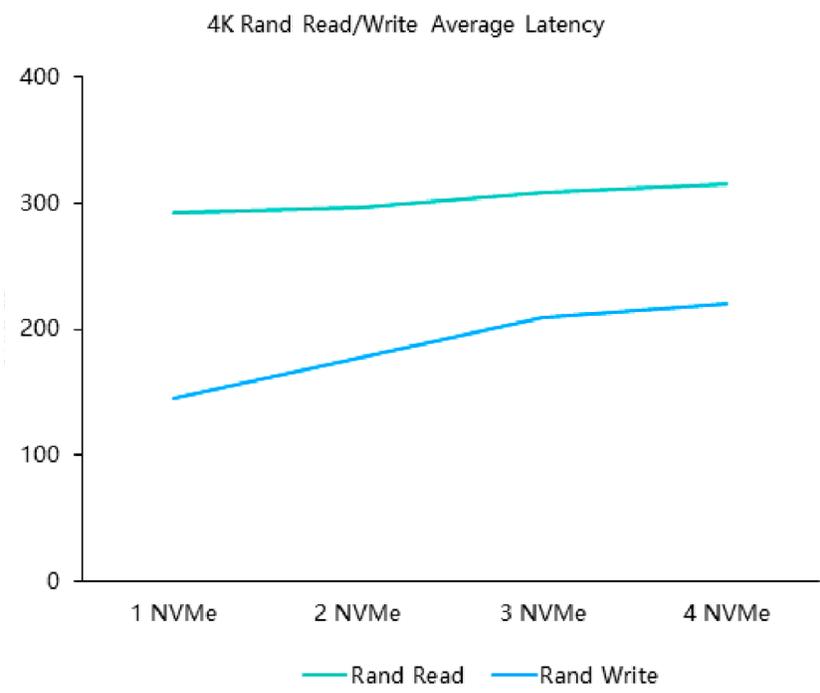


Figure 3: 4K Rand Read/Write Average Latency

In a 512K sequential read/write test, the bandwidth performance in a cluster architecture reaches 10GB/s and 6.5GB/s (the upper limit of the DUT disk’s performance), whether the environment supports two clients or 4 clients. With the support of an EDR 100Gb InfiniBand network, YRCloudFile can take full advantage of the physical performance of NVMe disk bandwidth. At the same time, 512K read/write latency remains under 2.5 milliseconds.

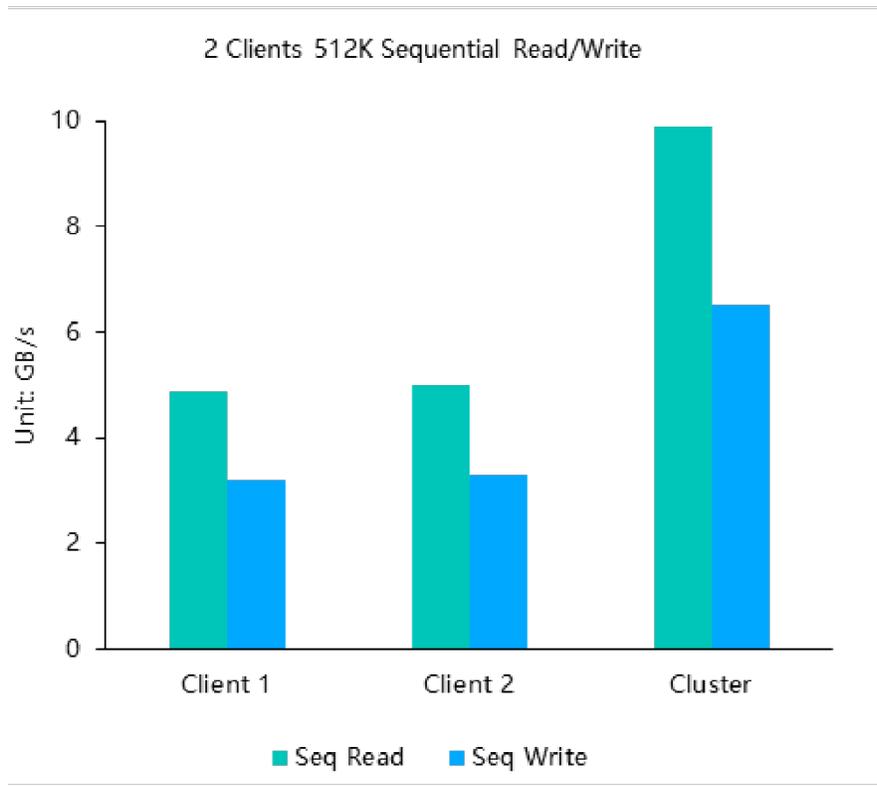


Figure 4: 512 sequential read/write test: 2 clients

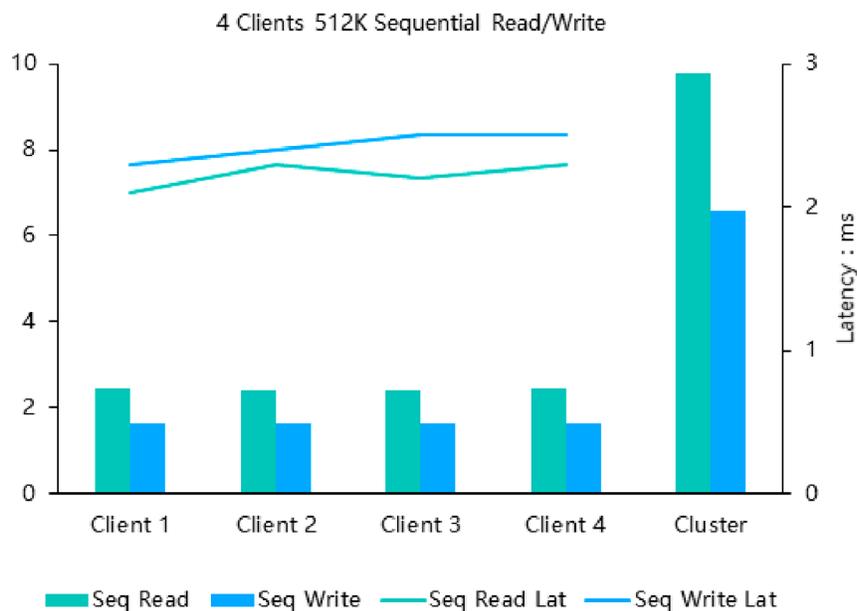


Figure 5: 512 sequential read/write test: 4 clients

Application Scenarios

YRCLOUDFILE and Mellanox jointly provide a high-performance distributed file storage solution, designed for container storage, high-performance computing, AI, and content management systems, etc. The solution greatly improves the performance of parallel access to small cluster storage files and simplifies management and maintenance.

Based on YRCLOUDFILE, the high-performance, high-availability, high-reliability unified storage platform provides users with the following technical advantages:

- High bandwidth, IOPS and low latency distributed file storage
- Support for multiple versions of Kubernetes, Docker Swarm, Mesos and other mainstream container orchestration platform; with lifecycle management of various files
- Support for standard POSIX interfaces

About YanRong Yun

YanRong is a high-tech enterprise with software-defined storage technology as its core competitiveness. It has independent intellectual property rights in key technologies such as distributed storage and is an industry leader of high-performance distributed cloud storage solutions. YanRong provides personalized solutions and one-stop services for various industries.

Based on self-developed distributed software-defined storage products, YanRong provides customers with efficient and stable storage products for their valuable data. The high performance and high reliability of YRCloudFile can satisfy the needs in different application scenarios, such as high performance computing, AI, container storage and so on.

More information is available at www.yanrongyun.com.

About Mellanox Technologies

Mellanox Technologies (NASDAQ: MLNX) is a leading supplier of end-to-end Ethernet and InfiniBand smart interconnect solutions and services for servers and storage. Mellanox interconnect solutions increase data center efficiency by providing the highest throughput and lowest latency, delivering data faster to applications, unlocking system performance and improving data security. Mellanox offers a choice of fast interconnect products: adapters, switches, software and silicon that accelerate application performance and maximize business results for a wide range of markets including cloud and hyperscale, high performance computing, artificial intelligence, enterprise data centers, cyber security, storage, financial services and more. More information is available at: www.mellanox.com.



350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085

Tel: 408-970-3400 • Fax: 408-970-3403

www.mellanox.com