Optimizing Red Hat Ceph Storage Performance

Objectives

After completing this section, you should be able to choose Red Hat Ceph Storage architecture scenarios and operate Red Hat Ceph Storage-specific performance analysis tools to optimize cluster deployments.

Defining Performance Tuning

Performance tuning is the process of tailoring system configurations, so that specific, critical applications have the best possible response time or throughput. Performance tuning for a Ceph cluster has three metrics: latency, IOPS (input/output operations per second), and throughput.

Latency

It is a common misconception that disk latency and response time are the same thing. Disk latency is a function of the device, but response time is measured as a function of the entire server.

For hard drives using spinning platters, disk latency has two components:

- Seek time: The time it takes to position the drive heads on the correct track on the platter, typically 0.2 to 0.8 ms.
- Rotational latency: The additional time it takes for the correct starting sector on that track to pass under the drive heads, typically a few milliseconds.

After the drive has positioned the heads, it can start transferring data from the platter. At that point, the sequential data transfer rate is important.

For solid-state drives (SSDs), the equivalent metric is the random access latency of the storage device, which is typically less than a millisecond. For non-volatile memory express drives (NVMes), the random access latency of the storage drive is typically in microseconds.

I/O Operations Per Second (IOPS)

The number of read and write requests that the system can process per second depends on the storage device capabilities and the application. When an application issues an I/O request, the operating system transfers the request to the device and waits until the request completes. As a reference, hard drives using spinning platters achieve between 50 and 200 IOPS, SSDs are on the order of thousands to hundreds of thousands IOPS, and NVMes achieve some hundreds of thousands IOPS.

Throughput

Throughput refers to the actual number of bytes per second the system can read or write. The size of the block and the data transfer rate affect the throughput. The higher the disk block size, the more you attenuate the latency factor. The higher the data transfer rate, the faster a disk can transfer data from its surface to a buffer.

As a reference value, hard drives using spinning platters have a throughput around 150 Mb/s, SSDs are around 500 Mbps, and NVMes are in the order of 2,000 Mb/s.

You can measure throughput for networks and the whole system, from a remote client to a server.