BlueStore Performance

FileStore writes to a journal and then writes from the journal to the block device. BlueStore avoids this double-write performance penalty by writing data directly to the block device and logging transactions to the write-ahead log simultaneously with a separate data stream. BlueStore write operations are approximately twice as fast as FileStore with similar workloads.

When using a mix of different cluster storage devices, customize BlueStore OSDs to improve performance. When you create a BlueStore OSD, the default is to place the data, block database, and write-ahead log all on the same block device. Many of the performance advantages come from the block database and the write-ahead log, so placing those components on separate, faster devices might improve performance.

Note

You might improve performance by moving BlueStore devices if the new device is faster than the primary storage device. For example, if object data is on HDD devices, then improve performance by placing the block database on SSD devices and the write-ahead log on NVMe devices.

Use service specification files to define the location of the BlueStore data, block database, and write-ahead log devices. The following example specifies the BlueStore devices for an OSD service.

```
service_type: osd
service_id: osd_example
placement:
  host_pattern: '*'
data_devices:
  paths:
    - /dev/vda
db_devices:
  paths:
    - /dev/nvme0
wal_devices:
  paths:
    - /dev/nvme1
```

The BlueStore storage back end provides the following features:

- Allows use of separate devices for the data, block database, and write-ahead log (WAL).
- Supports use of virtually any combination of HDD, SSD, and NVMe devices.
- Operates over raw devices or partitions, eliminating double writes to storage devices, with increased metadata efficiency.
- Writes all data and metadata with checksums. All read operations are verified with their corresponding checksums before returning to the client.

The following graphs show the performance of BlueStore versus the deprecated FileStore architecture.