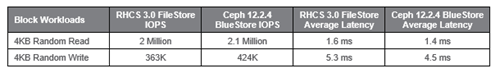
**Micron NVMe SSD 性能比较：Ceph BlueStore VS FileStore**

BlueStore是Ceph的新存储引擎，是社区版的默认配置。 BlueStore性能参数不包含在我们当前的Micron Accelerated Ceph存储解决方案参考架构中，因为Red Hat Ceph 3.0目前不支持它。 本文在Ceph参考架构硬件上对社区版Ceph Luminous（12.2.4）进行了性能测试，并将结果与同样硬件环境的RHCS 3.0中实现的FileStore性能进行比较。

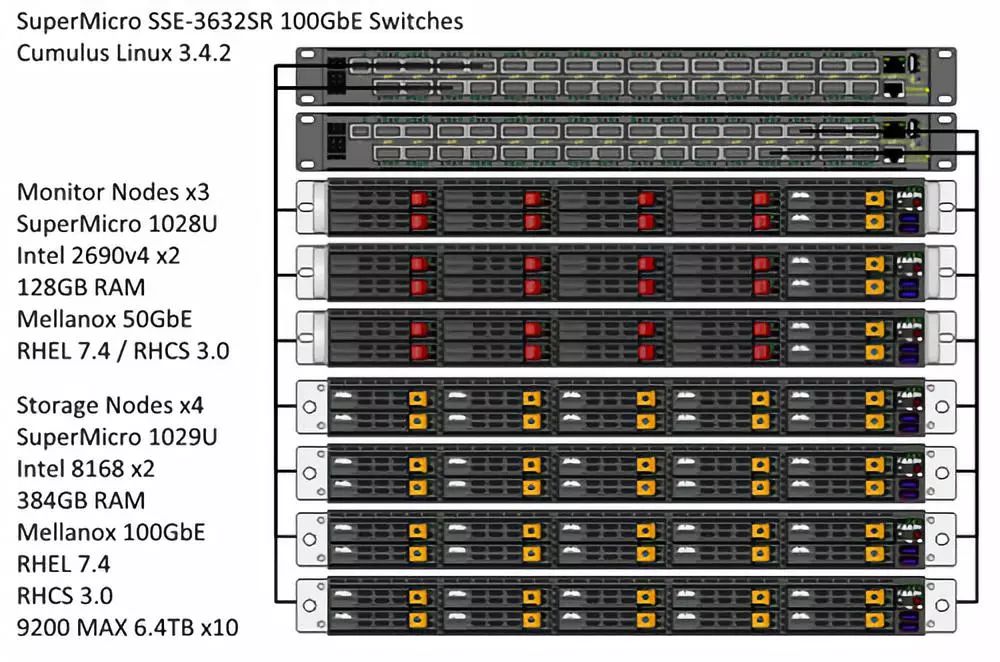
4KB随机写入IOPS性能提高18％，平均延迟降低15％，尾延迟降低99.99％高达80％。 使用BlueStore，在较高的队列深度下，4KB随机读取性能更好。



该解决方案针对块性能进行了优化。 使用Linux中的Rados Block Driver进行随机小块测试，在双插槽存储节点中使铂金级8168 Intel Purley处理器满负荷运行。

每个存储节点有10个驱动器，该架构具有232TB的可用存储容量，可通过添加额外的1U存储节点进行扩展。

**参考设计 - 硬件**



**测试结果和分析**

Ceph测试方法

Red Hat Ceph Storage 3.0（12.2.1）配置了FileStore，每个Micron 9200MAX NVMe SSD有2个OSD。 每个OSD使用20GB日志。

Ceph Luminous Community（12.2.4）配置了BlueStore，每个Micron 9200MAX NVMe SSD有2个OSD。 RocksDB和WAL数据与数据存储在同一分区中。

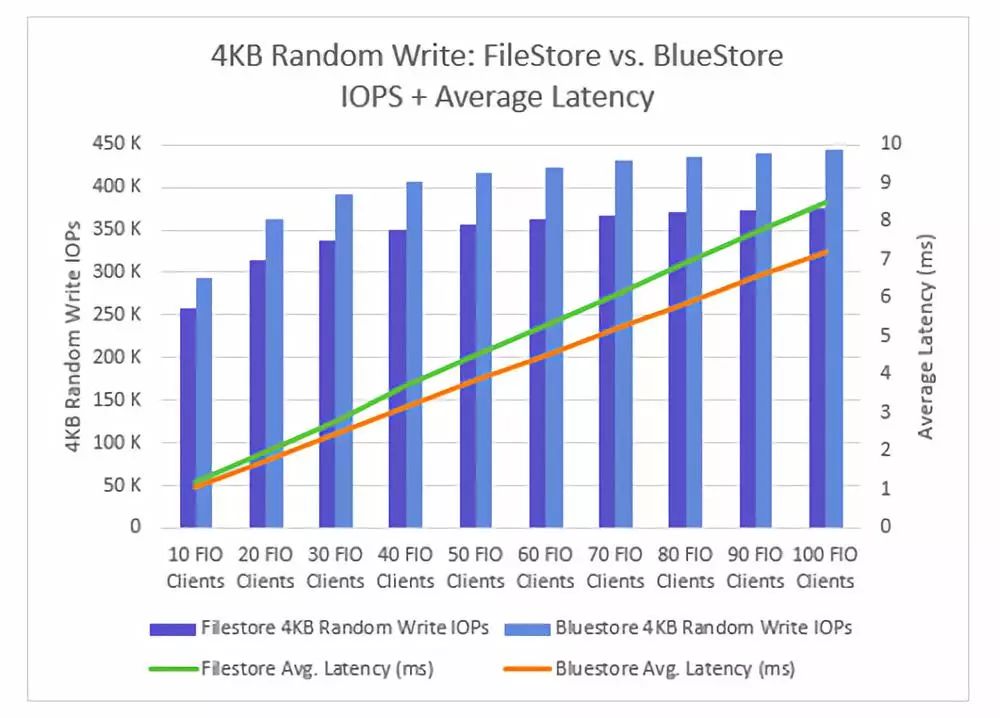
在这两种配置中，每个存储节点有10个驱动器，每个驱动器有2个OSD，总共80个OSD，可用容量为232TB。

测试的Ceph存储池是使用8192个pg和2副本创建的。 使用100个RBD快文件（每个75GB）测试性能，在2副本存储池上提供7.5TB数据，总数据为15TB。

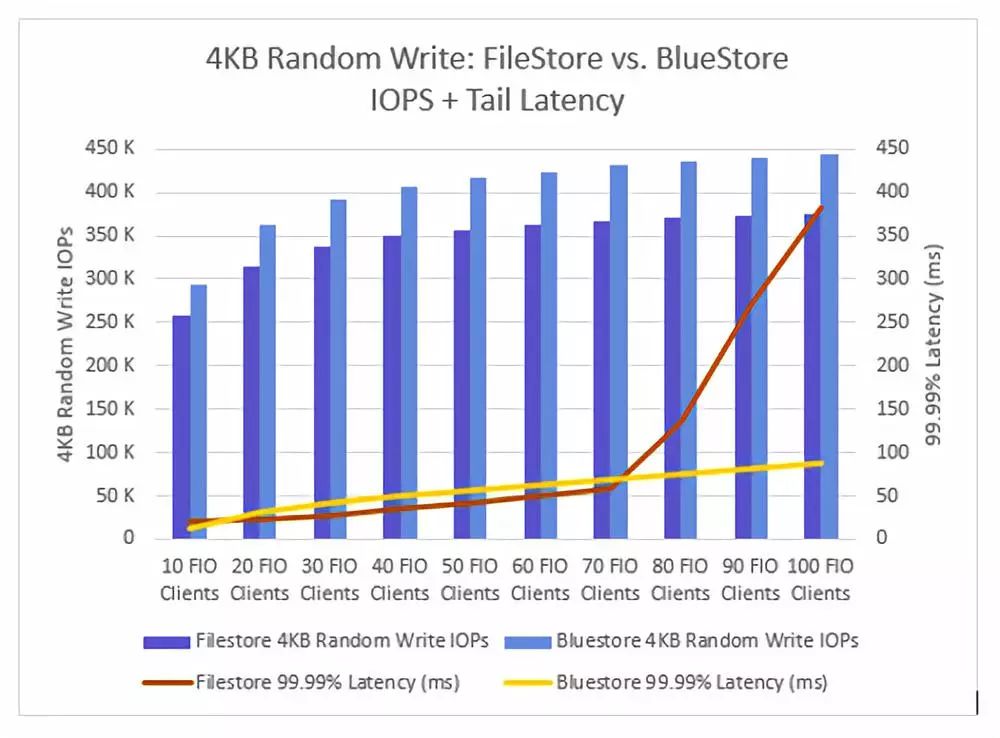
使用FIO针对Rados Block Driver测试4KB随机块性能。 我们在所有测试中都受CPU限制，即使每个存储节点使用2x Intel 8168 CPU也是如此。

**RBD FIO 4KB随机写性能：FileStore与BlueStore**

BlueStore的IOPS增加约18％，平均延迟减少约15％。

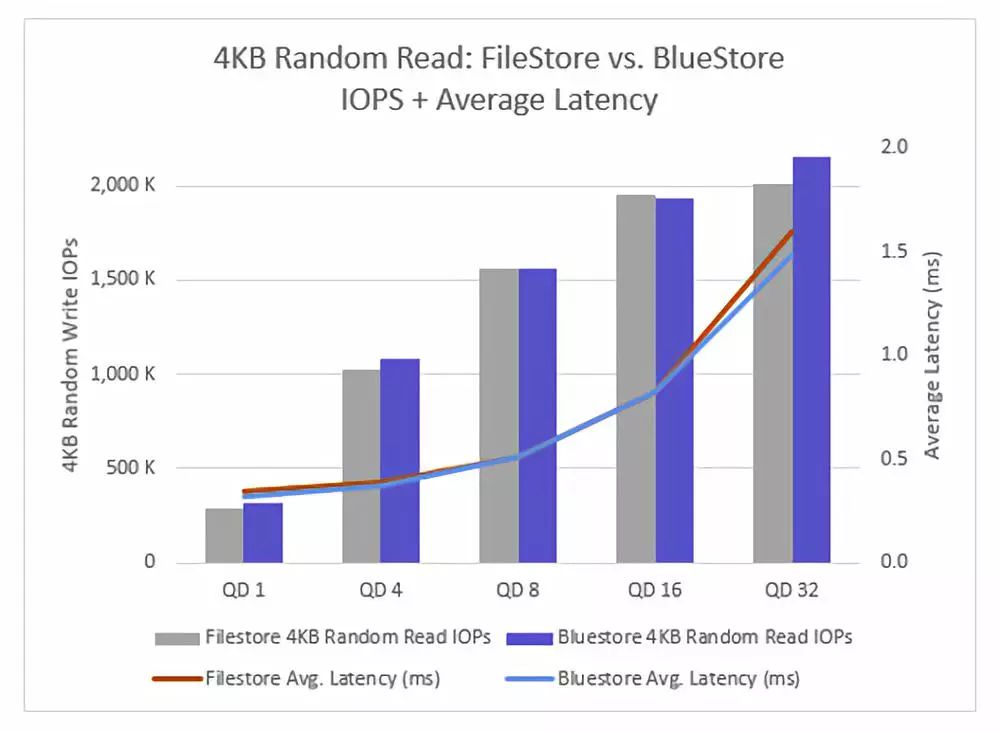


使用BlueStore，在更高的FIO客户端数量下，Ceph的尾延迟也大大减少。 在100个客户端，尾延迟减少了4.3倍。 在较低的客户端数量下，BlueStore的尾延迟高于FileStore，因为BlueStore正在推动更高的性能。

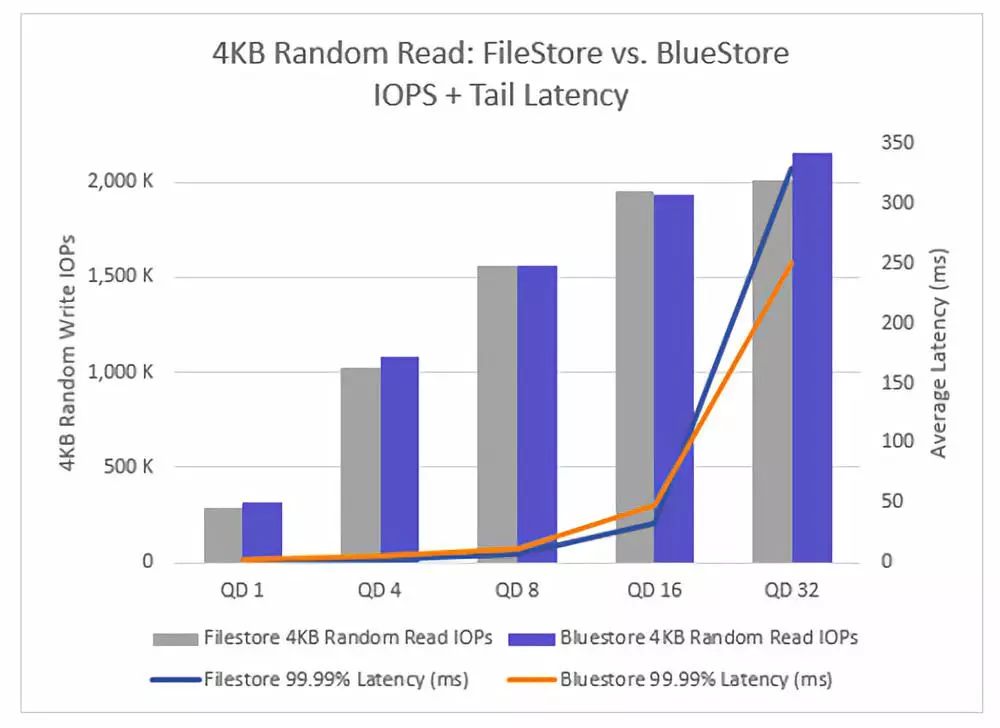


**RBD FIO 4KB随机读取性能：FileStore与BlueStore**

FileStore和BlueStore之间的4KB随机读取性能相似。 队列深度为32时，IOPS增加了5％。



尾延迟也类似于队列深度32，其中BlueStore表现更好。



原文链接：

https://www.micron.com/about/blogs/2018/may/ceph-bluestore-vs-filestoreblock-performance-comparison-when-leveraging-micron-nvme-ssds

来源：

本文授权【新钛云服公众号】