Crimson Seastore: performance scaling – progress report

José Juan Palacios Pérez
Ceph IBM,
Manchester UK

July 1, 2025

DRAFT

Abstract

In this brief report we show the performance scaling as incrementing the number of reactors for a single OSD, for Crimson SeaStore using the following configurations:

- we ranged the number of Seastar reactors as 4, 8, 16 and 56.
- All configurations used a single NUMA socket.
- We used the same ceph dev build from main branch (hash 87d2085) for all test runs.
- We only used the balanced OSD algorithm.

Our preliminary conclusions:

- The performance of the Crimson SeaStore on random workloads scales well with the number of reactors. However, sequential workloads don't scale well beyond 16 reactors.
- We are using flamegraphs data for 4 and 8 reactors to work out attribution on the hottest code paths.

Contents

DRAFT

1. Seastore performance scaling

In this Chapter we show the performance scaling of Seastore (build 87d2085) on longer duration test sets, producing response latency curves.

1. We used a single OSD, running within a single NUMA socket, we only increased the number of reactors as indicated.
2. We used the balanced OSD algorithm.
3. We used 32 RBD images, each of 2GB size, four jobs per image. All FIO processes were pinned to the other NUMA socket, so that the OSD processes were not affected by the FIO processes.
4. We disabled RBD coalescing so the sequential workloads would not be affected by the coalescing algorithm.

1.1 Random read 4k

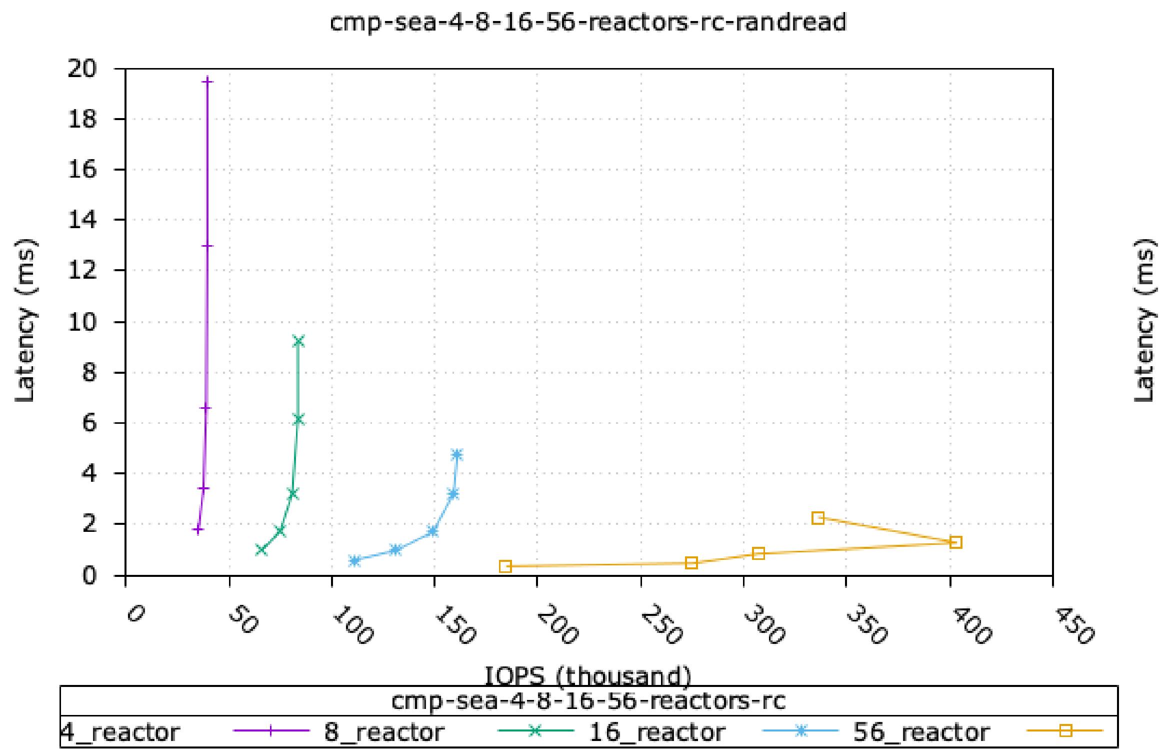
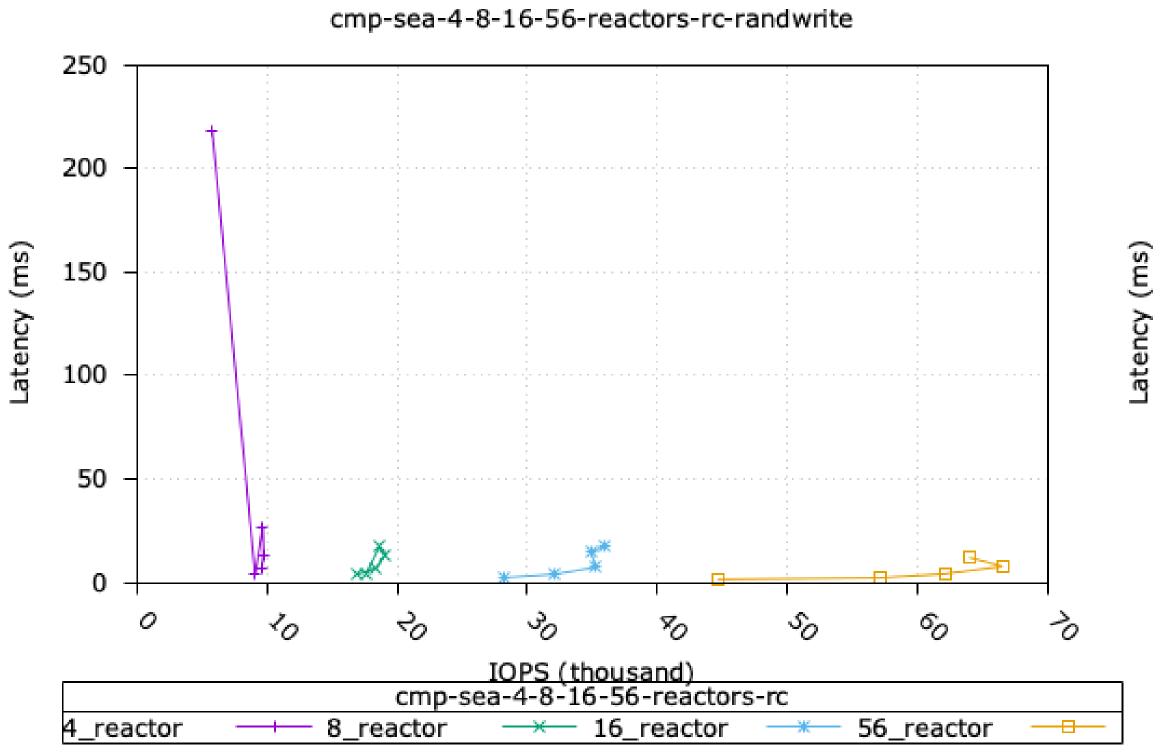


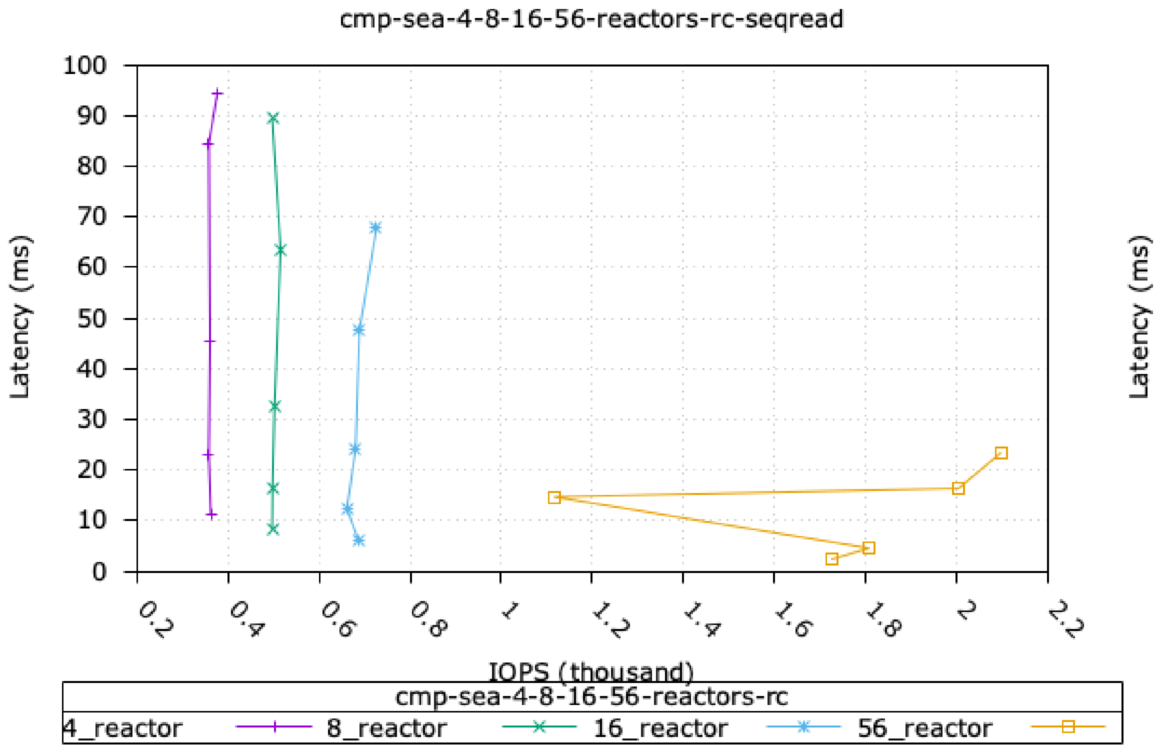
Figure 1.1: Response latency curves, IOPs vs. Latency. Each curve is for a different number of reactors, 4, 8, 16 and 56. Each datra point corresponds to an increasing IO depth, from 1 2 4 8 16 24 32 40 52 to 64. If the latency is higher than 1 sec, the data point is not shown. The x-axis is the IOPs, the y-axis is the latency in milliseconds.

1.2 Random write 4k



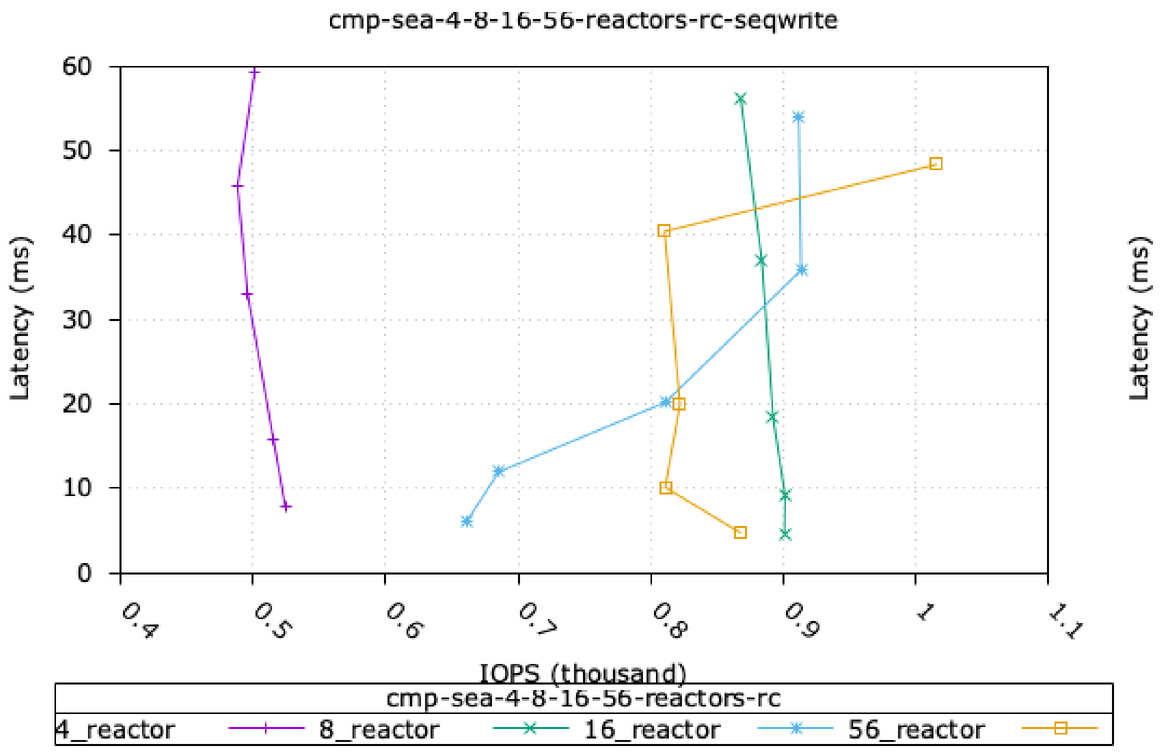
1.3 Sequential read 64k

DRAFT



1.4 Sequential write 64k

DRAFT


cmp-sea-4-8-16-56-reactors-rc-seqwrite
Latency (ms)
Latency (ms)
IOPS (thousand)
0
10
20
30
40
50
60
0.4
0.5
0.6
0.7
0.8
0.9
1
1.1
cmp-sea-4-8-16-56-reactors-rc
4_reactor
8_reactor
16_reactor
56_reactor