napp-it

napp-*it*

ZFS Epyc ESXi/ Storage Server

Some short performance tests
to compare

- Xeon Silver 4110 vs Epyc 7302
- Disk vs Flash vs Optane
- OmniOS 151034 vs 151036

- barebone vs virtualized

Content:

1.0 Intel Xeon vs AMD Epyc

This benchmark sequence was intended to answer some basic questions about how the new Epyc platform performs my current Intel Xeon silver on a disk pool vs Optane/Flash pool. Additionally we check encryption performance and OmniOS 151036 vs 151036 with FPU accelerated raid-Z. The Xeon system was my typical VM and storage server up to this year (without disks typically from the 3000 Euro region), the Epyc may become my next standard platform in this price range. While it is clear that a system from this year will always outperform one from last year at same price, I want to know how much and especially how fast is ZFS encryption now as encryption is a mandatory feature under the European data protection rules.

All tests are done via a filebench run that checks random io vs sequential io with sync and enc on/off
and barebone vs virtualised in napp-it menu Pools > Benchmark. Intension is to get a feeling about behaviours.

Intel Hardware:
SuperMicro X11SPH-NCTF, Xeon Silver 4110, 64 GB RAM, SAS 3008
7 x WD Ultrastar 8TB, 3 x Optane 900, 3 x Intel DC 750 (traditional Flash NVMe)

AMD Hardware:
SuperMicro H12SSL-C, Epyc 7302 128 GB RAM, SAS 3008 (BTO system)
7 x WD Ultrastar 8TB, 3 x Optane 900, 3 x Intel DC 750 (traditional Flash NVMe)

2.0: TLDR

First write results on Intel

151034	async	sync	sync+slog	async/enc	sync/enc	sync+slog/enc
Disk Pool	770-1195MB/s	40-50MB/s	640-783MB/s	907-935MB/s	22-24MB/s	19-24MB/s
Flash pool	1112-1388MB/s	430-613MB/s	-	950-954MB/s	112-129MB/s	-
Optane pool	1600-2000MB/s	550-770MB/s	-	930-960MB/s	118-130MB/s	-

151036	async	sync	sync+slog	async/enc	sync/enc	sync+slog/enc
Disk Pool	930-1170MB/s	48-50MB/s	779-1195MB/s	907-935MB/s	34-36MB/s	22-24MB/s
Flash pool	1112-1388MB/s	430-613MB/s	-	950-954MB/s	112-129MB/s	-
Optane pool	1600-2000MB/s	550-750MB/s	-	930-960MB/s	107-130MB/s	-

First result on AMD

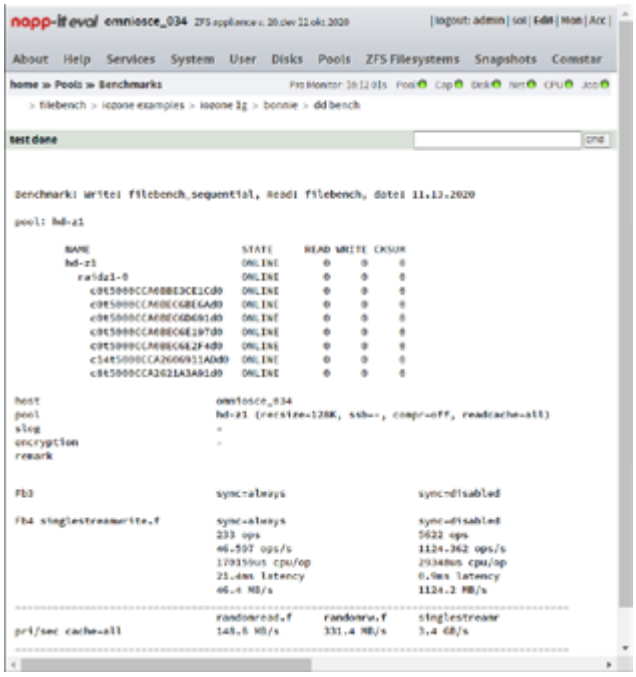
151036	async	sync	sync+slog	async/enc	sync/enc	sync+slog/enc
Disk Pool	1178-1197 MB/s	51-55 MB/s	908-956 MB/s	1203-1230 MB/s	36-40 MB/s	31-32 MB/s
Flash pool	1591-1877 MB/s	943-1134 MB/s	-	1541-1592 MB/s	201-206 MB/s	-
Optane pool	3339-3707 MB/s	1445-1471 MB/s	-	1720-1730 MB/s	202-207 MB/s	-

Some principles
Encryption is much faster than on my last tests
Storage + Encryption is perfect up to 10G,
Storage + sync is perfect up to 10G,
Storage + encryption + sync is much slower due encryption
OmniOS 1510034/36 perform quite similar (despite hardware accelerated Raid-Z)

Improvements due a faster CPU (Epyc): 150-200%

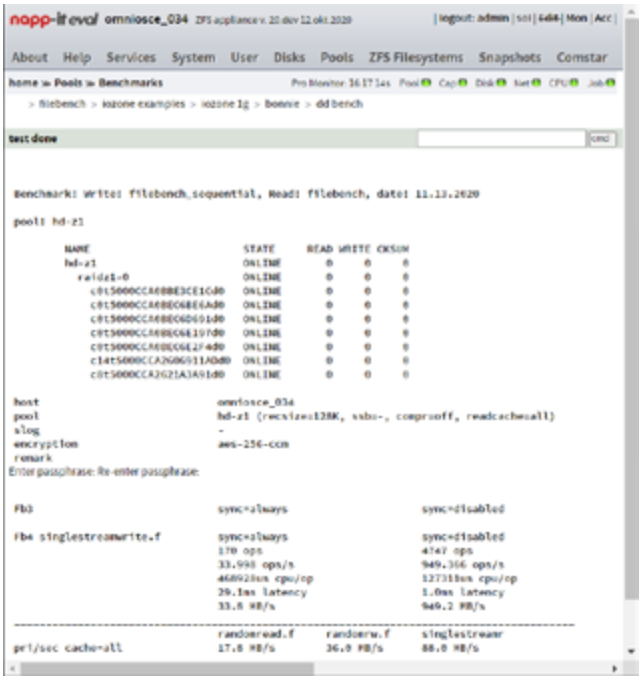
2.1: Disk pool, Intel Xeon on OmniOS 151034

Disk Pool: without slog, enc off



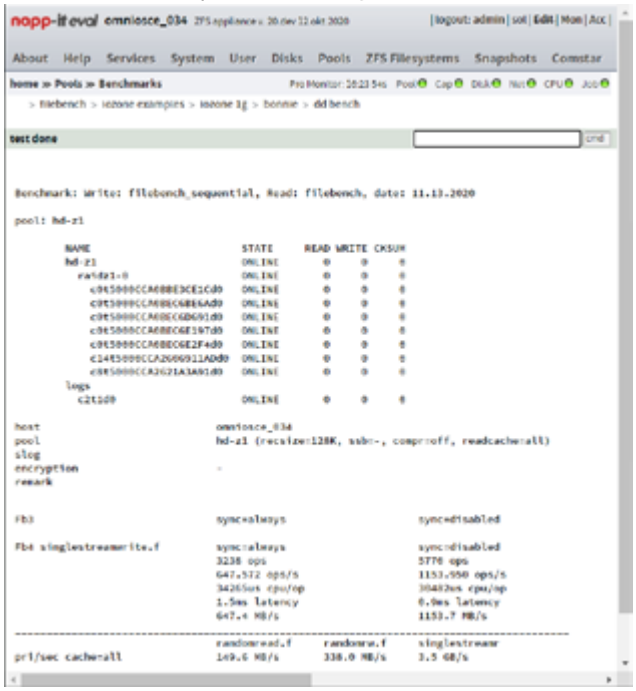
I startet the write benchmark multiple times:
sync was between 40 MB/s and 50 MB/s
non-sync between 770 MB/s and 1195 MB/s

without slog, enc on (aes-256ccm)



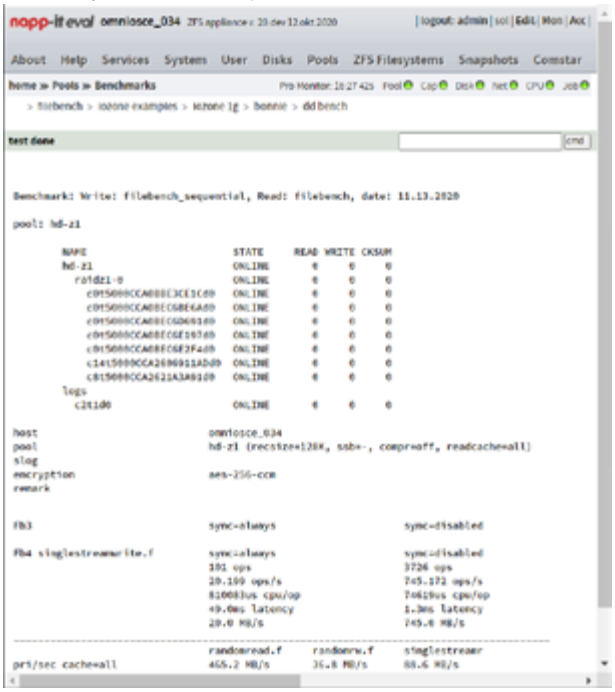
sync was between 22 MB/s and 24 MB/s and
non-sync was between 907 MB/s and 935 MB/s

Disk Pool: with Optane 900 slog, enc off



I startet the write benchmark multiple times:
sync was between 640 MB/s and 783 MB/s
non-sync was between 1002 MB/s and 1195 MB/s

with Optane 900 slog, enc on (aes-256ccm)



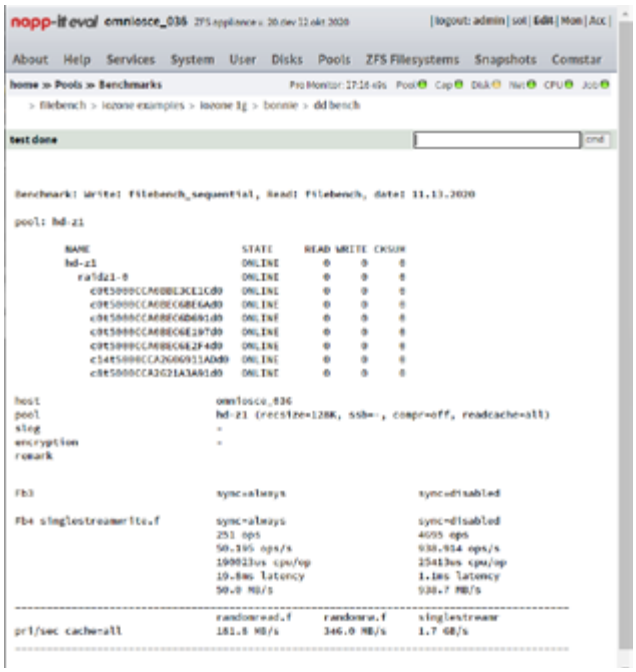
sync was between 19 MB/s and 24 MB/s and
non-sync was between 745 MB/s and 767 MB/s

A Xeon SMB filer with sync=disabled is a perfect solution for a 10G network.
If you need to force sync either because you use it for VM/databases or want an SMB filer with a perfect behaviour in a crash/ power outage scenario, you need a good Slog like the Optane or a WD SS530 (12G SAS).

If you need encryption
This is ok for a SMB filer without sync. If you need sync, do not pair with encryption.

2.2: Disk pool, Intel Xeon on OmniOS 151036

Disk Pool: without slog, enc off

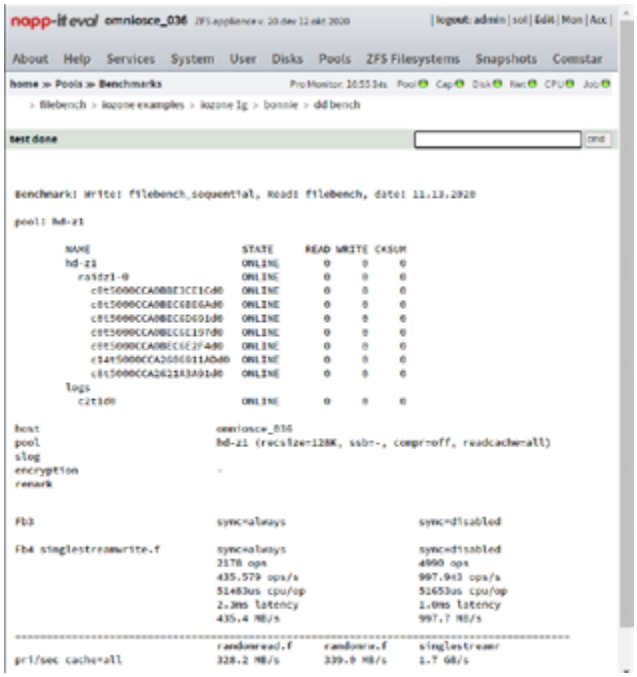


I startet the write benchmark multiple times:
sync was between 48 MB/s and 50 MB/s
non-sync between 930 MB/s and 1170 MB/s

without slog, enc on (aes-256ccm)

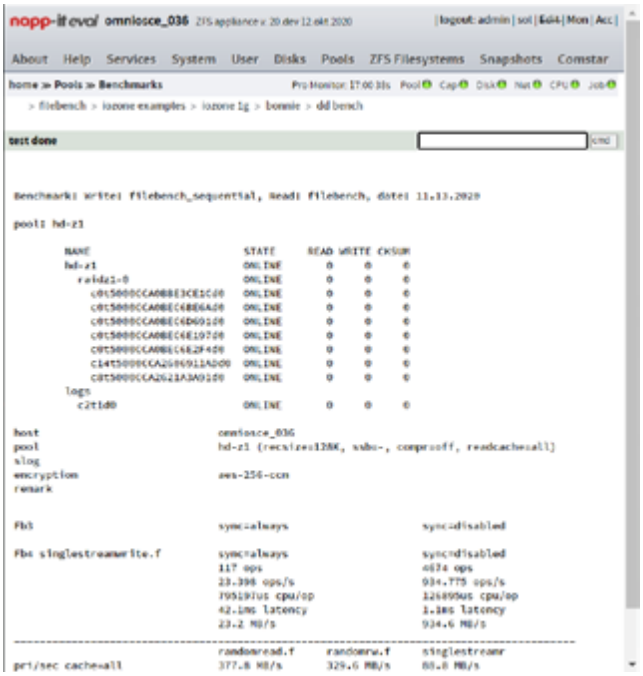
sync was between 34 MB/s and 36 MB/s
non-sync was between 890 MB/s and 924 MB/s

Disk Pool: with Optane 900 slog, enc off



I startet the write benchmark multiple times:
sync was between 359 MB/s and 435 MB/s
non-sync between 770 MB/s and 1195 MB/s

with Optane 900 slog, enc on (aes-256ccm)

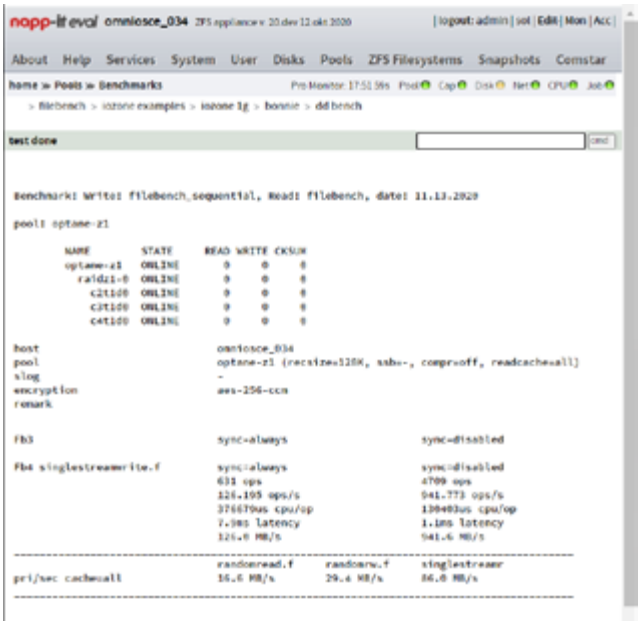


sync was between 22 MB/s and 24 MB/s and
non-sync was between 907 MB/s and 935 MB/s

2.3: Optane pool, Intel Xeon on OmniOS 151034 vs 151036

Optane Pool: 151034, enc off

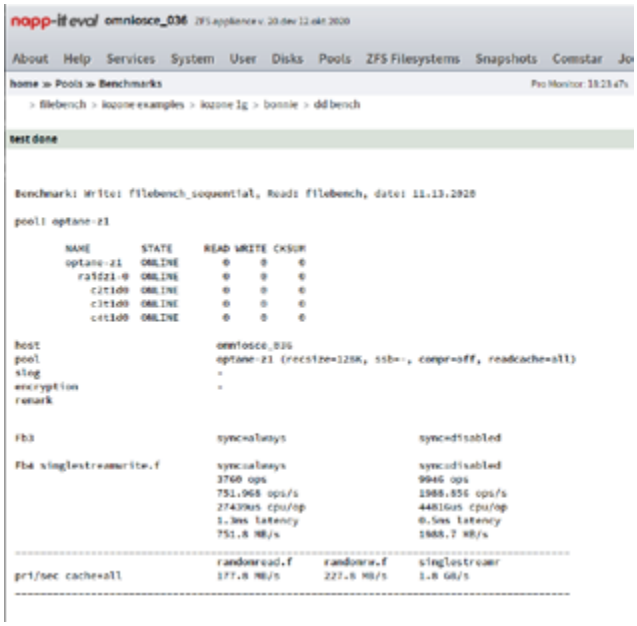
Optane Pool: 151034, enc on (aes-256ccm)



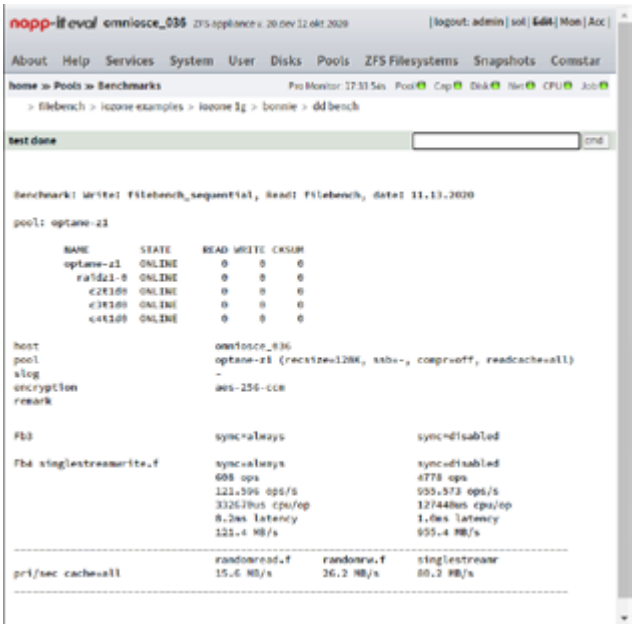
I startet the write benchmark multiple times:
sync was between 550 and 770 MB/s
non-sync between 1600 and 2000 MB/s

sync was between 118 and 130 MB/s
non-sync was between 930 and 960 MB/s

Optane Pool: 151036 enc off



Optane Pool: 151036, enc on (aes-256ccm)

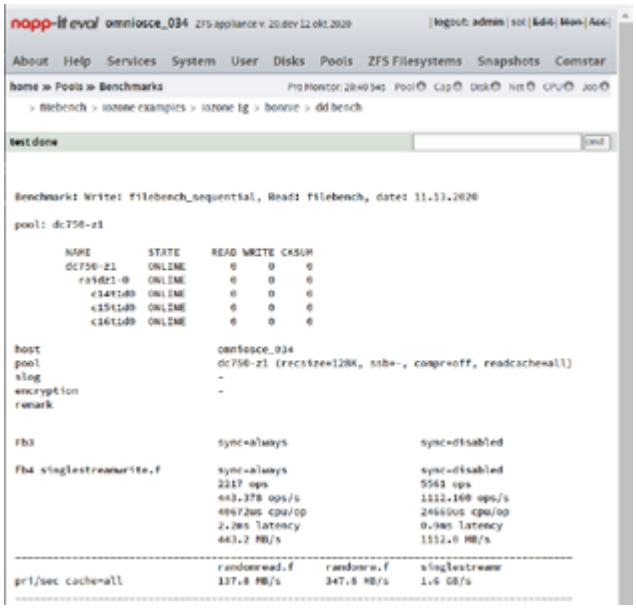


I startet the write benchmark multiple times:
sync was between 550 and 770 MB/s
non-sync between 1600 and 2000 MB/s

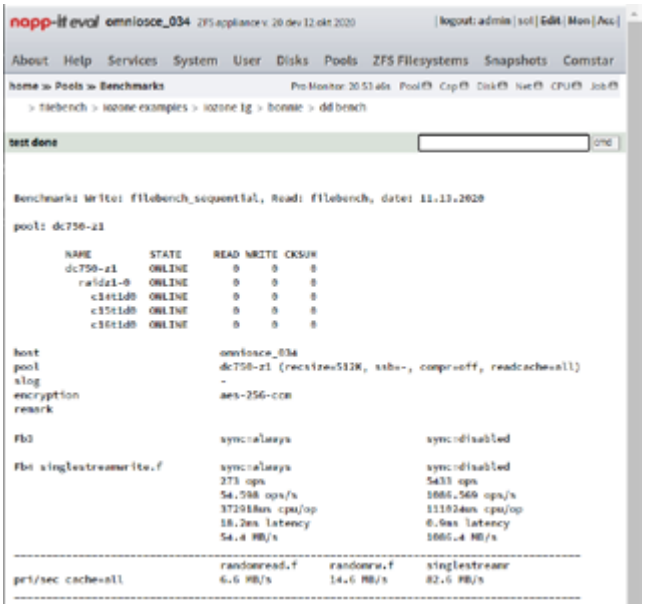
sync was between 107 and 130 MB/s
non-sync was between 930 and 960 MB/s

2.4: Traditional Flash NVMe (Intel DC750) pool, Intel Xeon on OmniOS 151034 vs 151036

DC 750 Pool: 151034, enc off



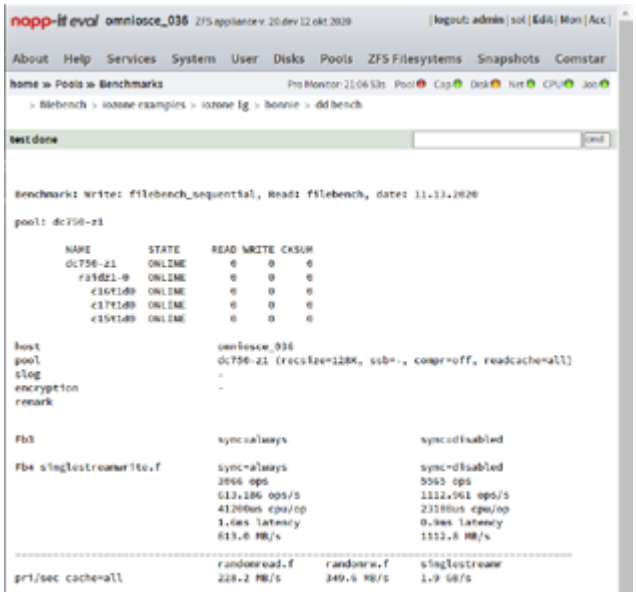
DC 750 Pool: 151034, enc on (aes-256ccm)



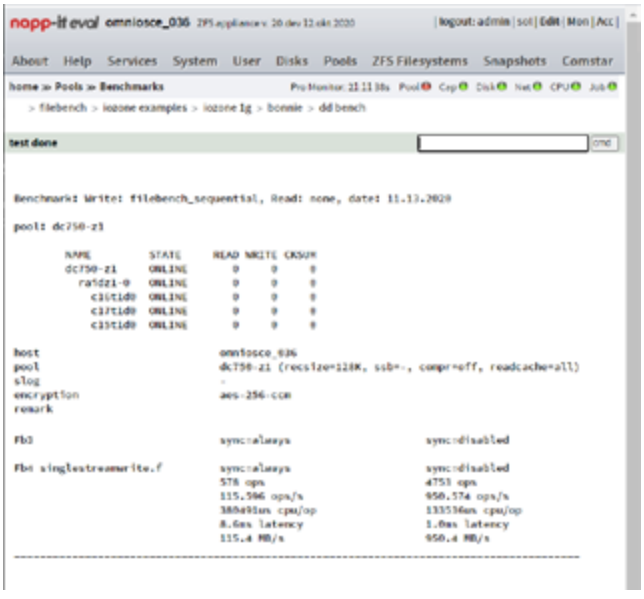
I startet the write benchmark multiple times:
sync was between 359 and 570 MB/s
non-sync between 1320 and 1490 MB/s

sync was between 52 and 56 MB/s
non-sync was between 1086 and 1093 MB/s

DC 750 Pool: 151036, enc off



DC 750 Pool: 151036, enc on (aes-256ccm)



I startet the write benchmark multiple times:
sync was between 430 and 613 MB/s
non-sync between 1112 and 1388 MB/s

sync was between 112 and 129 MB/s
non-sync was between 950 and 954 MB/s

3.0 Disk pool, Epyc on OmniOS 151036

I had to switch mpio off in menu Disk > Details > SAS, test date was Dec 03, 2020

Disk Pool: without slog, enc off

```
napp-it pro omniosce_036

About  Help  Services  System  User  Disks  Pools  ZFS Filesystems  Snapshots  Comstar

home » Pools » Benchmarks

test done

Benchmark: Write: filebench_sequential, Read: filebench, date: 08.19.2020

pool: hd-z1

        NAME                       STATE     READ WRITE CKSUM
        hd-z1                      ONLINE       0     0     0
          raidz1-0                 ONLINE       0     0     0
            c27t5000CCA08BE3CE1Cd0 ONLINE       0     0     0
            c29t5000CCA08EC6BE6Ad0 ONLINE       0     0     0
            c31t5000CCA08EC6D691d0 ONLINE       0     0     0
            c28t5000CCA08EC6E197d0 ONLINE       0     0     0
            c30t5000CCA08EC6E2F4d0 ONLINE       0     0     0
            c25t5000CCA2686911ADd0 ONLINE       0     0     0
            c26t5000CCA2621A3A91d0 ONLINE       0     0     0

host                    omniosce_036
pool                    hd-z1 (recsize=128K, ssb=-, compr=off, readcache=all)
slog                    -
encryption              -
remark

Fb3                     sync=always                 sync=disabled

Fb4 singlestreamwrite.f sync=always                 sync=disabled
                        258 ops                     5947 ops
                        51.598 ops/s                1189.365 ops/s
                        72281us cpu/op              27451us cpu/op
                        19.3ms latency              0.8ms latency
                        51.6 MB/s                   1189.2 MB/s

                        randomread.f    randomrw.f      singlestreamr
pri/sec cache=all       434.8 MB/s      572.2 MB/s      4.1 GB/s
```

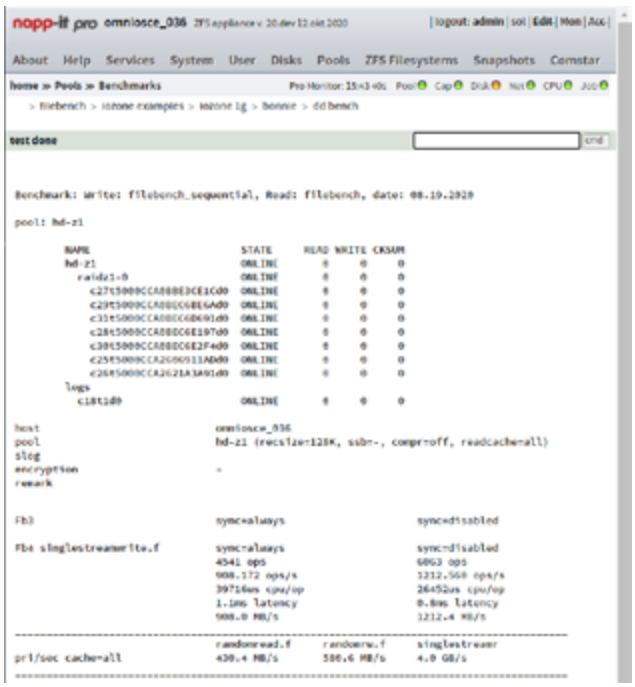
without slog, enc on (aes-256ccm)

```
napp-it pro omniosce_036

About  Help  Services  System  User  Disks  Pools  ZFS Filesystems  Snapshots  Comstar

home » Pools » Benchmarks

test done

Benchmark: Write: filebench_sequential, Read: filebench, date: 08.19.2020

pool: hd-z1

        NAME                       STATE     READ WRITE CKSUM
        hd-z1                      ONLINE       0     0     0
          raidz1-0                 ONLINE       0     0     0
            c27t5000CCA08BE3CE1Cd0 ONLINE       0     0     0
            c29t5000CCA08EC6BE6Ad0 ONLINE       0     0     0
            c31t5000CCA08EC6D691d0 ONLINE       0     0     0
            c28t5000CCA08EC6E197d0 ONLINE       0     0     0
            c30t5000CCA08EC6E2F4d0 ONLINE       0     0     0
            c25t5000CCA2686911ADd0 ONLINE       0     0     0
            c26t5000CCA2621A3A91d0 ONLINE       0     0     0

host                    omniosce_036
pool                    hd-z1 (recsize=128K, ssb=-, compr=off, readcache=all)
slog                    -
encryption              aes-256-ccm
remark

Fb3                     sync=always                 sync=disabled

Fb4 singlestreamwrite.f sync=always                 sync=disabled
                        185 ops                     6043 ops
                        36.999 ops/s                1208.563 ops/s
                        289290us cpu/op             133165us cpu/op
                        26.9ms latency              0.8ms latency
                        36.8 MB/s                   1208.4 MB/s

                        randomread.f    randomrw.f      singlestreamr
pri/sec cache=all       148.8 MB/s      268.4 MB/s      2.1 GB/s
```

I startet the write benchmark multiple times:
sync was between 51 and 55 MB/s
non-sync between 1178 and 1197 MB/s

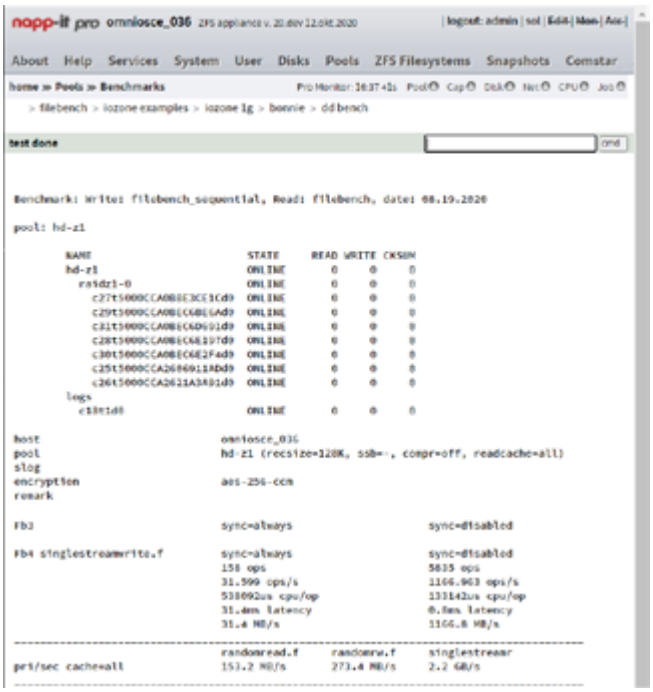
sync was between 36 and 40 MB/s
non-sync was between 1203 and 1230 MB/s

Disk Pool: with Optane 900 slog, enc off



```
napp-it pro omniosce_036

About  Help  Services  System  User  Disks  Pools  ZFS Filesystems  Snapshots  Comstar

home » Pools » Benchmarks

test done

Benchmark: Write: filebench_sequential, Read: filebench, date: 08.19.2020

pool: hd-z1

        NAME                       STATE     READ WRITE CKSUM
        hd-z1                      ONLINE       0     0     0
          raidz1-0                 ONLINE       0     0     0
            c27t5000CCA08BE3CE1Cd0 ONLINE       0     0     0
            c29t5000CCA08EC6BE6Ad0 ONLINE       0     0     0
            c31t5000CCA08EC6D691d0 ONLINE       0     0     0
            c28t5000CCA08EC6E197d0 ONLINE       0     0     0
            c30t5000CCA08EC6E2F4d0 ONLINE       0     0     0
            c25t5000CCA2686911ADd0 ONLINE       0     0     0
            c26t5000CCA2621A3A91d0 ONLINE       0     0     0
        logs
          c18t1d0                  ONLINE       0     0     0

host                    omniosce_036
pool                    hd-z1 (recsize=128K, ssb=-, compr=off, readcache=all)
slog                    -
encryption              -
remark

Fb3                     sync=always                 sync=disabled

Fb4 singlestreamwrite.f sync=always                 sync=disabled
                        4541 ops                    6063 ops
                        908.172 ops/s               1212.560 ops/s
                        39716us cpu/op              26452us cpu/op
                        1.1ms latency               0.8ms latency
                        908.0 MB/s                  1212.4 MB/s

                        randomread.f    randomrw.f      singlestreamr
pri/sec cache=all       430.4 MB/s      586.6 MB/s      4.0 GB/s
```

with Optane 900 slog, enc on (aes-256ccm)



```
napp-it pro omniosce_036

About  Help  Services  System  User  Disks  Pools  ZFS Filesystems  Snapshots  Comstar

home » Pools » Benchmarks

test done

Benchmark: Write: filebench_sequential, Read: filebench, date: 08.19.2020

pool: hd-z1

        NAME                       STATE     READ WRITE CKSUM
        hd-z1                      ONLINE       0     0     0
          raidz1-0                 ONLINE       0     0     0
            c27t5000CCA08BE3CE1Cd0 ONLINE       0     0     0
            c29t5000CCA08EC6BE6Ad0 ONLINE       0     0     0
            c31t5000CCA08EC6D691d0 ONLINE       0     0     0
            c28t5000CCA08EC6E197d0 ONLINE       0     0     0
            c30t5000CCA08EC6E2F4d0 ONLINE       0     0     0
            c25t5000CCA2686911ADd0 ONLINE       0     0     0
            c26t5000CCA2621A3A91d0 ONLINE       0     0     0
        logs
          c18t1d0                  ONLINE       0     0     0

host                    omniosce_036
pool                    hd-z1 (recsize=128K, ssb=-, compr=off, readcache=all)
slog                    -
encryption              aes-256-ccm
remark

Fb3                     sync=always                 sync=disabled

Fb4 singlestreamwrite.f sync=always                 sync=disabled
                        158 ops                     5835 ops
                        31.599 ops/s                1166.963 ops/s
                        538092us cpu/op             133142us cpu/op
                        31.4ms latency              0.8ms latency
                        31.4 MB/s                   1166.8 MB/s

                        randomread.f    randomrw.f      singlestreamr
pri/sec cache=all       153.2 MB/s      273.4 MB/s      2.2 GB/s
```

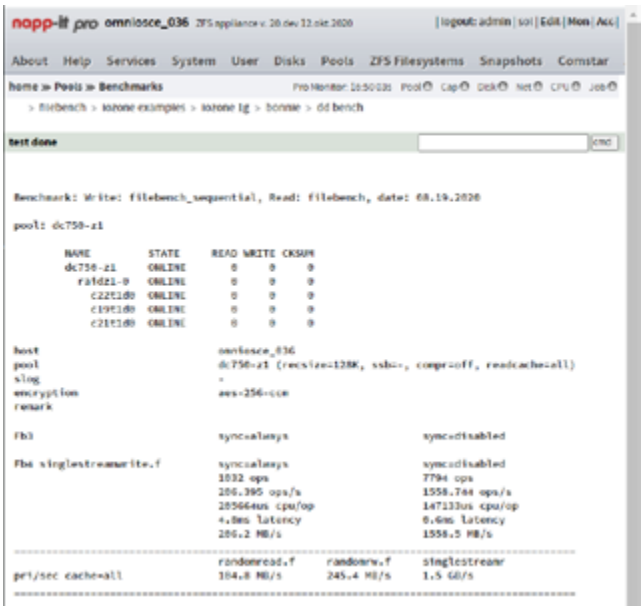
I startet the write benchmark multiple times:
sync was between 908 MB/s and 956 MB/s
non-sync between 1151 MB/s and 1212 MB/s

sync was between 31 MB/s and 32 MB/s and
non-sync was between 1166 MB/s and 1213 MB/s

3.1: Traditional Flash NVMe (Intel DC750) pool, Epyc on OmniOS 151036

DC 750 Pool: 151036, enc off

DC 750 Pool: 151036, enc on (aes-256ccm)



I startet the write benchmark multiple times:
sync was between 943 and 1134 MB/s
non-sync between 1591 and 1877 MB/s

sync was between 201 and 206 MB/s
non-sync was between 1541 and 1592MB/s

3.2: Optane 900 pool, Epyc on OmniOS 151036

Optane Pool: 151036, enc off

Optane Pool: 151036, enc on (aes-256ccm)

I startet the write benchmark multiple times:
sync was between 1445 and 1471 MB/s
non-sync between 3339 and 3707MB/s

sync was between 202 and 207 MB/s
non-sync was between 1720 -1730 MB/s

4.0 Virtualized NAS, ESXi 7.0u1, 8Core, 64GB RAM, SAS HBA and NVMe pass-through

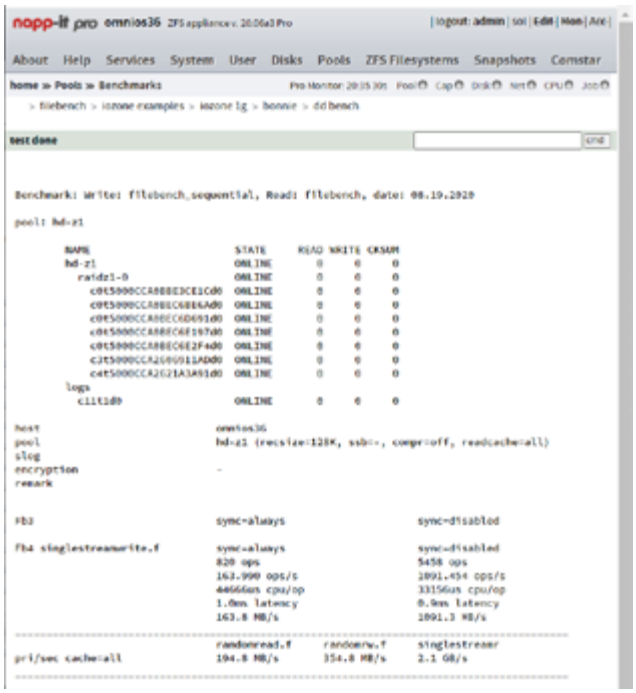
Disk Pool: without slog, enc off

without slog, enc on (aes-256ccm)

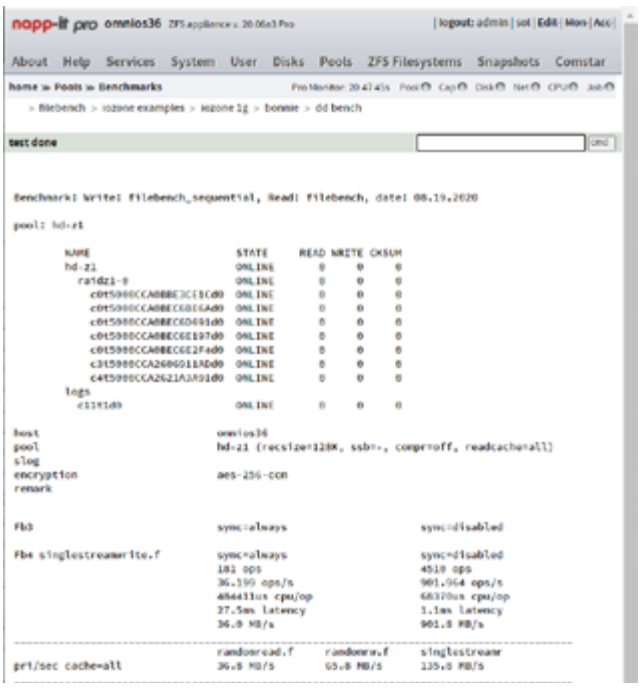
I startet the write benchmark multiple times:
sync was between 47 and 52 MB/s
non-sync between 1178 and 1225 MB/s

sync was between 37 and 40 MB/s
non-sync was between 901 and 905 MB/s

Disk Pool: with Optane 900 slog, enc off

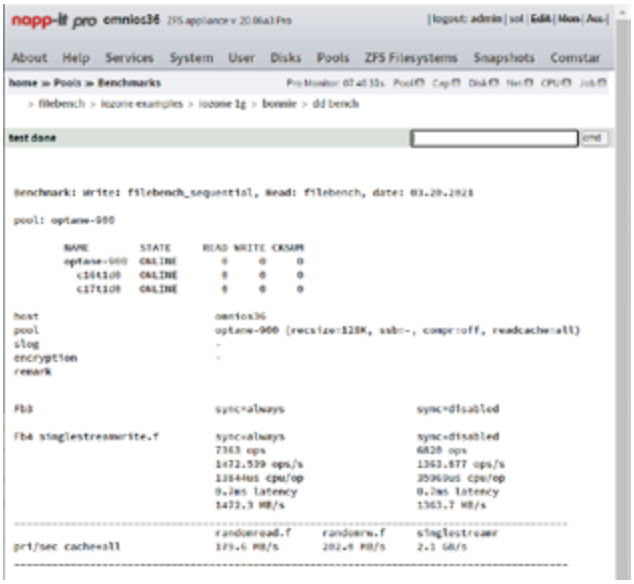


with Optane 900 slog, enc on (aes-256ccm)



I startet the write benchmark multiple times:
sync was between 163 and 281 MB/s
non-sync between 1091 and 1176 MB/s

sync was between 33 and 36 MB/s and
non-sync was between 901 and 902 MB/s



Needed steps for Optane 900 or DC750 and Passthrough/ ESXi

- ssh to ESXi
- edit /etc/vmware/passthru.map
- add following lines at the end of the file:

Intel Optane 900P pci-e device
8086 2700 d3d0 false
Intel DC750 pci-e device
8086 370D d3d0 false
- restart hypervisor and add the pci device to OmniOS

check device id ex 2700 in ESXi > Management >PCI devices
Either try id ex 2700 or sub id ex 3900 (Optane 900)
on problems with passhrough

more device id ex Optane 900 2,5", Intel DC 750 etc, see
https://pci-ids.ucw.cz/read/PC/8086

more
For some NVMe there is another hint on problems:
add the following to the .vmx file of a VM:
pciPassthru0.msiEnabled = „FALSE"

https://tinkertry.com/search?s=passhrough#gsc.tab=0&gsc.q=passthrough&gsc.sort=
https://kb.vmware.com/s/article/2142307