napp-it

napp-*it*

ZFS Storage Server

Some performance tests
to decide about some storage setups
on different use cases

- How important is RAM
- Differences between HD, SSD, NVMe
- What about Slog
- AiO or barebone setup

Content:

4. All-In-One setup 2vcpu, 24GB Ram
   - pass-through vs vdisk
   - the Optane as vdisk Slog
   - Perfomance vs RAM (24/16/8/4/2 GB)

5. iSCSI vs NFS vs SMB
   Oracle Solaris ZFS v.37 vs OmniOS OpenZFS
   Oracle Solaris 11.4 with ZFS v.43

5.3. OmniOS and a pool from 4 x 900P
5.4. Solaris 11.4 and a pool from 4 x 900P
5.4.1 Solaris 11.4 and a pool from 2 x 800P
vs a pool from 2 x 900P

6. Sata vs LSI SAS 12G vs ATTO 12G

7. U.2 NVMe HBA
   Cables, Disk enclosures and backplanes

1.0 TL;DR - „too long; didn't read"

This benchmark sequence was intended to answer some basic questions about disks, SSDs, NVMe/Optane, the effect of RAM and the difference between native ZFS in Oracle Solaris v.37 vs OpenZFS in the free Solaris fork OmniOS. If you want to build a ZFS system this may help to optimize.

1. The most important factor is RAM

Whenever your workload can be mainly processed within your RAM, even a slow HD pool is nearly as fast as an ultimate Optane pool. Calculate around 2 GB for your OS. Then add the wanted rambased write cache (OmniOS default: 10% of RAM, max 4GB) and add the RAM that you want as readcache.

If your workload exceed your RAM capabilities or cannot use the RAM like with sync-write performance can dramatically go down. If you cannot add more RAM (ex AiO), add an L2Arc (5x - max 10x RAM). In general you can say, with less RAM you need much faster disks for a similar overall performance compared to more RAM.
In a homeserver/ mediaserver/ SoHo filer environment with a few users and 1G networks 4-8GB RAM is ok.
In a multiuser environment or with large amount of random data (VMs, larger databases) use 16-32GB RAM. If you have a faster network (10/40G) add more RAM and use 32-64G and more.

2. Even a pure HD pool can be nearly as fast as a NVMe pool.

In my tests I used a pool from 4 x HGST HE8 disks with a combined raw sequential read/write performance of more than 1000 MB/s. As long as you can process your workload mainly from RAM, it is tremendous fast. The huge fallback when using sync-write can be nearly eleminated by a fast Optane Slog like the 900P. Such a combination can be nearly as fast as a pure SSD pool at a fraction of the cost with higher capacity.
Even an SMB filer with a secure write behaviour (sync-write=always) is now possible as a 4 x HGST HE8 pool (Raid-0) and an Optane 900P Slog offered around 500-700 MB/s (needed for 10G networks) on OmniOS. Solaris with native ZFS was even faster.

3. Critical workloads (many user, many random data, VM storage, databases)

Prefer SSD only pools. A dedicated Slog is not needed but prefer SSDs with powerloss protection when you want sync write. For HD pools and sync write, add an Slog, prefer Optane 32G or 900P, add much RAM.

4. Ultra critical or performance sensitive workloads

Intel Optane is unbeaten! Compared to a fast NVMe like Intel P3700 it reduces latency from 30us down to 10us and increases iops from 80k to 500k. While on most workloads you will not see much difference as most workloads are more sequential or the RAM takes the load some are different. If you really need small random read write performance that hits disks you do not have much alternatives.

Additionally Optane is more organized like RAM. This means no trim or garbage collection or erase prior write like on Flash is needed. Even a concurrent read/write workload does not affect performance in the same way as it was on Flash. For mission critical workloads use the Enterprise Optane 4800x with guaranteed powerloss protection. While there is no cache or capacitor on any, Intel guarantees this not for the cheaper Optane. Unsure if this is related to performance, firmware or marketing.

5. Oracle Solaris with native ZFS v.37 beats OpenZFS

OmniOS, a free Solartis fork is known to be one of the fastest OpenZFS systems but native ZFS v.37 on Solaris plays in a different ligue when you check pool performance as well when you check services like SMB.

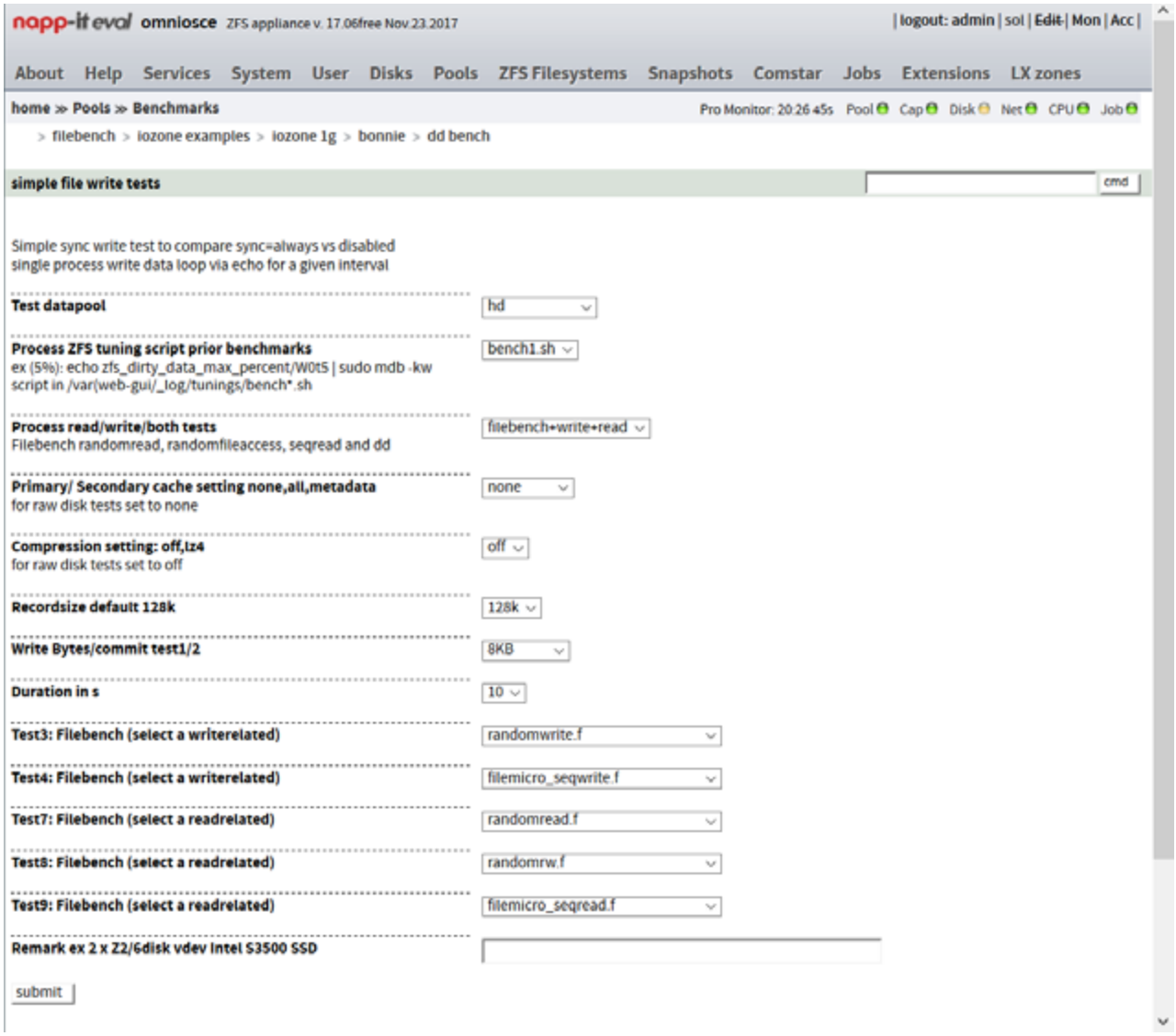
What I have found is that Solaris starts writes very fast and stalls then for a short time. OmniOS with its write throtteling seems not as fast regarding overall write performance but can guarantee a constant latency.

RAM efficiency regarding caching seems to be the major advantage on Solaris and even with low RAM for caching sync write performance even on harddisks is top.

1.1 The new Intel Optane NVMe 800P, 900P, 905P, 4801X

The new Intel Optane 3D Xpoint NVMe especially the new affordable 800P/900P is a game-changing technology . Unlike former Flash based NVMe they do not require Trim or Garbage Collection or erase prior write cycles. They are more adressed like RAM. Their write latency is 10us with 500k write iops. The Intel P3700, one of the best Flash based NVMe has 20us latency and around 80k write iops. Especially with low Queue depth they are much faster than the former NVMe what makes them a perfect solution for Ultra High performance pools or as an Slog/L2Arc device for slower disk based pools or SSD pools if they lack powerloss protection (the small Optane cache models lack powerloss protection what makes them a bad choice for an Slog. To give you an impression, I have done some write performance tests on Solaris and OmniOS 151024/26 with sync vs async. Unlike the Enterprise class Optane 4801X the 800/900P lacks guaranteed powerloss protection but for many use cases the risk when using as Slog seems acceptable. For production use always prefer the 4801X (100GB or 200GB) !!

1.2. Benchmark are done via napp-it menu Pools > Benchmarks



The benchmarks with different settings, RAM, pools or devices:

- random write performance with sync=always vs sync=disabled
- sequential write performance with sync=always vs sync=disabled

- random read performance
- mixed load (random read/write)
- sequential read

1.3 Result: Barebone vs AiO

I have made several test series over more than a week with different settings, configs and environments to decide how to build optimized ZFS systems. After some barebone tests I added tests with a virtualized NAS under ESXi that gaves me the best flexibility to answer question about use cases, needed or wanted RAM, disk types and settings. You can compare the detailled benchmarks to answer more question but my main findings are:

Performance differences between barebone 8Core/32GB RAM (3.2) and AiO (2 vcpu, 24GB RAM,)

Barebone setup, allcache write	**sync_random**	**async_random**	**sync_stream**	**async_stream**
4xHD, Optane Slog P900	16.0 MB/s	77.8 MB/s	479.6 MB/s	1109.9 MB/s
4xSSD Pool, no Slog	16.0 MB/s	55.6 MB/s	370 MB/s	951.6 MB/s
Single Optane Pool	13.2 MB/s	127.8 MB/s	682.8 MB/s	1611.0 MB/s

AiO setup, allcache write	**sync_random**	**async_random**	**sync_stream**	**async_stream**
4xHD, Optane Slog P900/20G vdisk	30.4 MB/s	64.4 MB/s	634.7 MB/s	1015.8 MB/s
4xSSD Pool, no Slog	30.4 MB/s	51.4 MB/s	656.6 MB/s	866.8 MB/s
Single Optane vdisk Pool 200G	25.6 MB/s	69.4 GB/s	767.8 MB/s	1249.5 MB/s

Barebone setup, allcache read	**random_read**	**random_r/w**	**stream_read**
4xHD Pool,	225.6 MB/s	252.6 MB/s	2.6 GB/s
4xSSD Pool,	209.0 MB/s	266.2 MB/s	2.5 GB/s
Single Optane	243.2 MB/s	301.6 MB/s	2.6 GB/s

AiO setup, allcache read	**random_read**	**random_r/w**	**stream_read**
4xHD Pool,	179.8 MB/s	184.1 MB/s	2.5 GB/s
4xSSD Pool	176.0 MB/s	204.8 MB/s	2.5 GB/s
Single Optane vdisk	225.6 MB/s	193.6 MB/s	2.6 MB/s

Conclusions

- AiO is nearly as fast as barebone using same or similar RAM
- Optane as vdisk is nearly as fast as barebone on reads
- Optane is faster with writes on vdisk, so I asume some caching in ESXi

As Optane is uncritical against concurrent read/write load
a suggested ESXi AiO setup is

- boot from USB or other bootdevice
- place your storage VM and Slog/L2Arc onto the Optane datastore
- use an Slog with HD pools or slow SSD pools
- Slog not needed with faster multi-SSD pools

- optionally use 2 Optane as vdisk for a high performance ZFS mirror on vdisk

Remains a small unsecurity about powerloss protection of Optane and
using Optane over the ESXi NVMe driver especially as the random sync write
values under ESXi are better than on barebone (ESXi cache?)

If you wonder why I use vdisks with Optane:
Pass-through of Optane is currently not working and the above results make vdisk
with Optane as datastore + slog + l2arc promising.

1.3 Result: RAM and pooltype, use cases

If you look at the AiO benchmarks with different RAM and disk type settings there are some clear relations between RAM as the main factor and disktype, random and sequential loads.

AiO 4GB setup, allcache write	**sync_random**	**async_random**	**sync_stream**	**async_stream**
4xHD, Optane Slog P900/20G vdisk	1.8 MB/a	1.4 MB/s	455.7 MB/s	665.6 MB/s
4xSSD Pool, Slog P900/20G vdisk	12.2 MB/s	11.4 MB/s	584 MB/s	594 MB/s
Single Optane vdisk Pool 200G	14.0 MB/s	39.4 MB/s	562.0 MB/s	755.6 MB/s

AiO 4GB setup, allcache read	**random_read**	**random_r/w**	**stream_read**
4xHD Pool,Slog P900/20G vdisk	1.4 MB/s	3.0 MB/s	604.6 MB/s
4xSSD Pool	17.8 MB/s	20.8 MB/s	1.2 GB/s
Single Optane	48.6 MB/s	57.0 MB/s	1.3 GB/s

AiO 8GB setup, allcache write	**sync_random**	**async_random**	**sync_stream**	**async_stream**
4xHD, Optane Slog P900/20G vdisk	31.0 MB/s	40.8 MB7s	126.2 MB/s	776 MB/s
4xSSD Pool, Slog P900/20G vdisk	31.0 MB/s	42.6 MB/s	612 MB/s	635 MB7s
Single Optane vdisk Pool 200G	27.2 MB/s	55.6 MB/s	659.2 MB/s	961 MB/s

AiO 8GB setup, allcache read	**random_read**	**random_r/w**	**stream_read**
4xHD Pool	234 MB/s	224,2 MB/s	2.1 GB/s
4xSSD Pool	185 MB/s	170.0 MB/s	2.4 GB/s
Single Optane	186.2 MB/s	153.0 MB/s	2.6 GB/s

AiO 24GB setup, allcache write	**sync_random**	**async_random**	**sync_stream**	**async_stream**
4xHD, Optane Slog P900/20G vdisk	30.4 MB/s	64.4 MB/s	634 MB/s	1015 MB/s
4xSSD Pool, Slog P900/20G vdisk	12.4 MB/s	48.4 MB/s	343 MB/s	869.8 MB/s
Single Optane vdisk Pool 200G	25.6 MB/s	69.4 MB/s	767.8 MB/s	1249 MB/s

AiO 24GB setup, allcache read	**random_read**	**random_r/w**	**stream_read**
4xHD Pool,	179.8 MB/s	184.1 MB/s	2.5 GB/s
4xSSD Pool	184 MB/s	245.4 MB/s	2.5 GB/s
Single Optane	225.6 MB/s	193.6 MB/s	2.6 GB/s

Conclusions:

Solarish works with 2 GB RAM but performance is then pure iops limited disk performance.
This is really bad with disks and good with Optane NVMe
For a small filer, 4 GB is ok for a mainly sequential read/write workload and few users.
If you want good random read/write performance you must add more RAM or use faster disks.

In general you find that even with slow disks read/write performance is very good as long as you have enough RAM to process your random read/ write requests. The value with lower RAM shows clearly that in the moment you must go to disk for reads (and writes must read metadata prior write as well and need the rambase writecache to improve small random writes) If you really want random read/write performance, add as much RAM as possible.

Pool Layout:
Disks are really bad with sync writes but can perform well with many vdev and an Slog.
A pool from slower SSDs gain much from a very fast Optane Slog. A pool from faster NVMe does not require an Slog. A single Optane is faster than a 4 vdev SSD pool or a 4 vdev HD pool with Optane Slog. On an Optane only Pool the 900P is twice as fast than the 800P for sync writes.

1.3 Result: Slog

If you want sync write on a disk-pool, you must use an Slog. Even a good SSD can help. A Dram based ZeusRAM can heavily improve write performance but Intel Optane as Slog is the Best.

If you have a pool from some faster SSD or NVMe vdevs, a dedicated Slog does not help unless its the Optane. The bigger the performance difference between pooldisks and Slog the more helpful is the Slog. If an Slog is as fast or only slightly faster than a pooldisk it is not helpful.

2. Barebone Setup (SM X11SPH-nCTF, Xeon Silver 8Core 4110, 32G RAM)

2.1 Single disk values, effect of Arc rambased read cache

Single HGST HD HE 8 nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	single-hd (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	208 ops		18 ops
	41.597 ops/s		3.600 ops/s
	7000us cpu/op		198872us cpu/op
	23.9ms latency		230.8ms latency
	0.2 MB/s		0.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	175 ops		3459 ops
	34.997 ops/s		691.756 ops/s
	8384us cpu/op		1407us cpu/op
	28.3ms latency		1.4ms latency
	34.8 MB/s		691.6 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.0 MB/s	0.0 MB/s	1.6 MB/s

hostname	omniosce Memory size: 32429 Megabytes		
pool	single-hd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	525 ops		12438 ops
	104.989 ops/s		2487.459 ops/s
	2172us cpu/op		416us cpu/op
	9.5ms latency		0.4ms latency
	0.8 MB/s		19.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	170 ops		3463 ops
	33.997 ops/s		692.550 ops/s
	8925us cpu/op		1529us cpu/op
	29.3ms latency		1.4ms latency
	33.8 MB/s		692.3 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	267.8 MB/s	252.0 MB/s	2.7 GB/s

2.2 Dual disk/ basic vdev values, effect of Arc rambased read cache

Dual basic HGST HD HE 8 nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	dual-hd (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	285 ops		8 ops
	56.995 ops/s		1.600 ops/s
	3911us cpu/op		359979us cpu/op
	17.5ms latency		621.9ms latency
	0.4 MB/s		0.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	186 ops		4102 ops
	37.197 ops/s		820.363 ops/s
	11112us cpu/op		1792us cpu/op
	26.6ms latency		1.2ms latency
	37.0 MB/s		820.2 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.0 MB/s	0.0 MB/s	7.0 MB/s

---

hostname	omniosce Memory size: 32429 Megabytes		
pool	dual-hd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	894 ops		24213 ops
	178.786 ops/s		4842.251 ops/s
	1514us cpu/op		244us cpu/op
	5.6ms latency		0.2ms latency
	1.2 MB/s		37.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	189 ops		4194 ops
	37.796 ops/s		838.744 ops/s
	7338us cpu/op		1576us cpu/op
	26.2ms latency		1.2ms latency
	37.6 MB/s		838.5 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	190.0 MB/s	272.0 MB/s	2.5 GB/s

2.3 Quad disk/ basic vdev values, effect of Arc rambased read cache

4 x basic HGST HD HE 8 nocache vs allcache

```
hostname                        omniosce  Memory size: 32429 Megabytes
pool                            quad-hd (recsize=128k, compr=off, readcache=none)
slog                            -
remark


Fb3 randomwrite.f               sync=always                   sync=disabled
                                177 ops                       391 ops
                                35.396 ops/s                  78.194 ops/s
                                7552us cpu/op                 6682us cpu/op
                                27.9ms latency                12.7ms latency
                                0.2 MB/s                      0.6 MB/s

Fb4 singlestreamwrite.f         sync=always                   sync=disabled
                                207 ops                       5428 ops
                                41.396 ops/s                  1085.533 ops/s
                                7713us cpu/op                 1901us cpu/op
                                24.0ms latency                0.9ms latency
                                41.2 MB/s                     1085.3 MB/s
```

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.4 MB/s	1.2 MB/s	230.4 MB/s

```
hostname                        omniosce  Memory size: 32429 Megabytes
pool                            quad-hd (recsize=128k, compr=off, readcache=all)
slog                            -
remark


Fb3 randomwrite.f               sync=always                   sync=disabled
                                1067 ops                      45348 ops
                                213.381 ops/s                 9069.205 ops/s
                                1884us cpu/op                 212us cpu/op
                                4.7ms latency                 0.1ms latency
                                1.6 MB/s                      70.8 MB/s

Fb4 singlestreamwrite.f         sync=always                   sync=disabled
                                179 ops                       5419 ops
                                35.796 ops/s                  1083.750 ops/s
                                7972us cpu/op                 2009us cpu/op
                                27.8ms latency                0.9ms latency
                                35.6 MB/s                     1083.5 MB/s
```

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	216.8 MB/s	271.4 MB/s	2.6 GB/s

2.4 Single SSD values, effect of Arc rambased read cache

Single Intel DC 3510-120 nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	singledc3510 (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1740 ops		331 ops
	347.983 ops/s		66.196 ops/s
	1775us cpu/op		8382us cpu/op
	2.9ms latency		15.1ms latency
	2.6 MB/s		0.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	682 ops		3324 ops
	136.392 ops/s		664.754 ops/s
	5203us cpu/op		1283us cpu/op
	7.3ms latency		1.5ms latency
	136.2 MB/s		664.6 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.6 MB/s	1.6 MB/s	36.4 MB/s

hostname	omniosce Memory size: 32429 Megabytes		
pool	singledc3510 (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	4445 ops		4772 ops
	888.945 ops/s		954.315 ops/s
	869us cpu/op		425us cpu/op
	1.1ms latency		0.6ms latency
	6.8 MB/s		7.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	675 ops		3313 ops
	134.992 ops/s		662.565 ops/s
	5214us cpu/op		1264us cpu/op
	7.3ms latency		1.5ms latency
	134.8 MB/s		662.4 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	264.2 MB/s	282.4 MB/s	2.7 GB/s

2.5 Dual SSD values, effect of Arc rambased read cache

Dual basic Intel DC 3510-120 nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	dual-dc3510 (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1842 ops		562 ops
	368.379 ops/s		112.393 ops/s
	1882us cpu/op		7243us cpu/op
	2.7ms latency		8.9ms latency
	2.8 MB/s		0.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1268 ops		3797 ops
	253.589 ops/s		759.362 ops/s
	4994us cpu/op		1786us cpu/op
	3.9ms latency		1.3ms latency
	253.4 MB/s		759.2 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	1.0 MB/s	2.6 MB/s	61.2 MB/s

---

hostname	omniosce Memory size: 32429 Megabytes		
pool	dual-dc3510 (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	6182 ops		18388 ops
	1236.333 ops/s		3677.364 ops/s
	861us cpu/op		296us cpu/op
	0.8ms latency		0.3ms latency
	9.6 MB/s		28.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1265 ops		3812 ops
	252.988 ops/s		762.364 ops/s
	4942us cpu/op		1645us cpu/op
	3.9ms latency		1.3ms latency
	252.8 MB/s		762.2 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	231.8 MB/s	262.8 MB/s	2.5 GB/s

2.6 Quad SSD values, effect of Arc rambased read cache

Quad basic Intel DC 3510-120 nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes	
pool	quad-dc3510 (recsize=128k, compr=off, readcache=none)	
slog	-	
remark		
Fb3 randomwrite.f	sync=always	sync=disabled
	1921 ops	1582 ops
	384.170 ops/s	316.379 ops/s
	1699us cpu/op	2567us cpu/op
	2.6ms latency	3.1ms latency
	3.0 MB/s	2.4 MB/s
Fb4 singlestreamwrite.f	sync=always	sync=disabled
	2007 ops	4813 ops
	401.374 ops/s	962.537 ops/s
	4359us cpu/op	2039us cpu/op
	2.5ms latency	1.0ms latency
	401.2 MB/s	962.3 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	3.0 MB/s	7.2 MB/s	217.2 MB/s

hostname	omniosce Memory size: 32429 Megabytes	
pool	quad-dc3510 (recsize=128k, compr=off, readcache=all)	
slog	-	
remark		
Fb3 randomwrite.f	sync=always	sync=disabled
	8432 ops	34067 ops
	1686.347 ops/s	6813.093 ops/s
	699us cpu/op	214us cpu/op
	0.6ms latency	0.1ms latency
	13.0 MB/s	53.2 MB/s
Fb4 singlestreamwrite.f	sync=always	sync=disabled
	1980 ops	4918 ops
	395.978 ops/s	983.554 ops/s
	4372us cpu/op	1794us cpu/op
	2.5ms latency	1.0ms latency
	395.8 MB/s	983.4 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	216.6 MB/s	249.6 MB/s	2.5 GB/s

2.7 Single SSD values, effect of Arc rambased read cache

Single Intel DC 3610-480 (Sata) nocache vs allcache

```
hostname                        omniosce  Memory size: 32429 Megabytes
pool                            singledc-3610 (recsize=128k, compr=off, readcache=none)
slog                            -
remark


Fb3 randomwrite.f               sync=always                   sync=disabled
                                2214 ops                      1273 ops
                                442.775 ops/s                 254.587 ops/s
                                1921us cpu/op                 3332us cpu/op
                                2.2ms latency                 3.9ms latency
                                3.4 MB/s                      1.8 MB/s

Fb4 singlestreamwrite.f         sync=always                   sync=disabled
                                1585 ops                      4335 ops
                                316.986 ops/s                 866.959 ops/s
                                4582us cpu/op                 2168us cpu/op
                                3.1ms latency                 1.1ms latency
                                316.8 MB/s                    866.8 MB/s
_____________________________________________________________________________________

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=none              2.6 MB/s         5.0 MB/s       137.6 MB/s
_____________________________________________________________________________________


hostname                        omniosce  Memory size: 32429 Megabytes
pool                            singledc-3610 (recsize=128k, compr=off, readcache=all)
slog                            -
remark


Fb3 randomwrite.f               sync=always                   sync=disabled
                                6466 ops                      28669 ops
                                1293.145 ops/s                5733.464 ops/s
                                791us cpu/op                  284us cpu/op
                                0.8ms latency                 0.2ms latency
                                10.0 MB/s                     44.6 MB/s

Fb4 singlestreamwrite.f         sync=always                   sync=disabled
                                1581 ops                      4360 ops
                                316.186 ops/s                 871.954 ops/s
                                4631us cpu/op                 1724us cpu/op
                                3.1ms latency                 1.1ms latency
                                316.0 MB/s                    871.8 MB/s
_____________________________________________________________________________________

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=all               225.4 MB/s       274.2 MB/s     2.6 GB/s
_____________________________________________________________________________________
```

2.8 Single NVMe values, effect of Arc rambased read cache

Single Intel P 3610-400 (NVMe) nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	singlep-360 (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1878 ops		1557 ops
	375.581 ops/s		311.380 ops/s
	2054us cpu/op		3217us cpu/op
	2.6ms latency		3.2ms latency
	2.8 MB/s		2.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2196 ops		4465 ops
	439.181 ops/s		892.953 ops/s
	4361us cpu/op		1941us cpu/op
	2.2ms latency		1.1ms latency
	439.0 MB/s		892.8 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	3.4 MB/s	7.0 MB/s	108.4 MB/s

---

hostname	omniosce Memory size: 32429 Megabytes		
pool	singlep-360 (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	7030 ops		30324 ops
	1405.892 ops/s		6064.535 ops/s
	725us cpu/op		246us cpu/op
	0.7ms latency		0.2ms latency
	10.8 MB/s		47.2 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2175 ops		4425 ops
	434.981 ops/s		884.954 ops/s
	4439us cpu/op		1976us cpu/op
	2.3ms latency		1.1ms latency
	434.8 MB/s		884.8 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	201.6 MB/s	260.0 MB/s	2.6 GB/s

2.9 Single NVMe values, effect of Arc rambased read cache

Single Intel 32G Optane M.2 Cache Modul (NVMe) nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	optane32 (recsize=128k, compr=off, readcache=none)		
slog	-		
remark	32G M.2 Cache Optane		
Fb3 randomwrite.f	sync=always		sync=disabled
	2184 ops		5056 ops
	436.769 ops/s		1011.146 ops/s
	1944us cpu/op		1236us cpu/op
	2.1ms latency		1.0ms latency
	3.4 MB/s		7.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1069 ops		2018 ops
	213.791 ops/s		403.575 ops/s
	5055us cpu/op		2747us cpu/op
	4.6ms latency		2.2ms latency
	213.6 MB/s		403.4 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	9.6 MB/s	21.0 MB/s	498.2 MB/s

---

hostname	omniosce Memory size: 32429 Megabytes		
pool	optane32 (recsize=128k, compr=off, readcache=all)		
slog	-		
remark	32G M.2 Cache Optane		
Fb3 randomwrite.f	sync=always		sync=disabled
	7932 ops		11952 ops
	1586.312 ops/s		2390.274 ops/s
	734us cpu/op		391us cpu/op
	0.6ms latency		0.3ms latency
	12.2 MB/s		18.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1068 ops		2018 ops
	213.586 ops/s		403.582 ops/s
	5828us cpu/op		2747us cpu/op
	4.6ms latency		2.2ms latency
	213.4 MB/s		403.4 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	241.8 MB/s	155.0 MB/s	2.7 GB/s

2.10 Single NVMe values, effect of Arc rambased read cache

Single Intel 900P (NVMe) nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	single900p (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	3525 ops		6915 ops
	704.953 ops/s		1382.936 ops/s
	1638us cpu/op		1062us cpu/op
	1.4ms latency		0.7ms latency
	5.4 MB/s		10.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3408 ops		8487 ops
	681.576 ops/s		1697.348 ops/s
	3295us cpu/op		1789us cpu/op
	1.4ms latency		0.6ms latency
	681.4 MB/s		1697.1 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	10.8 MB/s	39.2 MB/s	791.2 MB/s

hostname	omniosce Memory size: 32429 Megabytes		
pool	single900p (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	8562 ops		81358 ops
	1712.326 ops/s		16270.999 ops/s
	671us cpu/op		196us cpu/op
	0.6ms latency		0.1ms latency
	13.2 MB/s		127.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3415 ops		8056 ops
	682.975 ops/s		1611.156 ops/s
	3438us cpu/op		3335us cpu/op
	1.4ms latency		0.6ms latency
	682.8 MB/s		1611.0 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	243.2 MB/s	301.6 MB/s	2.6 GB/s

2.11 Dual NVMe values, effect of Arc rambased read cache

2 x vdev Intel 900P (NVMe) nocache vs allcache

```
hostname                        omniosce  Memory size: 32429 Megabytes
pool                            dual-900p (recsize=128k, compr=off, readcache=none)
slog                            -
remark


Fb3 randomwrite.f               sync=always                   sync=disabled
                                2552 ops                      6891 ops
                                510.372 ops/s                 1378.139 ops/s
                                1790us cpu/op                 1102us cpu/op
                                1.8ms latency                 0.7ms latency
                                3.8 MB/s                      10.6 MB/s

Fb4 singlestreamwrite.f         sync=always                   sync=disabled
                                4428 ops                      10680 ops
                                885.578 ops/s                 2135.931 ops/s
                                2937us cpu/op                 1940us cpu/op
                                1.1ms latency                 0.5ms latency
                                885.4 MB/s                    2135.7 MB/s
```

```
read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=none              9.4 MB/s         37.2 MB/s      758.4 MB/s
```

```
hostname                        omniosce  Memory size: 32429 Megabytes
pool                            dual-900p (recsize=128k, compr=off, readcache=all)
slog                            -
remark


Fb3 randomwrite.f               sync=always                   sync=disabled
                                7232 ops                      102006 ops
                                1446.342 ops/s                20400.488 ops/s
                                705us cpu/op                  219us cpu/op
                                0.7ms latency                 0.0ms latency
                                11.2 MB/s                     159.2 MB/s

Fb4 singlestreamwrite.f         sync=always                   sync=disabled
                                4153 ops                      8431 ops
                                830.575 ops/s                 1686.118 ops/s
                                2916us cpu/op                 3150us cpu/op
                                1.2ms latency                 0.6ms latency
                                830.4 MB/s                    1685.9 MB/s
```

```
read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=all               248.2 MB/s       257.4 MB/s     2.7 GB/s
```

2.12 Quad NVMe values, effect of Arc rambased read cache

4 x vdev Intel 900P (NVMe) nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	quad-900p (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	3496 ops		6503 ops
	699.165 ops/s		1300.548 ops/s
	1705us cpu/op		1177us cpu/op
	1.3ms latency		0.8ms latency
	5.4 MB/s		10.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	4382 ops		11876 ops
	876.369 ops/s		2374.887 ops/s
	3075us cpu/op		2320us cpu/op
	1.1ms latency		0.4ms latency
	876.2 MB/s		2374.7 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	12.0 MB/s	39.4 MB/s	789.6 MB/s

---

hostname	omniosce Memory size: 32429 Megabytes		
pool	quad-900p (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	9193 ops		89720 ops
	1838.495 ops/s		17943.234 ops/s
	614us cpu/op		269us cpu/op
	0.5ms latency		0.1ms latency
	14.2 MB/s		140.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	4383 ops		10174 ops
	876.575 ops/s		2034.736 ops/s
	3216us cpu/op		2539us cpu/op
	1.1ms latency		0.5ms latency
	876.4 MB/s		2034.5 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	242.6 MB/s	253.4 MB/s	2.7 GB/s

2.13 Single Optane 800P vs Single 900P

1 x Intel Optane 800P-118 vs 1 x Intel 900P-280 (allcache, OmniOS 151026)

```
hostname                        omni26  Memory size: 32429 Megabytes
pool                            single-800p (recsize=128k, compr=off, readcache=all)
slog                            -
remark                          800p


Fb3 randomwrite.f               sync=always                     sync=disabled
                                8694 ops                        38890 ops
                                1738.718 ops/s                  7777.688 ops/s
                                702us cpu/op                    192us cpu/op
                                0.6ms latency                   0.1ms latency
                                13.4 MB/s                       60.6 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                1015 ops                        3450 ops
                                202.985 ops/s                   689.967 ops/s
                                5787us cpu/op                   2049us cpu/op
                                4.9ms latency                   1.4ms latency
                                202.8 MB/s                      689.8 MB/s
```

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	254.0 MB/s	318.4 MB/s	2.7 GB/s

```
hostname                        omni26  Memory size: 32429 Megabytes
pool                            single-900p (recsize=128k, compr=off, readcache=all)
slog                            -
remark                          800p


Fb3 randomwrite.f               sync=always                     sync=disabled
                                9201 ops                        95278 ops
                                1840.113 ops/s                  19051.392 ops/s
                                637us cpu/op                    186us cpu/op
                                0.5ms latency                   0.0ms latency
                                14.2 MB/s                       148.8 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                3373 ops                        9724 ops
                                674.573 ops/s                   1944.735 ops/s
                                3576us cpu/op                   2008us cpu/op
                                1.5ms latency                   0.5ms latency
                                674.4 MB/s                      1944.5 MB/s
```

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	262.8 MB/s	341.6 MB/s	2.9 GB/s

2.14 Dual Optane 800P vs Dual 900P

2 x Intel Optane 800P-118 vs 2 x Intel 900P-280 (Raid-0, allcache, OmniOS 151026)

hostname	omni26 Memory size: 32429 Megabytes		
pool	dual-800p (recsize=128k, compr=off, readcache=all)		
slog	-		
remark	dual 800P-118 in Raid-0		
Fb3 randomwrite.f	sync=always		sync=disabled
	8229 ops		53876 ops
	1645.696 ops/s		10774.661 ops/s
	794us cpu/op		246us cpu/op
	0.6ms latency		0.1ms latency
	12.8 MB/s		84.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1524 ops		5383 ops
	304.788 ops/s		1076.546 ops/s
	5674us cpu/op		2158us cpu/op
	3.2ms latency		0.9ms latency
	304.6 MB/s		1076.3 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	256.6 MB/s	306.8 MB/s	2.8 GB/s

hostname	omni26 Memory size: 32429 Megabytes		
pool	dual-900p (recsize=128k, compr=off, readcache=all)		
slog	-		
remark	dual 900P-118 in Raid-0		
Fb3 randomwrite.f	sync=always		sync=disabled
	9905 ops		105152 ops
	1980.931 ops/s		21029.654 ops/s
	643us cpu/op		199us cpu/op
	0.5ms latency		0.0ms latency
	15.4 MB/s		164.2 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	4122 ops		8548 ops
	824.372 ops/s		1708.648 ops/s
	3258us cpu/op		3578us cpu/op
	1.2ms latency		0.6ms latency
	824.2 MB/s		1708.4 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	260.8 MB/s	288.0 MB/s	2.9 GB/s

3.0 A disk based pool build from 4 basic HGST 7200rpm Ultrastar HE8 disks

No Slog vs ZeusRAM Slog, allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	quad-hd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1060 ops		48677 ops
	211.983 ops/s		9735.087 ops/s
	2001us cpu/op		227us cpu/op
	4.7ms latency		0.1ms latency
	1.6 MB/s		76.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	206 ops		5614 ops
	41.197 ops/s		1122.756 ops/s
	12313us cpu/op		2018us cpu/op
	24.0ms latency		0.9ms latency
	41.0 MB/s		1122.6 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	219.2 MB/s	262.8 MB/s	2.6 GB/s

---

hostname	omniosce Memory size: 32429 Megabytes		
pool	quad-hd (recsize=128k, compr=off, readcache=all)		
slog	ZeusRAM 8 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	7423 ops		48020 ops
	1484.554 ops/s		9603.579 ops/s
	680us cpu/op		240us cpu/op
	0.6ms latency		0.1ms latency
	11.4 MB/s		75.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1472 ops		5352 ops
	294.387 ops/s		1070.352 ops/s
	6781us cpu/op		2145us cpu/op
	3.4ms latency		0.9ms latency
	294.2 MB/s		1070.2 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	212.6 MB/s	240.0 MB/s	2.6 GB/s

3.1 A disk based pool build from 4 x HGST 7200rpm Ultrastar HE8 disks

Solidata (Sandforce) 120G SSD Slog vs Intel DC 3700 Slog (100GB overprovisioned to 60GB)

```
hostname                        omniosce  Memory size: 32429 Megabytes
pool                            quad-hd (recsize=128k, compr=off, readcache=all)
slog                            Solidata SSD 120 GB
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                9525 ops                        45940 ops
                                1904.912 ops/s                  9187.375 ops/s
                                797us cpu/op                    218us cpu/op
                                0.5ms latency                   0.1ms latency
                                14.8 MB/s                       71.6 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                879 ops                         5432 ops
                                175.790 ops/s                   1086.342 ops/s
                                7415us cpu/op                   2067us cpu/op
                                5.6ms latency                   0.9ms latency
                                175.6 MB/s                      1086.1 MB/s
_____________________________________________________________________________________

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=all               213.2 MB/s       271.0 MB/s     2.6 GB/s
_____________________________________________________________________________________


hostname                        omniosce  Memory size: 32429 Megabytes
pool                            quad-hd (recsize=128k, compr=off, readcache=all)
slog                            INTEL SSDSC2BA10 62.7 GB
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                11353 ops                       47994 ops
                                2270.511 ops/s                  9598.452 ops/s
                                781us cpu/op                    214us cpu/op
                                0.4ms latency                   0.1ms latency
                                17.6 MB/s                       74.8 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                951 ops                         5440 ops
                                190.188 ops/s                   1087.936 ops/s
                                7460us cpu/op                   2027us cpu/op
                                5.2ms latency                   0.9ms latency
                                190.0 MB/s                      1087.7 MB/s
_____________________________________________________________________________________

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=all               219.0 MB/s       281.4 MB/s     2.5 GB/s
_____________________________________________________________________________________
```

3.2 A disk based pool build from 4 x HGST 7200rpm Ultrastar HE8 disks

No Slog vs Intel Optane Slog M.2/ 32G Slog (readcache=none)

```
hostname                        omniosce  Memory size: 32429 Megabytes
pool                            hd (recsize=128k, compr=off, readcache=none)
slog                            -
remark


Fb3 randomwrite.f               sync=always                   sync=disabled
                                245 ops                       379 ops
                                48.996 ops/s                  75.794 ops/s
                                5982us cpu/op                 7264us cpu/op
                                20.3ms latency                13.1ms latency
                                0.2 MB/s                      0.4 MB/s

Fb4 singlestreamwrite.f         sync=always                   sync=disabled
                                201 ops                       5410 ops
                                40.195 ops/s                  1081.947 ops/s
                                12272us cpu/op                1948us cpu/op
                                24.7ms latency                0.9ms latency
                                40.0 MB/s                     1081.7 MB/s
```

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.6 MB/s	1.2 MB/s	304.6 MB/s

```
hostname                        omniosce  Memory size: 32429 Megabytes
pool                            hd (recsize=128k, compr=off, readcache=none)
slog                            blkdev 29.3 GB
remark


Fb3 randomwrite.f               sync=always                   sync=disabled
                                552 ops                       381 ops
                                110.389 ops/s                 76.194 ops/s
                                3134us cpu/op                 7908us cpu/op
                                9.0ms latency                 13.1ms latency
                                0.8 MB/s                      0.4 MB/s

Fb4 singlestreamwrite.f         sync=always                   sync=disabled
                                1089 ops                      5595 ops
                                217.791 ops/s                 1118.950 ops/s
                                7074us cpu/op                 2021us cpu/op
                                4.5ms latency                 0.9ms latency
                                217.6 MB/s                    1118.8 MB/s
```

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.6 MB/s	1.2 MB/s	319.6 MB/s

The 32G Optane M.2 Cachemodul performs similar to Sata SSDs and far below a 900P but seems a cheaper but good Slog device for a disk based pool when increased sync performance is needed. Main advantage is that Optane does not slow down under steady load as there is no trim or erase regions prior write what slows down traditional flash after some time.

As Intel does not guarantee powerloss protection on the cheaper Optane 32G or 900P I would not recommend them in mission critical environments.

The performance degration of Flash over time or depending of fillrate is not tested.

3.3 A disk based pool build from 4 x HGST 7200rpm Ultrastar HE8 disks

No Slog vs Intel Optane Slog M.2/ 32G Slog (readcache=all)

hostname	omniosce Memory size: 32429 Megabytes		
pool	hd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1096 ops		46064 ops
	219.176 ops/s		9212.411 ops/s
	2409us cpu/op		226us cpu/op
	4.5ms latency		0.1ms latency
	1.6 MB/s		71.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	201 ops		5583 ops
	40.196 ops/s		1116.547 ops/s
	9038us cpu/op		2065us cpu/op
	24.7ms latency		0.9ms latency
	40.0 MB/s		1116.3 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	227.8 MB/s	235.4 MB/s	2.6 GB/s

hostname	omniosce Memory size: 32429 Megabytes		
pool	hd (recsize=128k, compr=off, readcache=all)		
slog	blkdev 29.3 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	8826 ops		46980 ops
	1765.087 ops/s		9395.620 ops/s
	613us cpu/op		222us cpu/op
	0.6ms latency		0.1ms latency
	13.6 MB/s		73.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1080 ops		5653 ops
	215.993 ops/s		1130.539 ops/s
	7458us cpu/op		1922us cpu/op
	4.6ms latency		0.9ms latency
	215.8 MB/s		1130.3 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	215.0 MB/s	241.0 MB/s	2.6 GB/s

3.4 A disk based pool build from 4 x HGST 7200rpm Ultrastar HE8 disks

No Slog vs Intel Optane 900P Slog (readcache=none)

hostname	omniosce Memory size: 32429 Megabytes		
pool	hd (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	238 ops		400 ops
	47.597 ops/s		79.996 ops/s
	4939us cpu/op		6032us cpu/op
	21.0ms latency		12.5ms latency
	0.2 MB/s		0.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	199 ops		5546 ops
	39.796 ops/s		1109.152 ops/s
	11441us cpu/op		2052us cpu/op
	24.9ms latency		0.9ms latency
	39.6 MB/s		1109.0 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.6 MB/s	1.2 MB/s	327.0 MB/s

hostname	omniosce Memory size: 32429 Megabytes		
pool	hd (recsize=128k, compr=off, readcache=none)		
slog	blkdev 280.1 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	549 ops		431 ops
	109.790 ops/s		86.195 ops/s
	3370us cpu/op		6967us cpu/op
	9.1ms latency		11.5ms latency
	0.8 MB/s		0.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2139 ops		5548 ops
	427.783 ops/s		1109.533 ops/s
	4769us cpu/op		1995us cpu/op
	2.3ms latency		0.9ms latency
	427.6 MB/s		1109.3 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.6 MB/s	1.2 MB/s	283.2 MB/s

3.5 A disk based pool build from 4 x HGST 7200rpm Ultrastar HE8 disks

No Slog vs Intel Optane 900P Slog (readcache=all)

hostname	omniosce Memory size: 32429 Megabytes		
pool	hd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1099 ops		46717 ops
	219.777 ops/s		9343.024 ops/s
	2001us cpu/op		223us cpu/op
	4.5ms latency		0.1ms latency
	1.6 MB/s		72.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	213 ops		5607 ops
	42.596 ops/s		1121.353 ops/s
	8105us cpu/op		1969us cpu/op
	23.3ms latency		0.9ms latency
	42.4 MB/s		1121.2 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	223.4 MB/s	268.8 MB/s	2.6 GB/s

hostname	omniosce Memory size: 32429 Megabytes		
pool	hd (recsize=128k, compr=off, readcache=all)		
slog	blkdev 280.1 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	9653 ops		46498 ops
	1930.536 ops/s		9299.254 ops/s
	582us cpu/op		206us cpu/op
	0.5ms latency		0.1ms latency
	15.0 MB/s		72.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2241 ops		5446 ops
	448.184 ops/s		1089.137 ops/s
	4636us cpu/op		2064us cpu/op
	2.2ms latency		0.9ms latency
	448.0 MB/s		1088.9 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	210.6 MB/s	270.6 MB/s	2.6 GB/s

3.6 A disk based pool build from 4 x HGST 7200rpm Ultrastar HE8 disks

Intel P3600-400 Slog vs Intel Optane 900P-280 Slog

hostname	omniosce Memory size: 32429 Megabytes		
pool	quad-hd (recsize=128k, compr=off, readcache=all)		
slog	blkdev 400.1 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	10690 ops		47297 ops
	2137.873 ops/s		9458.962 ops/s
	622us cpu/op		227us cpu/op
	0.4ms latency		0.1ms latency
	16.6 MB/s		73.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1726 ops		5494 ops
	345.185 ops/s		1098.750 ops/s
	5904us cpu/op		2071us cpu/op
	2.9ms latency		0.9ms latency
	345.0 MB/s		1098.5 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	218.8 MB/s	263.8 MB/s	2.6 GB/s

---

hostname	omniosce Memory size: 32429 Megabytes		
pool	quad-hd (recsize=128k, compr=off, readcache=all)		
slog	blkdev 280.1 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	10288 ops		49868 ops
	2057.456 ops/s		9972.930 ops/s
	661us cpu/op		215us cpu/op
	0.5ms latency		0.1ms latency
	16.0 MB/s		77.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2399 ops		5551 ops
	479.779 ops/s		1110.148 ops/s
	4576us cpu/op		2113us cpu/op
	2.1ms latency		0.9ms latency
	479.6 MB/s		1109.9 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	225.6 MB/s	252.6 MB/s	2.6 GB/s

3.7 A SSD based pool build from 4 SSD Intel DC 3510-120 SSD

No Slog vs ZeusRAM Slog, allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	quad-dc3510 (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	8033 ops		34060 ops
	1606.510 ops/s		6811.684 ops/s
	834us cpu/op		239us cpu/op
	0.6ms latency		0.1ms latency
	12.4 MB/s		53.2 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1979 ops		4862 ops
	395.785 ops/s		972.345 ops/s
	4542us cpu/op		2061us cpu/op
	2.5ms latency		1.0ms latency
	395.6 MB/s		972.1 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	206.0 MB/s	271.6 MB/s	2.5 GB/s

---

hostname	omniosce Memory size: 32429 Megabytes		
pool	quad-dc3510 (recsize=128k, compr=off, readcache=all)		
slog	ZeusRAM 8 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	9184 ops		34296 ops
	1836.710 ops/s		6858.840 ops/s
	654us cpu/op		251us cpu/op
	0.5ms latency		0.1ms latency
	14.2 MB/s		53.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1458 ops		4793 ops
	291.589 ops/s		958.553 ops/s
	7127us cpu/op		2220us cpu/op
	3.4ms latency		1.0ms latency
	291.4 MB/s		958.4 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	210.6 MB/s	271.2 MB/s	2.5 GB/s

---

4 x SSD Pool with ZeusRAM Slog slower than without Slog!

3.8 A SSD based pool build from 4 SSD Intel DC 3510-120 SSD

P3600-400 Slog vs Optane 900P Slog, allcache

```
hostname                        omniosce  Memory size: 32429 Megabytes
pool                            quad-dc3510 (recsize=128k, compr=off, readcache=all)
slog                            blkdev 400.1 GB
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                11029 ops                       36176 ops
                                2205.691 ops/s                  7234.911 ops/s
                                700us cpu/op                    224us cpu/op
                                0.4ms latency                   0.1ms latency
                                17.2 MB/s                       56.4 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                1835 ops                        4745 ops
                                366.981 ops/s                   948.933 ops/s
                                5743us cpu/op                   2056us cpu/op
                                2.7ms latency                   1.0ms latency
                                366.8 MB/s                      948.7 MB/s
-------------------------------------------------------------------------------

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=all               215.4 MB/s       265.8 MB/s     2.6 GB/s
-------------------------------------------------------------------------------

hostname                        omniosce  Memory size: 32429 Megabytes
pool                            quad-dc3510 (recsize=128k, compr=off, readcache=all)
slog                            blkdev 280.1 GB
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                10250 ops                       35707 ops
                                2049.898 ops/s                  7141.142 ops/s
                                700us cpu/op                    250us cpu/op
                                0.5ms latency                   0.1ms latency
                                16.0 MB/s                       55.6 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                1854 ops                        4757 ops
                                370.782 ops/s                   951.358 ops/s
                                5245us cpu/op                   2262us cpu/op
                                2.7ms latency                   1.0ms latency
                                370.6 MB/s                      951.2 MB/s
-------------------------------------------------------------------------------

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=all               209.0 MB/s       266.2 MB/s     2.5 GB/s
-------------------------------------------------------------------------------


no real advantage against 4 SSD without slog!
```

3.9 An NVMe based pool build from 2 Intel P750 basic vdev

nocache vs allcache

hostname	omniosce Memory size: 32429 Megabytes		
pool	dual-p750 (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1980 ops		3415 ops
	395.970 ops/s		682.963 ops/s
	1620us cpu/op		1166us cpu/op
	2.5ms latency		1.4ms latency
	3.0 MB/s		5.2 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	4717 ops		11021 ops
	943.364 ops/s		2204.126 ops/s
	2664us cpu/op		1695us cpu/op
	1.0ms latency		0.4ms latency
	943.2 MB/s		2203.9 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	6.2 MB/s	15.6 MB/s	347.0 MB/s

hostname	omniosce Memory size: 32429 Megabytes		
pool	dual-p750 (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	8235 ops		93094 ops
	1646.895 ops/s		18618.382 ops/s
	587us cpu/op		173us cpu/op
	0.6ms latency		0.0ms latency
	12.8 MB/s		145.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	4536 ops		8587 ops
	907.176 ops/s		1717.321 ops/s
	2632us cpu/op		3743us cpu/op
	1.1ms latency		0.6ms latency
	907.0 MB/s		1717.1 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	249.2 MB/s	269.4 MB/s	2.6 GB/s

3.10 SSD based pool build from 16 x Sandisk Pro Extreme-960 basic vdev (Raid-0)

no Slog vs Optane 800P-118 Slog

hostname	OMNI26 Memory size: 96940 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	2430 ops		160551 ops
	485.968 ops/s		32108.865 ops/s
	1634us cpu/op		165us cpu/op
	2.0ms latency		0.0ms latency
	3.6 MB/s		250.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	347 ops		8800 ops
	69.395 ops/s		1759.930 ops/s
	11768us cpu/op		2546us cpu/op
	14.3ms latency		0.6ms latency
	69.2 MB/s		1759.7 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	404.4 MB/s	624.6 MB/s	2.9 GB/s

hostname	OMNI26 Memory size: 96940 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=all)		
slog	blkdev 118.4 GB		
remark	optane 800p slog		
Fb3 randomwrite.f	sync=always		sync=disabled
	11399 ops		154345 ops
	2279.634 ops/s		30867.228 ops/s
	632us cpu/op		132us cpu/op
	0.4ms latency		0.0ms latency
	17.8 MB/s		241.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1733 ops		10619 ops
	346.585 ops/s		2123.730 ops/s
	6352us cpu/op		2198us cpu/op
	2.9ms latency		0.5ms latency
	346.4 MB/s		2123.5 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	404.0 MB/s	674.2 MB/s	2.9 GB/s

3.11 SSD based pool build from 16 x Sandisk Pro Extreme-960 basic vdev (Raid-0)

no Slog vs Optane 900P-280 Slog

hostname	OMNI26 Memory size: 96940 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	2430 ops		160551 ops
	485.968 ops/s		32108.865 ops/s
	1634us cpu/op		165us cpu/op
	2.0ms latency		0.0ms latency
	3.6 MB/s		250.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	347 ops		8800 ops
	69.395 ops/s		1759.930 ops/s
	11768us cpu/op		2546us cpu/op
	14.3ms latency		0.6ms latency
	69.2 MB/s		1759.7 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	404.4 MB/s	624.6 MB/s	2.9 GB/s

---

hostname	OMNI26 Memory size: 96940 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=all)		
slog	blkdev 118.4 GB		
remark	900p		
Fb3 randomwrite.f	sync=always		sync=disabled
	11675 ops		71466 ops
	2334.882 ops/s		14291.960 ops/s
	686us cpu/op		94us cpu/op
	0.4ms latency		0.1ms latency
	18.2 MB/s		111.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1745 ops		11692 ops
	348.988 ops/s		2338.283 ops/s
	5947us cpu/op		2037us cpu/op
	2.8ms latency		0.4ms latency
	348.8 MB/s		2338.1 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	398.8 MB/s	691.2 MB/s	2.9 GB/s

---

see also
https://forums.servethehome.com/index.php?threads/ssd-performance-issues-again.19625/page-2

In this thread a pool build from 14 x Intel DC 3700 in a 7x raid-10 setup:
sync write in a Filenench Singlestreamwrite with sync enabled: 357 MB/s wrize

same setup but with an Optane 900P as Slog
sync write in a Filenench Singlestreamwrite with sync enabled: 477 MB/s write

Result:
Even with quite the fastest Sata SSD: + around 30% sync write performance

3.12 SSD based pool build from 16 x Sandisk Pro Extreme-960 basic vdev (Raid-0)

two slog vs 3 slogs (adding more slogs will do a load balancing between them)

hostname	OMNI26 Memory size: 96940 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=all)		
slog	blkdev 118.4 GB		
remark	2 x 800P as Slog load balance		
Fb3 randomwrite.f	sync=always		sync=disabled
	13598 ops		123570 ops
	2703.368 ops/s		24712.634 ops/s
	459us cpu/op		124us cpu/op
	0.4ms latency		0.0ms latency
	21.1 MB/s		193.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2381 ops		11903 ops
	476.181 ops/s		2380.476 ops/s
	4630us cpu/op		1524us cpu/op
	2.1ms latency		0.4ms latency
	476.0 MB/s		2380.3 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	382.8 MB/s	554.0 MB/s	2.7 GB/s

---

346 MB/s -> 476 MB/s sequential sync write, around 20% better

hostname	OMNI26 Memory size: 96940 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=all)		
slog	blkdev 118.4 GB		
remark	2 x 800P + 900P as Slog load balance		
Fb3 randomwrite.f	sync=always		sync=disabled
	17572 ops		161653 ops
	3514.259 ops/s		32328.721 ops/s
	451us cpu/op		134us cpu/op
	0.3ms latency		0.0ms latency
	27.4 MB/s		252.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2722 ops		11824 ops
	544.377 ops/s		2364.725 ops/s
	3917us cpu/op		1646us cpu/op
	1.8ms latency		0.4ms latency
	544.2 MB/s		2364.5 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	385.6 MB/s	549.4 MB/s	2.7 GB/s

---

476 MB/s -> 544 MB/s, another 15%

Overall
You can mirror Slogs. This keeps full performance even when one dies and avoids a dataloss in case of a crash during a write with a damaged Slog at this point. An Slog failure at any other time is uncritical as ZFS then reverts to the onpool ZIL for logging. Think of a mirrored Slog like a hardwareraid with cache and two battery units.

If you add more than one Slog to a pool, you will do a load balancing between them so each must only do a part of the load with the result of a better performance.

3.13 HD based pool build from 11 x HGST 2TB basic vdev (Raid-0)

Effect of multiple Slogs from partitions on a single Optane 900P: no Slog vs one Slog part

hostname	omni26 Memory size: 32429 Megabytes		
pool	hd-11 (recsize=128k, compr=off, readcache=all)		
slog	-		
remark	11 x HGST 2TB in Raid-0		
Fb3 randomwrite.f	sync=always		sync=disabled
	525 ops		64988 ops
	104.991 ops/s		12997.092 ops/s
	3636us cpu/op		327us cpu/op
	9.5ms latency		0.1ms latency
	0.8 MB/s		101.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	176 ops		6059 ops
	35.197 ops/s		1211.750 ops/s
	13789us cpu/op		1976us cpu/op
	28.2ms latency		0.8ms latency
	35.0 MB/s		1211.5 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	266.4 MB/s	308.0 MB/s	2.8 GB/s

hostname	omni26 Memory size: 32429 Megabytes		
pool	hd-11 (recsize=128k, compr=off, readcache=all)		
slog	25.2 GB		
remark	11 x HGST 2TB in Raid-0 + one Slog (20G part from 900P)		
Fb3 randomwrite.f	sync=always		sync=disabled
	18358 ops		64034 ops
	3671.436 ops/s		12806.095 ops/s
	438us cpu/op		201us cpu/op
	0.3ms latency		0.1ms latency
	28.6 MB/s		100.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2699 ops		6631 ops
	539.775 ops/s		1326.128 ops/s
	4495us cpu/op		1953us cpu/op
	1.8ms latency		0.7ms latency
	539.6 MB/s		1325.9 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	243.6 MB/s	311.4 MB/s	2.8 GB/s

performance boost is dramatic from 35 MB/s seq sync to 539 MB/s

3.14 HD based pool build from 11 x HGST 2TB basic vdev (Raid-0)

Effect of multiple Slogs from partitions from one Optane 900P: two Slog vs three Slog part

hostname	omni26 Memory size: 32429 Megabytes		
pool	hd-11 (recsize=128k, compr=off, readcache=all)		
slog	25.2 GB		
remark	11 x HGST 2TB in Raid-0 + two Slog (20G part from 900P)		
Fb3 randomwrite.f	sync=always		sync=disabled
	18563 ops		66267 ops
	3712.461 ops/s		13253.006 ops/s
	451us cpu/op		204us cpu/op
	0.3ms latency		0.1ms latency
	29.0 MB/s		103.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2738 ops		6905 ops
	547.581 ops/s		1380.933 ops/s
	4168us cpu/op		1875us cpu/op
	1.8ms latency		0.7ms latency
	547.4 MB/s		1380.7 MB/s
read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	257.4 MB/s	318.8 MB/s	2.7 GB/s

hostname	omni26 Memory size: 32429 Megabytes		
pool	hd-11 (recsize=128k, compr=off, readcache=all)		
slog	25.2 GB		
remark	11 x HGST 2TB in Raid-0 + three Slog (20G part from 900P)		
Fb3 randomwrite.f	sync=always		sync=disabled
	17123 ops		65281 ops
	3424.481 ops/s		13055.691 ops/s
	456us cpu/op		194us cpu/op
	0.3ms latency		0.1ms latency
	26.6 MB/s		102.0 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2478 ops		6599 ops
	495.579 ops/s		1319.739 ops/s
	4751us cpu/op		2132us cpu/op
	2.0ms latency		0.7ms latency
	495.4 MB/s		1319.5 MB/s
read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	245.8 MB/s	316.8 MB/s	2.8 GB/s

Multiple Slog parttions cannot give an improvement regarding endurance and there is no performance improvment si this does not make sense.

This is different to multiple Slogs from multiple Optane 800P where you get a better endurance and a noticable better performance.

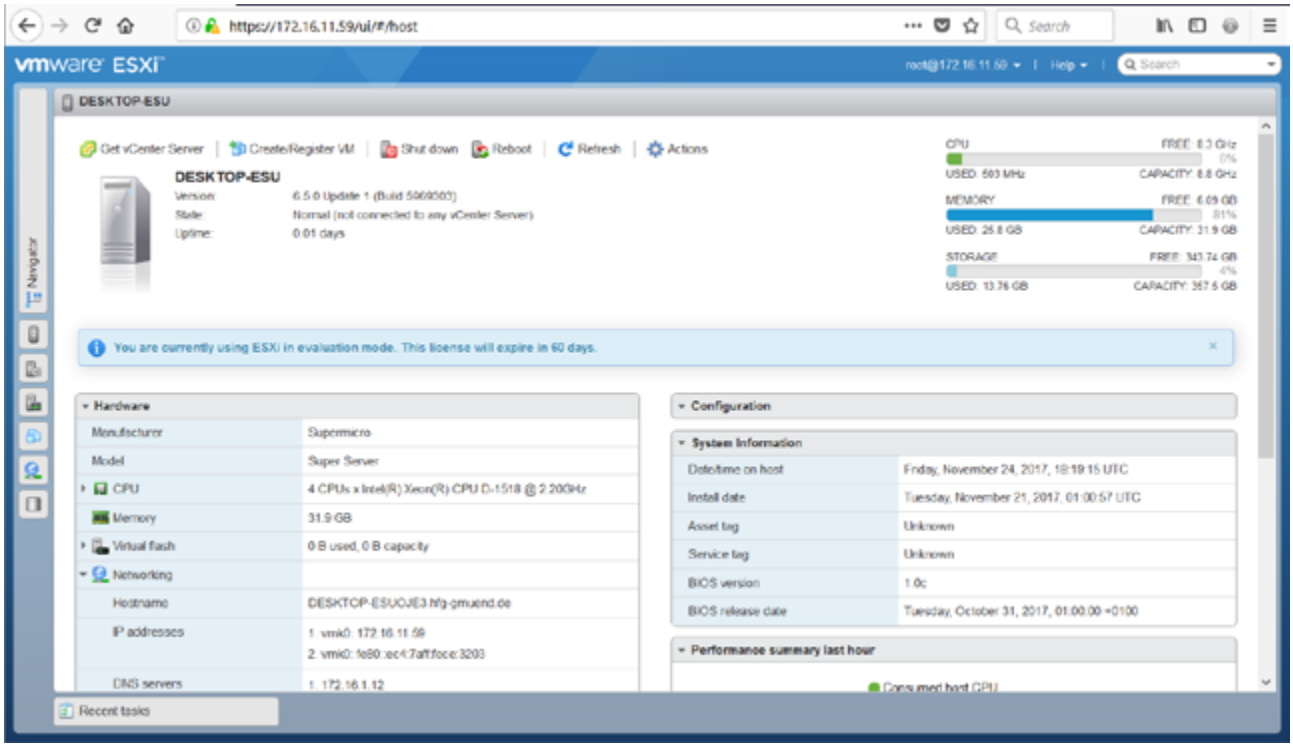
4. AiO setup (ESXi with a virtualized ZFS NAS/SAN)
SM X11SPH-nCTF, Xeon Silver 8Core 4110, 32G RAM

see http://napp-it.org/doc/downloads/napp-in-one.pdf

The new Intel Optane NVMe are currently not working under ESXi in pass-through mode. But what about using them over the ESXi driver as virtual disks? You will not be able to use Smartmontools from OmniOS and there is a slighty reduced performance when using ZFS on files on VMFS over the NVMe driver of ESXi.

But given the performance of the Optane and as there is no controller or disk cache that you must fear of, this may be an option unless Optane pass-through is working in ESXi.

In the following benchmark I use the same hardware as for my barebone tests with 24 GB RAM/ 2vcpu for the OmniOS 151024ce storage VM. The LSI HBA is in pass-througmode for disks and Optane 900P is used as an ESXi datastore for virtual disks.



The most common AiO setups are diskbased and ssd based pools with an additional Optane as Slog or L2Arc. Both can be delivered from one Optane as the additional readload affects write performance not in the same way like on current Flash disks.

The Optane currently does not work as a pass-through device under ESXi.

As the Optane does not contain a cache I would expect that using it as a virtual disks is acceptable from a security view and fast enough due its extreme performance. Following some benchmarks without Optane.

For the following tests, I gave 24/16/8/4/2 GB RAM to the storage VM with always 2 vcpu

4.1 A Disk based pool via LSI pass-through (4 x HGST HE8 vdev)

nocache vs allcache, 24G RAM

hostname	napp-it-san024		
pool	quad-hd (recsize=128k, compr=off, readcache=none)		
slog	-		
remark	4 x HE8, no cache, no slog		
Fb3 randomwrite.f	sync=always		sync=disabled
	241 ops		435 ops
	48.197 ops/s		86.995 ops/s
	7203us cpu/op		4445us cpu/op
	16.5ms latency		9.2ms latency
	0.2 MB/s		0.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	198 ops		5084 ops
	39.596 ops/s		1016.756 ops/s
	9682us cpu/op		1181us cpu/op
	25.0ms latency		1.0ms latency
	39.4 MB/s		1016.6 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.6 MB/s	1.2 MB/s	430.4 MB/s

---

hostname	napp-it-san024		
pool	quad-hd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark	4 x HE8, allcache. no slog		
Fb3 randomwrite.f	sync=always		sync=disabled
	966 ops		40472 ops
	193.188 ops/s		8078.390 ops/s
	2142us cpu/op		124us cpu/op
	4.1ms latency		0.1ms latency
	1.4 MB/s		63.1 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	194 ops		5142 ops
	38.797 ops/s		1028.301 ops/s
	8955us cpu/op		1033us cpu/op
	25.6ms latency		1.0ms latency
	38.6 MB/s		1028.1 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	187.4 MB/s	184.8 MB/s	2.1 GB/s

4.2 A Disk based pool via LSI pass-through (4 x HGST HE8 vdev)

nocache vs allcache and Slog (vdisk on Optane 900P) as vdisk, 24G RAM

hostname	napp-it-san024		
pool	quad-hd (recsize=128k, compr=off, readcache=none)		
slog	Virtual disk 21.5 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	361 ops		442 ops
	72.194 ops/s		88.394 ops/s
	4087us cpu/op		4466us cpu/op
	11.1ms latency		9.0ms latency
	0.4 MB/s		0.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2958 ops		5015 ops
	591.573 ops/s		1002.978 ops/s
	2456us cpu/op		1137us cpu/op
	1.7ms latency		1.0ms latency
	591.4 MB/s		1002.8 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	0.6 MB/s	1.4 MB/s	447.4 MB/s

---

hostname	napp-it-san024		
pool	quad-hd (recsize=128k, compr=off, readcache=all)		
slog	Virtual disk 21.5 GB		
remark	4 x HE8, slog, allcache		
Fb3 randomwrite.f	sync=always		sync=disabled
	19554 ops		41275 ops
	3910.645 ops/s		8254.787 ops/s
	273us cpu/op		138us cpu/op
	0.2ms latency		0.1ms latency
	30.4 MB/s		64.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3175 ops		5080 ops
	634.943 ops/s		1015.957 ops/s
	2404us cpu/op		1046us cpu/op
	1.6ms latency		1.0ms latency
	634.7 MB/s		1015.8 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	179.8 MB/s	184.1 MB/s	2.5 GB/s

4.3 A Disk based pool via LSI pass-through (4 x HGST HE8 vdev)

allcache, slog, 16GB vs 8GB RAM

```
hostname                        napp-it-san024  Memory size: 16384 Megabytes

pool                            quad-hd (recsize=128k, compr=off, readcache=all)
slog                            Virtual disk 21.5 GB
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                18941 ops                       34035 ops
                                3788.059 ops/s                  6768.878 ops/s
                                271us cpu/op                    147us cpu/op
                                0.2ms latency                   0.1ms latency
                                29.4 MB/s                       52.7 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                3228 ops                        3765 ops
                                644.146 ops/s                   752.965 ops/s
                                2321us cpu/op                   1232us cpu/op
                                1.5ms latency                   1.1ms latency
                                643.9 MB/s                      752.8 MB/s
-------------------------------------------------------------------------------

read fb 7-9 + dd (opt)          randomread.f    randomrw.f      singlestreamr
pri/sec cache=all               187.2 MB/s      230.0 MB/s      2.1 GB/s
-------------------------------------------------------------------------------


hostname                        napp-it-san024  Memory size: 8192 Megabytes

pool                            quad-hd (recsize=128k, compr=off, readcache=all)
slog                            Virtual disk 21.5 GB
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                19955 ops                       26116 ops
                                3990.885 ops/s                  5222.946 ops/s
                                260us cpu/op                    183us cpu/op
                                0.2ms latency                   0.1ms latency
                                31.0 MB/s                       40.8 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                646 ops                         3881 ops
                                126.408 ops/s                   776.179 ops/s
                                15781us cpu/op                  1259us cpu/op
                                7.8ms latency                   1.3ms latency
                                126.2 MB/s                      776.0 MB/s
-------------------------------------------------------------------------------

read fb 7-9 + dd (opt)          randomread.f    randomrw.f      singlestreamr
pri/sec cache=all               234.0 MB/s      224.2 MB/s      2.1 GB/s
-------------------------------------------------------------------------------
```

4.4 A Disk based pool via LSI pass-through (4 x HGST HE8 vdev)

allcache, slog, 4GB vs 2GB RAM

hostname	napp-it-san024 Memory size: 4096 Megabytes		
pool	quad-hd (recsize=128k, compr=off, readcache=all)		
slog	Virtual disk 21.5 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1180 ops		941 ops
	235.993 ops/s		188.181 ops/s
	1627us cpu/op		1754us cpu/op
	4.2ms latency		4.2ms latency
	1.8 MB/s		1.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2374 ops		3328 ops
	455.934 ops/s		665.586 ops/s
	3520us cpu/op		1262us cpu/op
	2.2ms latency		1.5ms latency
	455.7 MB/s		665.4 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	1.4 MB/s	3.0 MB/s	604.6 MB/s

hostname	napp-it-san024 Memory size: 2048 Megabytes		
pool	quad-hd (recsize=128k, compr=off, readcache=all)		
slog	Virtual disk 21.5 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	586 ops		814 ops
	117.187 ops/s		162.784 ops/s
	3314us cpu/op		2262us cpu/op
	6.8ms latency		6.1ms latency
	0.8 MB/s		1.2 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2364 ops		2694 ops
	472.753 ops/s		538.780 ops/s
	3163us cpu/op		2338us cpu/op
	2.1ms latency		1.8ms latency
	472.6 MB/s		538.6 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	1.0 MB/s	1.6 MB/s	613.0 MB/s

4.5 A SSD based pool via LSI pass-through (4 x Intel DC 3510 vdev)

nocache vs allcache, no slog, 24G RAM

hostname	napp-it-san024		
pool	quad-ssd (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1941 ops		2458 ops
	387.962 ops/s		491.588 ops/s
	1031us cpu/op		935us cpu/op
	2.1ms latency		1.6ms latency
	3.0 MB/s		3.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2455 ops		4239 ops
	490.991 ops/s		847.781 ops/s
	1654us cpu/op		1172us cpu/op
	2.0ms latency		1.2ms latency
	490.8 MB/s		847.6 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	4.4 MB/s	9.4 MB/s	341.2 MB/s

hostname	napp-it-san024		
pool	quad-ssd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark	4 x ssd, allcache, no slog		
Fb3 randomwrite.f	sync=always		sync=disabled
	7997 ops		31044 ops
	1599.363 ops/s		6208.635 ops/s
	343us cpu/op		124us cpu/op
	0.5ms latency		0.1ms latency
	12.4 MB/s		48.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	1716 ops		4350 ops
	343.184 ops/s		869.982 ops/s
	1960us cpu/op		1130us cpu/op
	2.3ms latency		1.1ms latency
	343.0 MB/s		869.8 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	184.8 MB/s	245.4 MB/s	2.5 GB/s

4.6 A SSD based pool via LSI pass-through (4 x Intel DC 3510 vdev)

nocache vs allcache, slog, 24G RAM

hostname	napp-it-san024		
pool	quad-ssd (recsize=128k, compr=off, readcache=none)		
slog	Virtual disk 21.5 GB		
remark	slog, nocache		
Fb3 randomwrite.f	sync=always		sync=disabled
	2315 ops		2164 ops
	462.988 ops/s		432.791 ops/s
	989us cpu/op		1857us cpu/op
	1.7ms latency		1.8ms latency
	3.6 MB/s		3.2 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2053 ops		4425 ops
	382.583 ops/s		884.980 ops/s
	4572us cpu/op		1138us cpu/op
	2.6ms latency		1.1ms latency
	382.4 MB/s		884.8 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	4.8 MB/s	9.2 MB/s	349.2 MB/s

---

hostname	napp-it-san024		
pool	quad-ssd (recsize=128k, compr=off, readcache=all)		
slog	Virtual disk 21.5 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	19575 ops		32994 ops
	3914.854 ops/s		6598.563 ops/s
	264us cpu/op		147us cpu/op
	0.2ms latency		0.1ms latency
	30.4 MB/s		51.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3285 ops		4345 ops
	656.768 ops/s		867.048 ops/s
	2360us cpu/op		1164us cpu/op
	1.5ms latency		1.1ms latency
	656.6 MB/s		866.8 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	176.0 MB/s	204.8 MB/s	2.5 GB/s

4.7 A SSD based pool via LSI pass-through (4 x Intel DC 3510 vdev)

allcache, slog, 16G RAM vs 8GB RAM

hostname	napp-it-san024 Memory size: 16384 Megabytes		
pool	quad-ssd (recsize=128k, compr=off, readcache=all)		
slog	Virtual disk 21.5 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	18664 ops		28754 ops
	3732.615 ops/s		5750.627 ops/s
	302us cpu/op		141us cpu/op
	0.2ms latency		0.1ms latency
	29.0 MB/s		44.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3339 ops		3825 ops
	667.744 ops/s		764.966 ops/s
	2286us cpu/op		1123us cpu/op
	1.5ms latency		1.3ms latency
	667.5 MB/s		764.8 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	225.6 MB/s	203.2 MB/s	2.6 GB/s

hostname	napp-it-san024 Memory size: 8192 Megabytes		
pool	quad-ssd (recsize=128k, compr=off, readcache=all)		
slog	Virtual disk 21.5 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	19911 ops		27367 ops
	3982.068 ops/s		5473.222 ops/s
	265us cpu/op		160us cpu/op
	0.2ms latency		0.2ms latency
	31.0 MB/s		42.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3066 ops		3180 ops
	612.776 ops/s		635.986 ops/s
	2353us cpu/op		1427us cpu/op
	1.6ms latency		1.6ms latency
	612.6 MB/s		635.8 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	185.6 MB/s	170.0 MB/s	2.4 GB/s

4.8 A SSD based pool via LSI pass-through (4 x Intel DC 3510 vdev)

allcache, slog, 4G RAM vs 2GB RAM

hostname	napp-it-san024 Memory size: 4096 Megabytes		
pool	quad-ssd (recsize=128k, compr=off, readcache=all)		
slog	Virtual disk 21.5 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	7843 ops		7392 ops
	1568.547 ops/s		1478.355 ops/s
	425us cpu/op		411us cpu/op
	0.6ms latency		0.5ms latency
	12.2 MB/s		11.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2925 ops		2972 ops
	584.960 ops/s		594.389 ops/s
	2647us cpu/op		1410us cpu/op
	1.7ms latency		1.7ms latency
	584.8 MB/s		594.2 MB/s
read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	17.8 MB/s	20.8 MB/s	1.2 GB/s

hostname	napp-it-san024 Memory size: 2048 Megabytes		
pool	quad-ssd (recsize=128k, compr=off, readcache=all)		
slog	Virtual disk 21.5 GB		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	4467 ops		4981 ops
	893.373 ops/s		996.121 ops/s
	625us cpu/op		506us cpu/op
	0.9ms latency		0.8ms latency
	6.8 MB/s		7.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	2251 ops		2729 ops
	450.172 ops/s		545.740 ops/s
	3473us cpu/op		2187us cpu/op
	2.2ms latency		1.8ms latency
	450.0 MB/s		545.5 MB/s
read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	14.6 MB/s	14.5 MB/s	1.2 GB/s

4.9 A Optane 200G based pool build from a vdisk

Single 900P nocache vs allcache, 24G RAM

hostname	napp-it-san024		
pool	sungle-optane (recsize=128k, compr=off, readcache=none)		
slog	-		
remark	single 900P vdisk, nocache		
Fb3 randomwrite.f	sync=always		sync=disabled
	6586 ops		7555 ops
	1317.086 ops/s		1510.959 ops/s
	640us cpu/op		560us cpu/op
	0.6ms latency		0.5ms latency
	10.2 MB/s		11.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3950 ops		5049 ops
	789.955 ops/s		1009.699 ops/s
	1434us cpu/op		1218us cpu/op
	1.3ms latency		0.8ms latency
	789.8 MB/s		1009.5 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	15.8 MB/s	29.2 MB/s	1.1 GB/s

hostname	napp-it-san024		
pool	sungle-optane (recsize=128k, compr=off, readcache=all)		
slog	-		
remark	900P vdisk, allcache		
Fb3 randomwrite.f	sync=always		sync=disabled
	16385 ops		44496 ops
	3276.901 ops/s		8898.625 ops/s
	328us cpu/op		130us cpu/op
	0.2ms latency		0.1ms latency
	25.6 MB/s		69.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3840 ops		6249 ops
	767.960 ops/s		1249.658 ops/s
	1402us cpu/op		1271us cpu/op
	1.3ms latency		0.8ms latency
	767.8 MB/s		1249.5 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	225.6 MB/s	193.6 MB/s	2.6 GB/s

4.10 A Optane 200G based pool build from a vdisk

Single 900P allcache, 16GB RAM vs 8GB RAM

hostname	napp-it-san024 Memory size: 16384 Megabytes		
pool	sungle-optane (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	16002 ops		37470 ops
	3200.299 ops/s		7493.433 ops/s
	309us cpu/op		146us cpu/op
	0.2ms latency		0.1ms latency
	25.0 MB/s		58.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3590 ops		5571 ops
	717.969 ops/s		1112.385 ops/s
	1556us cpu/op		1215us cpu/op
	1.4ms latency		0.9ms latency
	717.8 MB/s		1112.2 MB/s
read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	173.8 MB/s	201.2 MB/s	2.5 GB/s

hostname	napp-it-san024 Memory size: 8192 Megabytes		
pool	sungle-optane (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	17515 ops		35643 ops
	3502.907 ops/s		7128.387 ops/s
	333us cpu/op		169us cpu/op
	0.3ms latency		0.1ms latency
	27.2 MB/s		55.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	3297 ops		4806 ops
	659.385 ops/s		961.171 ops/s
	1552us cpu/op		1331us cpu/op
	1.5ms latency		1.0ms latency
	659.2 MB/s		961.0 MB/s
read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	186.8 MB/s	153.0 MB/s	2.6 GB/s

4.11 A Optane 200G based pool build from a vdisk

Single 900P allcache, 4GB RAM vs 2GB RAM

```
hostname                        napp-it-san024  Memory size: 4096 Megabytes

pool                            single-optane (recsize=128k, compr=off, readcache=all)
slog                            -
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                8982 ops                        25308 ops
                                1796.354 ops/s                  5061.458 ops/s
                                422us cpu/op                    241us cpu/op
                                0.4ms latency                   0.2ms latency
                                14.0 MB/s                       39.4 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                2811 ops                        3779 ops
                                562.187 ops/s                   755.782 ops/s
                                1749us cpu/op                   1557us cpu/op
                                1.8ms latency                   1.3ms latency
                                562.0 MB/s                      755.6 MB/s
_____________________________________________________________________________________

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=all               48.6 MB/s        57.0 MB/s      1.3 GB/s
_____________________________________________________________________________________


hostname                        napp-it-san024  Memory size: 2048 Megabytes

pool                            sungle-optane (recsize=128k, compr=off, readcache=all)
slog                            -
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                9116 ops                        12350 ops
                                1823.150 ops/s                  2469.919 ops/s
                                579us cpu/op                    433us cpu/op
                                0.5ms latency                   0.3ms latency
                                14.2 MB/s                       19.2 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                2367 ops                        3954 ops
                                473.387 ops/s                   790.616 ops/s
                                2340us cpu/op                   2114us cpu/op
                                2.1ms latency                   1.2ms latency
                                473.2 MB/s                      790.4 MB/s
_____________________________________________________________________________________

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=all               36.2 MB/s        35.4 MB/s      1.5 GB/s
_____________________________________________________________________________________
```

5.0 Oracle Solaris 11.4 ZFS v.43

I cannot publish results due Oracle restrictions.

But I am impressed of the good pool and SMB values on Solaris.
You nmust do your own tests

5.5 OmniOS OpenZFS iSCSI (AJA vs Atto vs Crystal Benchmark) XL710, 32G RAM

AJA 4k full, 16 bit RGB, 4GB file, sync disabled, compress off, iSCSI LUN, 40G, Jumbo, newest Win/Intel driver

4 x HD Pool

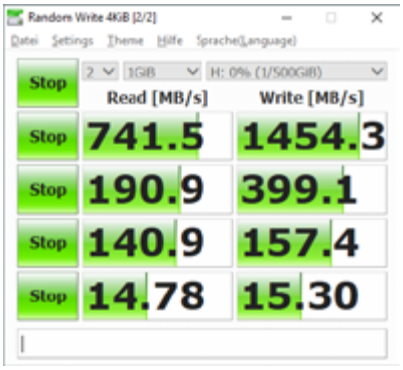
4 x SSD Pool

4 x Optane 900P Pool

Values are extremely sensitive to Windows driver (use newest Win10 driver from Intel) and settings like int throtteling. If I disable, write values are around 30-40% better while read values are worse. Jumboframes increses performance at around 20%

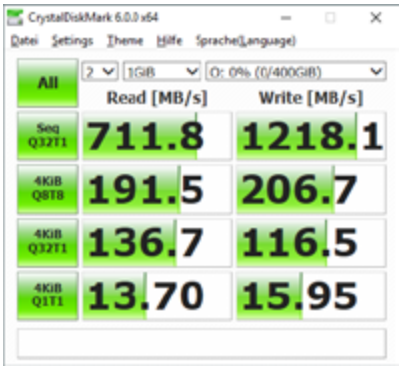
Crystal Diskmark, sync disabled, compress off, iSCSI LUN

4 x HD Pool



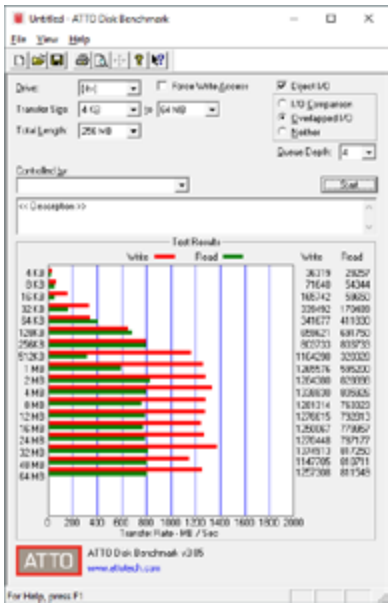
4 x SSD Pool

4 x Optane 900P Pool

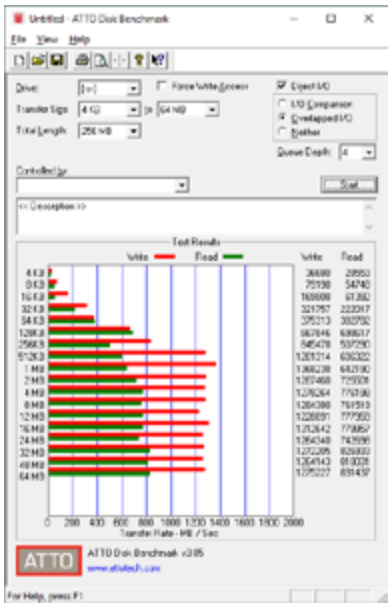


Atto QD2, sync disabled, compress off, iSCSI LUN

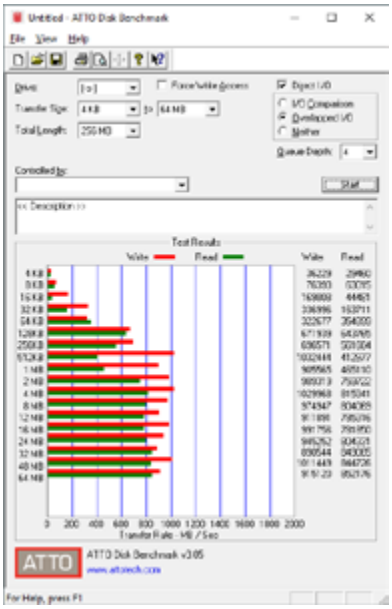
4 x HD Pool



4 x SSD Pool



4 x Optane 900P Pool



The general rersults are quite similar, indicating that RAM on the Server is the limiting factor beside network and drivers. Especially in the Atto, Optane shows its quality with the very similar write/read values over blocksize.

To verify, I redo all tests with 16G and 8GB RAM. Values with 10G nics are quite the same but as I am near the 10G limit i have used 40G

Do not forget to deactivate your virus scanner on Windows during benchmarks !
Prior tests I executed a default tuning (System - Appliance Tuning)

5.6 OmniOS OpenZFS iSCSI (AJA vs Atto vs Crystal Benchmark) XL710, 32G RAM

AJA 4k full, 16 bit RGB, 4GB file, sync disabled, compress off, iSCSI LUN, 40G, Jumbo, newest Win/Intel driver

Sync write enabled!

4 x HD Pool

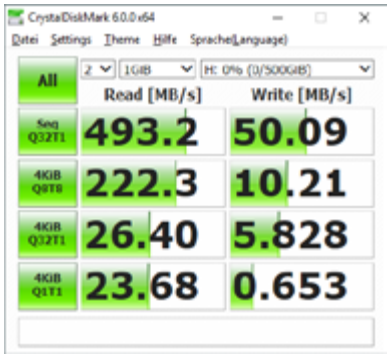
4 x SSD Pool

4 x Optane 900P Pool

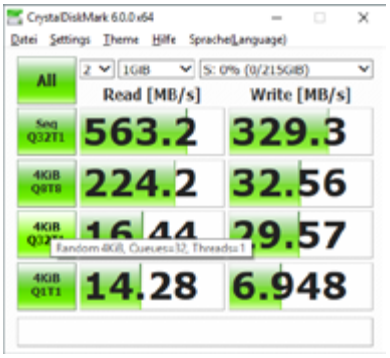
Sync write and HD only pools is a NoGo. SSD are ok, Optane writes 30% faster than SSD

Crystal Diskmark, sync enabled, compress off, iSCSI LUN

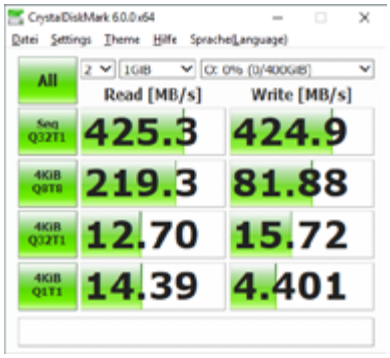
4 x HD Pool



4 x SSD Pool

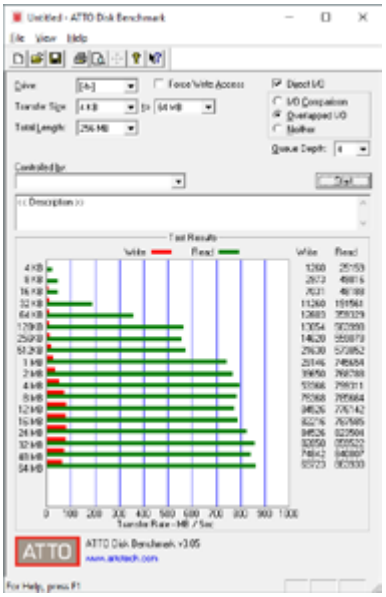


4 x Optane 900P Pool

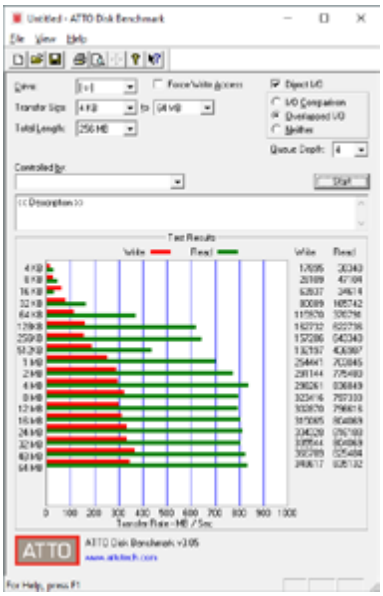


Atto QD2, sync enabled , compress off, iSCSI LUN

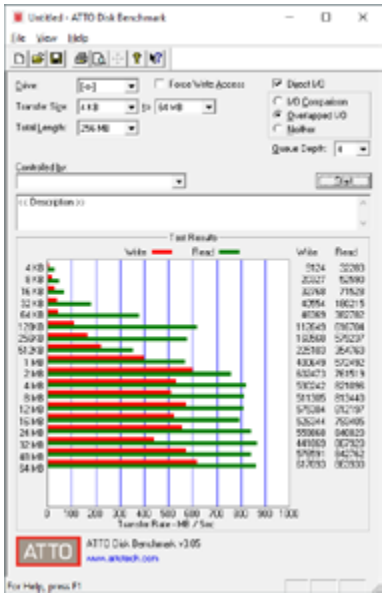
4 x HD Pool



4 x SSD Pool



4 x Optane 900P Pool



Whenever your workload cannot be processes mainly from RAM like with sync write enabled, you aware the quality of disks. While a HD Pool can give only 40 MB/s sync writes the Optane Pool can give up to 500 MB/s

Now let`s check RAM effects.

5.7 OmniOS OpenZFS iSCSI (AJA vs Atto vs Crystal Benchmark) XL710, 16G RAM

AJA 4k full, 16 bit RGB, 4GB file, sync disabled, compress off, iSCSI LUN, 40G, Jumbo, newest Win/Intel driver

4 x HD Pool

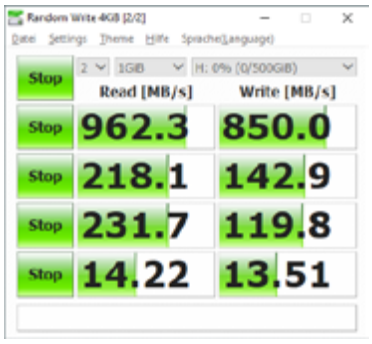
4 x SSD Pool

4 x Optane 900P Pool

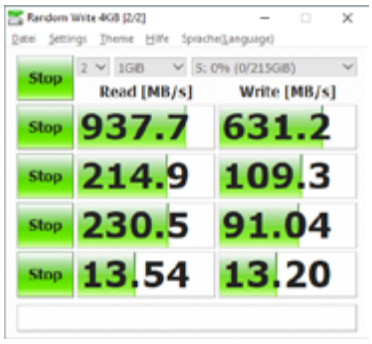
Values are extremely sensitive to Windows driver (use newest Win10 driver from Intel) and settings like int throtteling. If I disable, write values are around 30-40% better while read values are worse. Jumboframes increses performance at around 20%

Crystal Diskmark, sync disabled, compress off, iSCSI LUN

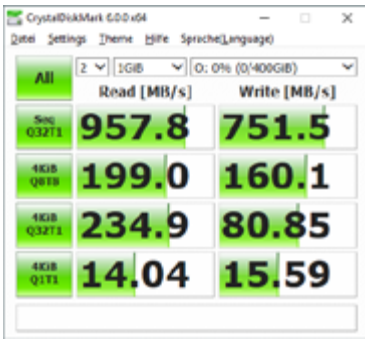
4 x HD Pool



4 x SSD Pool

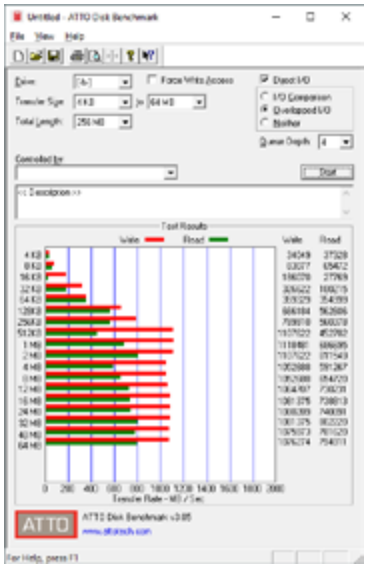


4 x Optane 900P Pool

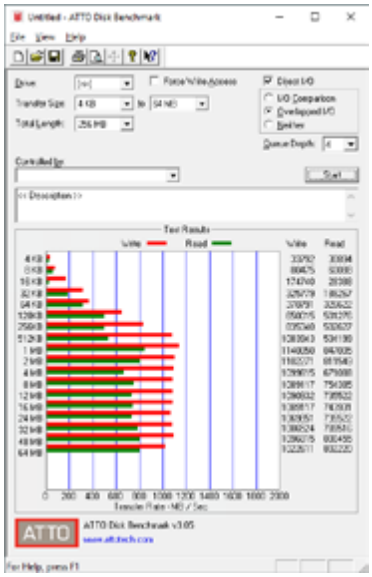


Atto QD2, sync disabled, compress off, iSCSI LUN

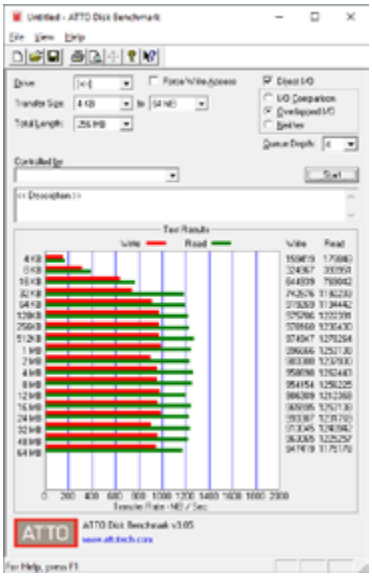
4 x HD Pool



4 x SSD Pool



4 x Optane 900P Pool



The general rersults are quite similar, indicating that RAM on the Server is the limiting factor beside network and drivers. Especially in Atto, Optane shows its quality with very similar write/read values over blocksize.

To verify, I redo all tests with 16G RAM

5.8 OmniOS OpenZFS SMB XL710, 8 G RAM

AJA 4k full, 16 bit RGB, 4GB file, sync disabled, compress off, iSCSI LUN, 40G, Jumbo, newest Win/Intel driver

4 x HD Pool

4 x SSD Pool

4 x Optane 900P Pool

same but with sync=always

This test eleminates the advantage of the rambased write cache. Especially with the HD pool it becomes clear, how bad disks are regarding small random writes. When your workload is sensitive to small random writes or if you want the superiour security of sync writes on a fileserver, tha pure disk pool is a nogo.

If you add a 900p as slog?

If you additionally add a 900P as L2Arc (Prefetch enabled)

This shows the real value of Optane.

With a 900P the HD pools is nearly as fast as the SSD/NVMe pools even with sync enabled. This makes a pure SMB filer with secure sync

5.9 OmniOS OpenZFS SMB XL710, 16 G RAM

AJA 4k full, 16 bit RGB, 4GB file, sync disabled, compress off, iSCSI LUN, 40G, Jumbo, newest Win/Intel driver

4 x HD Pool

4 x SSD Pool

4 x Optane 900P Pool

5.10 OmniOS OpenZFS SMB XL710, 32 G RAM

AJA 4k full, 16 bit RGB, 4GB file, sync disabled, compress off, iSCSI LUN, 40G, Jumbo, newest Win/Intel driver

4 x HD Pool over SMB

4 x SSD Pool

4 x Optane 900P Pool

same but with sync=always

This test eleminates the advantage of the rambased write cache. Especially with the HD pool it becomes clear, how bad disks are regarding small random writes. When your workload is sensitive to small random writes or if you want the superiour security of sync writes on a fileserver, the pure disk pool is a nogo. Optane with enough RAM is here the clear winner. A SMB filer with around 800 MB/s with sync write.

Add a 900p as slog to the HD?

If you additionally add a 900P as L2Arc (Prefetch enabled)

Clear results compared to 8GB

Whenever your workload can be processed mainly by RAM, performance is quite independent from disks. Whenever your workload must use disks directly disk quality matters.

As even a write requires a prior read (at least metadata) write performance is also dependent on RAM.

Secure SMB filers with a HD pool and an Optane Slog is now possible and even fast HD pools for VMs or databases are ok with an Optane Slog.

Add as much RAM as possible.

5.11 Oracle Solaris native ZFS

I cannot publish results due Oracle restrictions.

But I am impressed of the good pool and SMB values on Solaris.
You nmust do your own tests

6.0 Supermicro 4-Port NVme HBA AOC-SLG3-4E4T on a SM X11SPH-nCTF

First option: Attach 4 x Optane to 2 x Oculink and 2 x M.2 ports

hostname	omniosce Memory size: 32429 Megabytes		
pool	optane (recsize=128k, compr=off, readcache=all)		
slog	-		
remark	4 x 900P: 2 x Mainboard OcuLink + 2 x M.2 Adapter		
Fb3 randomwrite.f	sync=always		sync=disabled
	7134 ops		91442 ops
	1426.697 ops/s		18287.516 ops/s
	595us cpu/op		249us cpu/op
	0.7ms latency		0.0ms latency
	11.0 MB/s		142.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	4159 ops		8520 ops
	831.772 ops/s		1703.964 ops/s
	2781us cpu/op		2645us cpu/op
	1.2ms latency		0.6ms latency
	831.6 MB/s		1703.8 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	253.8 MB/s	218.4 MB/s	2.8 GB/s

Second option: 4 x Optane to the NVMe HBA SLG3-4E4T (4 x 4x mode, Oculink inserted in the 16x PCI-e slot) with Mainboard setting auto in Bios > Advanced > Chipset > North > IIO > CPU > IOUO: Auto

hostname	omniosce Memory size: 32429 Megabytes		
pool	optane (recsize=128k, compr=off, readcache=all)		
slog	-		
remark	4 x 900P: 4 x Oculink HBA, 16x slot, 4x4 mode		
Fb3 randomwrite.f	sync=always		sync=disabled
	8977 ops		94523 ops
	1795.286 ops/s		18903.578 ops/s
	534us cpu/op		258us cpu/op
	0.5ms latency		0.0ms latency
	14.0 MB/s		147.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	4188 ops		8516 ops
	837.572 ops/s		1703.120 ops/s
	2884us cpu/op		2575us cpu/op
	1.2ms latency		0.6ms latency
	837.4 MB/s		1702.9 MB/s

---

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	251.6 MB/s	248.8 MB/s	2.8 GB/s

NVMe HBA is working with same performance over the 4 x NVMe Oculink HBA Adapter
So I would not expect problems with this board and up to 10 x U.2 NVMe:

4 x U.2 NVMe connected to a SLG3-4E4T NVMe HBA in the 16x slot
2 x U.2 NVMe connected to a SLG3-2E4T NVMe HBA in the 8x slot

1 x U.2 NVMe connected with a M.2 adapter (4x slot)

2 x U.2 NVMe connected to the two onboard Oculink connectors
1 x U.2 NVMe connected with a M.2 cable (onboard M.2 slot)

6.1 Sata Performance

8 x Sandisk Pro 960 in Raid-0 allcache vs nocache, 96GB RAM, same mainboard as above

hostname	OMNI24 Memory size: 96941 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	3643 ops		95179 ops
	728.563 ops/s		19034.532 ops/s
	966us cpu/op		160us cpu/op
	1.4ms latency		0.0ms latency
	5.6 MB/s		148.6 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	755 ops		8446 ops
	150.992 ops/s		1689.144 ops/s
	5336us cpu/op		3549us cpu/op
	6.6ms latency		0.6ms latency
	150.8 MB/s		1688.9 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	248.0 MB/s	323.4 MB/s	2.8 GB/s

hostname	OMNI24 Memory size: 96941 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1361 ops		2444 ops
	272.176 ops/s		488.763 ops/s
	1953us cpu/op		1609us cpu/op
	3.6ms latency		2.0ms latency
	2.0 MB/s		3.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	783 ops		10651 ops
	156.594 ops/s		2130.115 ops/s
	5962us cpu/op		1701us cpu/op
	6.3ms latency		0.5ms latency
	156.4 MB/s		2129.9 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	4.4 MB/s	10.8 MB/s	251.8 MB/s

6.2 LSI 9003 IR SAS12G Performance

8 x Sandisk Pro 960 in Raid-0 allcache vs nocache, 96GB RAM, same mainboard as above

hostname	OMNI24 Memory size: 96941 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	3109 ops		102136 ops
	621.768 ops/s		20426.626 ops/s
	1058us cpu/op		203us cpu/op
	1.6ms latency		0.0ms latency
	4.8 MB/s		159.4 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	735 ops		10507 ops
	146.991 ops/s		2101.327 ops/s
	5467us cpu/op		1845us cpu/op
	6.7ms latency		0.5ms latency
	146.8 MB/s		2101.1 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	243.0 MB/s	286.6 MB/s	2.8 GB/s

hostname	OMNI24 Memory size: 96941 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1332 ops		2499 ops
	266.387 ops/s		499.765 ops/s
	1920us cpu/op		1590us cpu/op
	3.7ms latency		2.0ms latency
	2.0 MB/s		3.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	783 ops		12949 ops
	156.587 ops/s		2589.720 ops/s
	6255us cpu/op		1593us cpu/op
	6.3ms latency		0.4ms latency
	156.4 MB/s		2589.5 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	4.6 MB/s	11.6 MB/s	422.2 MB/s

6.2 LSI 9305 SAS12G Performance

8 x Sandisk Pro 960 in Raid-0 allcache vs nocache, 96GB RAM, same mainboard as above

hostname	OMNI24 Memory size: 96941 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=all)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	3748 ops		103927 ops
	749.545 ops/s		20784.742 ops/s
	931us cpu/op		203us cpu/op
	1.3ms latency		0.0ms latency
	5.8 MB/s		162.2 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	750 ops		8869 ops
	149.990 ops/s		1773.673 ops/s
	5384us cpu/op		2116us cpu/op
	6.6ms latency		0.6ms latency
	149.8 MB/s		1773.5 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=all	258.8 MB/s	286.4 MB/s	2.8 GB/s

hostname	OMNI24 Memory size: 96941 Megabytes		
pool	ssd (recsize=128k, compr=off, readcache=none)		
slog	-		
remark			
Fb3 randomwrite.f	sync=always		sync=disabled
	1373 ops		2527 ops
	274.588 ops/s		505.362 ops/s
	1950us cpu/op		1311us cpu/op
	3.6ms latency		2.0ms latency
	2.0 MB/s		3.8 MB/s
Fb4 singlestreamwrite.f	sync=always		sync=disabled
	687 ops		12886 ops
	137.391 ops/s		2577.128 ops/s
	6243us cpu/op		1596us cpu/op
	7.2ms latency		0.4ms latency
	137.2 MB/s		2576.9 MB/s

read fb 7-9 + dd (opt)	randomread.f	randomrw.f	singlestreamr
pri/sec cache=none	4.6 MB/s	11.8 MB/s	424.6 MB/s

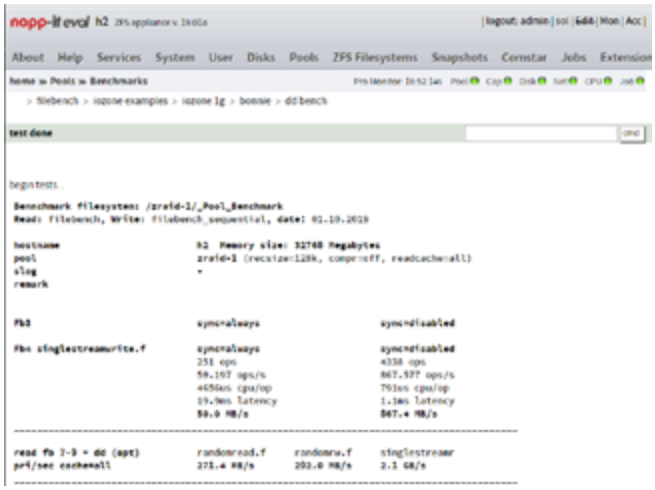
6.3 ATTO ESAH 1288 SAS12G Performance (new: ATTO now supports Illumos)

8 x Sandisk Pro 960 in Raid-0 allcache vs nocache, 96GB RAM, same mainboard as above

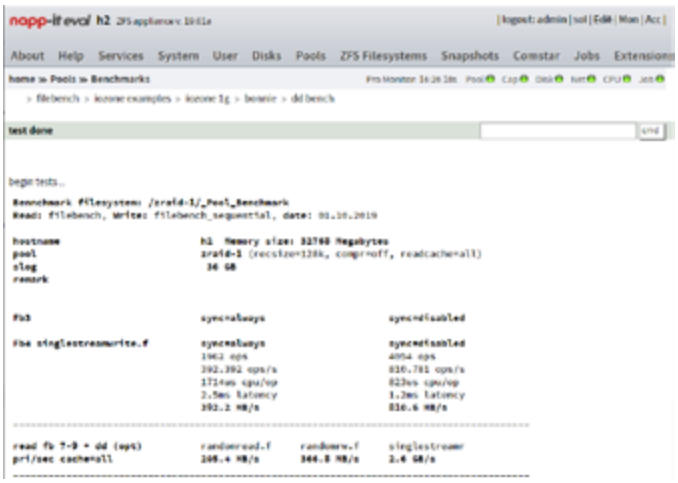
```
Pool layout
       NAME         STATE     READ WRITE CKSUM
       ssd          ONLINE       0     0     0
         c26t0d0    ONLINE       0     0     0
         c26t1d0    ONLINE       0     0     0
         c26t2d0    ONLINE       0     0     0
         c26t3d0    ONLINE       0     0     0
         c26t4d0    ONLINE       0     0     0
         c26t5d0    ONLINE       0     0     0
         c26t6d0    ONLINE       0     0     0
         c26t7d0    ONLINE       0     0     0

hostname                        OMNI24  Memory size: 96941 Megabytes
pool                            ssd (recsize=128k, compr=off, readcache=all)
slog                            -
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                2993 ops                        98072 ops
                                598.556 ops/s                   19613.891 ops/s
                                1142us cpu/op                   205us cpu/op
                                1.6ms latency                   0.0ms latency
                                4.6 MB/s                        153.2 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                760 ops                         11650 ops
                                151.990 ops/s                   2329.931 ops/s
                                5931us cpu/op                   1860us cpu/op
                                6.5ms latency                   0.4ms latency
                                151.8 MB/s                      2329.7 MB/s
_____________________________________________________________________________________

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=all               247.0 MB/s       280.6 MB/s     2.7 GB/s
_____________________________________________________________________________________


hostname                        OMNI24  Memory size: 96941 Megabytes
pool                            ssd (recsize=128k, compr=off, readcache=none)
slog                            -
remark


Fb3 randomwrite.f               sync=always                     sync=disabled
                                1352 ops                        2472 ops
                                270.378 ops/s                   494.363 ops/s
                                2484us cpu/op                   1393us cpu/op
                                3.7ms latency                   2.0ms latency
                                2.0 MB/s                        3.8 MB/s

Fb4 singlestreamwrite.f         sync=always                     sync=disabled
                                743 ops                         13173 ops
                                148.594 ops/s                   2634.522 ops/s
                                5717us cpu/op                   1603us cpu/op
                                6.7ms latency                   0.4ms latency
                                148.4 MB/s                      2634.3 MB/s
_____________________________________________________________________________________

read fb 7-9 + dd (opt)          randomread.f     randomrw.f     singlestreamr
pri/sec cache=none              5.4 MB/s         11.6 MB/s      419.4 MB/s
_____________________________________________________________________________________

Performance of Atto is quite similar with slight advantages in sequential writes
Disk detection is not based on WWN but on controller ports what makes identification easy.
```

6.4 High performance HA SAS storage/ Slog alternative to Optane

You can build a high capacity/ high performance HA storage using dualpath SAS disks that is very fast sequentially and with enough RAM for read/write caching also fast on random loads but is worse on secure sync write.

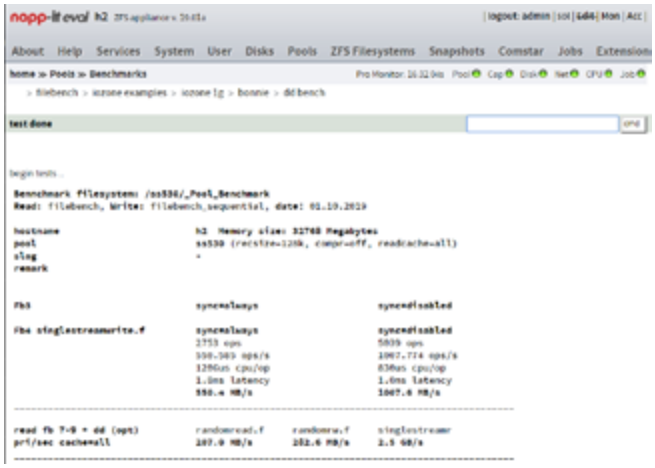


3 disks in a Raid-0 without Slog
Async write: >800 MB/s, sync write: 50 MB/s



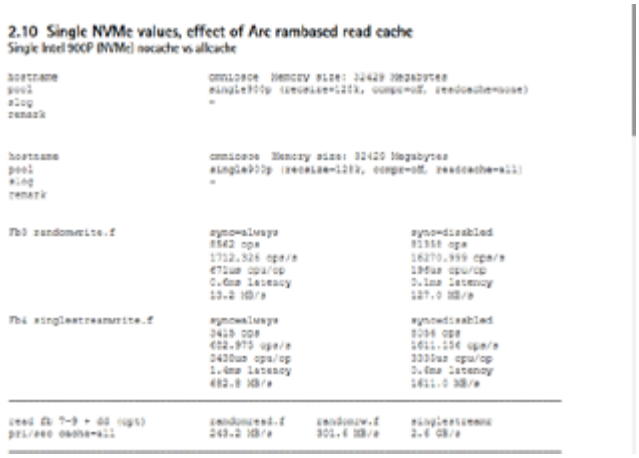
same 3 disks in a Raid-0 with an Slog (SS530)
Async Write > 800 MB/s, sync write nearly 400 MB/s

If you need HA storage with a very high random sync or async write performance, you can build an array from dualpath 12G SAS SSDs WD Ultrastar SS530. This is one of the fastest dualpath enterprise SAS SSDs, see datasheet. These disks are available with a different random write performance/ write endurance, as a 400GB model (use it as an Slog) or up to 15 TB for regular storage. With these SAS SSDs (3DW or 10 DW) you can get around 70-80% of the Intel Optane regarding sync write (currently the best Slog). As Optane is not available as dualpath SAS, the WD SS530 400 GB is a perfect Slog. Use a 12G HBA as on a 6G HBA they are around 20% slower.



A pool from a single DC SS 530 (3DW, 400 GB)
Async write: 1007 MB/s, sync write: 550 MB/s

napp-it AiO under ESXI
using
with disks in pass-through mode



A pool drom a single Intel Optane 900 NVMe
Async Write 1611 MB/s, sync write 680 MB/s

Optane in a barebone setup (there are problems

Optane as pass.through device

Fazit:
WD Ultrastar DC SS 530 is a perfect alternative for Optane when you need dualpath/ HA

7. Supermicro 4-Port NVme HBA AOC-SLG3-4E4T on a SM X11SPH-nCTF

good to know

The X11SPH offers 1 x 16 PCI-Slot that can be used with 1 x 16 lanes. In this case the 8x Slot beside must not be used. If you insert a PCI-e adpater into this 8x slot the 16x HBA reverts to 8x mode so only 2 x Oculink port are active.

Lets hope for Supermicro or others to offer something like a 8x U.2 backplane + 8-16 SAS to allow lowcost high performance systems as this board can the offer 10G network, 8x 12G SAS and up to 9 U.2 NVMe.

U.2/ NVMe Connectors/cables
I prefer OcuLink cables/ connectors as this is NVMe only. The alternative SFF 8643 connectors are also used for SAS so they are not obvious.

Parts for U.2 NVMe Systems

OcuLink NVMe HBA:
SuperMicro AOC-SLG3-4E4T (4x Oculink, PCI-16x adapter),
http://www.supermicro.com/manuals/other/AOC-SLG3-4E4T.pdf

SuperMicro AOC-SLG3-2E4T (2x Oculink, PCI-8x adapter)
http://www.supermicro.com/manuals/other/AOC-SLG3-2E4T.pdf

OcuLink Adaptercable (OcuLink to U.2)
SuperMicro CBL-SAST-0956

2.5" HDD cages for U.2 NVMe
Icy dock MB601VK-B (1 x U.2 NVMe)
Icy dock MB699VP-B (4 x U.2 NVMe)

Backplane
see SuperMicro NVMe solutions

Optionally:
U.2 NVMe connected to M.2 with an Adapter (Intel Optane 900P can be ordered with this adapter)
either to an onboard M.2 slot or to a M.2/ PCI-e adapter

Attention: Solaris
On my Solaris 11.3 system, only 2 of 4 Optane 900P were detected.
On Solaris 11.4 all Optane were detected