SAS disks for a
napp-it ZFS Storageserver

High performance, high capacity,
high number of disks, hotplug

Hardware:
SuperMicro H12SSL-C, Epyc 7302
128 GB RAM, 2 x SAS 3008
passthrough (mpio)

ESXi 7.0U3
Storage VM OmniOS 151038
8 vCores, 80G RAM
Multipath SAS disks

1. Samsung PM 1643a – 960GB
2. Seagate Nitro 3732 – 800GB
3. WD SS530–800 10 DPWD
High performance, high capacity, high number of disks and hotplug capability are quite common demands for storage. For high performance the first thought is NVMe. But when it comes to number of disks say dozens or hundreds or hotplug capability you may consider 2x12G Multipath SAS or 24G SAS in future.

I have tested some SAS disks in a multipath 2 x 12G configuration. Not only from specs but also from benchmarks, the best SAS disks like Seagate Nytro or WD SS530 are near to NVMe performance with a much easier handling and expandability especially in combination with 2 x 12G SAS Multipath that offers over 2000 MB/s throughput. This is 4 times the performance of 6G Sata.

**Specs**

<table>
<thead>
<tr>
<th></th>
<th>Samsung 1643</th>
<th>Seagate 3732</th>
<th>WD SS530</th>
<th>compare: Intel Optane 900</th>
</tr>
</thead>
<tbody>
<tr>
<td>sequ read</td>
<td>2100 MB/s</td>
<td>1100 MB/s</td>
<td>2150 MB/s</td>
<td>2600 MB/s</td>
</tr>
<tr>
<td>sequ write</td>
<td>2000 MB/s</td>
<td>1000 MB/s</td>
<td>2120 MB/s</td>
<td>2200 MB/s</td>
</tr>
<tr>
<td>random read</td>
<td>400k iops</td>
<td>180k iops</td>
<td>440k iops</td>
<td>575k iops</td>
</tr>
<tr>
<td>random write</td>
<td>70k iops</td>
<td>200k iops</td>
<td>320k iops</td>
<td>550k iops</td>
</tr>
<tr>
<td>DWPD</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>MTBF in h</td>
<td>2 Mio</td>
<td>2,5 Mio</td>
<td>2,5 Mio</td>
<td>1,6 Mio</td>
</tr>
<tr>
<td>total write</td>
<td>-</td>
<td>14,6 PB</td>
<td>15,22 PB</td>
<td>-</td>
</tr>
</tbody>
</table>

From results:
- Dualpath SAS improves performance up to 100% compared to singlepath
- The best SAS SSD (WD SS530) is near to the NVMe
- The Seagate Nytro is quite as good as the WD
- Both are much faster than the cheaper Samsung 1643

SAS multipath is worth the effort.
1. Samsung PM 1643a (960GB) Singlepath and Multipath

Benchmark: Write: filebench_sequential, Read: filebench, date: 10.07.2021

pool: tank

<table>
<thead>
<tr>
<th>NAME</th>
<th>STATE</th>
<th>READ</th>
<th>WRITE</th>
<th>CKSUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>tank</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mirror-0</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c3t5002538B007B77B2d0</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c4t5002538B007B77E2d0</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

host: napp-it-038
pool: tank (recsize=128K, ssb=-, compr=off, readcache=all)
slog: -
encryption: -
remark: -

Fb3                     sync=always            sync=disabled
Fb4 fivestreamwrite.f   sync=always            sync=disabled
  2240 ops              5605 ops
  447.275 ops/s         1051.827 ops/s
  13539us cpu/op        9437us cpu/op
  11.0ms latency        4.2ms latency
  446.3 MB/s            1050.9 MB/s

randomread.f   randomrw.f   fivestreamread
pri/sec cache=all                   179.9 MB/s       94.2 MB/s      1.6 GB/s

If something happens, reboot and select a former bootenvironment. Or reinstall napp-it via wget command. This preserves all setting.

same pool with mpio (2 HBA)

Fb3                     sync=always            sync=disabled
Fb4 fivestreamwrite.f   sync=always            sync=disabled
  4640 ops              5673 ops
  927.913 ops/s         1134.497 ops/s
  24536us cpu/op        25328us cpu/op
  5.4ms latency         4.4ms latency
  926.9 MB/s            1133.5 MB/s

randomread.f   randomrw.f   fivestreamread
pri/sec cache=all                   281.2 MB/s       489.2 MB/s      16.0 GB/s
2. Seagate Nytro 3732 – 800mpio

Benchmark: Write: filebench_sequential, Read: filebench, date: 10.07.2021

pool: tank

<table>
<thead>
<tr>
<th>NAME</th>
<th>STATE</th>
<th>READ</th>
<th>WRITE</th>
<th>CKSUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>tank</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mirror-0</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c0t5000C500A19F6007d0</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c0t5000C500A19F6027d0</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

host: napp-it-038

pool: tank (recsize=128K, ssb=-, compr=off, readcache=all)

slog: -

encryption: -

remark: mpio

Fb3

sync=always           sync=disabled

Fb4 fivestreamwrite.f

sync=always           sync=disabled

5536 ops              8440 ops
1107.100 ops/s       1687.618 ops/s
24028us cpu/op       33203us cpu/op
4.5ms latency       2.9ms latency
1106.1 MB/s         1686.6 MB/s

pri/sec cache=all

randomread.f          randomrw.f         fivestreamread

209.2 MB/s           464.3 MB/s         13.6 GB/s
### 3. WD SS530 – 800 10DWPD mpio

**Benchmark:** Write: filebench_sequential, Read: filebench, date: 10.07.2021

```
<table>
<thead>
<tr>
<th>NAME</th>
<th>STATE</th>
<th>READ</th>
<th>WRITE</th>
<th>CKSUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>tank</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mirror-0</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c0t5000CCA0AB743E88d0</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>c0t5000CCA0AB753CE0d0</td>
<td>ONLINE</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**host** napp-it-038
**pool** tank (recsize=128K, ssb=-, compr=off, readcache=all)
**slog** -
**encryption** -
**remark** mpio

**Fb3**

- **sync=always**
- **sync=disabled**

**Fb4** fivestreamwrite.f

- **sync=always**
- **sync=disabled**

<table>
<thead>
<tr>
<th></th>
<th>6316 ops</th>
<th>8509 ops</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1263.157 ops/s</td>
<td>1701.694 ops/s</td>
</tr>
<tr>
<td></td>
<td>24045us cpu/op</td>
<td>32470us cpu/op</td>
</tr>
<tr>
<td>3.9ms latency</td>
<td>2.9ms latency</td>
<td></td>
</tr>
<tr>
<td>1262.2 MB/s</td>
<td>1700.7 MB/s</td>
<td></td>
</tr>
</tbody>
</table>

**pri/sec cache=all**

- **randomread.f** 200.2 MB/s
- **randomrw.f** 472.0 MB/s
- **fivestreamread** 13.4 GB/s