

Memory and Storage System - HW1 FIO

Yi-Ting Shih (111550013)
National Yang Ming Chaio Tung University
ytshih@cs.nycu.edu.tw

1. Specifications

All the read operations except bonus are tested on ssd.

Model: Micron CT1000P3PSSD8[1] (Crucial P3 Plus)

All the write operations are tested on disk files, which are in a RAID 1 btrfs[2] filesystem.

The tested operating system was Arch Linux[3], which was running Linux kernel 6.14.2.arch1-1.

2. Questions

Q0. Example

Result

```
job1: (g=0): rw=read, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync, iodepth=1
fio-3.39
```

Starting 1 process

```
job1: (groupid=0, jobs=1): err= 0: pid=22642: Sat Apr 12 19:28:44 2025
read: IOPS=60.8k, BW=238MiB/s (249MB/s)(1024MiB/4311msec)
  clat (usec): min=9, max=2086, avg=16.20, stdev= 9.84
  lat (usec): min=9, max=2087, avg=16.23, stdev= 9.84
  clat percentiles (nsec):
    | 1.00th=[11712], 5.00th=[11840], 10.00th=[12096], 20.00th=[12224],
    | 30.00th=[12352], 40.00th=[12352], 50.00th=[12480], 60.00th=[12736],
    | 70.00th=[13504], 80.00th=[14912], 90.00th=[25728], 95.00th=[36608],
    | 99.00th=[48896], 99.50th=[64256], 99.90th=[71168], 99.95th=[71168],
    | 99.99th=[88576]
  bw ( KiB/s): min=210008, max=250232, per=99.85%, avg=242862.00, stdev=13417.57, samples=8
  iops        : min=52502, max=62558, avg=60715.50, stdev=3354.39, samples=8
  lat (usec)  : 10=0.03%, 20=81.94%, 50=17.13%, 100=0.90%, 250=0.01%
  lat (usec)  : 500=0.01%, 750=0.01%
  lat (msec)  : 4=0.01%
  cpu         : usr=2.39%, sys=7.68%, ctx=262148, majf=0, minf=9
  IO depths   : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
    submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    issued rwts: total=262144,0,0,0 short=0,0,0,0 dropped=0,0,0,0
    latency   : target=0, window=0, percentile=100.00%, depth=1
```

Run status group 0 (all jobs):

READ: bw=238MiB/s (249MB/s), 238MiB/s-238MiB/s (249MB/s-249MB/s), io=1024MiB (1074MB), run=4311-4311msec

Disk stats (read/write):

nvme1n1: ios=251884/14, sectors=2015072/536, merge=0/0, ticks=3534/6, in_queue=3543, util=81.46%

Q1. Read vs. Randread

Result

```
read: (g=0): rw=read, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync, iodepth=1
```

```
randread: (g=1): rw=randread, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync, iodepth=1
fio-3.39
```

Starting 2 processes

```
read: (groupid=0, jobs=1): err= 0: pid=23955: Sat Apr 12 19:36:41 2025
read: IOPS=62.1k, BW=242MiB/s (254MB/s)(64.0MiB/264msec)
  clat (usec): min=10, max=913, avg=15.86, stdev=11.07
  lat (usec): min=10, max=913, avg=15.89, stdev=11.07
  clat percentiles (nsec):
    | 1.00th=[10176], 5.00th=[11840], 10.00th=[12096], 20.00th=[12224],
    | 30.00th=[12224], 40.00th=[12224], 50.00th=[12352], 60.00th=[12480],
    | 70.00th=[12992], 80.00th=[14016], 90.00th=[25472], 95.00th=[37632],
    | 99.00th=[48896], 99.50th=[51456], 99.90th=[65280], 99.95th=[69120],
```

```

| 99.99th=[93696]
lat (usec) : 20=83.84%, 50=15.50%, 100=0.66%, 1000=0.01%
cpu : usr=3.42%, sys=6.84%, ctx=16384, majf=0, minf=9
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=16384,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1
randread: (groupid=1, jobs=1): err= 0: pid=23956: Sat Apr 12 19:36:41 2025
read: IOPS=15.4k, BW=60.0MiB/s (62.9MB/s)(64.0MiB/1067msec)
clat (usec): min=10, max=2157, avg=64.80, stdev=22.68
lat (usec): min=10, max=2157, avg=64.83, stdev=22.68
clat percentiles (usec):
| 1.00th=[ 11], 5.00th=[ 63], 10.00th=[ 65], 20.00th=[ 65],
| 30.00th=[ 65], 40.00th=[ 65], 50.00th=[ 65], 60.00th=[ 65],
| 70.00th=[ 70], 80.00th=[ 71], 90.00th=[ 72], 95.00th=[ 72],
| 99.00th=[ 73], 99.50th=[ 75], 99.90th=[ 113], 99.95th=[ 165],
| 99.99th=[ 914]
bw ( KiB/s): min=61312, max=61784, per=100.00%, avg=61548.00, stdev=333.75, samples=2
iops : min=15328, max=15446, avg=15387.00, stdev=83.44, samples=2
lat (usec) : 20=4.03%, 100=95.65%, 250=0.27%, 500=0.02%, 1000=0.01%
lat (msec) : 4=0.01%
cpu : usr=0.66%, sys=2.06%, ctx=16390, majf=0, minf=8
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=16384,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1

Run status group 0 (all jobs):
READ: bw=242MiB/s (254MB/s), 242MiB/s-242MiB/s (254MB/s-254MB/s), io=64.0MiB (67.1MB), run=264-264msec

Run status group 1 (all jobs):
READ: bw=60.0MiB/s (62.9MB/s), 60.0MiB/s-60.0MiB/s (62.9MB/s-62.9MB/s), io=64.0MiB (67.1MB),
run=1067-1067msec

Disk stats (read/write):
nvme1n1: ios=30190/5, sectors=241528/376, merge=0/0, ticks=1083/69, in_queue=1153, util=72.01%

```

Question

Is there any significant difference between read and randread? Why or why not? Please justify your answer in brief.

Answer

read is significantly faster than *randread*.

The SSD controller might be optimized for sequential access, such as leveraging internal parallelism more effectively, or have lower command overhead per unit of data for sequential reads.

Q2. Write vs. Randwrite

Result

```

write: (g=0): rw=write, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync, iodepth=1
randwrite: (g=1): rw=randwrite, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync, iodepth=1
fio-3.39
Starting 2 processes

```

```

write: (groupid=0, jobs=1): err= 0: pid=12418: Sat Apr 12 18:34:15 2025
write: IOPS=33.1k, BW=129MiB/s (136MB/s)(64.0MiB/495msec); 0 zone resets
clat (usec): min=18, max=2971, avg=29.41, stdev=26.49
lat (usec): min=19, max=2972, avg=29.55, stdev=26.49
clat percentiles (usec):
| 1.00th=[ 24], 5.00th=[ 26], 10.00th=[ 26], 20.00th=[ 27],
| 30.00th=[ 27], 40.00th=[ 28], 50.00th=[ 28], 60.00th=[ 28],
| 70.00th=[ 29], 80.00th=[ 31], 90.00th=[ 32], 95.00th=[ 32],
| 99.00th=[ 118], 99.50th=[ 131], 99.90th=[ 208], 99.95th=[ 225],
| 99.99th=[ 619]
lat (usec) : 20=0.01%, 50=98.83%, 100=0.10%, 250=1.01%, 500=0.03%
lat (usec) : 750=0.01%
lat (msec) : 4=0.01%
cpu : usr=3.04%, sys=36.44%, ctx=37026, majf=0, minf=6
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=0,16384,0,0 short=0,0,0,0 dropped=0,0,0,0

```

```

latency : target=0, window=0, percentile=100.00%, depth=1
randwrite: (groupid=1, jobs=1): err= 0: pid=12423: Sat Apr 12 18:34:15 2025
write: IOPS=39.7k, BW=155MiB/s (162MB/s)(64.0MiB/413msec); 0 zone resets
clat (usec): min=16, max=33415, avg=24.41, stdev=261.81
lat (usec): min=16, max=33415, avg=24.53, stdev=261.81
clat percentiles (usec):
| 1.00th=[ 18], 5.00th=[ 22], 10.00th=[ 22], 20.00th=[ 23],
| 30.00th=[ 23], 40.00th=[ 23], 50.00th=[ 23], 60.00th=[ 23],
| 70.00th=[ 23], 80.00th=[ 23], 90.00th=[ 23], 95.00th=[ 23],
| 99.00th=[ 25], 99.50th=[ 25], 99.90th=[ 28], 99.95th=[ 31],
| 99.99th=[ 2835]
lat (usec) : 20=3.01%, 50=96.97%, 250=0.01%
lat (msec) : 4=0.01%, 50=0.01%
cpu : usr=4.13%, sys=31.55%, ctx=16389, majf=0, minf=7
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=0,16384,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1

```

Run status group 0 (all jobs):

```
WRITE: bw=129MiB/s (136MB/s), 129MiB/s-129MiB/s (136MB/s-136MB/s), io=64.0MiB (67.1MB), run=495-495msec
```

Run status group 1 (all jobs):

```
WRITE: bw=155MiB/s (162MB/s), 155MiB/s-155MiB/s (162MB/s-162MB/s), io=64.0MiB (67.1MB), run=413-413msec
```

Question

Is there any significant difference between write and randwrite? Why or why not? Please justify your answer in brief.

Answer

There is no significant difference between *write* and *randwrite*.

The SSD controller will go through the page / block translation process no matter there is write or random write access. Therefore there is no significant difference.

Also, I use disk file to test the write operation, which is on a btrfs filesystem. There are cache on btrfs[2] or page cache so the data might not be actually write to the SSD.

Q3. Forward write vs. Backward write

Result

```

forward: (g=0): rw=write, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync, iodepth=1
backward: (g=1): rw=write, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync, iodepth=1
fio-3.39
Starting 2 processes

```

forward: (groupid=0, jobs=1): err= 0: pid=12464: Sat Apr 12 18:34:17 2025

```
write: IOPS=42.8k, BW=167MiB/s (175MB/s)(64.0MiB/383msec); 0 zone resets
```

```
clat (usec): min=15, max=2975, avg=22.54, stdev=23.29
```

```
lat (usec): min=16, max=2976, avg=22.71, stdev=23.30
```

```
clat percentiles (usec):
```

```

| 1.00th=[ 17], 5.00th=[ 18], 10.00th=[ 22], 20.00th=[ 23],
| 30.00th=[ 23], 40.00th=[ 23], 50.00th=[ 23], 60.00th=[ 23],
| 70.00th=[ 23], 80.00th=[ 23], 90.00th=[ 24], 95.00th=[ 24],
| 99.00th=[ 28], 99.50th=[ 37], 99.90th=[ 48], 99.95th=[ 58],
| 99.99th=[ 322]

```

```
lat (usec) : 20=5.47%, 50=94.43%, 100=0.08%, 500=0.01%
```

```
lat (msec) : 4=0.01%
```

```
cpu : usr=4.19%, sys=31.41%, ctx=16387, majf=0, minf=6
```

```
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
```

```
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
```

```
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
```

```
issued rwts: total=0,16384,0,0 short=0,0,0,0 dropped=0,0,0,0
```

```
latency : target=0, window=0, percentile=100.00%, depth=1
```

backward: (groupid=1, jobs=1): err= 0: pid=12474: Sat Apr 12 18:34:17 2025

```
write: IOPS=41.6k, BW=162MiB/s (170MB/s)(64.0MiB/394msec); 0 zone resets
```

```
clat (usec): min=15, max=726, avg=23.17, stdev=10.51
```

```
lat (usec): min=15, max=727, avg=23.34, stdev=10.51
```

```
clat percentiles (usec):
```

```

| 1.00th=[ 17], 5.00th=[ 22], 10.00th=[ 22], 20.00th=[ 23],
| 30.00th=[ 23], 40.00th=[ 23], 50.00th=[ 23], 60.00th=[ 23],
| 70.00th=[ 23], 80.00th=[ 24], 90.00th=[ 26], 95.00th=[ 26],
| 99.00th=[ 38], 99.50th=[ 41], 99.90th=[ 59], 99.95th=[ 85],
| 99.99th=[ 693]

```

```

lat (usec)   : 20=4.46%, 50=95.36%, 100=0.14%, 250=0.02%, 750=0.02%
cpu          : usr=4.33%, sys=32.82%, ctx=16389, majf=0, minf=7
IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit       : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete     : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=0,16384,0,0 short=0,0,0,0 dropped=0,0,0,0
latency      : target=0, window=0, percentile=100.00%, depth=1

```

Run status group 0 (all jobs):

```
WRITE: bw=167MiB/s (175MB/s), 167MiB/s-167MiB/s (175MB/s-175MB/s), io=64.0MiB (67.1MB), run=383-383msec
```

Run status group 1 (all jobs):

```
WRITE: bw=162MiB/s (170MB/s), 162MiB/s-162MiB/s (170MB/s-170MB/s), io=64.0MiB (67.1MB), run=394-394msec
```

Question

Is there any significant difference between forward and backward write? Why or why not? Please justify your answer in brief.

Answer

There is no significant difference between *forward write* and *backward write*.

The reason is same as the previous question.

Q4. Buffered read vs. Nonbuffered read

Result

```

sequential_buffered: (g=0): rw=read, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync,
iodepth=1
sequential_nonbuffered: (g=1): rw=read, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync,
iodepth=1
rand_buffered: (g=2): rw=randread, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync,
iodepth=1
rand_nonbuffered: (g=3): rw=randread, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync,
iodepth=1
fio-3.39
Starting 4 processes

```

sequential_buffered: (groupid=0, jobs=1): err= 0: pid=225892: Sun Apr 13 16:32:18 2025

```
read: IOPS=364k, BW=1422MiB/s (1491MB/s)(64.0MiB/45msec)
```

```
clat (nsec): min=414, max=244928, avg=2552.07, stdev=17447.08
```

```
lat (nsec): min=434, max=244948, avg=2573.40, stdev=17447.28
```

```
clat percentiles (nsec):
```

```

| 1.00th=[ 426], 5.00th=[ 434], 10.00th=[ 438], 20.00th=[ 446],
| 30.00th=[ 450], 40.00th=[ 454], 50.00th=[ 462], 60.00th=[ 474],
| 70.00th=[ 490], 80.00th=[ 506], 90.00th=[ 556], 95.00th=[ 652],
| 99.00th=[118272], 99.50th=[154624], 99.90th=[211968], 99.95th=[224256],
| 99.99th=[238592]

```

```
lat (nsec)   : 500=77.34%, 750=19.23%, 1000=1.57%
```

```
lat (usec)   : 2=0.21%, 4=0.06%, 10=0.01%, 20=0.01%, 50=0.12%
```

```
lat (usec)   : 100=0.31%, 250=1.15%
```

```
cpu          : usr=6.82%, sys=34.09%, ctx=246, majf=0, minf=8
```

```
IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
```

```
submit       : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
```

```
complete     : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
```

```
issued rwts: total=16384,0,0,0 short=0,0,0,0 dropped=0,0,0,0
```

```
latency      : target=0, window=0, percentile=100.00%, depth=1
```

sequential_nonbuffered: (groupid=1, jobs=1): err= 0: pid=225893: Sun Apr 13 16:32:18 2025

```
read: IOPS=35.0k, BW=137MiB/s (143MB/s)(64.0MiB/468msec)
```

```
clat (usec): min=10, max=310, avg=28.03, stdev=24.76
```

```
lat (usec): min=10, max=310, avg=28.07, stdev=24.76
```

```
clat percentiles (usec):
```

```

| 1.00th=[ 11], 5.00th=[ 13], 10.00th=[ 13], 20.00th=[ 13],
| 30.00th=[ 13], 40.00th=[ 13], 50.00th=[ 14], 60.00th=[ 16],
| 70.00th=[ 25], 80.00th=[ 64], 90.00th=[ 68], 95.00th=[ 73],
| 99.00th=[ 85], 99.50th=[ 122], 99.90th=[ 182], 99.95th=[ 194],
| 99.99th=[ 239]

```

```
lat (usec)   : 20=65.66%, 50=11.58%, 100=22.11%, 250=0.65%, 500=0.01%
```

```
cpu          : usr=2.36%, sys=5.35%, ctx=16388, majf=0, minf=8
```

```
IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
```

```
submit       : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
```

```
complete     : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
```

```
issued rwts: total=16384,0,0,0 short=0,0,0,0 dropped=0,0,0,0
```

```
latency      : target=0, window=0, percentile=100.00%, depth=1
```

rand_buffered: (groupid=2, jobs=1): err= 0: pid=225899: Sun Apr 13 16:32:18 2025

```

read: IOPS=14.8k, BW=57.9MiB/s (60.7MB/s)(64.0MiB/1105msec)
  clat (usec): min=11, max=740, avg=67.02, stdev=18.23
    lat (usec): min=11, max=740, avg=67.06, stdev=18.23
  clat percentiles (usec):
    | 1.00th=[ 12], 5.00th=[ 64], 10.00th=[ 65], 20.00th=[ 67],
    | 30.00th=[ 67], 40.00th=[ 67], 50.00th=[ 68], 60.00th=[ 68],
    | 70.00th=[ 72], 80.00th=[ 73], 90.00th=[ 74], 95.00th=[ 75],
    | 99.00th=[ 92], 99.50th=[ 106], 99.90th=[ 351], 99.95th=[ 429],
    | 99.99th=[ 586]
  bw ( KiB/s): min=54936, max=58928, per=95.99%, avg=56932.00, stdev=2822.77, samples=2
  iops       : min=13734, max=14732, avg=14233.00, stdev=705.69, samples=2
  lat (usec) : 20=4.00%, 50=0.02%, 100=95.34%, 250=0.51%, 500=0.12%
  lat (usec) : 750=0.02%
  cpu        : usr=0.91%, sys=3.71%, ctx=16386, majf=0, minf=7
  IO depths  : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
    submit   : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    issued rwts: total=16384,0,0,0 short=0,0,0,0 dropped=0,0,0,0
    latency   : target=0, window=0, percentile=100.00%, depth=1
rand_nonbuffered: (groupid=3, jobs=1): err= 0: pid=225901: Sun Apr 13 16:32:18 2025
read: IOPS=15.1k, BW=59.1MiB/s (62.0MB/s)(64.0MiB/1082msec)
  clat (usec): min=10, max=161, avg=65.44, stdev=11.95
    lat (usec): min=10, max=161, avg=65.48, stdev=11.95
  clat percentiles (usec):
    | 1.00th=[ 11], 5.00th=[ 63], 10.00th=[ 64], 20.00th=[ 65],
    | 30.00th=[ 65], 40.00th=[ 66], 50.00th=[ 67], 60.00th=[ 68],
    | 70.00th=[ 71], 80.00th=[ 72], 90.00th=[ 73], 95.00th=[ 74],
    | 99.00th=[ 86], 99.50th=[ 94], 99.90th=[ 112], 99.95th=[ 115],
    | 99.99th=[ 153]
  bw ( KiB/s): min=60152, max=60984, per=100.00%, avg=60568.00, stdev=588.31, samples=2
  iops       : min=15038, max=15246, avg=15142.00, stdev=147.08, samples=2
  lat (usec) : 20=4.01%, 50=0.02%, 100=95.66%, 250=0.31%
  cpu        : usr=1.20%, sys=2.41%, ctx=16388, majf=0, minf=7
  IO depths  : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
    submit   : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    issued rwts: total=16384,0,0,0 short=0,0,0,0 dropped=0,0,0,0
    latency   : target=0, window=0, percentile=100.00%, depth=1

Run status group 0 (all jobs):
  READ: bw=1422MiB/s (1491MB/s), 1422MiB/s-1422MiB/s (1491MB/s-1491MB/s), io=64.0MiB (67.1MB), run=45-45msec

Run status group 1 (all jobs):
  READ: bw=137MiB/s (143MB/s), 137MiB/s-137MiB/s (143MB/s-143MB/s), io=64.0MiB (67.1MB), run=468-468msec

Run status group 2 (all jobs):
  READ: bw=57.9MiB/s (60.7MB/s), 57.9MiB/s-57.9MiB/s (60.7MB/s-60.7MB/s), io=64.0MiB (67.1MB),
run=1105-1105msec

Run status group 3 (all jobs):
  READ: bw=59.1MiB/s (62.0MB/s), 59.1MiB/s-59.1MiB/s (62.0MB/s-62.0MB/s), io=64.0MiB (67.1MB),
run=1082-1082msec

Disk stats (read/write):
  nvme1n1: ios=60308/601, sectors=612976/18856, merge=67/21, ticks=3076/2672, in_queue=5750, util=75.21%

```

Question

1. Is there any significant difference between buffered and nonbuffered sequential read? Why or why not? Please justify your answer in brief.
2. Replace sequential read with random read. Is there any significant difference between buffered and nonbuffered random read? Why or why not? Please justify your answer in brief.

Answer

- *buffered sequential read* is significantly faster than *non-buffered sequential read*.
- There is no significant difference between *buffered random read* and *non-buffered random read*.

The data can be buffered by *read-ahead* for sequential read operation. This cannot be done for random read operation, because the next read operation cannot be predicted.

Q5. LBA

Result

```
job1: (g=0): rw=read, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=psync,
iodepth=1
job2: (g=1): rw=read, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=psync,
iodepth=1
job3: (g=2): rw=read, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=psync,
iodepth=1
job4: (g=3): rw=read, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=psync,
iodepth=1
job5: (g=4): rw=read, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=psync,
iodepth=1
fio-3.39
Starting 5 processes
```

```
job1: (groupid=0, jobs=1): err= 0: pid=307878: Mon Apr 14 07:49:52 2025
read: IOPS=2985, BW=2985MiB/s (3130MB/s)(1024MiB/343msec)
  clat (usec): min=328, max=1158, avg=333.33, stdev=26.18
    lat (usec): min=328, max=1158, avg=333.41, stdev=26.19
  clat percentiles (usec):
    | 1.00th=[ 330],  5.00th=[ 330], 10.00th=[ 330], 20.00th=[ 330],
    | 30.00th=[ 330], 40.00th=[ 330], 50.00th=[ 330], 60.00th=[ 334],
    | 70.00th=[ 334], 80.00th=[ 334], 90.00th=[ 334], 95.00th=[ 338],
    | 99.00th=[ 351], 99.50th=[ 371], 99.90th=[ 400], 99.95th=[ 1156],
    | 99.99th=[ 1156]
  lat (usec)   : 500=99.90%
  lat (msec)   : 2=0.10%
  cpu          : usr=0.58%, sys=7.31%, ctx=1037, majf=0, minf=264
  IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
    submit     : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    complete   : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    issued rwts: total=1024,0,0,0 short=0,0,0,0 dropped=0,0,0,0
    latency    : target=0, window=0, percentile=100.00%, depth=1
job2: (groupid=1, jobs=1): err= 0: pid=307879: Mon Apr 14 07:49:52 2025
read: IOPS=2737, BW=2738MiB/s (2871MB/s)(1024MiB/374msec)
  clat (usec): min=327, max=2373, avg=363.76, stdev=97.13
    lat (usec): min=327, max=2373, avg=363.84, stdev=97.13
  clat percentiles (usec):
    | 1.00th=[ 330],  5.00th=[ 330], 10.00th=[ 330], 20.00th=[ 330],
    | 30.00th=[ 330], 40.00th=[ 330], 50.00th=[ 330], 60.00th=[ 334],
    | 70.00th=[ 334], 80.00th=[ 343], 90.00th=[ 490], 95.00th=[ 515],
    | 99.00th=[ 553], 99.50th=[ 586], 99.90th=[ 1500], 99.95th=[ 2376],
    | 99.99th=[ 2376]
  lat (usec)   : 500=91.89%, 750=7.91%
  lat (msec)   : 2=0.10%, 4=0.10%
  cpu          : usr=0.27%, sys=6.97%, ctx=1024, majf=0, minf=264
  IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
    submit     : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    complete   : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    issued rwts: total=1024,0,0,0 short=0,0,0,0 dropped=0,0,0,0
    latency    : target=0, window=0, percentile=100.00%, depth=1
job3: (groupid=2, jobs=1): err= 0: pid=307880: Mon Apr 14 07:49:52 2025
read: IOPS=1996, BW=1996MiB/s (2093MB/s)(1024MiB/513msec)
  clat (usec): min=330, max=1471, avg=499.92, stdev=66.04
    lat (usec): min=330, max=1472, avg=500.01, stdev=66.05
  clat percentiles (usec):
    | 1.00th=[ 334],  5.00th=[ 461], 10.00th=[ 469], 20.00th=[ 478],
    | 30.00th=[ 486], 40.00th=[ 494], 50.00th=[ 498], 60.00th=[ 502],
    | 70.00th=[ 510], 80.00th=[ 523], 90.00th=[ 537], 95.00th=[ 545],
    | 99.00th=[ 652], 99.50th=[ 709], 99.90th=[ 1270], 99.95th=[ 1467],
    | 99.99th=[ 1467]
  lat (usec)   : 500=53.81%, 750=45.70%, 1000=0.10%
  lat (msec)   : 2=0.39%
  cpu          : usr=0.78%, sys=4.49%, ctx=1025, majf=0, minf=262
  IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
    submit     : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    complete   : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    issued rwts: total=1024,0,0,0 short=0,0,0,0 dropped=0,0,0,0
    latency    : target=0, window=0, percentile=100.00%, depth=1
job4: (groupid=3, jobs=1): err= 0: pid=307881: Mon Apr 14 07:49:52 2025
read: IOPS=2000, BW=2000MiB/s (2097MB/s)(1024MiB/512msec)
  clat (usec): min=337, max=1452, avg=499.02, stdev=39.01
    lat (usec): min=338, max=1452, avg=499.10, stdev=39.01
  clat percentiles (usec):
    | 1.00th=[ 461],  5.00th=[ 465], 10.00th=[ 469], 20.00th=[ 478],
```

```

| 30.00th=[ 486], 40.00th=[ 490], 50.00th=[ 498], 60.00th=[ 502],
| 70.00th=[ 506], 80.00th=[ 519], 90.00th=[ 537], 95.00th=[ 545],
| 99.00th=[ 570], 99.50th=[ 603], 99.90th=[ 644], 99.95th=[ 1450],
| 99.99th=[ 1450]
lat (usec) : 500=60.16%, 750=39.75%
lat (msec) : 2=0.10%
cpu : usr=0.00%, sys=5.28%, ctx=1024, majf=0, minf=262
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=1024,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1
job5: (groupid=4, jobs=1): err= 0: pid=307882: Mon Apr 14 07:49:52 2025
read: IOPS=2000, BW=2000MiB/s (2097MB/s)(1024MiB/512msec)
clat (usec): min=369, max=1216, avg=499.36, stdev=33.86
lat (usec): min=369, max=1216, avg=499.45, stdev=33.87
clat percentiles (usec):
| 1.00th=[ 429], 5.00th=[ 465], 10.00th=[ 469], 20.00th=[ 478],
| 30.00th=[ 486], 40.00th=[ 494], 50.00th=[ 498], 60.00th=[ 502],
| 70.00th=[ 506], 80.00th=[ 519], 90.00th=[ 537], 95.00th=[ 545],
| 99.00th=[ 553], 99.50th=[ 586], 99.90th=[ 627], 99.95th=[ 1221],
| 99.99th=[ 1221]
bw ( MiB/s): min= 2002, max= 2002, per=100.00%, avg=2002.00, stdev= 0.00, samples=1
iops : min= 2002, max= 2002, avg=2002.00, stdev= 0.00, samples=1
lat (usec) : 500=56.45%, 750=43.46%
lat (msec) : 2=0.10%
cpu : usr=0.00%, sys=5.09%, ctx=1025, majf=0, minf=264
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=1024,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1

Run status group 0 (all jobs):
READ: bw=2985MiB/s (3130MB/s), 2985MiB/s-2985MiB/s (3130MB/s-3130MB/s), io=1024MiB (1074MB), run=343-343msec

Run status group 1 (all jobs):
READ: bw=2738MiB/s (2871MB/s), 2738MiB/s-2738MiB/s (2871MB/s-2871MB/s), io=1024MiB (1074MB), run=374-374msec

Run status group 2 (all jobs):
READ: bw=1996MiB/s (2093MB/s), 1996MiB/s-1996MiB/s (2093MB/s-2093MB/s), io=1024MiB (1074MB), run=513-513msec

Run status group 3 (all jobs):
READ: bw=2000MiB/s (2097MB/s), 2000MiB/s-2000MiB/s (2097MB/s-2097MB/s), io=1024MiB (1074MB), run=512-512msec

Run status group 4 (all jobs):
READ: bw=2000MiB/s (2097MB/s), 2000MiB/s-2000MiB/s (2097MB/s-2097MB/s), io=1024MiB (1074MB), run=512-512msec

Disk stats (read/write):
nvme1n1: ios=19144/231, sectors=9801728/7256, merge=0/4, ticks=6025/353, in_queue=6382, util=61.60%

```

Question

Is there any bandwidth trend between these jobs? Why or why not? Please justify your answer in brief. You can also experiment with the size of the file.

Answer

The job where offset=80% is faster than offset=60%, and offset=60% is faster than the rest of the jobs.

The identified response presents an analysis of experimental results; the underlying theoretical framework remains unclear, necessitating further investigation.

Q6. Blocksize

Result

```

1k_block: (g=0): rw=read, bs=(R) 1024B-1024B, (W) 1024B-1024B, (T) 1024B-1024B, ioengine=psync, iodepth=1
4k_block: (g=1): rw=read, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=psync, iodepth=1
1m_block: (g=2): rw=read, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=psync,
iodepth=1
2m_block: (g=3): rw=read, bs=(R) 2048KiB-2048KiB, (W) 2048KiB-2048KiB, (T) 2048KiB-2048KiB, ioengine=psync,
iodepth=1
4m_block: (g=4): rw=read, bs=(R) 4096KiB-4096KiB, (W) 4096KiB-4096KiB, (T) 4096KiB-4096KiB, ioengine=psync,
iodepth=1
8m_block: (g=5): rw=read, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB, ioengine=psync,
iodepth=1

```

16m_block: (g=6): rw=read, bs=(R) 16.0MiB-16.0MiB, (W) 16.0MiB-16.0MiB, (T) 16.0MiB-16.0MiB, ioengine=psync, iodepth=1
32m_block: (g=7): rw=read, bs=(R) 32.0MiB-32.0MiB, (W) 32.0MiB-32.0MiB, (T) 32.0MiB-32.0MiB, ioengine=psync, iodepth=1
fio-3.39
Starting 8 processes

1k_block: (groupid=0, jobs=1): err= 0: pid=310156: Mon Apr 14 08:12:27 2025

read: IOPS=61.9k, BW=60.5MiB/s (63.4MB/s)(64.0MiB/1058msec)
clat (usec): min=9, max=828, avg=15.88, stdev= 6.29
lat (usec): min=9, max=828, avg=15.92, stdev= 6.29
clat percentiles (usec):
| 1.00th=[10], 5.00th=[15], 10.00th=[15], 20.00th=[15],
| 30.00th=[15], 40.00th=[15], 50.00th=[15], 60.00th=[16],
| 70.00th=[16], 80.00th=[17], 90.00th=[18], 95.00th=[24],
| 99.00th=[25], 99.50th=[25], 99.90th=[67], 99.95th=[137],
| 99.99th=[176]
bw (KiB/s): min=62446, max=62446, per=100.00%, avg=62446.00, stdev= 0.00, samples=1
iops : min=62446, max=62446, avg=62446.00, stdev= 0.00, samples=1
lat (usec) : 10=3.78%, 20=89.12%, 50=6.87%, 100=0.15%, 250=0.07%
lat (usec) : 500=0.01%, 1000=0.01%
cpu : usr=3.12%, sys=8.33%, ctx=65539, majf=0, minf=9
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=65536,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1

4k_block: (groupid=1, jobs=1): err= 0: pid=310157: Mon Apr 14 08:12:27 2025

read: IOPS=61.4k, BW=240MiB/s (251MB/s)(64.0MiB/267msec)
clat (usec): min=10, max=893, avg=15.99, stdev=11.09
lat (usec): min=10, max=893, avg=16.02, stdev=11.09
clat percentiles (usec):
| 1.00th=[11], 5.00th=[12], 10.00th=[13], 20.00th=[13],
| 30.00th=[13], 40.00th=[13], 50.00th=[13], 60.00th=[13],
| 70.00th=[14], 80.00th=[15], 90.00th=[26], 95.00th=[38],
| 99.00th=[49], 99.50th=[50], 99.90th=[67], 99.95th=[133],
| 99.99th=[182]
lat (usec) : 20=83.81%, 50=15.75%, 100=0.38%, 250=0.05%, 1000=0.01%
cpu : usr=2.63%, sys=8.27%, ctx=16384, majf=0, minf=8
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=16384,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1

1m_block: (groupid=2, jobs=1): err= 0: pid=310160: Mon Apr 14 08:12:27 2025

read: IOPS=2000, BW=2000MiB/s (2097MB/s)(64.0MiB/32msec)
clat (usec): min=372, max=1191, avg=494.23, stdev=91.47
lat (usec): min=372, max=1191, avg=494.31, stdev=91.49
clat percentiles (usec):
| 1.00th=[371], 5.00th=[461], 10.00th=[461], 20.00th=[469],
| 30.00th=[474], 40.00th=[478], 50.00th=[482], 60.00th=[490],
| 70.00th=[494], 80.00th=[502], 90.00th=[515], 95.00th=[523],
| 99.00th=[1188], 99.50th=[1188], 99.90th=[1188], 99.95th=[1188],
| 99.99th=[1188]
lat (usec) : 500=79.69%, 750=18.75%
lat (msec) : 2=1.56%
cpu : usr=0.00%, sys=6.45%, ctx=64, majf=0, minf=262
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=64,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1

2m_block: (groupid=3, jobs=1): err= 0: pid=310161: Mon Apr 14 08:12:27 2025

read: IOPS=1280, BW=2560MiB/s (2684MB/s)(64.0MiB/25msec)
clat (usec): min=642, max=979, avg=763.00, stdev=45.56
lat (usec): min=642, max=980, avg=763.10, stdev=45.58
clat percentiles (usec):
| 1.00th=[644], 5.00th=[725], 10.00th=[750], 20.00th=[758],
| 30.00th=[758], 40.00th=[758], 50.00th=[758], 60.00th=[766],
| 70.00th=[766], 80.00th=[766], 90.00th=[766], 95.00th=[783],
| 99.00th=[979], 99.50th=[979], 99.90th=[979], 99.95th=[979],
| 99.99th=[979]
lat (usec) : 750=6.25%, 1000=93.75%
cpu : usr=0.00%, sys=4.17%, ctx=32, majf=0, minf=518
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%


```

complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=32,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1
4m_block: (groupid=4, jobs=1): err= 0: pid=310162: Mon Apr 14 08:12:27 2025
read: IOPS=695, BW=2783MiB/s (2918MB/s)(64.0MiB/23msec)
clat (usec): min=1326, max=2065, avg=1403.20, stdev=176.93
lat (usec): min=1326, max=2065, avg=1403.32, stdev=177.04
clat percentiles (usec):
| 1.00th=[ 1319], 5.00th=[ 1319], 10.00th=[ 1352], 20.00th=[ 1352],
| 30.00th=[ 1352], 40.00th=[ 1352], 50.00th=[ 1352], 60.00th=[ 1369],
| 70.00th=[ 1369], 80.00th=[ 1369], 90.00th=[ 1385], 95.00th=[ 2073],
| 99.00th=[ 2073], 99.50th=[ 2073], 99.90th=[ 2073], 99.95th=[ 2073],
| 99.99th=[ 2073]
lat (msec) : 2=93.75%, 4=6.25%
cpu : usr=0.00%, sys=4.55%, ctx=17, majf=0, minf=520
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=16,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1
8m_block: (groupid=5, jobs=1): err= 0: pid=310163: Mon Apr 14 08:12:27 2025
read: IOPS=363, BW=2909MiB/s (3050MB/s)(64.0MiB/22msec)
clat (usec): min=2545, max=3297, avg=2648.25, stdev=262.22
lat (usec): min=2545, max=3297, avg=2648.38, stdev=262.28
clat percentiles (usec):
| 1.00th=[ 2540], 5.00th=[ 2540], 10.00th=[ 2540], 20.00th=[ 2540],
| 30.00th=[ 2540], 40.00th=[ 2540], 50.00th=[ 2540], 60.00th=[ 2573],
| 70.00th=[ 2573], 80.00th=[ 2573], 90.00th=[ 3294], 95.00th=[ 3294],
| 99.00th=[ 3294], 99.50th=[ 3294], 99.90th=[ 3294], 99.95th=[ 3294],
| 99.99th=[ 3294]
lat (msec) : 4=100.00%
cpu : usr=4.76%, sys=0.00%, ctx=8, majf=0, minf=522
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=8,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1
16m_block: (groupid=6, jobs=1): err= 0: pid=310164: Mon Apr 14 08:12:27 2025
read: IOPS=190, BW=3048MiB/s (3196MB/s)(64.0MiB/21msec)
clat (usec): min=4946, max=5717, avg=5146.10, stdev=380.86
lat (usec): min=4946, max=5717, avg=5146.43, stdev=380.85
clat percentiles (usec):
| 1.00th=[ 4948], 5.00th=[ 4948], 10.00th=[ 4948], 20.00th=[ 4948],
| 30.00th=[ 4948], 40.00th=[ 4948], 50.00th=[ 4948], 60.00th=[ 4948],
| 70.00th=[ 4948], 80.00th=[ 5735], 90.00th=[ 5735], 95.00th=[ 5735],
| 99.00th=[ 5735], 99.50th=[ 5735], 99.90th=[ 5735], 99.95th=[ 5735],
| 99.99th=[ 5735]
lat (msec) : 10=100.00%
cpu : usr=0.00%, sys=5.00%, ctx=4, majf=0, minf=525
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=4,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1
32m_block: (groupid=7, jobs=1): err= 0: pid=310165: Mon Apr 14 08:12:27 2025
read: IOPS=90, BW=2909MiB/s (3050MB/s)(64.0MiB/22msec)
clat (usec): min=10075, max=11211, avg=10643.69, stdev=803.40
lat (usec): min=10075, max=11212, avg=10643.99, stdev=803.39
clat percentiles (usec):
| 1.00th=[10028], 5.00th=[10028], 10.00th=[10028], 20.00th=[10028],
| 30.00th=[10028], 40.00th=[10028], 50.00th=[10028], 60.00th=[11207],
| 70.00th=[11207], 80.00th=[11207], 90.00th=[11207], 95.00th=[11207],
| 99.00th=[11207], 99.50th=[11207], 99.90th=[11207], 99.95th=[11207],
| 99.99th=[11207]
lat (msec) : 20=100.00%
cpu : usr=0.00%, sys=9.52%, ctx=2, majf=0, minf=533
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=2,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1

```

Run status group 0 (all jobs):

READ: bw=60.5MiB/s (63.4MB/s), 60.5MiB/s-60.5MiB/s (63.4MB/s-63.4MB/s), io=64.0MiB (67.1MB), run=1058-1058msec

```

Run status group 1 (all jobs):
  READ: bw=240MiB/s (251MB/s), 240MiB/s-240MiB/s (251MB/s-251MB/s), io=64.0MiB (67.1MB), run=267-267msec

Run status group 2 (all jobs):
  READ: bw=2000MiB/s (2097MB/s), 2000MiB/s-2000MiB/s (2097MB/s-2097MB/s), io=64.0MiB (67.1MB), run=32-32msec

Run status group 3 (all jobs):
  READ: bw=2560MiB/s (2684MB/s), 2560MiB/s-2560MiB/s (2684MB/s-2684MB/s), io=64.0MiB (67.1MB), run=25-25msec

Run status group 4 (all jobs):
  READ: bw=2783MiB/s (2918MB/s), 2783MiB/s-2783MiB/s (2918MB/s-2918MB/s), io=64.0MiB (67.1MB), run=23-23msec

Run status group 5 (all jobs):
  READ: bw=2909MiB/s (3050MB/s), 2909MiB/s-2909MiB/s (3050MB/s-3050MB/s), io=64.0MiB (67.1MB), run=22-22msec

Run status group 6 (all jobs):
  READ: bw=3048MiB/s (3196MB/s), 3048MiB/s-3048MiB/s (3196MB/s-3196MB/s), io=64.0MiB (67.1MB), run=21-21msec

Run status group 7 (all jobs):
  READ: bw=2909MiB/s (3050MB/s), 2909MiB/s-2909MiB/s (3050MB/s-3050MB/s), io=64.0MiB (67.1MB), run=22-22msec

Disk stats (read/write):
  nvme1n1: ios=71570/134, sectors=179344/4416, merge=0/3, ticks=948/104, in_queue=1053, util=62.66%

```

Question

1. Is there any significant difference between 4k and 1k read? Why or why not? Please justify your answer in brief.
2. If you want to achieve the best performance in the condition above, how would you modify blocksize? Explain it briefly.

Answer

- *4k read* is significantly faster than *1k read* in terms of bandwidth and latency.
- The bandwidth can be increased by changing blocksize to up to 8m.

Larger blocksize might be useful when doing large sequential read.

Q7. Fastest nonbuffered read

Result

```

psync: (g=0): rw=read, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB, ioengine=psync,
iodepth=1
libaio: (g=1): rw=read, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB, ioengine=libaio,
iodepth=1
libaio-iod32: (g=2): rw=read, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB, ioengine=libaio,
iodepth=32
libaio-iod64: (g=3): rw=read, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB, ioengine=libaio,
iodepth=64
libaio-iod128: (g=4): rw=read, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB,
ioengine=libaio, iodepth=128
fastest: (g=5): rw=read, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB, ioengine=libaio,
iodepth=1
fio-3.39
Starting 6 processes

```

```

psync: (groupid=0, jobs=1): err= 0: pid=314338: Mon Apr 14 08:40:53 2025
read: IOPS=411, BW=3293MiB/s (3453MB/s)(1024MiB/311msec)
  clat (usec): min=2412, max=3248, avg=2426.43, stdev=73.58
  lat (usec): min=2412, max=3248, avg=2426.55, stdev=73.59
  clat percentiles (usec):
    | 1.00th=[ 2409],  5.00th=[ 2409], 10.00th=[ 2409], 20.00th=[ 2409],
    | 30.00th=[ 2409], 40.00th=[ 2409], 50.00th=[ 2409], 60.00th=[ 2409],
    | 70.00th=[ 2409], 80.00th=[ 2409], 90.00th=[ 2442], 95.00th=[ 2442],
    | 99.00th=[ 2442], 99.50th=[ 3261], 99.90th=[ 3261], 99.95th=[ 3261],
    | 99.99th=[ 3261]
  lat (msec)  : 4=100.00%
cpu           : usr=0.00%, sys=2.90%, ctx=130, majf=0, minf=522
IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit      : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency     : target=0, window=0, percentile=100.00%, depth=1
libaio: (groupid=1, jobs=1): err= 0: pid=314339: Mon Apr 14 08:40:53 2025
read: IOPS=410, BW=3282MiB/s (3441MB/s)(1024MiB/312msec)
  slat (usec): min=62, max=651, avg=75.86, stdev=53.47
  clat (usec): min=2219, max=2586, avg=2352.36, stdev=25.94

```

```

lat (usec): min=2414, max=3238, avg=2428.22, stdev=72.91
clat percentiles (usec):
| 1.00th=[ 2343], 5.00th=[ 2343], 10.00th=[ 2343], 20.00th=[ 2343],
| 30.00th=[ 2343], 40.00th=[ 2343], 50.00th=[ 2343], 60.00th=[ 2343],
| 70.00th=[ 2343], 80.00th=[ 2343], 90.00th=[ 2376], 95.00th=[ 2376],
| 99.00th=[ 2409], 99.50th=[ 2573], 99.90th=[ 2573], 99.95th=[ 2573],
| 99.99th=[ 2573]
lat (msec) : 4=100.00%
cpu : usr=0.00%, sys=2.89%, ctx=131, majf=0, minf=522
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1
libaio-iod32: (groupid=2, jobs=1): err= 0: pid=314340: Mon Apr 14 08:40:53 2025
read: IOPS=418, BW=3346MiB/s (3509MB/s)(1024MiB/306msec)
slat (usec): min=48, max=597, avg=500.10, stdev=162.08
clat (msec): min=11, max=146, avg=73.69, stdev=30.45
lat (msec): min=11, max=147, avg=74.19, stdev=30.48
clat percentiles (msec):
| 1.00th=[ 12], 5.00th=[ 15], 10.00th=[ 27], 20.00th=[ 52],
| 30.00th=[ 77], 40.00th=[ 77], 50.00th=[ 77], 60.00th=[ 77],
| 70.00th=[ 77], 80.00th=[ 88], 90.00th=[ 120], 95.00th=[ 133],
| 99.00th=[ 144], 99.50th=[ 148], 99.90th=[ 148], 99.95th=[ 148],
| 99.99th=[ 148]
lat (msec) : 20=7.03%, 50=12.50%, 100=64.84%, 250=15.62%
cpu : usr=0.00%, sys=5.90%, ctx=177, majf=0, minf=136
IO depths : 1=0.8%, 2=1.6%, 4=3.1%, 8=6.2%, 16=12.5%, 32=75.8%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=99.0%, 8=0.0%, 16=0.0%, 32=1.0%, 64=0.0%, >=64=0.0%
issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=32
libaio-iod64: (groupid=3, jobs=1): err= 0: pid=314341: Mon Apr 14 08:40:53 2025
read: IOPS=418, BW=3346MiB/s (3509MB/s)(1024MiB/306msec)
slat (usec): min=49, max=7156, avg=1794.84, stdev=2134.18
clat (msec): min=18, max=263, avg=138.99, stdev=70.33
lat (msec): min=19, max=266, avg=140.78, stdev=70.87
clat percentiles (msec):
| 1.00th=[ 35], 5.00th=[ 60], 10.00th=[ 62], 20.00th=[ 66],
| 30.00th=[ 79], 40.00th=[ 100], 50.00th=[ 129], 60.00th=[ 155],
| 70.00th=[ 186], 80.00th=[ 232], 90.00th=[ 236], 95.00th=[ 249],
| 99.00th=[ 262], 99.50th=[ 264], 99.90th=[ 264], 99.95th=[ 264],
| 99.99th=[ 264]
lat (msec) : 20=0.78%, 50=1.56%, 100=39.06%, 250=53.91%, 500=4.69%
cpu : usr=0.00%, sys=10.49%, ctx=221, majf=0, minf=265
IO depths : 1=0.8%, 2=1.6%, 4=3.1%, 8=6.2%, 16=12.5%, 32=25.0%, >=64=50.8%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=98.5%, 8=0.0%, 16=0.0%, 32=0.0%, 64=1.5%, >=64=0.0%
issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=64
libaio-iod128: (groupid=4, jobs=1): err= 0: pid=314344: Mon Apr 14 08:40:53 2025
read: IOPS=418, BW=3346MiB/s (3509MB/s)(1024MiB/306msec)
slat (usec): min=371, max=7565, avg=1792.97, stdev=2067.63
clat (msec): min=76, max=304, avg=231.90, stdev=60.80
lat (msec): min=82, max=305, avg=233.69, stdev=59.54
clat percentiles (msec):
| 1.00th=[ 83], 5.00th=[ 104], 10.00th=[ 132], 20.00th=[ 197],
| 30.00th=[ 215], 40.00th=[ 220], 50.00th=[ 226], 60.00th=[ 255],
| 70.00th=[ 284], 80.00th=[ 296], 90.00th=[ 300], 95.00th=[ 305],
| 99.00th=[ 305], 99.50th=[ 305], 99.90th=[ 305], 99.95th=[ 305],
| 99.99th=[ 305]
lat (msec) : 100=3.91%, 250=54.69%, 500=41.41%
cpu : usr=0.00%, sys=17.38%, ctx=198, majf=0, minf=522
IO depths : 1=0.8%, 2=1.6%, 4=3.1%, 8=6.2%, 16=12.5%, 32=25.0%, >=64=50.8%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=50.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=50.0%
issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=128
fastest: (groupid=5, jobs=1): err= 0: pid=314345: Mon Apr 14 08:40:53 2025
read: IOPS=411, BW=3293MiB/s (3453MB/s)(1024MiB/311msec)
slat (usec): min=60, max=630, avg=72.57, stdev=51.51
clat (usec): min=2214, max=2634, avg=2353.93, stdev=29.25
lat (usec): min=2413, max=3265, avg=2426.50, stdev=75.19
clat percentiles (usec):
| 1.00th=[ 2311], 5.00th=[ 2343], 10.00th=[ 2343], 20.00th=[ 2343],
| 30.00th=[ 2343], 40.00th=[ 2343], 50.00th=[ 2343], 60.00th=[ 2343],

```

```

| 70.00th=[ 2343], 80.00th=[ 2343], 90.00th=[ 2376], 95.00th=[ 2376],
| 99.00th=[ 2442], 99.50th=[ 2638], 99.90th=[ 2638], 99.95th=[ 2638],
| 99.99th=[ 2638]
lat (msec) : 4=100.00%
cpu : usr=0.00%, sys=2.90%, ctx=128, majf=0, minf=522
IO depths : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
submit : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
complete : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
latency : target=0, window=0, percentile=100.00%, depth=1

Run status group 0 (all jobs):
READ: bw=3293MiB/s (3453MB/s), 3293MiB/s-3293MiB/s (3453MB/s-3453MB/s), io=1024MiB (1074MB), run=311-311msec

Run status group 1 (all jobs):
READ: bw=3282MiB/s (3441MB/s), 3282MiB/s-3282MiB/s (3441MB/s-3441MB/s), io=1024MiB (1074MB), run=312-312msec

Run status group 2 (all jobs):
READ: bw=3346MiB/s (3509MB/s), 3346MiB/s-3346MiB/s (3509MB/s-3509MB/s), io=1024MiB (1074MB), run=306-306msec

Run status group 3 (all jobs):
READ: bw=3346MiB/s (3509MB/s), 3346MiB/s-3346MiB/s (3509MB/s-3509MB/s), io=1024MiB (1074MB), run=306-306msec

Run status group 4 (all jobs):
READ: bw=3346MiB/s (3509MB/s), 3346MiB/s-3346MiB/s (3509MB/s-3509MB/s), io=1024MiB (1074MB), run=306-306msec

Run status group 5 (all jobs):
READ: bw=3293MiB/s (3453MB/s), 3293MiB/s-3293MiB/s (3453MB/s-3453MB/s), io=1024MiB (1074MB), run=311-311msec

Disk stats (read/write):
nvme1n1: ios=22560/1, sectors=11454960/176, merge=128/0, ticks=940215/0, in_queue=940215, util=51.56%

```

Question

Please explain how you achieve the fastest 1G nonbuffered read in brief.

Answer

According to the result in the previous questions, namely, Q5 and Q6, the fastest non-buffered read happens when offset=80% and blocksize=8m.

Furthermore, I have tested the difference between ioengine=psync, which is the default for fio, and ioengine=libaio, which is Linux native asynchronous I/O. The result shows that ioengine=libaio is better in terms of bandwidth.

For asynchronous I/O, I also tested the impact of the iodepth. The result shows that iodepth=64 will reach the peak bandwidth.

Therefore, the fastest non-buffered 1G read should be the combination of offset=80%, blocksize=8m, ioengine=libaio, and iodepth=64.

Bonus

Result

```

ssd-read: (g=0): rw=read, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB, ioengine=libaio,
iodepth=64
ssd-randread: (g=1): rw=randread, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB,
ioengine=libaio, iodepth=64
hdd-read: (g=2): rw=read, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB, ioengine=libaio,
iodepth=64
hdd-randread: (g=3): rw=randread, bs=(R) 8192KiB-8192KiB, (W) 8192KiB-8192KiB, (T) 8192KiB-8192KiB,
ioengine=libaio, iodepth=64
fio-3.39
Starting 4 processes

```

```

ssd-read: (groupid=0, jobs=1): err= 0: pid=314662: Mon Apr 14 08:43:57 2025
read: IOPS=412, BW=3303MiB/s (3464MB/s)(1024MiB/310msec)
  slat (usec): min=46, max=7314, avg=1366.16, stdev=1757.98
  clat (msec): min=14, max=255, avg=140.27, stdev=65.17
    lat (msec): min=14, max=257, avg=141.63, stdev=65.69
  clat percentiles (msec):
    | 1.00th=[ 17], 5.00th=[ 44], 10.00th=[ 80], 20.00th=[ 85],
    | 30.00th=[ 89], 40.00th=[ 94], 50.00th=[ 110], 60.00th=[ 167],
    | 70.00th=[ 207], 80.00th=[ 209], 90.00th=[ 226], 95.00th=[ 241],
    | 99.00th=[ 253], 99.50th=[ 255], 99.90th=[ 255], 99.95th=[ 255],
    | 99.99th=[ 255]
  lat (msec) : 20=1.56%, 50=6.25%, 100=36.72%, 250=53.12%, 500=2.34%

```

```

cpu      : usr=0.00%, sys=10.36%, ctx=176, majf=0, minf=775
IO depths : 1=0.8%, 2=1.6%, 4=3.1%, 8=6.2%, 16=12.5%, 32=25.0%, >=64=50.8%
  submit  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
  complete : 0=0.0%, 4=98.5%, 8=0.0%, 16=0.0%, 32=0.0%, 64=1.5%, >=64=0.0%
  issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
  latency  : target=0, window=0, percentile=100.00%, depth=64
ssd-randread: (groupid=1, jobs=1): err= 0: pid=314663: Mon Apr 14 08:43:57 2025
  read: IOPS=415, BW=3325MiB/s (3486MB/s)(1024MiB/308msec)
    slat (usec): min=44, max=7084, avg=1425.85, stdev=1528.22
    clat (msec): min=12, max=236, avg=139.74, stdev=49.47
    lat (msec): min=12, max=238, avg=141.17, stdev=49.75
    clat percentiles (msec):
      | 1.00th=[ 26], 5.00th=[ 79], 10.00th=[ 84], 20.00th=[ 96],
      | 30.00th=[ 108], 40.00th=[ 118], 50.00th=[ 123], 60.00th=[ 176],
      | 70.00th=[ 176], 80.00th=[ 178], 90.00th=[ 207], 95.00th=[ 222],
      | 99.00th=[ 234], 99.50th=[ 236], 99.90th=[ 236], 99.95th=[ 236],
      | 99.99th=[ 236]
  lat (msec) : 20=0.78%, 50=2.34%, 100=21.09%, 250=75.78%
cpu      : usr=0.00%, sys=9.45%, ctx=222, majf=0, minf=265
IO depths : 1=0.8%, 2=1.6%, 4=3.1%, 8=6.2%, 16=12.5%, 32=25.0%, >=64=50.8%
  submit  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
  complete : 0=0.0%, 4=98.5%, 8=0.0%, 16=0.0%, 32=0.0%, 64=1.5%, >=64=0.0%
  issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
  latency  : target=0, window=0, percentile=100.00%, depth=64
hdd-read: (groupid=2, jobs=1): err= 0: pid=314664: Mon Apr 14 08:43:57 2025
  read: IOPS=20, BW=167MiB/s (175MB/s)(1024MiB/6140msec)
    slat (usec): min=342, max=528428, avg=44674.51, stdev=44631.06
    clat (msec): min=419, max=3260, avg=2302.17, stdev=775.39
    lat (msec): min=466, max=3304, avg=2346.84, stdev=776.16
    clat percentiles (msec):
      | 1.00th=[ 468], 5.00th=[ 684], 10.00th=[ 936], 20.00th=[ 1485],
      | 30.00th=[ 2072], 40.00th=[ 2668], 50.00th=[ 2769], 60.00th=[ 2769],
      | 70.00th=[ 2769], 80.00th=[ 2802], 90.00th=[ 2970], 95.00th=[ 3071],
      | 99.00th=[ 3205], 99.50th=[ 3272], 99.90th=[ 3272], 99.95th=[ 3272],
      | 99.99th=[ 3272]
  bw ( KiB/s): min=114688, max=1048576, per=100.00%, avg=299644.71, stdev=331555.65, samples=7
  iops       : min= 14, max= 128, avg=36.57, stdev=40.48, samples=7
  lat (msec) : 500=1.56%, 750=4.69%, 1000=4.69%, 2000=17.97%, >=2000=71.09%
cpu      : usr=0.00%, sys=0.52%, ctx=837, majf=0, minf=265
IO depths : 1=0.8%, 2=1.6%, 4=3.1%, 8=6.2%, 16=12.5%, 32=25.0%, >=64=50.8%
  submit  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
  complete : 0=0.0%, 4=98.5%, 8=0.0%, 16=0.0%, 32=0.0%, 64=1.5%, >=64=0.0%
  issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
  latency  : target=0, window=0, percentile=100.00%, depth=64
hdd-randread: (groupid=3, jobs=1): err= 0: pid=314666: Mon Apr 14 08:43:57 2025
  read: IOPS=20, BW=161MiB/s (169MB/s)(1024MiB/6366msec)
    slat (usec): min=358, max=108902, avg=46194.96, stdev=14404.86
    clat (msec): min=451, max=3702, avg=2565.65, stdev=871.70
    lat (msec): min=494, max=3754, avg=2611.84, stdev=870.24
    clat percentiles (msec):
      | 1.00th=[ 493], 5.00th=[ 735], 10.00th=[ 1036], 20.00th=[ 1687],
      | 30.00th=[ 2366], 40.00th=[ 2836], 50.00th=[ 2937], 60.00th=[ 3004],
      | 70.00th=[ 3104], 80.00th=[ 3239], 90.00th=[ 3373], 95.00th=[ 3473],
      | 99.00th=[ 3675], 99.50th=[ 3708], 99.90th=[ 3708], 99.95th=[ 3708],
      | 99.99th=[ 3708]
  bw ( KiB/s): min=114688, max=163840, per=92.36%, avg=152137.14, stdev=18230.43, samples=7
  iops       : min= 14, max= 20, avg=18.57, stdev= 2.23, samples=7
  lat (msec) : 500=1.56%, 750=3.91%, 1000=3.91%, 2000=15.62%, >=2000=75.00%
cpu      : usr=0.02%, sys=0.49%, ctx=835, majf=0, minf=264
IO depths : 1=0.8%, 2=1.6%, 4=3.1%, 8=6.2%, 16=12.5%, 32=25.0%, >=64=50.8%
  submit  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
  complete : 0=0.0%, 4=98.5%, 8=0.0%, 16=0.0%, 32=0.0%, 64=1.5%, >=64=0.0%
  issued rwts: total=128,0,0,0 short=0,0,0,0 dropped=0,0,0,0
  latency  : target=0, window=0, percentile=100.00%, depth=64

Run status group 0 (all jobs):
  READ: bw=3303MiB/s (3464MB/s), 3303MiB/s-3303MiB/s (3464MB/s-3464MB/s), io=1024MiB (1074MB), run=310-310msec

Run status group 1 (all jobs):
  READ: bw=3325MiB/s (3486MB/s), 3325MiB/s-3325MiB/s (3486MB/s-3486MB/s), io=1024MiB (1074MB), run=308-308msec

Run status group 2 (all jobs):
  READ: bw=167MiB/s (175MB/s), 167MiB/s-167MiB/s (175MB/s-175MB/s), io=1024MiB (1074MB), run=6140-6140msec

Run status group 3 (all jobs):
  READ: bw=161MiB/s (169MB/s), 161MiB/s-161MiB/s (169MB/s-169MB/s), io=1024MiB (1074MB), run=6366-6366msec

```

Disk stats (read/write):

```
nvme1n1: ios=8036/19, sectors=4114432/504, merge=0/1, ticks=658566/8, in_queue=658578, util=61.58%  
sda: ios=1781/0, sectors=4169216/0, merge=0/0, ticks=767527/0, in_queue=767527, util=97.16%
```

Question

1. Compare with multiple kinds of storage devices.

Is there any significant difference between read/randread on storage devices? Why or why not? Please justify your answer in brief.

2. Find the specs of your own hardware that you tested on.

Is your hardware running to spec? If not, could you come up with a possible theory?

Answer

This experiment include two storage devices, Crucial P3 Plus M.2 SSD[1] (Disk 1) and WD Blue HDD[4] (Disk 2). The hardware sequential read rate for Disk 1 is 5000MB/s. And the hardware transfer rate for Disk 2 is up to 150MB/s.

The result for the experiment shows that

1. Disk 1 (SSD) is significantly faster than Disk 2 (HDD).
2. Disk 1 is quite a bit slower than the hardware spec (about 3400MB/s).
3. Disk 2 is faster than the hardware spec (about 170MB/s).

The reason for the underperformance of Disk 1 might be

- I'm using Disk 1 to run my OS at the same time, which interferes the result.
- Disk 1 had been used for over 4 years and had used up 50% of the capacity. In construct, Disk 2 was seldom used and didn't used up any of the capacity.

3. Bibliography

- [1] "Crucial P3 Plus 1TB PCIe M.2 2280 SSD." [Online]. Available: <https://www.crucial.tw/ssd/p3-plus/ct1000p3pssd8>
- [2] "Btrfs, a modern copy on write filesystem for Linux." [Online]. Available: <https://btrfs.readthedocs.io/en/latest/>
- [3] "ArchLinux, a simple, lightweight distribution." [Online]. Available: <https://archlinux.org/>
- [4] "WD Blue PC Desktop Hard Drive - 1TB." [Online]. Available: <https://www.westerndigital.com/products/internal-drives/wd-blue-desktop-sata-hdd?sku=WD10EZEX>