MyDiskBench

All experiments have been conducted under **Haswell** Computing Environment.

Theoretical IOPS and Throughput

According to the storage data sheet, it has the maximum sustained transferred rate at 115MB/sec.

Throughput

Work-load	Con-currency	Record Size	MyDiskBench Measured Throughput (MB/sec)	IOZone Measured Throughput (MB/sec)	Theoretical Throughput (MB/sec)	MyDiskBench Efficiency (%)	IOZone Efficiency (%)
WS	1	64KB	1.309666	6.89396	115	1.13884	5.99474783
WS	1	1MB	16.739024	50.70584	115	14.55567	44.0920348
WS	1	16MB	65.406324	66.77512	115	56.87506	58.0653217
WS	2	64KB	1.783234	7.33769	115	1.550638	6.3806
WS	2	1MB	17.56324	49.844	115	15.27238	43.3426087
WS	2	16MB	51.630662	55.25018	115	44.89623	48.0436348
WS	4	64KB	2.865091	6.99113	115	2.491383	6.07924348
WS	4	1MB	30.820498	43.52433	115	26.80043	37.8472435
WS	4	16MB	77.448739	35.02915	115	67.34673	30.4601304
WS	8	64KB	5.923607	7.09202	115	5.150963	6.16697391
WS	8	1MB	44.280661	39.9964	115	38.50492	34.7794783
WS	8	16MB	56.92976	27.10066	115	49.50414	23.5657913
WS	12	64K	7.602244	7.15759	115	6.610647	6.2239913
WS	12	1MB	45.760608	37.26656	115	39.79183	32.4057043
WS	12	16M	43.797484	47.01933	115	38.08477	40.8863739
WS	24	64K	10.604236	7.10709	115	9.221075	6.18007826
WS	24	1MB	33.838973	33.78625	115	29.42519	29.3793478
WS	24	16M	32.605079	56.35912	115	28.35224	49.0079304
WS	48	64K	31.264833	7.27335	115	27.18681	6.32465217
WS	48	1MB	27.999229	31.79085	115	24.34716	27.6442174
WS	48	16M	11.196105	36.663	115	9.735743	31.8808696

Table 1.a : workload WS

Work-load	Con-currency	Record Size	MyDiskBench Measured Throughput (MB/sec)	IOZone Measured Throughput (MB/sec)	Theoretical Throughput (MB/sec)	MyDiskBench Efficiency (%)	IOZone Efficiency (%)
RS	1	64KB	82.503778	91.74283	115	71.74242	79.7763739
RS	1	1MB	82.226985	88.2225	115	71.50173	76.7152174
RS	1	16MB	81.47361	68.94078	115	70.84662	59.9485043
RS	2	64KB	71.153573	82.32854	115	61.87267	71.5900348
RS	2	1MB	70.332559	76.96215	115	61.15875	66.9236087
RS	2	16MB	69.254856	79.5108	115	60.22161	69.1398261
RS	4	64KB	69.357771	81.57924	115	60.31111	70.9384696
RS	4	1MB	68.173308	74.50421	115	59.28114	64.7862696
RS	4	16MB	68.47359	61.95663	115	59.54225	53.8753304
RS	8	64KB	62.095008	72.67486	115	53.99566	63.1955304
RS	8	1MB	65.610477	69.18772	115	57.05259	60.1632348
RS	8	16MB	67.26986	54.1667	115	58.49553	47.1014783
RS	12	64K	55.131995	65.31374	115	47.94087	56.7945565
RS	12	1MB	59.923282	63.83407	115	52.1072	55.507887
RS	12	16M	63.310737	76.05155	115	55.05281	66.1317826
RS	24	64K	49.002699	58.17361	115	42.61104	50.5857478
RS	24	1MB	57.262499	58.82318	115	49.79348	51.1505913
RS	24	16M	61.053379	72.35659	115	53.08989	62.9187739
RS	48	64K	49.25086	60.37182	115	42.82683	52.4972348
RS	48	1MB	45.378603	54.4137	115	39.45965	47.3162609
RS	48	16M	61.851761	64.72254	115	53.78414	56.2804696

Table 1.b workload RS

Work-load	Con-currency	Record Size	MyDiskBench Measured Throughput (MB/sec)	IOZone Measured Throughput (MB/sec)	Theoretical Throughput (MB/sec)	MyDiskBench Efficiency (%)	IOZone Efficiency (%)
WR	1	64KB	1.368257	7.4068	115	1.189789	6.440696
WR	1	1MB	18.831789	52.239	115	16.37547	45.42522
WR	1	16MB	60.609687	65.38772	115	52.70408	56.85889
WR	2	64KB	1.976504	8.2206	115	1.718699	7.148348
WR	2	1MB	22.498049	48.83943	115	19.56352	42.46907
WR	2	16MB	54.955272	55.30316	115	47.78719	48.0897
WR	4	64KB	2.942604	9.99638	115	2.558786	8.692504
WR	4	1MB	25.517163	37.46561	115	22.18884	32.57879
WR	4	16MB	50.985315	29.47134	115	44.33506	25.62725
WR	8	64KB	3.923891	12.02452	115	3.412079	10.4561
WR	8	1MB	25.474171	34.23673	115	22.15145	29.77107
WR	8	16MB	31.839103	24.46402	115	27.68618	21.27306
WR	12	64K	4.442537	12.25162	115	3.863076	10.65358
WR	12	1MB	22.727848	34.41053	115	19.76335	29.9222
WR	12	16M	28.389191	35.1721	115	24.68625	30.58443
WR	24	64K	4.905433	12.71277	115	4.265594	11.05458
WR	24	1MB	21.874355	32.35183	115	19.02118	28.13203
WR	24	16M	47.608886	32.83172	115	41.39903	28.54932
WR	48	64K	6.415418	12.64371	115	5.578624	10.99453
WR	48	1MB	26.126099	27.61703	115	22.71835	24.01481
WR	48	16M	40.937718	31.27122	115	35.59802	27.19237

Table 1.c workload WR

Work-load	Con-currency	Record Size	MyDiskBench Measured Throughput (MB/sec)	IOZone Measured Throughput (MB/sec)	Theoretical Throughput (MB/sec)	MyDiskBench Efficiency (%)	IOZone Efficiency (%)
RR	1	64KB	8.565218	8.29067	115	7.448016	7.209278
RR	1	1MB	62.594892	53.0727	115	54.43034	46.15017

RR	1	16MB	68.881413	66.98451	115	59.89688	58.2474
RR	2	64KB	10.533499	10.18753	115	9.159564	8.858722
RR	2	1MB	65.821897	48.83943	115	57.23643	42.46907
RR	2	16MB	65.914339	74.83618	115	57.31682	65.07494
RR	4	64KB	11.10533	12.74473	115	9.656809	11.08237
RR	4	1MB	65.859456	54.32561	115	57.26909	47.23966
RR	4	16MB	67.727293	58.25953	115	58.8933	50.66046
RR	8	64KB	14.933232	13.94191	115	12.98542	12.1234
RR	8	1MB	74.072025	55.94969	115	64.41046	48.6519
RR	8	16MB	79.753122	52.65417	115	69.35054	45.78623
RR	12	64K	14.968671	13.94168	115	13.01624	12.1232
RR	12	1MB	67.939222	54.84068	115	59.07758	47.68755
RR	12	16M	72.92795	73.39134	115	63.41561	63.81856
RR	24	64K	17.968962	15.68393	115	15.62518	13.6382
RR	24	1MB	69.191116	54.17761	115	60.16619	47.11097
RR	24	16M	78.279191	69.92173	115	68.06886	60.8015
RR	48	64K	20.571546	17.66937	115	17.8883	15.36467
RR	48	1MB	71.482322	52.73947	115	62.15854	45.86041
RR	48	16M	82.648571	65.06518	115	71.86832	56.57842

Table 1.d workload RR

IOPS

According to the hdd specification from the manufacturer, its maximum throughput, average latency, average access time read/write are 115MB/s, 4.16ms, 8.5ms/9.5ms respectively. Theoretical IOPS = 1/(0.009+ 0.00416) = 75.9878419453 iops/sec

Work- load	Con-currency	Record Size	MyDiskBench Measured IOPS (OPS/sec)	IOZone Measured IOPS (OPS/sec)	Theoretical IOPS (OPS/sec)	MyDiskBench Efficiency (%)	IOZone Efficiency (%)
WR	1	4KB	22.04416	122.79	75.98784195	29.0101146	161.59164
WR	2	4KB	37.392347	135.44	75.98784195	49.2083287	178.23904
WR	4	4KB	46.206124	157.99	75.98784195	60.8072592	207.91484
WR	8	4KB	54.679133	194.2	75.98784195	71.957739	255.5672
WR	12	4KB	62.894161	196.2	75.98784195	82.7687159	258.1992
WR	24	4KB	70.911762	220.18	75.98784195	93.3198788	289.75688
WR	48	4KB	79.582829	246.7	75.98784195	104.731003	324.6572

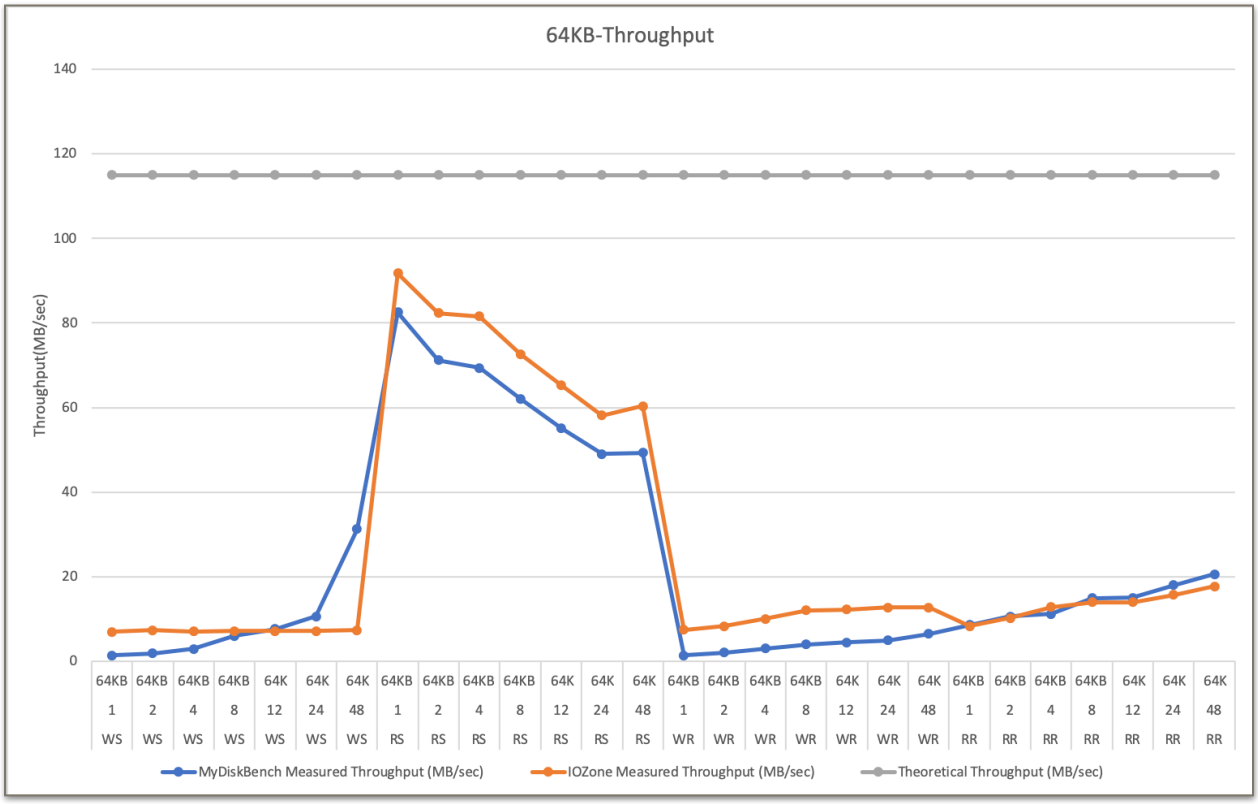
Table 2.a workload WR

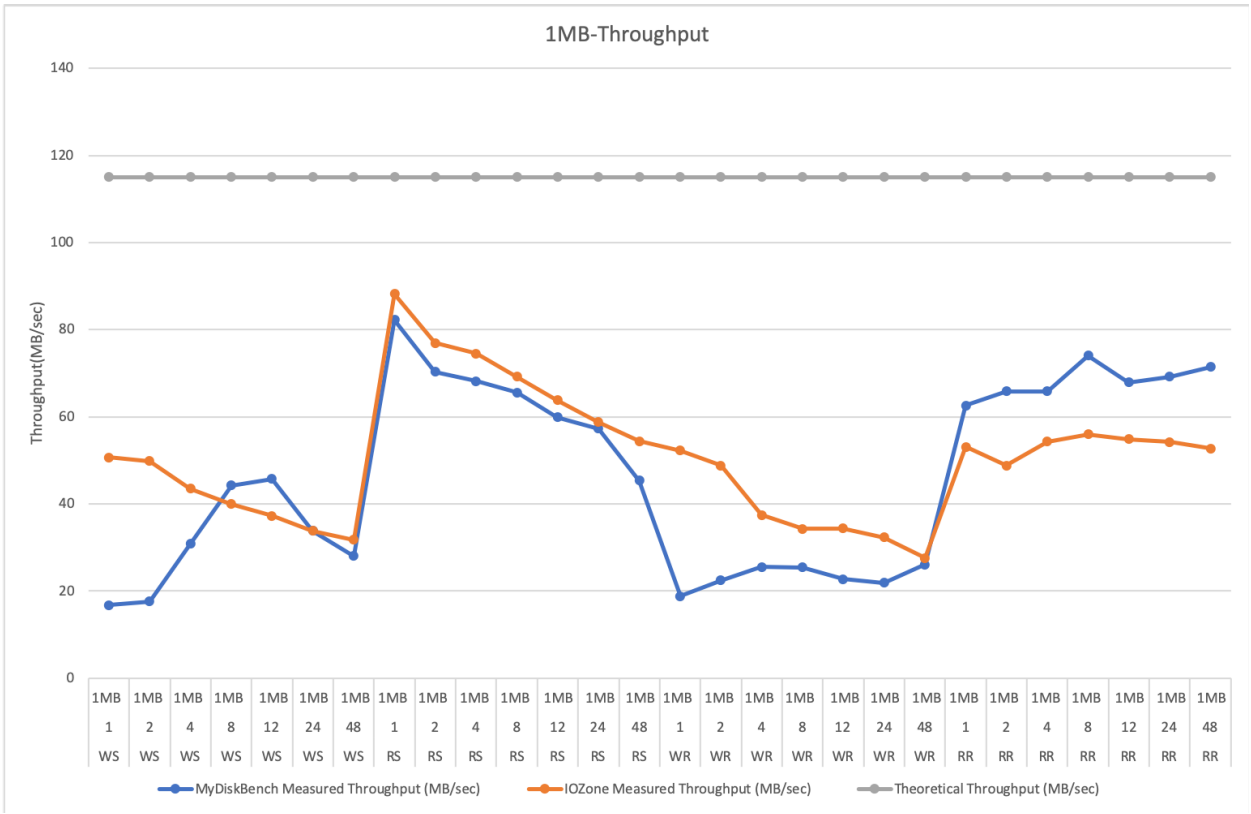
Work- load	Con-currency	Record Size	MyDiskBench Measured IOPS (OPS/sec)	IOZone Measured IOPS (OPS/sec)	Theoretical IOPS (OPS/sec)	MyDiskBench Efficiency (%)	IOZone Efficiency (%)
RR	1	4KB	136.978424	139.14	75.98784195	180.263606	183.10824
RR	2	4KB	161.948698	161.59	75.98784195	213.124487	212.65244
RR	4	4KB	152.106552	239.06	75.98784195	200.172222	314.60296
RR	8	4KB	197.512191	220.14	75.98784195	259.926043	289.70424
RR	12	4KB	258.19139	264.3	75.98784195	339.779869	347.8188
RR	24	4KB	253.108637	302.12	75.98784195	333.090966	397.58992
RR	48	4KB	291.87248	360.53	75.98784195	384.104184	474.45748

Table 2.b workload RR

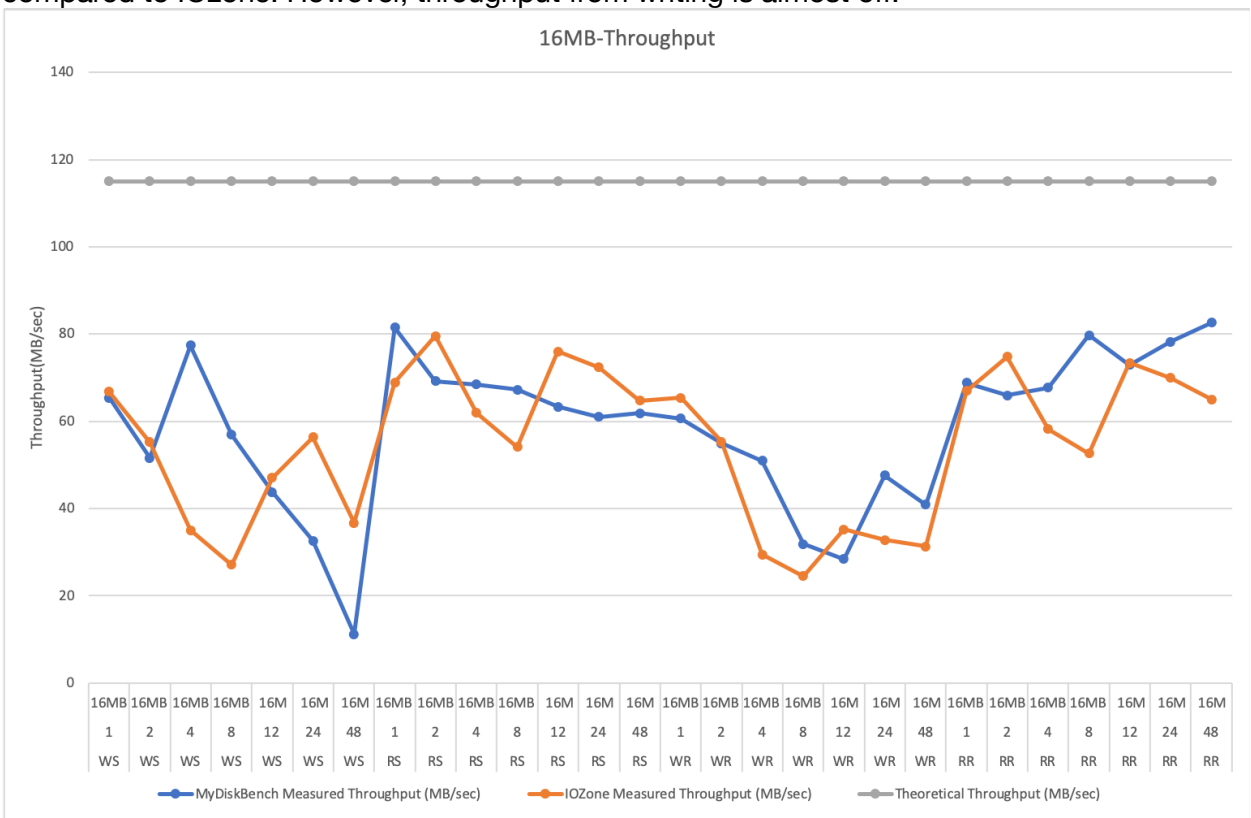
Graphical Figures

With 64KB record size, the MyDiskBench performance is a slightly off compared to IOzone, but we can notice that it has similar trend in general.



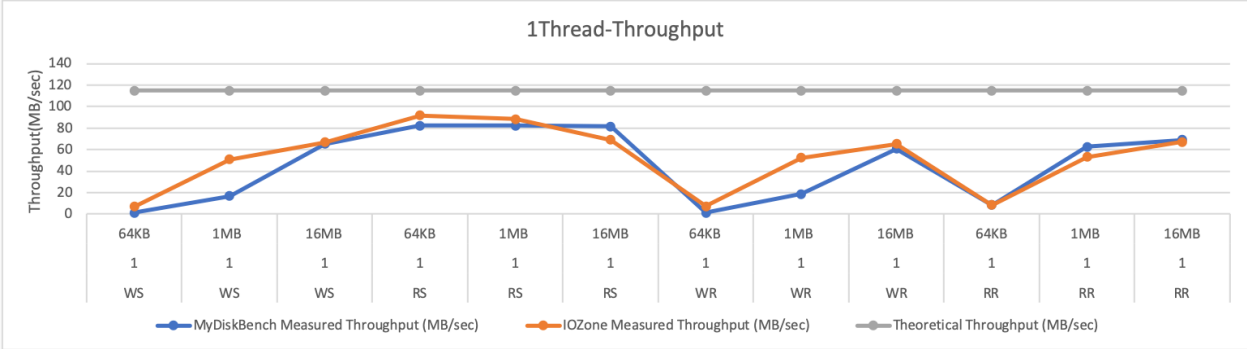


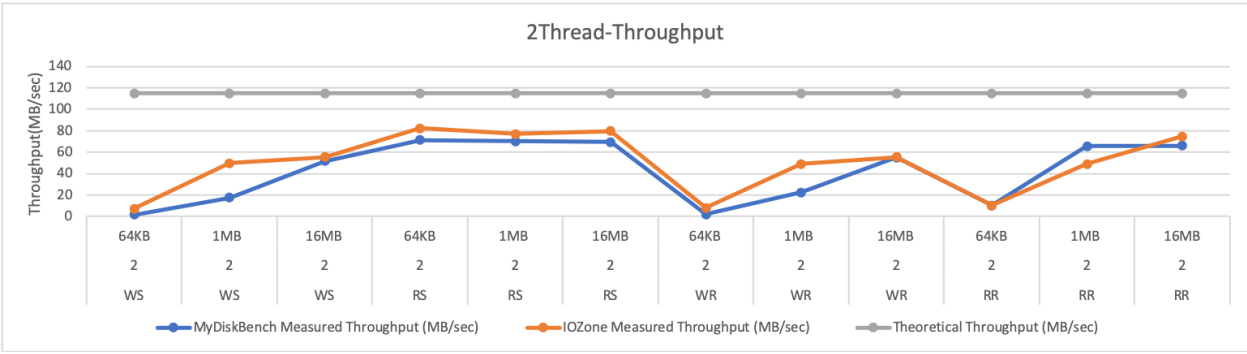
When the record size goes up to 1MB, throughput from reading keeps similar throughput compared to IOzone. However, throughput from writing is almost off.

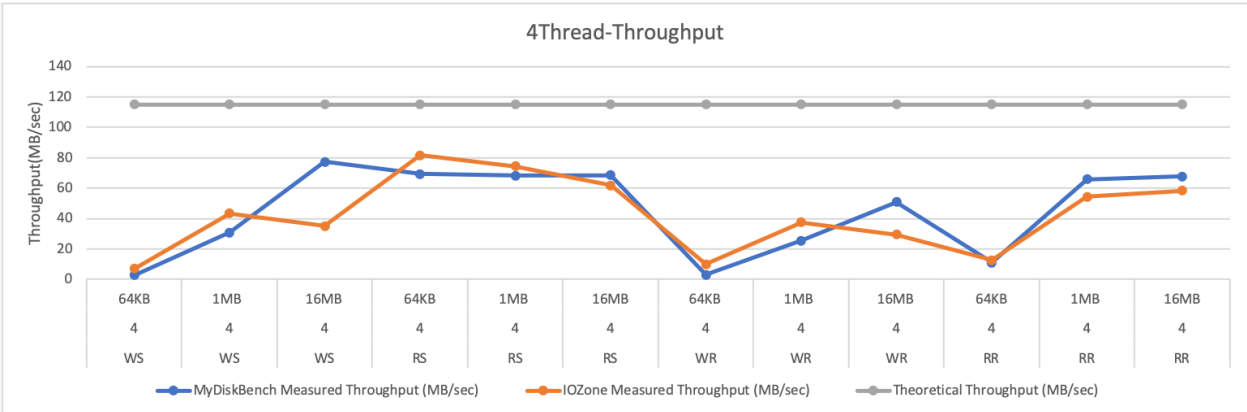


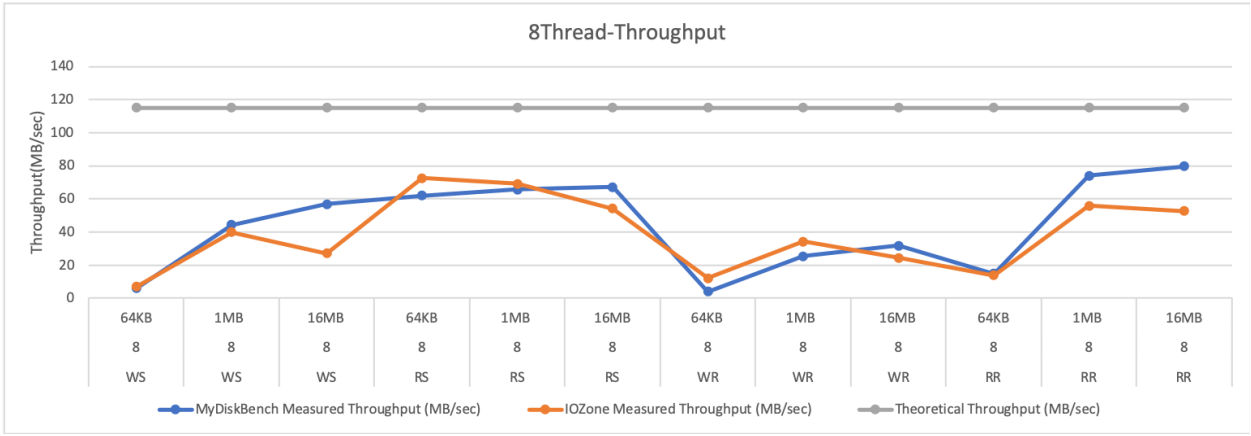
The inefficiency of MyDiskBench becomes huge when its record size grow up to 16MB. Especially at writing. In order to precisely describe and investigate why these happened, we decided to plot based on the number of threads, workload types and workload types plus record sizes.

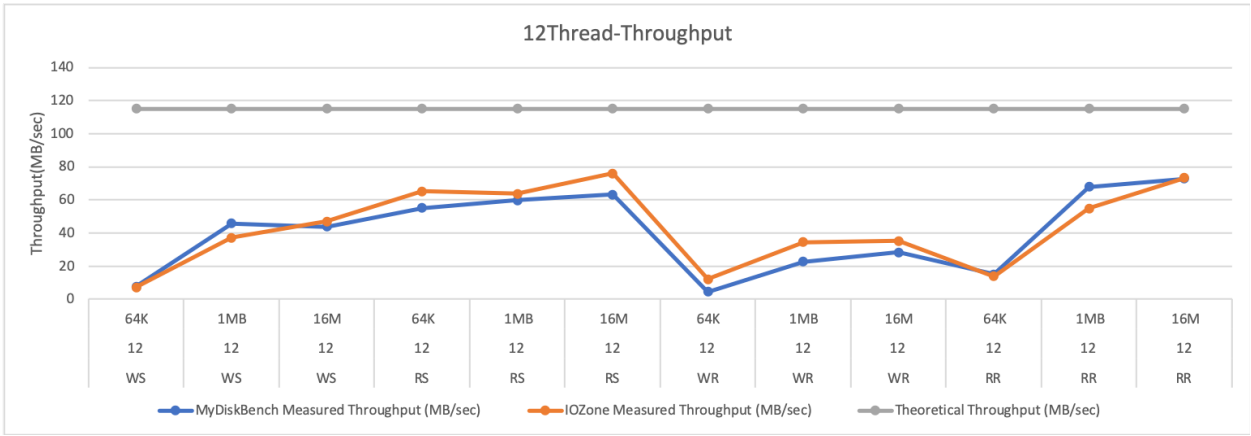
THREADS

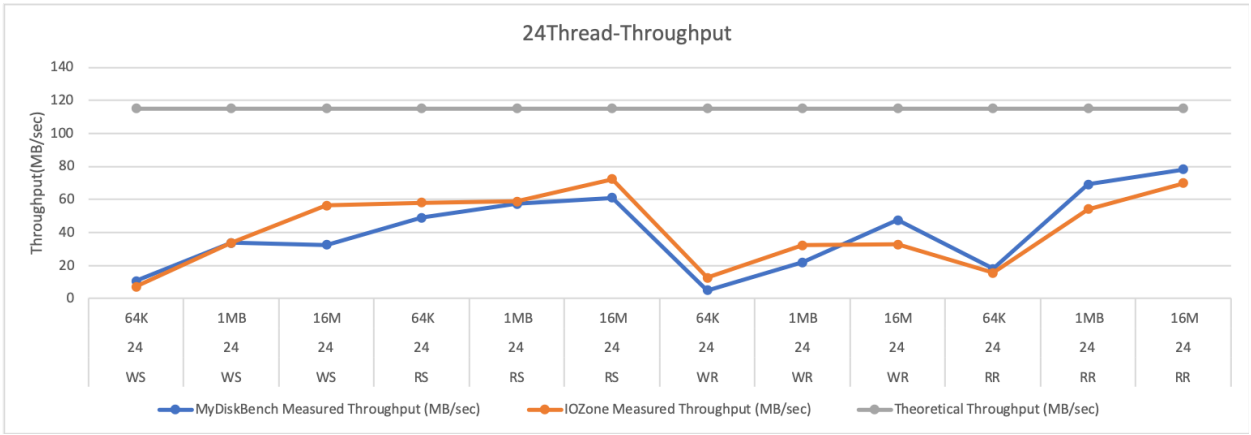


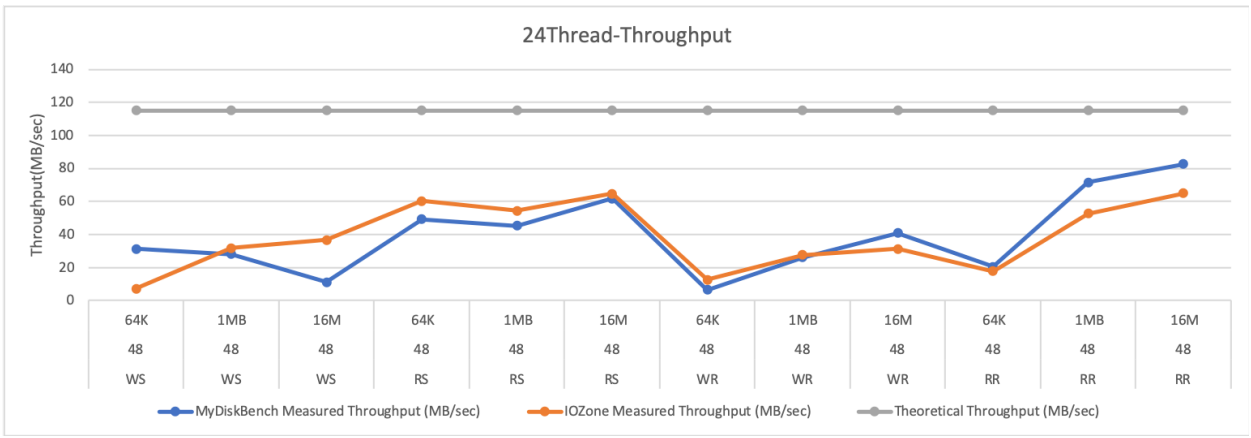





8Thread-Throughput
Throughput(MB/sec)
140
120
100
80
60
40
20
0
64KB 1MB 16MB 64KB 1MB 16MB 64KB 1MB 16MB 64KB 1MB 16MB
8 8 8 8 8 8 8 8 8 8 8 8
WS WS WS RS RS RS WR WR WR RR RR RR
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)

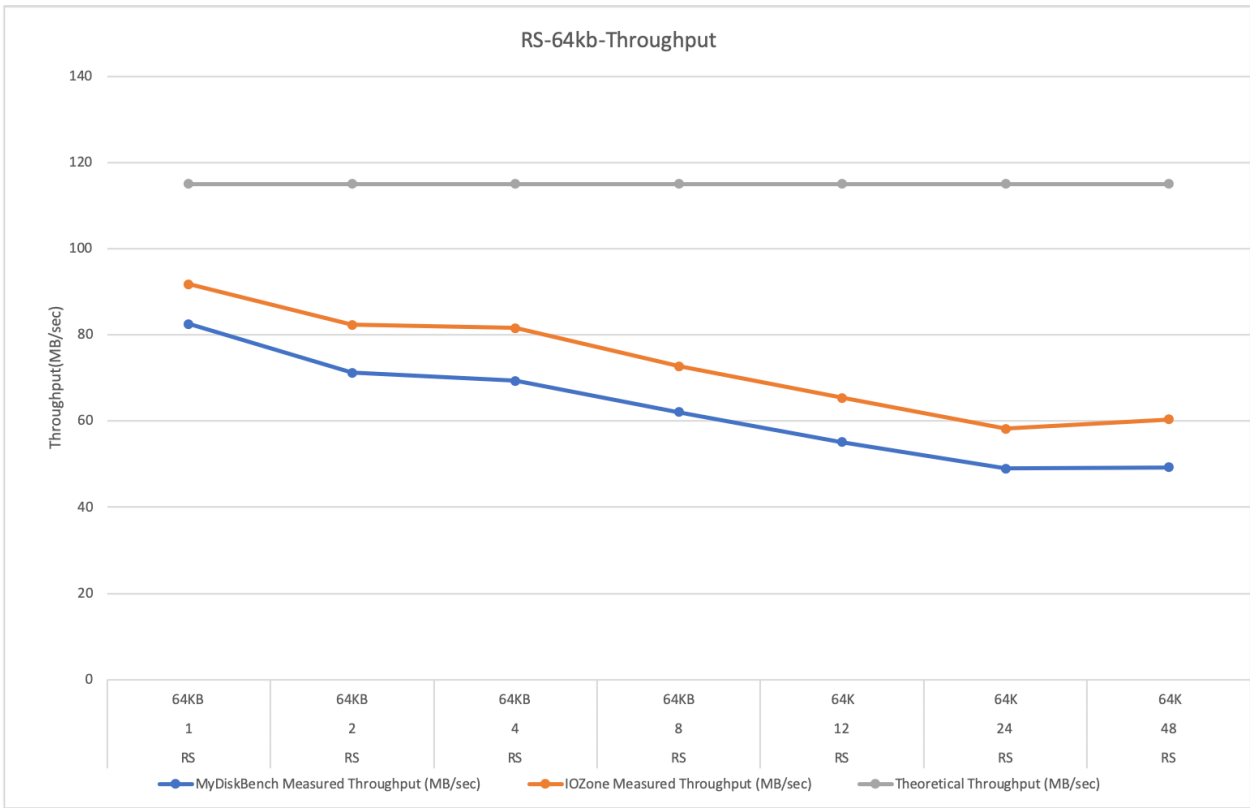

12Thread-Throughput
Throughput(MB/sec)
140
120
100
80
60
40
20
0
64K 1MB 16M 64K 1MB 16M 64K 1MB 16M 64K 1MB 16M
12 12 12 12 12 12 12 12 12 12 12 12
WS WS WS RS RS RS WR WR WR RR RR RR
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)


24Thread-Throughput
Throughput(MB/sec)
140
120
100
80
60
40
20
0
64K 1MB 16M 64K 1MB 16M 64K 1MB 16M 64K 1MB 16M
24 24 24 24 24 24 24 24 24 24 24 24
WS WS WS RS RS RS WR WR WR RR RR RR
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)

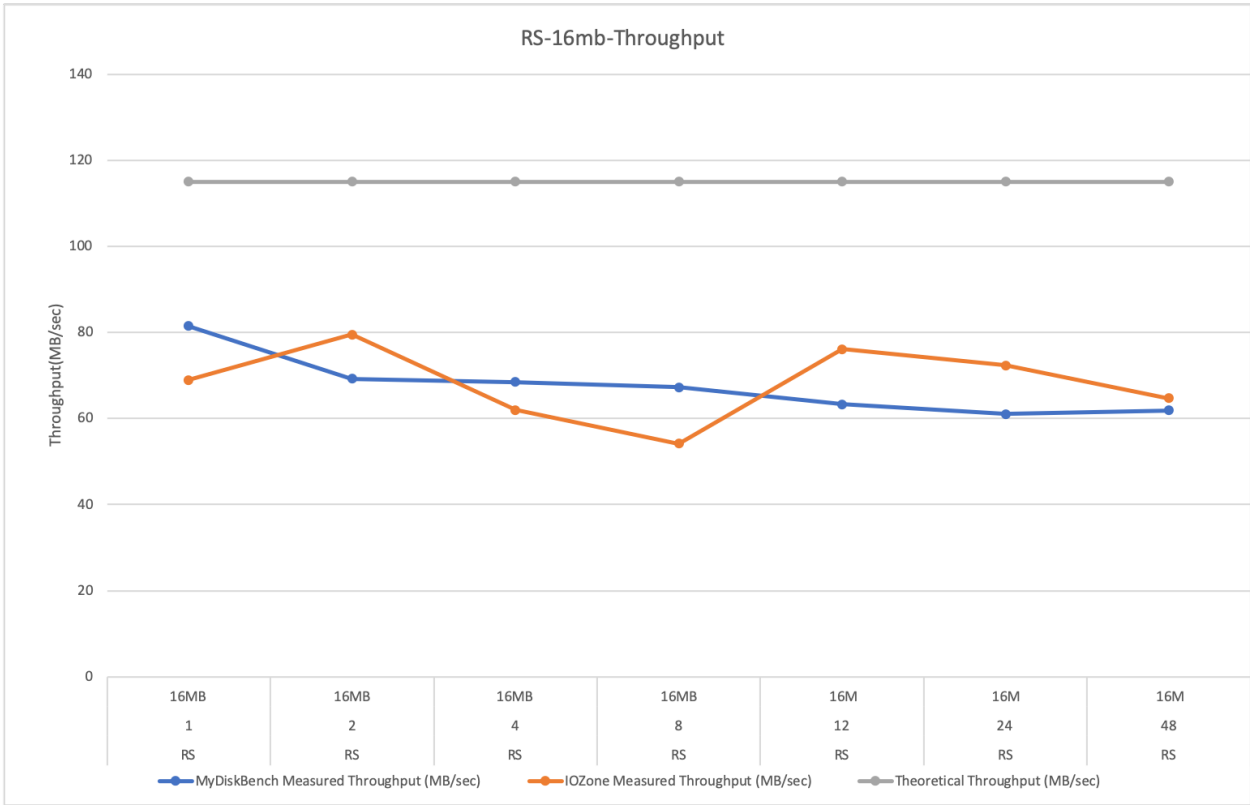


From threads plot, we can easily notice that WS, WR at 1MB and 16MB are the most one where emerge being unaligned. Our guess of this situation is that we invoke o_sync() after write() in order to ensure that our I/Os become synchronized. In addition, the different size of I/O in OS could be the one causing this inefficiency. In order to dig more why different workload types have different efficiency, we start investigating Sequential-Read first.

SEQUENTIAL-READ

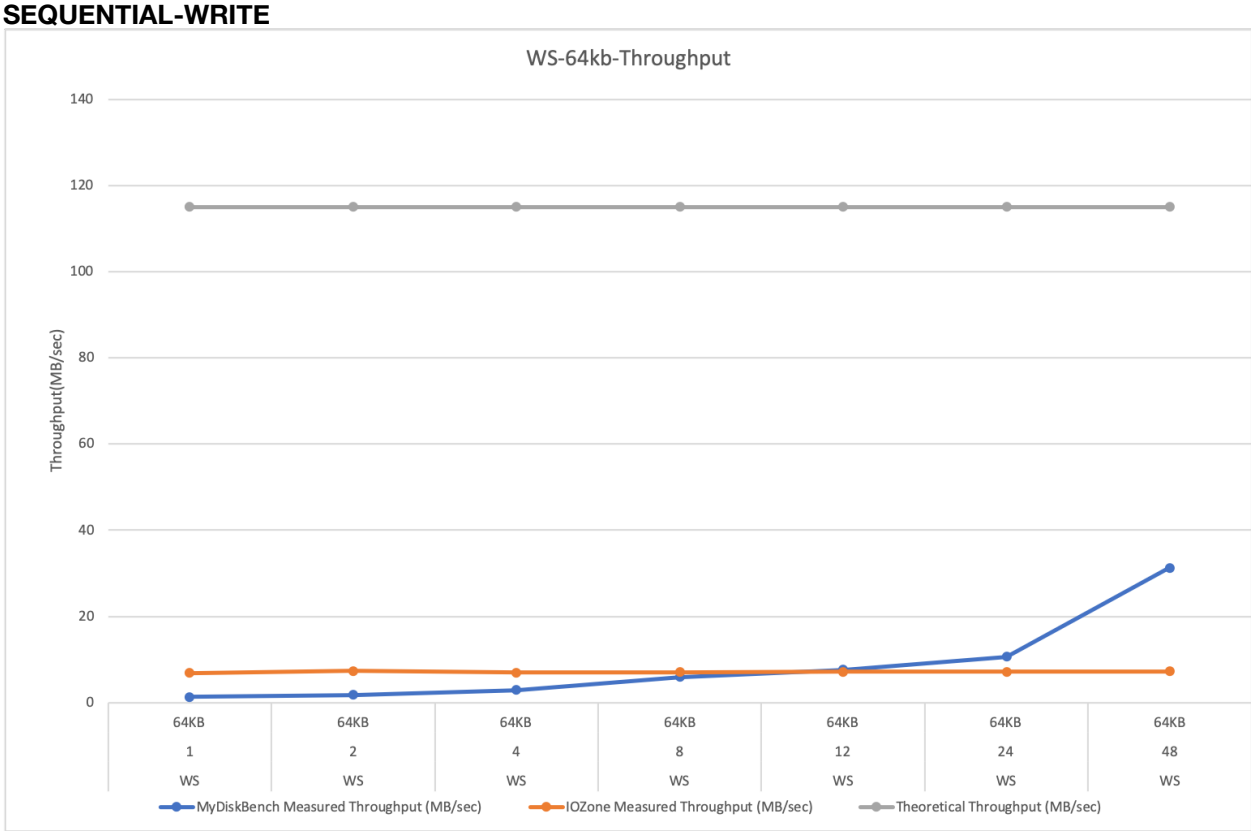


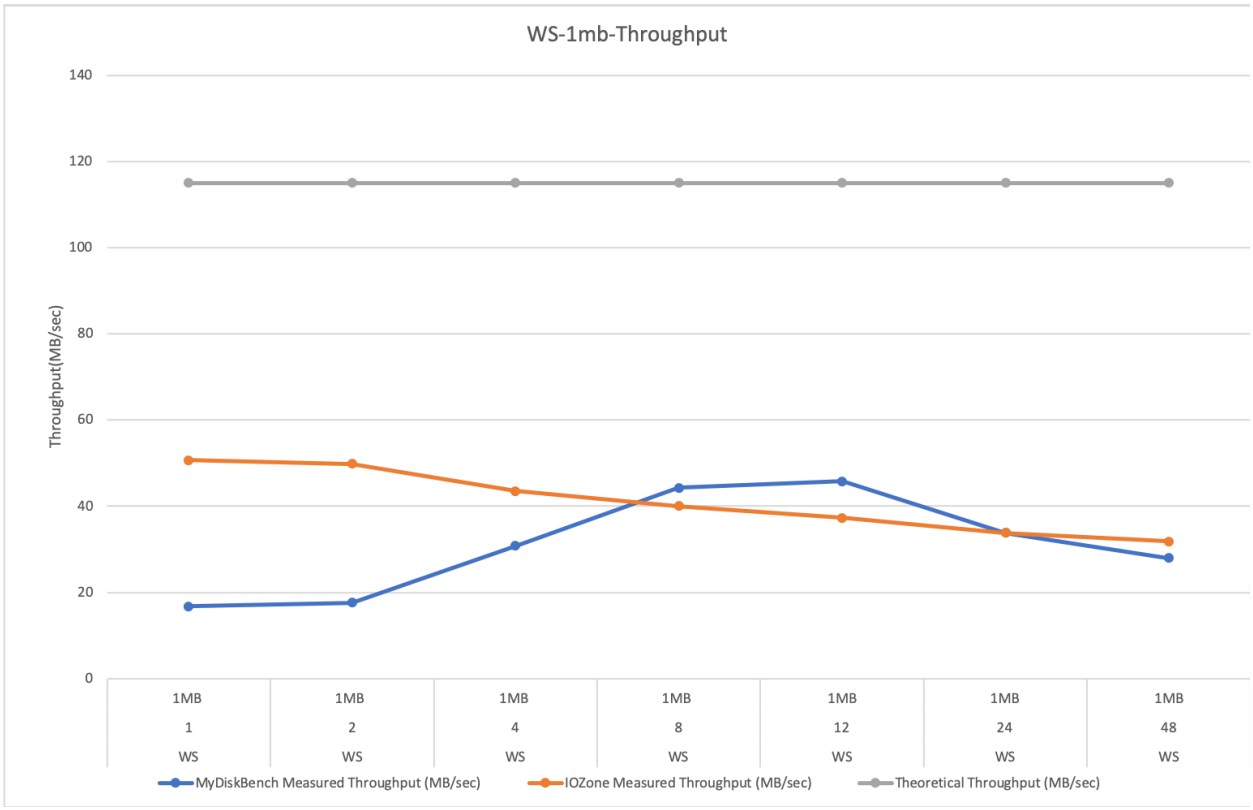

RS-1mb-Throughput
140
120
100
80
60
40
20
0
Throughput(MB/sec)
1MB 1 RS
1MB 2 RS
1MB 4 RS
1MB 8 RS
1MB 12 RS
1MB 24 RS
1MB 48 RS
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)

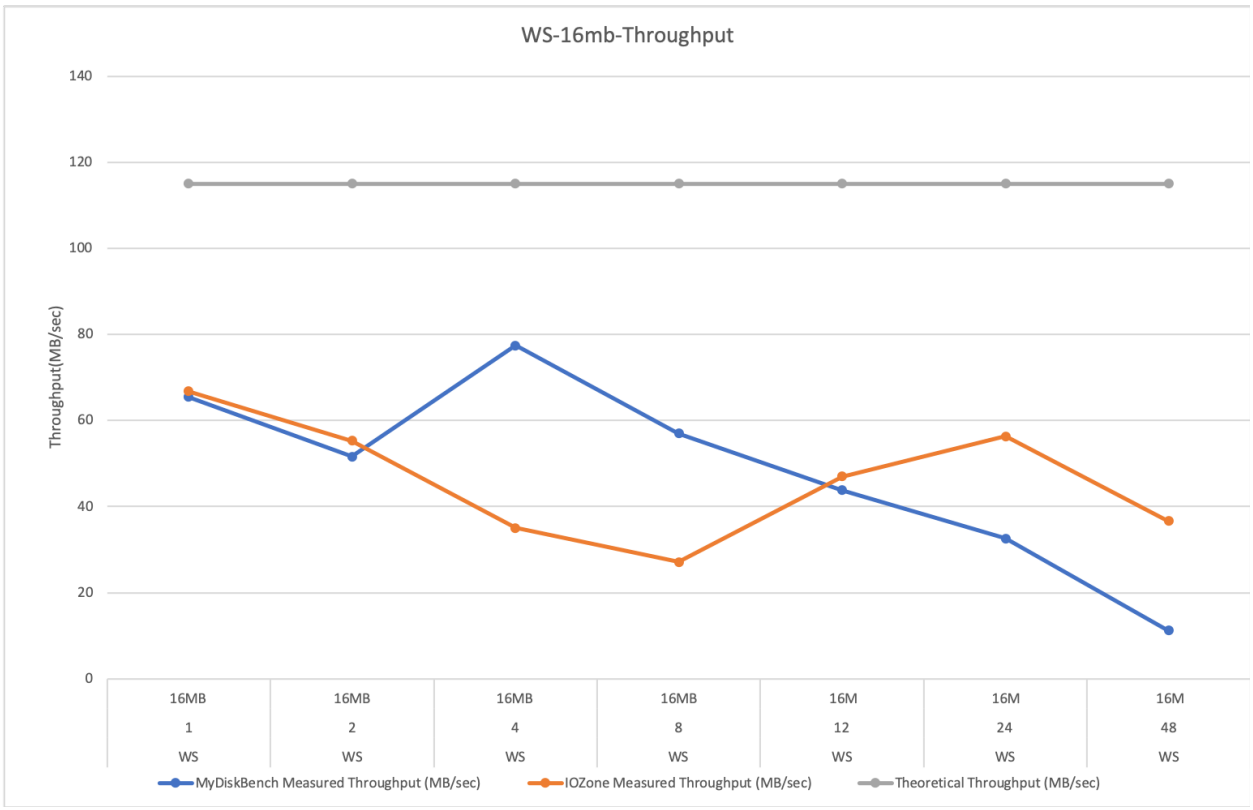

RS-16mb-Throughput
140
120
100
80
60
40
20
0
Throughput(MB/sec)
16MB 1 RS
16MB 2 RS
16MB 4 RS
16MB 8 RS
16M 12 RS
16M 24 RS
16M 48 RS
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)

As these above plots depict, our MyDiskBench performance is stable over different record sizes, but performance of Iozone at 16mb fluctuate irregularly. We can conclude that we can ensure that reading is definitely doing synchronized I/O operation because it must go to the disk to read. Therefore, there is no doubt.

SEQUENTIAL-WRITE

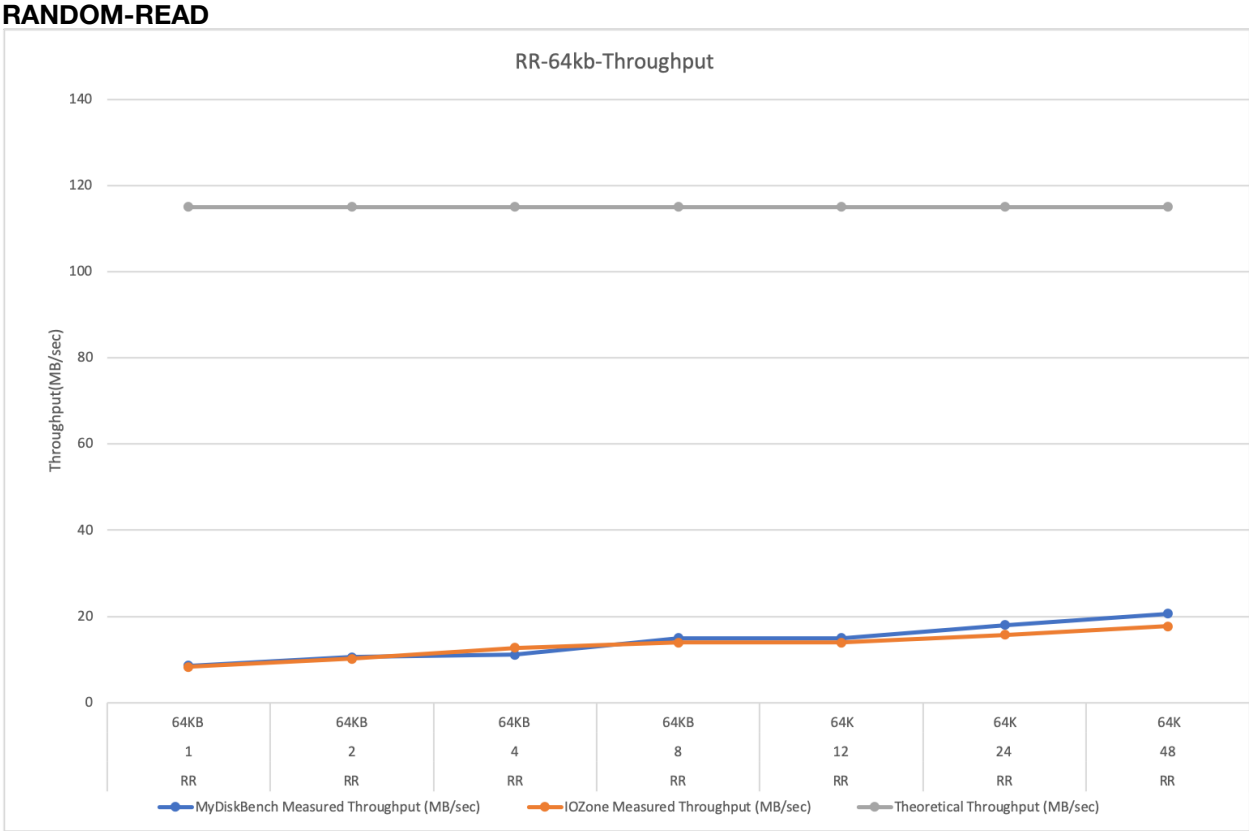


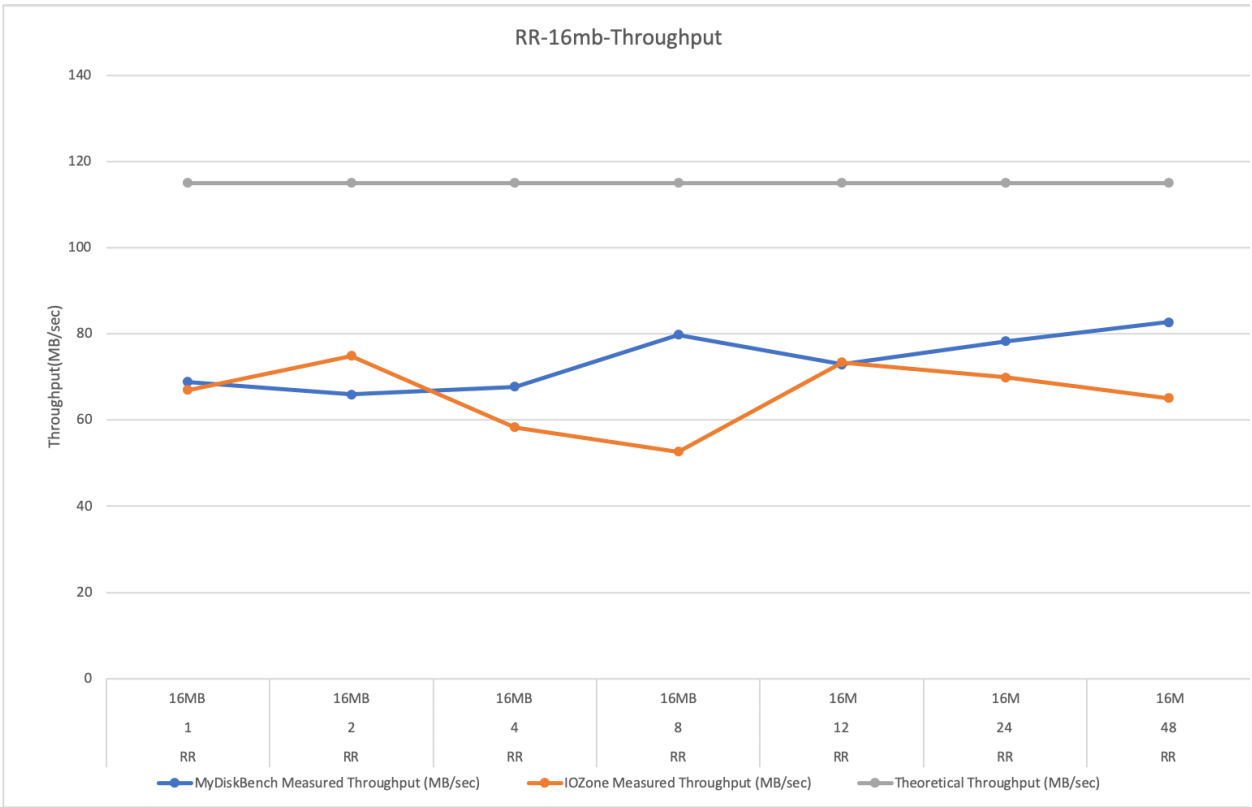

WS-1mb-Throughput
140
120
100
80
60
40
20
0
Throughput(MB/sec)
1MB 1 WS
1MB 2 WS
1MB 4 WS
1MB 8 WS
1MB 12 WS
1MB 24 WS
1MB 48 WS
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)

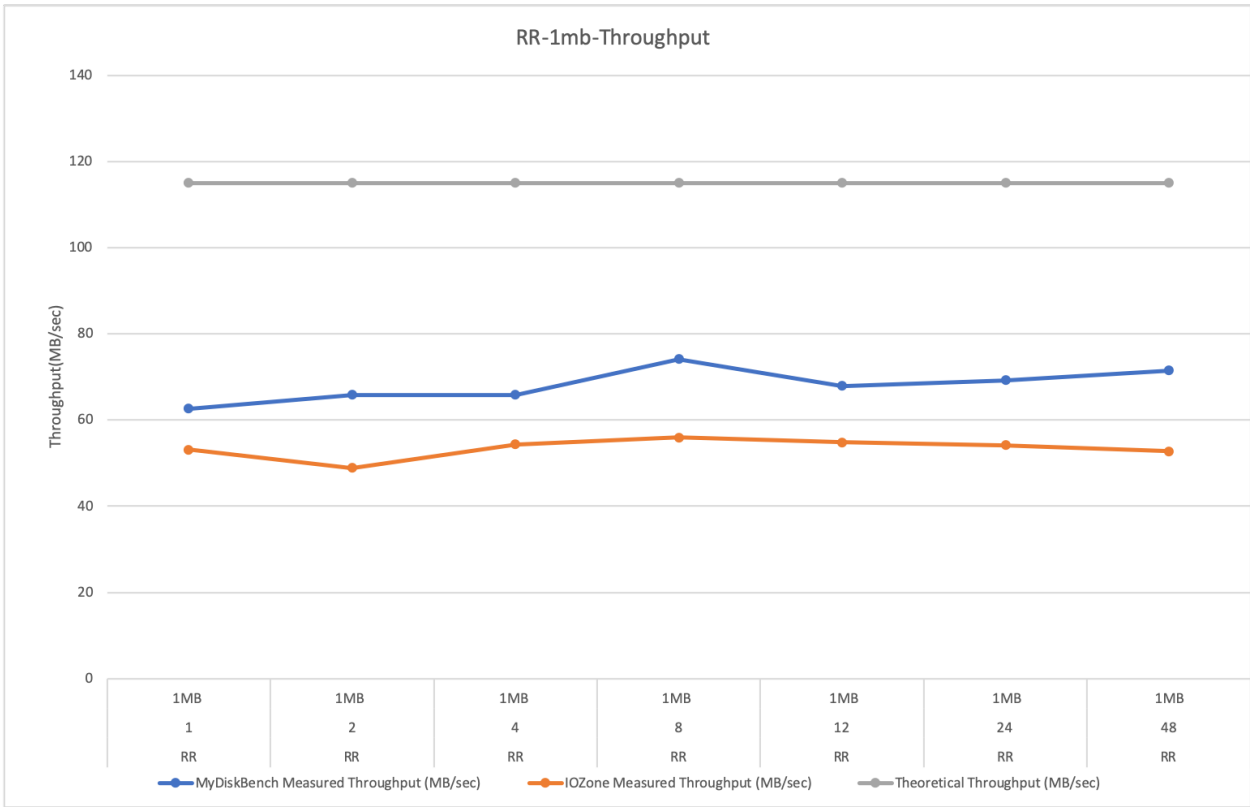

WS-16mb-Throughput
140
120
100
80
60
40
20
0
Throughput(MB/sec)
16MB 1 WS
16MB 2 WS
16MB 4 WS
16MB 8 WS
16M 12 WS
16M 24 WS
16M 48 WS
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)

Sequential-Write is one of the one that our team has been disappointed with the result. As the number indicates, it does not measure the speed of memory. However, we can't say that it did not get any help from memory at all because there are some points that numbers get doubled. Although we tried O_DIRECT flag and ran echo 3 > /sys/proc/vm/drop_caches both, it implies that there may be more things we need to control in order to bypass cache effect especially writing.

RANDOM-READ

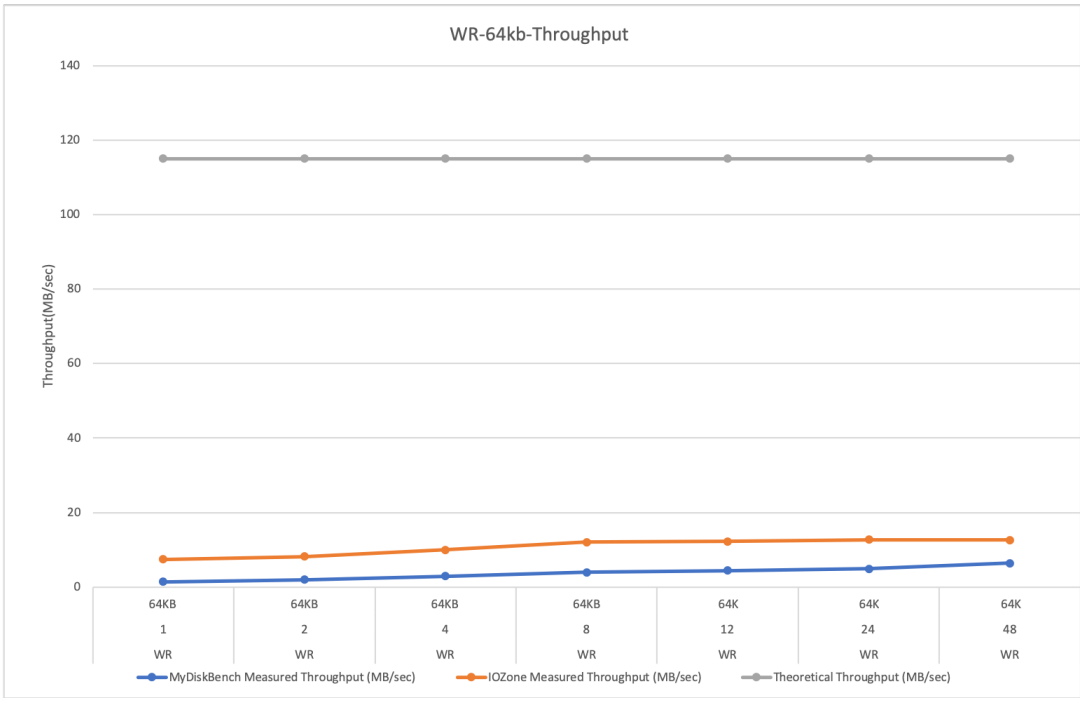

RR-64kb-Throughput
Throughput(MB/sec)
140
120
100
80
60
40
20
0
64KB 1 RR
64KB 2 RR
64KB 4 RR
64KB 8 RR
64K 12 RR
64K 24 RR
64K 48 RR
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)

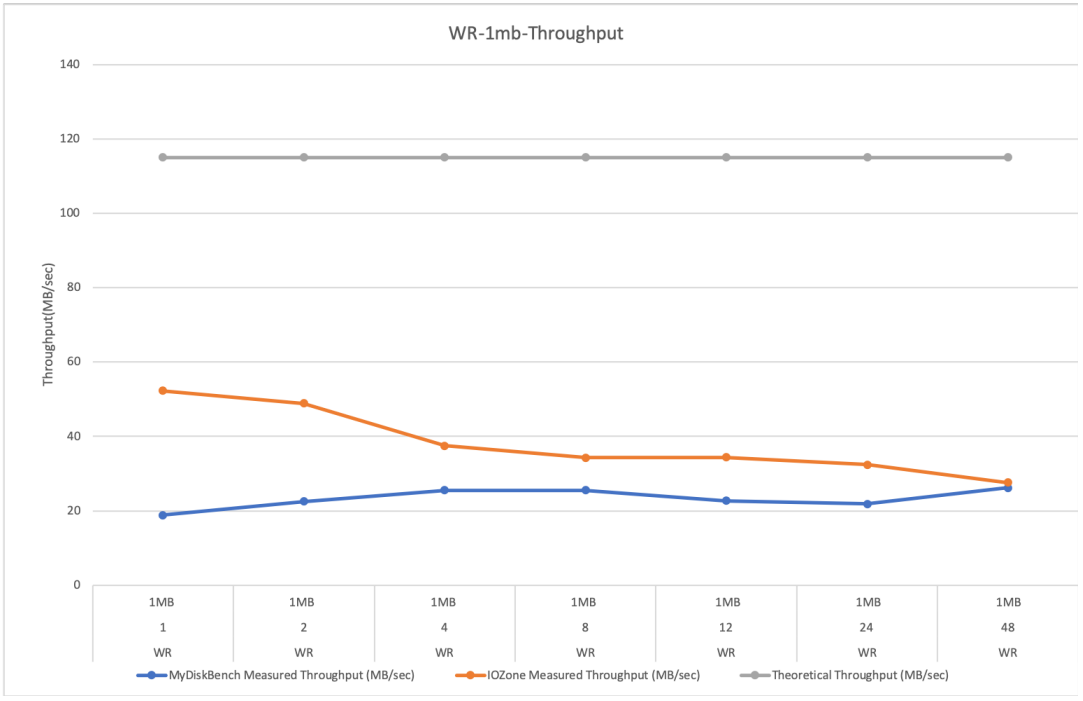

RR-16mb-Throughput
Throughput(MB/sec)
140
120
100
80
60
40
20
0
16MB 1 RR
16MB 2 RR
16MB 4 RR
16MB 8 RR
16M 12 RR
16M 24 RR
16M 48 RR
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)

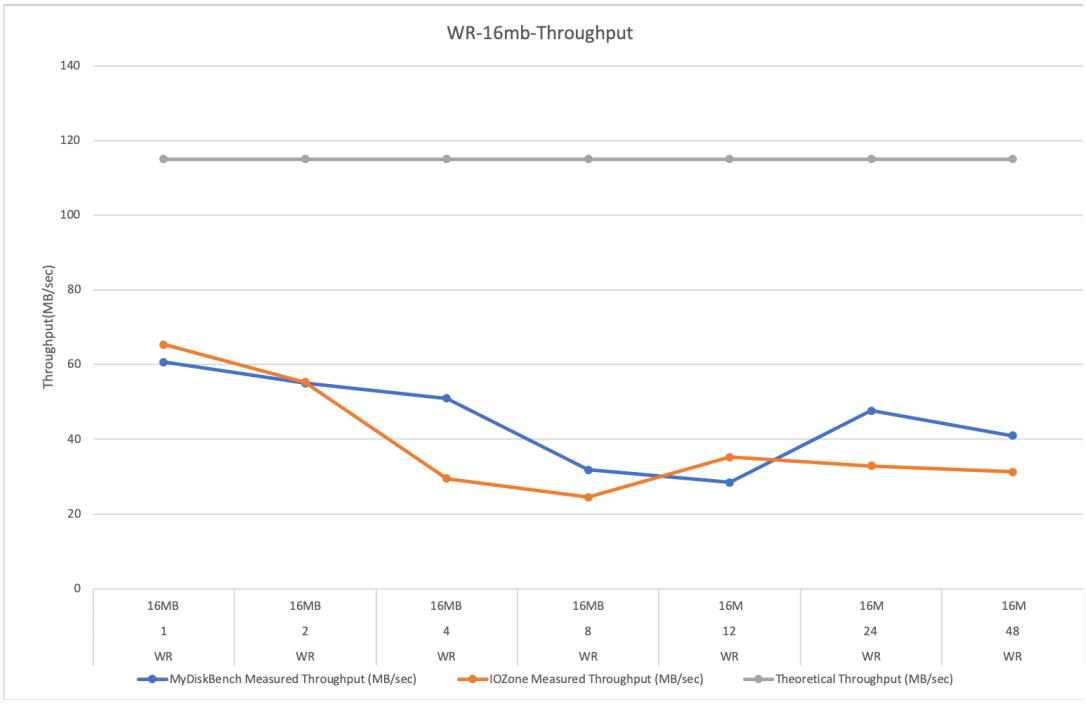


Although its trend goes off a little bit with 16MB record size, it produces stable throughput as we expected.

RANDOM-WRITE

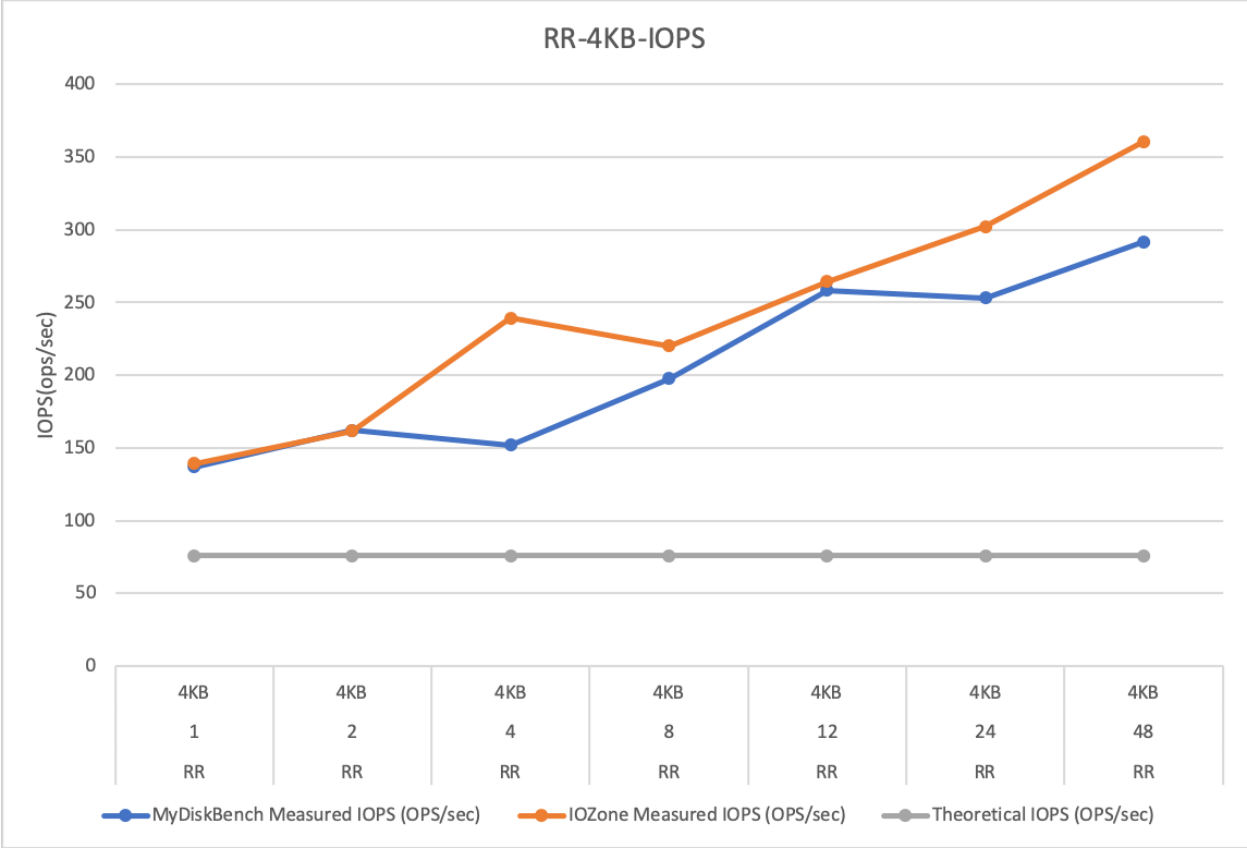

WR-64kb-Throughput
140
120
100
80
60
40
20
0
Throughput(MB/sec)
64KB 1 WR
64KB 2 WR
64KB 4 WR
64KB 8 WR
64K 12 WR
64K 24 WR
64K 48 WR
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)


WR-1mb-Throughput
140
120
100
80
60
40
20
0
Throughput(MB/sec)
1MB 1 WR
1MB 2 WR
1MB 4 WR
1MB 8 WR
1MB 12 WR
1MB 24 WR
1MB 48 WR
MyDiskBench Measured Throughput (MB/sec)
IOZone Measured Throughput (MB/sec)
Theoretical Throughput (MB/sec)



From random-write, we can notice that the total throughput goes up when the number of threads goes up as well. However, our MyDiskBench performance is really really lower than IOzone with 64KB. We guess this is because of the fail of random seeking. We have encountered that write operations often fail by generating error “Invalid argument”. We could reproduce this error when open has O_DIRECT flag. Thus, our solution was doing seek operation until write doesn’t fail. In addition, the lower record size is the often it could happen. Therefore, 64KB random read performance is significantly lower than others. We can see this situation at 4KB oops experiments as well.

RANDOM-READ IOPS



RANDOM-WRITE IOPS