Test of Intel's X25 Flash SSD

February 2009

By Solid Data Systems, Inc.

Contents:

Overview

There has been much said and debated over the promise of Flash-based SSD. The vendors of these drives tout the superior performance of Flash over rotating disk and herald them as the heir apparent for enterprise data centers where high performance is critical. This white paper provides a closer examination of the true performance capabilities of Flash drives, illustrating the effects of high-volume transactions and the inherent "wear-leveling" that Flash disks incorporate to prolong their inevitable chip degradation.

Test of Intel X25 80 Gigabyte Flash Drive

Table 1 shows the performance of the Intel X25 Flash drive under various test conditions. Interpretation and discussion of these test results follow.

	Read IOPS New	Read IOPS EoT	Read MB/s New	Read MB/s EoT	Write IOPS New	Write IOPS EoT	Write MB/s New	Write MB/s EoT	Write IOPS New	Write MB/s New
IOPS / Process	10K	10K	10K	10K	10K	10K	10K	10K	250K	250K
Block Size										
512	9115	4050	4.6	2.0	5508	446	2.8	0.2	851	0.42 MBs
1024	8205	3597	8.2	3.6	4432	843	4.4	0.8	400	0.40 MBs
2048	6878	3036	13.8	6.1	3761	891	7.5	1.8	223	0.45 MBs
4096	5156	2255	20.6	9.0	3038	630	12.2	2.5	203	0.81 MBs
8192	4538	2006	36.3	16.1	2490	460	19.9	3.7	177	1.42 MBs
16384	3856	1854	61.7	29.7	1949	280	31.2	4.5	133	2.13 MBs
32768	2953	1626	94.5	52.1	1369	172	43.8	5.5	50	1.61 MBs
65536	1873	1244	119.9	79.6	831	147	53.2	9.4	29	1.89 MBs
131072	604	404	77.4	51.8	471	76	60.4	9.8	18	2.33 MBs

Table 1

	Write IOPS New	Write IOPS New
IOPS / Process	10K	250K
Block Size		
512	5508	851
1024	4432	400
2048	3761	223
4096	3038	203
8192	2490	177
16384	1949	133
32768	1369	50
65536	831	29
131072	471	18

Table 2

Performance as a Function of Transaction Volume

The initial read and write tests showed good performance. In fact, the write performance was too good for what could realistically be expected for the underlying flash memory. In order to evaluate the performance under a higher volume of writes, we increased the volume of IOPS from 10K to 250k per process. This increased volume made our test environment more indicative of the loads that could be generated in a high-transaction rate application and allowed the performance of the underlying flash to be evaluated. The implication of this, as shown in Table 2, is that as long as the amount of write data is small, performance is good. However, when the transaction volume is increased the write performance suffers dramatically — and at standard database block sizes (4096 and 8192) is no better than rotating disk.

Effect of Wear-Leveling and Fragmentation

After running several write tests, we noticed that both read and write performance deteriorated. Flash storage blocks can only be written so many times before they fail and, as computers can re-write the same block at high rates, the technique of wear-leveling must be used. To manage the wear-leveling capability, a table containing the number of times a block has been written must be maintained. So rather than writing to the same block, a processor directs the writes around to the least written block. Thus, writes get distributed all over the drive and their locations are unpredictable, as the location of the least written block depends on past write patterns. Reads also slow down, as the table must be accessed to find the new location of the desired block. The more the drive is written to, the more fragmented the flash storage blocks become. So over time, it takes longer and longer to read and write them. Table 3 shows both reads and writes when the drive was first tested and at the End of Test (EoT). Note that the performance degradation due to this fragmentation occurred even at light data volumes. In particular, note that the performance of the drive on writes deteriorated dramatically. So a client will get great performance during initial tests, only to see the performance deteriorate the longer the drive is in production.

	Read IOPS New	Read IOPS EoT	Write IOPS New	Write IOPS EoT
IOPS / Process	10K	10K	10K	10K
Block Size				
512	9115	4050	5508	446
1024	8205	3597	4432	843
2048	6878	3036	3761	891
4096	5156	2255	3038	630
8192	4538	2006	2490	460
16384	3856	1854	1949	280
32768	2953	1626	1369	172
65536	1873	1244	831	147
131072	604	404	471	76

Table 3

Because the benchmark tool can issue pure reads or writes, the read performance often looks very good. This can lead to the mistaken view that the drive could be used in a highly read-oriented environment. What is missed, though, is that even a small percentage of writes dramatically lowers the read performance. Running tests in which a combination of read/write commands were issued, the IOPS performance dropped from 2006 IOPS in pure read mode to 708 IOPS in 80/20 read/write on 8k blocks.

DRAM vs. Flash

It is clear that the benefits of flash are very dependent on the write volume, and both read and write performance deteriorate over time. The change in the write performance occurred by writing continuously for only one weekend. Solid Data DRAM-based SSDs have none of these performance irregularities and provide the same high performance, regardless of application characteristics. They also do not have any wear out characteristics and operate at full performance for their service life, which often exceeds five years.

	Read IOPS Solid Data	Write IOPS Solid Data	Read IOPS EoT Intel	Write IOPS EoT Intel
IOPS / Process	10K	10K	10K	10K
Block Size				
512	23584	20768	4050	446
1024	22068	17391	3597	843
2048	17066	14988	3036	891
4096	12641	11365	2255	630
8192	8299	7400	2006	460
16384	3980	3582	1854	280
32768	2689	2371	1626	172
65536	1288	1140	1244	147
131072	685	604	404	76

Table 4