HOW (AND WHY) NEW QLC SATA SSDs
DELIVER 8X FASTER MACHINE LEARNING

Tony Yoo
Colfax International

October 10, 2018

Abstract

This paper reports the performance of a machine learning application relying on the new Micron® 5210 ION enterprise SATA SSD. The test workload involves intensive parallel reading from the drive, which represents the data preprocessing pipeline for machine learning. Micron 5210 ION SSDs are based on the novel Quad Level Cell (QLC) NAND technology, which enables a capacity of 8 TB in a 2.5" form-factor SSD

Search: test results and analysis for typical ML workloads: image classification, data pre-processing, ingest

Table of Contents

WHO WE ARE

Colfax Research is a department of Colfax Internationa, a Silicon Valley-based provider of novel computing systems. Our research team works to help you leverage new hardware and software tools to harness the full power of computational innovations.

WHAT WE DO

We work independently as well as collaborate with other researchers in science and industry to produce case studies, white papers, and educational materials with the goal of developing a wide knowledge base of the applications of current and future computational technologies. In addition, we run educational programs, provide consulting services, and offer specialized hosting for technology adoption programs.

PUBLICATIONS

colfaxresearch.com/research

TRAINING

colfaxresearch.com/training

SERVICES

colfaxresearch.com/services

1. Introduction

For years, 7200 RPM hard disk drives (HDDs) have been the standard media on which Machine Learning (ML) training data sets have been stored. These traditional HDDs have been preferred due to their low cost and easy to adopt, familiar SATA interfaces. However, their relatively slow throughput and resultant data exgest rates were key tradeoffs. Solid State Drives (SSDs) have been too expensive to justify their potential gain. Despite SSD cost / GB decreases, their overall cost has remained just out of reach for all but the most demanding (and expensive) ML platforms. This barrier is starting to fall with the introduction of the world's first QLC SSD, the Micron® 5210 ION enterprise SATA SSD.

In our paper, we record performance measurements on Micron 5210 SSD related to ML workflow. Although ML is highly CPU intensive, fast storage can lower training time though faster file pre-processing and serialization, particularly when the size of a data set exceeds the amount of installed memory. A popular format for datasets is TFRecord, and in our performance measurements, we will be comparing the throughput speed and completion time of a TFRecord on a 7.68TB Micron 5210 ION SSD versus that of an 8TB Seagate 7200RPM HDD.

2. Micron QLC SSDs

Micron QLC SSDs are a whole new type of SSD that uses Quad Level Cell (QLC) NAND technology. QLC flash stores 4 bits of data per cell enabling higher density. While more bits per flash cell does decrease the amount of write cycles available, it adds the benefit of higher capacity and the fast read speeds. For read-intensive ML workloads such as those analyzed here, lower write cycles are far less important than the cost decrease that QLC technology enables.

QLC SSDs are aimed to eventually replace standard HDDs in read intensive applications by providing SSD read speed with HDD capacity. They come in 2.5-inch form factor and the capacities range from 1.92TB to 7.68TB at an approachable cost.

3. Test System Configuration

Our test configuration is a standard, x86 server platform equipped as noted in Table 1.

Component	Details
CPU	2x Intel® Xeon® Platinum 8168 CPU @ 2.70GHz
Cores	48 cores with 2-way Intel® Hyper-Threading Technology
Memory	96 GB DDR4
Drives	7.68TB Micron® 5210 QLC SSD 8TB Seagate® 7200 RPM HDD

Table 1: System Configuration

4. Test Workload: TFRecord

TFRecord is the standard format for TensorFlow. It is used to store large amounts of data (for example, a collection of images) in a single TFRecord file, which can be read from storage faster than individual files and loaded into TensorFlow in batches for training.

In our test workload, we read large high-resolution images in the TIFF format from the drive using OpenCV, rescale each image to low resolution, append the resulting image to a TFRecord, and write the resultant TFRecord object back onto the drive. This process simulates the preprocessing of real-world data into a format suitable for deep learning training. To reach the performance limits of our system, we designed our TFRecord creation script to ingest images in parallel using Python multiprocessing. When we run the code, we can specify the number of concurrent processes that read and resize the images. Parallel processing of data reading and conversion taps into the additional performance of both the drive and the multi-core CPU. We can tune this parallelism to utilize and maximize available read speed and CPU processing power.

5. Test Results

First, we fixed the number of processes and increased the number of images to show scale in performance and utilization. Each run was tested 3 times to minimize variation in speed and time. This resulted in consistent results over various runs.

Table 2 shows the bandwidth and time for each test and drive by process count. The sample size is 1,000 images at 22MB each, totaling 23GB data set size.

Number of Processes	Micron 5210 ION 7.68TB		Typical 7200 RPM HDD 8TB	
	Bandwidth, MB/s	Time, s	Bandwidth, MB/s	Time, s
1	24.4	981	24.2	898
2	47.7	501	39.4	607
3	69.2	345	49.2	523
4	91.6	261	47.5	507
6	116	206	50.1	507
8	174	137	52.1	504
10	208	115	53.5	501
14	287	83.2	54.5	494
18	320	74.6	55.5	487
22	348	68.7	55.8	485
26	362	66.1	56.2	478
30	368	64.9	57.0	459
40	353	67.7	57.9	448
50	355	67.3	47.5	439
60	364	65.6	49.4	431
70	367	65.1	47.2	429
80	376	63.6	48.5	426
90	377	63.3	45.8	420
100	375	64.0	47.7	413

Table 2: Bandwidth and Test Time

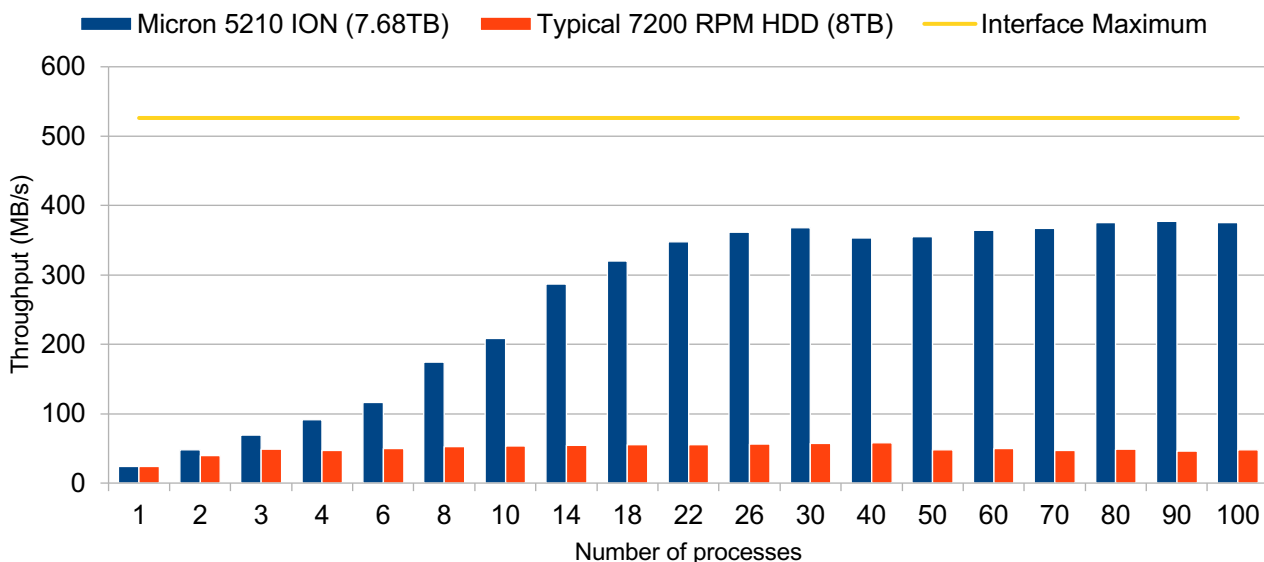## TFRecord Creation: Throughput vs Processes

Figure 1: Large image resizing and TFRecord formation

At a single process, the Micron 5210 and the HDD performed similarly (a rate of 1 image per second), resulting in a run time of 980 seconds. With one process, the performance is limited by the CPU and not the drive read speed. As the number of concurrent processes increased, the throughput of the Micron 5210 SSD also increased until reaching a maximum sustained read speed at around 30 processes.

Parallel processing did not significantly accelerate image conversion on the HDD, which achieved at best, a 2.3x speedup compared to single-process performance. We attribute that to the nature of a spinning drive, which cannot efficiently handle parallel read operations. In contrast, the quick read speed and the parallel nature of the Micron 5210 SSD managed many more queues compared to the HDD with a commensurate performance improvement displayed. At 90 concurrent processes, the Micron 5210 was 8.2x faster than the HDD and 15.5x faster than either drive when running only 1 process.

Next to fine tune our results, we tested all the process counts between 90 and 100 to find the optimal parallelism (observed when speed reached its maximum). These results are shown in Table 3.

Number of Processes	Micron 5210 ION 7.68TB	
	Bandwidth, MB/s	Time, s
91	388.9	61.5
92	382.3	62.6
93	391.1	61.2
94	386.0	62.0
95	391.3	61.2
96	391.7	61.1
97	390.1	61.3
98	391.1	31.1
99	389.6	31.4

Table 3: 5210 process count > 90

Performance results in Table 3 demonstrate how the process count affects bandwidth. Here we noted that the performance peaked at 96 processes. Running the test script multiple times showed some variation where 97 processes was faster, but 96 processes consistently showed better results.

Because 96 processes showed the highest throughput for any number of images, we noted this and used this value as the base for our next performance measurement. Note that 96 processes may have showed the highest throughput given that it's related to the number of logical CPUs available in the server (48 cores each running 2 hyper-threads).

Using the optimal process count of 96, we then ran TFRecord on various numbers of images. Given that many ML workloads involve very large numbers of images, we tested to ensure real-world applicability. For example, skin cancer detection algorithms (that are more accurate than doctors) have been built doing ML training on 100,000 images, so we used 100,000 images as a barometer of real-world results due to the size of the dataset.

The results for image sets above 10,000 are the key numbers to focus on below, as in all instances, storage is effectively isolated as the test variable because the size of the dataset exceeds the available DRAM, ensuring that memory has no bearing on these results. Given that many ML datasets are often far larger than a system's installed memory capacity, these results are a good indicator of the actual performance differences between the storage drives being compared. Results are shown below in Table 4.

Test Parameters			7.68TB Micron 5210		8TB HDD		Performance Delta	
Number of Images	Number of Processes	Size of Dataset*	Bandwidth, MB/s*	Completion Time, s*	Bandwidth, MB/s*	Completion Time, s*	Speedup	Time Saved
1	96	23MB	23	1	22	1	1x	-
2	96	46 MB	24	1	23	1	1x	-
4	96	92 MB	46	1	42	1	1x	-
8	96	184 MB	113	1	49	3	2x	2 seconds
16	96	368 MB	255	1	48	7	5x	6 seconds
32	96	736 MB	340	2	44	16	9x	14 seconds
64	96	2GB	371	4	44	34	9x	30 seconds
128	96	3GB	378	8	47	65	8x	1 minute
256	96	6GB	389	16	44	138	9x	2 minutes
512	96	12GB	389	32	40	305	10x	5 minutes
1,000	96	23 GB	394	61	43	605	10x	9 minutes
10,000	96	230 GB	376	636	44	5,423	9x	1.3 hours
50,000	96	1150 GB	370	3,234	44	27,258	8x	6.7 hours
100,000	96	2300 GB	354	6,764	43	54,617	8x	13.3 hours

Table 4: Varying image counts tested. All data rounded to the nearest whole number.

Table 4 shows TFRecord reaches near max throughput when the dataset is greater than 64 images and stays fairly consistent all the way to 100,000 images. Comparing these results to an HDD, the time difference is substantial (especially at 100,000 as the Micron 5210 SSD is able to complete the workload in 6,764 seconds (113 minutes, or about 2 hours) versus the 54,617 seconds that the HDD takes (910 minutes, or about 15 hours). The delta between these times was 13.3 hours, and for image sets above 10,000, the QLC SSD performed 8x faster than the HDD.

6. Summary

In our study, with the Micron 5210 ION SSD, a read-intensive transformation of an image dataset with the purpose of a TFRecord file creation was accelerated by about 8x compared to a similar-sized HDD. In a 100,000 image dataset at 23MB per image, our test HDD took 15.17 hours to resize the images and pack them into a single TFRecord file while the Micron 5210 ION took only 1.88 hours to do the same task – 13 additional hours for the HDD to complete the same work.

Faster completion times can have dramatic effects on overall project value, delivery and expense. When analyzing the potential benefits of the Micron 5210 ION, a cost per GB analysis at the time of acquisition may be short-sighted. One should consider the total cost of ownership, the value of keeping expensive CPU assets busy, and the potential reductions in yearly operating expenses related to power and cooling. How often you do machine learning is also very important, as performance deltas begin to compound over time, particularly given how long many ML workflows typically take to complete.

Overall, the Micron 5210 ION is a high capacity, read-intensive SSD that can add significant value to data pre-processing to speed up machine learning workflows at an approachable price.