Improving Test Procedures for Consumer SSDs

Andrew Mierau – Micron Technology

and

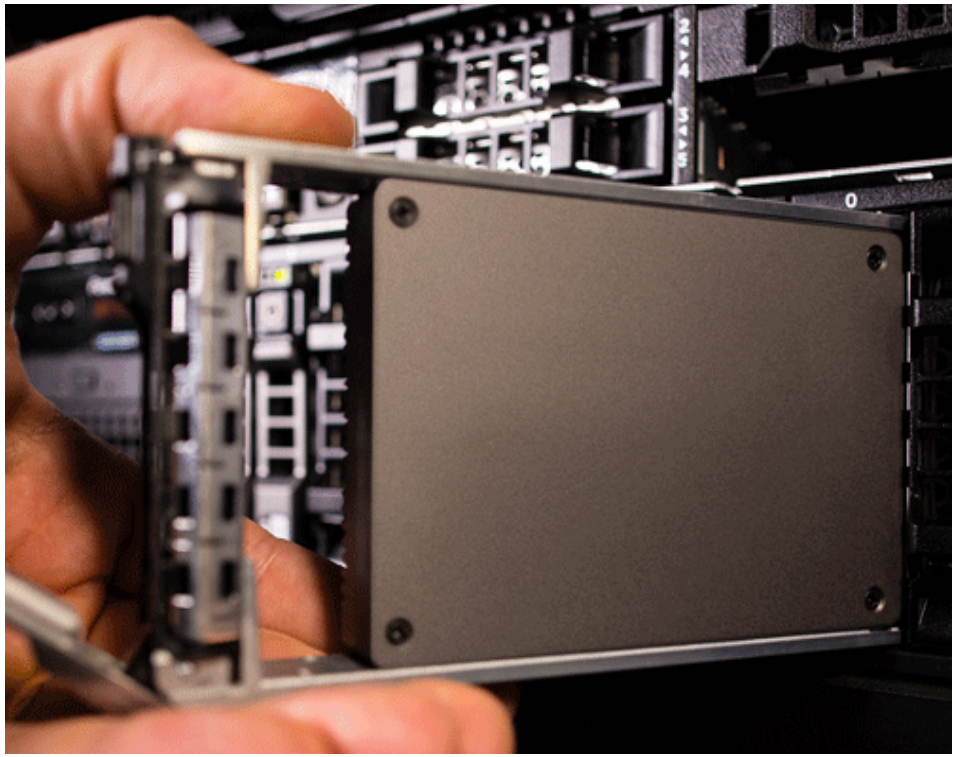
Allyn Malventano – PC Perspective / Shrout Research

The Problem with Benchmarks

- Many of the common benchmarks and test scripts do not accurately represent the common consumer use cases:
  - Test Consumer SSDs like Enterprise SSDs
  - Heavily rely on averages
  - “Real World” benchmarks that are nothing like the real world

Enterprise Testing

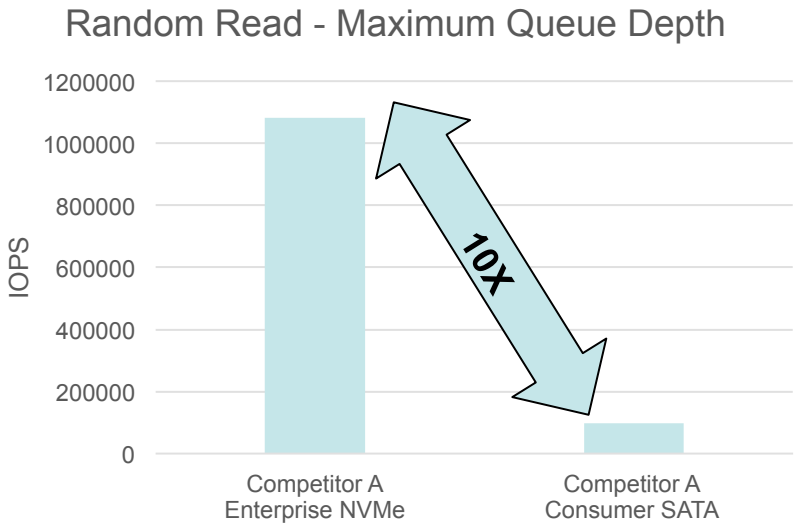
- Most SSD benchmarks were designed around enterprise SSD specs/workloads
  - 4K Random Writes
  - Server/Database Profiles
  - Sustained/Stead-State

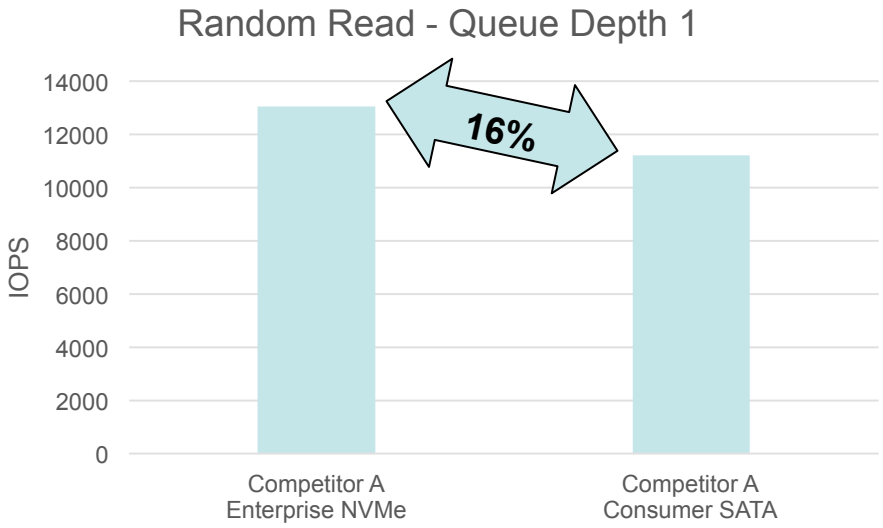


Consumer Workloads

- Consumer workloads are less strenuous than most people think
  - Mostly Idle/Wait, ”Bursty”
  - Mostly Read
  - Mostly Low Queue Depth
- Modern caching SSDs take advantage of this workload

Does that really make a difference?





YES!

Benchmark Types

- Synthetic
  - IOMeter/FIO
  - CrystalDM, ATTO, HD Tune, etc
- Application
  - SYSmark
- Trace
  - PCMark

Allyn Malventano

- United States Navy (Retired)
  - Submarines, Cryptography, RE
- PC Perspective
  - Storage Editor
- Shrout Research
  - Storage Analysis Lead

SHROUT
RESEARCH

Objectives

- Workload Application
  - Preconditioning consumer SSDs
  - Zones, burst and paced workloads
- Data Acquisition
  - Enterprise methods evolved into client testing
- Putting it all together
  - Result examples

Preconditioning

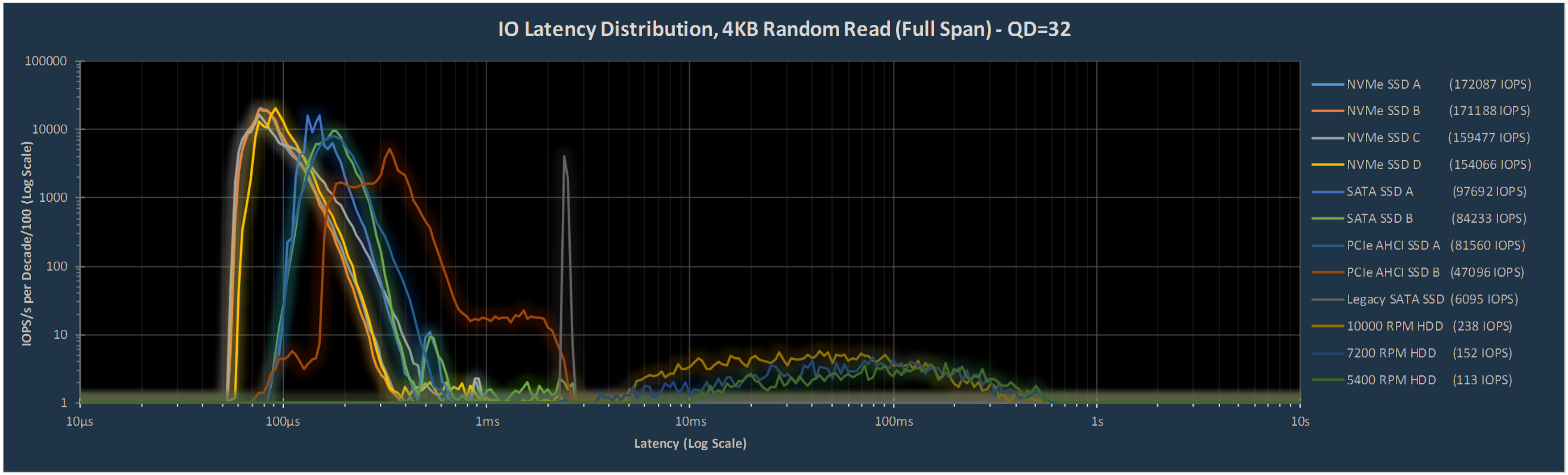
- Closely approximate a client, not a server
  - Random to a portion of the SSD, else sequential
    - Do not leave SSD mostly empty! (FOB)
  - Random and sequential areas must remain so
  - Each zone must be at steady state prior to test run
    - Complicated by SLC caching SSDs
    - Complicated by some FTL / Flash management types

Workload Application

- Closely approximate a client, not a server
  - Ideally:
    - Short intermittent bursts
      - Emulates true client usage
      - Delay time sufficient for cache flush
    - Some bursts should be paced (IOPS limited)
      - CPU bound applications

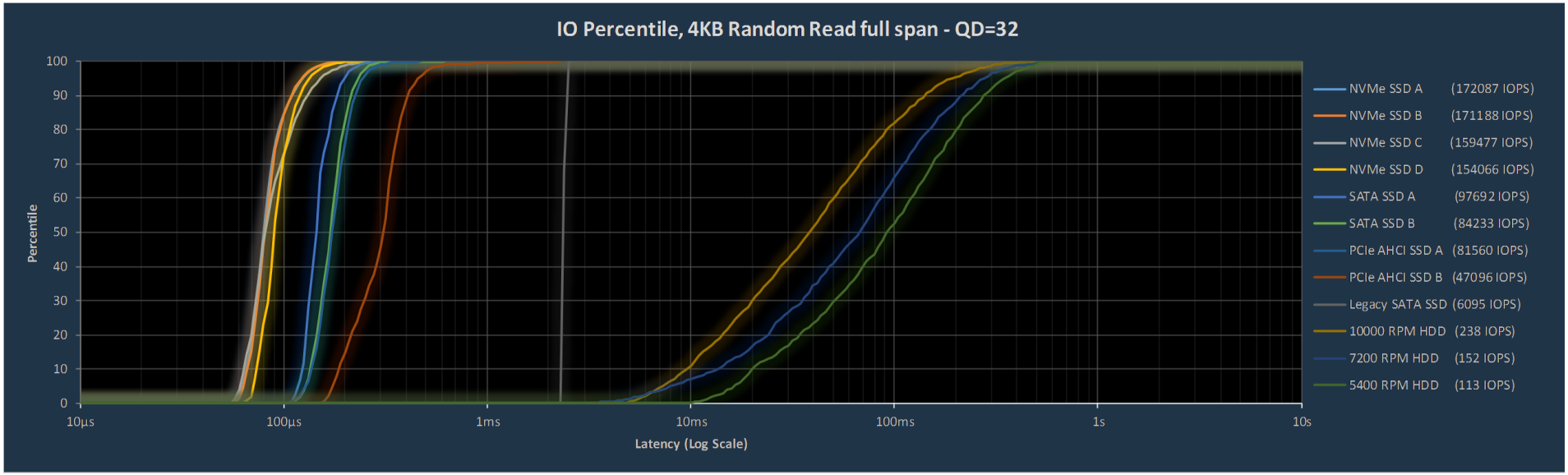
Data Acquisition

- Log Every IO to catch outliers

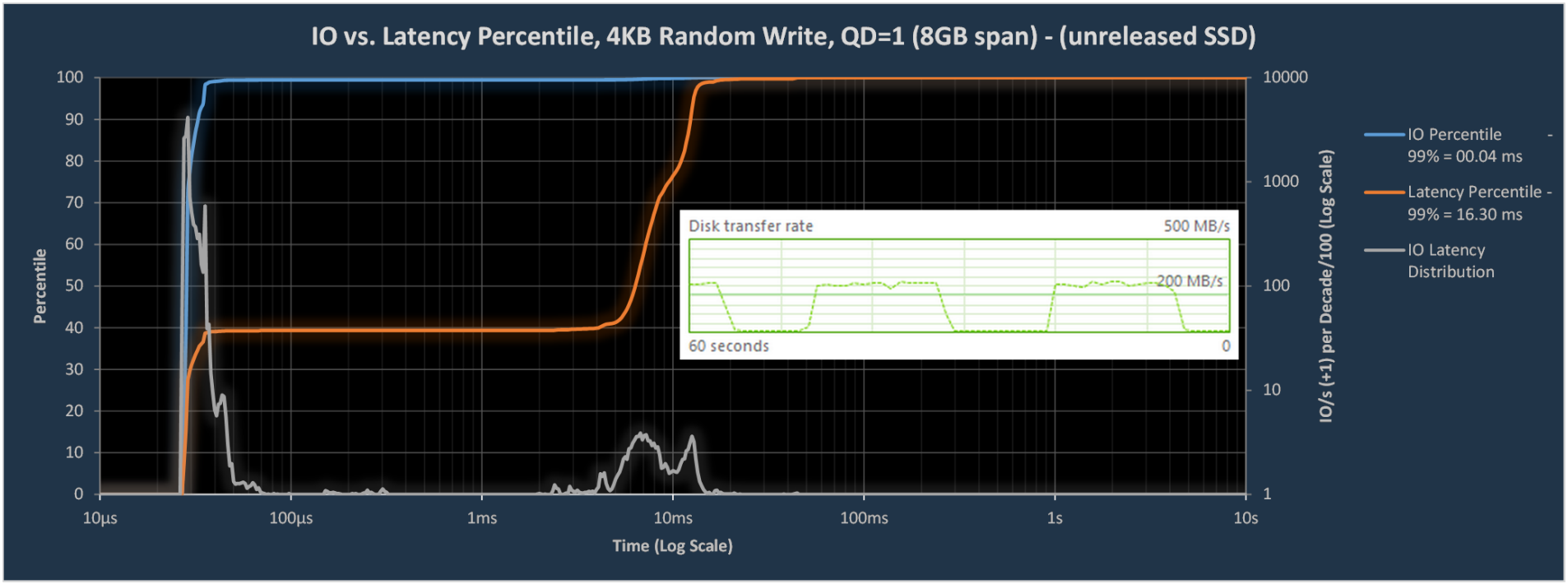


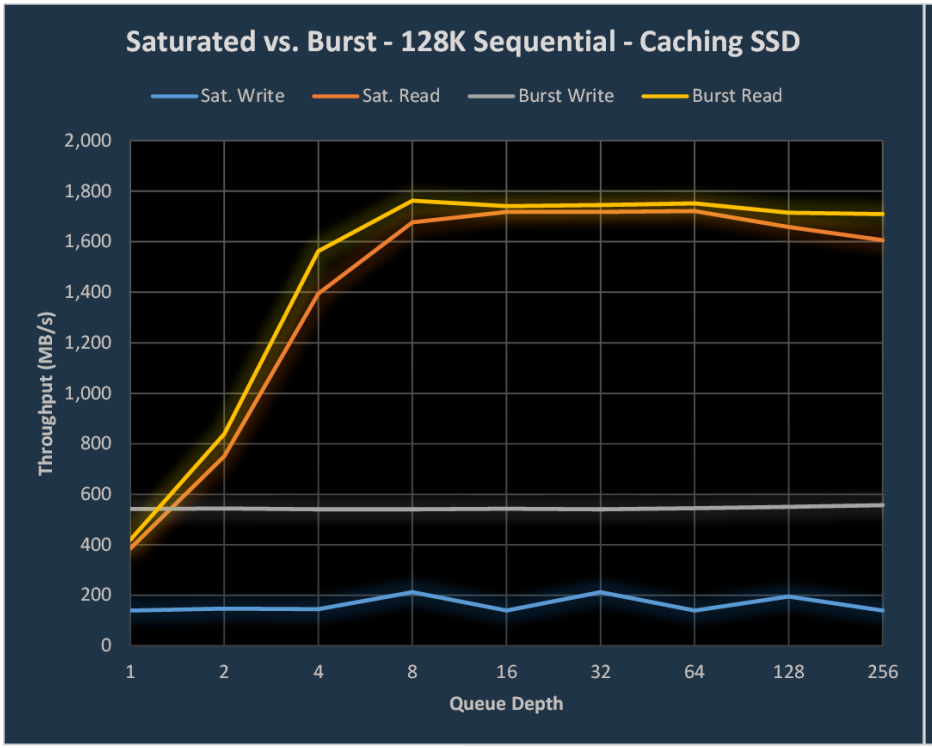
Data Acquisition

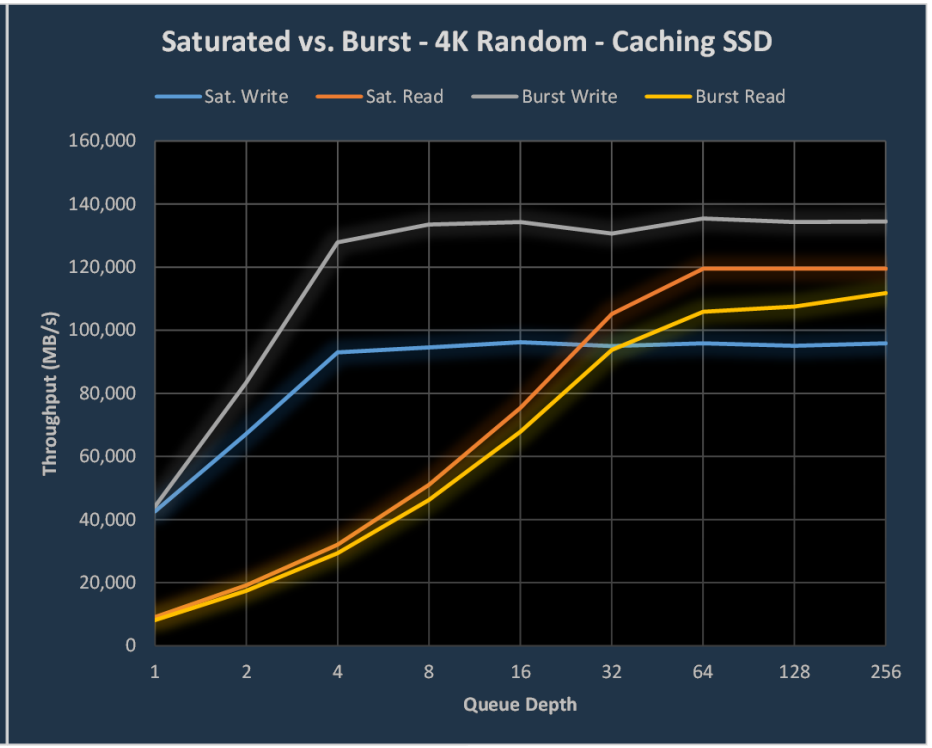
- Enterprise Method – IO Percentile



IO Percentile Failures (and a solution)

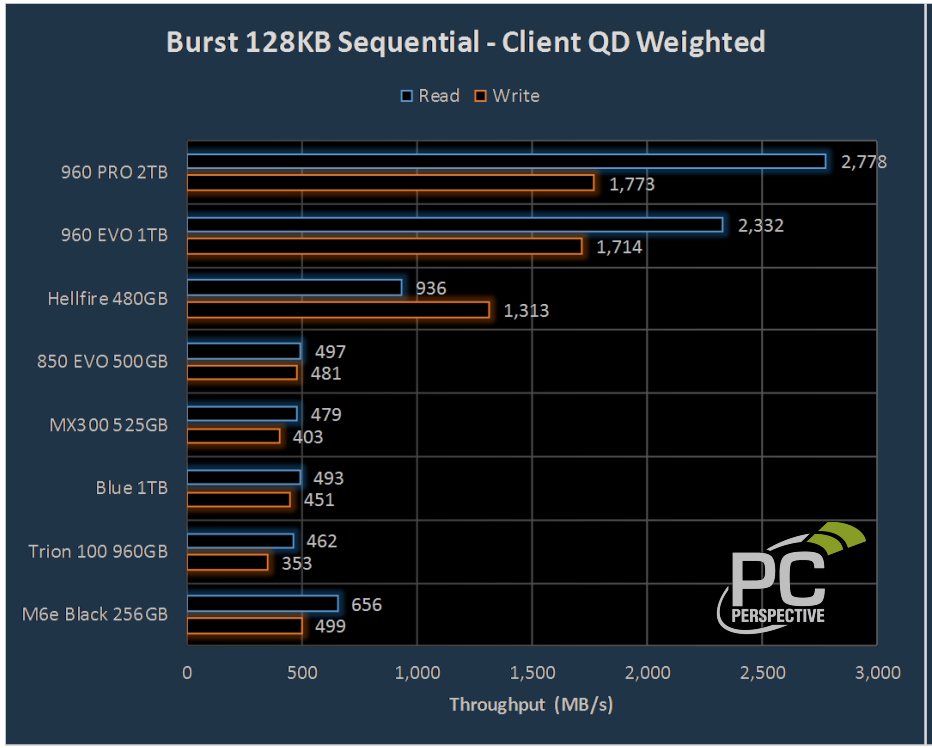


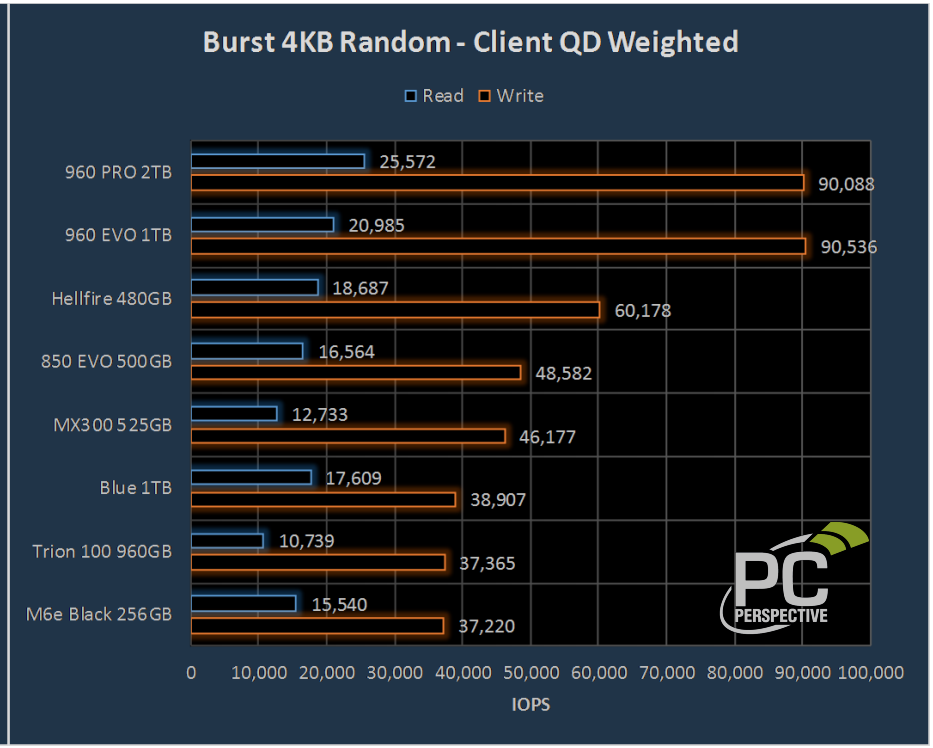




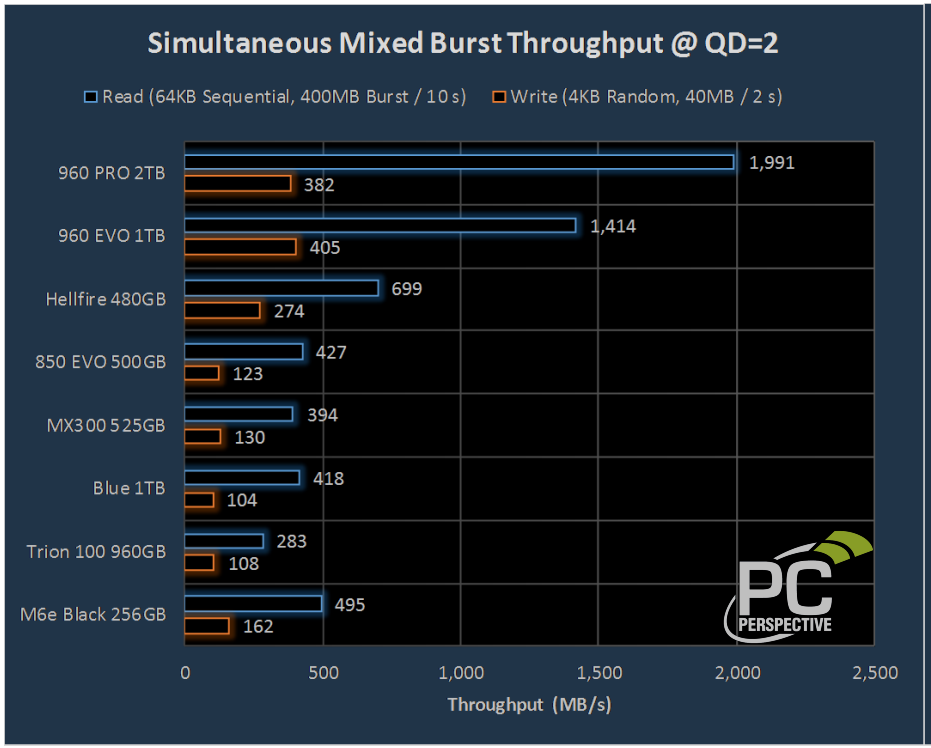
Example: Burst Workload

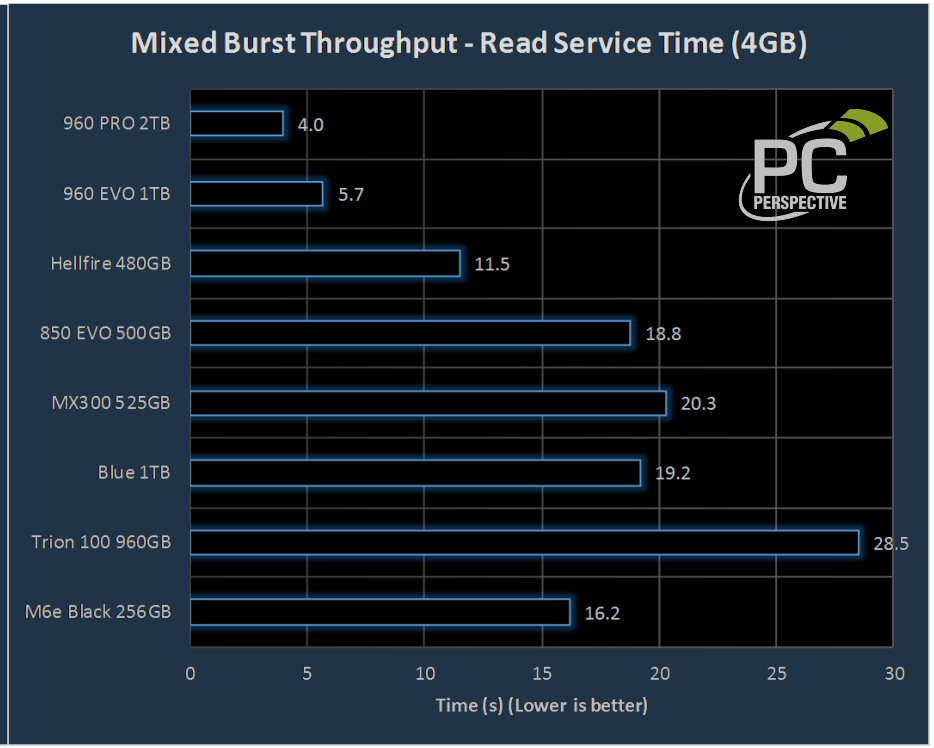
Example: Weighting



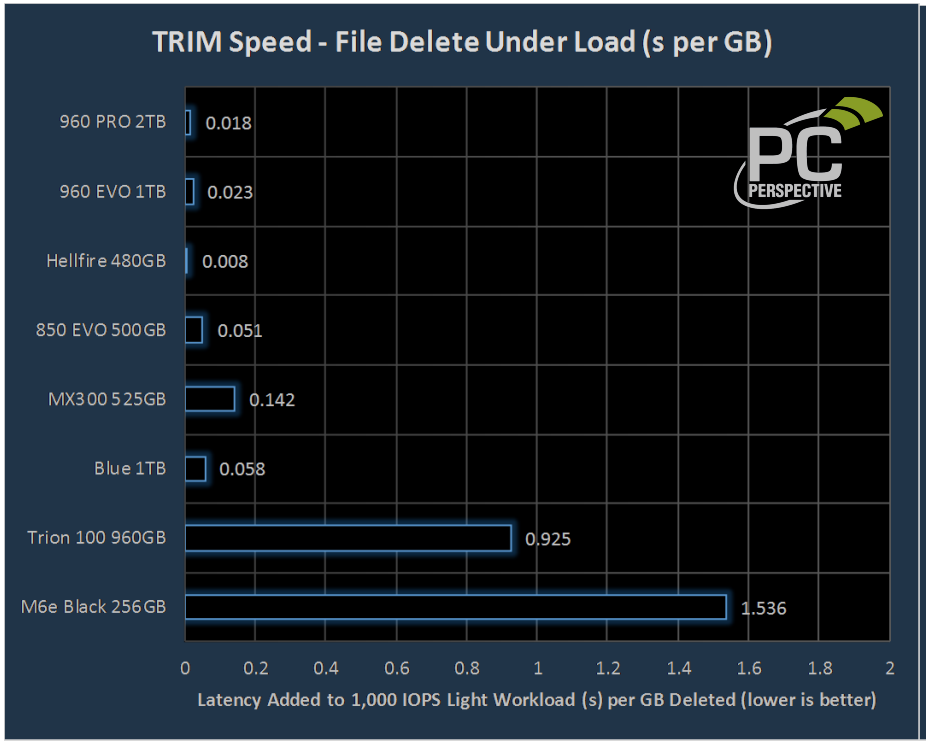


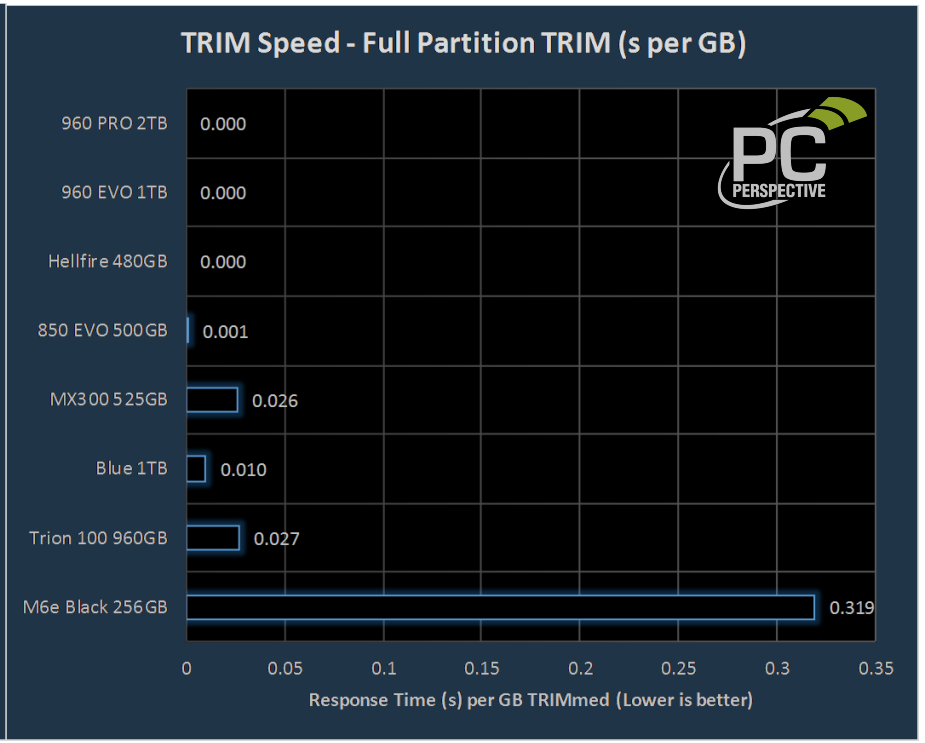
Example: Mixed Burst





Advanced Techniques: TRIM





Summary

- Workloads matter
  - Consumer workloads are inherently different than enterprise
  - Read-centric, ‘bursty’ and low queue depths
- Look beyond averages
  - Averages can hide bad behavior
  - Inspect data more granularly, looking at distributions over time
- Waiting in important
  - Modern consumer SSDs are designed around idle time
  - Condensing tests by removing idle can severely impact performance and is not representative of an actual consumer workload