F2A: Measuring and Understanding True SSD Performance and Consistency

Kent Smith

Sr. Dir. Product Marketing

Agenda

- Benchmarking objectives
- Preconditioning review
- Benchmarking 100% reads
- TRIM and Over Provisioning
- Secure Erase Between Tests

Benchmarking Objectives

- Simulate real-world conditions to accurately predict real product performance
- Compare real performance of various products
- Make purchasing decisions based on results

Preconditioning Review from FMS 2009

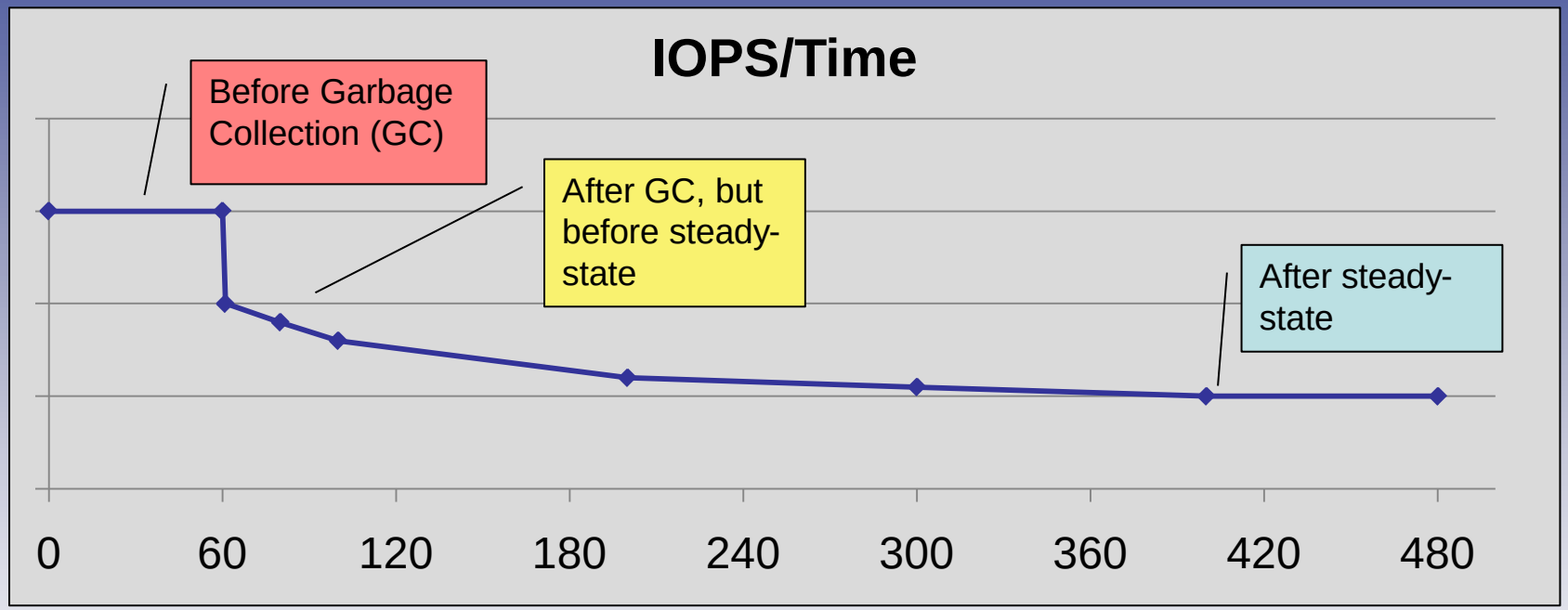
- SSDs and HDDs are affected by the way data comes in (sequential vs. random, queue depth, transfer size, AHCI drivers, etc.)
- Only SSDs are affected by what was previously written to the drive
- Preconditioning assures predictability and repeatability

- http://www.flashmemorysummit.com/English/Collaterals/Proceedings/2009/20090811_F2A_Smith.pdf

Preconditioning Review

- All Flash-based SSDs need to Garbage Collect (GC) after each block has been written once
- Preconditioning just writes data to the SSD to get it into a GC state after Secure Erase
- The best data to write is the data desired to be tested
- Write the desired data to the SSD until performance no longer drops
  - From 2 to 48 hrs depending on SSD and test

Preconditioning Review – When to Sample

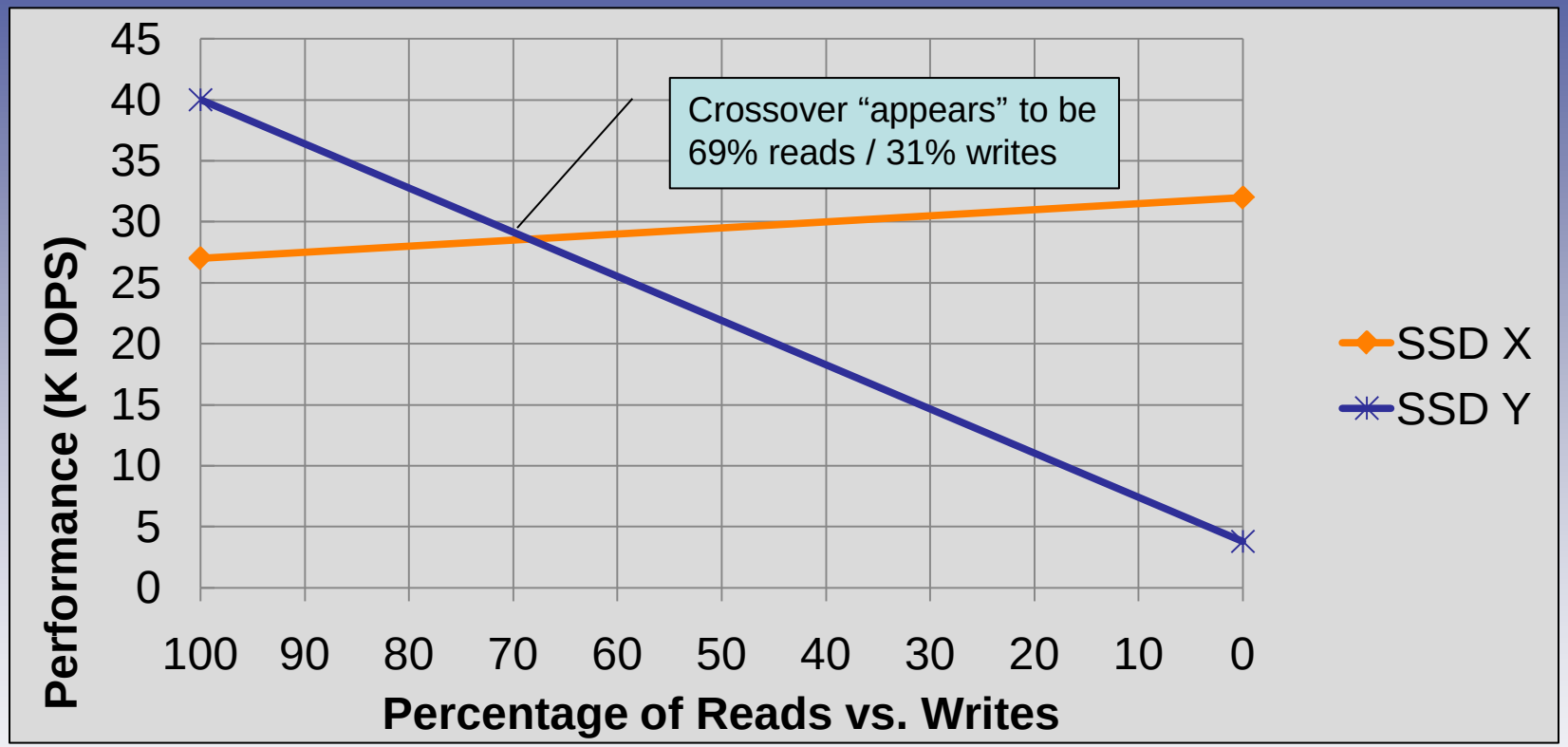


- GC is present during no less than 99.98% of drive life
- Steady-state comes after drive performance levels out
- Measurements before steady-state are misleading
  - Up to 10x higher, but not sustainable

Benchmarking 100% Reads

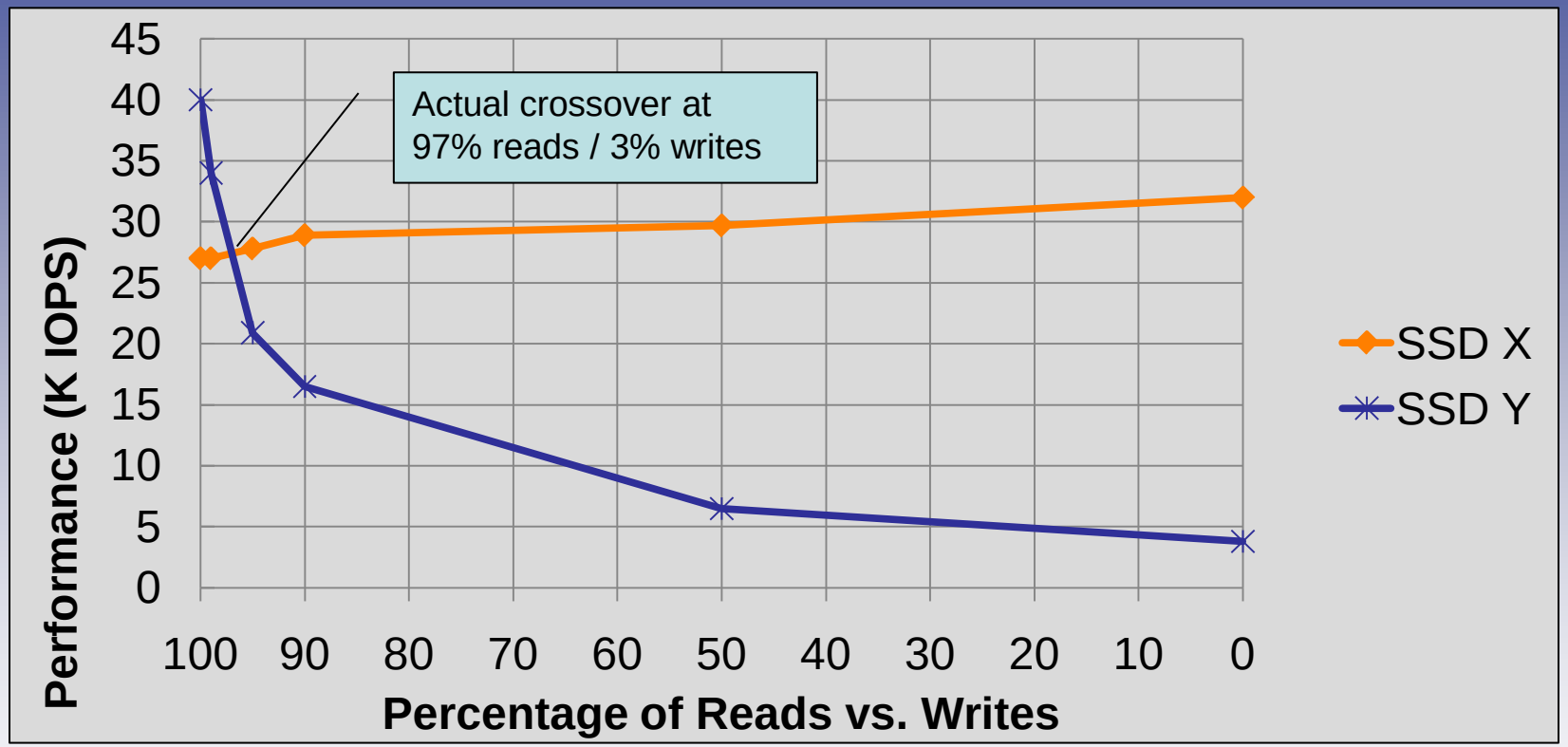
- Many benchmarkers focus on testing 100% reads to see what an SSD can do
- This is unrealistic in a writable drive
  - At some point data will need to be written or updated in any usage environment
- It would be more appropriate to measure the average expected ratio
  - Example R/W ratios: 60/40, 70/30, 90/10
  - Even a cashing appliance would be no better than 19:1 R/W ratio (95/5)

Comparing 100% Reads or Writes



- With only two data points, it looks like SSD Y is better at reads even down to 69% reads / 31% writes

Now Drill Deeper…
Test Various Mix Ratios

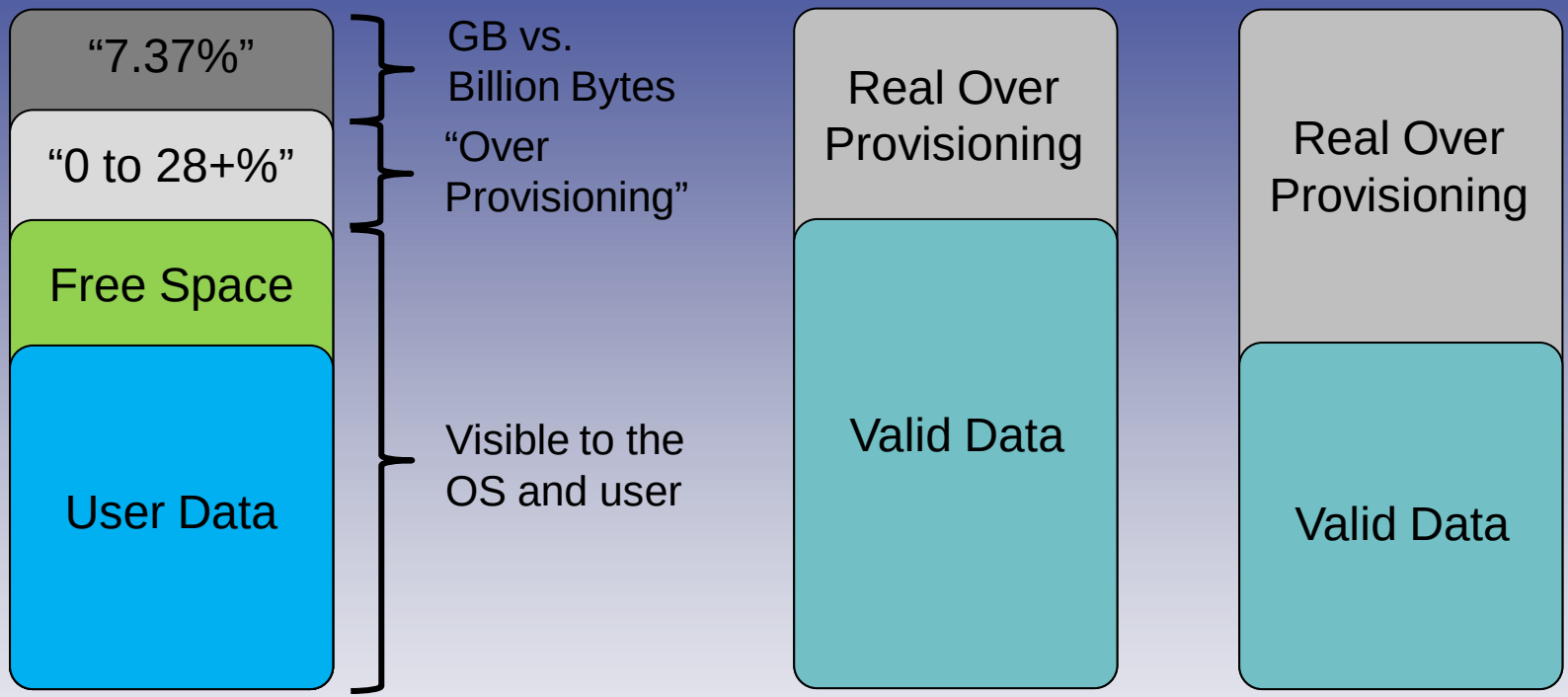


- Low write performance greatly impacts read performance even at small mix ratios

Logical Block Address (LBA) vs. Physical Address (PA)

Operation	OS (or HDD) View	SSD View
Write a new file	New valid data to new LBA	New valid data to new PA
Overwrite file	Replace old data with new valid data in same LBA	Marks original data in original PA as invalid, and writes new data to new PA
Move file to trash	Valid data in same LBA	Valid data in same PA
Empty trash (no TRIM)	Free space in same LBA	Valid data in same PA
Empty trash (w/TRIM)	Free space in same LBA	Invalid data in same PA (can GC)
Garbage Collection (GC)	Invisible to OS	Converts invalid data in PA to free space in same PA

TRIM and Over Provisioning (OP)



Higher OP =

- Lower Write Amplification
- Higher performance
- Longer Flash Life

Benchmarking and TRIM

- User performance is based on capacity used
  - More free space = More OP (if TRIM is supported)
- Many benchmarks will fill the drive to capacity
  - This is a worst case test; minimal OP
  - Most users do not run at 100% drive capacity
- Any benchmark can simulate a lower percentage consumed user capacity
  - Create a smaller partition before starting the test
  - The extra space is automatically used by the SSD as OP

Secure Erase Between Tests

- Prior data written to SSD changes future performance
- Secure Erase between benchmarking tests:
  - To eliminate unknown variables in test
  - To quickly return SSD to a known state for next test
  - To test the pre-GC performance if desired
  - Is great for benchmarking since data is not “valuable”
- Secure Erase periodically in real life to “restore” SSD performance is:
  - Very burdensome to backup all data, erase, and restore
  - Unnecessary on a true enterprise grade SSD

Summary

- Remember why the benchmark is being run – to simulate the real world as best as possible
- Preconditioning is important for SSDs for many reasons
- Look at the data points just below 100% reads to see the true nature of mixed workloads
- TRIM and Over Provisioning are key to performance and low write amplification
- Secure Erase between tests to isolate variables

SandForce SSD Processors

World-class reliability, performance, & power efficiency for enterprise, client, and industrial SSD applications

- Visit us at **booth #407** to see our DuraClass™ technology in action with the latest 30nm MLC flash technology
  - MO297 – Ultra compact high-density industrial standard SATA form factor with up to 30K read/write IOPS* per watt
  - PCIe – Industry leading IOPS* for read, write, and mixed workloads
  - SAS – 25K sustained write IOPS* through Emulex's high-performance SAS-SATA embedded storage bridge technology
  - SATA – Updates on SSDs, testing, and infrastructure equipment collaboration
- Stop by to enter our **free drawing** for one of three Corsair Force 120GB SandForce Driven SSDs
- See other SandForce Driven SSDs in our partners' booths as well

***Random 4K transfers**