SSD Performance Tutorial

SSDs and SSSI Testing

Developing Test Standards for SSD Performance

Eden Kim, CEO
Easen Ho, CTO

Santa Clara, CA USA
August 2009

Outline

- Introduction
- SSSI Performance Test Suite Specification
- Factors Affecting SSD Performance
  - Pre-Conditioning
  - Effect of Spares on Performance (TRIM Simulation)
  - Idle Effects
- Summary

Spotlight is now on….

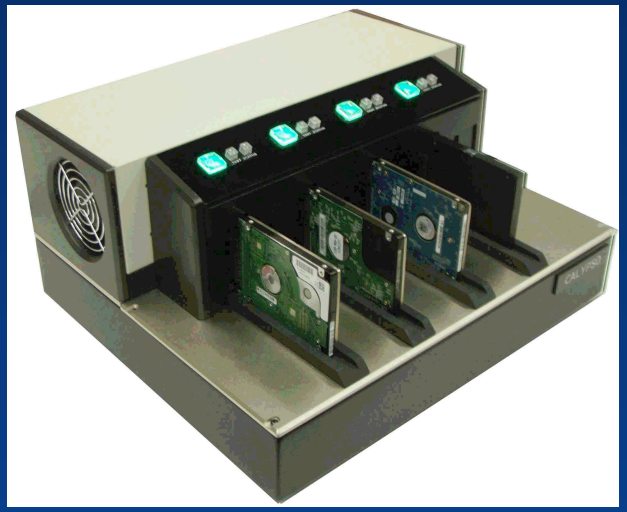
- Innovation in architecture to manage flash performance characteristics
- Matching device characteristics to application needs
- Defining device-level reliability and data-integrity metrics
- Understanding performance issues

Current SSD Landscape

- Lots of variability
  - in specification
  - in actual performance
  - in cost metrics
- NO standard for specification disclosure
- NO standard for performance measurements
- Technology/products keeps changing
  - more controllers; more interface choices
  - new 3X-nm flash brings new issues
- Need for Performance Test Standards!

SSD Blind Surveys for SNIA SSSI

- Common hardware & software environment provides basis for performance comparison
- SSSI Reference Test Platforms – Multiple OS Capability, Standard Test Methodology
- Multiple, Parallel, Asynchronous Test Bays
- SSS TWG Performance Test Metrics
- Both synthetic RAW Device & System Level Composite Tests



Scope of Blind Study

- SSD Blind Survey No. 1 (BS1) – 2.5" SATAII SSDs circa Dec. 2008
  - Conducted by Calypso Systems & SSSI
  - Presented at SV09 & SNIA Symposium Jan 2009
  - Incorporated into SSS TWG Performance Test Suite Draft
- 2nd & 3rd Generation SSDs under BS2 investigation – Aug 2009
- Performance Test Data provided to SNIA Tech Working Group for development of SSD Performance Specifications & Standards
- Work is on-going
- Call to Vendors to submit SSDs for Evaluation / Inclusion in Studies
- Plans for SNIA Tech Center SSD Test Lab

Outline

- Introduction
- SSSI Performance Test Suite Specification
- Factors Affecting SSD Performance
  - Pre-Conditioning
  - Effect of Spares on Performance (TRIM Simulation)
  - Idle Effects
- Summary

Goal of Any Performance Measurement Specification

Repeatable	common starting point; common procedures
Stable	test at steady state
Applicable	results relevant to user's conditions
Comparable	fair device-to-device comparison
Practical	completes with reasonable time and effort
Accessible	open specification; 3rd party validation

SNIA TWG Working to Define Performance Tests

- **Methodology & Metrics**
  - Defines test preparations, test procedures, and reporting requirements
- **IOPS Test**
  - Generates 3D surfaces for IOPS, Bandwidth, and Response Time
  - Data taken at various Block Sizes, R/W Mixes, Queue Depths, and Threads
  - Various Pre-Conditioning Regimes being investigated
  - Includes procedure for TRIM Simulation
- **Idle Test**
  - Looks at effect of Idle on garbage collection & performance recovery
- **Latency Test - at different Queue Depths & Threads**
- **Composite Application Tests (TBD)**

Other Activity Areas

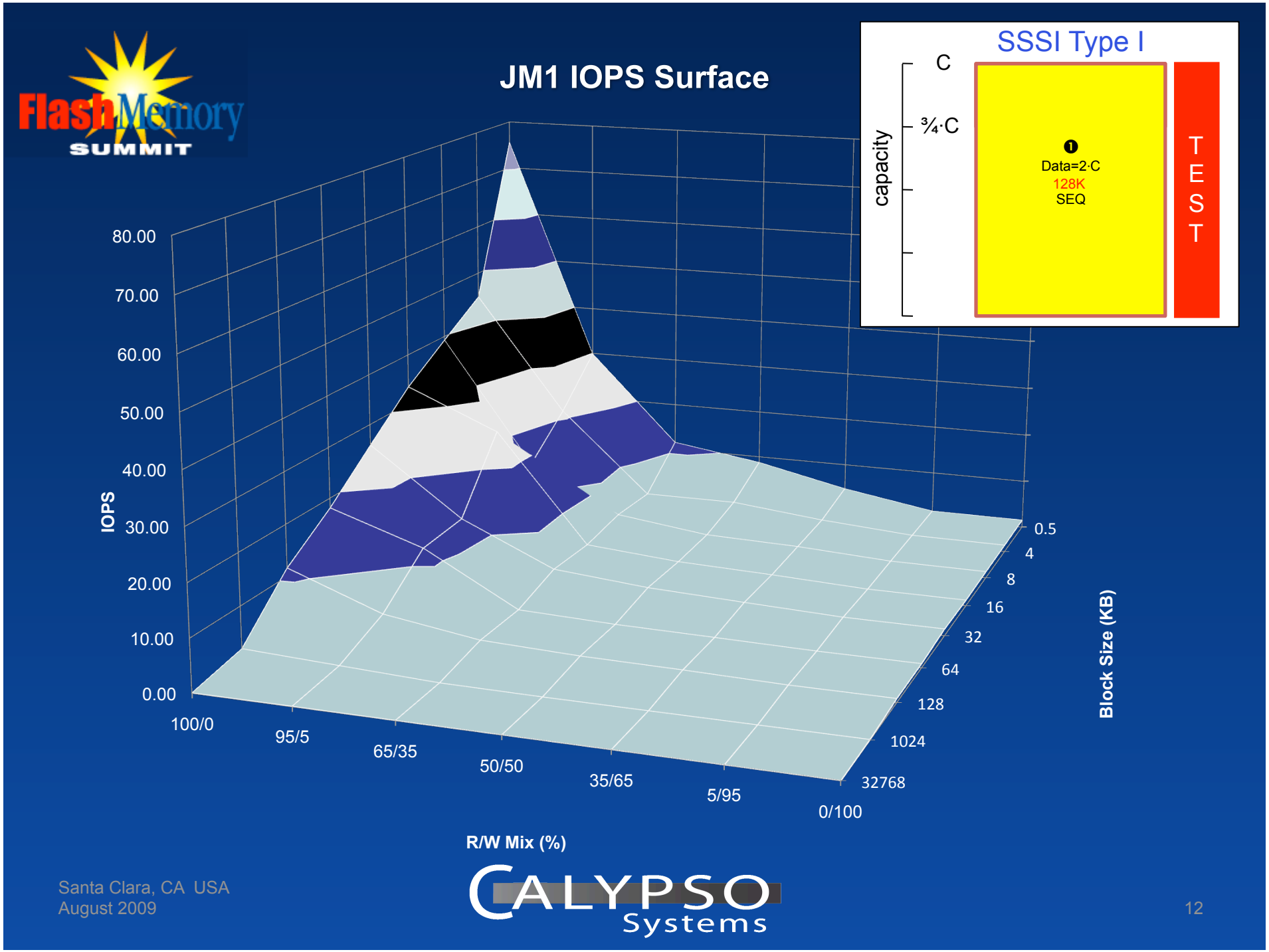
- **"Purge" or SE for SSDs**
- **Workload capture** – What are the "Real World" Uses and performance under those use cases?
- Are SSDs Performance Measurements "Real World?"

Thought on Real world Workloads

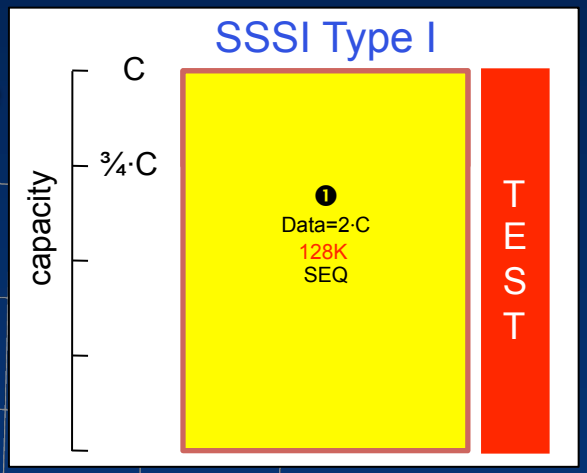
"the Methodology used by the EPA is somewhat abstract"

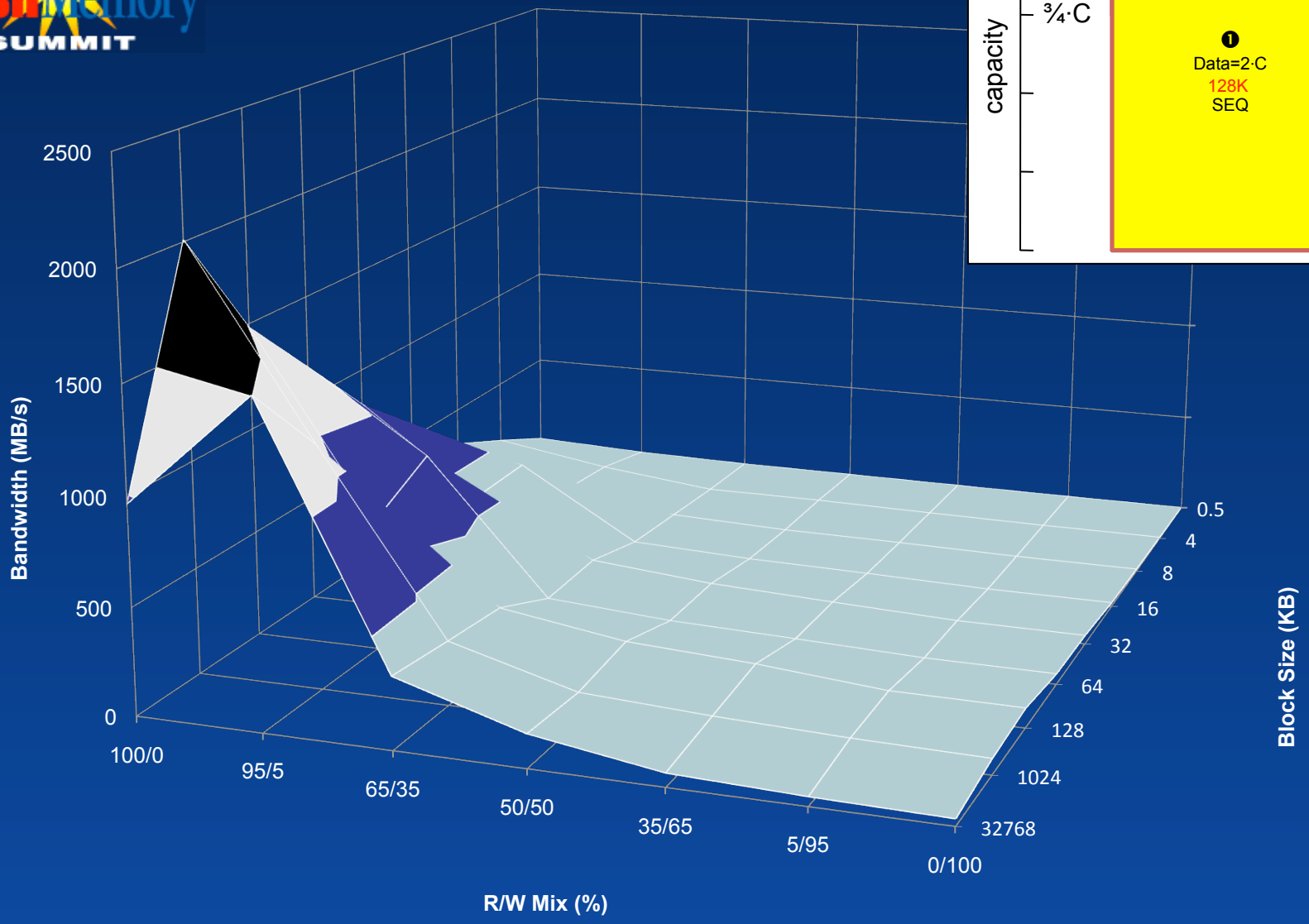
"The high mileage figure, though, is not a meaningful number to many consumers as it doesn't represent real world driving…."

Chevrolet Volt introduction Aug 11, 2009

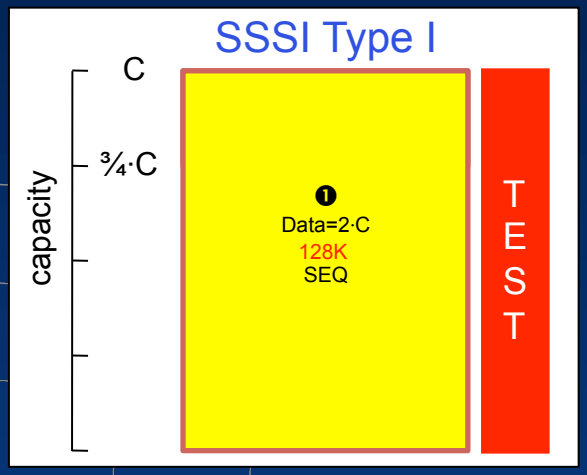

JM1 IOPS Surface
SSSI Type I
C
¾·C
capacity
❶
Data=2·C
128K
SEQ
TEST
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
IOPS
100/0
95/5
65/35
50/50
35/65
5/95
0/100
R/W Mix (%)
0.5
4
8
16
32
64
128
1024
32768
Block Size (KB)

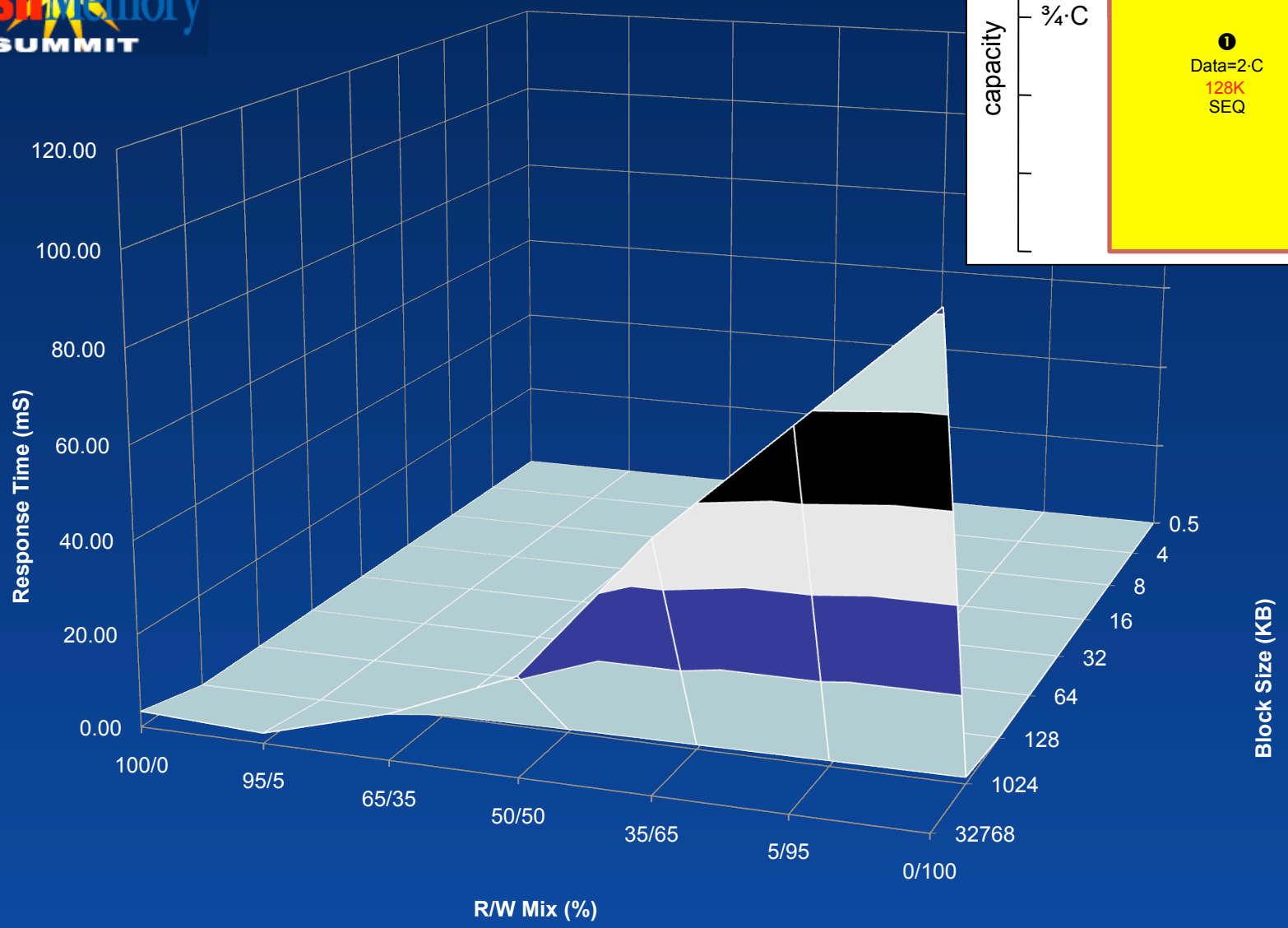
JM1 IOPS Bandwidth Surface





JM1 Response Time Surface





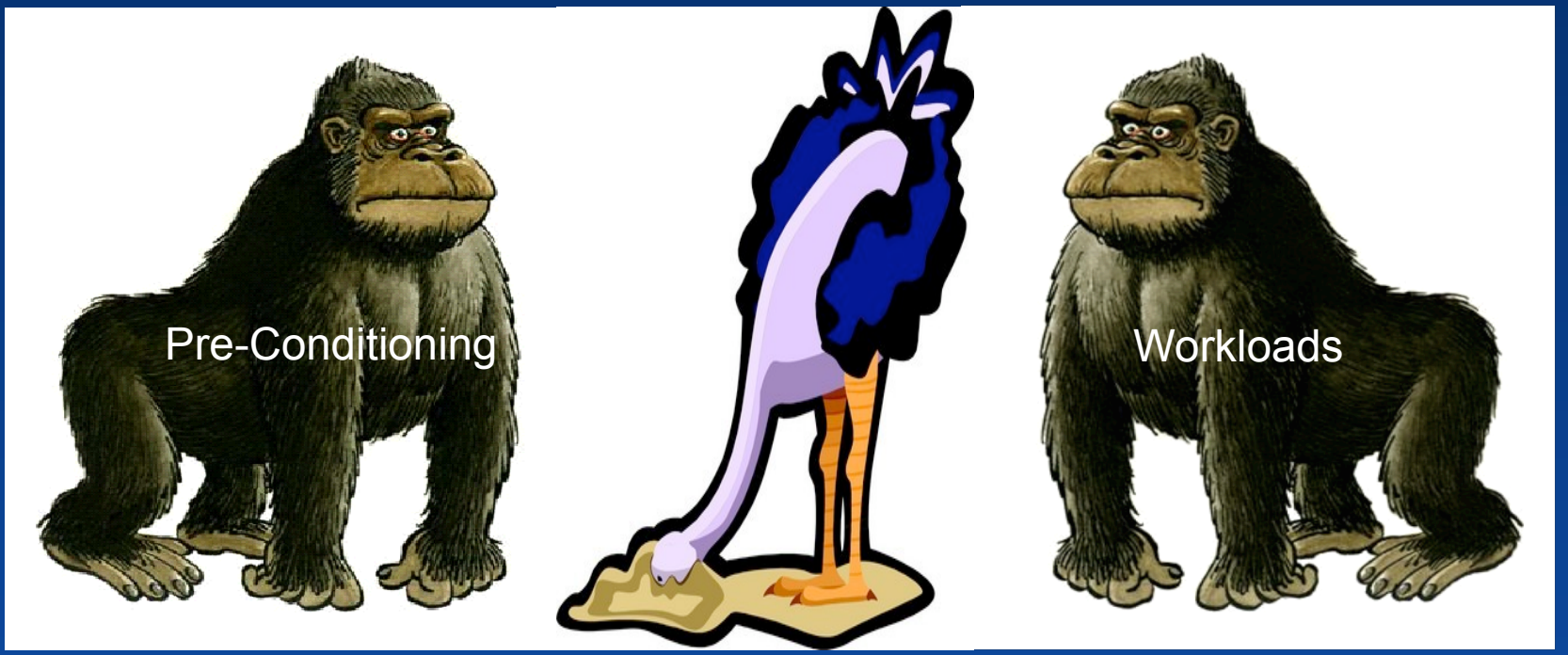
Outline

- Introduction
- SSSI Performance Test Suite Specification
- Factors Affecting SSD Performance
  - Pre-Conditioning
  - Effect of Spares on Performance (TRIM Simulation)
  - Idle Effects
- Summary

Many Factors Impact Performance

- Hardware (CPU, interface, chipset…)
- Software (OS, application, drivers, various caches, SSD-specific “TRIM”, “Purge”…)
- Device (flash generation, parallelism, caching strategy, wear-leveling, garbage collection, warranty strategy…)
- Write History (TGW, spares…)
- Workload (RND, SEQ, read/write mix, queues, threads…)
- Pre-Conditioning (RND, SEQ, amount…)

The 800-lb Gorillas…

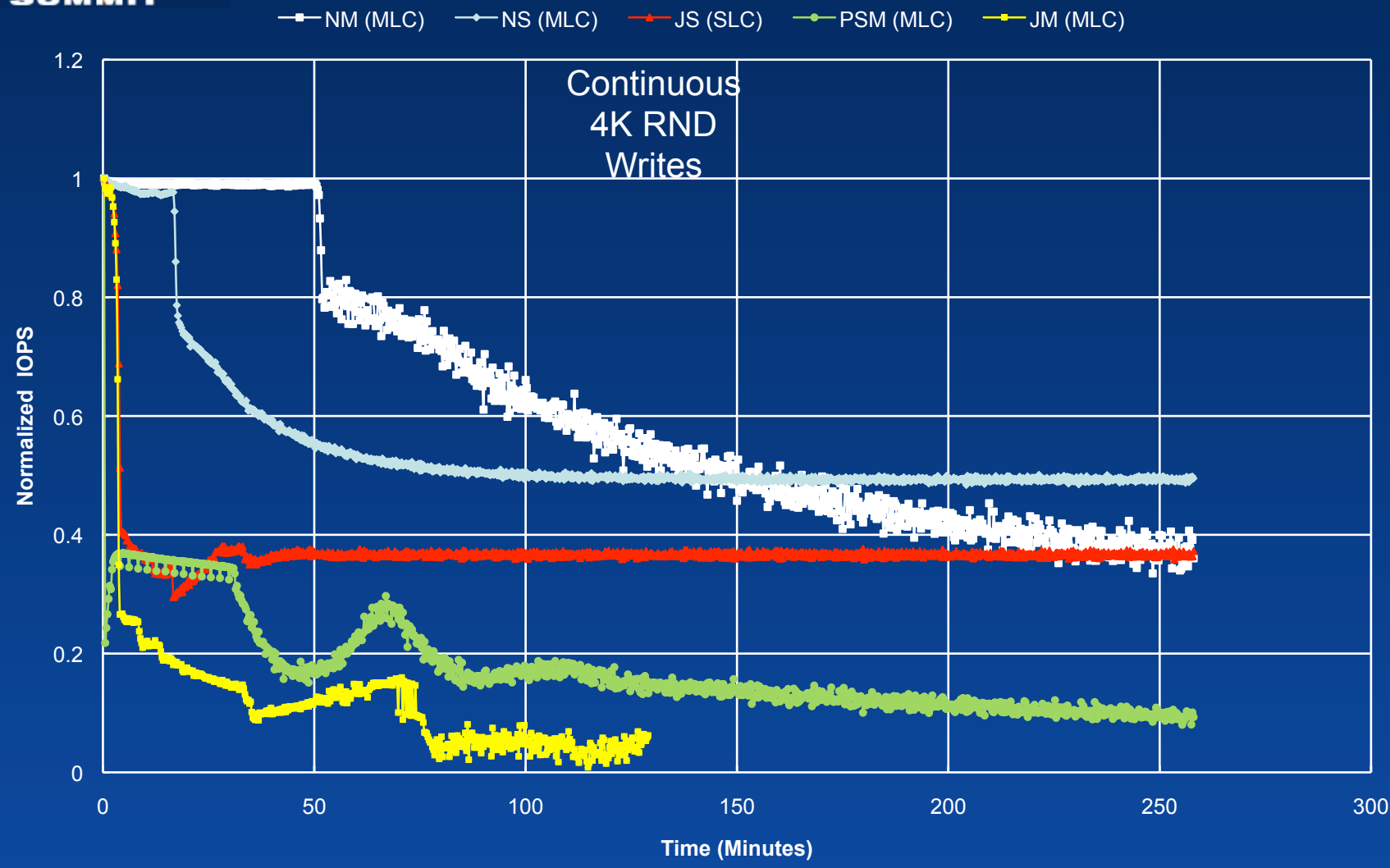


Workloads & Pre-Conditioning are Difficult Issues

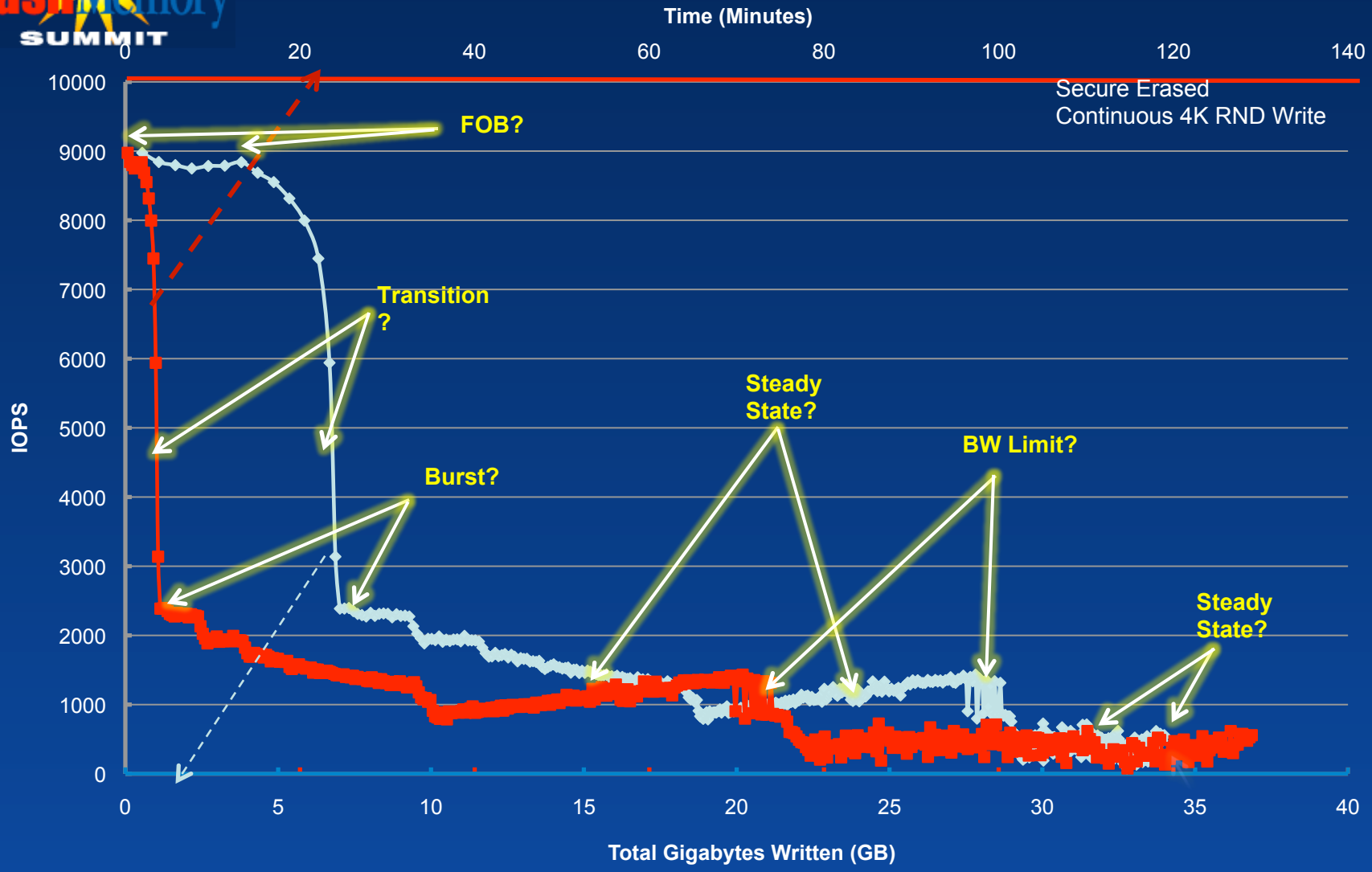
- Workloads
  - "My Workload is better than your Workload"
  - Captured Workloads: hard to get and to use; narrow (but accurate) applicability
  - Synthetic workloads: easy to get and to use; generally useful (while making no one completely happy)
- Pre-Conditioning
  - What kind?
  - How much?
  - Does it need to match Workloads?

Pre-Conditioning

Why Pre-Condition???



Performance States Can Be Complex….



Need for Pre-Conditioning

- Previous slide points out:
  - Need for common terminology
  - Short “Burst” performance when FOB
    - FOB state not important unless drive can return to FOB-like performance somehow
  - Performance can change dramatically with time
    - Can have many transition phases
    - performance comparison valid only under same conditions

Pre-Conditioning Necessary To Get Drive Into Known State

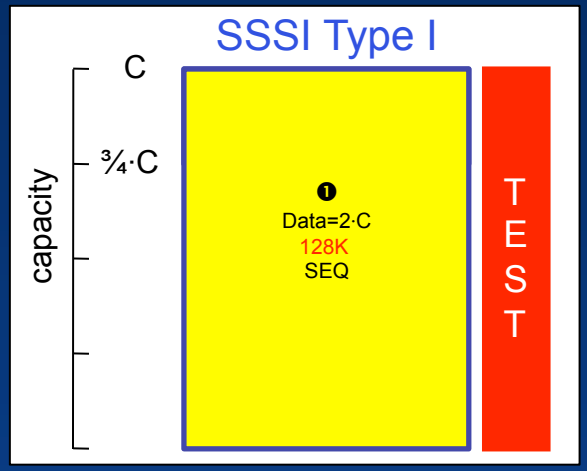
Dimensions of Pre-Conditioning

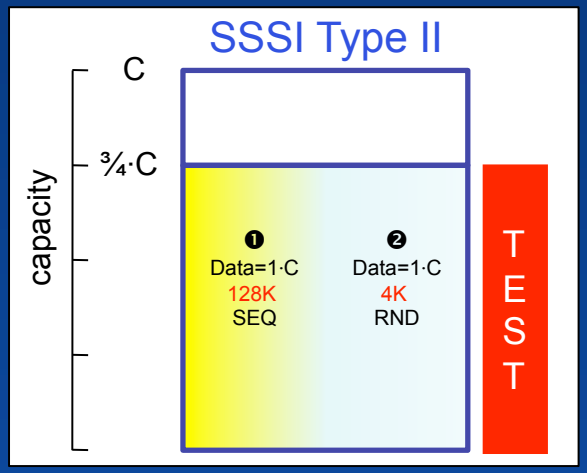
- Access Pattern
  - RND or SEQ
  - Workload-based
- Access Range
  - Full LBA or limited LBA
- Data Amount
- Data Pattern
  - Uniform (0,1)
  - RND
  - Workload-based

SSSI TWG Looking At Various Pre-Conditionings

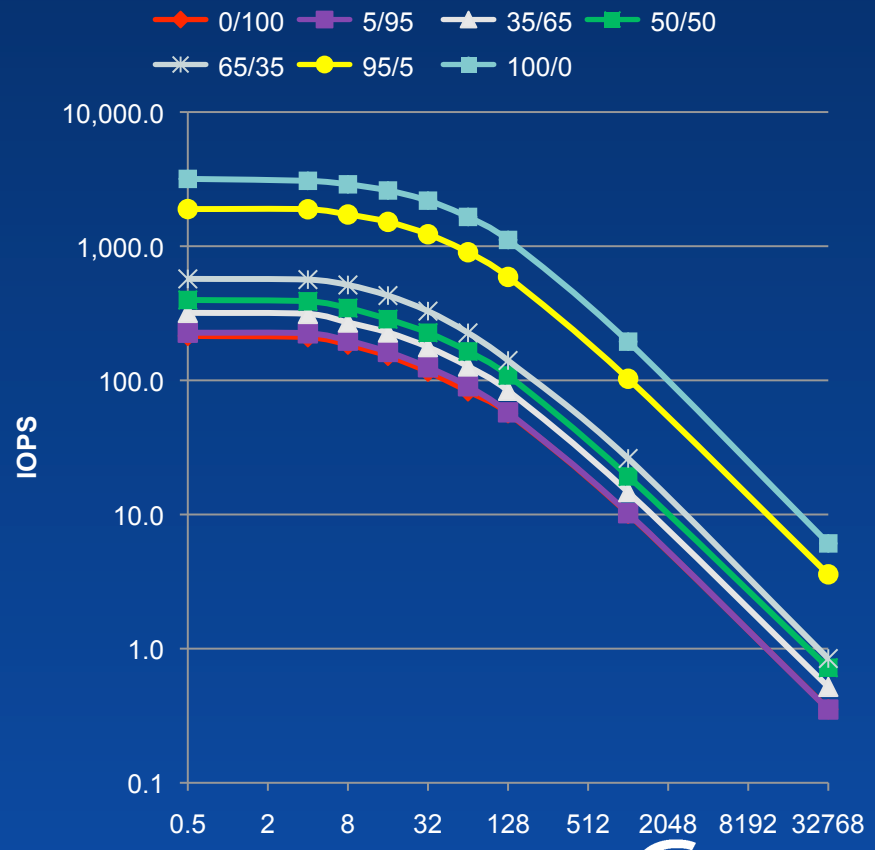
- Type I Pre-Conditioning (100% Full Fill)
  - Writes starting LBA=0, using entire LBA range available to user
    - e.g. Write 128K sequential, total amount written equaling 2X User Capacity

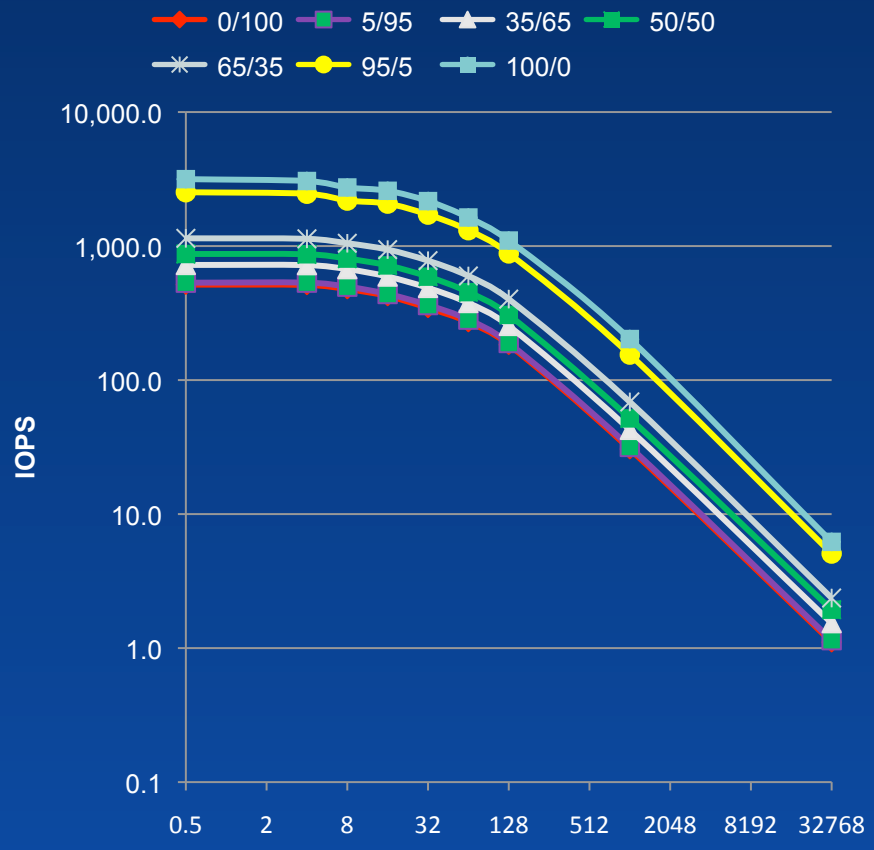
- Type II Pre-Conditioning (75/25 Trim Simulation)
  - Writes restricted to 75% of LBA range available to user
    - 1. Write 128K sequential, total amount written equaling 1X User Capacity
    - 2. Write 4K random, total amount written equaling 1X User Capacity





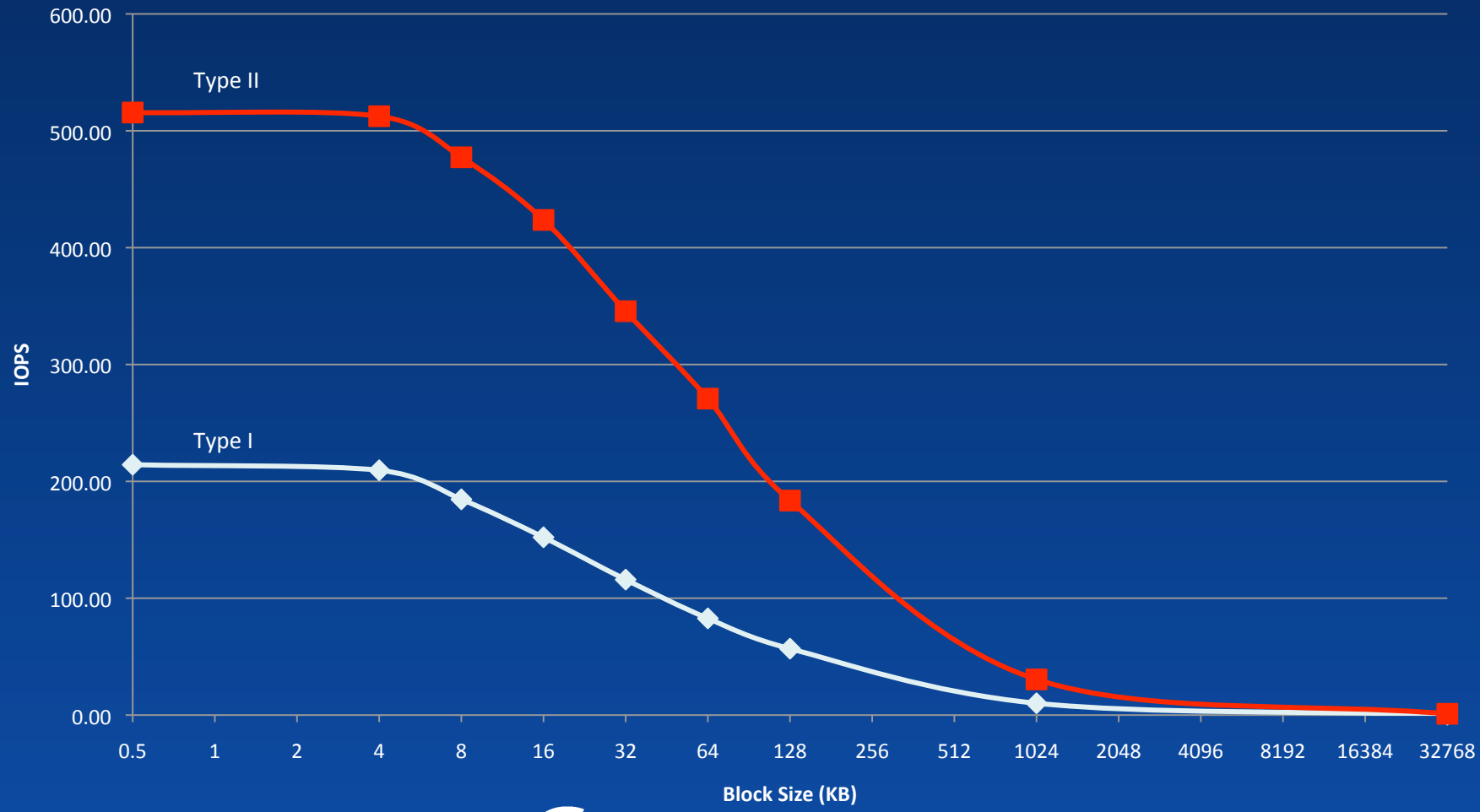
150G-Class MLC Trim/Spares Are Good…



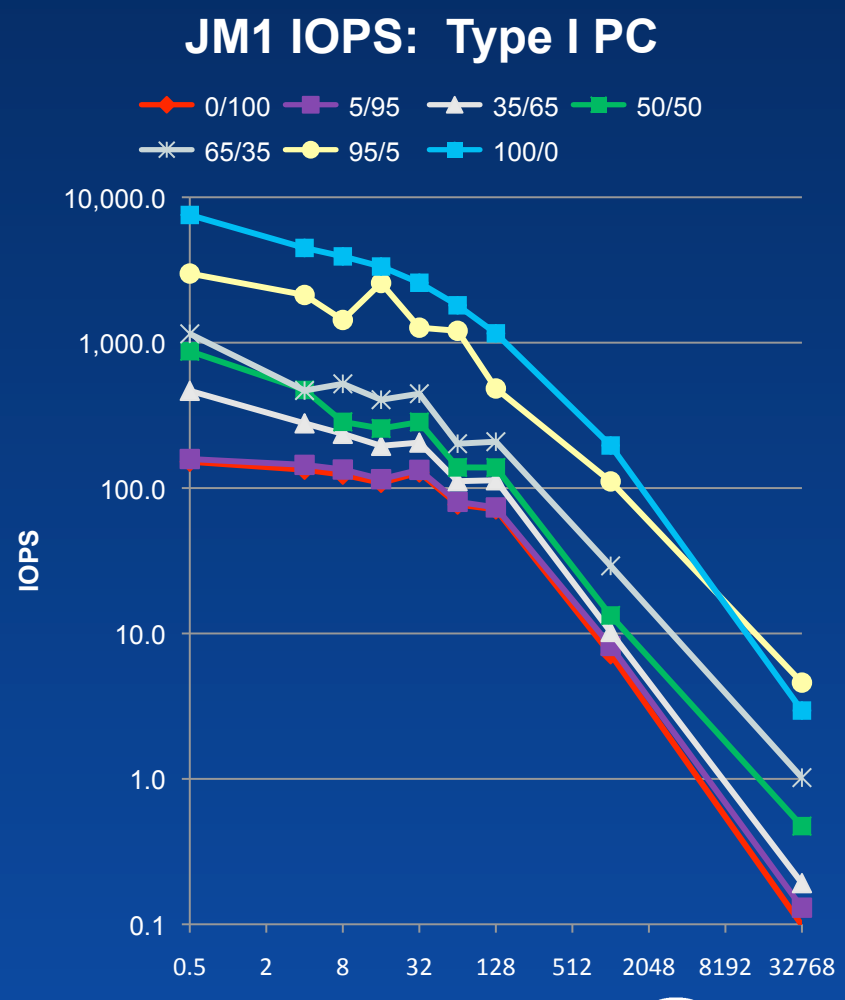


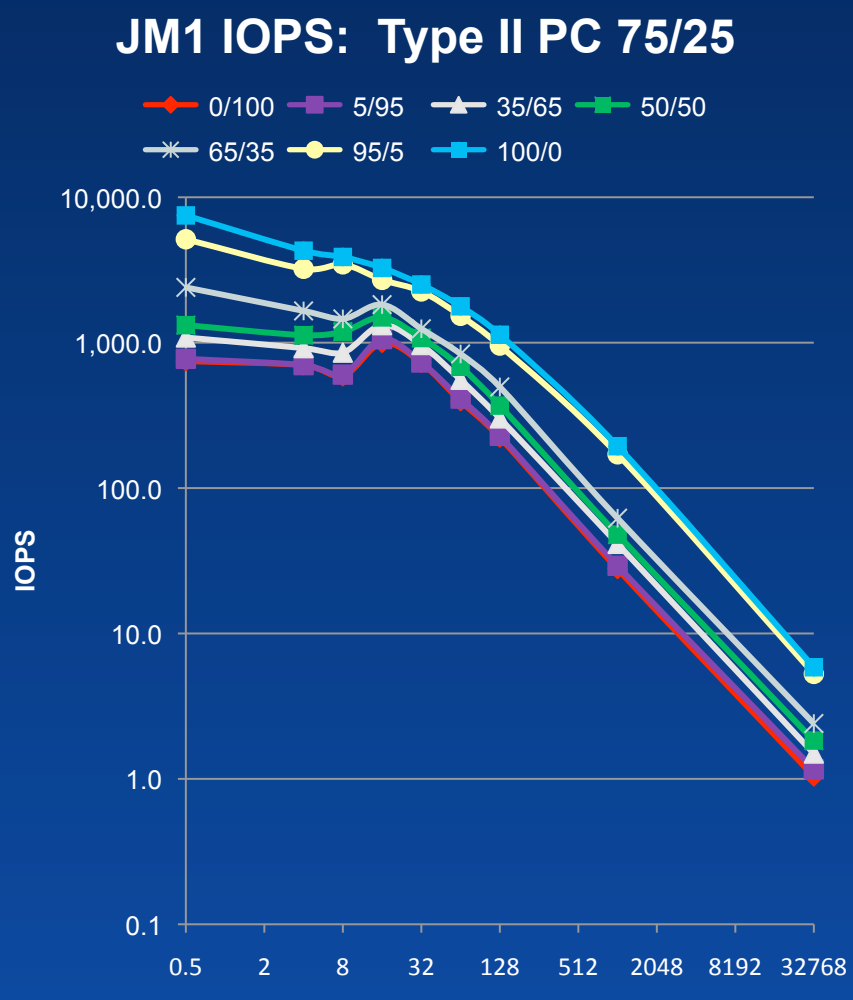
150G-Class MLC
Significant Effect on Writes…

NM12 IOPS: Type I and Type II, 100% RND Write



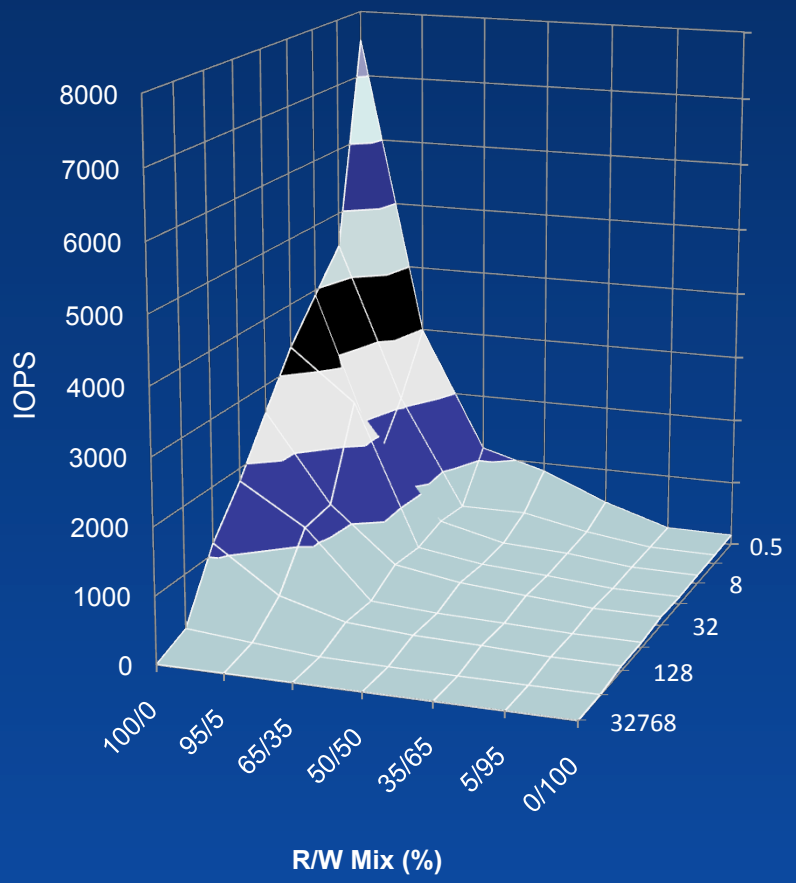
100G-Class MLC
Trim/Spares are good…



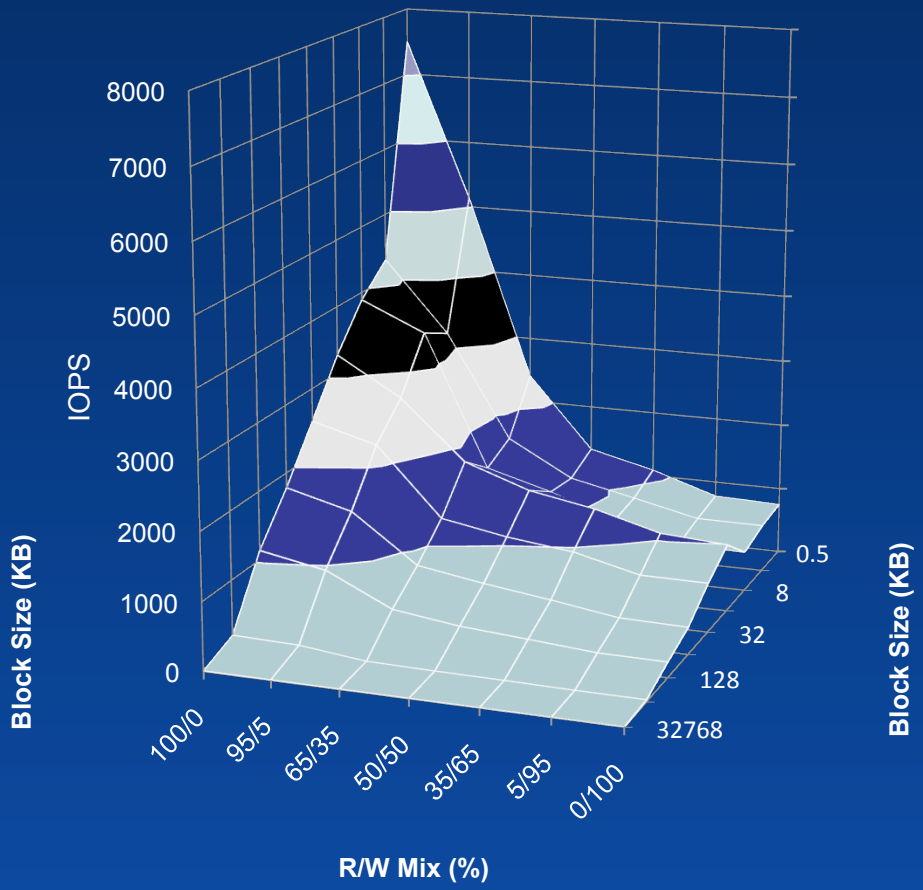


100G-Class MLC
Overall write improvements…

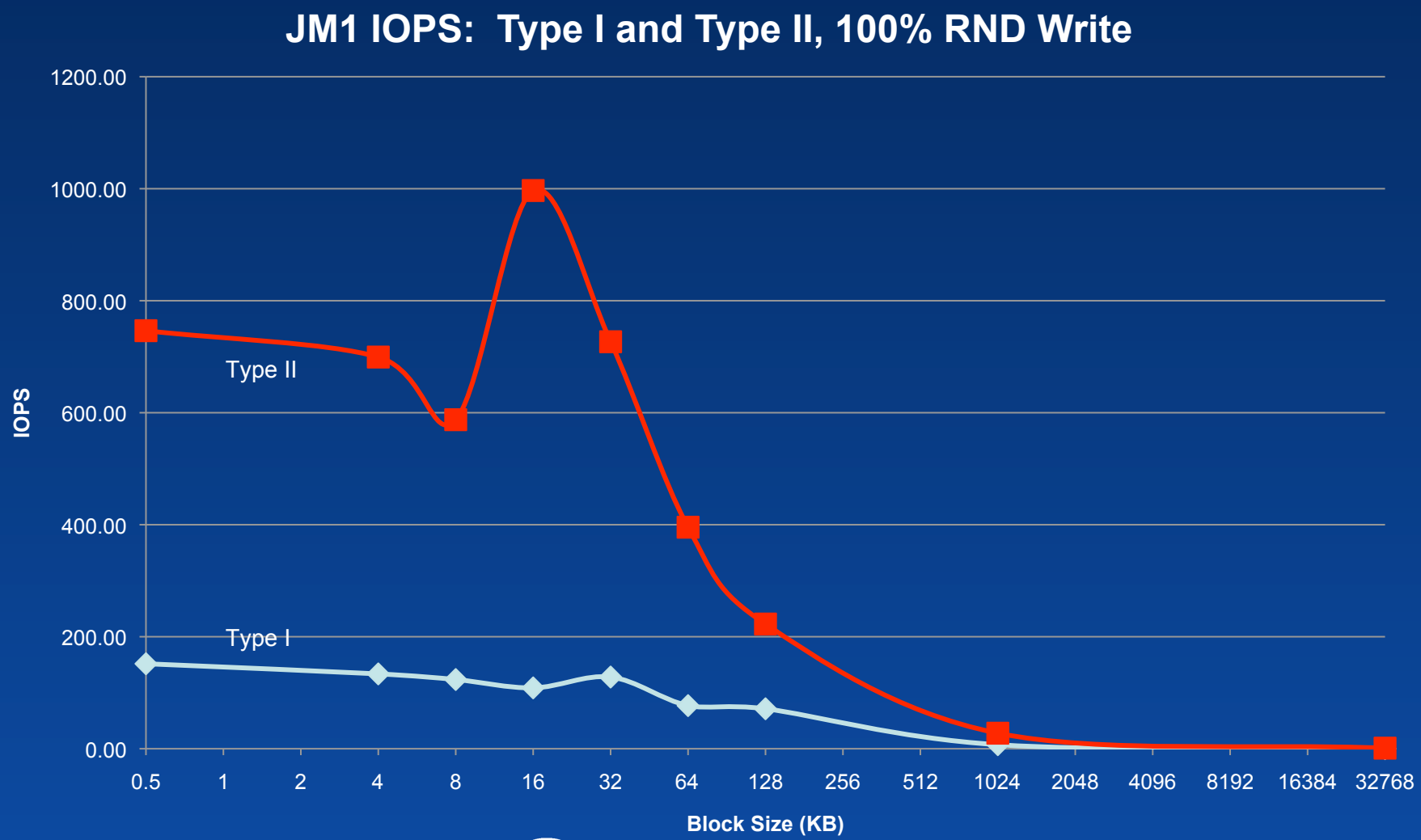
JM1 IOPS: Type I



JM1 IOPS: Type II 75/25

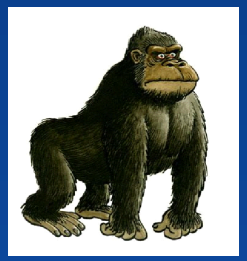


100G-Class MLC



Summary

- Clear need for performance standards
- Need to identify both client and enterprise performance issues
- SSSI is actively working in this area! Stay tuned for more!



Join SSSI!