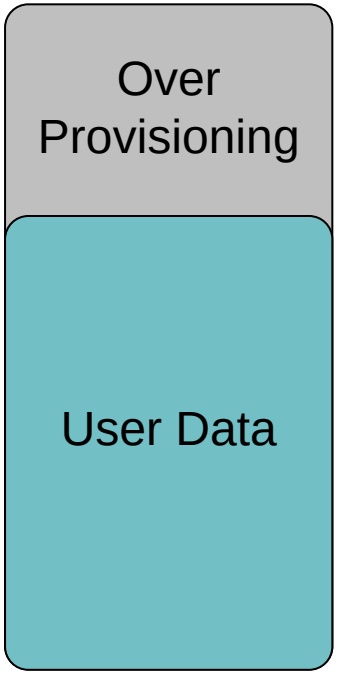
Understanding SSD Over Provisioning

Kent Smith
Sr. Director Product Marketing
Flash Components Division, LSI

What is SSD Over Provisioning (OP)

- It is part of all SSDs that use Flash memory
  - Required due to the inability to overwrite Flash without it first being erased
  - Therefore not part of HDDs
- The portion of the SSD capacity held in reserve (unavailable to the user):
  - Garbage collection (the major use)
  - SSD controller Firmware storage (small %)
  - Spare blocks (small %)
  - Some SSDs include other data protection beyond ECC, like RAISE™ technology (space requirement varies)



Flash-based SSD

Why is More OP a Good Thing?

- OP will consume part of the storage capacity otherwise available to users
- Most users presume that a **lower** OP is better to provide maximum user storage capacity
- Most users do not understand that **higher** OP generally provides:
  - Higher write performance
  - Lower “Write Amplification”
  - Longer Flash life (endurance)
  - Space for data protection beyond ECC

How is OP Calculated?

- The ratio of OP vs. total user capacity

$$\left(\frac{\text{Physical Capacity} - \text{User Capacity}}{\text{User Capacity}}\right) = \text{Over Provisioning}$$

- 128GB physical flash capacity
- 120GB user capacity
- 7% (6.67%) OP
- However, the “true” physical capacity is usually misunderstood…

What is a Gigabyte?

	Binary	Decimal
Exponential Notation	2^30	10^9
Actual Number of Bytes	1,073,741,824	1,000,000,000
Naming Convention	Gibibyte[IEC]	Gigabyte[SI]
Typical Uses for That Radix	System Memory	Storage/Networking

IEC – International Electrotechnical Commission
SI – International System of Units

- Some confusion due to the use of two different base numbers (radix)
- Binary GB is 7.37% more than Decimal GB
- Most OS's display the "binary" representation for all categories (system memory, storage, networking, etc.)

True Physical OP on SSDs

Over Provisioning Percentages				
Marketed OP*	0%	7%	16%	28%
True Physical OP*	7%	15%	25%	37%
SSD Physical Cap	Resulting SSD User Capacity			
64	64	60	55	50
128	128	120	110	100
256	256	240	220	200

- An SSD listed with 128GB is marketed as “0% OP”, but in reality the true physical OP is ~7%

*Rounded results

Performance Test – Environment

- Test result data points are based on post-garbage collection, steady state operation
- All preconditioning uses the same transfer size and type as the test result
  - E.g., random 4K results are preconditioned with random 4K transfers until it reaches steady state operation
- Test conducted on a single SSD to isolate the OP variable

Hardware:

- Intel Core i5-2500K 3.30 GHz
- 4 GB RAM 1333 MHz
- Intel H67 Express Chipset
- Intel RST 10.1.0.1008 (AHCI Enabled)
- Windows 7 Professional (32-bit)

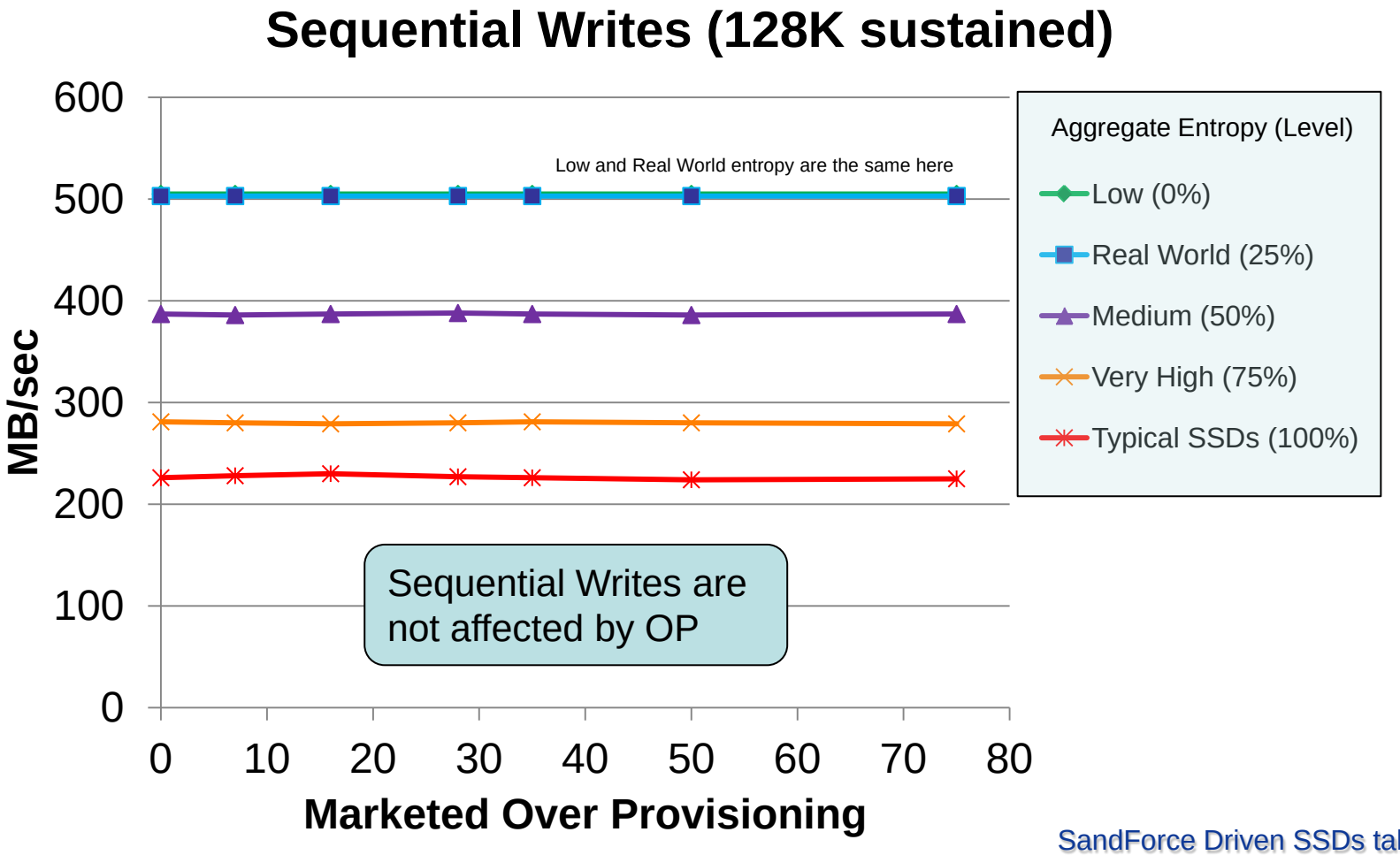
Software:

- VDBench V5.02 (main test SW)
- IOMeter V1.1.0 (cross check)

SSD:

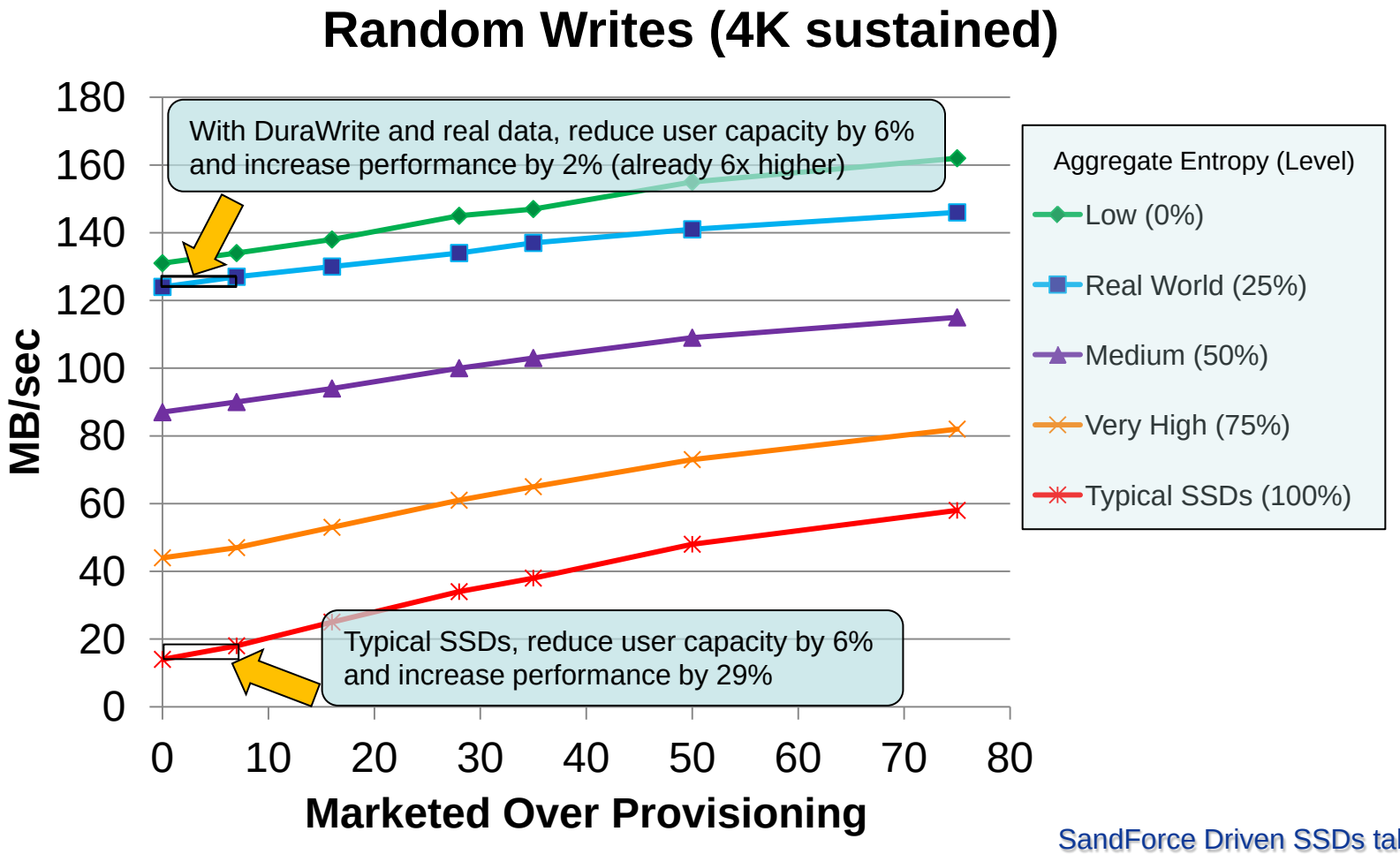
- MLC 24nm Toshiba NAND Flash
- SF-2281 FSP

Performance Test Results



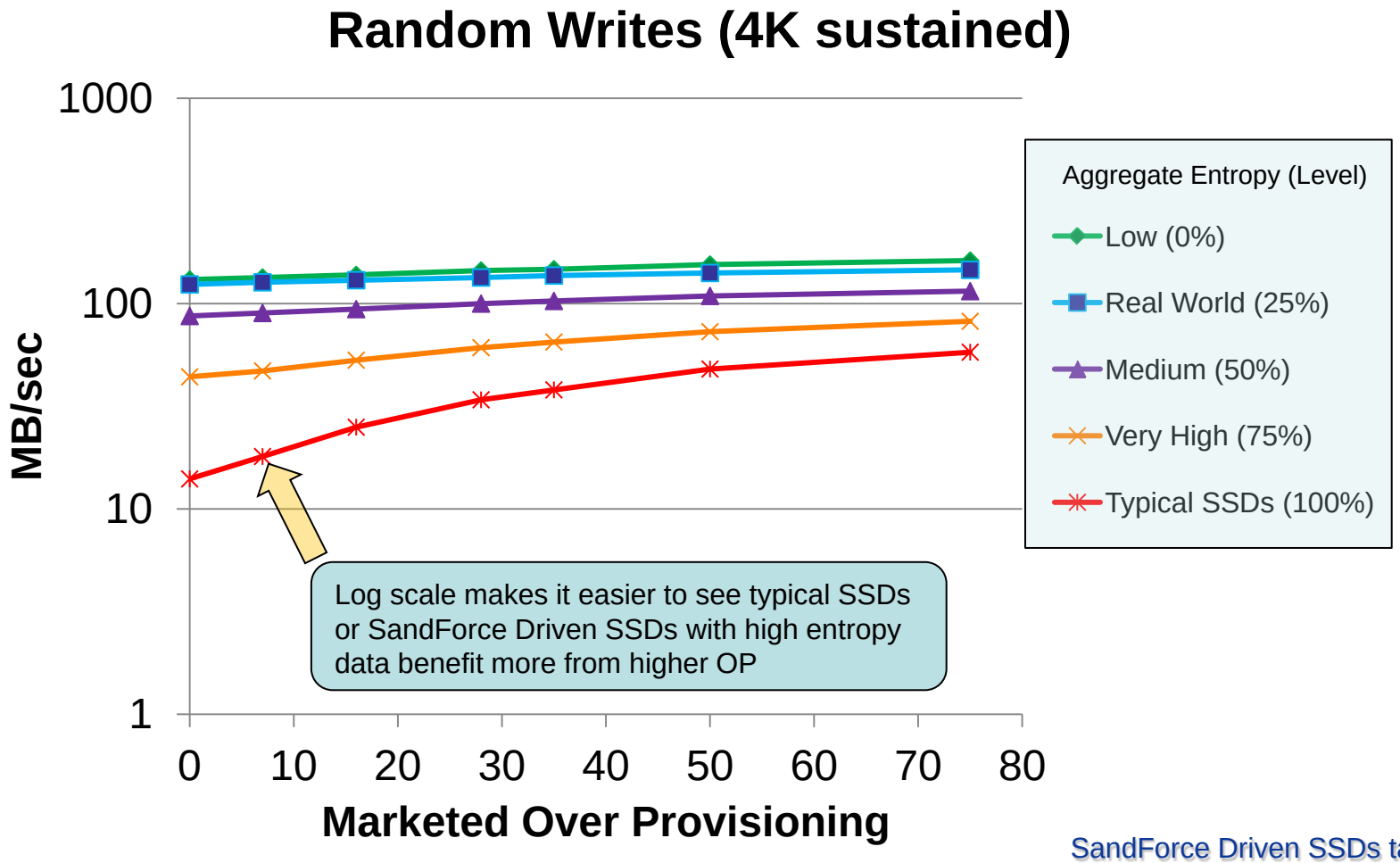
SandForce Driven SSDs take advantage of different entropy levels with DuraWrite™

Performance Test Results



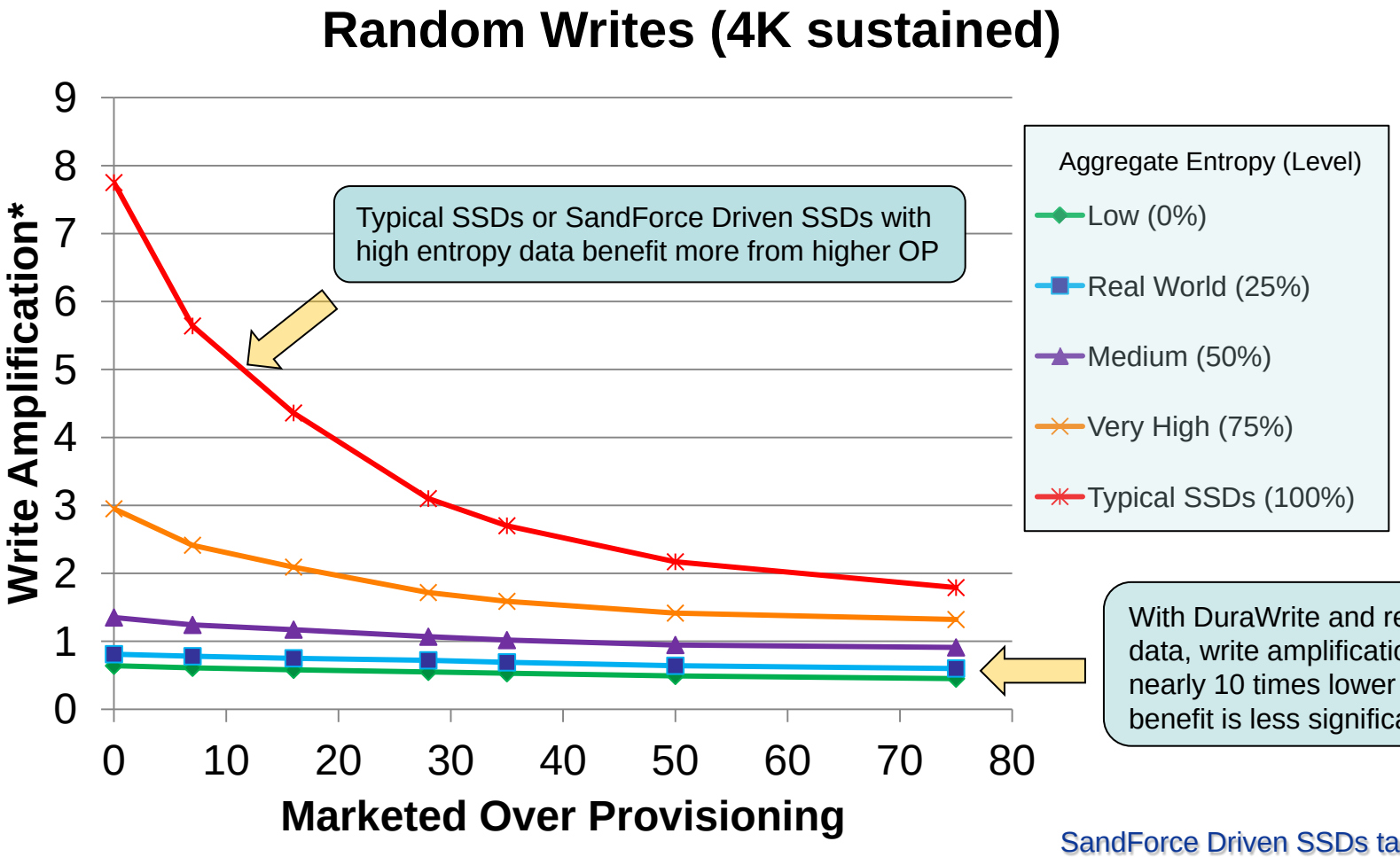
SandForce Driven SSDs take advantage of different entropy levels with DuraWrite™

Performance Test Results



SandForce Driven SSDs take advantage of different entropy levels with DuraWrite™

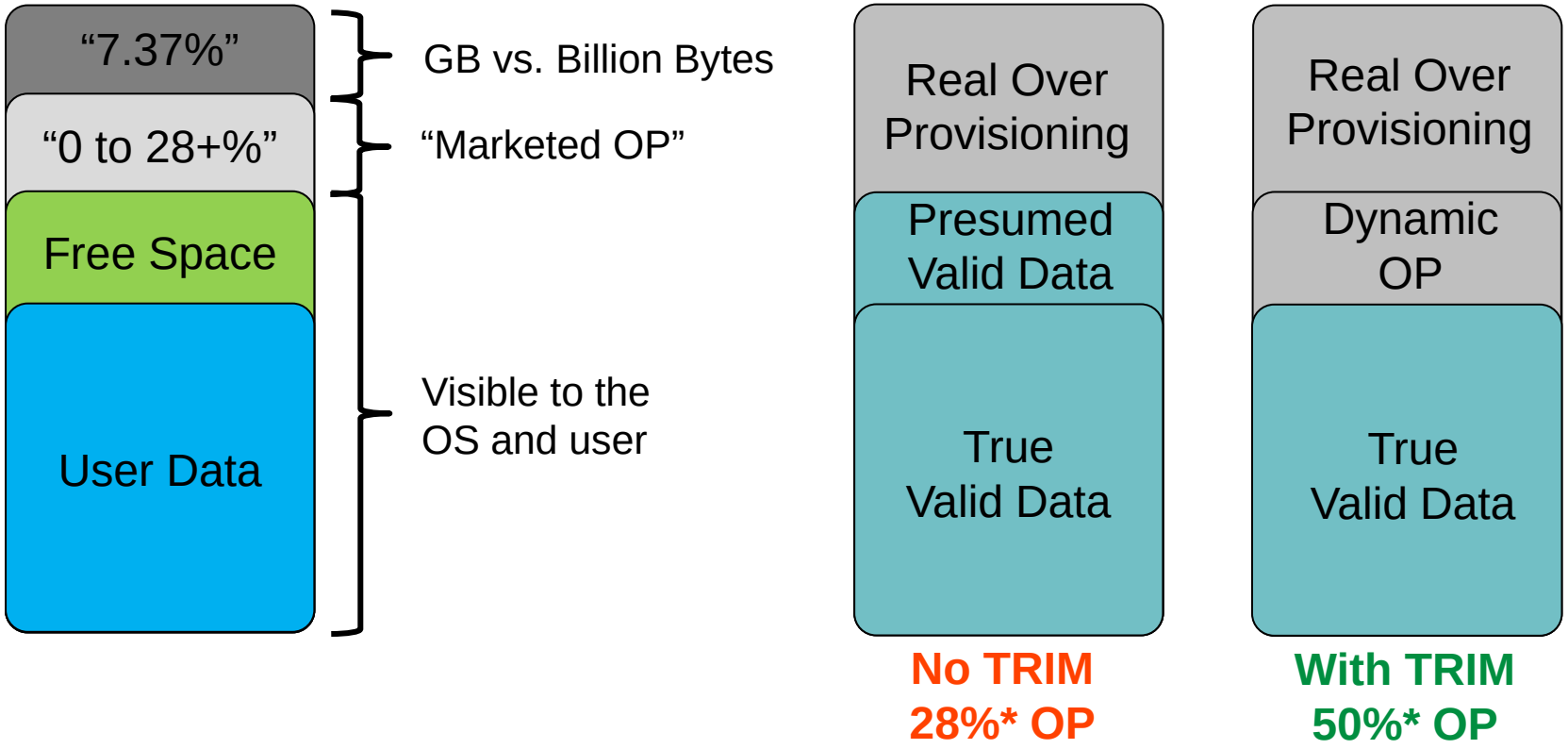
Write Amplification Test Results



*(GB Written to Flash / GB Written from Host)

SandForce Driven SSDs take advantage of different entropy levels with DuraWrite™

TRIM and Over Provisioning



- **Presumed Valid Data** – Data deleted by the OS or user, but the SSD is not aware of it because TRIM was not present
- **Dynamic OP** – TRIM expands the OP when the OS or user erases data

*Illustrative numbers only

DuraWrite, TRIM and Over Provisioning

DuraWrite provides additional dynamic OP like TRIM

	1. OS w/o TRIM or 2. RAID Environment*	1. OS with TRIM and 2. No RAID*
SSDs without DuraWrite	Free Space (OP) Presumed Valid Data True Valid Data	Free Space (OP) Dynamic OP True Valid Data
SSDs with DuraWrite	Free Space (OP) Larger due to DuraWrite Presumed Valid Data True Valid Data	Free Space (OP) Larger due to DuraWrite Dynamic OP True Valid Data

*Some RAID 0 starting to support TRIM

User Controlled Higher OP

- Users can increase the OP, but not decrease it
- During initial setup and formatting, allocate a smaller partition (don't use the full space)
  - SSD must be either "Fresh Out of Box" (FOB) or secure erased
- Leave the extra space unallocated
- The SSD controller automatically uses this as additional dynamic OP

Summary

- Over provisioning is a key component of any SSD
- Higher OP provides:
  - Higher write performance
  - Lower “Write Amplification”
  - Longer Flash life (endurance)
  - Space for data protection beyond ECC
- TRIM and DuraWrite contribute to OP
- Users can easily increase the OP if desired

LSI is Accelerating Mobile Computing at the FMS 2012!

- **Visit us at the booth 624-626 to:**
  - Experience the latest SandForce Driven™ Ultrabook™ systems
  - See a live demo of the SandForce Driven Kingston USB SSD with Windows To Go
  - Enter to win:
    - SandForce Driven SSDs from **SuperSSpeed, Kingston, Corsair, Adata, EDGE, Mushkin, OWC, PNY, DMS, Patriot Memory and Wintec**

      Up to 4 winners every 30 minutes!

    - **Grand Prize – the latest SandForce Driven ASUS Zenbook Prime!**

      Don't miss the drawing on Wednesday 8/22 @ 6:30pm