Accelerating SSD Performance with HLNAND

Roland Schuetz

MOSAID Technologies Inc.

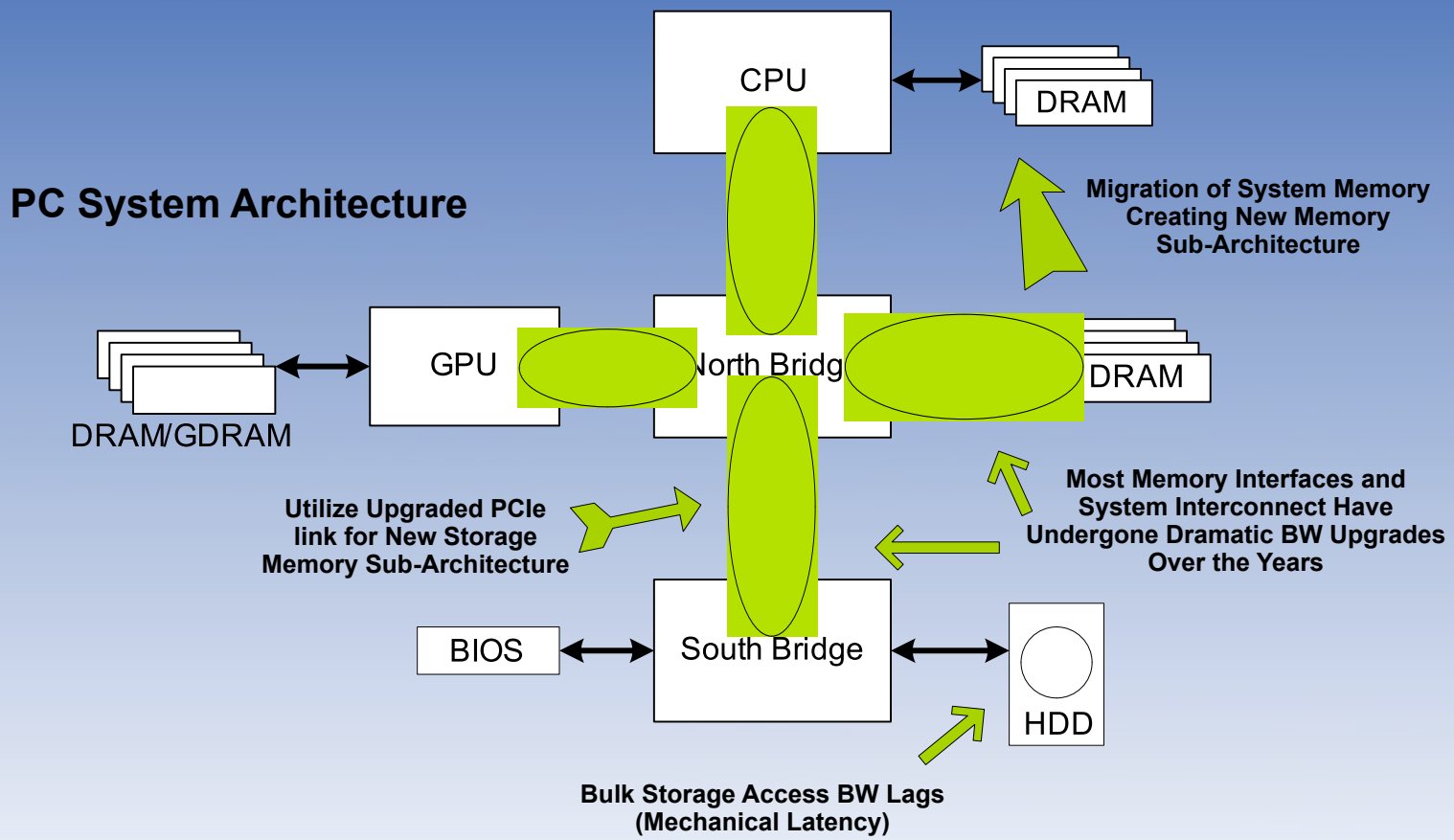
David Won, INDILINX

Santa Clara, CA USA
August 2008

Presentation Outline

- PC Architecture and the Solid State Drive
- HLNAND Introduction
- HLNAND Enhances SSD Performance
- Conclusions

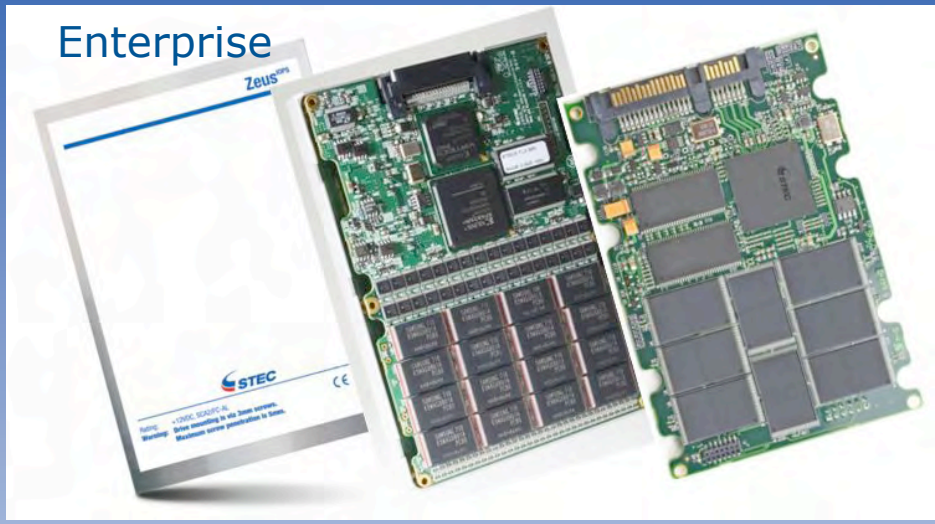
PC System Architecture
Where are the Bottlenecks?



SSD Market Segment

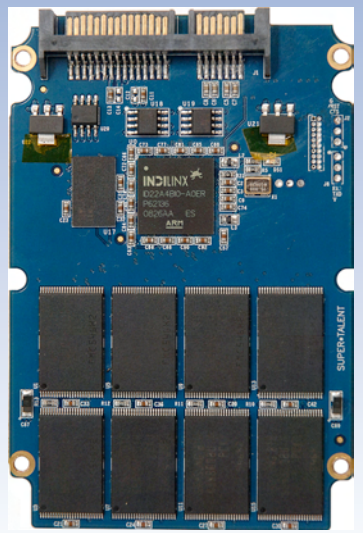
- Three categories:
  - Enterprise
  - Notebook / PC
  - NetBook / Ultraportable

Enterprise



Source: STEC

Notebook



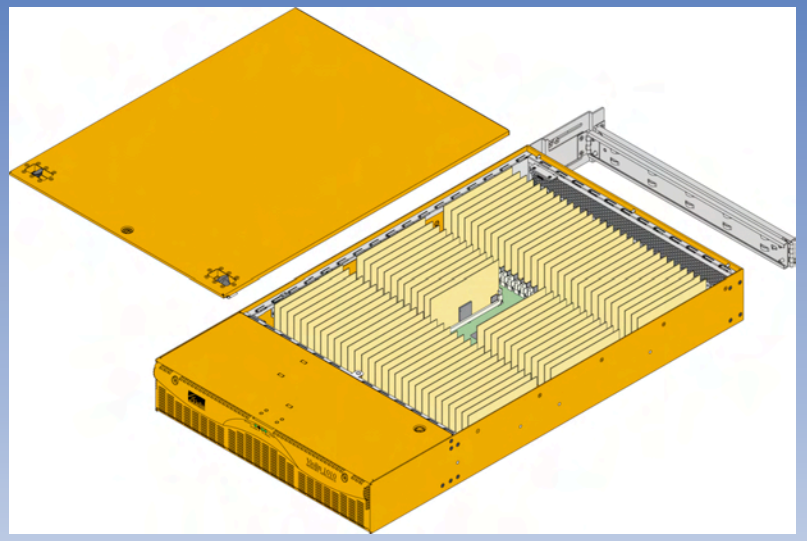
Source: INDILINX & Super Talent

NetBook / Ultraportable

Source: STEC

Enterprise Flash Storage Excellent Market for HLNAND SSDs

- Enterprise application move huge amounts of data and are IO bound
- These include:
  - File, web, transaction servers
  - Multimedia editing systems
  - Simulation servers/ workstations
- Current Flash based enterprise applications based on conventional, 40 – 100 MBps sub-systems



Source: Violin 1010 Memory Appliance, Violin Memory, Inc.

Current SSD Offerings (1st Generation)

- Leading NAND, memory module, and specialized SSD manufacturers offer conventional NAND flash based product
- Cost is pivotal for SSD adoption
- Little product and architectural differentiation
  - 4 channel
  - 4-way interleave
- Similar cost structure
  - NAND flash constitutes the majority of the BOM
  - Vertically integrated manufacturers have pricing advantage
- Similar performance
  - 30 ~ 100MB/s Read/Write performance
  - No one competitor has performance lock on the market

How are Current SSDs Stacking Up?

GIZMODO TEAM

Tip your editors:
tips@gizmodo.com

DISAPPOINTMENT

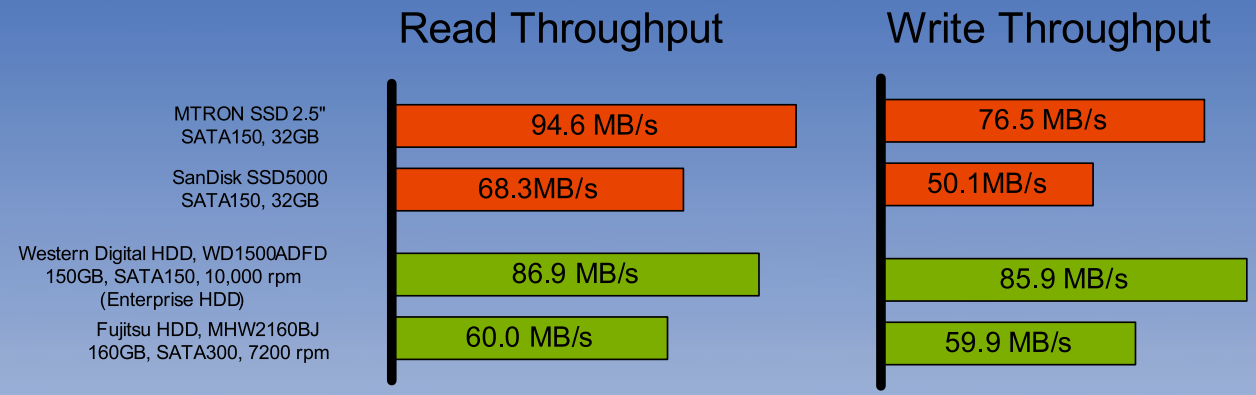
« | next »

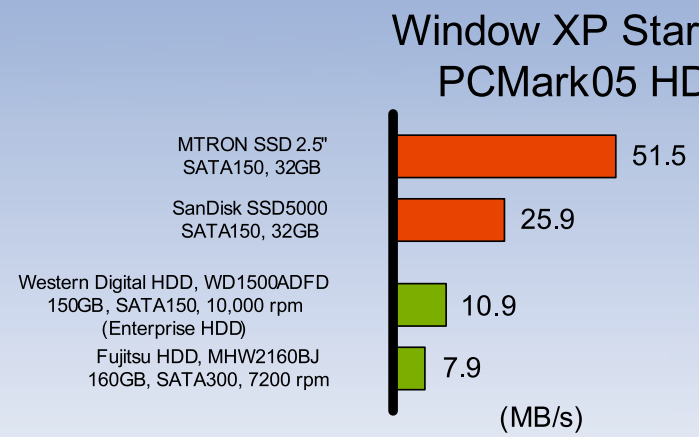
The Macbook Air SSD Performance Boost: Pretty Much Non-Existent

The gang over at Ars Technica got their hands on both an SSD-packed Macbook Air as well as a regular, HDD-equipped model and then put both through their paces to see if the $1,300 difference in price was worth it. So they had two computers, both exactly the same save a 200Mhz bump in processor speed and a swapping out of a 4200rpm PATA drive for a hot, hot SSD drive. So did the performance shoot through the roof, confirming all of our wildest wet dreams about SSDs? Uh, no, not really.

Source: Engadget

Current SSDs vs. HDDs



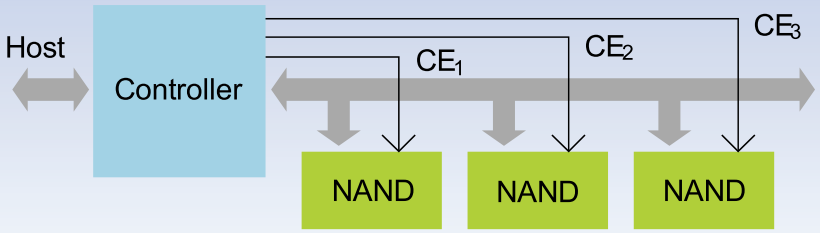


Sources:
Tom's Hardware, **Mtron SSD 32 GB: Performance with a Catch**, Patrick Schmid, Achim Roos, November 21, 2007;
Tom's Hardware, **Flash-Based Hard Drives Cometh**, Patrick Schmid, Achim Roos, August 13, 2007;

HLNAND – New High Speed NAND Standard

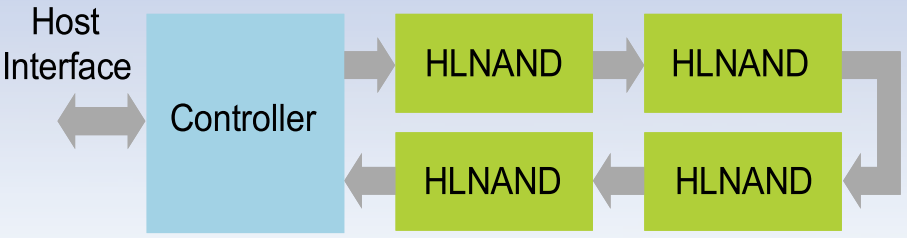
Conventional NAND

- 8 bit, bidirectional, multi-drop bus
- Asynchronous LVTTL signaling up to 40Mb/s/pin
- Speed degradation with more than 4 devices on bus
- Chip Enable (CE) signal required for each device
- Power hungry 3.3V I/O



HLNAND

- Unidirectional, point-to-point, daisy-chain cascade with programmable link width; 1- 8 bits
- Synchronous DDR signaling up to 800Mb/s/pin
- Each ring supports up to 255 devices with no bandwidth degradation
- Single CE per ring enables pin controller count reduction
- Low Power 1.8V I/O



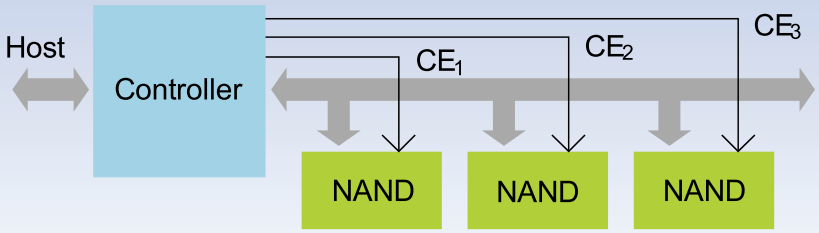
HLNAND – New High Speed NAND Standard (cont'd)

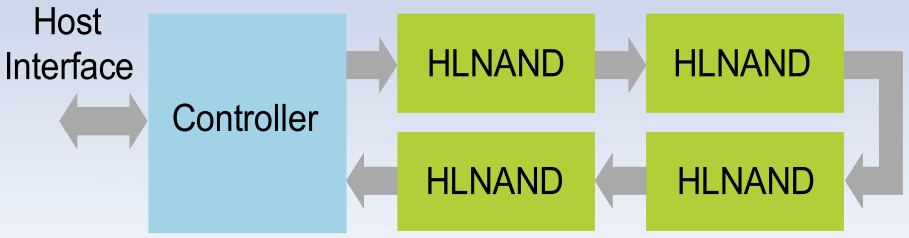
Conventional NAND

- No new features to enhance flash performance or simplify controller and SSD design

HLNAND

- New device features enhance performance and simplify controller and SSD design
- New, low-stress program scheme enables:
- Random page program
- Page-pair erase
- Multi-page & Multi-block erase





PC Demands on System Storage

- Certain types of PC use are IO intensive -> frequent HDD/SSD access; Ex: creating slide presentation, virus scan, etc.
- Booting PC is very IO intensive since the OS must load a large amount of data from bulk storage to DRAM

IO Operation Example
Can SSD Deliver "Instant Boot"

- Populate 1GB DRAM from Flash Based Bulk Storage
- Conventional Flash SSD with 60MB/s BW takes ~17sec
- Optimized conventional flash SSD with 95MB/s* takes ~10 sec

10 Seconds is very observable time!

* **Mtron SSD 32 GB: Performance with a Catch Patrick Schmid, Achim Roos,** November 21, 2007; SSD: Mtron MSDSATA6025032NA

IO Operation Example
Can SSD Deliver "Instant Boot"

- Populate 1GB DRAM from Flash Based Bulk Storage
- HLNAND SSD with one Ring and 266MB/s BW takes ~3.8sec
- Optimized HLNAND SSD with 4 Rings and ~1.1GB/s BW takes ~.9sec

HLNAND based SSDs offer 260% - 1100% IO rate improvement.
0.9 seconds is virtually instant!!

IOPS are Holly Grail in Enterprise Applications

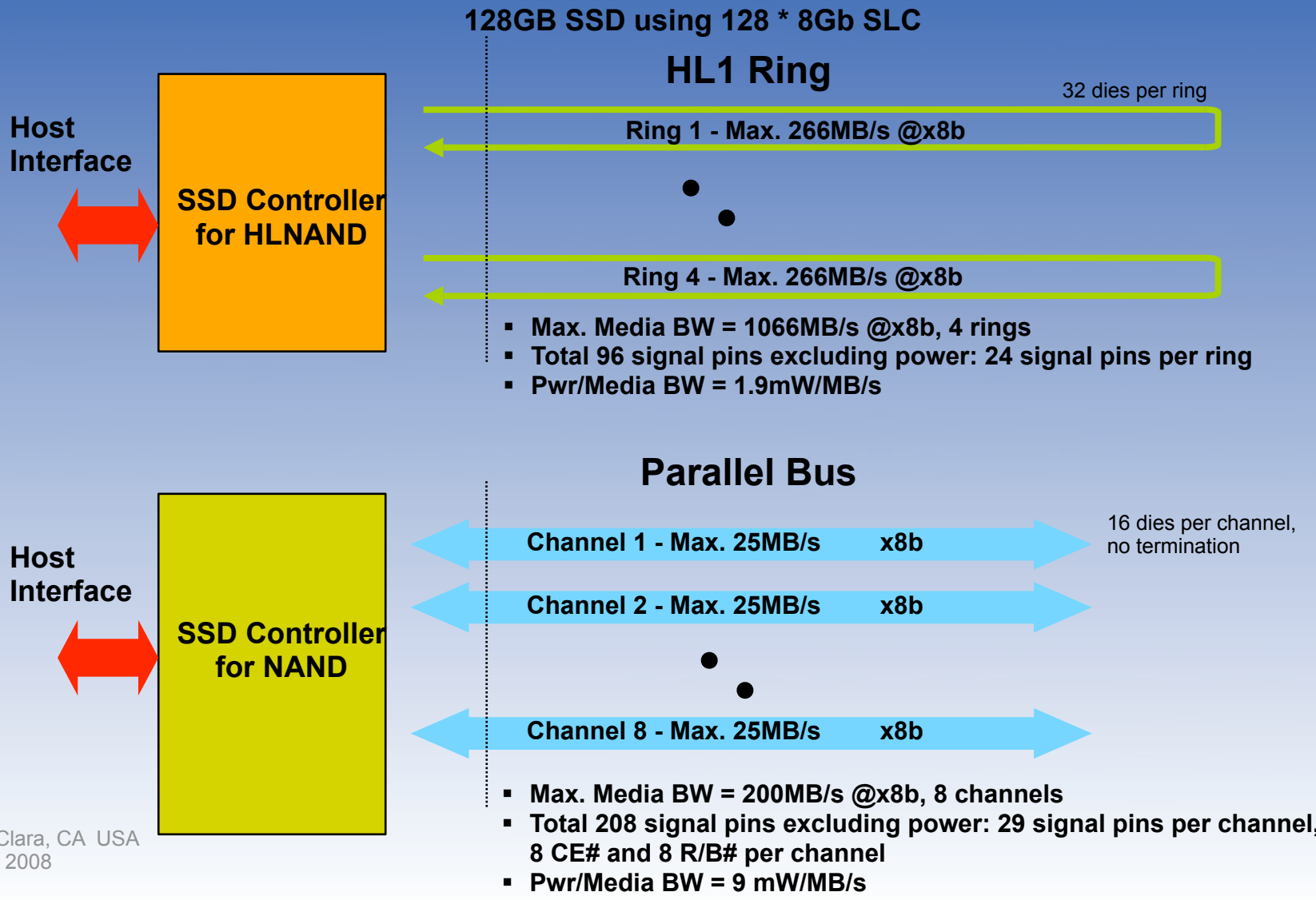
Holy Grail of the Monastery of Xenophontos, Macedonian Heritage, 2000-2008

- SSDs offer significantly higher I/O (inputs/outputs) per second Than HDDs: HLNAND will shine
- HDD (typical enterprise class) ~150
- SSD, SATA (advertised by Sandisk) 7,000
- SSD, 4G Fiber Channel (advertised by STEC) 45,000
- HLNAND 300,000* IOPS

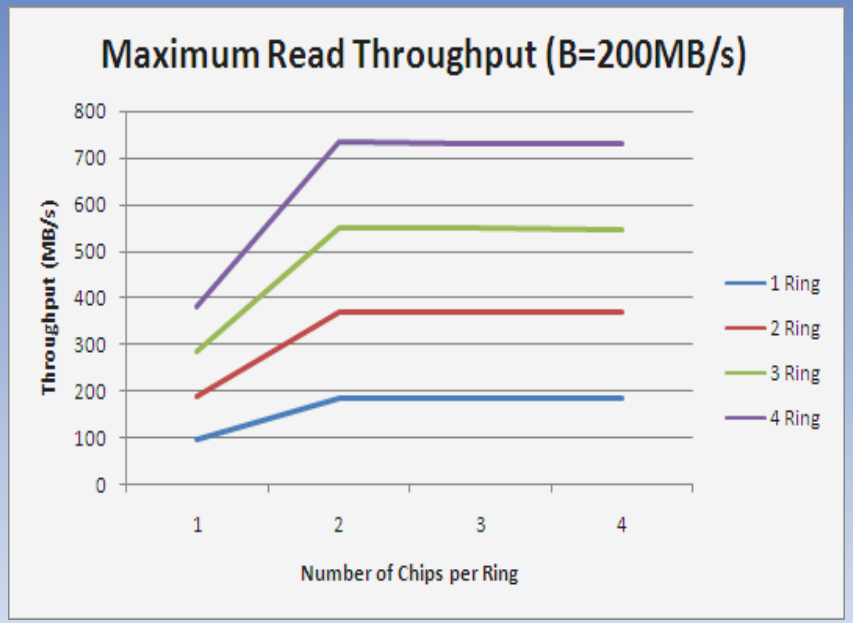
Source: Sandisk, STEC, and Deutsche Bank estimates

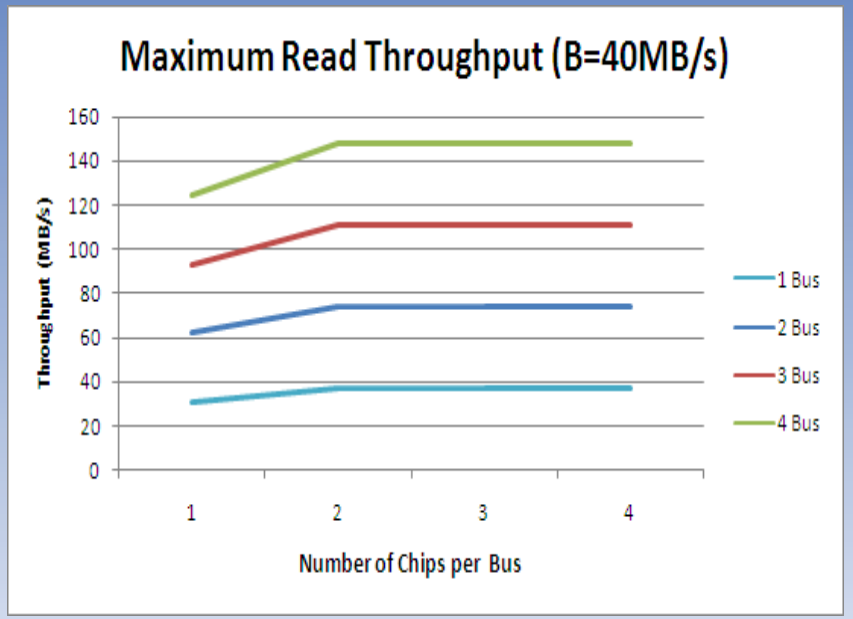
* Assume similar design to STEC SSD and multiply by media speed improvement; 45,000 * (266/40).Not including 4GFC saturation limit.

SSD Performance Comparison
HLNAND-based vs. NAND-based



HLNAND SSD vs. NAND SSD
Read Throughput, Simulation Results





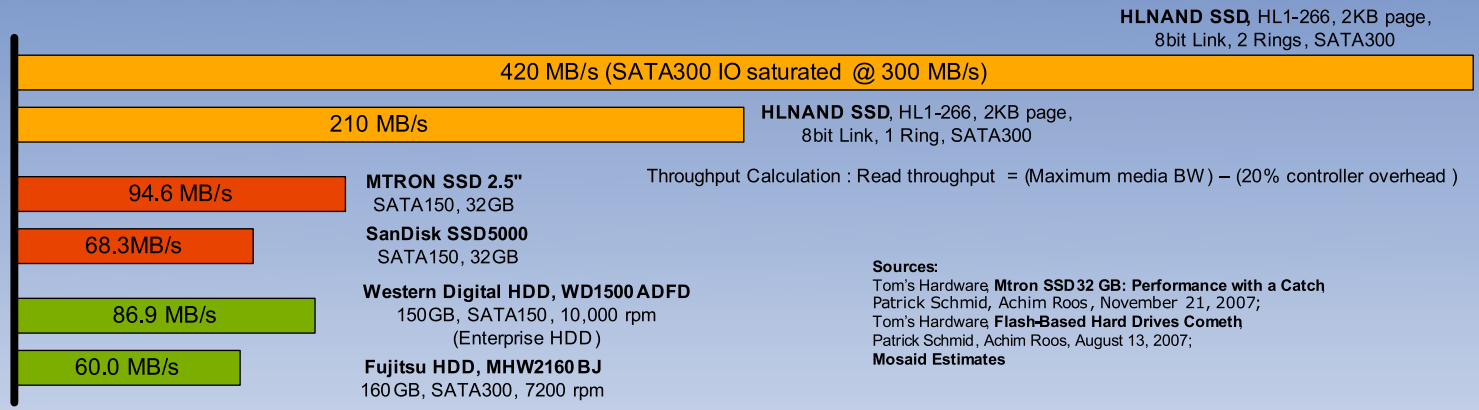
HLNAND-based SSD, Media Rate

NAND-based SSD, Media Rate

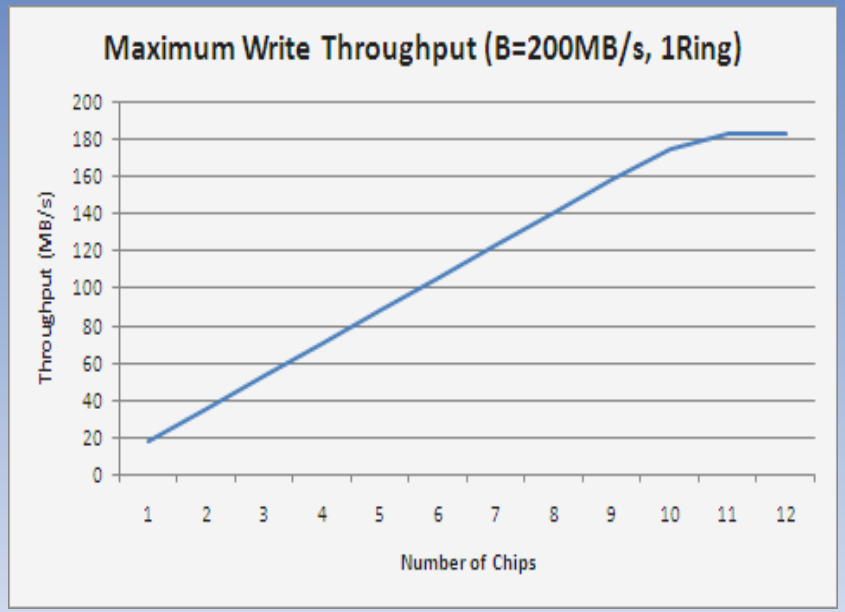
Plots from Mobile Embedded Lab, University of Seoul, 2008

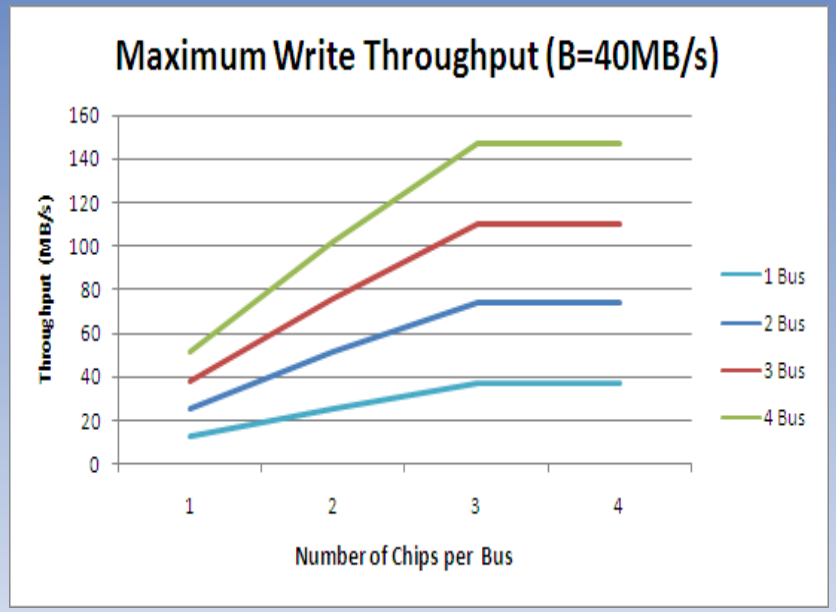
HLNAND SSD vs. SSD vs. HDD
Estimates and Measured Performance

Read Throughput



HLNAND SSD vs. NAND SSD
Write Throughput, Simulation Results





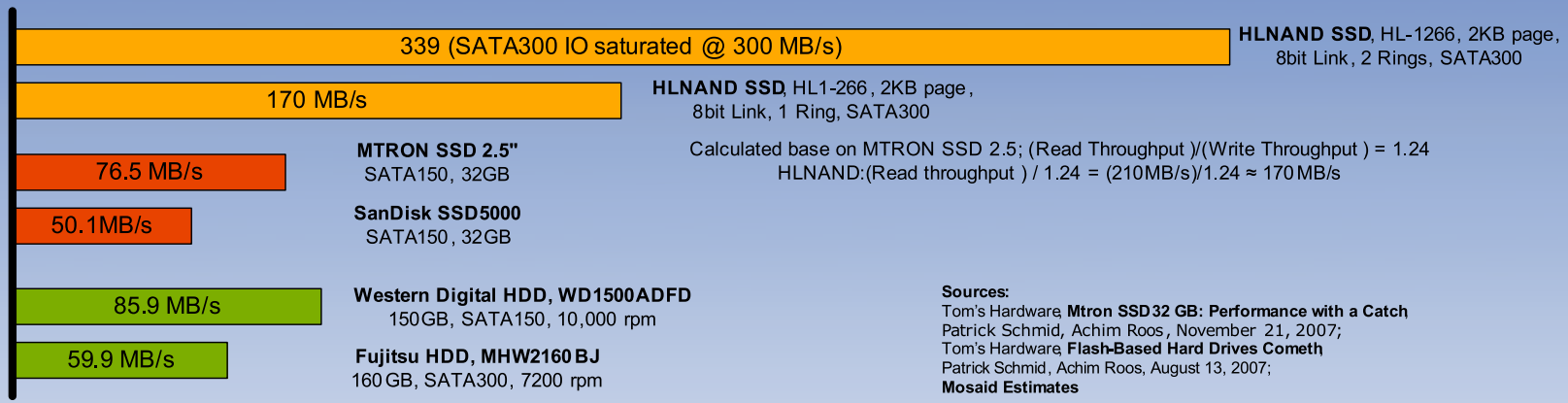
HLNAND-based SSD, Media Rate

NAND-based SSD, Media Rate

Plots from Mobile Embedded Lab, University of Seoul, 2008

HLNAND SSD vs. SSD vs. HDD
Estimates and Measured Performance

Write Throughput



Sources:
Tom's Hardware, **Mtron SSD 32 GB: Performance with a Catch**
Patrick Schmid, Achim Roos, November 21, 2007;
Tom's Hardware, **Flash-Based Hard Drives Cometh**
Patrick Schmid, Achim Roos, August 13, 2007;
Mosaid Estimates

SSD Controller Design

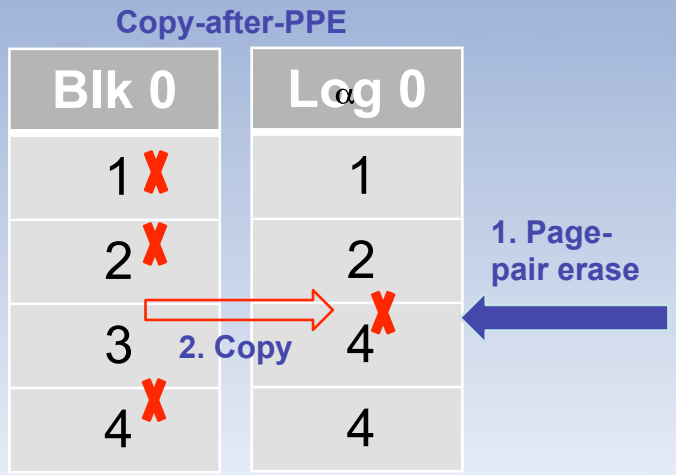
- Controller responsible for high cost jobs of wear-leveling and garbage collection
- Wear leveling operations are block based and include
  - Block Merge
  - Block Switch
  - Block Switch after copy
  - Block Migration
- HLNAND provides means of doing page-pair (wordline) based wear-leveling reducing background operations

HLNAND Improves Performance and Reliability

- New HLNAND features reduce block recycling costs
  - Page-Pair Erase
  - Random page program
  - Partial block erase
  - Multi block erase
- Erase size matches prog/read size thereby reducing program erase cycles related to wear-leveling

Wear Leveling Enhancement: Copy-After-Page-Pair-Erase

- Page-pair erase introduces low cost wear-leveling opportunities
- Translates into greater system longevity and less controller overhead



$$\frac{\frac{C_E}{M} + C_E + k \times C_{cp} + \alpha}{N_p}$$

C_E: Erase cost

C_{cp} : Copy cost

M : Number of blocks erased concurrently

N_p: Number of pages in a block

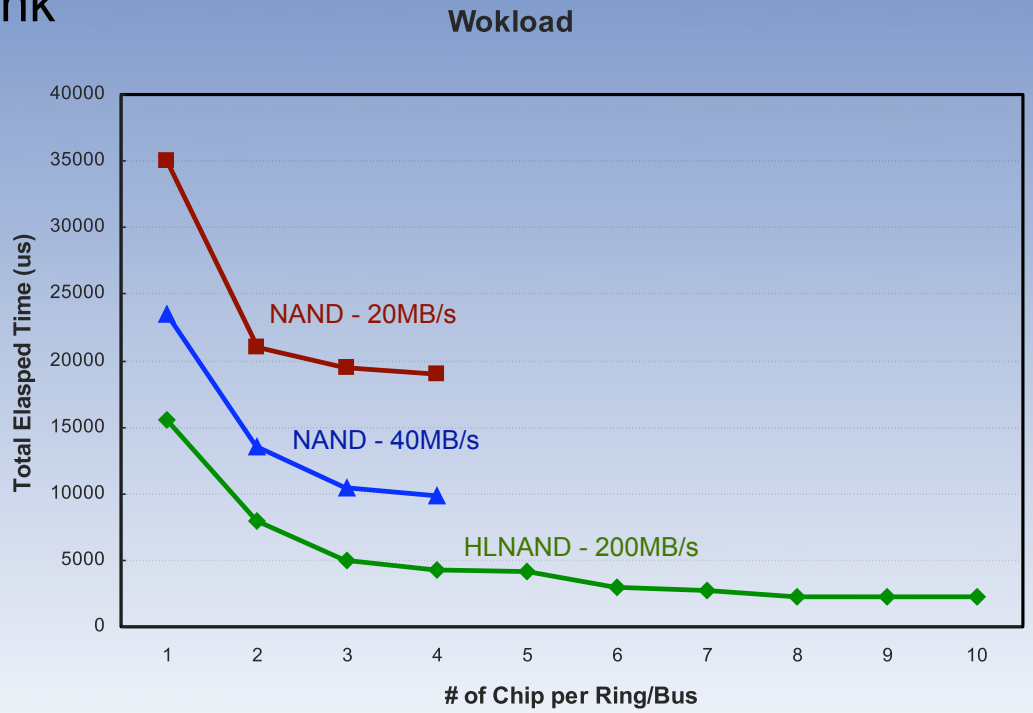
k: Number of page-copies

α: Additional copy overhead

Synthetic Workload Simulation

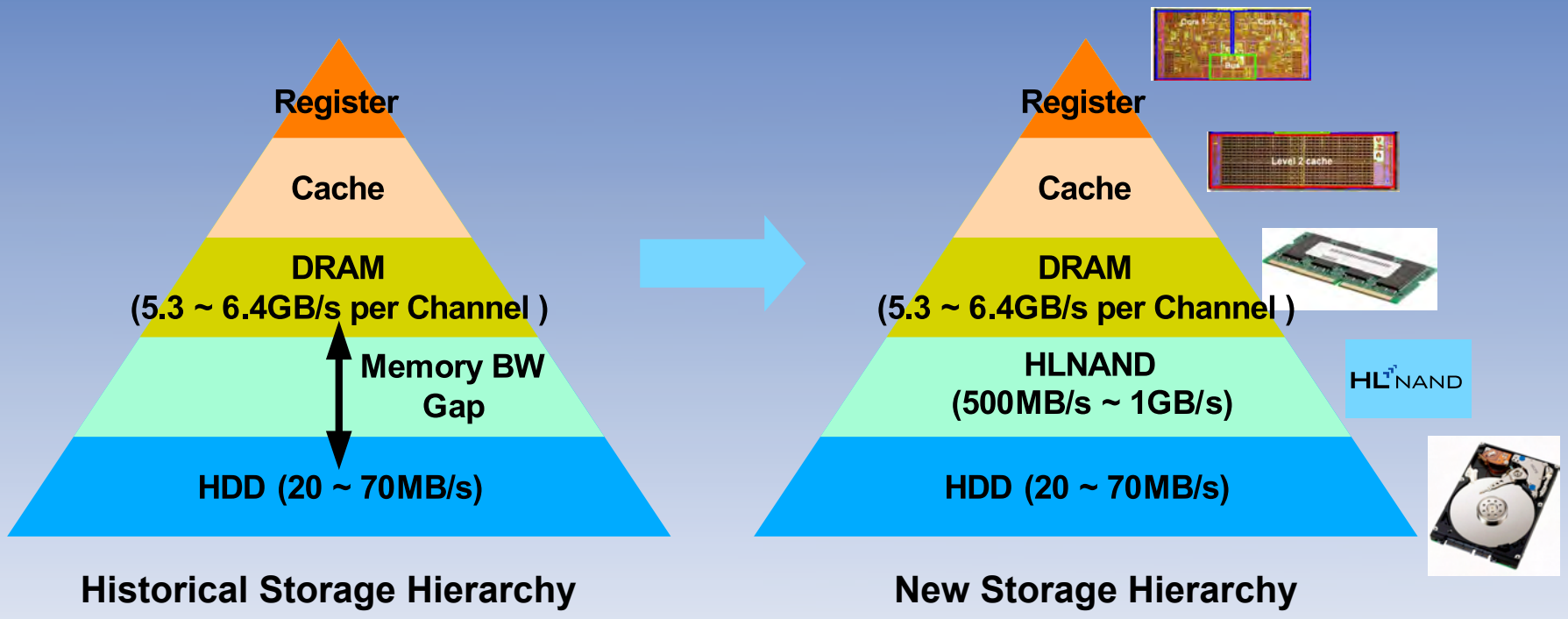
- X8b
- 4KB page size
- 256KB block size
- 2048 blocks/bank
- 1 bank

Read	Program	Copyback	Erase
60%	25%	10%	5%
617	263	82	38



Experiment performed by Mobile Embedded Lab, University of Seoul, 2008

New Hierarchy for New User Experience



Conclusions

- HLNAND's features contribute to the acceleration of SSD satisfaction and adoption

High Speed Interface

Low Power Consumption

High Scalability

Interface Extensibility

Advance Core Features

Reduced Overall Cost with Increased Performance