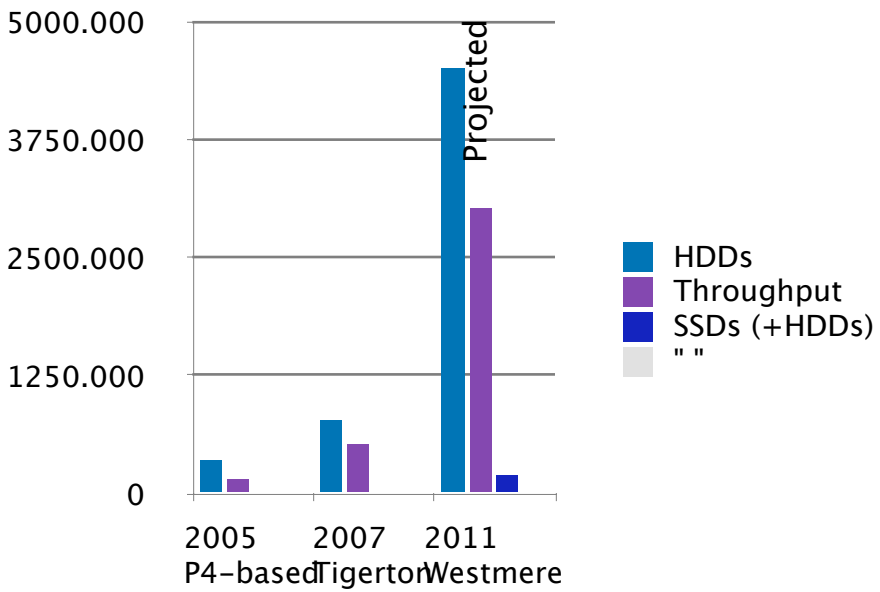

intel
Next Frontier in SSD Performance
Knut Grimsrud
Aug. 10, 2011
STG
Storage Technologies Group

Legal Notices and Disclaimers

- INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL® PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. INTEL PRODUCTS ARE NOT INTENDED FOR USE IN MEDICAL, LIFE SAVING, OR LIFE SUSTAINING APPLICATIONS.
- Intel may make changes to specifications and product descriptions at any time, without notice.
- All products, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice.
- Intel, processors, chipsets, and desktop boards may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available on request.
- Any code names featured are used internally within Intel to identify products that are in development and not yet publicly announced for release. Customers, licensees and other third parties are not authorized by Intel to use code names in advertising, promotion or marketing of any product or services and any such use of Intel's internal code names is at the sole risk of the user.
- Intel product plans in this presentation do not constitute Intel plan of record product roadmaps. Please contact your Intel representative to obtain Intel's current plan of record product roadmaps.
- Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to http://www.intel.com/performance
- Intel, Intel Inside, the Intel logo, Centrino, Intel Core, Intel Atom, Pentium and UltraBook are trademarks of Intel Corporation in the United States and other countries.
- **Material in this presentation is intended as product positioning and not approved end user messaging.**
- **This document contains information on products in the design phase of development.**
- *Other names and brands may be claimed as the property of others.
-

SSD Storage Already Revolutionary

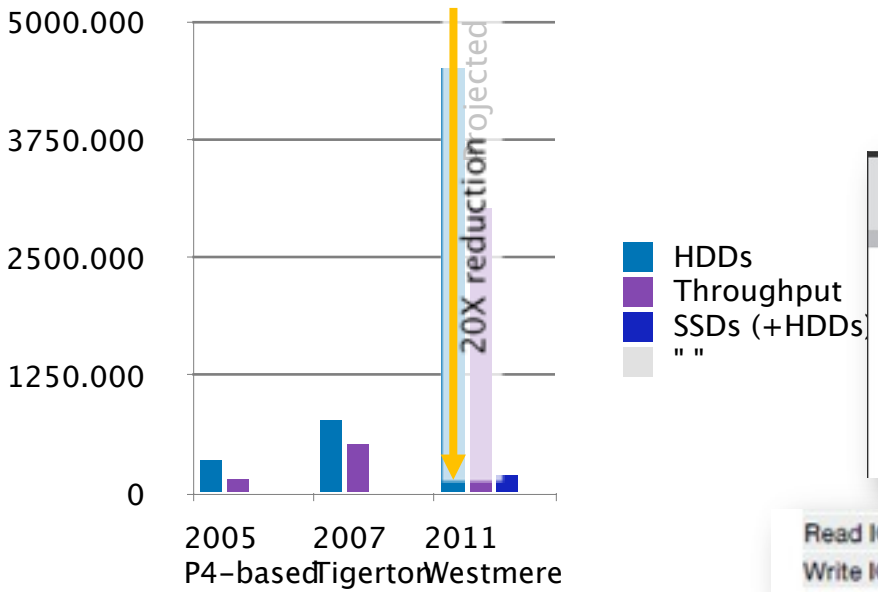
- In some cases SSDs have already realized 20X improvements in enterprise environments

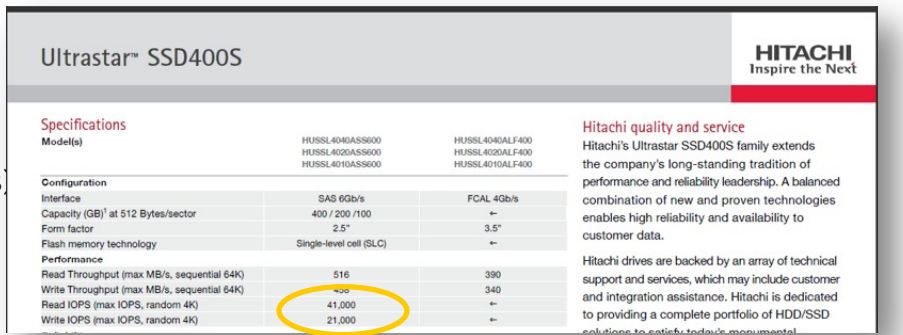


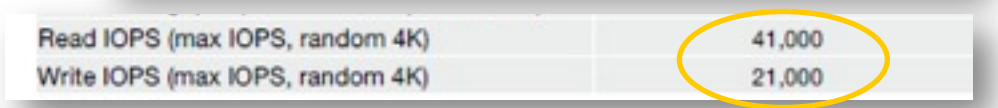
SSD Storage Already Revolutionary

- In some cases SSDs have already realized 20X improvements in enterprise environments

I/O Intensive OLTP Workload







The first 20X performance increase having dramatic impacts

*Third party names and brands are the property of their respective owner
**Datasheet image used with permission from Hitachi GST
***OLTP system configuration data from Intel OLTP benchmarking lab

What About the Next 20X?

- Let me show you

* Prototype: Supports 8 NVMe queues, uses DRAM to highlight interface/stack capability, workload is 512B random reads (aligned to 4KB), System: Intel Core i7-2600K Sandy Bridge 3.4GHz (3.8GHz Turbo Boost) 4 x 256KB L2 Cache 8MB L3 Cache LGA 1155 95W Quad-Core Desktop Processor with Intel HD 3000 graphics and 8GB DRAM. Windows 7 Ultimate 64-bit. Rendering engine Iron Galaxy via Digital Extremes, Bink video encoder

What About the Next 20X?

- Let ou

1M+ IOPs
from one slot in mainstream platform

* Prototype: Supports 8 NVMe queues, uses DRAM to highlight interface/stack capability, workload is 512B random reads (aligned to 4KB), System: Intel Core i7-2600K Sandy Bridge 3.4GHz (3.8GHz Turbo Boost) 4 x 256KB L2 Cache 8MB L3 Cache LGA 1155 95W Quad-Core Desktop Processor with Intel HD 3000 graphics and 8GB DRAM. Windows 7 Ultimate 64-bit. Rendering engine Iron Galaxy via Digital Extremes, Bink video encoder

What About the Next 20X?

* Prototype: Supports 8 NVMe queues, uses DRAM to highlight interface/stack capability, workload is 512B random reads (aligned to 4KB), System: Intel Core i7-2600K Sandy Bridge 3.4GHz (3.8GHz Turbo Boost) 4 x 256KB L2 Cache 8MB L3 Cache LGA 1155 95W Quad-Core Desktop Processor with Intel HD 3000 graphics and 8GB DRAM. Windows 7 Ultimate 64-bit. Rendering engine Iron Galaxy via Digital Extremes, Bink video encoder

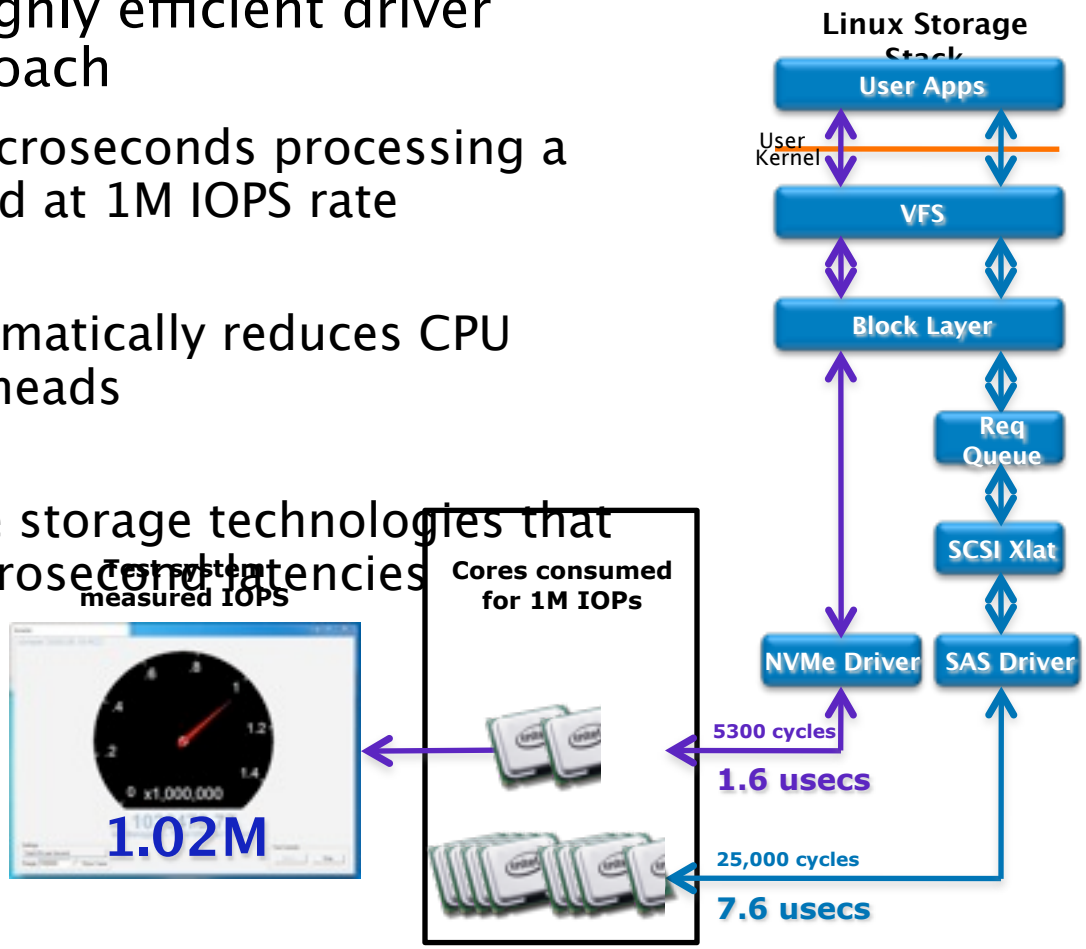
Efficient Interface Key Enabler

- PCIe physical interface for scalable b
  - 8X PCIe Gen3 provides up to ~8GE
- Efficient NVMe logical programming interface
  - NVMe supports efficient all-hardware controller in device performance path to avoid firmware
- Spec available and broad support
  - NVMe 1.0 published March and freely
  - Broad 80+ member adoptio
  - Infrastructure support maturing
    - NVMe now standard feature in all LeCroy PCIe protocol analyzers
  - Linux driver published, Windows baseline driver in development

For more information, visit http://nvmexpress.org

Stack Must Also Be Efficient

- 1M IOPS calls for highly efficient driver approach
  - Can't spend multiple microseconds processing a storage command at 1M IOPS rate
  - Streamlined stack dramatically reduces CPU overheads
  - Important for some future storage technologies that may have sub microsecond latencies

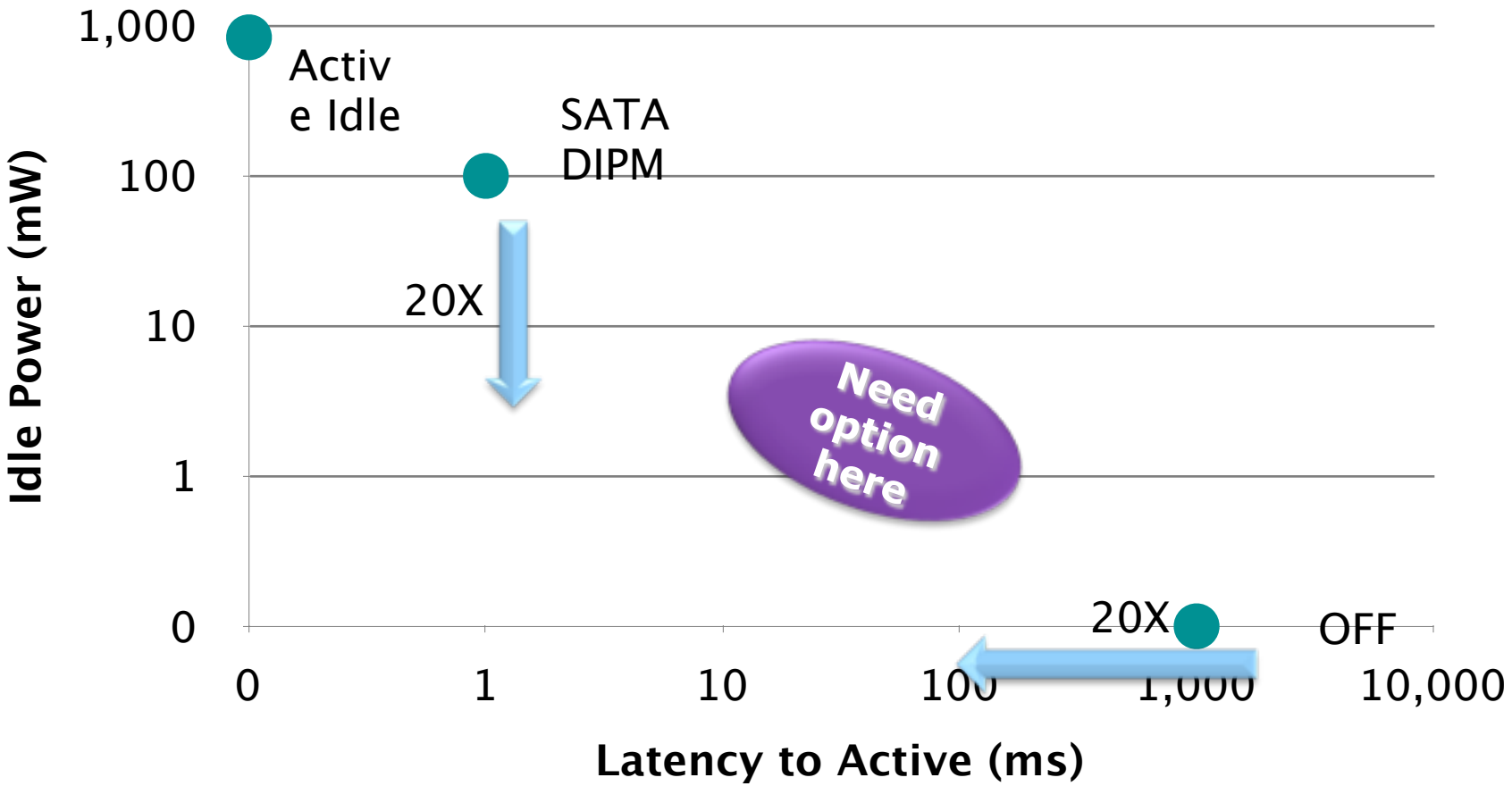


* Measurement taken on Intel Core i5-2500K Sandy Bridge 3.3GHz 6MB L3 Cache Quad-Core Desktop Processor using Linux RedHat EL6.0 2.6.32-71 Kernel using FIO with raw IO

It's Not All About Enterprise

- Coolest consumer and client devices use non-volatile memory for storage

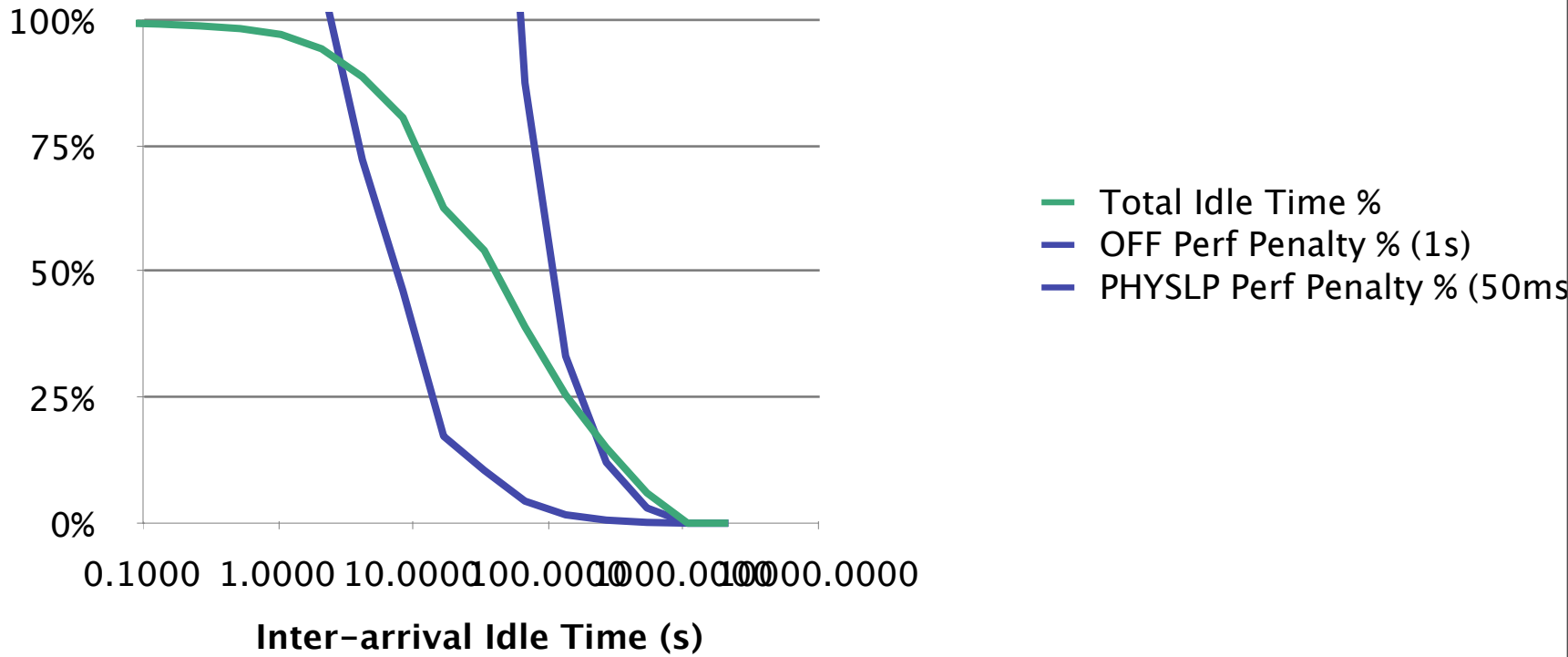
Low Idle Power SATA SSD Requirements



Target 5mW at 50ms. 20X lower power than DIPM or 20X faster than OFF.

Resume Latencies & Lost Time

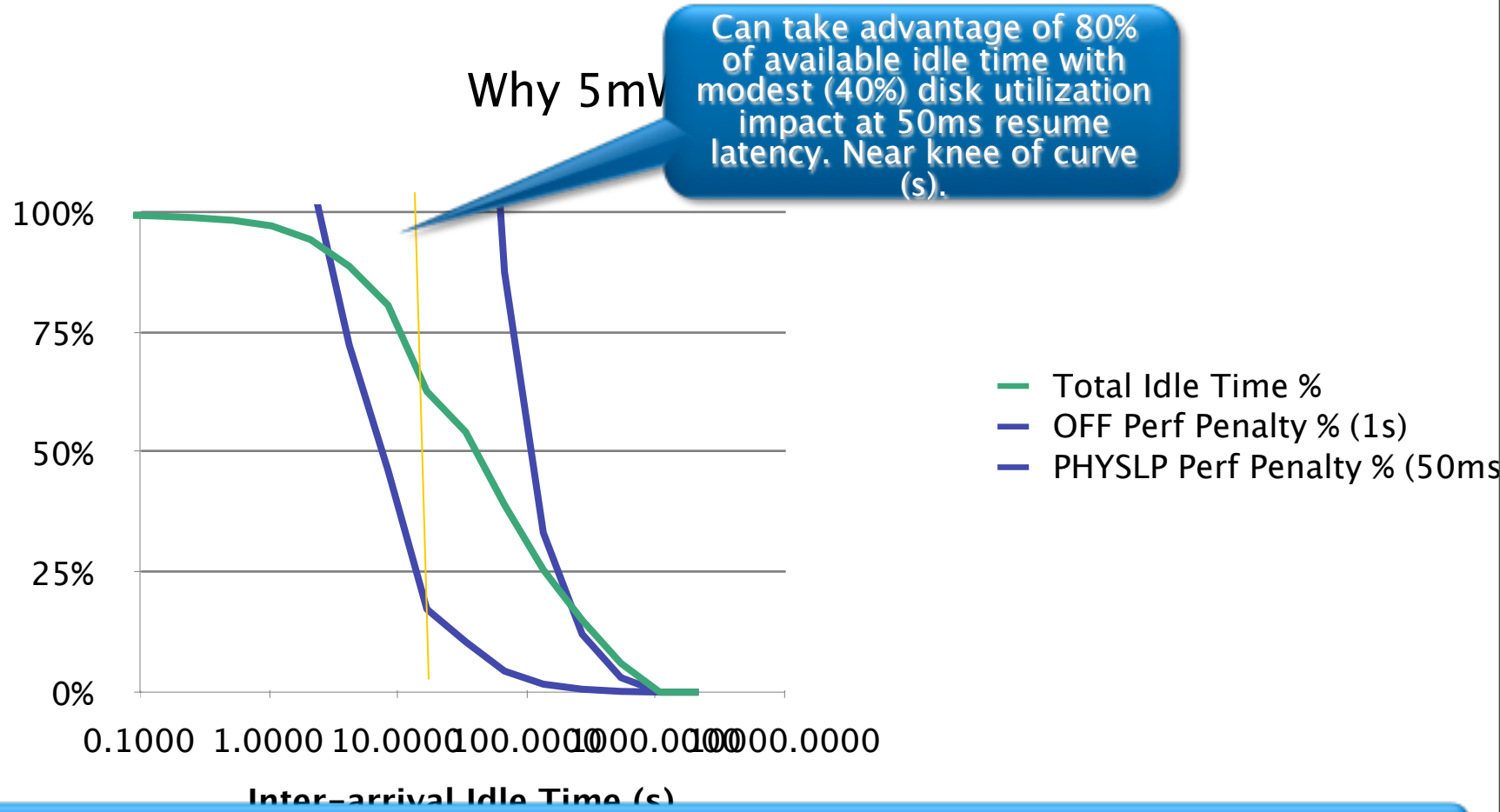
- Why 5mW and 50ms?



* Measurement taken from MobileMark 07** workload using an Intel160GB Postville SSD
** Other brands and names are the property of their respective owners

Resume Latencies & Lost Time

- Why 5mV

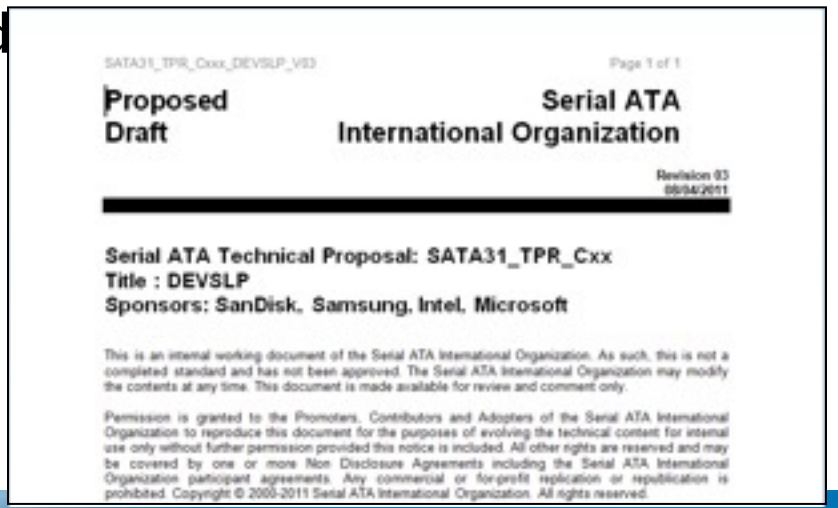


Power models suggest net halving of overall workload power feasible

* Measurement taken from MobileMark 07** workload using an Intel160GB Postville SSD
** Other brands and names are the property of their respective owners

New SATA Mechanism Key Enabler

- 20X idle power reduction requires enhanced mechanisms to support
  - The challenge is not going to sleep, but waking up
- Proposal for SATA key enabler in progress to address
  - Collaboration among lead



SATA31_TPR_Cxxx_DEVSLP_V03 Page 1 of 1

Proposed Draft **Serial ATA International Organization**

Revision 03
09/04/2011

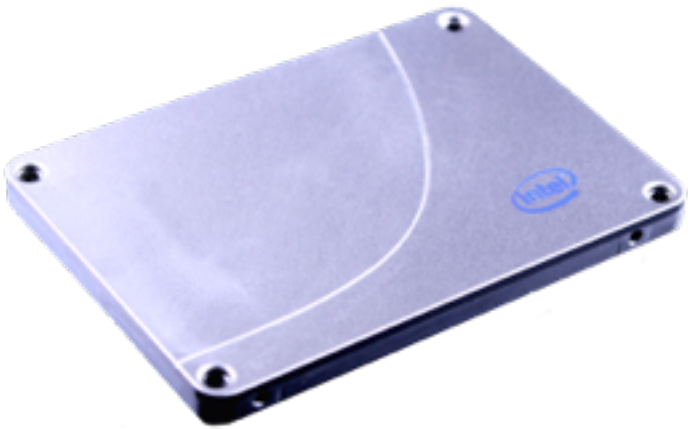
Serial ATA Technical Proposal: SATA31_TPR_Cxx
Title : DEVSLP
Sponsors: SanDisk, Samsung, Intel, Microsoft

This is an internal working document of the Serial ATA International Organization. As such, this is not a completed standard and has not been approved. The Serial ATA International Organization may modify the contents at any time. This document is made available for review and comment only.

Permission is granted to the Promoters, Contributors and Adopters of the Serial ATA International Organization to reproduce this document for the purposes of evolving the technical content for internal use only without further permission provided this notice is included. All other rights are reserved and may be covered by one or more Non Disclosure Agreements including the Serial ATA International Organization participant agreements. Any commercial or for-profit replication or republication is prohibited. Copyright © 2005-2011 Serial ATA International Organization. All rights reserved.

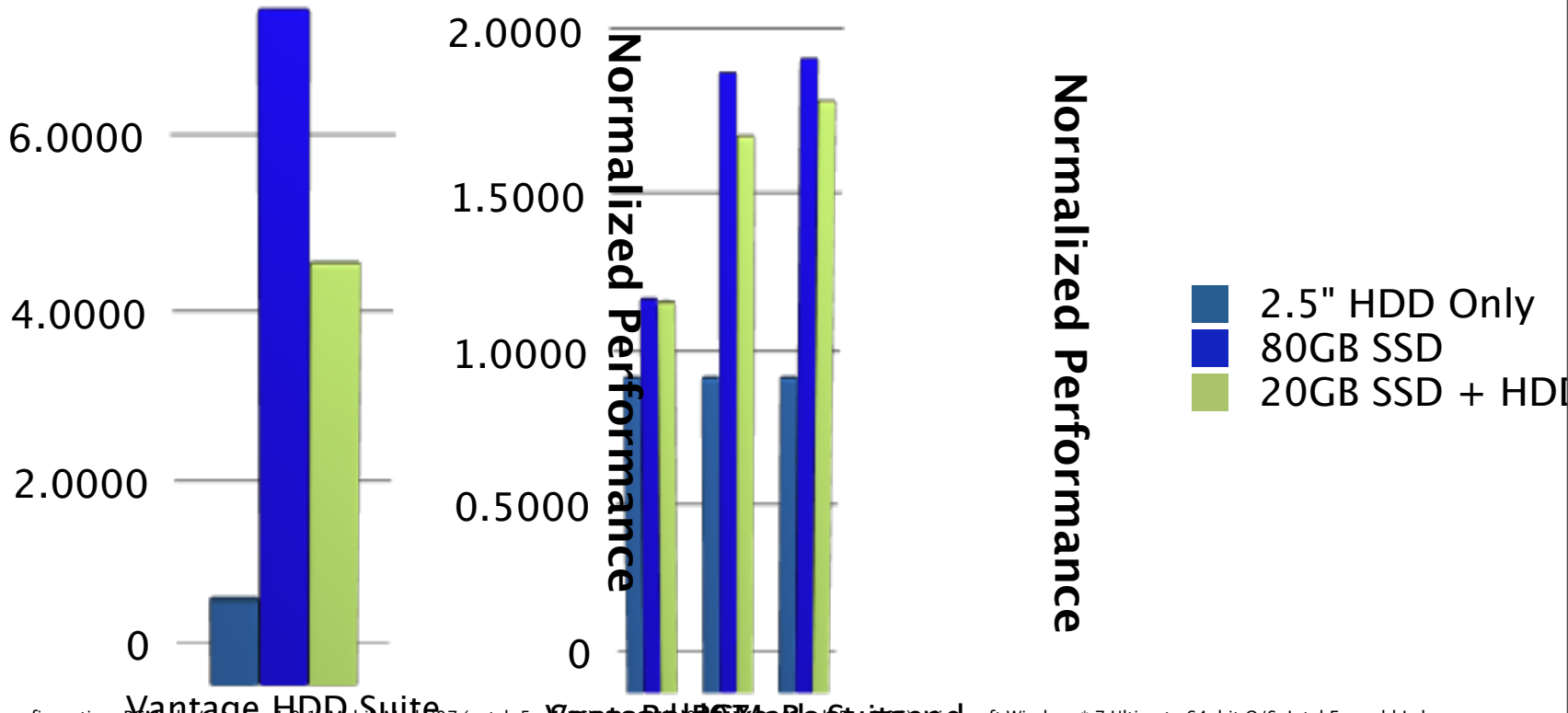
Not Every Solution Needs 100% Non-Volatile Memory (NVM)

- NVM + HDD complements NVM-only solutions
  - Get most of the SSD performance & responsiveness benefits with HDD capacities



Non-Volatile Memory + Mechanical Disk

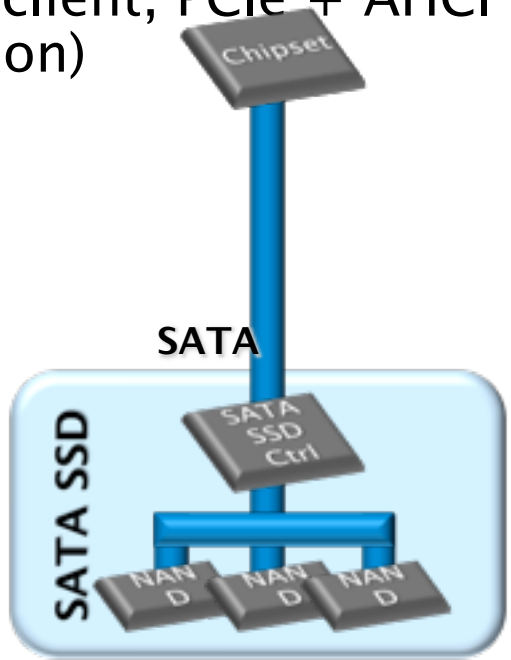
- Smart Response Technology is an Intel Rapid Storage Technology (iRST) feature found on Z68 platforms in the latest 10.5 driver



System configuration: PCMark Vantage v1.0.2, MobileMark* 07 (patch 5, w[illegible] Microsoft Windows* 7 Ultimate 64-bit O/S, Intel Emerald Lake CRB (CPT mobile), Fab 2 board, B2 silicon, CPU: Sandy Bridge D2 2.4GHz 4+2, Chipset: Intel 6 Series (Cougar Point), SATA 2 for both SSD and HDD, Hitachi 7200RPM 320GB HDD, Postville 20GB SLC, 80GB MLC SSDs, Integrated Graphics, 4GB 1066MHz DDR3 DRAM (MB).
System performance improvement on platforms is configuration-dependent; system performance claim as measured by PCMark* Vantage tests. Boot times taken with Microsoft Velocity v4.3 and Microsoft PwrTest (included in Microsoft WDK, for S4 times only).
*Other names and brands may be claimed as the property of others.

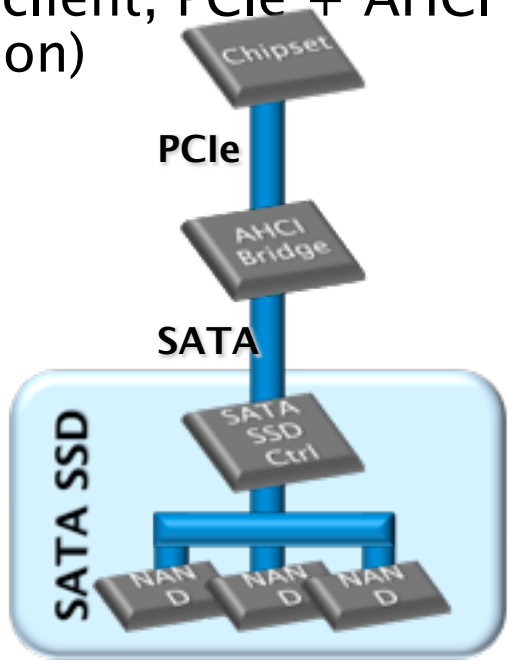
PCIe Client Interfaces

- For solutions calling for bandwidth greater than supported by SATA, PCIe is the ubiquitous scalable physical interface option
- AHCI is ubiquitous logical interface supported by client infrastructure
- For client, PCIe + AHCI is obvious option (in addition to NVMe option)



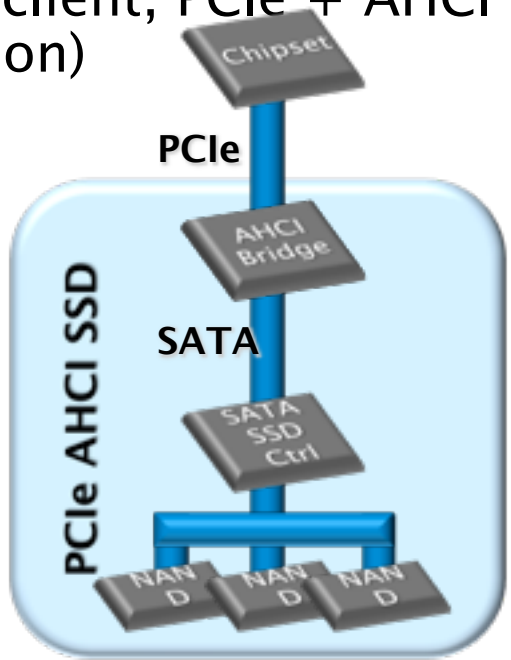
PCIe Client Interfaces

- For solutions calling for bandwidth greater than supported by SATA, PCIe is the ubiquitous scalable physical interface option
- AHCI is ubiquitous logical interface supported by client infrastructure
- For client, PCIe + AHCI is obvious option (in addition to NVMe option)



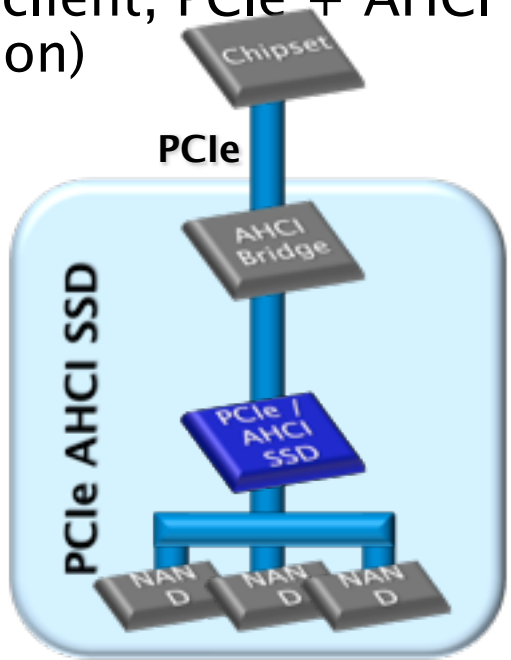
PCIe Client Interfaces

- For solutions calling for bandwidth greater than supported by SATA, PCIe is the ubiquitous scalable physical interface option
- AHCI is ubiquitous logical interface supported by client infrastructure
- For client, PCIe + AHCI is obvious option (in addition to NVMe option)



PCIe Client Interfaces

- For solutions calling for bandwidth greater than supported by SATA, PCIe is the ubiquitous scalable physical interface option
- AHCI is ubiquitous logical interface supported by client infrastructure
- For client, PCIe + AHCI is obvious option (in addition to NVMe option)



Future Outlook

- First 20X performance improvement in enterprise segments will be followed by another 20X improvement
  - Will call for efficient interfaces and storage stacks to realize
- 20X improvement in low power client idle power
  - New mechanisms required to enable
- NVM in client platforms will be 20X more prevalent
  - NVM+HDD solutions will complement NVM-only solutions
- Both physical and logical interfaces for storage will transition and host infrastructure will evolve