Client PCIe* Storage

Smaller FFs and Lower Power

John Carroll

Storage Architect

Intel Corporation

Legal Disclaimer

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.

A "Mission Critical Application" is any application in which failure of the Intel Product could result, directly or indirectly, in personal injury or death. SHOULD YOU PURCHASE OR USE INTEL'S PRODUCTS FOR ANY SUCH MISSION CRITICAL APPLICATION, YOU SHALL INDEMNIFY AND HOLD INTEL AND ITS SUBSIDIARIES, SUBCONTRACTORS AND AFFILIATES, AND THE DIRECTORS, OFFICERS, AND EMPLOYEES OF EACH, HARMLESS AGAINST ALL CLAIMS COSTS, DAMAGES, AND EXPENSES AND REASONABLE ATTORNEYS' FEES ARISING OUT OF, DIRECTLY OR INDIRECTLY, ANY CLAIM OF PRODUCT LIABILITY, PERSONAL INJURY, OR DEATH ARISING IN ANY WAY OUT OF SUCH MISSION CRITICAL APPLICATION, WHETHER OR NOT INTEL OR ITS SUBCONTRACTOR WAS NEGLIGENT IN THE DESIGN, MANUFACTURE, OR WARNING OF THE INTEL PRODUCT OR ANY OF ITS PARTS.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

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.

Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.

No computer system can provide absolute security. Requires an enabled Intel® processor, enabled chipset, firmware and/or software optimized to use the technologies. Consult your system manufacturer and/or software vendor for more information.

The products described in this document 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.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.
Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or go to: http://www.intel.com/design/literature.htm

All products, computer systems, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice.

Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

Outline

- Client Storage Form Factors
- Transitioning from SATA to PCIe*
- Performance impact from low power state entry
- Power usage from a low power PCIe SSD

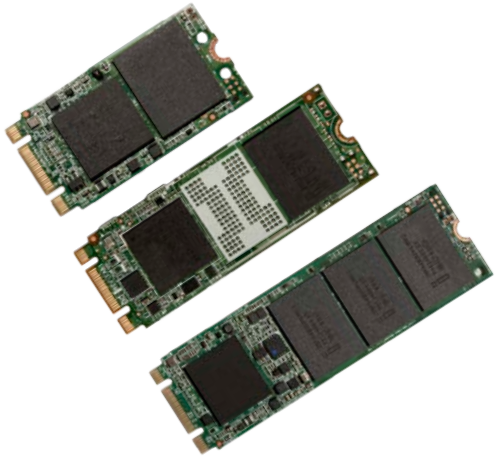
2013: PCIe* Storage Arrived in Client

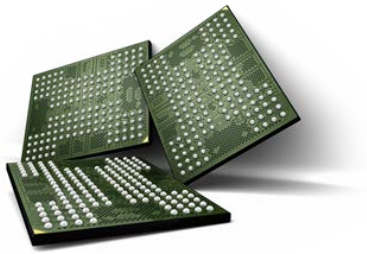
Source: FMS '13

… but device power higher than SATA

Client Storage Form Factors

- 2.5” SATA Express* connector for hosts to interchangeably support SSD, SSHD, or HDD
- M.2 is an optimized SSD only form factor
- BGA solution under definition in PCI-SIG for smallest devices

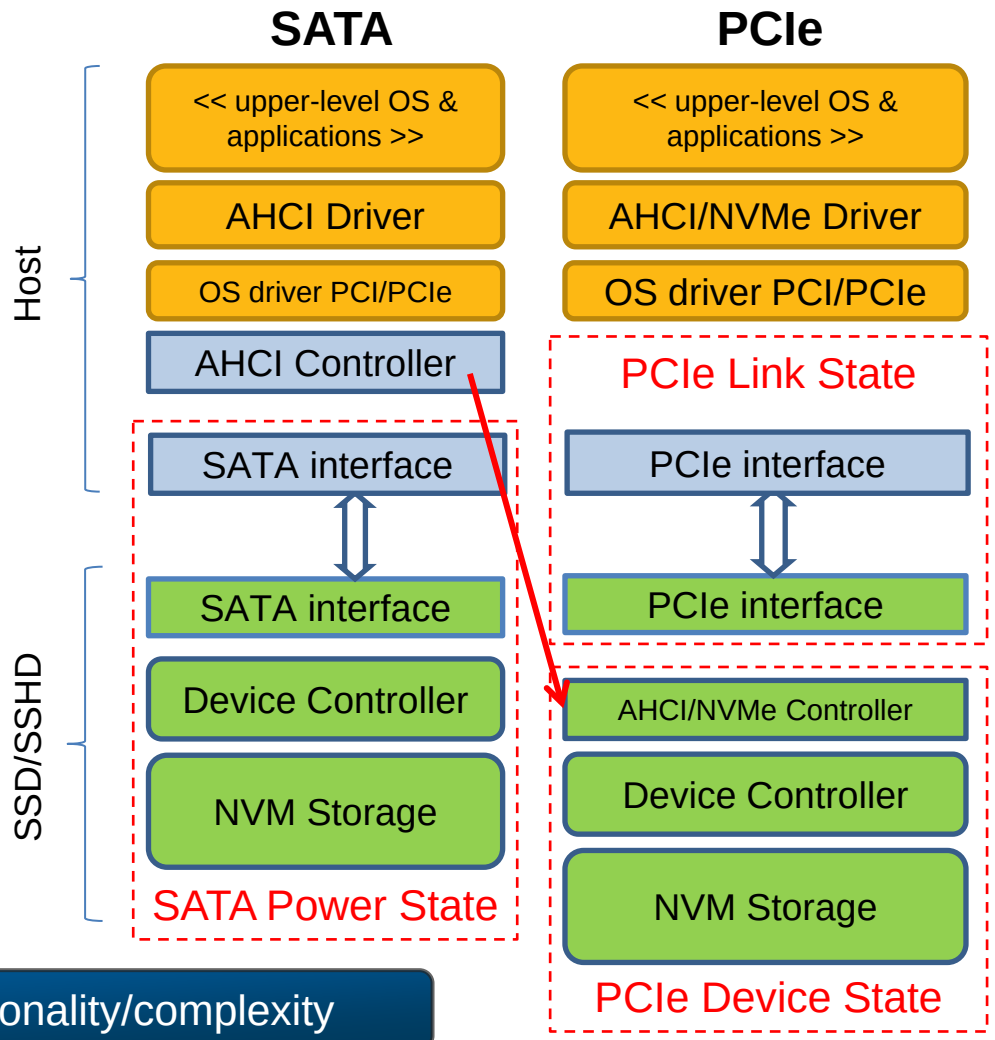




A variety of form factors available for client PCIe* storage adoption; smaller form factors have less surface area to dissipate heat

How PCIe* is Different

- Controller moves to the device
  - New functionality in the device (LTR and L1.2)
  - Power state splits into 2 independent states
- Aggressively enter low power PCIe link states to reduce PHY power
- Enter low power device states after an extended idle period



PCIe adds power management functionality/complexity

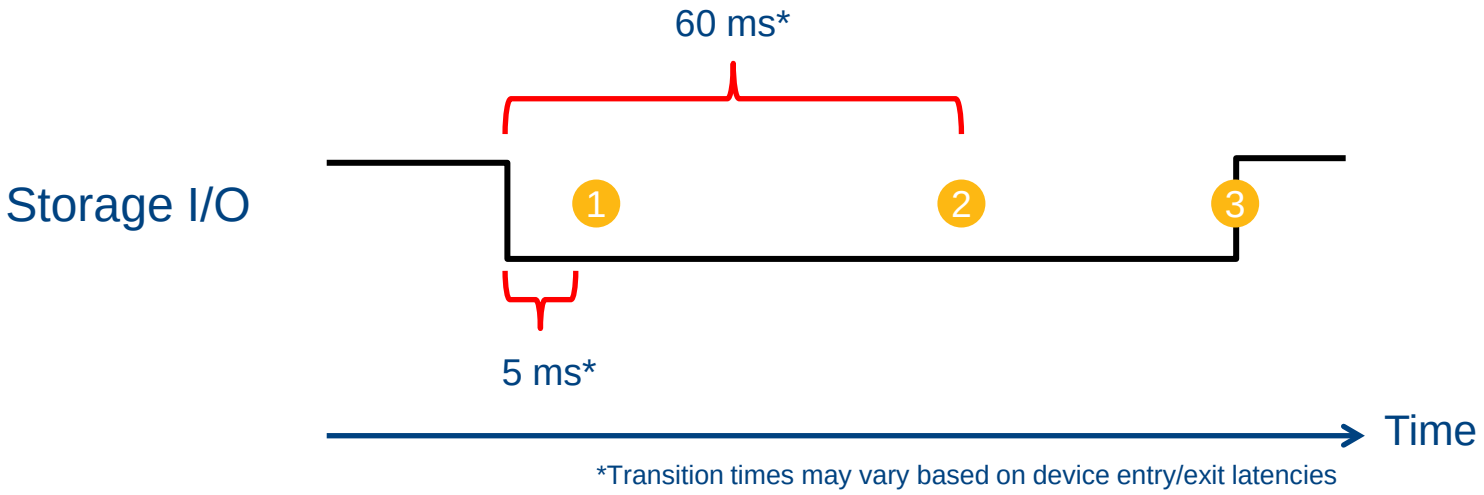
Recommended Device Power States

Power State	Description	Relative Performance	Power	Resume Latency
PS0	Default Operational	100% Performance	High	-
PS1	Light Thermal Throttle	~75% performance	Thermal workload < 2W	-
PS2	Heavy Thermal Throttle	< 50% performance	Thermal workload < 1W	-
PS3	Non-operational with fast recovery	-	Idle < 50 - 100 mW	< 1 ms
PS4	Lowest non-zero power state	-	Idle < 2.5 mW	< 50 ms

Vendors may choose to add additional power states
(e.g., additional intermediate non-operational states, or throttling levels)

Example Power State Transitions

Initial config: PCIe* link in L0 and device active

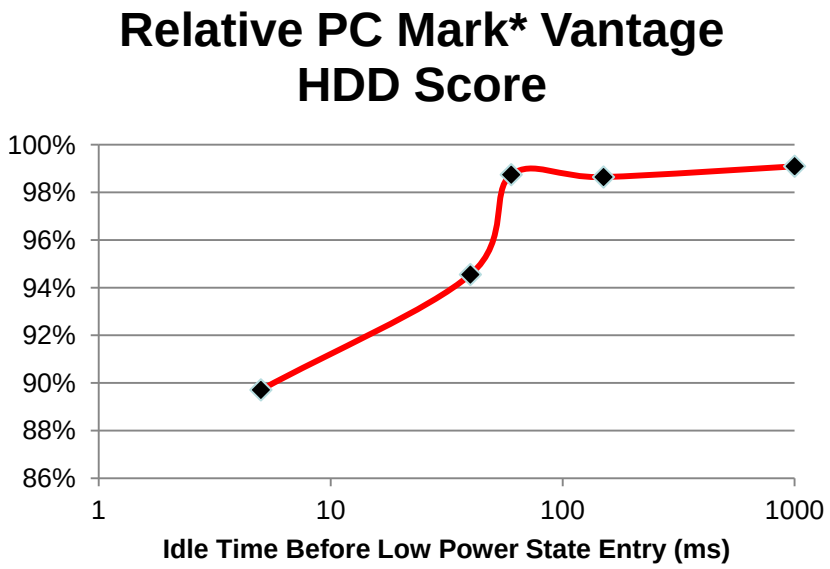


*Transition times may vary based on device entry/exit latencies

1. Controller initiates link transition to L1.2; controller stays in PS0
2. After 60ms without an I/O command, controller transitions to PS3
3. The host transitions link to L0, sends an I/O request; controller transitions back to PS0

Performance Impact of Low Power Entry

- Aggressive low power state entry negatively impacts device performance
- 10 – 15% performance loss for 5ms low power state entry, as tested in the configuration noted below
- Very aggressive autonomous power state entry also tested - showing 20-25% performance loss

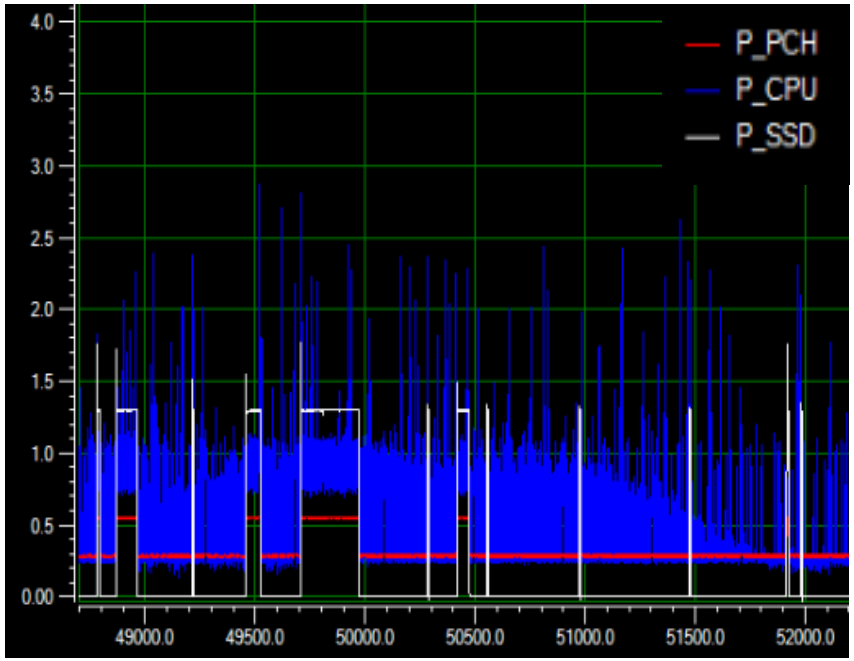


NOTE: Measurements made on next gen Intel Core * SIP platform running Windows 8.1 Standard O/S, Intel RST prototype driver, prototype SSD, data collected by PC Mark Vantage* tool. Scores averaged over 3 runs. 100% = performance on benchmark on same platform with no power management.
* Other names and brands are property of their respective owners

Measuring the Power Impact

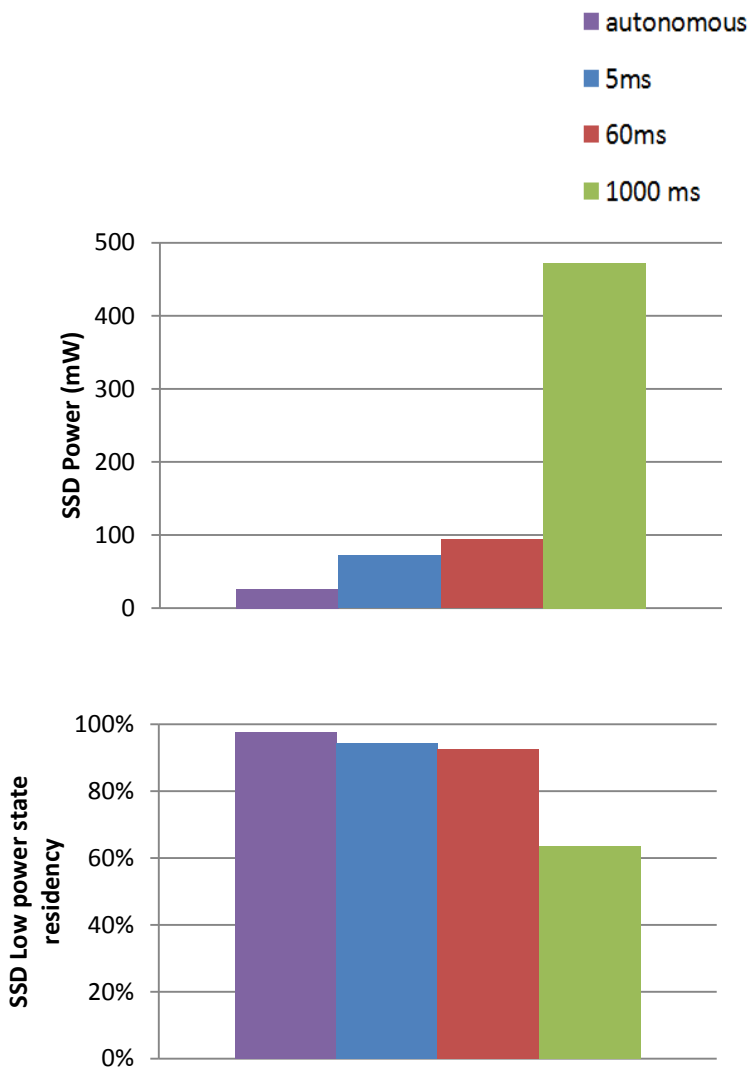
- Sample CPU, SSD, and overall platform power while running real workloads

- Focus on 2 workloads:
  - Video Playback (1080p)
  - Windows* Connected Standby



Video Playback

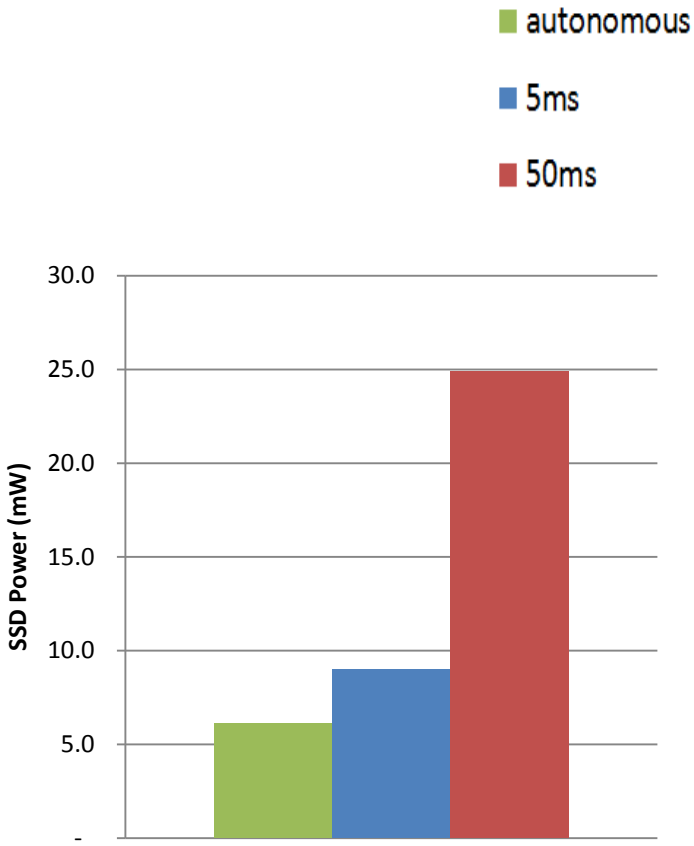
- CPU active during the workload; storage wakes up to retrieve data periodically
- Storage power on this workload similar for 5ms and 60ms entry
- Opportunity for even lower power using autonomous power transitions and aggressive RTD3



Note: Measurements made on next gen Intel® Core* SIP platform running Windows* 8.1 Standard O/S, Intel RST prototype driver, prototype SSD,. 1080p Video Playback workload description available in Ultrabook* logo documentation. Power data averaged over 3 runs.
* Other names and brands are property of their respective owners

Connected Standby

- CPU mostly idle during workload; I/O activity limited with extended idle periods
- Noticeable power difference between 5ms and 50ms entry
- Drivers may need to adjust power policies with workloads like Connected Standby



NOTE: Measurements made on next gen Intel® Core* SIP platform running Windows* 8.1 Standard O/S, Intel prototype driver, prototype SSD. Connected Standby workload description available in Ultrabook* logo documentation. Power data averaged over 3, 4-hour runs.
* Other names and brands are property of their respective owners

Summary

- Smaller client platforms reducing storage form factor size
  - Small size creates thermal challenges with high power
- With low power functionality (e.g., L1.2, LTR, and low power device states), PCIe* power comparable to SATA
- Implementing low power functionality mitigates thermal challenges while enabling higher performance PCIe storage