Interconnect Bus Extensions for Energy-Efficient Platforms

Sonesh Balchandani,
Product Marketing Manager, Intel Corporation

Jaya Jeyaseelan,
Client Platform Architect, Intel Corporation

EBLS001

Sponsors of Tomorrow.™ intel

Agenda

- Platform energy-efficiency - Overview
- Introduction to the interconnect bus extensions
- Implementation guidelines for devices using these bus extensions
  - PCI Express*(PCIe) Devices
  - USB2 Devices
  - USB3 Devices
  - SATA Devices
- Summary and next steps

Agenda

- Platform energy-efficiency - Overview
- Introduction to the interconnect bus extensions
- Implementation guidelines for devices using these bus extensions
  - PCI Express*(PCIe) Devices
  - USB2 Devices
  - USB3 Devices
  - SATA Devices
- Summary and next steps

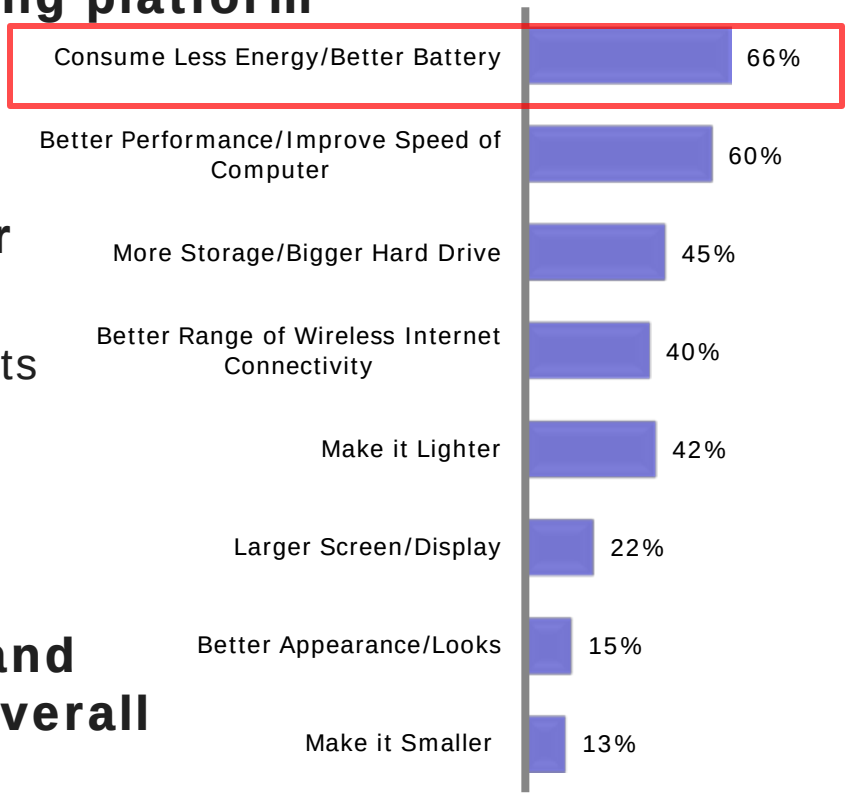
Platform Energy-Efficiency Overview

- **Intel is planning a new framework to dramatically reduce platform power**
  - Focus on reducing idle power
- **Dynamic idle power reductions benefit most common user workloads**
  - Entertainment, social networking, media, web, email, etc.
- **Maximum platform energy efficiency depends on well behaved devices and applications**
- **Intel has extensive collateral to help increase energy efficiency**

Increased energy efficiency by reducing platform idle power; enabled ecosystem makes significant contribution

Device Expectations for Improving Platform Energy-Efficiency

- **Beyond individual device power reductions, optimize device behavior for reducing platform power**
- **Intel has worked with industry groups to extend bus standards for energy-efficiency**
  - Dynamically indicate service requirements to platform as a function of workload
  - Align device traffic to platform activity whenever possible
- **Devices supporting bus extensions and following Intel guidelines improve overall platform energy-efficiency**
  - Enhances platform battery life
  - Enables smaller, thinner, compact designs



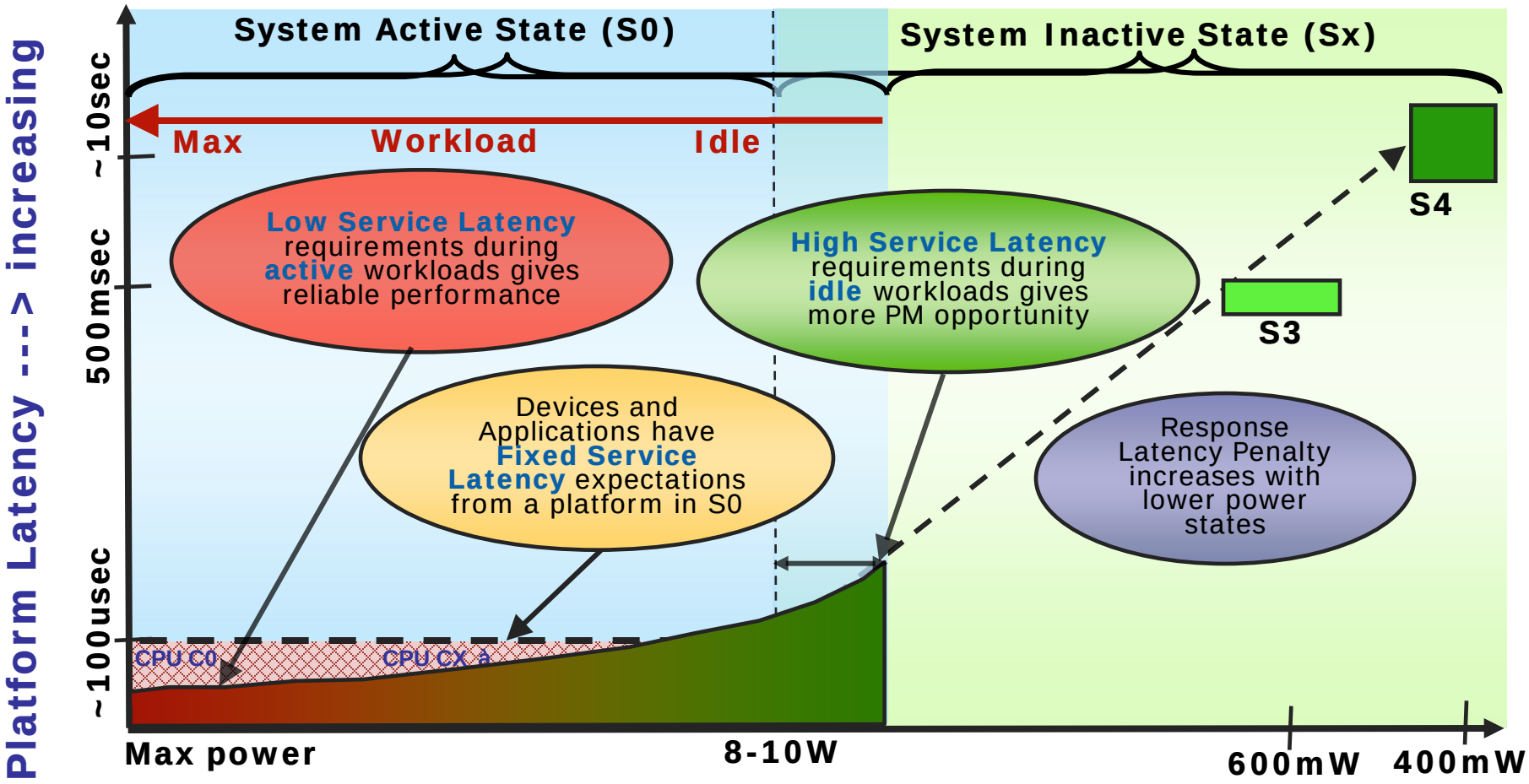
Source : Internal Worldwide Market Research, 2010

Opportunity for devices to improve platform energy-efficiency!

Agenda

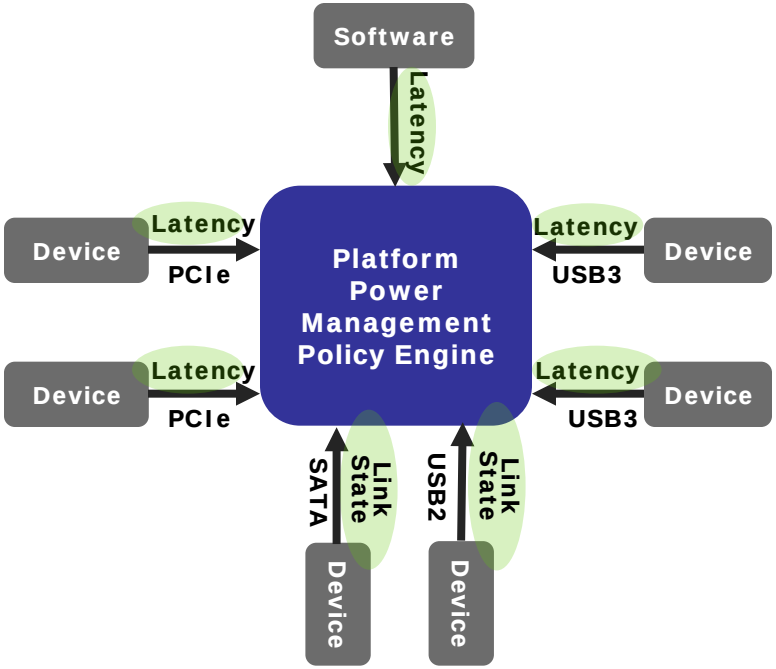
- Platform energy-efficiency - Overview
- Introduction to the interconnect bus extensions
- Implementation guidelines for devices using these bus extensions
  - PCI Express*(PCIe) Devices
  - USB2 Devices
  - USB3 Devices
  - SATA Devices
- Summary and next steps

Power vs. Response Latency (Mobile)



Variable service latency indication from devices required for aggressive, yet robust power management

Dynamic Latency Based Infrastructure



Explicit Latency Messages

- Refers to DMA access latency tolerance for reads and writes
  - Ø PCI Express* Gen2/Gen3 Latency Tolerance Reporting (LTR)
  - Ø USB3 Latency Tolerance Messaging (LTM)

Advantages

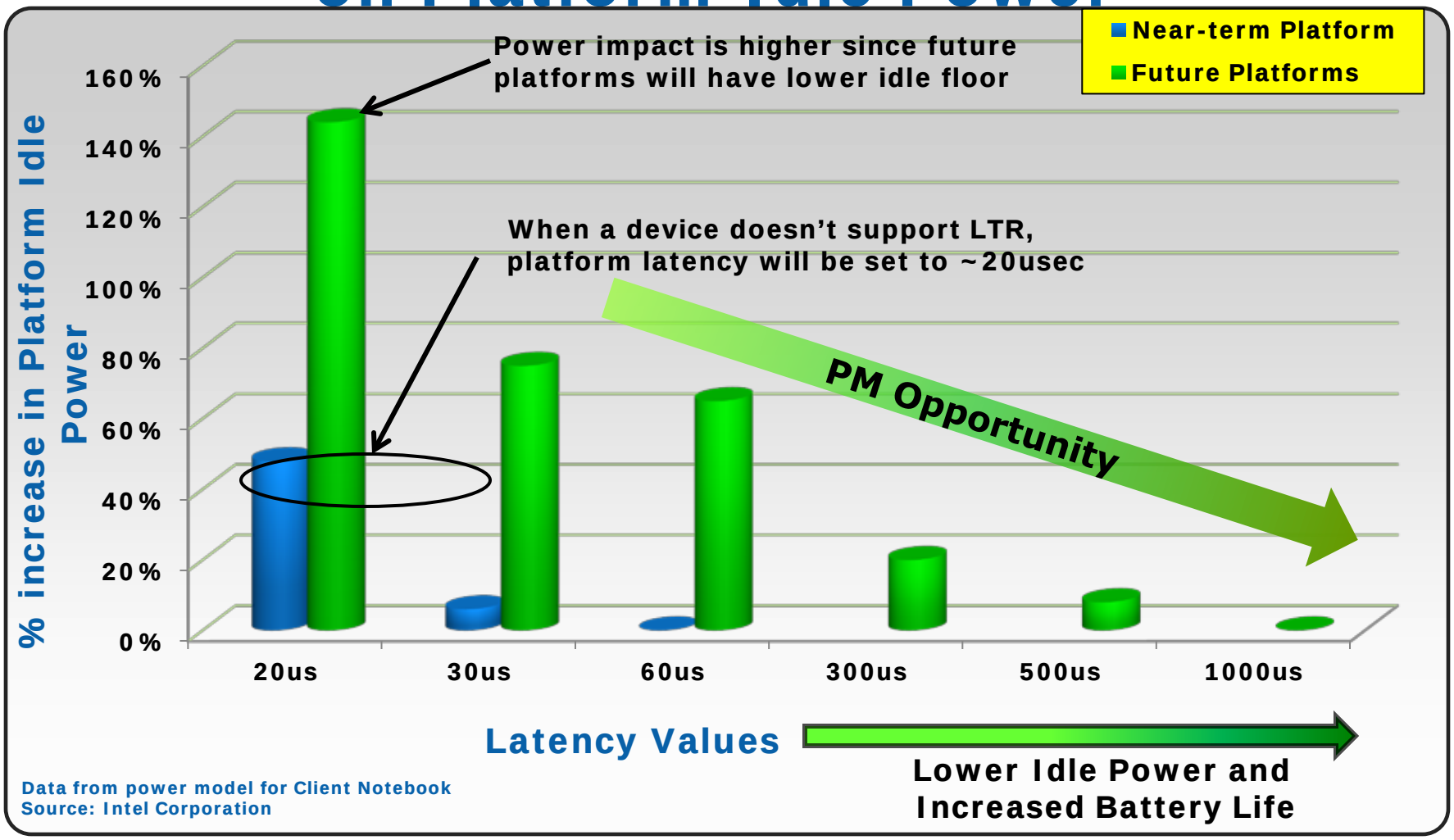
- Allows for aggressive PM without sacrificing performance or reliability
- Provides opportunity to reduce average power when workload is mostly idle

Implicit Link States

- Host controller will translate link states to latency requirements
  - Ø USB2 LPM L1 and Selective Suspend
  - Ø SATA Partial and Slumber link states

New interconnect extensions and link states dynamically convey device latency requirements to platform

Impact of Device Latency Tolerance Value on Platform Idle Power

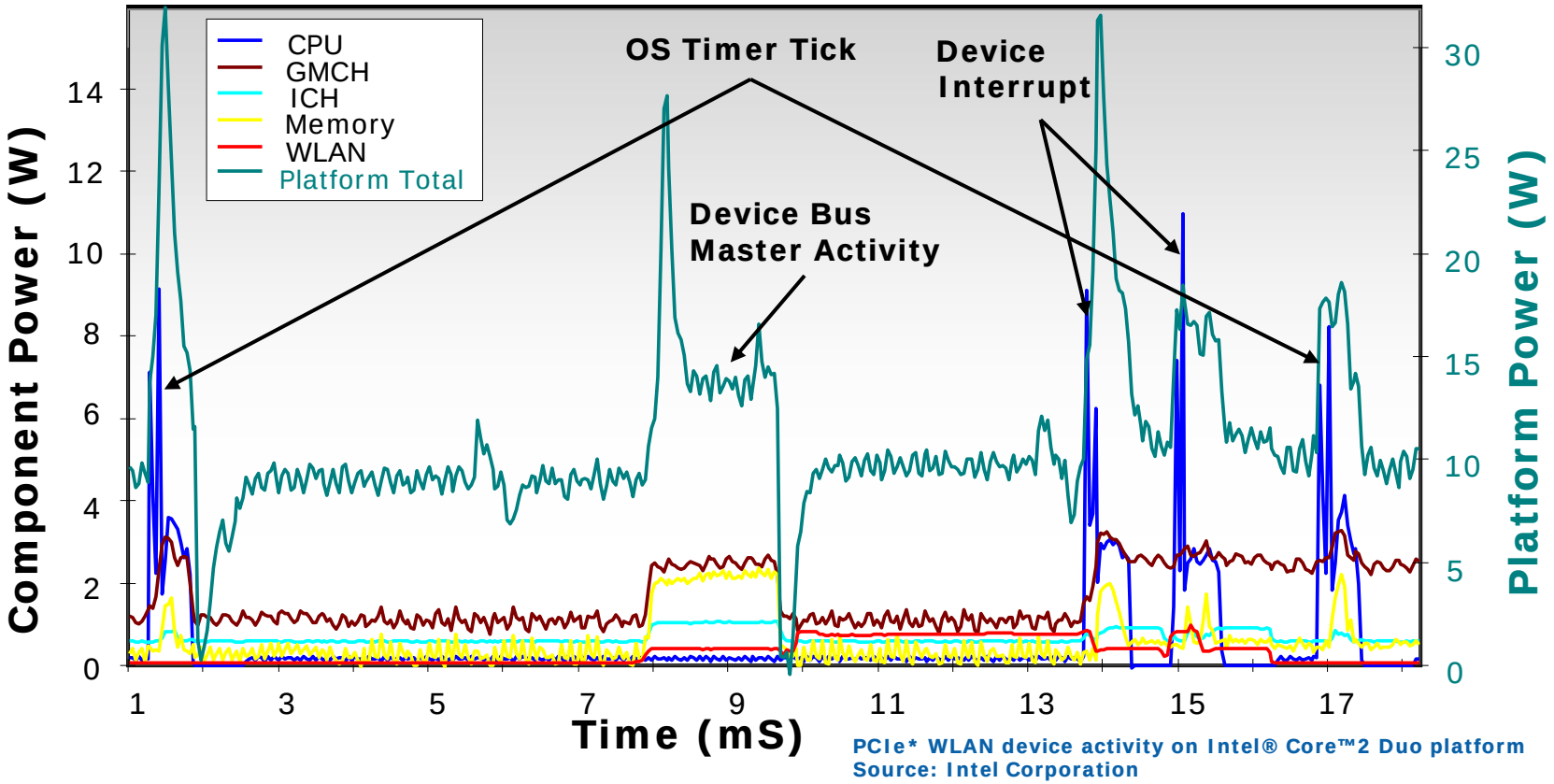


+ This data is for illustration purposes only & actual data will be available as platforms become available
+ All products, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice

Crucial that all devices indicate latency tolerance for maximum platform power savings

Power Impact of Device Activity

- **Frequent and random device activity bringing platform components out of low power states can have significant power impact**
  - Ø E.g. 100 bus master transactions per second = ~200mW



PCIe* WLAN device activity on Intel® Core™2 Duo platform
Source: Intel Corporation

Opportunity to reduce platform power by aligning device activity
Platform power savings of ~>200mW

Agenda

- Platform energy-efficiency - Overview
- Introduction to the interconnect bus extensions
- Implementation guidelines for devices using these bus extensions
  - PCI Express*(PCIe) Devices
  - USB2 Devices
  - USB3 Devices
  - SATA Devices
- Summary and next steps

LTR Recommendation for Client Devices

Latency Tolerance Reporting (LTR) Mechanism

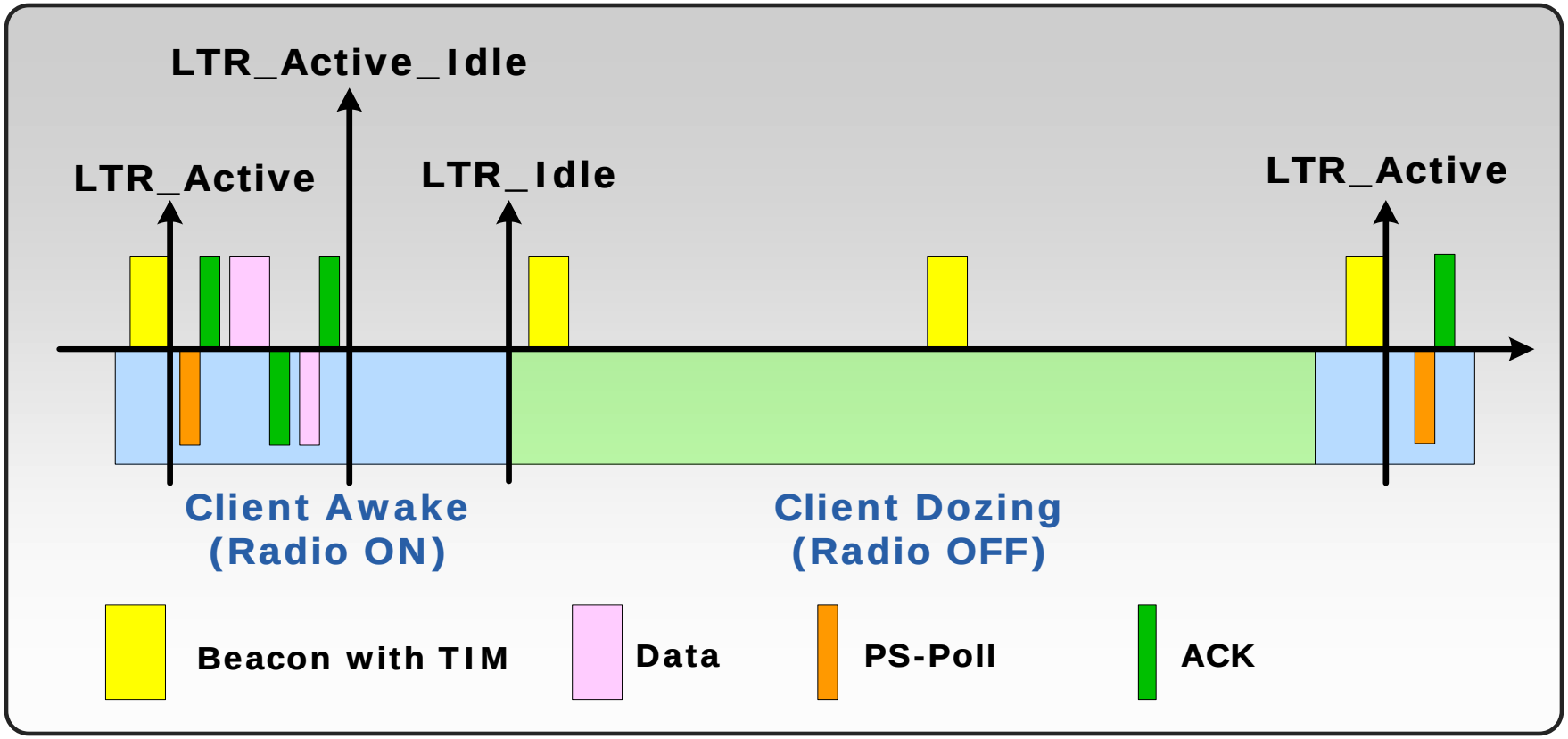
- **LTR message (TLP) sent by device dynamically as a function of workload**
  - Ø Smaller values during active workloads, larger value when idle

Devices	LTR_Active	LTR_Act_Idle	LTR_Idle	Comments
WLAN	60usec	300usec (minimum)	LTR_No_Req (unassociated) LTR_MaxPlatLat (associated and radio off)	Device Initiated
Ethernet LAN (1Gb or lower)	60usec	300usec (minimum)	LTR_No_Req (Link Disconnected) LTR_MaxPlatLat (LPI mode)	Device Initiated LPI – Low Power Idle mode in IEEE 802.3az standard
Graphics	60usec	Optional	LTR_MaxPlatLat	Can be SW guided
Client Storage (e.g. memory card reader)	60usec	Optional	LTR_MaxPlatLat	Can be SW guided

+ BIOS programs LTR Extended Capability Structure field with LTR_MaxPlatLat **(1msec)**
+ These numbers are preliminary. Monitor the following link for updates:
  http://developer.intel.com/technology/pciexpress/devnet/index.htm

Request values <60usec only when necessary–for short durations

Example: WLAN Device



Latency information with Wi-Fi Power Save

Use of device PM states to give Latency guidance

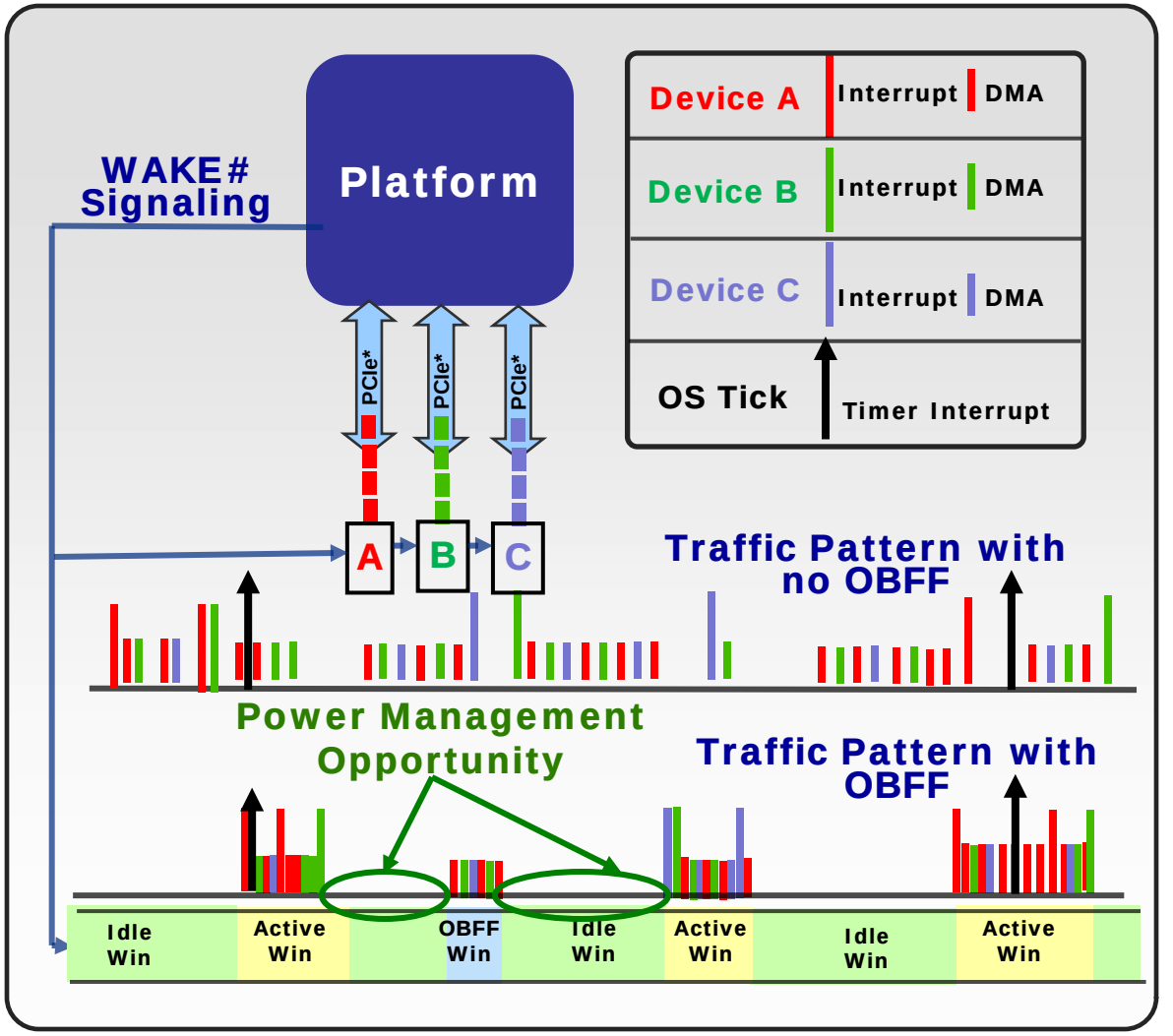
Optimized Buffer Flush / Fill (OBFF)

OBFF Mechanism

- Indicates optimal windows for bus mastering and interrupt activity
  - Ø ***Intel chipsets will drive WAKE# at the root complex for OBFF***

Optimal Windows

- **Active Window** – Platform fully active. Optimal for bus mastering and interrupts
- **OBFF Window** – Platform memory path available for memory reads and writes
- **Idle Window** – Platform is in low power state

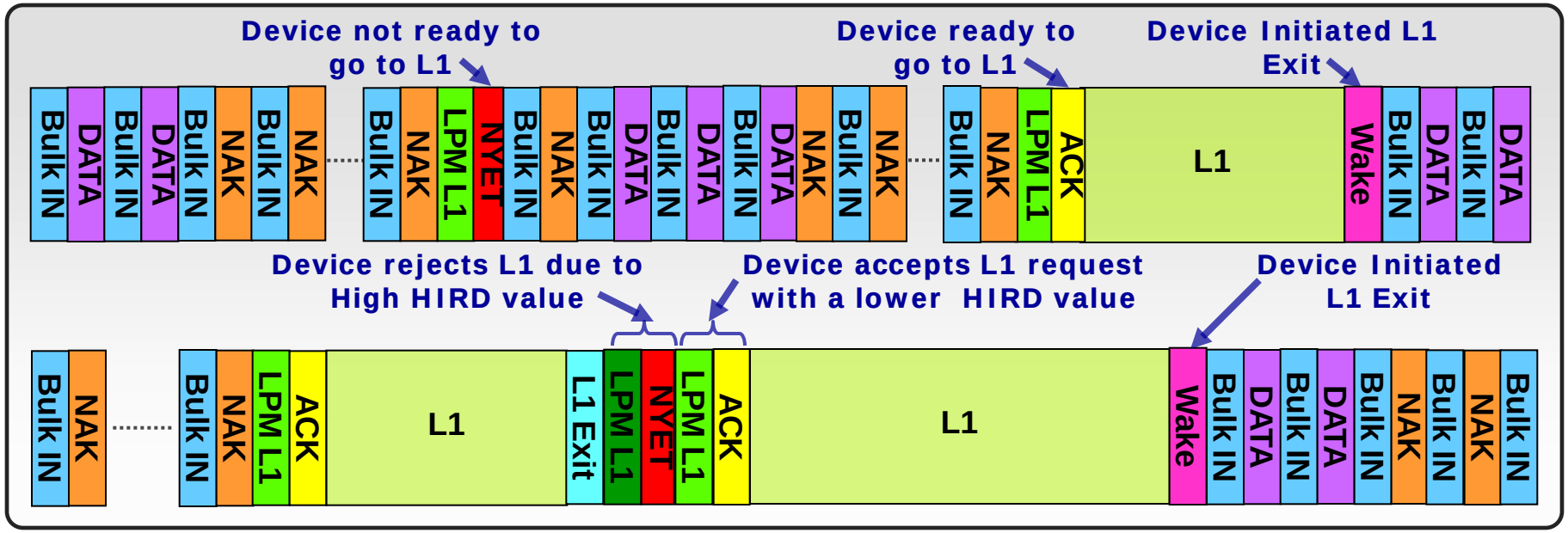


Agenda

- Platform energy-efficiency - Overview
- Introduction to the interconnect bus extensions
- Implementation guidelines for devices using these bus extensions
  - PCI Express*(PCIe) Devices
  - USB2 Devices
  - USB3 Devices
  - SATA Devices
- Summary and next steps

USB2.0 Link Power Management (LPM L1)

- **New low-latency L1 low power state intended for dynamic use**
  - Power savings at link and building-block for the energy efficiency latency infrastructure



- **Host Controller will initiate an LPM L1 transaction after some period of inactivity**
  - Device can ACK or NYET the L1 entry request based on knowledge of its activity
  - Device can also reject L1 Entry request if the value of HIRD (Host Initiated Resume Duration) is high and device requires lower exit latencies
    - Ø Larger HIRD values implies that the platform can go to deeper idle states

USB2 LPM L1 implicitly provides latency requirements

Latency Tolerance Recommendation for USB2 Devices

Devices	Active	Active_Idle	Idle	Comments
Bluetooth	L0 125usec	LPM L1 HIRD =300usec (minimum)	LPM L1, HIRD=1025usec (when connected and idle) Selective suspend (When not connected)	Support Remote-wake
3G/WLAN/ Wimax	L0 125usec	LPM L1 HIRD = 300usec (minimum)	LPM L1, HIRD=1025usec (when connected and idle) Selective suspend (When not connected)	Support Remote-wake
Mouse	L0	LPM L1 between polls (moving data)	LPM L1, HIRD=1025usec Selective suspend optional	Support Remote-wake
Storage devices (e.g. memory card reader)	L0 125usec	Optional	Selective suspend	

- **USB 3.0 xHCI-based Peripheral Development Kit (PDK) available from USB eStore**
  - http://www.usb.org/developers/ssusb/ssusbtools/xhcipdk
  - Supports USB2 LPM L1

Agenda

- Platform Energy-Efficiency - Overview
- Introduction to the Energy-Efficiency Bus extensions
- Implementation guidelines for devices using these bus extensions
  - PCI Express*(PCIe) Devices
  - USB2 Devices
  - USB3 Devices
  - SATA Devices
- Summary and Next Steps

USB3.0 Link Power Management (LPM) and Latency Tolerance Messaging (LTM)

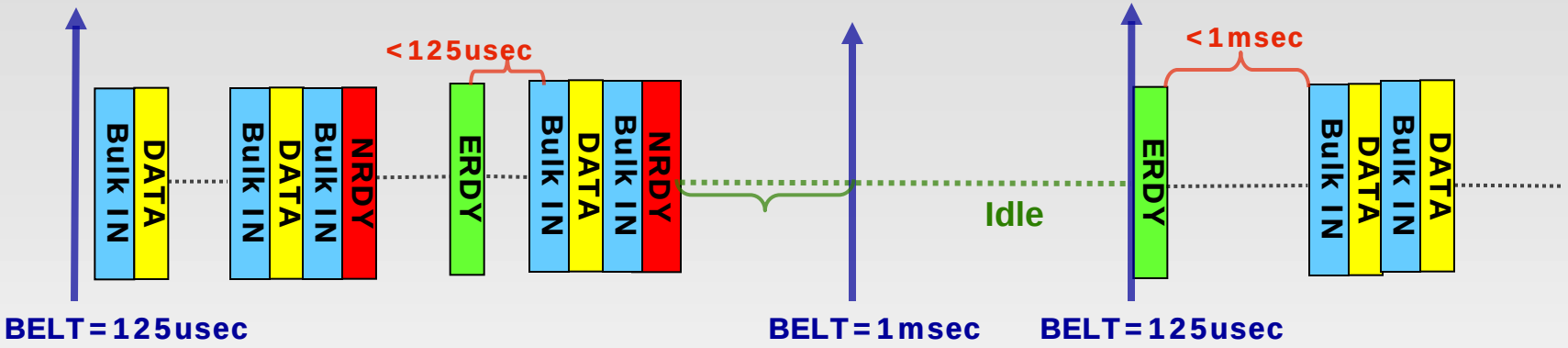
- **USB3.0 eliminates polling and supports multiple hardware driven link power states**
  - U0: Operational
  - U1: link idle with fast exit (PLL remains on)
  - U2: link idle with slow exit (PLL may be off)
  - U3: Suspend (Software driven)

- **USB3.0 defines a Device Notification Transaction Packet for the LTM scheme**
  - The Best Effort Latency Tolerance (BELT) value defines how much latency a device can tolerate from the platform

Support USB3.0 LPM and LTM for maximum power savings

Latency Tolerance Messaging (LTM)

Asynchronous Endpoints



- The BELT value is represented by the time between ERDY, and the host responding with an IN/OUT transaction associated with that ERDY
- Indicate smaller BELT value when active and larger value when idle

Agenda

- Platform energy-efficiency - Overview
- Introduction to the interconnect bus extensions
- Implementation guidelines for devices using these bus extensions
  - PCI Express*(PCIe) Devices
  - USB2 Devices
  - USB3 Devices
  - SATA Devices
- Summary and next steps

SATA Link Power Management

		Host-Initiated	Device-Initiated
Slumber Timeout	Between Commands	10msec	10msec
	Within Commands	None	Optional
Partial Timeout	Between Commands	<1usec (Immediate)	5usec (allows host to transition first)
	Within Commands	None	Entry decision made by device assuming 100usec system resume latency

- **Host is best at initiating LPM transitions between commands**
  - Transitions link to partial as soon as command completes, no timeout
- **Device is best at initiating LPM transitions within commands**
  - Knows how long the device is going to take to respond (e.g. head seek)
- **Link when in Active or Partial state will inject tighter latency requirements into platform**
  - Hold link in partial when commands are pending. Addresses any performance issues (e.g. SSD)

Host Controller will translate link state to latency requirements

Agenda

- Platform energy-efficiency - Overview
- Introduction to the interconnect bus extensions
- Implementation guidelines for devices using these bus extensions
  - PCI Express*(PCIe) Devices
  - USB2 Devices
  - USB3 Devices
  - SATA Devices
- Summary and next steps

Summary and Next Steps

- **Summary**
  - Well-behaved devices optimize platform idle power
    - Improves battery life for all client usage models
  - Every device in the ecosystem must support the bus extensions and Intel guidelines for maximum power benefit

- **Next Steps for IHVs**
  - Start architecting devices with a view towards using the new energy-efficient bus extensions
  - Work with Intel and your OEMs to understand requirements and timeline

Early implementation provides first mover advantage and opportunity to be showcased at launch

Additional sources of information on this topic:

- **Other Sessions:**
  - EBLS002: Impact of “Idle” Software on Battery Life
  - EBLS003: Mobile Platform Idle Power Optimization – Methodologies and Tools
  - PCIS002: Device guidelines for PCI Express* technology extensions

- **More web based info:**
  - http://www.intel.com/technology/mobility/notebooks.htm
    - Whitepapers under the Energy efficiency section
      - Energy-efficient platform devices
      - Designing power friendly devices
      - Designing energy efficient SATA devices
  - http://www.pci-sig.org
  - http://www.usb.org

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 PROPETY 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 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.
- Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance.
- Intel, Intel Sponsors of Tomorrow. and Intel Sponsors of Tomorrow. logo and the Intel logo are trademarks of Intel Corporation in the United States and other countries.
- *Other names and brands may be claimed as the property of others.
-

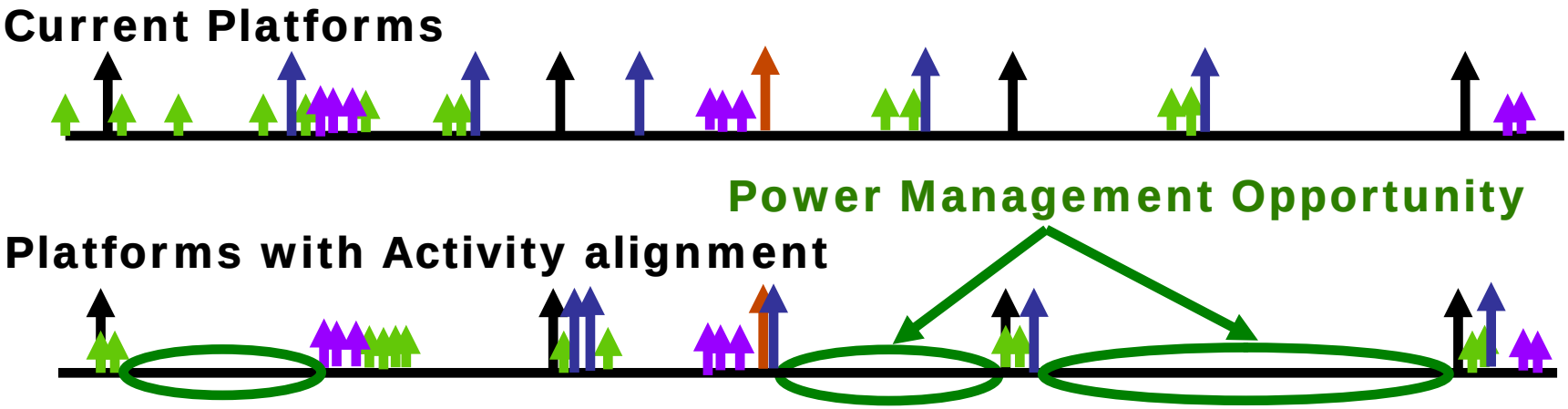
Risk Factors

The above statements and any others in this document that refer to plans and expectations for the second quarter, the year and the future are forward-looking statements that involve a number of risks and uncertainties. Many factors could affect Intel's actual results, and variances from Intel's current expectations regarding such factors could cause actual results to differ materially from those expressed in these forward-looking statements. Intel presently considers the following to be the important factors that could cause actual results to differ materially from the corporation's expectations. Demand could be different from Intel's expectations due to factors including changes in business and economic conditions; customer acceptance of Intel's and competitors' products; changes in customer order patterns including order cancellations; and changes in the level of inventory at customers. Intel operates in intensely competitive industries that are characterized by a high percentage of costs that are fixed or difficult to reduce in the short term and product demand that is highly variable and difficult to forecast. Additionally, Intel is in the process of transitioning to its next generation of products on 32nm process technology, and there could be execution issues associated with these changes, including product defects and errata along with lower than anticipated manufacturing yields. Revenue and the gross margin percentage are affected by the timing of new Intel product introductions and the demand for and market acceptance of Intel's products; actions taken by Intel's competitors, including product offerings and introductions, marketing programs and pricing pressures and Intel's response to such actions; defects or disruptions in the supply of materials or resources; and Intel's ability to respond quickly to technological developments and to incorporate new features into its products. The gross margin percentage could vary significantly from expectations based on changes in revenue levels; product mix and pricing; start-up costs, including costs associated with the new 32nm process technology; variations in inventory valuation, including variations related to the timing of qualifying products for sale; excess or obsolete inventory; manufacturing yields; changes in unit costs; impairments of long-lived assets, including manufacturing, assembly/test and intangible assets; the timing and execution of the manufacturing ramp and associated costs; and capacity utilization. Expenses, particularly certain marketing and compensation expenses, as well as restructuring and asset impairment charges, vary depending on the level of demand for Intel's products and the level of revenue and profits. The majority of our non-marketable equity investment portfolio balance is concentrated in the flash memory market segment, and declines in this market segment or changes in management's plans with respect to our investment in this market segment could result in significant impairment charges, impacting restructuring charges as well as gains/losses on equity investments and interest and other. Intel's results could be impacted by adverse economic, social, political and physical/infrastructure conditions in countries where Intel, its customers or its suppliers operate, including military conflict and other security risks, natural disasters, infrastructure disruptions, health concerns and fluctuations in currency exchange rates. Intel's results could be affected by the timing of closing of acquisitions and divestitures. Intel's results could be affected by adverse effects associated with product defects and errata (deviations from published specifications), and by litigation or regulatory matters involving intellectual property, stockholder, consumer, antitrust and other issues, such as the litigation and regulatory matters described in Intel's SEC reports. An unfavorable ruling could include monetary damages or an injunction prohibiting us from manufacturing or selling one or more products, precluding particular business practices, impacting our ability to design our products, or requiring other remedies such as compulsory licensing of intellectual property. A detailed discussion of these and other factors that could affect Intel's results is included in Intel's SEC filings, including the report on Form 10-Q for the quarter ended March 27, 2010.

Rev. 5/7/10

Backup Slides

Platform Activity Alignment



OS tick interrupt

Device interrupts (critical)

- Time Critical
- Buffer replenish
- Performance / Throughput

Device interrupts (deferrable)

- Not time critical
- Status Notifications
- User Command Completions
- Debug, Statistics

Device traffic (critical)

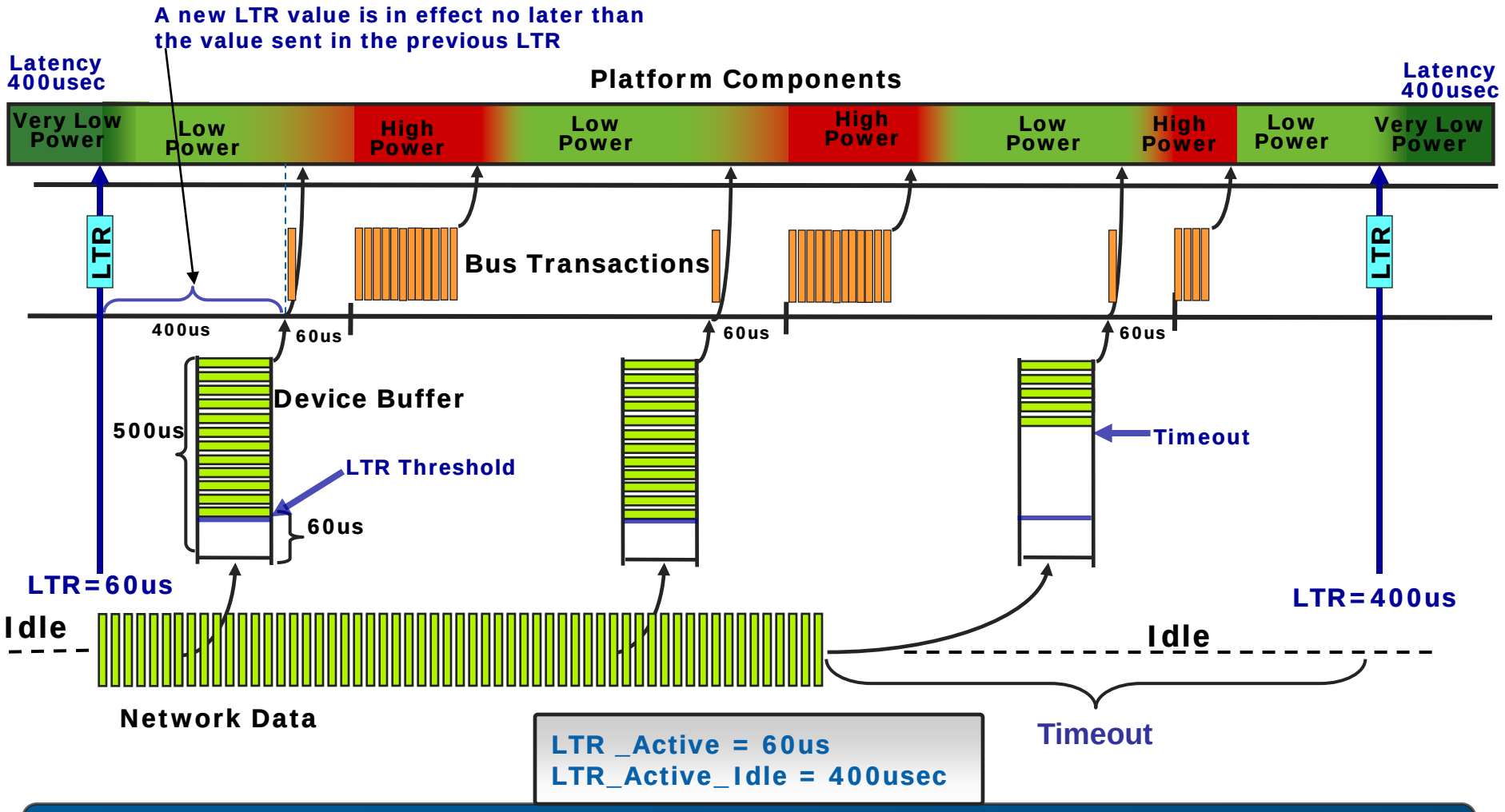
- Buffer overflow
- Throughput / Performance

Device traffic (deferrable)

- No Buffering constraints
- Debug dumps

Creates PM opportunities for semi-active workloads
Platform power savings of ~>200mW

Example: Active Ethernet NIC



Latency guidance based on buffer size and utilization