Volume 09 Issue 01 Published, February 17, 2005 ISSN 1535-864X DOI: 10.1535/itj.0901

Intel® Technology Journal

Second-Generation Intel® Centrino™ Mobile Technology

The second-generation platform built on Intel® Centrino™ mobile technology continues to drive new capabilities. This issue of Intel Technology Journal (Volume 9, Issue 1) describes new features, interfaces, and performance that enable new usages for mobile computing.

Inside you'll find the following articles:

More information, including current and past issues of Intel Technology Journal, can be found at:
http://developer.intel.com/technology/itj/index.htm

Volume 09 Issue 01 Published, February 17, 2005 ISSN 1535-864X DOI: 10.1535/itj.0901

Intel® Technology Journal

Second-Generation Intel® Centrino™ Mobile Technology

Articles

THIS PAGE INTENTIONALLY LEFT BLANK

Preface

Second-Generation Intel® Centrino™ Mobile Technology
by Lin Chao
Publisher, *Intel Technology Journal*

As we enter the ninth year of publication of the Intel Technology Journal, we bring to our readers an updated web layout and easier navigation for a better reader experience. We continue to provide pdf files for each paper and a combined pdf for the entire issue. For the first time, we are using Digital Object identifiers (DOIs), content identifiers, for web documents similar to ISSN numbers for serial print publications. DOIs provide persistent tagging of online content, addressing actual content rather than location.

This issue of Intel Technology Journal (Volume 9, Issue 1) focuses on the second-generation platform built on Intel® Centrino™ mobile technology. For many users, a laptop is their primary computer and a large part of their daily experience. This includes mobility usage models that fall into three main categories as defined by Intel: Mobile Digital Office, Mobile On-the-Go, and mobile entertainment.

Mobile Digital Office connects the IT enterprise to a wirelessly connected mobile notebook, an always-available virtual office with a primary focus of manageability and security. For example, malicious buffer-overflow attacks pose a significant security threat; the Pentium® M processor supports the XD (execute disable) bit that prevents buffer-overflow virus attacks. Another aspect of digital office is being able to connect wirelessly as easily as with a cell phone.

Mobile On-the-Go is mobility in or out of the office or home with emphasis on productivity, entertainment, and communication. Highly mobile users prefer flexible thin and light mobile systems with optimized battery life. Extended Mobile Access (EMA) technology offers a better thin and light experience with a small screen in the lid of the notebook similar to some cellular telephone models. Even with the notebook lid closed, the EMA display will show new e-mail messages, up-to-date calendars, or daily task lists. EMA technology reduces notebook power and consumption and saves battery life by turning off the main LCD.

Mobile Entertainment is optimized for entertainment–viewing movies, watching TV, sharing and editing photos, and playing games anywhere in the home. The Intel® 915 Express Chipset has a TV output feature, which is important for merging the personal computer and the television into a single platform.

The eight papers in this issue of Intel Technology Journal (Volume 9, Issue 1) examine the newest technologies for second-generation platforms built on Intel® Centrino™ mobile technology. The introductory paper provides an overview of the platform including the new Intel® Pentium® M processors, the Mobile Intel® 915 Express Chipset family (formerly codenamed Alviso), and Intel® PRO/Wireless 2915ABG or 2200BG wireless LAN components.

™ Centrino is a trademark of Intel Corporation or its subsidiaries in the United States and other countries.

® Pentium is a registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

The next four papers in this Journal look at the new features and capabilities in the platform system. The second explores the electrical packaging, manufacturing, and motherboard routing issues associated with enabling thinner, lighter laptop designs. The third looks at PCI Express*, a high-speed bit serial bus for I/O, graphics, and networking interconnections. The paper focuses on PCI Express architecture, power management, and mobile applications such as graphics, networking, and form factors including the ExpressCard* module and future form factors such as the PCI Express Wireless Card and the PCI Express Mini Card. The fourth paper considers two major enhancements to Intel's graphics memory controller hubs in the Intel® 915 Express Chipset family. First, it discusses the microarchitecture of the 3D pipeline and the steps taken to optimize it for peak performance. Secondly, the TV output feature–which is important for merging PCs and TVs into a single platform–is described. The fifth paper focuses on two key Input/Output (I/O) technologies provided by the chipset: integrated audio and the interface to the storage device (hard drive). First it examines the architectural features of Intel® High Definition (HD) Audio. Next it discusses the architectural benefits of Serial AT Attachment (SATA) which will replace Parallel AT Attachment as the interface for hard drives.

The next section is on power and thermal management. The sixth paper takes a look at performance and power consumption under common use models. In this paper we discuss some of these new features and the impact they have on platform performance and power as observed while executing industry benchmarks. The seventh examines thermal interface material (TIM) selection and improving the platform component performance/power efficiency and platform cooling capability when given finite heat budgets. By using better material for component packages, component cooling is improved.

The last paper focuses on future and current new usages for laptops. Here we explore new capabilities of the platform, and the interfaces and the wireless ecosystem used to enable new usage models: Extended Mobile Access (EMA), Voice over Internet Protocol (VoIP), Simplified Network Selection (SNS), and One Bill Roaming (OBR).

These papers reveal the exciting new mobile technologies and supporting standards already here or coming soon to our next laptop–technologies and standards that complement our multi-tasking, on-the-go computing lifestyle.

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

® Intel is a registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

Technical Reviewers

Ketan Bhat, Mobility Group
Bala Cadambi, Digital Enterprise Group
Chia-Pin Chiu, Technology and Manufacturing Group
Les Cline, Mobility Group
Peter Doyle, Mobility Group
Terry Fletcher, Mobility Group
Bill Huffman, Mobility Group
Jon Inouye, Mobility Group
Bob Jacobs, Mobility Group
Mark Jamieson, Mobility Group
Rob Kubick, Mobility Group
Animesh Mishra, Mobility Group
Ramin Neshati, Digital Enterprise Group
Himanshu Pokharna, Mobility Group
Ed Pole, Mobility Group
Krishnan Rajamani, Mobility Group
Vijay Rao, Mobility Group
Jeff Reilly, Digital Enterprise Group
Bill Sadler, Mobility Group
Allen Sampson, Corporate Technology Group
Eric Samson, Mobility Group
Brad Saunders, Mobility Group
Nirav Shah, Mobility Group
Tom Shewchuk, Mobility Group
Sanjaya Tayal, Digital Enterprise Group
Jim Trethewey, Mobility Group
Murali Veeramoney, Mobility Group
Feng (Frank) Yang, Mobility Group

THIS PAGE INTENTIONALLY LEFT BLANK

Second-Generation Intel® Centrino™ Mobile Technology Platform

Shreekant (Ticky) Thakkar, Mobility Group, Intel Corporation

Index words: CPU, GMCH, ICH, DDR2, WLAN, Audio, Power, Performance, Mobile, Notebook, Laptop, Centrino, Florence, Sonoma, Dothan, Alviso

ABSTRACT

In this paper, we provide an overview of the second-generation Intel® Centrino™ mobile technology platform. This represents a major revision of several new I/O and memory interfaces: Peripheral Component Interconnect (PCI) Express* (PCIe), Serial Advanced Technology Attachment (SATA), Intel® High Definition Audio (Intel® HD Audio), and Double Data Rate (DDR2), which all enable a range of computing and media capabilities. The second-generation Pentium® M processor, the i915 Graphics and Memory Controller Hub (GMCH), the 82801FM I/O Controller Hub (ICH) and the Intel® PRO/Wireless 2915 Network Interface Controller (NIC) with higher performance, new capabilities and interfaces enable exciting new usages: Mobile Digital Office, Mobile On-the-Go, and Mobile Entertainment. Intel has demonstrated these usages through innovative concepts based on the Intel Centrino mobile technology platform and is actively enabling our Original Equipment Manufacturers (OEMs) and Original Design Manufacturers (ODMs) to bring these to market. Intel has driven power reductions on both the platform and its silicon while still delivering new features and capabilities.

INTRODUCTION

The mobile Personal Computer (PC) market segment has entered a period of rapid innovation driven by an accelerated demand for mobile devices and advances in mobile technologies that enable new mobile usage models.

With the introduction of the first-generation Centrino mobile technology platform in 2003, the wireless LAN (WLAN) attached rate has climbed from under 20% to over 65% in 2004. According to industry analysts, IDC, more than 95% of all notebook PCs are expected to have integrated WLAN capability by 2006. This will complete the notebook transition from a portable to a wireless device.

Also, IDC expects that by the end of 2005, there will be over 150,000 WLAN hotspots around the world, nearly a 300% increase from 2003. Hotspot growth enables users to access critical data needed to make decisions, stay informed, and communicate while on-the-go. This growth in WLAN hotspots is expected to help drive a tenfold increase in the number of frequent WLAN hotspot users from less than 1 million in 2002 to over 10 million in 2004, and expanding to over 30 million by 2006.

Users view the notebook as the preferred mobile device to access wireless messaging content. Over the past few years computer mobility is part of many lifestyles: businesses, students, and home users have clearly recognized the flexibility and productivity benefits of having notebooks with wireless connectivity. Enterprise IT managers have been moving wireless deployments from prototype to production resources and enabling more notebooks with wireless access. According to industry analyst, Gartner (Dec. 2003), Mobile PC Notebook CAGR is 17% for 2002-2007.

Intel introduces the second-generation platform built on Intel Centrino mobile technology (previously codenamed Sonoma) in 2005 to continue driving growth in mobility through its vision. This platform represents a new building block with new interfaces, features, and performance that enable exciting new usages for second-generation mobile form factors. We briefly discuss technologies, features, and usage models in this paper.

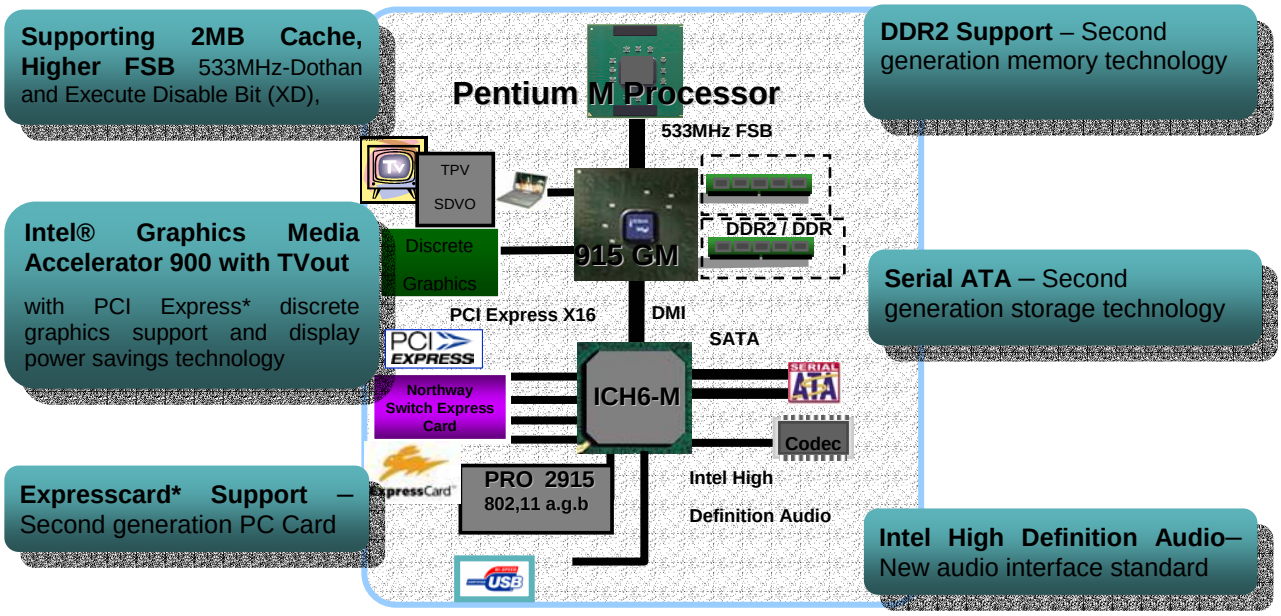


Figure 1: The second-generation platform built on Intel Centrino mobile technology

Intel's mobility vision revolves around the four vectors of mobility that drive mobile PC capabilities: breakthrough mobile performance, integrated Wireless LAN (WLAN) capability, great battery life, and thinner, lighter designs. Newer capabilities will continue to emerge to build on the foundation of these vectors and provide future users with an advanced and compelling mobile computing experience for the wireless communications environment and anywhere, anytime computing.

NEW PLATFORM

The second-generation platform built on Intel Centrino mobile technology (Figure 1) continues to excel on the four vectors of mobility. It comprises the Intel Pentium M processor, the i915 Graphics and Memory Controller Hub (GMCH), the IO Controller Hub (ICH) 6-M, and the Intel PRO/Wireless 2915 a/b/g Network Interface Controller (NIC). It introduces several new technologies that make it a new building block for this decade: mobile PCI Express (PCIe) and ExpressCard* technology, Serial Advanced Technology Attachment (SATA) interface, PCIe External Graphics (PEG) interface, Intel High Definition (HD) Audio, and Double Data Rate (DDR2) memory.

This new platform delivers both on CPU and graphics performance through scaling in core and bus frequencies as well as with microarchitecture enhancements including doubling processor cache size. The graphics performance improves by 2x over the previous-generation platform. GMCH enables an enriched media experience through incorporation of Intel HD Audio and higher-speed interfaces.

Various platform power-savings techniques are used in all Intel silicon to reduce the power consumption of the platform while still providing the new features that demand higher power. Intel has also been working with the Industry to drive lower power displays, reducing power from over 4.5 W for 14" XGA LCD to under 3 W. The display vendors have shipped over 4M such units in the past year. Thus, battery life on the second-generation platform built on Intel Centrino mobile technology stays the same as the previous generation with integrated graphics, even though graphics performance has been doubled in the second-generation platform.

Wireless connectivity is enhanced with industry-standard WLAN security support (such as WPA2 and Cisco* Compatible Extensions v.3), a new user interface for

* Other brands and names are the property of their respective owners.

* Other brands and names are the property of their respective owners.

ease of connectivity, and support of a wireless coexistence solution that mitigates interference with wireless Bluetooth[1] human interface devices (HIDs).

Smaller mobile form factors (mini and subnotebooks) are now enabled through the introduction of new smaller chipset packages.

NEW INTERFACES

PCIe introduces a scalable, point-to-point, power-managed, serial interface. Full isochronous data transfer support is provided to guarantee minimum service latency and bandwidth requirements. This mode is advantageous for isochronous or time-sensitive applications, such as streaming video, where it is more important to display frames in a timely fashion than it is to display every frame. The Mobile PCIe Mini Card and ExpressCard interfaces enable a new level of interconnect performance on mobile platforms upgrading the I/O bandwidth from a half-duplex 133 MB/s to a full-duplex 250 MB/s. The PCIe Mini Card replaces the Mini PCI card, typically used on the mobile motherboard for build-to-order optional functions. The ExpressCard module is a new add-in card that replaces the CardBus PC Card. Both of these new cards provide not only PCIe bus connectivity but also USB 2.0 connectivity through the same connector. The PCIe Mini Card (30 mm x 51 mm) is about half the size of the Mini PCI (61 mm x 51 mm) card, thus two PCIe Mini Cards can fit in the space for one Mini PCI card. The ExpressCard module comes in two sizes: One is 40% the size of the PC Card for small mobile small factors, and the other is 80% the size of the PC Card for larger mobility notebooks. The ExpressCard technology will also appear on desktops.

SATA is the evolution of the Parallel ATA (PATA) bus for connecting mass storage (hard and optical drives) on mobile platforms initially upgrading the transfer rate from 133 MB/s to 150 MB/s with ability to scale to 300 MB/s and 600 MB/s in future. SATA is also friendly to mobile form factors since it needs a much smaller connector than PATA. Also, both PCIe and SATA were designed from the ground up to be better power managed than their predecessors. Because of higher transfer rates they consume higher power during transfers. However, overall they will use less power since the work (transfer) is done quicker than it was done with previous-generation interfaces. From a power-management perspective, it is best to get work done quickly in mobile platforms and then power down the link versus taking a longer time using less power.

Finally, the second-generation platform built on Intel Centrino mobile technology introduces Intel HD Audio which delivers significant improvements over AC'97 previous-generation integrated audio and sound cards. Intel HD Audio hardware is capable of delivering the support and sound quality for up to eight channels at 192 kHz/32-bit quality, while the AC'97 specification can only support six channels at 48 kHz/20-bit. This new platform can support three external codecs with isochronous data transfers. In addition, Intel HD Audio is architected to prevent the occasional glitches or pops that other audio solutions can have by providing dedicated system bandwidth for critical audio functions. Intel HD audio offers considerable power savings (over 700 mW) for media-oriented workloads such as DVD and CD audio playback because it supports power management by allowing the processor to run in lower power states. Additionally, Intel HD Audio enables Dolby 5.1/7.1 surround sound audio out capability, which can be used with an SPDIF optical interface for component audio amplifiers.

To support PCIe, SATA, and Intel HD Audio, a newer cross-chip interconnect, Direct Media Interface (DMI), between the GMCH and the ICH, was developed. This upgraded the cross-chip transfer rate from 266 MB/s to 2 GB/s The DMI enables concurrent traffic and isochronous data transfer capabilities.

Platforms based on this new mobile technology support next-generation memory technology by implementing the DDR2 specification, an evolutionary technology that extends first-generation DDR (supported on the first Centrino mobile technology platforms). The new mobile technology platforms (i.e., GMCH) also introduce a second memory channel to system memory, effectively doubling the total available memory bandwidth especially for graphics' workloads. With dual-channel DDR2 memory support, these systems have increased peak bandwidth and lower per-SO-DIMM power consumption over first-generation Intel Centrino mobile technology platforms. The DDR2 specification allows increased clock rates over DDR while operating at 1.8 V vs. 2.5 V on earlier-generation technology. These platforms also support a peak bandwidth of 4.3 GB/s, a 60% increase over first-generation Intel Centrino mobile technology platforms, having a peak bandwidth of 2.7 GB/s (DDR 333 MHz). Together with the second channel, overall bandwidth of memory sub sub-systems will be around 8.5 GB/s.

[1] Bluetooth is a trademark owned by its proprietor and used by Intel under license.

NEW COMPONENTS

Intel's next-generation Pentium M (codename Dothan) processor implemented using Intel's 90nm technology has a number of new features. These include the L2 cache that has been doubled to 2 MB and at the same time optimized for reduced power. The Front-Side Bus (FSB) has been increased from 400 MHz to 533 MHz, and the available CPU frequencies have also been increased. Microarchitecture implementation has been optimized for performance and the Execute Disable Bit (XD) has been added to harden against buffer overflow virus attacks.

The i915 (codename Alviso) GMCH is the integrated GMCH for the second-generation platform built on Intel Centrino mobile technology delivering over 2x improvement in graphics performance over the previous 855GM controller. The i915 supports several newer and faster interfaces such as a 533 MHz FSB, two DDR2 Memory Channels, PEG (PCIe 16x) and a DMI link to ICH6-M. The integrated graphics controller includes support for Pixel Shader 2.0 and Microsoft Windows* DX9 Graphics API, integrated TVout, and improved power and performance techniques such as Display Power Savings Technique 2 (DPST2), Intel Dual Frequency Graphics (IDFG), and Frame Buffer Compression (FBC) as well as techniques to increase processor C3/4 residency, thus decreasing average power consumption. There is also use of a virtual thermal sensor to help manage thermal throttling and shutdown. GMCH comes in two packages (Intel 915GM and Intel 915GMS) supporting full and reduced functionality to accommodate different form factors.

The Intel 82801FM ICH (ICH6-M) comprises newer interfaces: PCIe, SATA with an Advanced Host Controller Interface (AHCI), and Intel HD Audio devices to connect current and the next generation of high-performance and lower-power peripherals. In addition to these new interfaces, ICH6-M provides support for USB 1.1/2.0 as on the previous-generation ICH4-M, but it adds two additional ports. The ICH6-M also provides the interface for the Intel PRO/Wireless 2915ABG wireless module and ExpressCard devices. The latter removes the need for an additional PCMCIA controller, thus saving cost while providing higher performance and form-factor-friendly connectivity.

Finally, WLAN communication is provided by the Intel PRO/Wireless 2915ABG NIC. It supports 802.11a, 802.11b, and 802.11g with the Intel Wireless Coexistence System (WCS) II that mitigates wireless interference with wireless Bluetooth devices, for e.g., headsets. This solution also supports industry-standard wireless security features including WPA2 and Cisco Compatible Extensions v3. The Intel PRO/Wireless 2915 NIC also supports the EAP-SIM protocol that allows SIM credentials to be accessed from the WWAN GSM/GPRS add-in card or add-in USB SIM reader for use in WLAN authentication and billing. This would enable carriers with the qualified WWAN add-in cards or USB SIM Reader to provide One-Bill Roaming (OBR) capability for WLAN hotspots and WWAN usage.

The increased convenience of on-demand connectivity will result in a new set of requirements to address users' needs to maintain the value of being connected while they are in transit or while the notebook lid is closed. To address these needs, Intel introduces Extended Mobile Access (EMA) capability, which enables the notebook to become a more useful tool to the user all day long.

In 2005, EMA features will include a small external display (Figure 1), similar to what has been available on some new cell phones, that provides the user with quick access to calendar, contact, and e-mail information, as well as alerts, reminders, and network availability and connection status. Running in a low-power state, the mobile notebook will provide critical information without unnecessary drain on the battery. We believe this capability will become a standard one especially with the introduction of support for auxiliary displays in the next generation of the PC client operating system.

NEW USAGES

The second-generation platform built on Intel Centrino mobile technology enables new mobile usages with the above technologies and features. We divide these usages into three key categories: Mobile Digital Office, Mobile On-the-Go, and Mobile Entertainment (Figures 2a, b, c).

* Other brands and names are the property of their respective owners.

Figure 2(a): Mobile Digital Office

Figure 2(b): Mobile On-the-Go

Figure 2(c): Mobile Entertainment

Mobile Digital Office

This usage is the evolution of the enterprise portable computer to a wirelessly connected mobile notebook. Our vision for this notebook is that it be an always-available virtual office with the following four pillars of the Digital Office:

- Embedded IT–making manageability and security transparent.
- Pervasive Connectivity–making wireless connectivity as easy as with a cell phone.
- Instant Teamwork–making collaboration spontaneous.
- Information Assistant–ready to relieve users from mundane tasks by anticipating their needs.

The second-generation platform built on Intel Centrino mobile technology takes a first step towards making this vision a reality by driving delivery on the key pillars of the Digital Office. The Pentium M processor supports the XD bit that prevents buffer overflow type virus attacks. The Intel PRO/Wireless 2915 NIC supports all security standards and offers OBR capability to make connectivity easier between WLAN hotspots and WWAN (GSM/GPRS) through the use of SIM credentials. The second-generation platform also delivers the next iteration of the Intel Stable Image Platform Program (single set of drivers) for facilitating better manageability by reducing client environment complexity, which consequently reduces total cost of ownership.

With rapid adoption of Voice-Over-IP (VoIP), the mobile PC is the platform of choice for providing access to the office phone while at home or on the road, and for voice and video conferencing. Intel HD Audio provides a better way of attaching array microphones to the platform, thus delivering a great audio experience for collaboration through VoIP calls, rivaling the $200 price tag of a

discrete microphone array used in high-end conference rooms. The CPU performance of the next-generation Pentium M processor is used for echo cancellation, ambient noise reduction, and optimization of codecs. Intel Wireless Coexistence System II in the Intel PRO/Wireless 2915ABG NIC mitigates interference resulting from overlapping harmonics of WLAN and Bluetooth radios operating in the same frequency spectrum. This enables the use of Bluetooth headsets for VoIP calls and for listening to music.

In order to inspire and demonstrate these usage models, Intel showcased the Florence Concept Family based on our new platform in the Spring of 2004 at the Intel Developers Forum. The first concept is the Florence 15" (Figure 3), and it delivers the promise of Digital Office as described with array microphones and camera for a great collaboration experience, a 15"-wide screen with a 16:9 aspect ratio for a great viewing experience, a finger print sensor for easier log in, and an EMA module for access to information on the go. This notebook also has an ExpressCard slot, a DVI connector, and a Smartcard slot for industry-standard smartcards. The system is based on the Intel Centrino mobile technology platform with a standard-voltage Pentium M processor, a 915 chipset, and an Intel PRO/Wireless 2915 NIC.

Intel is actively enabling ODMs and OEMS to make these Digital Office features part of their enterprise notebook offering starting in 2005.

Mobile On-the-Go

This usage is optimized for mobility in or out of the office or home, for productivity, entertainment, and communication. Highly mobile users prefer flexible thin and light mobile systems with optimized battery life. Flexibility allows the system to be used as a tablet or as a notebook either convertible or detachable. The system can be used to do all the work done on a Digital Office notebook while on the move. These systems can also acquire and view content such as photos, TV programs, music, or movies on a home network for consumption outside the home.

The second member of the Florence concept family is the 12" On-the-Go concept (Figure 4). This enables users to use it as a detachable tablet or as a laptop. It features array microphones, camera, finger print sensors, EMA display, and Bluetooth connectivity just like the Digital Office concept. Both the tablet and base with keyboard have lithium polymer battery packs for maximum battery life. The system also features a lower-power Low Temperature Poly Silicon (LTPS) display to enable lower average power, further enhancing the battery life of the platform. This concept is built on the Intel Centrino mobile technology platform and uses the ultra-low-voltage Pentium M processor, the 915 chipset, and Intel PRO/Wireless 2915 NIC, and it is passively cooled. It uses a 1.8 in. hard and optical drive.

Several OEMS, since the introduction of these concepts, have produced these kinds of detachable tablet/notebook systems.

Figure 3: The Florence 15”

Figure 4: The Florence 12”

Mobile Entertainment

This usage is optimized for entertainment: viewing movies, watching TV, sharing and editing photos, and playing games anywhere in the home. The Mobile Entertainment PC provides a consumer electronics experience with added PC functionality in a portable, all-in-one form factor allowing the user to bring the PC and entertainment experience to more communal areas of the home, such as the kitchen or living room. The mobile entertainment PC is also a rich communications platform allowing users to make voice or video calls and create a “virtual gathering” with friends and family far away. This platform represents the convergence of computing, communications, and entertainment.

Figure 5: The Florence 17”

The final member of the 2004 Businessweek/IDA Gold award-winning Florence concept family is a 17” Mobile Entertainment PC (Figure 5). This innovative design turns a PC into a Digital Home Consumer Device. Its main usage is entertainment and communications, but when needed is also a high-performance wireless portable computer. The industrial design of this system reflects this usage; it looks like a 17” LCD TV screen with a stand, but hidden behind the stand are a wireless Bluetooth keyboard, remote control, and VoIP handset. These peripherals can be detached from the base and moved 6-10 feet away for viewing. The system is based on the second-generation Intel Centrino mobile technology platform with a standard-voltage Pentium M processor, a 915 chipset, and the Intel PRO/Wireless 2915 NIC. It delivers a HDTV-quality video experience and a good gaming experience with 915 graphics performance. Additionally, it incorporates array microphones and a camera, just like the two other members of the Florence family, to deliver a great virtual gathering experience with family and friends. The latter is now possible with broadband services delivered to a large percentage of homes. The Florence 17” screen is also a wireless TV (without a tuner) that receives TV broadcast over Intel PRO/Wireless 2915 NIC from a wireless TV receiver connected to a cable or satellite source. Such devices are now sold by Sony and Sharp to work with wireless TVs operating over IEEE 802.11a/b/g WLAN. The Florence family also uses Intel HD Audio to provide Dolby 5.1/7.1 surround sound audio with an SPDIF optical interface that can be interfaced with a component audio amplifier.

Intel is enabling these kinds of consumer systems for introduction in 2005. We expect this trend to continue with larger screen sizes. Our vision is to see every TV integrated with Intel Centrino mobile technology to provide a rich entertainment, communication, and computing experience.

CONCLUSION

The second-generation platform built on Intel Centrino mobile technology excels on the four vectors of mobility delivering more performance through microarchitecture feature enhancements, 90nm process technology, and new interfaces while keeping the battery life of the platform the same as that of previous-generation technologies. These new interfaces, performance, and capabilities unleash exciting Mobile Digital Office, Mobile On-the-Go, and Mobile Entertainment usage models. Intel demonstrated these usage models through innovative concept platforms and are continuing to enable our OEMs and ODMs to include these capabilities in the next generation of mobile platforms.

ACKNOWLEDGMENTS

We acknowledge all the members of the second-generation Centrino mobile technology platform and the Mobile Concept team for their dedication in making this platform happen. And special thanks to Tom Shewchuk for his dedication and help in guiding the development of this issue of Intel Technology Journal**.** And an additional acknowledgment to Judith C. Anthony for coordinating a seamless issue.

AUTHOR’S BIOGRAPHY

Dr. Shreekant (Ticky) Thakkar is the director of mobile technology and is the Mobile visionary in Intel’s Mobility Group. He has over 26 years of experience in various development and planning positions at Intel and Sequent (now part of IBM). Dr. Thakkar was a key driver in establishing the direction for mobile notebook PCs to transition from portable to wireless computing with Intel’s Centrino mobile technology. Recently, he drove the convergence of computing, communications, and entertainment in the Mobility Group’s Florence Concept Platform. Dr. Thakkar also created the Mobile Platform Architecture team within the Mobility Group. Prior to the Mobility Group, he was the general manager of a new business unit, Persona, which delivered smart proactive services over wired and wireless devices. At Intel, he also led the team that developed the Pentium Pro MP (Intel®

Xeon™ processor) as well as teams that developed Multimedia/Graphics Media (SSE) extensions to the Pentium III and 4 processors. Among Dr. Thakkar's many accomplishments was bringing the Pentium Pro MP (Intel Xeon) processor to market in record time, 11 months from first silicon to production. He has led the development of security functionality in Intel's processors and chipsets and fills the roadmap planning role for Intel's microprocessor family. Prior to Intel, Dr. Thakkar pioneered the development of Shared Memory Multiprocessors and Databases on these systems at Sequent Computer Systems

Dr. Thakkar holds a Ph.D. degree, an M.S. degree in Computer Science/Engineering from the University of Manchester (England), and a B.S. degree in Computer Science/Statistics from the University of London. He holds or has pending applications for over 50 patents. Dr. Thakkar has published numerous articles and edited special editions of IEEE journals. His e-mail is ticky.thakker at intel.com.

™ Xeon is a trademark of Intel Corporation or its subsidiaries in the United States and other countries.

Intel® 915GMS Chipset: In Mobile Platforms, Smaller is Better

David W. Browning, Mobility Group, Intel Corporation
Vinay Kumar C., Mobility Group, Intel Corporation
Praveen G., Mobility Group, Intel Corporation
Ananth V. Gopal, Digital Enterprise Group, Intel Corporation
Tom Shewchuk, Mobility Group, Intel Corporation
Suresh V. Subramanyam, Mobility Group, Intel Corporation

Index words: Intel 915GMS chipset, package, manufacturing, motherboard, routing

ABSTRACT

The Intel® Centrino™ mobile technology is based on four basic vectors of mobility: breakthrough mobile performance, integrated Wireless LAN (WLAN) capability, great battery life, and thinner, lighter designs. Historically, Intel's focus has been on increasing performance and, to a lesser extent, on improving power consumption. Intel Centrino mobile technology sharpens this focus and also allows Intel to increase the feature set of the platform to include wireless solutions. In this paper, however, we focus on the mostly overlooked promise of the fourth vector of mobility: thinner, lighter designs.

The packaging, manufacturing, and motherboard routing issues associated with enabling thinner lighter designs are explored in detail. We discuss the package electrical and motherboard breakout challenges, exploring the ideas and concepts used to ensure that reducing the package size will allow for high-volume manufacturing requirements. Finally, we discuss how Intel enables unique form factors through the support of specialized motherboard routing techniques. All of these solutions allow Intel Centrino mobile technology to provide a very convincing answer to the challenge of innovating compelling form factors.

INTRODUCTION

The Intel Centrino mobile technology was specifically developed for on-the-go computing, with a focus on the four vectors of mobility: breakthrough mobile performance, integrated Wireless LAN (WLAN) capability, great battery life, and thinner, lighter designs. When Centrino mobile technology platforms were introduced in March of 2003, notebooks became much smaller (X and Y) and much thinner (Z) compared to previous notebooks based on the Intel Pentium® 4 processor. This reduction was primarily due to the significantly lower Thermal Design Power (TDP) of the processor and the Graphics and Memory Controller Hub (GMCH) components compared to previous platforms.

However, even smaller form factors could not be developed as the size of the processor and GMCH components was still quite large; in fact these components were the largest two components on a mobile motherboard. For example, the Intel 855GM GMCH package was 37.5 mm square, even though the silicon die size was only 9 mm square. This was due to the very large number of pins on the package together with the board assembly technology required to get all these signals routed onto the motherboard.

As Intel started to develop the next-generation mobile PC platform built on Intel Centrino mobile technology, the GMCH package size actually grew by 7% to 37.5 mm x 40 mm. This growth was primarily due to the addition of a second channel of Double Data Rate (DDR2) system

memory, that was required for the highest performance mobile notebooks. In order to continue to drive the thinner, lighter designs vector of mobility, Intel wanted to develop a significantly smaller GMCH package, specifically designed for the smallest (X and Y) and thinnest (Z) systems. Thus the idea of the Intel 915GMS chipset was born.

This paper explains the steps Intel took to design this new product. We first explain how the number of signal and power pins was reduced on this product, and how the package ball pitch was optimized while maintaining effective board routing. Second, we address the manufacturing concerns that result from a significantly smaller package. The result of our work was the creation of a 27 mm square Intel 915GMS package, a 51% size reduction from the standard Intel 915 package.

Third, we explain additional board-routing guidelines that were developed to allow mobile Original Equipment Manufacturers (OEMs) to further compress their motherboard layout. It is hard to know what new form factors will result from this, but OEMs are very creative when it comes to smaller and sleeker notebooks. We end with some ideas on what new usage models might result from this work, and we discuss future challenges for the continued push toward compelling form factors.

INTEL 915GMS BREAKOUT ROUTING CHALLENGES AND SOLUTIONS

The standard Intel 915 chipset was developed in a 37.5 mm x 40 mm package. Many techniques were used to reduce the size of the Intel 915GMS package while still meeting the needs of the small form factor market.

Interface Removal

One of the easiest ways to reduce the size of a package is to remove unnecessary features. A detailed investigation of the small form factor market led to an understanding of what subset of features of the Intel 915 chipset were needed. As seen in Table 1, ~300 pins were removed on four key Intel 915GMS interfaces.

Table 1: Pin-reduction summary

Interface	915GM	915GMS	# Removed
PCIe to Ext.GFX	142	32	110
DDR to Memory	383	227	156
DMI to ICH	35	18	17
LVDS to Panel	41	25	16

During this analysis we found that the small form factor segment does not typically have room to accommodate an external graphics controller and therefore is primarily designed for internal graphics. Since the Intel 915 chipset family supports both internal and external graphics configurations, external graphics support was removed from the Intel 915GMS product in order to eliminate the majority of the Peripheral Component Interconnect (PCI) Express* interface and its associated signal pins. This allowed us to remove 50 signal pins and 60 ground pins.

Another thing we found was that most of the small form factor systems only require a single channel of memory support due to their reduced performance needs. Since the Intel 915 chipset family supports both dual-channel and single-channel system memory, one of the two channels was removed to save ~120 signal pins. This also allowed approximately 30 power pins to be removed since fewer power pins were needed to support a single-channel memory configuration.

In addition to the reduced memory performance requirements, it was found that the small form factor systems also have reduced I/O performance requirements. The Intel 915 chipset family supports both a 4-lane Direct Media Interface (DMI) and a 2-lane DMI between the GMCH and the I/O Controller Hub (ICH). The reduced I/O performance requirement of the small form factor segment allowed us to alter the DMI bus for this chipset to only support a 2-lane configuration, and we thus were able to remove 17 more pins.

Finally, we found that most of the small form factor LCDs selected by OEMs and Original Design Manufacturers (ODMs) utilize a single channel of Low Voltage Differential Signaling (LVDS). This finding allowed the design team to remove one of the LVDS channels when developing the Intel 915GMS chipset to save another 16 pins.

The removal of these pins associated with the unnecessary interfaces allowed for a major reduction in package size, resulting in a 31 mm x 31 mm package. We then moved on to investigate making pin pitch changes to the pinout in order to shrink the package size even further.

Parquet Package Technique

The standard Intel 915 chipset product utilizes a uniform grid pin pitch of 42 mils. This pin pitch was selected to allow two traces to route between the via field in the breakout region, assuming a via antipad diameter of 30 mils (a via size typically used in mobile high-volume manufacturing). After the vias' antipad area is accounted for there is only 12 mils of remaining room for trace routing. As illustrated in Figure 1, utilizing a standard

* Other brands and names are the property of their respective owners.

stripline trace width of 4 mils allows for two traces to fit between the vias with 4 mils of space between them (hereafter referred to as 4:4 routing).

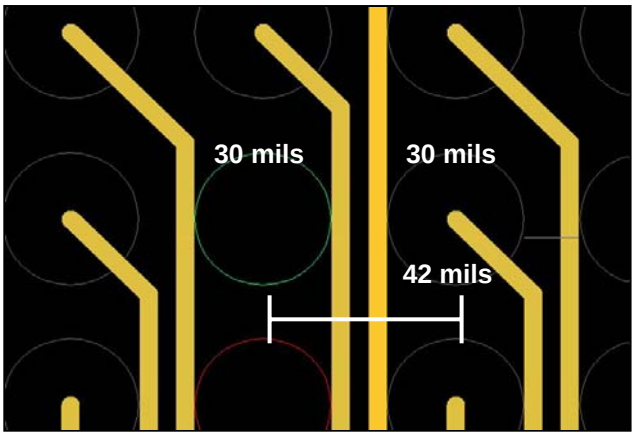


Figure 1: Breakout region routing dictated by via size

As can be seen, the minimum allowable pin pitch is directly related to the via size that is assumed for the platform. For the majority of mobile designs the standard via uses an antipad size of 30 mils so the resulting pin pitch is 42 mils (to allow for two tracks of 4:4 routing). However, the analysis of the small form factor market segment showed an inclination to utilize slightly more expensive motherboard technologies. This analysis showed that a 28 mil antipad was acceptable in this market and would allow for some pin compression. Therefore, the Intel 915GMS pin pitch was reduced from 42 mils to 40 mils (assuming a via with a 28 mil antipad).

Upon closer inspection of the breakout routing of the standard Intel 915 chipset package, it was determined that the perpendicular (or "transverse") routing channels were not being utilized. In effect, all of the signals were breaking out in a radial fashion away from the center of the package. This observation led to the idea of reducing the perpendicular pin pitch to something smaller than 40 mils to "compress" the package even further. In order to achieve an equal amount of compression along all four package sides, the parquet technique was developed.

The parquet package technique is composed of four periphery sections and a center region. Each of the periphery sections utilizes a wider pin pitch in the direction along the package edge, and they utilize a reduced pin pitch in the direction towards the package center (see Figure 2). The result is a 40 x 32 mil parquet pin pitch.

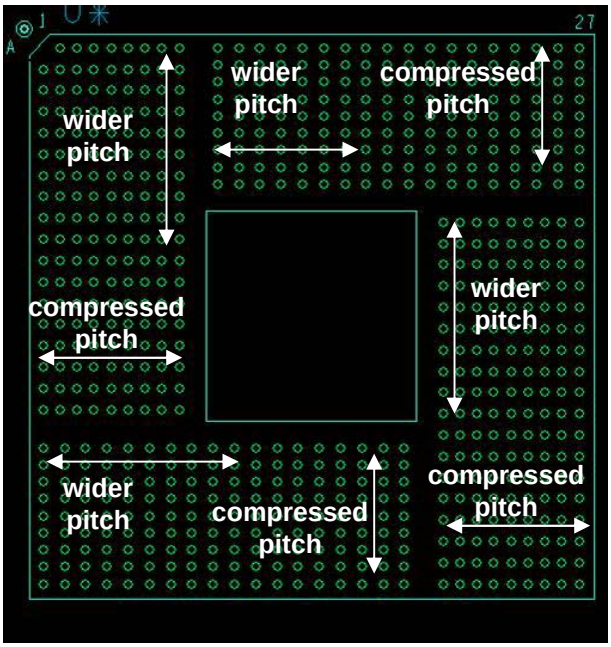


Figure 2: Parquet package illustration

The parquet package routing technique is not without challenges of its own. In particular, since one of the pin pitches is reduced to 32 mils, there is only enough room to route one signal in a perpendicular fashion to the typical breakout direction. This in itself presents no problem, because as the signal density increases so can the complexity of trying to find a breakout routing solution for each of the signals (see Figure 3).

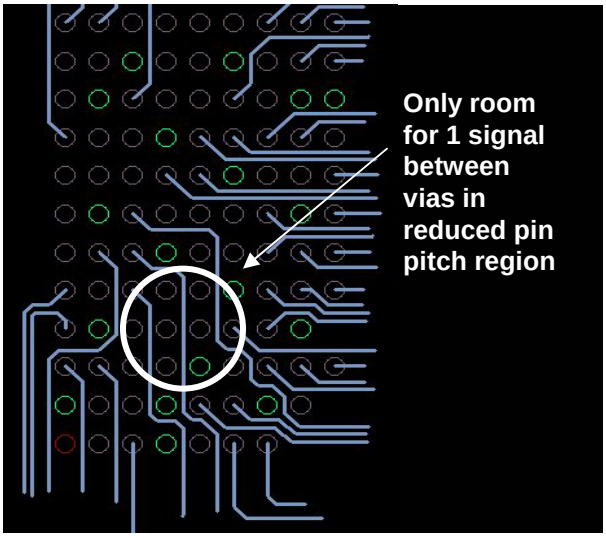


Figure 3: Transverse parquet routing

Corridor Routing

An additional benefit of the parquet package technique is that it provides natural "gaps" or "corridors" where two different parquet periphery sections meet. These corridors

have a wider pin pitch and therefore can be used to route more signals, or to provide a wider gap for power delivery (see gaps between periphery sections in Figure 2).

In addition to the natural corridors, the Intel 915GMS pinmap creates some artificial corridors with power pin assignments. The pinmap assigns power pins in corridor shapes from the edge of the package towards the center region. These "power corridors" have power delivery benefits since they allow a wider copper flood on the surface layer to connect to the package and therefore can accommodate higher currents and provide a lower DC-power loss routing path for power delivery (see Figure 4).

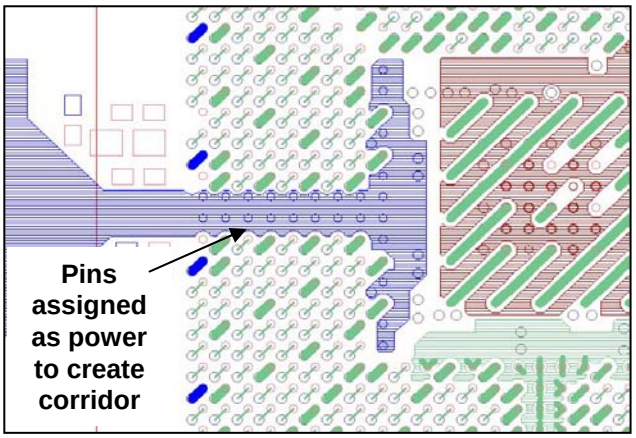


Figure 4: Power corridor on top-surface layer

Another benefit of the corridors is that they allow higher signal density breakout on the inner stripline layers due to that region being absent of vias. Without a via being associated with every package pin, these corridors can be seen on internal routing layers as a wide unused region through which many signals can be routed. This allows for greater signal density within a smaller package size (see Figure 5).

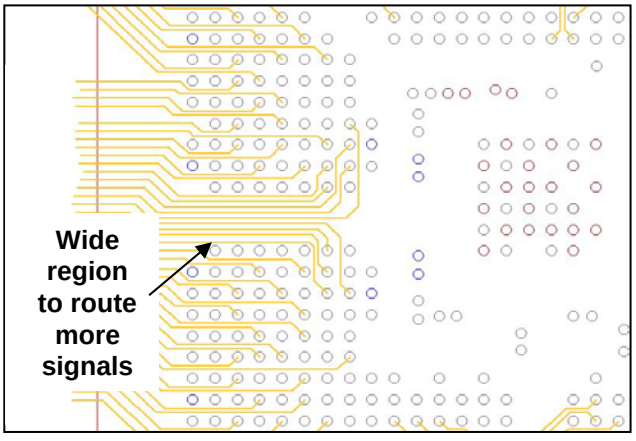


Figure 5: Signal routing under power corridor

The final benefit of power corridors is that the same advantage seen on internal layers is seen on the backside surface layer. Just like the wide open regions on the internal layers, the backside surface layer is open and can be used to deliver wide copper floods for additional power delivery (see Figure 6).

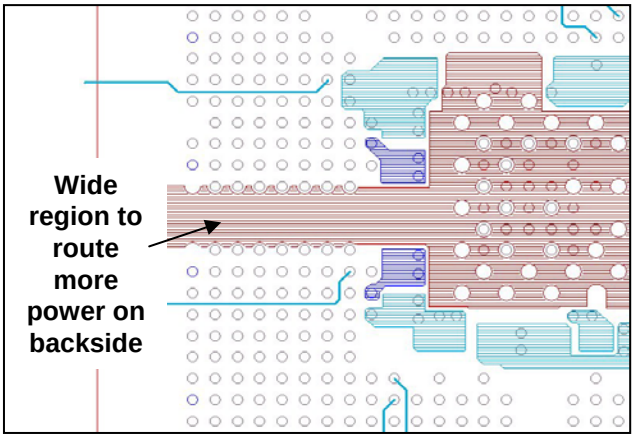


Figure 6: Power routing under power corridor

Together with interface removal, the parquet package technique, and corridor routing, the size of the Intel 915GMS package was reduced to 27 x 27 mm, a 51% area reduction from the original 37.5 x 40 mm package size of the standard Intel 915 chipset.

PACKAGE SUBSTRATE LAYOUT AND PACKAGE TECHNOLOGY CHALLENGES

The challenges faced by the package development team in delivering the Intel 915GMS package included pushing key assembly technology parameters.

Package Size and Layer Stack-up

The package development team was given the task of delivering the smallest package possible for the given product feature-set yet still meeting the electrical requirements.

Using a 6-layer package substrate, similar to the standard Intel 915 chipset package, would have resulted in a package size of at least 35 mm. Moving to an 8-layer package substrate would enable the bulk of the signals, mainly front-side bus (FSB) and DDR2, to be routed as striplines, and this would reduce the size of the package. As mentioned earlier, there were a couple of package sizes that were considered during the feasibility analysis phase.

Refer to Figure 7 and Figure 8 for definitions of microstrip and stripline routing.

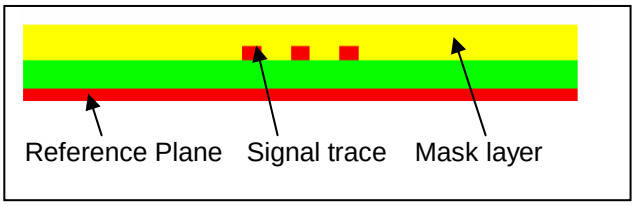


Figure 7: Microstrip signal routing

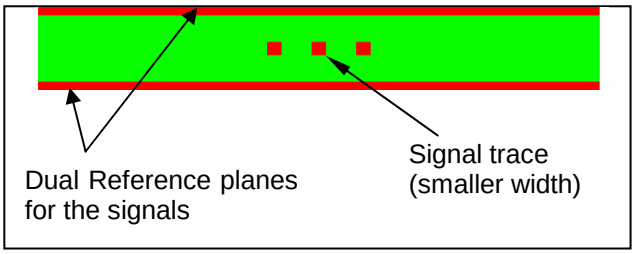


Figure 8: Stripline signal routing

The Intel 915 chipset used a 6-layer package substrate with all the signals routed on the surface layer as microstrip. Stripline routing with smaller trace width and spacing, when compared to microstrip routing, allowed us to achieve more routes per unit area, while maintaining the electrical integrity. Utilizing a 26 µm trace width for stripline results in a 50-ohms impedance as compared to a 55-ohm impedance which is found on the Intel 915 chipset. We decided to proceed with a 26 µm trace width for FSB and DDR2. The 5-ohms mismatch was deemed a non issue for FSB speeds @ 400/533 MHz. With DDR2 motherboard routing optimal at 40-ohms, the 50-ohms on the Intel 915GMS package was seen as an advantage over the 55-ohms on the standard Intel 915 chipset package.

Sub 1.0 mm Ball Pitch

One of the two package options required a minimum ball pitch of 0.8128 mm (32 mils) that was falling out of the Intel certified flip chip package technology envelope. Limited experimental reliability data was available on sub 1.0 mm ball pitch. The initial experimental data collected for the shock and vibrations test, for sub 1.0 mm ball pitch, showed that the solder joint reliability was a manageable risk. The product development team decided to continue working on the sub 1.0 mm pitch package design.

The package substrate design was intercepted, just before the design completion, to enlarge the ball grid array (BGA) pad size, an increase that was recommended based on the initial experimental results. For the subsequent shock and vibration experiments, Design of Experiments (DOEs) were planned with refined mobile shock tests and increased pad size on the substrate. Engineers from the Intel Technology and Manufacturing groups and the Mobile Chipset Products group worked closely to define a new experimental set-up including the Shock-Test Board design and the revised Shock-Test Spec. Interaction with key customers to understand the system-level designs and their capability for shock and vibration formed part of the DOEs' planning process. The data collected using the Intel 915GMS chipset Thermo-Mechanical Test Vehicle (TMTV) showed that the overall shock and vibration risk level of the Intel 915GMS package was manageable, and so the technology was certified.

Keep-Out Zone (KOZ) Reduction

Certain areas on the top surface of the package must be clear of any interferences: the die, decoupling capacitors, and underfill material. These are called Keep-Out Zones (KOZs) and are needed for handling of the package during the package assembly and testing processes.

Figure 9 shows the KOZs surrounding the die and the package periphery. The zones shown were defined as per the Intel Package Design specifications. As you can see, there is an overlap of KOZs at various locations. The overlapping of the KOZs violated the design specification and does not accommodate any on-package capacitors.

Experiments in the assembly and test areas were initiated to determine how to reduce the KOZs. The experiments yielded positive results. The Edge exclusion zone was reduced by over 25% of the specification and the Corner exclusion zone was reduced by 35%.

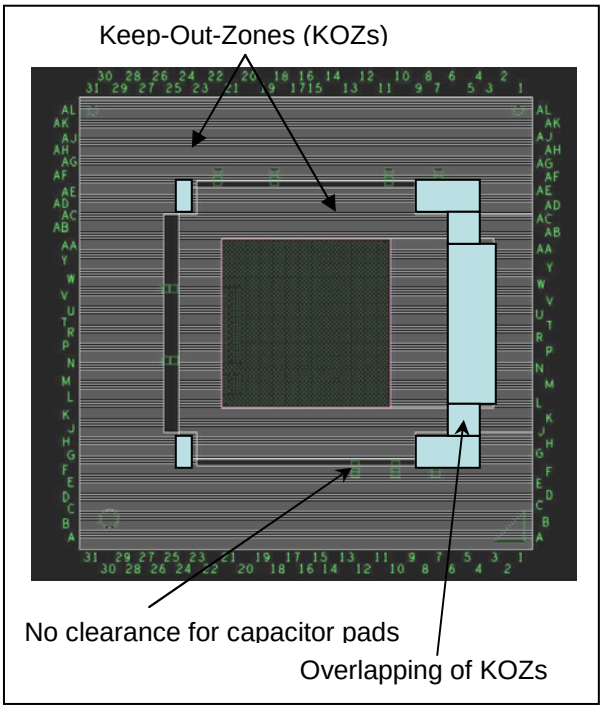


Figure 9: KOZs overlapping (before modifying the Intel Package Design spec)

Figure 10 shows the modified KOZs, which allowed the design to accommodate the on-package capacitors.

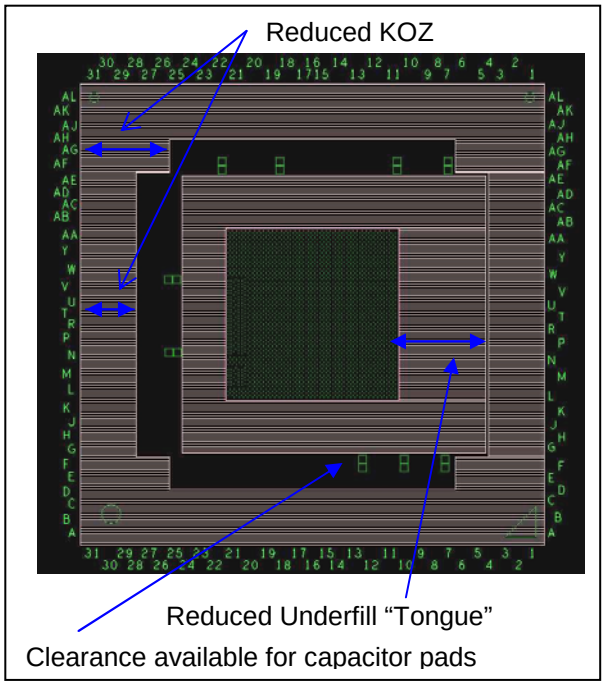


Figure 10: KOZs after modification of Intel Package Design spec

Also, the assembly line experiments showed that we could reduce the Underfill Tongue-width by 17%, which was needed to avoid overlap with the already "reduced" edge exclusion zone. A "trench" was cut on the top solder mask layer that further enabled us to reduce the tongue width. Since the routing on the east side was stripline, the trench was implemented without any signal traces passing across the trench. Figure 11 shows the trench design.

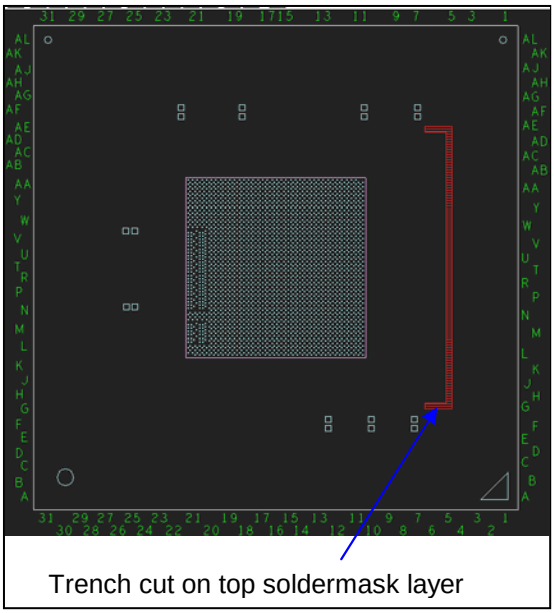


Figure 11: Location of the trench

MOTHERBOARD ROUTING IMPROVEMENTS

Along with providing an optimized chipset for small form-factor platforms, Intel has also put effort into developing guidelines to help in reducing the motherboard size and in reducing the effort required to design a small form-factor platform.

In the following sections, we discuss these guidelines and their benefits. The guidelines are optimized for two important areas: system memory and FSB.

System Memory Guidelines

With the Intel 915GMS chipset, Intel provided guidelines for implementing on-board memory that helps to reduce the size of the board. In a notebook design, there are many benefits if one can design the memory on the board rather than utilizing Small Outline Dual Inline Memory Modules (SO-DIMMs). By going for on-board memory, along with a reduction in the board size, a reduction in height can also be achieved.

Typically, the notebooks that are available in the market have two SO-DIMM slots. The product is sold with one SO-DIMM pre-installed; if consumers wish to increase the memory, they can buy an additional SO-DIMM and upgrade their notebooks. With Intel 915GMS platforms, designers can take advantage of the new on-board memory guidelines to change the pre-installed memory from SO-DIMM to on-board memory. Also, with products where a memory upgrade by the end user is not feasible, on-board memory can supply the total memory requirements of the platform.

Figure 12 illustrates the benefits achieved, with respect to board size and height, when designed with one SO-DIMM and on-board memory as compared to two SO-DIMMs. Benefits would even be greater if only on-board memory is supported.

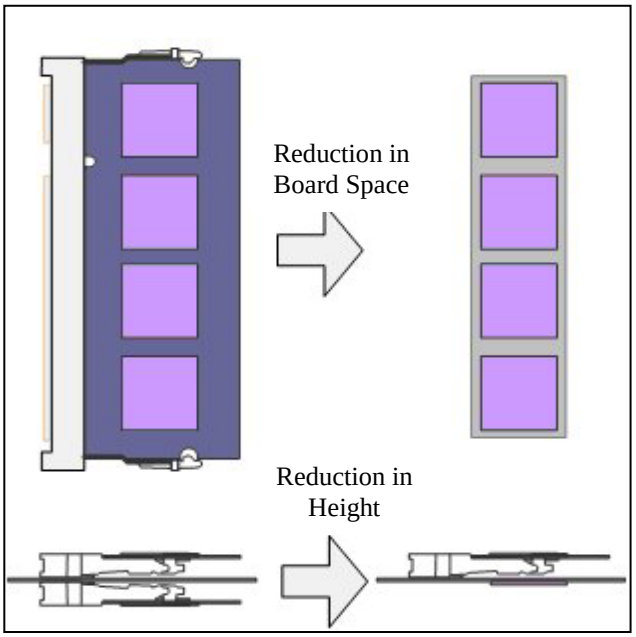


Figure 12: Benefits of on-board memory

Intel provides guidelines for various configurations with SO-DIMM and on-board memory. With these configurations, one can design the platform that can support up to 2 GB of memory. Table 2 lists these configurations and the maximum memory possible when 1 Gbit memory devices are used.

Table 2: Various configurations

Configuration (in space-saving order)	Max Memory (with 1 Gbit Technology)
1 rank (4 devices) on-board	512 MB
2 ranks (8 devices) on-board	1 GB
1 SO-DIMM	1 GB
1 SO-DIMM & 1 rank on-board	1.5 GB
1 SO-DIMM & 2 ranks on-board	2 GB
2 SO-DIMMs	2 GB

A configuration of four SDRAM components, where all four components are on the same side in one row, is the simplest configuration to route and yields the best signal integrity as well as timing budgets. When designing with eight SDRAM components, it is recommended to have four components on the top and four on the bottom of the board. This configuration is easier to route and has better signal integrity than designs where all of the eight components are on the same side: these may necessitate the use of more than 8-layer boards in order to route all of the memory signals.

For the configurations with on-board memory, the recommendations are for the SO-DIMM to be placed closest to the Intel 915GMS and the on-board memory SDRAM devices to be placed farthest away from the Intel 915GMS. Figure 13 shows the topology for one SO-DIMM and eight devices on the board.

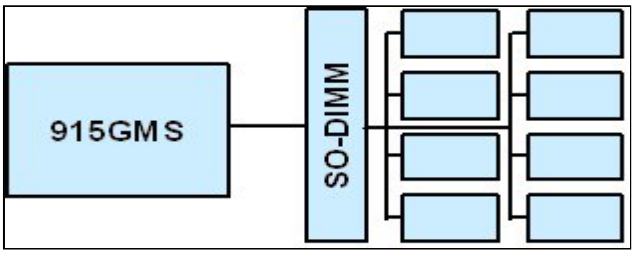


Figure 13: Simplified topology with eight memory devices

Front-Side Bus Guidelines

Intel 915GMS supports a 400 MT/s FSB to connect to low-voltage and ultra-low-voltage processors based on Intel Pentium M 90nm technology. The FSB guidelines were optimized to ease the routing on small form-factor platforms. With these optimized guidelines, the minimum routing length of the FSB signals can be as low as 0.1" as compared to 0.5" for earlier platforms. This provides the possibility of placing the processor and the Intel 915GMS very close to each other. Also, the minimum signal-to-signal spacing of the FSB interface can be as low as 1:1 (4 mil trace and 4 mil space) as compared to 1:2 (4 mil trace and 8 mil space) of earlier platforms. The reduced spacing requirement makes routing of FSB signals more flexible and also allows for denser routing of the signals. This leads to a reduction in the required routing space on the board.

INTEL 915GMS USAGE MODELS

Up to this point we have focused our discussion on what was done to reduce the size of the Intel 915GMS chipset. This reduced size delivers on the Intel Centrino brand promise of enabling unique form factors. However, what exactly will these products look like? How can this technology be used?

There are many vertical segments that might be interested in smaller and lighter form factors: health care, research, entertainment, and education, for example. These designs could be in the form of Tablet PCs, extremely thin-and-

light notebooks, or something radically new. Only time will tell what creative OEMs will do with these platforms, based on the Intel 915GMS platforms.

In the field of health care, for example, people move around a lot such as in hospitals where doctors and nurses need to gather information from many patients. Most of this information today is not digitized. Smaller tablet designs could provide the right combination of features to push digitization into this field and make this usage model real.

Similar to the health care profession, researchers in all disciplines require mobility in order to collect information in the field. For these individuals, smaller and lighter notebooks would provide greater ease of data collection. Notebooks, based on the Centrino mobile technology, also provide the ability to wirelessly upload these data to laboratories for analysis and storage.

OEMs targeting consumers who purchase mobile PCs to enjoy entertainment have many opportunities to take advantage of smaller form factors. The reduced size of the Intel 915GMS could be coupled with designed-in features like small wide-aspect ratio LCDs, and the capability to instantly wake-up the platform into a DVD-playback mode. This effort could result in a smaller entertainment focused mobile platform.

As notebook computers get smaller and lighter, their use in education by students will also increase. For example, students currently carry most of their text books to and from school. By making use of Intel's smaller and lighter notebooks, students would have less heavy books to carry; something that would also make it easier for younger students where the weight of books is often an issue.

As the market for personal computers expands so will the demanded usages from customers. The examples listed above are only a few of the areas where the Intel 915GMS chipset can provide for further development of the uniqueness that mobility provides.

Future Challenges

The Intel 915GMS chipset improves upon the form-factor promise of the Intel Centrino mobile technology brand. However, there are areas where future improvements could be made in order to allow form factors to become even smaller. The underlying design challenge is to reduce the size of the platform, while still providing performance and great battery life.

First and foremost, the techniques used on the Intel 915GMS product need to be utilized on other components on the platform, including the Intel Pentium M processor, ICH, and wireless solutions. Ball pitch and board routing technologies can be pushed to even further reduce package sizes. And additional integration of other platform components, removal of resistors and capacitors, and voltage plane reduction will all help to even further reduce the size of systems based on Intel Centrino mobile technology.

Advancement in some of these areas will be required to continue to push the vector of thinner, lighter designs, and some are being considered by Intel for use in future generations of small form-factor designs.

SUMMARY AND CONCLUSION

Together with signal and power-pin reduction, optimized package ball pitch, and resolution of manufacturing concerns, Intel was able to develop the Intel 915GMS product with a 27 mm square package, a 51% reduction from the standard Intel 915 package.

The excitement from the customer base is clear: there are more than a dozen Intel 915GMS platforms in development, all ready to launch in the March 2005 through September 2005 timeframe. Customers will be able to take advantage of this small package, and use the routing guidelines to further reduce the board area of their products. Intel expects that this will enable smaller and sleeker form factors, and potentially new usage models. And as the interest of platforms based on Intel Centrino mobile technology continues to grow, it is clear that the Intel 915GMS is just the first of many products focused on smaller and sleeker form factors.

ACKNOWLEDGMENTS

The authors acknowledge all of the Intel teams who worked together to develop the Intel 915GMS chipset. This list includes, but is not limited to, the Product Development Team, the Package Development Team, the Second Level Interconnect Team, the Future Platform Enabling Team, the Signal Integrity Analysis Team, the Power Delivery Team, the Assembly/Test/Development Team, the Manufacturing Team, the Board Development Team, the Post-Silicon Validation Team, and the Applications Engineering Team.

AUTHORS' BIOGRAPHIES

David W. Browning is a senior design engineer in the Mobile Product Group at Intel Corporation with a focus on advanced mobile platform architecture. He has been working for Intel for more than six years. David has a B.S. degree in Electrical Engineering from the University of Washington. His e-mail is david.w.browning at intel.com

Vinay Kumar C. is a platform application engineer and was in charge of the platform design guide document. He also provided technical support on Intel 915 to mobile customers. He has a B.E. degree in Electronics and

Communication from India. After a job at Force Computers as a board designer he joined Intel in 2003. His e-mail is vinay.k.c at intel.com.

Praveen G. is a platform application engineer and was in charge of the chipset datasheet. He also provided technical support on the Intel 915 chipset to mobile customers. He has a B.E. degree in Electronics and Communication from India. After a job at Force Computers as a board designer he joined Intel in 2003. His e-mail is praveen.g at intel.com.

Ananth V. Gopal is a software platform application engineer and provided technical support on the Intel 915 chipset to mobile customers. He has a B.E. degree in Computer Science and Engineering from India. He joined Intel in 2003. His e-mail is ananth.v.gopal at intel.com.

Tom Shewchuk was the Intel 915 chipset program manager and currently is managing Intel's next-generation mobile chipsets. He has B.S.E.E, M.S.E.E, and M.B.A. degrees, all from the University of Illinois at Urbana-Champaign. After jobs at Honeywell and Ford Microelectronics, he joined Intel in 1991, spending his first five years working on microprocessors, and the past nine years working on integrated graphics chipsets. His e-mail is tom.j.shewchuk at intel.com.

Suresh V. Subramanyam is the package design manager with Intel, India, supporting mobile chipsets. He has a B.S.M.E. degree from India and an M.S. degree in Materials Science & Engineering from Singapore. He joined Intel in 2003 after spending six years with HP/Agilent, Singapore. He started his career with Bharat Electronics, Bangalore, India, in 1988. His e-mail is suresh.v.subramanyam at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

The Emergence of PCI Express* in the Next Generation of Mobile Platforms

Mohammad Kolbehdari, Digital Enterprise Group, Intel Corporation
David Harriman, Digital Enterprise Group, Intel Corporation
Altug Koker, Mobility Group, Intel Corporation
Seh Kwa, Mobility Group, Intel Corporation
Brad Saunders, Mobility Group, Intel Corporation

Index words: PCI Express* architecture, ExpressCard*, mini-card, GMCH, ICH, SIOM, DLLP, TLP, PHY, power management (PM), reliability, high-end graphics, PC built on Intel® Centrino™ mobile technology

ABSTRACT

The PCI Express* architecture, both as a unified foundation of I/O, graphics, and networking interconnections, and as a preeminent building block on chip-to-chip, board-to-board, and system-to-system, is widely adopted by multiple market segments in the computing and communication industries. PCI Express architecture is a state-of-the-art serial interconnect technology that keeps pace with recent advances in processor and memory subsystems. From its initial release at 0.8 V, 2.5 GHz, the PCI Express technology has evolved to the general-purpose interconnect of choice for a wide range of applications, including graphics, storage, networking, etc. The PCI Express architecture retains the familiar PCI software and configuration interfaces for seamless migration and adoption in desktop and mobile PC platforms, enterprise servers and workstations, and, increasingly, a wide range of communication and embedded systems.

The PCI Express technology addresses requirements from multiple market segments in the computing and communication industries, and it supports chip-to-chip, board-to-board, and adapter solutions at an equivalent or lower cost than existing PCI designs. Currently, PCI Express architecture supports a 2.5 GT/s signaling rate that yields 500 MB/s bandwidth per lane and a maximum bandwidth of 16 GB/s in a 32-lane configuration. Consistent with the expected cadence of I/O performance progression, the next generation of the PCI Express interconnect will support a signaling rate of 5 GT/s, doubling the performance of the existing links.

The PCI Express interconnect provides numerous architectural improvements over existing I/O technologies. It defines a native hot-plug scheme, enables aggressive power management, provides advanced Reliability, Availability, and Serviceability (RAS) and Quality of Service (QoS) features, and simplifies PCB layouts. In this paper, we discuss the unique, universal capabilities and values of PCI Express technology emerging in the next generation of mobile platforms. We focus on PCI Express architecture, power management, and mobile applications such as graphics, networking, and form factors including the ExpressCard* module as well as future form factors such as the PCI Express Wireless Card and the PCI Express Mini Card. Finally, we cover the next generation of mobile PC platforms built on the Intel® Centrino™ mobile technology.

INTRODUCTION

The PCI Express technology has emerged as the platform I/O solution of choice in the computing and communications industries. Some significant features of the PCI Express technology are also geared towards the next generation of mobile platform designs especially the Active State Power Management (ASPM) capability that enables aggressive power management within the link

* Other brands and names are the property of their respective owners.

® Intel is a registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

™ Centrino is a trademark of Intel Corporation or its subsidiaries in the United States and other countries.

layer as defined in the PCI Express Base specification. Support for ASPM varies by form factor; however, all mobile electromechanical specifications require ASPM support by PCI Express applications. In this paper, we focus on PCI Express applications, PCI Express architecture and protocols, as well as PCI Express form factors for next-generation mobile platforms and high performance, such as ExpressCard module defined by the Personal Computer Memory Card International Association (PCMCIA) industry group and the PCI Express Mini Card defined by the PCI Special Interest Group (PCI-SIG).

Additionally, the next-generation Intel Centrino mobile technology platform contains a variant of the Intel 915 Express as its "root complex." This chipset is comprised of a PCI Express x16 link optimized for high-performance graphics applications chipsets, the Graphics Memory Controller Hub (GMCH), and the I/O Controller Hub (ICH), which provide four PCI Express x1 links for general-purpose I/O applications. These latter links can be used to implement a combination of PCI Express Mini Card sockets, slots for ExpressCard modules, and other PCI Express devices on the system board that enable communications, storage, or other peripheral applications. If additional ports are necessary to implement a higher-end platform, OEMs may add a PCI Express switch to the ICH to increase I/O fanout. To facilitate a smooth, gradual transition for legacy devices, the conventional, multi-drop PCI bus continues to be supported via the ICH.

PCI EXPRESS ARCHITECTURE

Architectural Overview

The PCI Express architecture uses familiar software and configuration interfaces, the conventional PCI bus architecture, by providing a new high-performance physical interface and numerous new and enhanced capabilities that are built into a framework that retains software compatibility with the existing conventional PCI infrastructure. The enhanced capability is new and different from the legacy, parallel PCI bus architecture and supports scalable link widths in 1-, 2-, 4-, 8-, 16- and 32-lane configurations.

PCI Express power management is built on the PCI Power Management (PCI-PM) software architecture defined for conventional PCI. Additional PCI Express specific power-management capabilities are defined that further extend power manageability by allowing direct hardware control of power states with entry and exit latencies that are low enough to be effectively invisible to software.

PCI Express supports a native hot-plug architecture. Hot-plug support requirements vary by form factor and are documented in the various PCI Express electromechanical specifications. The hot-plug "Toolkit" is defined in the PCI Express base specification to ensure a consistent, common interface for system software to correctly manage hot-plug operations.

PCI Express architecture supports numerous Reliability, Availability, and Serviceability (RAS) features such as error detection and reporting that can be detected using an End-to-end Cyclic Redundancy Check (ECRC) and a Cyclic Redundancy Check (CRC). Erroneous packets are corrected, and the reporting and logging of error conditions is considerably expanded. Mechanisms such as traffic service differentiation, including architecturally defined mechanisms for system control of arbitration, are also provided.

Support for multiple form factors was considered from the outset of the architecture definition. In addition to a card form factor (similar to the existing mainstream PCI cards/slots), there are other form factors defined such as the PCI Express Mini Card, also similar to the existing conventional Mini PCI form factor for notebooks, the ExpressCard standard, meant to replace the CardBus PC Card* form factor, a wireless card form factor designed to fit in the lid of notebook computers, an enterprise-class hot-plug module commonly called a Server I/O Module (SIOM), and several mezzanine cards and blade modules. Additional form factors are being defined.

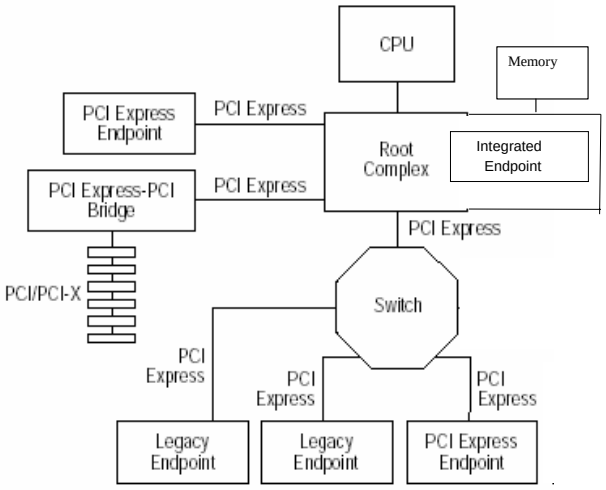


Figure 1: PCI Express architecture system

Figure 1 shows an example of the PCI Express architecture system. There are six types of PCI Express devices. Endpoints include devices such as network and disk interfaces, and they are categorized as either PCI Express Native endpoints or Legacy endpoints. The

* Other brands and names are the property of their respective owners.

Legacy endpoint category is intended for existing designs that have been migrated to PCI Express, and it allows legacy behaviors such as use of I/O space and locked transactions. PCI Express endpoints are not permitted to require the use of I/O space at runtime and must not use locked transactions. By distinguishing these categories, it is possible for a system designer to restrict or eliminate legacy behaviors that have negative impacts on system performance and robustness. PCI Express/PCI Bridges allow older PCI devices to be connected to PCI Express-based systems. Since PCI Express is a point-to-point interconnect, switches are used to increase connectivity. The Root Complex (RC), which includes the Graphics Memory Controller Hub (GMCH) and the I/O Controller Hub (ICH), is at the top of a PCI Express hierarchy, and it connects PCI Express to the CPU and main memory.

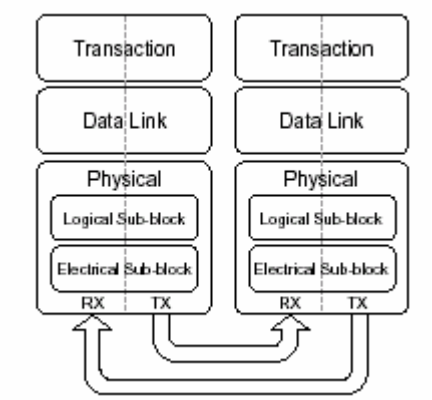


Figure 2: PCI Express link and layered architecture

Figure 2 shows two components connected by PCI Express and highlights the layered structure of the PCI Express interface at the upper component. It should be noted that the layering structure shown represents the way PCI Express is described, but is not an implementation requirement. The Physical Layer (PHY) initialized the link between the two components on a link, and it manages low-level aspects of data transfer and power management. The Data Link Layer (DLL) provides reliable data transfer service to the Transaction Layer (TL) and also a lower overhead communication mechanism for link management of flow control and power. Data packets generated and consumed by the DLL are called Data Link Layer Packets (DLLPs). The TL generates and consumes data packets used to implement load/store data transfer mechanisms and also manages the flow control of those packets between the two components on a link. Data packets generated and consumed by the TL are called Transaction Layer Packets (TLPs).

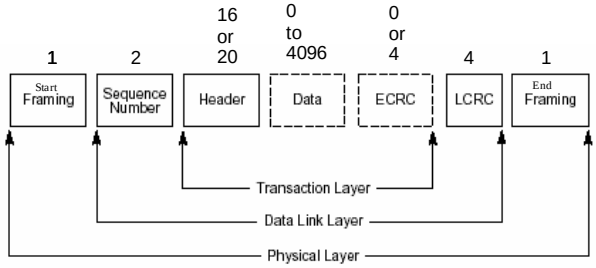


Figure 3: Representation of a Transaction Layer Packet as transmitted over a link

Figure 3 shows how a TLP generated by the TL is modified as it flows through the DLL and PHY for transmission. The Header describes the type of packet and includes information needed by the receiver to process the packet, including any needed routing information, and it is either 12 bytes or 16 bytes in size. Some TLPs include data that are included following the Header. Up to 4096 bytes of data can be included in a TLP, although in a specific platform the maximum size may be limited to a smaller value. The number of bytes of data is always a multiple of 4. The TL may append a 4-byte ECRC, which will travel with the TLP throughout the PCI Express hierarchy "end-to-end." The packet formed by the TL is delivered to the DLL that adds a Sequence Number and a local CRC called a "LCRC." Both of these are used by the receiving DLL to determine that the TLP was received without corruption and without packet loss. Finally, the TLP is delivered to the PHY, which converts it from a sequence of 8-bit bytes to a sequence of 10-bit symbols and adds framing symbols at the start and end. This sequence of symbols is then transmitted across the link to the other component, which checks and disassembles the TLP using a process that mirrors the assembly process at the transmitter.

We now consider in more detail the layered structure of a PCI Express interface.

Electrical Signaling

PCI Express preserves the PCI and PCI-X* architecture with additional new PHY electrical sub-block features as follows:

- *Embedded clock*. Unlike PCI and PCI-X, the clock is not an external data or strobe, but it is embedded in the data signal itself. This new capability enables PCI Express to reduce the clock timing skew.
- *Differential link*. This is the difference in the D+ and D- signals. It forms the signal amplitude and has the

* Other brands and names are the property of their respective owners.

advantage of low crosstalk, low Inter Symbol Interference (ISI), improved Electromagnetic Compatibility (EMC), low common mode, and shared reference ground, among others.

- Full-duplex transmitting with independent transmitter and receiver signal links.
- *AC coupling.* This isolates each transmitter and receiver enabling both hot-plug and hot-swap capabilities and also decodes the transmitter and receiver common mode voltage.
- *8B/10B encoding scheme*. This significantly improves EMC and electrical signal performance. This feature enables the frequency band range limit by encoding each 8-bit byte data to one of two defined 10-bit codes given in the specification. The numbering disparity for each symbol is controlled by this feature maintaining the maximum difference between the number of zeros or ones transmitted across multiple symbols at a delta of 2. This difference keeps a DC balance at both the transmitter and the receiver.
- *Data scrambling with a polynomial equation*. This transfers an 8-bit data byte to a different data byte value that reduces Electromagnetic Interference (EMI) radiation and emission.
- De-emphasis on functionality.
- Polarity inversion and lane reversal.

PCI Express architecture is attributed with ports, links, lanes, and detect. A port is a group of transmitters and receivers on a component that constitutes a link when connected to another component. A line is one set of transmit and receive differential pairs. Link manages an intelligent training sequence to detect and establish deterministic links between component ports. Detect is a feature when an upstream device, RC, or north bridge polls each transmit lane to detect an RC time constant that indicates that a load is present. Once the upstream device detects a load, it initiates a polling process to determine whether the load is an active device or a test load. Figure 4 shows the concepts of ports, links, and lanes.

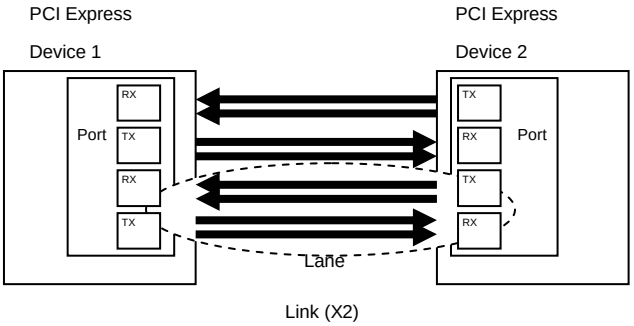


Figure 4: The concept of ports, links, and lane

The PCI Express base specification [1] requires AC coupling between the transmitter and receiver. It also requires each end of the link to be on-die terminated with a 100 Ohm differential. In addition, the PCI base specification requires each transmitter (TX) component to use on-die equalization or de-emphasis with a typical value of 3.5 dB +/- 0.5 dB. The PCI base specification defines the Unit Interval (UI) of 400 ps +/- 300 parts per million (ppm). The PCI Express timing specification defines eye diagrams, one at the transmitter end and one at the receiver end. Also the PCI Express Card Electromechanical Specification [2] defines a specific eye diagram with given height and width for the add-in card and system board. The PCI-base specification allows certain insertion and return losses to ensure that the transmitter signal has adequate amplitude at the receiver signal pin. It also defines a rise and fall time of 0.125 UI or 50 PS. For all the electrical and AC specifications see the PCI Express base specification and also the PCI Express Card Electromechanical Specification [1-2].

Protocol

The PHY Electrical sub-block, described above, is controlled by the Logical sub-block that implements link initialization and low-level control functions as well as packet framing and, for multi-lane links, distribution of packet information across the lanes. The Link Training and Status State Machine (LTSSM) is responsible for establishing the link between two components at the PHY level, including the width of a multi-lane link. When establishing a link, the LTSSM starts in the Detect state, moving to the Polling state when another component is found to be present on the link. Once the two components have established a data transfer connection, they enter the Configuration state where the configuration of the link itself is negotiated between the two components. Finally the L0 state is entered to establish normal link operation. The LTSSM is also capable of re-establishing the link if it is "lost" due to an electrical disturbance or a power state transition, using the Recovery state. It also manages link power state transitions at the PHY level using the L0s, L1, and L2 states.

Depending on system clocking, it is possible that the transmitter and receiver will operate at slightly different clock rates. The PHY uses a periodic transmission of a specific sequence of symbols to allow the receiver to compensate for any drift between the two components.

In a multi-lane link it is also necessary for the receiver to compensate for any differences in skew between the multiple lanes of the link. This is done by transmitting specific sequences of symbols simultaneously on all lanes such that the receiver can detect and compensate for any differences in the receipt of the sequences on a lane-by-lane basis. In this way, the other layers see the received data as all arriving at the same time.

Once the physical link is established by the PHY, the DLLs and TLs of the two components will then exchange flow control information to allow the transmission of TLPs between the two components. This initial exchange of flow control information is managed by the Data Link Control and Management State Machine of the DLL. Once both components have received the required information, TLP communication starts. The DLL implements a retry mechanism that stores transmitted TLPs until they are acknowledged by the receiver, presenting the view of a reliable link to the TL.

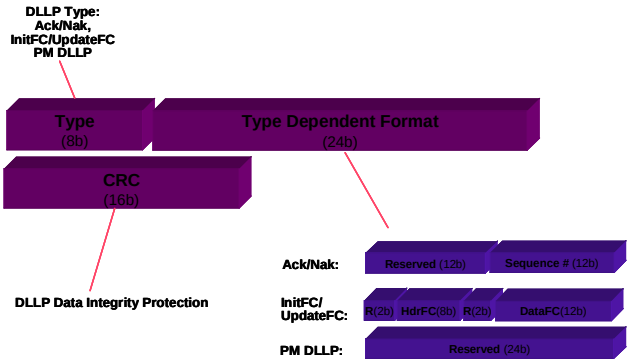


Figure 5: Data Link Layer Packet (DLLP) format

The retry mechanism, continued flow control information exchange, and certain power-management protocols use DLLPs, of which there are 15 specific types defined. Figure 5 shows the format of the 6-byte body of a DLLP, not including the framing information added by the transmitting PHY and removed by the receiving PHY. The first byte describes the type of the DLLP. The format of the following three bytes depends on the type of the DLLP. The final two bytes are a 16-bit CRC that allows the receiver to determine that a DLLP was received without corruption. DLLPs that are corrupted in transit are simply discarded by the receiver, and all protocols using DLLPs ensure that such loss will be corrected by a subsequent transmission.

There are two types of TLPs used by the TL: Request TLPs, and Completion TLPs, simply called "Requests" and "Completions," respectively.

A schematic representation of the Header of a Request is shown in Figure 6. Note that this figure does not represent the actual bit positions of the bits in the Header, but is simplified to highlight the functionality of the fields. The Fmt and Type fields encode the type of layout of the TLP, for example to encode a Configuration Write or a Memory Read. The Length field encodes the length of data included with the TLP, or the amount of data requested to be returned for a Read Request. The TC and Attr fields are used for traffic service differentiation. The TD field indicates TLP includes an ECRC. The EP field indicates that a TLP includes data that are "poisoned" i.e., known by the transmitter to be erroneous. The Requestor ID is used to route a Completion, if needed, back to the Requestor, and the Tag is used by the Requestor to distinguish different Completions for multiple outstanding Requests. The field labeled "BE/Msg. Code" is used for Configuration, I/O, and Memory Requests to encode the byte enables for the Request, and for Message Requests to encode the Message type. The remainder of the Header is used for routing information, either an address in Memory or I/O space, or the unique ID of a specific device as used for a Configuration Request.

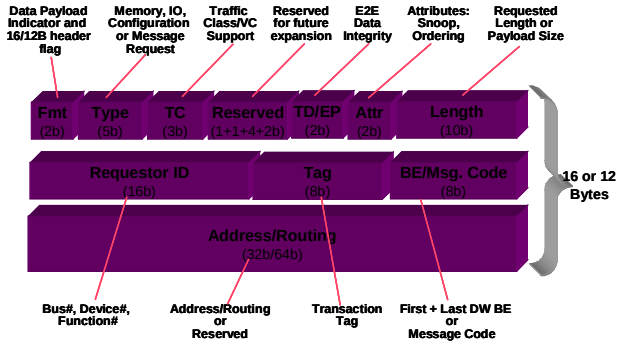


Figure 6: Representation of the format of a request TLP

Completion Headers are similar to Request Headers. Figure 7 shows a schematic representation of a Completion Header. Completions always use the 12-byte Header size, because full address information is not required. Descriptions of fields not present in the Request Header follow.

The Status field communicates to the Requester that its Request was completed successfully or with some type of error. In addition to a Successful Completion status, there are two error statuses: Unsupported Request and Completer Abort, which correspond to Master Abort and Target Abort in conventional PCI. There is also a

completion status used only in association with an initial configuration Request to a device (Configuration Retry status), which indicates that the Completer requires more time before it is able to service the Configuration request. Completion headers include two ID fields: the Requestor ID and Tag. These are the same as the Requestor ID and Tag values supplied with the corresponding Request. The Completer ID identifies the entity that serviced the Request to generate the Completion, and uses the same format as the Requestor ID. Byte Count and Lower Address can be used by Requestors to simplify processing of data returned in Completions for Memory Read requests.

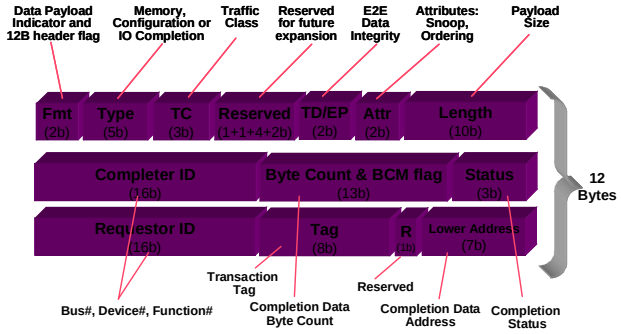


Figure 7: Representation of the format of a Completion TLP

PCI Compatibility and New Software Features

PCI Express devices include the same Configuration headers as corresponding conventional PCI devices. System software that comprehends conventional PCI works with PCI Express devices using these configuration mechanisms. For systems with updated system software, PCI Express expands the capabilities available to software to provide for native hot-plug support, improved data integrity, and error-reporting mechanisms, all of which are covered in this section, as well as more advanced power management, which is covered in the following section. For additional new capabilities not covered in this paper, refer to the PCI Express Specification [1]. PCI Express provides a new memory-mapped configuration access mechanism to replace the windowing mechanism used for conventional PCI in PC architecture systems, and it provides each device with an expanded configuration space of 4 KB compared to 256 B in conventional PCI.

PCI Express defines a register interface to support a "toolbox" of hot-plug capabilities that can be used by different form factors to allow different usage models with a consistent system software interface. Each element of the toolbox that a form factor may specify as a required or optional element of hot-plug support for that form factor has an architected capability bit associated with the port connected to the slot or connector. Table 1 lists the supported elements that form factors may employ to support a particular hot-plug usage model, although it should be noted that in the mobile environment only the Surprise Remove capability is used.

Table 1: Elements of the PCI Express hot-plug "toolbox"

Attention Button	Used to request hot-plug operations. Note: A software user interface may be used instead or in addition.
Attention Indicator	Indicates slot and/or adapter that requires user attention.
Power Indicator	Indicates that the slot has power.
Electromechanical Interlock	Prevents removal of an adapter from a slot under system software control.
Manually operated Retention Latch (MRL) Sensor	Allows software detection of MRL operation by the user, for example to indicate that the user is about to remove an adapter.
Power Controller (software controlled)	Allows software control of slot power.
Hot-Plug Surprise	Indicates that an adapter removal without advanced software notification is permitted by this form factor.

If an element is included, it must be used in a manner consistent with certain defined usages.

The ExpressCard form factor permits surprise removal of the adapter, and does not include a software-controlled power controller or support for the button/indicator user interface. The toolkit approach allows a common system software implementation to support all PCI Express hot-pluggable form factors.

POWER MANAGEMENT

PCI Power Management is an optional capability for conventional PCI devices, but is required for PCI Express devices. The PCI power management states are mapped to PCI Express link states to allow existing system software to power-manage PCI Express Links without having to comprehend the actual link states.

Additionally, the PCI Express specifications define an Active State Power Management (ASPM) capability that enables aggressive power management operating within the link layer. Support of ASPM is form-factor specification dependent, and all of the mobile platform

form-factor specifications call its implementation out as a mandatory feature of all PCI Express end-point links.

Device Power States

The Advanced Control and Power Interface (ACPI) specification defines component device power states (D-states) that allow the platform to establish and control power states for the component ranging from fully ON to fully OFF (drawing no power) and various in-between levels of power-saving states, annotated as D0-D3. Similarly, PCI Express defines a series of link power states (L-states) that work specifically within the link layer between the component and its upstream PCI Express port (typically in the host chipset). For a given component D-state, only certain L-states are possible as detailed below.

- *D0 (Fully on)*: The device is completely active and responsive during this D-state. The link may be in either L0 or a low-latency idle state referred to as L0s. Minimizing L0s exit latency is paramount for allowing frequent entry into L0s while facilitating performance needs via a fast exit. The next lower link power state, L1 state, may be achieved either by hardware-based ASPM or by requesting the link to enter L1 after the OS places the downstream device in a D1-D3 state.
- *D1 and D2*: There is no universal definition for these intermediate D-states. In general, D1 is expected to save less power but preserve more device context than D2. L1 state is the required link power state in both of these D-states.
- *D3 (Off)*: Primary power may be fully removed from the device ($D3_{cold}$), or not removed from the device ($D3_{hot}$). $D3_{cold}$ maps to L2 if auxiliary power is supported on the device with wake-capable logic, or to L3 if no power is delivered to the device. Sideband WAKE# mechanism is recommended to support wake-enabled logic on mobile platforms during the L2 state. $D3_{hot}$ maps to L1 to support clock removal on mobile platforms.

Table 2 summarizes the mapping from D-states to L-states for a PCI Express link.

Table 2: Mapping from D-states to L-states

Downstream Component D-state	Permissible Upstream Component D-state	Permissible L-state
D0	D0	L0, L0s, or L1
D1	D0-D1	L1
D2	D0-D2	L1
$D3_{hot}$	D0-$D3_{hot}$	L1
$D3_{cold}$	D0-$D3_{cold}$	L2 or L3

Active State Power Management

ASPM is the hardware-based capability to power-manage the PCI Express link. Only L0s and L1 are used during ASPM.

- *L0s*: This link state is a very low exit latency (<1 μs) link state intended to reduce power wastage during short intervals of logical idle between link activities. The power-saving opportunities during this state include, but are not limited to, most of the transceiver circuitry as well as the clock gating of at least the link layer logic. Devices must transition to L0s independently on each direction of the link.
- *L1*: This link state is a low exit latency (~2-4 μs) link state that is intended to reduce power when the device becomes aware of a lack of outstanding requests or pending transactions. The power-saving opportunities during this state include, but are not limited to, shutdown of almost all the transceiver circuitry, clock gating of most PCI Express architecture logic, and shutdown of the PLL. CLKFREQ# can be opportunistically de-asserted during L1 to allow for platform reference clock gating.

In a multi-lane PCI Express link, the detection of L0 or L1 exits can be communicated through lane #0 of a configured link. Also note that L1 exit signaling does not have to be derived from Phase Locked Loop (PLL). (If this were not the case, the opportunity to shut down the PLL during L1 would be lost.) More aggressive low-power implementations may consider the minimization of leakage power during L1 state.

PCI Express Link Power States

Configuration of devices into D-states will automatically cause the PCI Express links to transition to the appropriate L-states. Refer to Table 2 for the mapping.

- *L2/L3 Ready:* This link state prepares the PCI Express link for the removal of power and clock. The

device is in the $D3_{hot}$ state and is preparing to enter $D3_{cold}$. The power-saving opportunities for this state include, but are not limited to, clock gating of all PCI Express architecture logic, shutdown of the PLL, and shutdown of all transceiver circuitry.

- *L2*: This link state is intended to comprehend $D3_{cold}$ with Aux power support. Sideband WAKE# signaling is recommended to cause wake-capable devices to exit this state. The power-saving opportunities for this state include, but are not limited to, shutdown of all transceiver circuitry except detection circuitry to support exit, clock gating of all PCI Express logic, and shutdown of the PLL as well as appropriate platform voltage and clock generators.
- *L3 (link off):* Power and clock are removed in this link state, and there is no Aux power available. To bring the device and its link back up, the platform must go through a boot sequence where power, clock, and reset are reapplied appropriately.

L2/L3 Ready and L2 are the link states with the lowest power but longest exit latency whereas L0 is the link state with lower power but short exit latency.

Pipelining and Buffering to Minimize Power and Maximize Performance

Techniques to improve packet scheduling and ensure effective utilization of bandwidth are also essential to a power-optimized architecture. Pipelining packets effectively and providing additional buffering are effective methods to reduce power while sustaining higher levels of performance. Two scenarios are described below as illustrated in Figure 8.

- Scenario (A) shows an "ask for one, get one" approach of utilizing a PCI Express link. One Completion is submitted after its Request is issued. Flow Control (FC) update and Acknowledge packets alternate, assuming no non-recoverable link errors.
- Scenario (B) illustrates the streaming of multiple Requests and subsequently multiple Completions before FC update and Acknowledge packets are issued. Note that the transmitting lane of the Requester is idle more often in scenario (A) than scenario (B), thus presenting longer periods for residence in lower link power states, provided efficient power-management features are implemented. Longer idle periods allow for power down of circuitry that would require non-zero energizing and synchronization time. Scenario (B) also allows the Completer to submit an FC update packet and an Acknowledgement packet after a few TL packets have been transmitted. Since the Completer is aware of the pending transactions, it is able to minimize the overhead of FC update and Acknowledge packets to maximize bandwidth utilization.

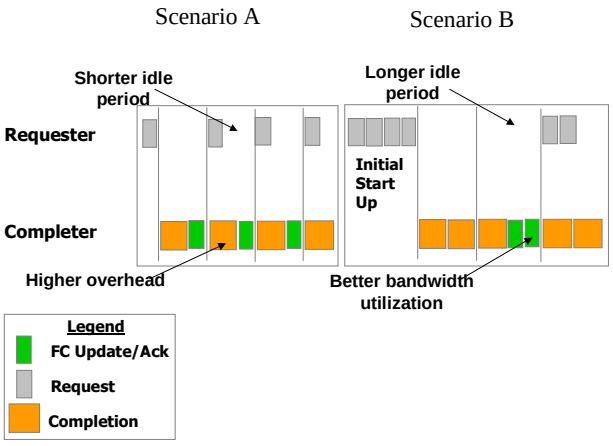


Figure 8: Scenarios to illustrate scheduling and bandwidth utilization impact

PCI EXPRESS MOBILE APPLICATIONS

PCI Express provides a high-performance, adaptable interface technology for the interconnection of applications peripheral to the platform's processor and memory subsystem. For mobile platforms, three key application areas are being initially deployed with PCI Express including the platform graphics and two add-in card formats for extending and upgrading the platform.

Similar to what is happening in the desktop PC space; PCI Express-based graphics are being established as the follow-on technology to existing Accelerated Graphics Port AGP-based graphics and is natively supported in the Intel host chipset (915PM). Also within the host chipset, the PCI Express I/O interconnect covers support for moving peripherals from the traditional PCI bus to PCI Express ports exposed via new add-in card connectors or alternatively available to devices to be integrated into the system board.

In the remainder of this section we focus on the two new add-in form-factors that are enabled by PCI Express technology.

PCI Express MiniCard Electromechanical (CEM) Form-Factor

The PCI Express Mini CEM add-in card shown in Figure 9 provides for expanding and upgrading the communications capabilities of the mobile platform. Replacing the conventional Mini PCI add-in card, this form factor is roughly half the size and is based on PCI

Express x1 and USB 2.0. The specification of add-in card and socket definition is given in the PCI-SIG.

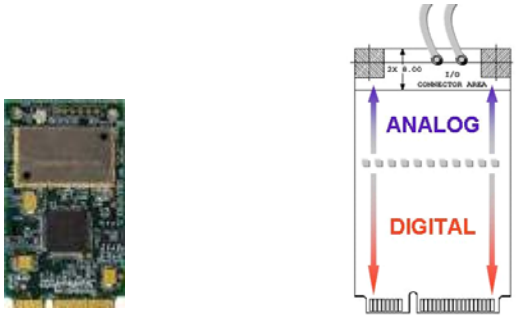


Figure 9: PCI Express mini CEM form factor

Both wired and wireless communications applications are comprehended by the Mini CEM specification. Examples include Local Area Network (LAN, e.g., 10/100/1000 Mbps Ethernet), Wide Area Network (WAN, e.g., V.90/V.92 modem), Wireless-LAN (W-LAN, e.g., 802.11b/g/a), Wireless-WAN (W-WAN, e.g., cellular data), and Wireless-Personal Area Network (W-PAN, e.g., Bluetooth*). Although not specifically considered, other applications may also find their way to this form factor.

BTO/CTO Form Factor

The PCI Express Mini Card specifically targets addressing system manufacturers' needs for build-to-order (BTO) and configure-to-order (CTO) applications rather than providing a general end-user-replaceable module. This form factor has characteristics more typical of an "embedded" application including the platform integration of the media interfaces such as communications connectors or wireless antennas.

The Mini Card and host socket is based on a single 52-pin card-edge type connector for its system interfaces. The host system connector is similar to a Small Outline Dual In-line Memory Module (SO-DIMM) connector and is modeled after the Mini PCI Type III connector but without side retaining clips, rather relying on two card-retention mounting points at the opposite end of the card. The placement of I/O connectors on a Mini Card is at the end opposite of the system connector. Depending on the application, one or more connectors may be required to provide for cabled access between the card and media interfaces such as LAN and modem line interfaces and/or RF antennas.

With both PCI Express x1 and USB 2.0 interfaces defined for the Mini Card and given the BTO/CTO nature of the application, the OEM has the option to only accommodate one of these interfaces in a given platform or socket. The OEM also dictates what I/O connector interfaces are built into the platform. Card vendors are obligated to work closely with each OEM to ensure that the application that they are designing will function properly in the target socket. Due to the mobile focus for this form factor, all of the advanced power-management features of these interfaces are expected to be used aggressively including support for PCI Express ASPM.

Application-Specific Interface Features

The PCI Express Mini CEM specification defines a number of system interface features helpful in the implementation of communications applications. These include defining a Light Emitting Diode (LED) indicator interface, a wireless radio transmitter disable control, remote UIM/SIM socket support, and System Management Bus (SMBus). Most of these features are specific to wireless applications.

The LED interface provides for three separate indicators, one each associated with WLAN, WWAN, and WPAN technologies. A recommended usage model covers indications for various states of radio operation including on, off, searching/associating, and transmit/receive activity. To support emerging requirements for controlling wireless RF transmitters in areas where interference may be a potential hazard, such as on commercial aircraft, a wireless disable signal is provided for so that the OEM can implement an on/off switch in the platform that directly controls the radio's ability to transmit.

For WWAN applications, an ISO standard UIM/SIM socket interface is defined to allow the OEM to support the SIM card needs for GSM/GPRS network account security and management. Due to the small size of the Mini Card, SIM card support on the radio card may not be practical.

The ExpressCard Standard Form Factor

Following the hot plug-and-play usage model successfully established by CardBus PC Card modular add-in cards for mobile platforms, the PCMCIA industry group has developed a PCI Express-based replacement technology released as the ExpressCard Standard. This new technology replaces conventional parallel buses for I/O devices with scalable, high-speed, serial interfaces, either PCI Express x1 for high-performance applications, or USB 2.0 to take advantage of the wide range of USB

* Bluetooth is a trademark owned by its proprietor and used by Intel Corporation under license.

silicon already available. Whichever interface is used; the end user experience is the same.

Module Form Factors

The ExpressCard Standard is designed to deliver high-performance, modular expansion to both desktop and mobile platforms in a lower cost, smaller form factor. Users are able to add a wide variety of applications including memory, wired and wireless communications, multimedia, and security features by inserting ExpressCard modules into compliant systems. At roughly half the size and lighter than today's PC Card, ExpressCard products also leverage the proven advantages of PC Card technology, including reliability, durability, and expansion flexibility while offering improved performance.

There are two standard formats of ExpressCard modules shown in Figure 10: the ExpressCard/34 module (34 mm x 75 mm) and the ExpressCard/54 module (54 mm x 75 mm). Both module formats are 5 mm thick, the same as the Type II PC Card, and at a standard length of 75 mm, are 10.6 mm shorter than a standard PC Card. A common, 26-pin, beam-on-blade style connector designed for high durability and reliability is used for both module formats, and the corresponding host connector accommodates the insertion of either module. The host slot for the ExpressCard/54 module features a novel guidance mechanism that also supports ExpressCard/34 modules by steering the narrower module into the connector socket. In any host system that implements multiple slots, all slots provide equivalent I/O interface functionality.

Figure 10: ExpressCard modules

The two ExpressCard module sizes give system manufacturers greater flexibility than in the past. While the ExpressCard/34 device is better suited to smaller systems, the wider ExpressCard/54 module can accommodate applications that do not physically fit into the narrower ExpressCard/34 form factor such as Smart Card readers, CompactFlash readers, and 1.8" disk drives. With its extra space for components and for spreading thermal energy, the ExpressCard/54 module format is also a natural choice for higher-performance and first-generation applications. The specification also allows for extended module formats to integrate features such as LAN and phone line connectors, or antennas for wireless cards into products.

Hot-Plug Functionality and Power Management

Leveraging PCI Express and USB built-in support for hot-plug functionality, ExpressCard technology is designed to allow users to install and remove modules at anytime, without having to switch off their systems. By relying on the auto-detection and configuration of the native I/O buses, ExpressCard technology can be implemented on a host system without an external slot controller. A device to control power to the slot is required, based on a simple, wired, module presence detection scheme.

Both PCI Express and USB natively support features that enable module applications to be placed in very low power states while maintaining the ability to detect and respond to wake-up requests. For example, an ExpressCard application can receive network messages via a wireless communications module even while the PC is in a sleep state. Additionally, PCI Express reference clock on/off control during ASPM L1/L2 states is supported. Effective use of these features is the key to creating high-performance applications that are both power and thermally efficient.

ExpressCard technology will typically require less electrical power than products built on previous PC Card standards although the slot power specification does provide adequate peak current to meet the needs of applications such as wireless transmitters. Independent of the amount of power drawn from the host system, the ExpressCard Standard also specifies thermal power dissipation at a maximum of 2.1 W for modules. Thermal dissipation limits are specific to the heat released within the slot discounting the energy that is dissipated in module extensions outside the confines of the slot.

User Benefits

Users will be able to identify modules and host systems that are compliant with the ExpressCard Standard by looking for the licensed ExpressCard logo, an energetic rabbit signifying mobility, fast performance, and ease-of-use, either directly on the product, in product documentation, or in the ExpressCard Compliant Products Directory available online at expresscard.org (see Figure 11). The compliance program intends to ensure interoperability between ExpressCard modules and systems using a two-step process consisting of self-compliance testing against a comprehensive requirements checklist and formal interoperability testing between modules and host systems. For product manufacturers, the compliance process also covers key ingredients such as

the connectors and power switch circuits used in building ExpressCard modules and host systems.

Figure 11: ExpressCard logo

By supporting both PCI Express and USB in all compliant module slots, ExpressCard technology brings new functionality to computer users not found in today's PC Card. The technology delivers a consistent, easy, reliable, and non-threatening way to connect devices into their systems. ExpressCard modules can be plugged in or removed at almost any time. With an expected strong adoption in desktop systems, driven by the desire to move these platforms to a sealed-box expansion model, users will find that they will be able to move modules between their desktop and mobile platforms so that they can easily share useful resources and transfer data between the platforms.

PCI EXPRESS MICROARCHITECTURE IN MOBILE PLATFORM

In next-generation mobile platforms, the Intel 915 family of chipsets, the GMCH and ICH incorporate a PCI Express-based interconnect. These two hubs form the RC of the platform. This means they connect a host CPU/memory subsystem to a PCI Express hierarchy. The GMCH contains a PCI Express x16 port that has been designed and optimized for high-end graphics usage. There are also 4x1 general-purpose PCI Express ports that that can be connected to form the rest of the PCI Express fabric such as PCI Express Mini CEM sockets, ExpressCard slots, and system board PCI Express devices that cover communications and storage applications.

PCI Express fabric in next-generation mobile platform microarchitecture implements multiple "virtual channels" that are described in the PCI Express specification-SIG. They are used in order to provide a reliable service for both bandwidth and latency. These virtual channels are controlled and assigned to different peripherals on the platform. Depending on the system needs, they are designed to utilize the system resources, CPU, and memory between I/O connections.

Due to the high-speed serial nature of the PCI Express links on the next-generation mobile platform, power management became one of the focus areas. The microarchitecture of the RC implements aggressive power-control policies to keep the consumed current to a minimum while performing to system needs.

PCI Express Microarchitecture Partitioning

The PCI Express architecture is specified in layers as shown in Figure 2. Compatibility with the PCI addressing model load-store architecture with a flat address space is maintained to ensure that all existing applications and drivers operate unchanged. The PCI Express configuration uses standard mechanisms as defined in the PCI Plug-and-Play specification. The software layers generate read and write requests that are transported by the TL to the I/O devices using a packet-based, split-transaction protocol. The link layer adds sequence numbers and CRC to these packets to create a highly reliable data transfer mechanism. The basic PHY consists of a full-duplex channel that is implemented as a transmit pair and a receive pair for each lane. The initial speed of 2.5 GHz (250 MHz internally) results in 2.5 Gb/s/direction that provides a 250 MB/s communications channel in each direction (500 MB/s total).

PCI Express uses packets to communicate information between components. Packets are formed in the TLs and DLLs to carry the information from the transmitting component to the receiving component. As the transmitted packets flow through the other layers, they are extended with additional information necessary to handle packets at those layers. At the receiving side the reverse process occurs, and packets get transformed from their PHY representation to the DLL representation and finally (for TLPs) to the form that can be processed by the TL of the receiving device.

Power Management Policies

Beyond supporting the already defined Active Link Power states, the next-generation mobile platform implements architecture-specific, power-saving options and policies. These policies in some cases take full advantage and aggressively pursue ASPM states as well as take advantage of notebook form factors.

Aggressive L0s Entry

PCI Express links on the next-generation mobile Platform pursue a quick entry to low-power ASPM states, especially while deciding on L0 to L0s transition. With this aggressive entry policy, in order to minimize the latency issues, the links also implement a very quick exit time in order to service the incoming requests. This policy is observed to be not only helpful in keeping the link power low, but also in packetizing the traffic to increase the ASPM state residency times.

Compressive Core Controls for Analog Logic
Due to the high-speed nature of PCI Express and clock recovery logic, most power on PCI Express links is being consumed at the interface level where analog controls reside. The next-generation mobile platform implements comprehensive controls from the core logic to minimize the use of power both during the active states as well as ASPM states. Some of these controls are outlined below:

- *Single Squelch Detection*: In the next-generation platform, built on Intel Centrino mobile technology, both during L0s and L1 ASPM states, PCI Express link controls enable single-lane squelch detection for exits. This feature reduces the unnecessary use of squelch detection logic across the links and saves significant power for wide PCI Express links.
- *Low-Power Circuits*: In order to save power during active states, PCI Express interfaces use optionally low-voltage swings with reduced drive strengths. For close interface links it is proven to produce significant power savings where the driver current is reduced to half (from 20 mA to 10 mA).

Clock Gating Microarchitecture
The Intel 915 family of chipsets in next-generation mobile platforms uses aggressive clock-gating policies and domains to minimize unnecessary switching across the micro-architecture. In previous microarchitectures, clock gating has been a proven method to reduce power by eliminating unnecessary toggles. For next-generation mobile platforms, we expanded the methodology by grouping different sections of the microarchitecture into separate clocking domains. These domains include

- Separate clocking for different layers.
    - Transaction Layer clocking.
    - Link Layer clocking.
    - Physical Layer clocking .
        - Logical Sub Block.
        - Electrical Sub Block.
- Separate Clocking for RX (receiver) and TX (transmitter).

All the listed clocking sections of the microarchitecture are controlled via a list of indicators that signals whether the block in question is needed or not.

All these power-saving features are included in the next-generation mobile platform to increase the battery life and reduce the thermal envelope on the system. They are implemented in such a way that keeps the balance between a very high-performance I/O connection system, like the PCI Express, and a very power-conscious mobile platform.

SUMMARY

The emergence of PCI Express technology as a unified foundation of I/O, graphics, and networking interconnections for the next generation of mobile platforms built on Intel Centrino mobile technology has been presented. PCI Express architecture as a state-of-the-art platform I/O solution of choice has been discussed with many values such as high performance, high-bandwidth scalability, easy networking connections, improved power management, interoperability, network reliability, native hot-plug capability, and software compatibility. In this paper we focused on areas such as integration of PCI Express technology into the next generation of mobile platforms, by describing the PCI Express architecture and protocols, PCI Express power management, advanced RAS and QoS features, PCI Express mobile applications such as graphics, networking, and form factors including ExpressCard standard, the PCI Express Mini Card as well as future form factors specifications such as the PCI Express Wireless Card and the PCI Express Cable. We also focused on ASPM capability that enables aggressive power-management within the link layer as defined in the PCI Express Base Specification. The next-generation Intel Centrino mobile platform based on the Intel 915 family Express chipsets has also been described with integration of PCI Express technology.

ACKNOWLEDGMENTS

We would like to acknowledge the encouragement and help that we received from our managers, particularly Valerio Jim, Ajay Bhatt, and Bala Cadambi. Additional thanks go out to individuals who reviewed the paper and provided valuable feedback: Ramin Neshati, Dave Puffer, and Aditya Sreenivas.

REFERENCES

[1] PCI Express Base Specification, 1.0a. Portland, OR: PCI-SIG.

[2] PCI Express card Electromechanical (CEM) Specification, Portland OR: PCI-SIG.

AUTHORS' BIOGRAPHIES

Mohammad Kolbehdari is a senior staff hardware engineer in the Digital Enterprise Group. He received B.Sc., M.Sc., and Ph.D. degrees in Electrical Engineering from Communication University Tehran, Iran, University of Mississippi Oxford MS, and Temple University, Philadelphia, Pennsylvania in 1986, 1991, and 1994,

respectively. He conducted research in the EM lab at Temple University prior to joining the Department of Electronics, Carleton University, Ottawa, Canada as a research associate faculty member from 1994 to 1998. He joined Intel in 1998 and has been involved in many product development teams including the Pentium® 4 FSB and PCI Express electrical designs. His areas of interest are EM modeling, power delivery, signal integrity, I/O serial differential design, system-level architecture, platform initiatives, and clock jitter characterization and measurements. His e-mail is mohammad.kolbehdari at intel.com.

David Harriman is a senior staff platform architect in the Digital Enterprise Group. He develops and promotes next-generation platform technologies, focusing on I/O interconnects. He has been a leading contributor to the PCI Express specification since its inception. He joined Intel in 1993, and he worked on the development of the first desktop chipset for the Pentium Pro processors and on the first AGP chipset. He was one of the primary developers of the Intel Hub architecture. He received a Bachelor's degree with dual majors in Electrical Engineering and Computer Science from the University of Wisconsin, Madison in 1993. He holds more than 20 patents relating to chipsets and interconnects. His e-mail is david.harriman at Intel.com.

Altug Koker is a senior staff engineer in the Mobility Group. He received a B.Sc. degree in Electrical Engineering from Bilkent University, Ankara Turkey in 1995 and an M.Sc. degree in Electrical and Computer Engineering from State University of New York at Buffalo in 1997. Altug conducted research on micro electronics while he was at Bilkent University, Ankara prior to joining the Department of Electronics and Computers at the State University of New York at Buffalo where he researched the design of high-performance computers and networks. He joined Intel in 1997 and has been involved in many chipset design projects including the initial Accelerated Graphics Ports (AGP), Pentium III and 4 FSB, and PCI Express ports. His e-mail is altug.koker at intel.com.

Seh Kwa is a staff platform architect in the Mobility Group of Intel Corporation. He joined Intel in 1998 as part of EPG 460GX chipset team before focusing on low-power Intel architecture and power management for the last four years. Prior to joining Intel, he was involved in low-power Digital Signal Processing as well as communication development at two startup/IPO companies. He received B.S.E.E., B.S.C.S, M.S.E.E. and D.Sc. E.E. degrees from Washington University. His e-mail is seh.kwa at intel.com.

Brad Saunders is a senior mobile systems architect focusing on I/O chipset definition in Intel Corporation's Mobility Group. Brad also leads the PCMCIA industry group responsible for the ExpressCard Standard and is the technical editor for both the PCI Express Mini and Wireless Form-Factor (WFF) electro-mechanical specifications within the PCI-SIG. Brad came to Intel Corporation as part of the Xircom, Inc. acquisition in early 2001, where he had spent two years in Xircom's chief technology office. Prior to that, he spent 21 years working in various fields of communications at Rockwell International, ranging from secure communications for defense applications to analog modems for mobile systems. Brad received his B.S.E.E. degree from the University of California, Irvine in 1976. His e-mail is brad.saunders at intel.com.

® Pentium is a registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

THIS PAGE INTENTIONALLY LEFT BLANK

High-Performance Graphics and TV Output Comes to the Second-Generation Intel® Centrino™ Mobile Technology Platform

Kam Leung, Mobility Group, Intel Corporation
Steve Spangler, Mobility Group, Intel Corporation
Todd Witter, Mobility Group, Intel Corporation
Kevin E. Arendt, Mobility Group, Intel Corporation
Satya Avadhanam, Mobility Group, Intel Corporation
Satyaki Koneru, Mobility Group, Intel Corporation
Val Cook, Mobility Group, Intel Corporation

Index words: 3D graphics, 3D graphics acceleration, memory controller hub, TV output, image filtering, color space conversion

ABSTRACT

This paper considers two major enhancements to Intel's graphics memory controller hubs, as embodied in the Intel® 915 Express Chipset Family. First, we discuss the microarchitecture of the 3D pipeline and the steps taken to optimize it for peak performance. Secondly, we present the TV output feature, which is important for merging the personal computer and the television into a single platform, in support of Intel's digital home initiative.

In the 3D graphics section, we present an overview of the microarchitecture in the context of the design challenges inherent in a next-generation integrated graphics accelerator. This section demonstrates the key roll benchmark analysis played in optimizing the performance of various components of the graphics pipeline such as command processing, primitive processing, pixel shader floating point units (FPUs), caching algorithms, etc. Obtaining a balanced pipeline is central to achieving maximum performance.

With the merging of the personal computer and the television into a single platform for the digital home, new requirements are emerging for the standard home PC. High among these features is the inclusion of the television output interface or the digital television encoder. Television signals are quite different from the standard computer monitor connections and therefore have very different requirements. The TV output function also needs to be highly flexible because television encoding algorithms are different depending on the country of origin.

3D MICROARCHITECTURE INTRODUCTION

Significant feature and performance enhancements have been incorporated into Intel's third-generation graphics processing unit. The performance of these platforms generally exceeds expectations, with the Intel 915 Express Chipset Family demonstrating approximately 2.4 times the performance of its predecessor, as measured using Future Mark's 3DMark*2001 [1] benchmark, and approximately 10 times the performance using the 2003 version of their benchmark [2].

This paper presents the key architectural challenges associated with attaining these performance levels; we intend that our methodologies and conclusions will provide valuable insight for other design projects. A cursory understanding of the state of the personal computer graphics industry is necessary to appreciate the task presented to the graphics team.

3D Graphics Background

The graphics and media industries can be segmented into many categories, but a few of the more defining characteristics are directly related to the device's memory subsystem and its cost, considering both power consumption and die area. Intel's integrated approach places the graphics accelerator and TV output logic on the same die as the Memory Controller Hub (MCH), where a single memory subsystem is shared with the CPU. Memory access arbitration policies are consciously tuned to favor the CPU, resulting in increased access latency for the graphics operations. Additionally, manufacturing costs and power consumption constraints severely limit the die size. However, discrete add-in cards generally have local memory subsystems and larger die size budgets, due to the premium price obtained for these devices. Working within these constraints is the overriding consideration in the development of the architecture.

Additionally, the graphics industry is at a crucial inflection point in the development of consumer-level graphics acceleration. All of the lighting and rasterization calculations were previously performed by fixed-function logic, with relatively few parameters, but are now being replaced with programmable execution units. These units are commonly called "shaders" implying the process of shading or lighting the rendered primitives. A Pixel Shader unit is programmed using an assembly language similar to a CPU. The transition from fixed-function logic to programmable shaders makes it much more difficult to select a single performance *sweet spot* and optimize the architecture for that point. The concept of the sweet spot has vanished in favor of many workload-dependant optimization points.

3D Microarchitecture Overview

A brief discussion of the graphics architecture aids in understanding the specific design challenges that are considered in detail in subsequent sections. From a system-level point of view a graphics application performs several tasks prior to engaging the graphics accelerator to synthesize an image. Initially, various surfaces are populated in memory, such as texture maps, constant buffers, and vertex buffers. Many state control structures are also created in memory, including compiled versions of the pixel shader programs. It is not uncommon for all of these entities to require 100's of megabytes of storage.

The application then begins the rendering process by making specific DirectX* [3] or OpenGL* [4] API calls to draw primitives on the render target surface. These commands are placed in a ring buffer, and the graphics accelerator is instructed to begin processing the buffer. As shown in the simplified block diagram of Figure 1, the commands are retrieved from the memory subsystem and when applicable are presented to the graphics pipeline for processing.

Primitive Processing and Iteration

Primitive processing consists of computing various constant attributes of a primitive, such as its axis-aligned bounding box and coefficients required by the Iterators. The Iterators traverse the interior of a primitive, such as a triangle, and produce interpolated values for each pixel location contained within the bounds of the primitive. It is important that the throughput of the primitive processing stage be such that the Iterators, which contain a significant amount of custom floating-point logic, are not starved for input. Correspondingly, the computation capability of the Iterators must be sufficient so as not to starve the even more expensive Texture Sampler and Pixel Shaders.

Dispatch

These interpolated values may then be used directly by the Pixel Shaders or presented to the Texture Sampler, as determined by the pixel shader program and managed by the Dispatch unit. The Dispatch unit also arbitrates sample requests from the Pixel Shader, where a programmatically generated sample location has been determined, rather than simply interpolated. This arbitration and scheduling task is particularly challenging.

Texture Sampler

When a sample request is delivered to the Texture Sampler one of texture maps that was initially loaded by the application is accessed and the resulting data are optionally filtered and presented to the Pixel Shaders. The Texture Sampler employs a multilevel caching architecture where the second-level cache is significantly larger and less capable than the L1 cache. The necessary format conversions, data swizzling, and output bandwidth of the L1 cache make it substantially more expensive than the L2 cache. Maximizing the utilization of the L1 cache and its supporting structures is essential to balancing the performance against the area and power consumption of the device.

* Other brands and names are the property of their respective owners.

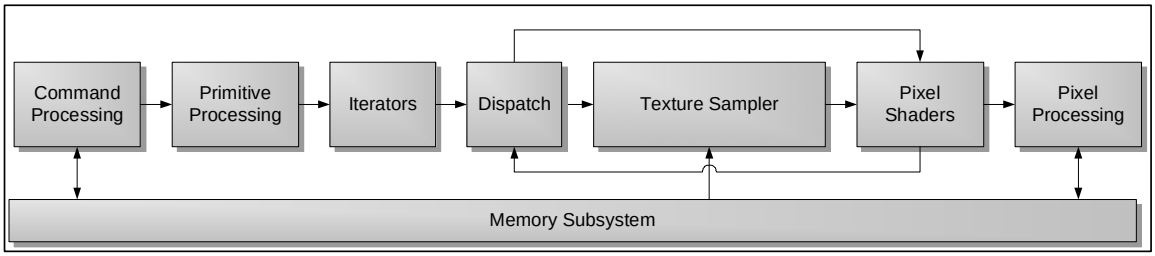


Figure 1: 3D graphics pipeline

Pixel Shaders

The Pixel Shaders consist of many floating-point units (FPUs) that are controlled in a lock step fashion. This Single Instruction Multiple Data (SIMD) approach has the advantage of being able to amortize the cost of the control logic across multiple FPUs, but it also reduces the overall efficiency of the more expensive resources in the architecture.

Pixel Processing

Finally, the Pixel Processing unit performs many fixed-function operations such as format conversion, color space conversion, and alpha blending. The resulting pixels are also assembled and delivered to the memory subsystem in this unit. The throughput of these operations is significant because the overall pixel fill-rate, one of the key performance metrics, is ultimately determined by this unit. However, when complex shading or large amounts of texture sampling is occurring this logic can be significantly underutilized.

Workload Visibility and Analysis Tools

In the presence of an industry-wide inflection point in the graphics programming model, requiring a substantially more dynamic hardware architecture, the need to have visibility into key workloads is essential. This necessarily results in the enhancement and development of several new tools and analysis capabilities. We focus on the following three primary areas.

Enhanced Driver Capture Capability

The ability to run an actual application or benchmark against the graphics simulator is achieved by capturing the stream of commands presented to the driver. Additionally, memory resident structures, such as texture maps, vertex buffers, and shader programs are also captured. These data are written to disk in a standardized binary file format that is then used as stimulus for our functional simulator. The simulator can also be installed as a virtual device, thus eliminating the file I/O from the simulation.

Timing Approximation from Functional Simulation

Our primary simulator provides very precise functional modeling, but contains no concurrency. This presented us with a significant challenge: how could the expected performance of a captured workload be predicted without a timing-aware simulator? Traditional instrumentation approaches are inadequate for our needs; however, by approximating the design as a fully pipelined, directed graph of logical units we could instrument the functional simulator to provide valuable timing information. Each unit is assumed to have an infinitely deep FIFO between it and the adjacent units in the graph. If a completely homogenous stream of work were flowing through the graph, and the minimum time that each quantum of work could reside in each unit were known, the throughput of the system would be limited by the slowest unit.

Unfortunately, real graphics workloads are not homogenous and each unit does not operate on the same input data. However, by selecting a quantum of work that naturally flows through the data path of the design, one can accumulate the time required to operate on each unit of work across multiple inputs. We chose to use a single primitive as our quantum of work. The functional simulator was then instrumented to determine the minimum amount of time that each primitive could reside in each unit. The maximum time required by any unit was then determined to be the pipeline throughput for that quantum of work. The sum of all of the quantums of work represents the theoretical peak processing time for the workload. In practice the time required to service each primitive varies radically and the FIFOs between the units are of finite size. However, we found this approach to correlate with compute bound workloads reasonably well. More traditional analytical techniques are used when the memory bandwidth is the dominant factor in the performance.

Database Driven Static Analysis

A single benchmark requires days to execute on the functional simulator. When changes to the statistics-gathering instrumentation or "what if" scenarios are desired many days of reacquiring the information from the

simulator are required. In order to minimize this impact we further instrumented the simulator to export control state, primitives, and timing information with primary keys. These data are then imported into a relational database, containing millions of rows in some tables. Relatively simple queries produced information in seconds, that would have otherwise required several days of simulation time. Additionally, stored procedures were used to perform complex computations that modeled proposed changes to the design, such as various state caching mechanisms and performance improvements in specific units. Most of these more complex queries completed in less than an hour.

These tools and capabilities assisted in providing valuable information for many of the challenges faced by the architecture team.

GRAPHICS ARCHITECTURAL CHALLENGES

The major architectural challenges of the Intel 915 Express Chipset Family, as briefly addressed in the overview section, are divided into six categories corresponding to the basic units of the block diagram in Figure 1. Each of these challenges are addressed in the following sections.

Choosing the Optimal Primitive Processing Rate

The time required to process a primitive, independent of its area, consists of a fixed overhead time and a per attribute computation time. The computations associated with the fixed overhead can be pipelined, but because there are varying numbers of attributes on a vertex, such as its position, diffuse color and texture coordinates, etc. the computational resources are reused. This is traditionally accomplished by re-circulating each attribute through the same logic, thus preventing efficient pipelining, but significantly reducing the cost. For example, the previous generation could process a primitive with a position, color, and 2D texture coordinate attribute at each vertex in approximately 50 clocks. This relatively large latency can be hidden if the primitive lights a sufficiently large number of pixels. However, today's graphics benchmarks and other important workloads consist of many very small triangles that do not consume much time processing pixels. These primitives are said to be "setup limited."

Our workload analysis suggests that even at our chosen processing rate, an average of 68% of the primitives are potentially setup limited. However, these same primitives represent only 7% of the total time required to render the workload, due to their small area. The challenge then becomes deciding how much computational logic to enlist in order to improve a relatively small portion of the overall rendering time.

As we considered primitive processing designs similar to our previous generations the workload analysis indicated that approximately 17% of the rendering time would be spent on setup limited primitives. Realizing the relatively large performance penalty associated with these primitives and the trend toward smaller and smaller primitives we chose to improve the throughput of the per attribute computations to one clock/attribute, thus attaining a theoretical 10% improvement over previous generations.

Selecting the Iteration Rate

The Iterators determine the pixels that are lit within a primitive. It implements a walking algorithm that minimizes the pixels that are considered for testing. This testing is carried out at the rate of 16 pixels/3 clocks to sustain the required throughput of 4 pixels/clock. Once the lit pixels are determined, an early depth test, whenever possible, is performed to eliminate the covered pixels. This significantly reduces the workload on the remaining part of the pipeline. The rate for these operations provides sufficient performance that these operations are rarely the bottleneck in the pipeline. However, determining the rate at which the vertex attributes should be interpolated is a significant design decision.

Because the interpolation operations and the shading operations occur concurrently, it is desirable to have the throughput of both units matched in order to achieve the maximum performance given the available die area. However, there is not an inherent relationship between these two operations. It is very possible to require few interpolated attributes while utilizing many shader instructions to produce the output. This case is typical of procedural shaders, such as wood, marble, and water, where the output color is determined algorithmically. However, it is also possible to require many interpolated values that utilize very few instructions. These shaders tend to drive much of the computation with tables or texture maps. Filtering techniques that require large filter kernels also tend to fall into this category.

Various shaders from prominent workloads were captured to obtain the number of attribute components to be interpolated vs. the number of arithmetic instructions. We considered providing four or six interpolators per pixel, which all operate in parallel. These interpolators require floating-point precision and are relatively expensive in terms of die area and power. However, they are not as expensive as the FPUs used in the Pixel Shaders. Each line in Figure 2 identifies the difference in the interpolation and Pixel Shader rate, normalized to

pixels/clock. When the value for a given shader is positive the interpolation rate exceeds the shader rate.

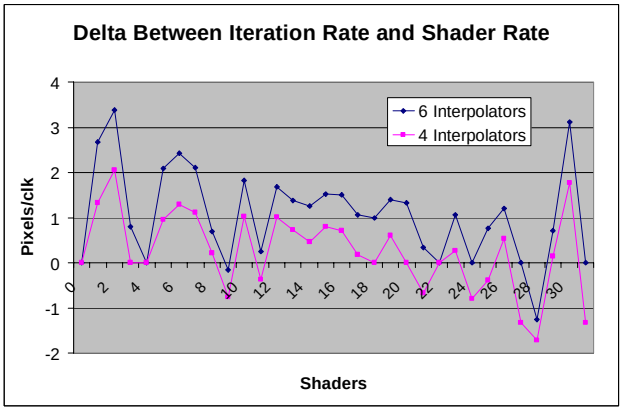


Figure 2: Difference between the iteration rate and the shader rate for two interpolator design options

As can be seen from the graph in Figure 2, four interpolators perform reasonably well; however, the penalty for underutilizing the Pixel Shaders is also substantial. Additionally, the 4-pixels/clock throughput was desired for the very typical scenario of a 4D diffuse color and a 2D texture coordinate. Thus, the number of interpolators was determined to be six. This also meant that the capability to process two coordinate sets at a time was needed. Any combination of a (1-4)D coordinate set and a (1-2)D coordinate set is allowed to be processed in parallel.

Determining the Dispatch Output Bandwidth

The interpolated coordinate sets may be used as is and/or passed through the Sampler before being delivered to the Pixel Shaders. Having previously decided to generate two coordinate sets every clock during iteration, the dispatcher was designed to deliver one coordinate set to the Sampler and one coordinate set directly to the Pixel Shader (through a bypass FIFO) simultaneously. In the case of dependent read scenarios, there are additional source coordinate sets produced by the Pixel Shaders. This allows the pipeline to operate in a balanced fashion.

Maximizing the Texture Sampler Utilization

Texture sampling is central to the overall performance of the graphics accelerator, and the caching microarchitecture is the heart of the Sampler. These structures are discussed in detail in the following sections.

Determining the L1 Texture Cache Read Port Count

The Intel 915 Express Chipset Family expanded the texture sampling engine to four pixels wide, from two pixels wide in the previous-generation sampler. The four pixels are always adjacent, forming a 2x2 orientation called a pixel quad. The width in pixels of the sampling engine determines, among other things, the amount of data that needs to be read from the L1 texture cache in parallel. Adding ports to the cache is expensive, so we needed to identify a way to get most of the performance benefit of a 4-pixel sampling engine without the cost.

Through the simulation of workloads, we determined that in most cases there is significant overlap of the texture data (texels) that the adjacent pixels fetch. Thus it is not necessary to have a dedicated independent port for each pixel on the L1 cache to efficiently fetch the required texel data.

A single-ported L1 cache was also considered, but the performance tradeoff was deemed too high. The following graph shows the relative efficiency of the 2-port and 1-port caches. The 4-port cache has 100% efficiency by definition.

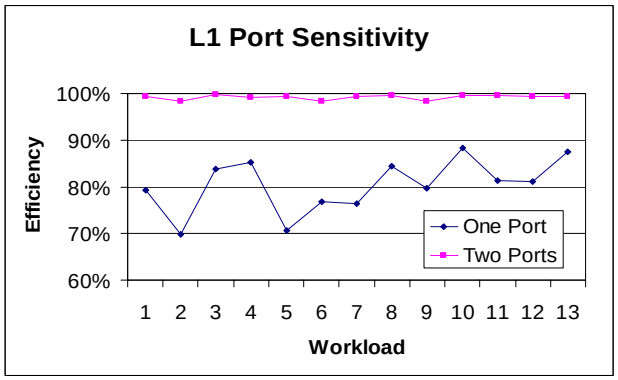


Figure 3: L1 texture cache read port sensitivity

Specific control logic checks each incoming pixel quad to determine whether it requires more than two cache ports. This process involves comparing the addresses of the data requested for each pixel. If the data can be retrieved through two cache ports, the pixel quad is allowed to proceed unmodified. If more than two ports are required, the quad is split into two 2-pixel packets.

Pixel to Port Mapping

The next consideration was, in cases where the pixel quad does not need to be split due to the above considerations, to determine which pixels are mapped to each of the two cache ports. Increased flexibility generally means increased cost, and this case is no different. Information on which port each pixel is using needs to be sent downstream to enable the correct data to be routed out of the cache. Full flexibility means that many bits indicate the port/pixel mapping.

Further performance simulations, again using workloads, resulted in a decision to allow only two port/pixel mappings, termed “horizontal” and “vertical.” Horizontal indicates that horizontally adjacent pixels share a port; vertical means vertically adjacent pixels share a port.

Cases in which neither of the allowed mappings work are split into two packets of two pixels each. This is a performance/cost tradeoff. The addition of more flexibility than this does not significantly increase performance and does not justify the additional cost. The relative efficiency of this approach exceeds 90%, averaged across the workloads.

The graph in Figure 4 shows the fully flexible case (equivalent to the "two ports" case in the prior graph), a fixed mapping case where the port/pixel mapping is fixed, and the compromise option that we ended up choosing.

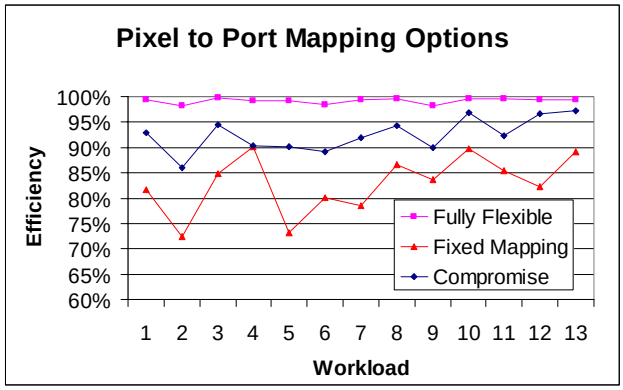


Figure 4: Various pixel-to-port mapping options

Line Size

The size of the L1 texture cache needed to be doubled; however, there were several approaches that we considered in order to accomplish this. The number of ways could be doubled, the number of sets could be doubled, or the line size could be doubled. We used workloads under simulation to compare doubling the line size to doubling the number of ways. The number of sets is tied to the ports, so we did not consider that vector. The analysis indicated that doubling the number of ways provided a better cost/performance tradeoff.

The graph in Figure 5 compares the miss rate on a 32-byte vs. a 64-byte cache line size. The overall cache size is the same for both cases. The 32-byte case simply doubles the number of cache lines. Although the miss rate is generally higher for the 32-byte cache size, the 64-byte line size has double the bandwidth requirement per miss, which became the driving factor in our decision. We selected the smaller cache line size with twice the number of lines.

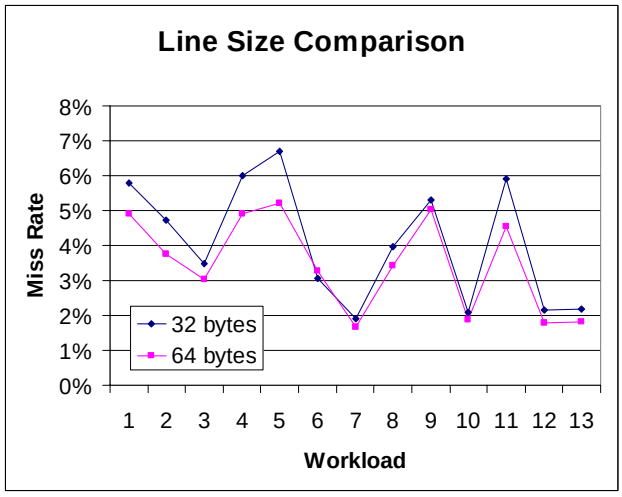


Figure 5: L1 cache line size comparison

Determining the Pixel Shader Throughput

Based on the ratios in the pixel-shader programs between the number of interpolated coordinates and the number of arithmetic instructions, it was determined to have eight FPUs in the Pixel Shader. The 2x2 pixel quad is the basic processing entity, these eight FPUs are organized as two sets of four FPUs. Each set of 4-FPUs operate in parallel, potentially executing different instructions. The four grouped FPUs operate in a SIMD fashion. The pixel quads are delivered to the two sets of FPUs in a ping-pong/alternating fashion and are output in the same manner to maintain a strict in-order behavior through the 3D pipeline.

The execution of the program is interleaved. Execution of an instruction on a pixel quad occurs only after the completion of execution of the previous instruction, thus eliminating data-dependency checking requirements. An empty slot in the FPU's pipeline is filled with a new pixel quad, if available. This multithreaded execution mechanism boosts the utilization significantly.

The sizing of the RAMs for storing the input, output, feedback, and temporary registers is determined by the register utilization in a pixel-shader program and the latency of the execution pipeline, as shown in Figure 6. To allow for streaming, additional storage is required for two cases: first, when a pixel quad's input payload is delivered and it is waiting to start execution and second, when a pixel quad has completed execution and is delivering its output payload. Taking this into consideration and the expectation that the register usage would grow over time we selected a storage size of 64 entries.

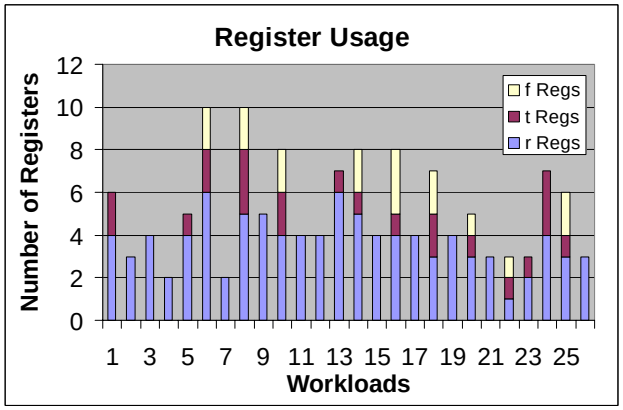


Figure 6: Register utilization

Determining the Pixel Processing Throughput

Pixel processing includes fog blending, alpha/stencil/depth testing, alpha blending, gamma correction, dithering, and logic operations. Our analysis suggests that on average 23% of the pixels rendered require fewer than six attributes to be interpolated, consisting of one texture sample, and two shader instructions. These data suggested that there is significant performance to be captured if the back-end pixel-processing functions also operate at 4-pixels/clock. To achieve this desired fill-rate four pixel-processing blocks operate in parallel on a pixel quad.

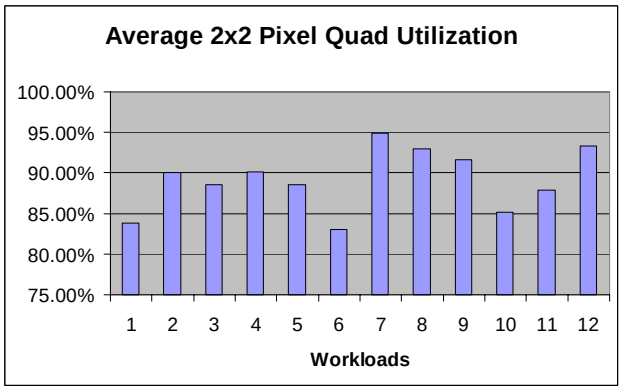


Figure 7: Average 2x2 pixel quad utilization

Grouping pixels in this fashion is somewhat inefficient, due to the fact that not all of the pixels may be lit, but will still consume the hardware resources. Our analysis indicates that this inefficiency is relatively small, as can be seen in Figure 7.

3D GRAPHICS RESULTS

The actual benchmark results indicate significant performance improvements over previous generations. Additionally, we outperform the two leading competitor's integrated devices.

Table 1: Benchmark results

Benchmark	865G	915G	915GM (200 MHz)	915GM (320 MHz)	Integrated Device A	Integrated Device B
3DMark01	159	1522	1054	1266	790	0
3DMark03	3100	5872	5101	5337	5224	4947

We credit much of this success to a design methodology that incorporates workload analysis as a key component of the decision-making process. Our experience suggests that investment in tools and infrastructure to provide workload-driven information during the architecture definition phase produces valuable results. The Intel 915 Express Chipset Family stands as evidence, amidst widely dispersed workload characteristics and an industry-wide programming paradigm inflection point, that advanced analysis tools and data-driven design methodologies are vitally important.

TVOUT

In recent years, more people are using their personal computers, or PCs, to play and store music and digital photos, and their PC's DVD drives to watch movies. Moreover, people are using their large-screen TV as a viewing device for movies, digital photos, and online gaming. Finally, the integration of digitally recording favorite TV programs onto a hard drive, combined with the television connection, has earned the PC a position in home entertainment systems. Advancements in networking have even made it possible to share the recorded programs with a remotely located PC. Such a wide variety of applications triggered the convergence of PC and consumer electronics. Initially this functionality was achieved by using an external television encoder device connected to a graphics accelerator, which is complex and expensive. The Intel 915 Express Chipset Family integrates the two functions to reduce the complexity and price.

The goal of this section of our paper is to define and explain the video encoder function and the required associated logic for the encoder to function in a PC environment. The design can be divided into three major sections, each having its own design requirements. When displaying computer graphics data on a television the data need to be converted to meet the more stringent requirements of the TV modulation.

Television video output is similar to computer video output in that it is composed of an analog signal with two parts. The first part is timing information that tells the display where to put the data on the screen, and the second

part is the image data encoded in one of the three output modes. Output modes are explained in detail later.

For the computer, the analog connection to the display device uses three connections utilizing red, green, and blue colors, commonly known as the RGB interface. In the RGB interface, color information is carried on three signals indicating the color intensities of red, green, and blue. This is convenient since that is (nearly) the same color space as the monitor and it is also (nearly) the native format of the data as they are stored in the computer's memory. The timing information is usually carried on separate signals for horizontal and vertical synchronization pulses; however, some monitors require the synchronization of information to be placed on the green channel. The vertical synchronization pulse indicates when to start a new frame of data; the horizontal synchronization pulse indicates when to start a new line. Vertical rates are usually in the 60-85 Hz range, but can go much higher. Horizontal rates depend on the total number of pixels being displayed and range from 30 kHz to over 100 kHz.

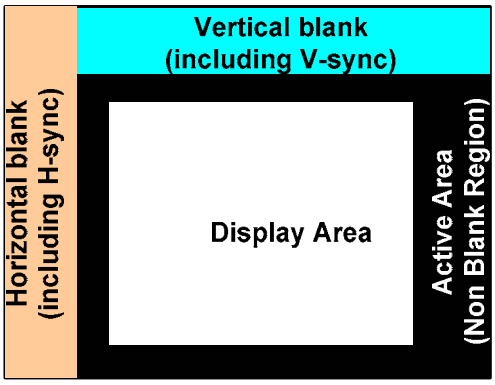


Figure 8: Television display timing areas

For connection to the television, originally all of the information (color and synchronization) was transmitted in just one signal (actually the signal was modulated onto the channel carrier and transmitted). This signal is called Composite Video Block Sync (CVBS) or usually just Composite Video. The data transmitted in this form use a different color space. The YUV color space represents picture information in Luminance (Y) and Chrominance (U and V). Luminance is simply the brightness portion of the image and was the only part transmitted until the advent of color television. The UV pair indicates the color with hue and saturation in vector form. The luminance information is encoded as a simple-level signal, and the chrominance information is encoded in a modulated carrier on top of the luminance. Different parts of the world use different modulation techniques.

More recently the composite signal has been separated out into two signals: this type of connection is called S-Video. In an S-Video connection, the luminance and chrominance are transmitted each on their own pair, eliminating crosstalk between the two signals. The luminance channel additionally carries the synchronization pulses.

In the last few years, televisions have appeared with a new analog connection that uses three pairs of wires called Component Video. This connection is much like S-Video except that the U and V signals are separated out onto their own pairs of wires, further improving the signal quality. Additionally the UV signals have their amplitudes scaled to match the luminance signal; this level matched interface goes by the convention YPrPb. Because of its higher quality, this type of signal is used for analog High Definition TV (HDTV) connections.

The actual timing signal and placement of the video data are different for the TV and PC video. The computer's video is drawn in line order or progressive scan. Each line of data on the screen is followed in time directly by the line beneath it. In most TV modes, the method is to draw all of the odd numbered lines in sequence followed by all of the even numbered lines. This method is called interlaced scan. The two groups of lines are called the fields: Field 1 has the odd lines while Field 2 has the even lines (counting from one, not zero). The two fields together compose a frame. In order to differentiate the two fields, the vertical synchronization pulse is skewed by a half line between Field 1 and 2 (but not between 2 and 1).

HDTV supports both interlaced and progressive video formats. The different modes of operation are typically indicated by adding a p or i to the number of displayed lines in the format, e.g., 480i, 720p, 1080i, etc.

PC Video is Not TV Video

There are several important differences to account for when converting PC video for use on a TV. The primary concern is the visual artifact known as "flicker" caused by the conversion from progressive scan to interlaced scan. Flicker is caused by the presence of single-pixel horizontal lines in the source material. When these lines are displayed on an interlaced monitor it appears to flicker on and off at the frame rate, because the line is only present in one of the two fields. These types of horizontal lines do not normally show up in normal video transmission because the source is generally a camera. However, since the computer video is a synthetic image, it is possible (in fact common) for a one-pixel tall horizontal black line to appear on a white background. Further, if the displayed area is in small font text (which usually has lots of little horizontal lines), the text can be difficult or impossible to read.

Another major difference is that the PC video is "underscanned" and the television video is "overscanned."

This means that the four corners of the PC's video image are visible to the viewer while the television allows the corners to fall behind the bezel of the monitor. In addition, many computer modes use pixel resolutions that are simply too big to fit on the television monitor in the standard definition mode.

One further unfortunate problem with television signals is that they are different depending on their country of origin. This is due in part to the European and Asian countries using a 50 Hz for power distribution while North America uses 60 Hz. The fallout of this is that the two systems are incompatible. Even worse is that many countries have developed their own versions of the two specifications resulting in a plethora of required formats. Fortunately, nearly all of the standards are similar enough to make it possible to generate them with one programmable logic block.

In summary, in order to correctly convert the PC's video signal for use in a television it is at the minimum necessary to perform three functions. The PC video color space must be changed from RGB to YUV; additionally it must be scaled to fit the television's display area accounting for underscan. Then that signal must be encoded into one or more of the three signal transmission standards mentioned above: Composite, S-Video, or Component. This last issue has two pieces: a digital portion and digital-to-analog converters (DACs). DACs are described in the next section.

Color Space Conversion

The television encoder works in the YUV color space, but computers operate in RGB color space so it is necessary to convert the RGB into YUV. Color space conversion is accomplished by putting the RGB data into a 3x3 matrix multiply. Unfortunately, not all television standards use the same YUV color space (or encoded space), so the converter needs to be completely programmable.

For NTSC Composite and S-Video

```
Y = 0.299R′ + 0.587G′ + 0.114B′
U = -0.147R′ - 0.289G′ + 0.436B′
V = 0.615R′ - 0.515G′ - 0.100B′
```

After color space conversion, there are two luminance controls that are typically made available to the user. These are *brightness* and *contrast*. Brightness is simply a constant value added to the luminance value. The contrast value is used to adjust the gain of the luminance. Similarly, there are two controls to chrominance. They are *saturation* and *hue*. Saturation adjustment manipulates the amplitude of the modulated chrominance, which is seen as color intensity. Hue adjustment manipulates the relative phase angle of the color component, which is seen as a color shift.

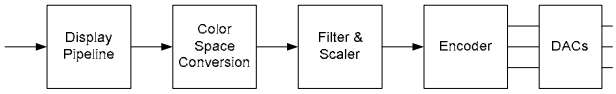


Figure 9: Data flow through the TVout pipeline

Scaling and Filtering

The conversion of progressive image data to interlaced data is essentially a scaling operation. The image is reduced to one half the original number of vertical lines. The simplest method to accomplish this would be to throw away every other line. However, this only works if the original image and the TV output image are the same size, which is not normally the case. So in this implementation a programmable scaler is placed in the pipeline to scale the incoming image to the TVout display size. This scaler can then also be used for the underscan correction (if that is desired) when displaying the PC content on a TV. There is a small one-half line offset in the scaling operation in the second field to account for the vertical offset between the fields and to preserve as much of the original data as possible.

At the same time the image is scaled for interlaced display, it should also be filtered to remove flicker. Flicker reduction is accomplished by softening the hard edges present in PC displays. Flicker reduction is a subjective feature, so several options are provided for users to select the amount of flicker or softness they desire.

The video filter is implemented as seven taps horizontally by five taps vertically, with each filter having its own set of coefficients. Five vertical taps are required in order to properly interlace the vertical data without causing flicker. The larger seven tap horizontal filter is required to properly filter the chrominance data before they are modulated. This prevents cross-color distortion.

Encoding

The final stage of the process before conversion to the analog signal is encoding. There are four major functions within the encoder. The timing generator is responsible for generating the internal timing reference signals as well as the external composite synchronization, which will eventually be added to the luminance channel data. It also generates timing signals that place the video data on the screen. The subcarrier generator is responsible for the proper generation of the reference subcarrier signal for the modulator and for ensuring a proper phase relationship to the horizontal synchronization. The luminance processor combines the composite synchronization and the Y data and generates the luminance signal that will be sent to the DAC. The chrominance processor block takes the subcarrier signal in along with the UV data and generates the chrominance signal; it also generates the subcarrier

burst signal, which is used by the television receiver to correctly demodulate the signal.

Creating a good subcarrier is of paramount importance to the generation of good quality video. The Intel 915 Express Chipset Family uses an 8x over-sampled clock when generating video for modulated formats. For NTSC, the color subcarrier is about 3.58 MHz; with a 1x clock of 13.5 MHz there would only be three or four samples per cycle. This is just barely enough to reconstruct the signal. At 8x over sampling (108 MHz) there are 30 or 31 samples per subcarrier cycle. This provides a very accurate representation of the carrier, which results in better color recovery at the receiver.

In order to generate quality output, the data output from the encoder is in a 10-bit per channel format requiring 10-bit DACs. The peak voltage of the CVBS DAC needed to be 1.3 V; the Y channel of the S-Video and Component Video should be 1.1 V max. The CRT DAC implementation for PC video is only 8 bits with a peak voltage of 0.7 volts. There is no way to generate a quality television signal with only an 8-bit resolution on the DAC. This is partly because of the accuracy required to properly construct the subcarrier, but mostly because a large number of the values are used in creating the synchronization pulse. For Composite Video, of the 1.3 volt swing available, only about 0.7 volts are available for luminance. This means that of the 256 levels in the DAC only 138 can be used for luminance; this severely quantizes the picture data causing banding and contouring errors in the image. The 10-bits per channel used in our DAC provides the accuracy needed for good quality video.

CHALLENGES AND SOLUTIONS

Firstly, design to support multiple TV standards:

- Each TV standard has its own specifications with respect to video bandwidth, sub-carrier frequency, color space conversion coefficients, etc. To meet all the requirements this function is designed using many programmable registers.

Secondly, the quality of the filter is crucial:

- The sharpness of the image on TV depends on the transfer function of the horizontal and vertical filters, and since this control is subjective, the filter coefficients are made programmable so that the sharpness can be adjusted by the user.
- Pixels are processed in YUV422 format in the filters. In this format luminance is sampled on every pixel, and chrominance is sampled once every two pixels. Due to this, the filter output incurs a delay between luminance and chrominance called group delay. This delay varies based on the scale ratio and the over-sampling rate. Unwanted visual artifacts are seen on the screen if the delay is not offset. The internal logic compensates for this delay by a programmed number of samples.

Validation of the design is a major challenge.

- TVout was validated at three different levels: RTL using simulations, FPGA, and emulation. Each one of the levels has its own advantages and disadvantages. A combination of the three levels gave good coverage. For better testability of the digital logic and the DACs, 13 different test patterns are implemented in the design.

TVout Digital-to-Analog Converter

A key component of the integrated TV encoder is the DAC embedded block. The TVout embedded block consists of three, 10-bit resolution channels that are fully integrated and custom designed for consumer quality television. Each channel can be programmed to one of three different video voltage swings to support Composite Video, S-Video, and Component Video format.

Figure 10 shows the block diagram of a DAC embedded block that interfaces to the TV.

The DAC embedded block consists of three identical, 10-bit channels, a first-order temperature-compensated voltage reference circuit, local biasing and decoding for output video voltage-level programming, clock distribution, and circuitry for video cable detection.

The DAC is designed to directly drive the TV without the need for external video buffers. There is a passive low-pass video filter at each channel output designed with a nominal bandwidth of 30 MHz to limit the video bandwidth for HDTV. The TV connection to the motherboard is made through a video connector that is compatible with standard S-Video cables and RCA-type cable connections with the use of a dongle. This interface was designed to minimize the number of video connectors and components on the motherboard as well as to minimize the complexity of routing several video signals to various RCA and S-Video-type connectors on the motherboard.

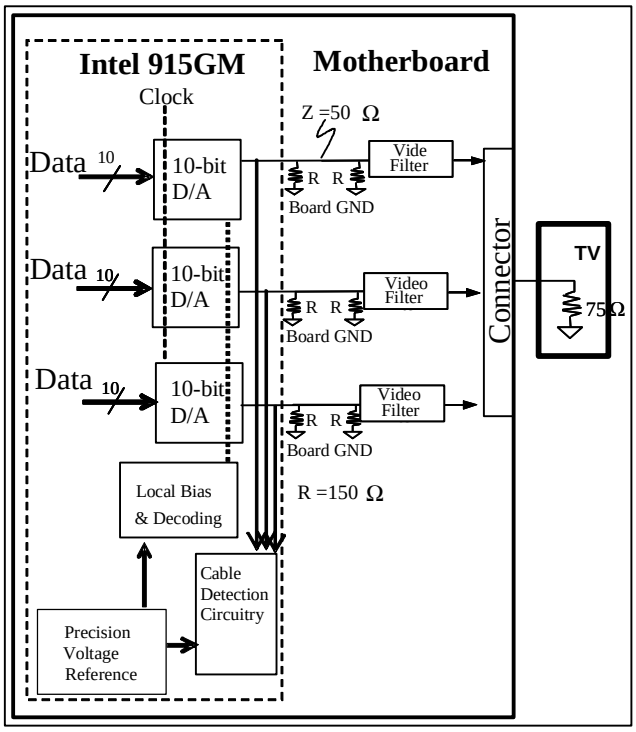


Figure 10: Block diagram of the DAC embedded block and associated board-level components

DAC Circuit Architecture

The performance of the DAC is critical for achieving excellent video quality. The required accuracy of the DAC is based on the differential gain and phase distortion specifications for TV. The architecture is designed as a current-steering architecture to achieve high accuracy and low distortion of the analog video signal from the motherboard ground level (zero volts) to the maximum nominal analog video signal swing of 1.3 V.

The digital input to each DAC is latched on the rising edge of each clock and converted to an analog current. For a given digital input, the current source outputs are either summed together and directed to the output pin or directed to the motherboard ground plane by the differential current switches. An analog video voltage is created from the DAC output current flowing into the termination resistors. The video level specifications for the various video formats along with the effective load termination determine the required output current of the DAC circuit. The least significant bit (LSB) output voltage is a function of the video format supported and normally ranges between 684 μV to 1.27 mV above the motherboard ground.

One of the key performance parameters of the DAC is the static accuracy characterized by the differential (DNL) and integral (INL) non-linearity. Both the DNL and INL must be less than +/- 0.5 LSB for all video modes over the entire analog output range and over all operating conditions in order to achieve monotonic, intrinsic 10-bit static accuracy. Therefore, the current sources controlled by each bit of the digital input word must have an error of less than +/- 0.5 LSB. This specification is very difficult to achieve for a DAC implemented with a binary weighted current source array in a standard CMOS technology while simultaneously meeting dynamic specifications such as signal-to-noise and distortion. Therefore, to reduce the matching constraint for the most significant bit (MSB) current sources while improving the dynamic performance of the converter, the DAC circuit architecture is typically segmented into two current source arrays, one binary weighted and one linear weighted. A DAC implemented with only a linear weighted array will guarantee monotonic behavior at the expense of silicon area. Therefore, there is a tradeoff between silicon area and accuracy of the DAC.

In order to achieve the DAC performance requirements with high yield and without calibration in high-volume manufacturing, the allowable mismatch of the unit current source or LSB was determined using statistical analysis based on the 0.13 μm CMOS process technology. The random and systematic mismatch among the transistors along with the voltage drop of the power routing and the current source output resistance are the dominant terms that affect the accuracy of the DAC.

An additional constraint on the DAC design arises from the dynamic accuracy that is characterized by signal-to-noise and distortion (SINAD), and the minimum effective number of bits required for analog video bandwidths. For this design, the DAC is segmented where the six MSBs of the digital input code are thermometer decoded into 63, 16 LSB weighted current sources. Bits 2, 3, and 4 of the digital input code are also thermometer decoded into 7, 2 LSB weighted current sources. The 71 total current source outputs are summed together to create the DAC current output that corresponds to the digital input. At full-scale or digital input code = 3FFh, the channel output is 1023*I where I is the unit current source magnitude based on a particular video-level format.

DAC Circuit and Layout Design

The DAC current source arrays are implemented with PMOS transistors with cascode transistors to obtain high output impedance. Design techniques were applied to the DAC circuitry to minimize noise and distortion. Special layout techniques were applied throughout the circuitry to minimize the effects of random and systematic mismatch on circuit performance.

The layout of the DAC was custom designed and drawn to fit in the I/O ring utilizing the area of a standard I/O buffer

height while minimizing the overall width of the embedded block. The placement of the current sources for the DAC was designed as a two-dimensional array and the sequence of switching the current sources according to the digital input was designed to consider the spatial arrangement of the current sources to cancel, to the first-order, linear and non-linearity gradients that may occur. A custom package power plane was designed for a low-resistance, symmetric power grid for each channel.

Referring to Figure 10, the nominal DC termination in the TV is a resistive 75Ω termination, and the effective DC resistance at each DAC output is therefore 37.5Ω. Since 75Ω routing impedances are difficult to achieve on low-cost board fabrication, each channel is double terminated with a 150Ω resistors on the motherboard, with one termination resistor placed near the chip and the other placed near the video connector for the TV. This termination scheme results in a 50Ω nominal routing impedance between the termination resistors on the motherboard. As a result, the 75Ω routing impedance is eliminated from the chip output to the input of the video filter, and optimal signal integrity is achieved, which is essential for excellent video quality.

The TVout DAC embedded circuit block accuracy performance along with optimal signal integrity of the analog video signals is essential to obtaining excellent visual quality. The selection of the DAC circuit architecture along with the circuit design and layout techniques applied resulted in a 10-bit DAC without the need for calibration in high-volume manufacturing. The performance of each channel was characterized with static and dynamic test conditions over a wide range of supply voltages (3.0 V to 3.7 V) and a wide temperature range (-20° C to +120° C) for the three supported analog video formats. The excellent performance of the TVout DAC embedded circuit block along with the topology of the termination on the motherboard results in an optimal analog video channel that achieves excellent visual quality.

CONCLUSIONS

With the introduction of the Intel 915 Express Chipset Family we have brought high-performance graphics to the digital home. Our workload-driven design methodologies have delivered best-of-class performance to our customers.

The TVOut feature provides extremely accurate and programmable color space conversion, high-quality multitap flicker filters, and high-precision DACs in order to produce outstanding image quality. When these features are combined with the encoder's support of multiple standards, the Intel 915 Express Chipset Family provides an excellent television experience.

Coupling these two features together brings the Second-Generation Intel® Centrino™ mobile technology Platform into the digital home as never before.

REFERENCES

[1] 3DMark2001 SE, Futuremark Corporation, http://futuremark.com/.*

[2] 3DMark*03 Pro, Futuremark Corporation, http://futuremark.com/.*

[3] DirectX Specification, Microsoft Corporation, http://www.microsoft.com/windows/directx/.*

[4] OpenGL Specification, http://opengl.org/.*

AUTHORS' BIOGRAPHIES

Kam Leung was the Intel 915 Express Chipset Family 3D graphics design manager, and currently is managing Intel's next-generation of 3D graphics development. He has B.S.E.E. and M.S.E.E. degrees from Utah State University. He joined Intel in 1990 and has spent the past 14 years working on chipset and graphics development. His e-mail is kam.leung at intel.com.

Steve Spangler is a 3D Graphics microarchitect working on Intel's next-generation mobile and desktop integrated chipsets. He has B.S.E.E and M.S.E.E. degrees from the University of Wisconsin at Madison. He joined Intel in 1987, working on a variety of chipsets and microprocessors with the last six years being devoted to graphics chipsets. His e-mail is steve.j.spangler at intel.com.

Todd Witter is a senior staff engineer in Intel's Mobility Group. Todd is responsible for the microarchitecture of the graphics display system of Intel's integrated graphics controllers. Todd joined Intel in 1997, after working as a design engineer at the Interactive Television team at Philips Semiconductors. He received his B.S.E.E. degree from California State University at Chico in 1989. His e-mail is todd.m.witter at intel.com.

Kevin E. Arendt is a senior staff design engineer working in Intel's Mobility Group. He is responsible for designing custom embedded circuits for chipsets for desktop and

---

™ Centrino is a trademark of Intel Corporation or its subsidiaries in the United States and other countries.

* Other brands and names are the property of their respective owners.

mobile platforms. Kevin received his B.S.E.E. and M.S.E.E. degrees from Southern Illinois University at Carbondale in 1990 and 1992, respectively. Kevin joined Intel in 1995. Prior to joining Intel, he worked as a design engineer for AT&T's Global Information Solutions Division where he worked in the Analog Cell Library Group. Kevin is a member of the IEEE and his interest is in analog integrated circuit design. His e-mail is kevin.e.arendt at intel.com.

Satya Avadhanam is a senior design engineer working in Intel's Mobility Group. He received his B-Tech. degree from Sri Venkateswara University, India and his M.S.E.E. degree from Arizona State University at Tempe in 1995 and 1997, respectively. Satya joined Intel in 1998 and worked on various projects in 2D and 3D graphics areas. His e-mail is satyanarayana.avadhanam at intel.com.

Satyaki Koneru received his B-Tech. degree in Electrical Engineering in 1993 from the Indian Institute of Technology, Bombay and his M.S. degree in Computer Engineering in 1998 from Iowa State University. He joined Intel in 1998 and has been working in the field of 3D graphics since then. He is currently a microarchitect in the Mobile Graphics Engineering group in the Mobility Group. He was involved in the architecture of the graphics core built into the Intel 915 Express Chipset Family chipsets for the Second-Generation Intel Centrino mobile technology platform. His e-mail is satyaki.koneru at intel.com.

Val Cook is a senior staff engineer working in Intel's Mobility Group. He received his B.S.E.E and M.S.C.S. degrees from Brigham Young University at Provo in 1989 and 1995, respectively. He joined Intel in 1995 working on their 3D graphics products, helped architect the Intel 915 Express Chipset Family 3D graphics and continues to be involved with the development of future graphics architectures. His interests include 3D rendering algorithms and architectures. His e-mail is val.g.cook at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Low-Power Audio and Storage Input/Output Technologies for the Second-Generation Intel® Centrino™ Mobile Technology Platform

Tom Clark, Mobility Group, Intel Corporation
Feng (Frank) Yang, Mobility Group, Intel Corporation
Greg Kaine, Mobility Group, Intel Corporation
Brian Leete, Mobility Group, Intel Corporation
Sriram Ranganathan, Mobility Group, Intel Corporation

Index words: High Definition Audio, Serial ATA, SATA

ABSTRACT

One significant challenge in new platform design is how to add new and improved capabilities without increasing the power required to provide a particular function. This paper focuses on two key Input/Output (I/O) technologies provided by the chipset: integrated audio and the interface to the storage device (hard drive). First, we examine the architectural features of Intel® High Definition (HD) Audio and compare them to the previous integrated audio solution, in terms of power. Second, we examine the architectural benefits of Serial AT Attachment (SATA), and we describe the power-saving techniques that enable increasing functionality while maintaining power.

INTRODUCTION

A key requirement for the next-generation platform built on Intel® Centrino™ mobile technology is increasing the functionality provided by the platform in general, and the chipset in particular, while not using more power than that used by the previous-generation platforms. Intel chipsets provide many I/O interfaces on the Input/Output Controller Hub (ICH). This platform uses the sixth generation of the ICH, known as the ICH6-M. Two key interfaces provided by the ICH6-M are audio and storage. Both interfaces require increased functionality and throughput. Increased functionality usually comes at the cost of a larger design (increased gate count), which translates into higher leakage power and higher active power. Higher throughput implies faster activity and/or wider data paths, which also imply higher power.

Intel HD Audio improves audio quality with a higher bit rate and wider data words than the previous implementation. Other features improve the end-user experience by adding new functionality, such as support for array microphones, audio jack retasking, and multiple audio streams. Intel HD Audio mitigates the power impact of these improvements at a system level by enabling the processor to enter deeper power states and power-savings states more frequently, even while the audio interface is actively employed, such as during CD or DVD playback.

Serial AT Attachment (SATA) will eventually replace the previous storage interface, known as Parallel AT Attachment (PATA) as the interface of choice for hard drives. SATA provides an immediate increase in storage interface headroom by increasing the maximum theoretical bandwidth from 100 MBps to 150 MBps. This ensures that the rate at which the platform transfers data to and from the hard drive will be limited by hard drive mechanics, not by the electrical interface. SATA achieves this high transfer rate with a high-speed serial interface, which reduces component product cost and which can decrease platform manufacturing costs. To overcome the power demand of a high-speed serial interface, SATA uses power-reduction protocols on the host (chipset) as well as the device (hard drive).

HIGH DEFINITION AUDIO ARCHITECTURE

Audio Codec '97 (AC'97) was the first standard for integrated audio. AC'97 is capable of delivering up to 96 kHz/20-bit playback in stereo and 48 kHz/20-bit in multichannel playback modes. AC'97 first appeared in Intel chipsets starting in 1999. However, as consumers increasingly use PCs for their entertainment and communication needs, the demand for audio capability is much higher than AC'97 can deliver. Intel HD Audio is the next-generation integrated audio solution for PCs.

Intel HD Audio is expected to allow the integrated audio solution to keep pace with PC audio advances for the next ten years or more. The next-generation PC platform built on Intel Centrino mobile technology will be the first mobile platform with Intel HD Audio. It will gradually replace AC'97 over the next few years. Table 1 provides a summary of the key advantages of Intel HD Audio over AC'97.

Table 1: Feature comparison of AC'97 to Intel HD Audio

Feature	AC'97	Intel HD Audio
Streams	single stream (in and out)	up to 15 input and 15 output streams at one time and up to 16 channels per stream
Bandwidth	11.5 Mbps maximum	48 Mbps per SDO 24 Mbps per SDI
Bit Rate	up to 20-bit, 96 kHz	up to 32-bit, 192 kHz
DMA Usage	predefined	dynamic
Bandwidth Assignment	fixed	dynamic
Analog Plug-and-Play	limited	comprehensive
Power Management	none	idle detection, fine-grained power control
Microphone Array	2-elements	up to 16-elements
Driver SW	developed by audio codec supplier	OS native bus driver and Independent Hardware Vendor value-added function driver
Docking	none	enabled
HW Configuration	fixed	dynamic device enumeration

Intel HD Audio Hardware Architecture

In this section we provide a conceptual overview of Intel HD Audio. Our intention is not to provide any detailed or quantitative definition. For more information, refer to [6].

Intel HD Audio hardware architecture consists of three major components: a controller, a link, and a codec. The hardware building blocks of the architecture are shown in Figure 1.

Intel HD Audio Controller

The controller is connected to the system memory via a Peripheral Component Interconnect (PCI) Express*

* Other brands and names are the property of their respective owners.

interface. It implements a memory mapped register interface for standard programming access. One or more Direct Memory Access (DMA) engines transfer audio "streams" between the system memory and codec(s) connected to the Intel HD Audio link. Each DMA engine in the controller can be set up to transfer a single "stream" between the system memory and one or more codecs.

Intel HD Audio Link

The audio link is a digitized serial interface that conveys data between the Intel HD Audio controller and the codecs. The controller is connected to one or more codecs via the audio link. Figure 2 shows how to connect an Intel HD Audio controller to codecs.

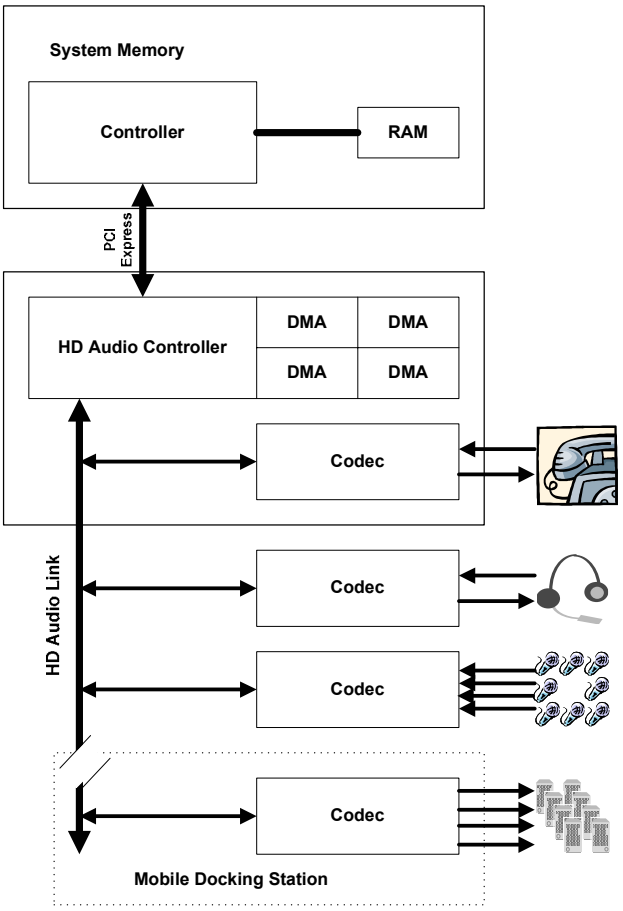


Figure 1: Intel HD Audio architecture block diagram

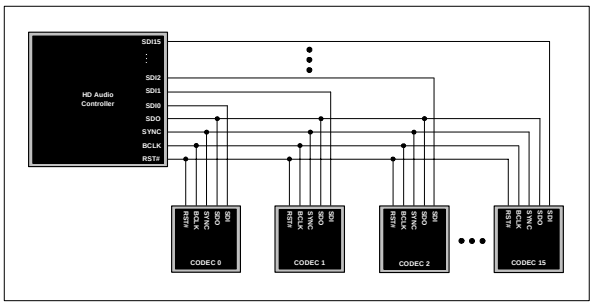


Figure 2: Example Intel HD Audio system

The link protocol is synchronized by the Intel HD Audio controller. It is purely isochronous with a 48 kHz framing period.

The Intel HD Audio link uses existing AC'97 signals with some changes:

- *Reset*. Also used to supply codec addresses.
- *Clock*. Rate is increased to and from the Intel HD Audio controller to codecs.
- *Sync to codecs*. Aligns frames and delineates outbound stream.
- *Serial Data Out (SDO).* Broadcast to all codecs.
- *Serial Data In (SDI).* Point to point, one or more per codec.

Figure 3 illustrates an example transaction on the Intel HD Audio interface.

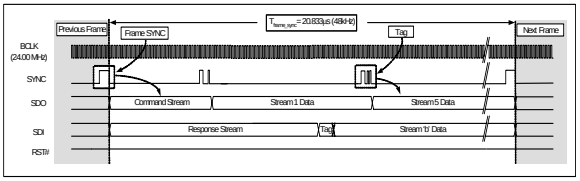


Figure 3: Intel HD Audio frame showing output command and data stream and input response stream

Codec

A codec is a physical device connected to the Intel HD Audio link. A single codec may contain one or more function groups.

- For output, it extracts one or more audio streams from an Intel HD Audio frame and converts them to an output stream through one or more converters.
- For input, it collects the data streams from one or more converters, frames them into Intel HD Audio frames, and sends them to the Intel HD Audio controller for processing.

Function Group
A function group is a collection of widgets that are all common to a single application/purpose and that are controlled by a single function driver. A widget is the smallest enumerable and addressable module within a function group. For each widget, there is a defined set of standard parameters and controls.

Intel HD Audio Software Architecture

Intel HD Audio software architecture is a layered architecture as shown in Figure 4.

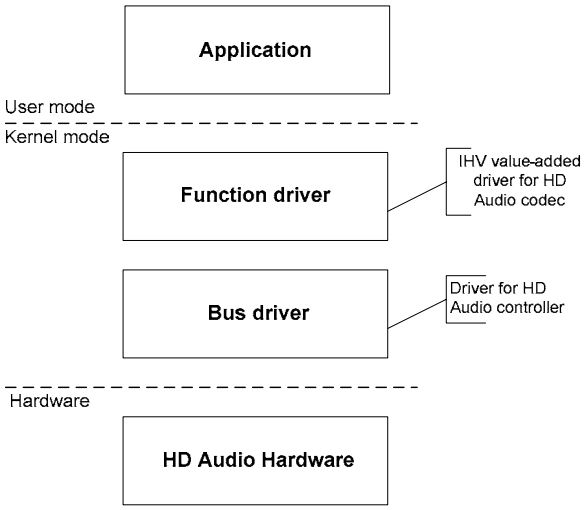


Figure 4: Intel HD Audio Software Driver Stack

Intel HD Audio Bus Driver
The Intel HD Audio Bus Driver has direct access to the HD controller register interface. It is responsible for the following:

- Initialization of the Intel HD Audio controller and link.
- Enumeration of all codecs on an Intel HD Audio link.
- Run-time management of Intel HD Audio controller and link, including power management.

Intel HD Audio defines a robust initialization model wherein the operating system can discover all of the codecs present and then can allocate system resources, such as memory, I/O space, and interrupts to create an optimal system environment. Intel HD Audio also defines a robust discovery mechanism for the codecs that are added at run time.

Intel HD Audio Function Driver
In most cases, the Intel HD Audio function driver is provided by codec vendors. It must implement a set of the standardized commands as defined in Intel HD Audio Specification 1.0. In general, it is responsible for the following:

- Initialization/reset of codec.
- Codec command processing.
- Management of codec topology and configuration.
- Management of streaming operation.

Intel HD Audio Power Management Architecture

Intel HD Audio provides building blocks for fine-grained power management. It defines power states for links and codecs. In most cases, all power-management controls are driven by software. Intel HD Audio also defines power widgets for power-management purposes. A power widget provides a single point of power state control for a predefined group of audio widgets. Power control to this widget effectively places all member widgets in the group in the prescribed power state. Fine-grained power management can be achieved with definitions of power widgets.

To enable more effective power management, the platform enables the CPU being put into sleep state while there is still an active stream on the Intel HD Audio link.

Intel HD Audio Docking Architecture

Intel HD Audio architecture is designed for docking support. It supports dynamic codec enumeration and defines a mechanism for heart beat detection in the link protocol. With appropriate chipset and software support, Intel HD Audio can be set up to have one or more Dolby 5.1/7.1 setups on a docking station.

Intel HD Audio Architecture for Digital Home and Digital Office

To meet emerging requirements for providing a surround sound experience with a variety of speaker configurations, Intel is working with Dolby Laboratories. Intel HD Audio can support all the Dolby technologies, including the latest Dolby Pro Logic IIx,* which makes it possible to enjoy older stereo content in 7.1 channel surround sound.

Intel HD Audio also supports the Sony/Philips Digital Interface (S/PDIF) that enables users to connect PCs to high end digital audio equipment.

Intel HD Audio not only provides an enriched playback experience, but it can also deliver a better quality input for voice and communication applications that are critical to a

* Other brands and names are the property of their respective owners.

digital office usage model. Due to its new architecture and increased available bandwidth, Intel HD Audio can support up to a 16-element array microphone for higher quality input. One business application for array microphones is video conferencing, which could employ a voice-tracking camera, with noise cancellation over a Voice over Internet Protocol (VoIP) phone.

Jack retasking enables the host controller to redirect the audio stream to the correct device. In the case of a user plugging a headphone into a microphone jack, jack retasking will direct the headphone stream to the microphone jack.

Designed for Future

Intel HD Audio is a key building block of Microsoft's United Audio Architecture (UAA)*. UAA is a feature of the next Windows* Operating System (OS) codenamed "Longhorn." Intel HD Audio architecture is also extensible and can support any non-audio device connected to an Intel HD Audio link, as long as the device is compliant with the Intel HD Audio specification.

Intel HD Audio Architecture Summary

Intel HD Audio defines a new architecture that provides a significant performance increase compared to AC'97. It brings a consumer-electronics level of audio experience to PCs. Furthermore, it provides building blocks to enable emerging usage models for the PC as a family entertainment hub, i.e., a collaboration and communication device.

INTEL HIGH DEFINITION AUDIO POWER FEATURES AND RESULTS

Intel HD Audio System Power-Savings Design

The battery life of a mobile device is dependent upon platform power consumption. Intel HD Audio has been implemented with this understanding and enables techniques that can provide significant power savings under certain usage models. Intel HD Audio does this by allowing the CPU to enter a lower power ACPI idle state (C4) under certain conditions. In contrast, AC'97 will only allow the CPU to go as deep as C2, a state that consumes greater power than C4. Workloads that contain continuous audio streams such as CD audio playback and DVD playback can take advantage of this deeper power-saving state. Power savings are also achieved during idle periods with no audio playback. Intel HD Audio supports an efficient low-power ACPI idle state (D3) that it enters when no audio streams are being processed for an amount of time defined by the device driver.

When playing an audio stream using Intel HD Audio the CPU is allowed to enter C4, a low power CPU state. For AC'97 audio the depth of the CPU C-state is limited to C2 which consumes more power than C4. The ability of the CPU to enter this deeper sleep state with Intel HD Audio enables substantial power savings over AC'97. This power savings has been evaluated on an Intel reference board designed for taking power measurements.

Where the data stream resides is what permits this C-state behavior to occur. Intel HD Audio keeps data stream traffic in non-cached system memory and uses DMI VC1 to permit the Intel HD Audio controller to write or read these data without needing to snoop them on the processor. AC'97 keeps data stream traffic in cached system memory that must be snooped when the AC'97 controller writes or reads them. As a result the processor must always be in a snoopable C-state (C0-C2).

Intel HD Audio Power-Savings Results

CD audio playback and DVD movie playback are two common usage models that show the C4 power savings with Intel HD Audio as opposed to AC'97 audio. Both usage models generate isochronous data traffic thus allowing the benefits described above.

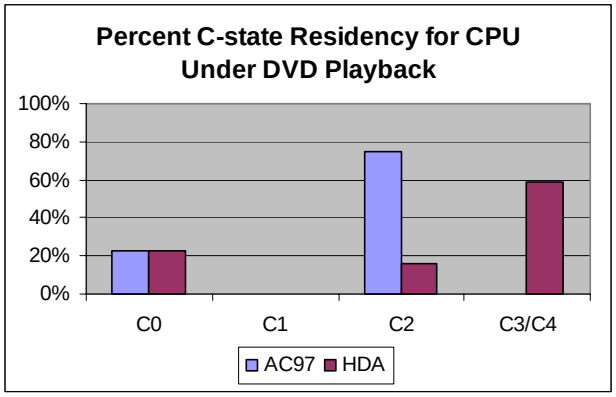


Figure 5: DVD playback CPU C-state residency

DVD playback shows a large shift in C2 residency to C4 as illustrated in Figure 5. C2 residency decreases from 75% with AC'97 audio to 16% for Intel HD Audio. This also results in an increase of C4 residency from 0% with AC'97 audio to 58% for Intel HD Audio. The shift in the distribution of C-states can result in a power savings of ~700 mW (see Figure 7). This power savings is almost entirely provided by the increased CPU residency in the lower power C4 state. Figure 5 also shows a significant CPU overhead in DVD decoding. Much of this overhead

* Other brands and names are the property of their respective owners.

is due to video frames and not related to decoding audio. C1 residency is largely unaffected by either audio codec.

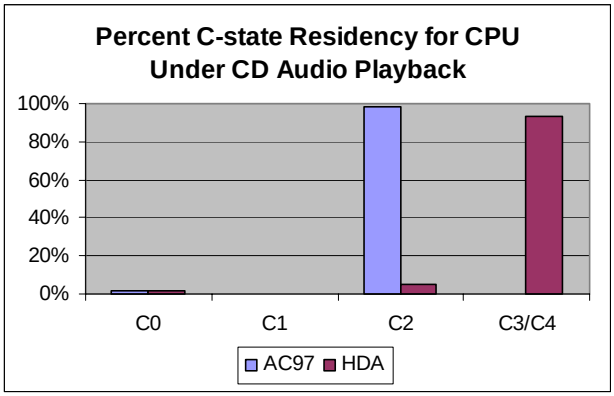


Figure 6: CD audio playback CPU C-state residency

CD audio (no video decoding overhead) shows an even greater shift in the CPU residency from C2 to C4. Figure 6 shows close to a complete shift of C2 residency to C4. C4 residency goes from 0% with AC'97 to 93% with Intel HD Audio, and there is a similar decrease in C2 residency from 98% (AC'97) to only 5% with Intel HD Audio. This large shift in CPU C-state residency results in power savings of ~850 mW as measured on the reference platform (see Figure 7).

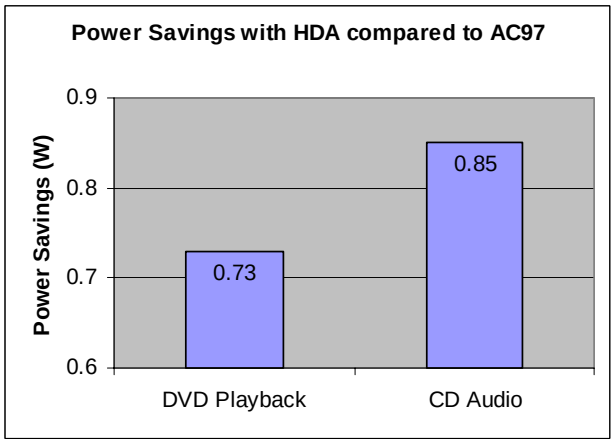


Figure 7: DVD and CD audio power savings

The impact on battery life of these power savings may be determined given the battery capacity and the average power for the given workload. For this analysis, platform average power for DVD playback is assumed to be in the range of 13 W to 17 W with Intel HD Audio. The percentage increase in battery life resulting from Intel HD Audio is directly related to the 700 mW decrease in power. Intel HD Audio could therefore potentially increase in battery life by as much as 8-17 minutes in DVD playback over a system configured with AC'97 audio.[1]

Similarly, for battery-life savings, analysis of CD audio playback power consumption in the range of 9 W to 13 W is assumed. The percentage increase in battery life resulting from Intel HD Audio is again dependent on the power savings of the workload. CD audio sees power savings of ~850 mW with Intel HD Audio. This power savings could potentially result in battery-life savings as much as 35-42 minutes in CD playback over a system configured with AC'97 audio.[1]

In addition to the power savings provided by Intel HD Audio during active workloads, there is also savings provided by the D3 low-power ACPI state when no audio stream is being processed. This low-power state brings Intel HD Audio power consumption down to levels on the order of tens of milliwatts. Without this feature, power consumption for the device could be as much as 10x greater when the device is idle. This can provide a ~5% decrease in power consumption for a platform at idle. A period of time is defined by the device driver that looks for any audio activity occurring during this period. If no activity is encountered the device is sent to the D3 state. A more power-friendly device driver would implement this period to be as small as possible in order to maximize power reduction. At the time of writing this paper the Windows driver for Intel HD Audio sets this value to a fairly aggressive period of five seconds.

SERIAL ATA ARCHITECTURE

Serial ATA (SATA) is the next generation of storage interface technology that is expected to replace Parallel ATA (PATA) in mobile systems. SATA provides many new features including: increased data transfer rates with increased protocol efficiency; low pin count interconnect with thinner flexible cables and connectors; decreased signaling voltage; enhanced error detection; and a point-to-point topology with the elimination of master/slave. It also includes support for Native Command Queuing, Native Hot Plug, Staggered Spin Up, and Port Multipliers. These new features are designed to provide both immediate value to the SATA subsystem today, and allow for future growth in storage technology over the next ten years. The following table compares SATA to PATA.

[1] Battery-life savings will vary with system configuration and platform power for a given workload. This range is based on battery capacities ranging from 54 Wh to 72 Wh with platform power for the workload at the minimum of the assumed range of power consumption.

Table 2: Feature comparison of SATA and PATA interfaces

Feature	PATA	SATA
Effective Data Rate	133 MB/Sec	150 MB/Sec 300 MB/Sec (Future) 600 MB/Sec (Future)
Connector Size	40	7
ICH6-M I/O Pins	27 (1 bus)	4 (per port)
Signaling Voltage	3.3 V	500 mV differential 250 mV common mode (nominal)
Error Detection	Data Only	Data and Commands
Point to Point Topology	No	Yes
Master/Slave Elimination	No	Yes
Native Command Queuing	No	Yes
Native Hot Plug	No	Yes
Staggered Spin Up	No	Yes
Port Multipliers	No	Yes
Link Power Management	N/A	Yes

In the ICH6-M, each SATA port consumes only 4 I/O pins (vs. 27 for PATA). Freeing up I/O pins allows additional functionality to be added to the chipset without increasing die size. Also, removing the requirement to provide the 3.3 V signaling that PATA requires, allows greater flexibility in adding future features to new chipsets.

On the motherboard, only four data traces to the drive need be provided to the drive for signaling purposes, and the SATA connector is correspondingly smaller, allowing more compact and efficient motherboard designs.

The SATA host controller in the ICH6-M can be configured to export one of two distinct software interfaces, a fully ATA backwards-compatible interface and a new SATA-specific Advanced Host Controller Interface (AHCI).

The AHCI specification supports the new SATA-specific features specified in SATA I and SATA II specifications. These include Native Command Queuing, Native Hot Plug, Staggered Spin Up, Port Multipliers, and Aggressive Link Power Management.

One of the basic challenges always present in introducing a new technology into the mobile platform is the additional power consumed by the added functionality. SATA Link Power Management can be used to limit the additional power consumption of the new SATA interface, providing the additional capability of SATA at the minimum power cost.

Link Power Management Overview

SATA Link Power Management puts the physical layer (PHY) of the link into a low-power state, independent of the ATA protocol power state of the disk, and as such complements the existing power management capabilities provided by the ATA command set. The ATA command set reduces the power consumption of the attached device by issuing protocol-level power state change requests to the disk. These requests typically instruct the device to spin down the media to save power.

Independent intelligent PHY power management has shown a significant reduction in the overall power consumption of the SATA subsystem, both in the platform and in the drive itself.

Link Power Management States

SATA provides two link power management states, in addition to the “active” state. These states are Partial and Slumber, that by specification, differ only by return latency. Partial has a maximum return latency of 10 microseconds, while Slumber has a maximum return latency of 10 milliseconds. Typical host and device hardware implementations should be able to realize greater power savings in Slumber with its longer specified return latency. Partial, therefore, is designed to allow link power state transitions continually with minimal impact on performance. Slumber is designed to be used only when the link is expected to be idle for an extended period of time. It is not possible to transition the link directly from Partial to Slumber without passing through the active state, as shown in Figure 8.

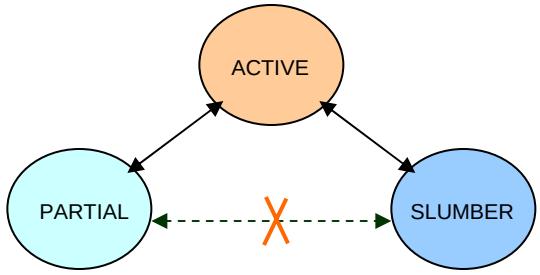


Figure 8: SATA power state transitions

Host- and Device-Initiated Power Management

SATA Link Power Management can be broken up into two basic types: Host Initiated Link Power Management (HIPM) and Device Initiated Link Power Management (DIPM).

SATA Link Power Management requires cooperation between the host and the device. Either the host or the device can request the link to enter a low-power state, but the corresponding host or device must accept or reject the link state change request. Each of these provides power savings by themselves; maximum power savings, however, are achieved when both are implemented together.

Host-initiated power management can be implemented either in the host hardware or the host software. In the first case, the host controller requests a link power management transition immediately after all outstanding commands to the drive have been completed. This allows the link to enter a low-power state immediately upon completion of the commands to the disk. Since the host has the best knowledge of what commands have been posted, or will be posted to the device, the host is able to make an immediate link power state change without invoking a time-out period.

Device-initiated power management is implemented by the drive. The drive knows best how long a specific command might take to complete, and is best equipped to request a link power management state change while processing the command.

The host controller on the host can automatically put the link into either Slumber or Partial after the command completes. Typically, for performance reasons, this will be Partial. However, after some period of idleness, it is generally assumed that the link will be inactive for an extended period of time and it is desirable to transition the link from Partial to Slumber. This can be done either by the host software or the device.

Since the host is best equipped to manage the PHY between commands and the best device within a command, the best power management is obtained when the host and device cooperate.

Link Power Management and Device State Interaction

The operating system can put devices into a D-state as defined by the Advanced Configuration and Power Interface (ACPI) specification [4]. Devices can include the host controller and the drive itself. The link, however, can be put into its power management state independent of the D-state of the host controller or the device.

Figure 9 illustrates the interaction between the link and the device D-states.

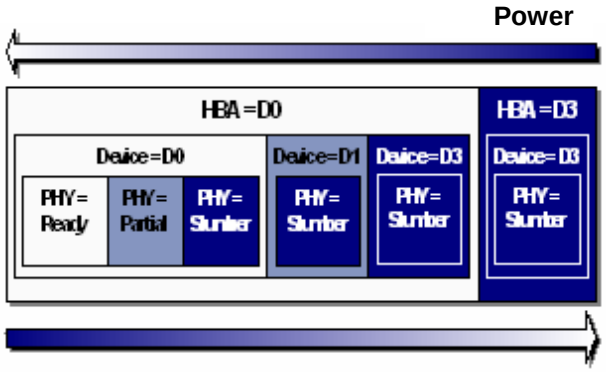


Figure 9: SATA Link and Device Power Management states

Here we see that all link power states are enabled when the device is in the D0 active state. When the device is in one of the lower power states, the link should be in Slumber.

Power Management Protocol and Results

The protocol for entering and exiting Partial and Slumber states is different. To enter one of the low-power states, the standard communication protocol between the host and device can be used. Once the PHY has been placed into a low-power state, the standard communication protocol between the host and drive cannot be used; another Out of Band (OOB) mechanism is required.

Entry Signaling Protocol

Either the host or the device can initiate a request to enter into the Partial or Slumber power state. The request is signaled by transmitting either the PMREQ_P or PMREQ_S primitive. PMREQ_P is used for a request to transition to the Partial power state, and PMREQ_S is used to transition to the Slumber power state. The

corresponding host or device must then respond with a PMACK acknowledge or PMNAK negative acknowledge primitive. If the request is acknowledged with a PMACK, both the host and device transition into the corresponding power state. To ensure the PMACK received is reliable without having to handshake the handshake, the target sends several PMACK's before entering a low-power state. If a PMNAK is received, no power state change occurs.

Exit Signaling Protocol

Once the PHY is in a low-power state, SATA specifies a low-level signaling mechanism to bring the interface back to active state. An OOB signal "COMWAKE" is sent to the device that acts as a wake-up call and causes communication to be reestablished.

Hardware/Software Control

Link Power Management is only enabled when the host controller and the device reports that it is capable of issuing or receiving requests, and the driver is capable of enabling the host hardware and drive. The host controller reports this capability to software in the Capabilities Register of the AHCI host controller (See [3] for details). The drive reports the ability to support these commands in the IDENTIFY_DEVICE data structure returned by the device during device enumeration (see the ATA [5] and SATAII [2] specifications for details). The host must specifically enable DIPM on the drive via the ATA SET_FEATURES command upon initialization.

Host/Device Design Recommendations and Interaction

The host hardware should transition the link into Partial after every command. Partial provides the best tradeoff between power savings and performance. The host should also transition the link into the Slumber state after requesting the drive to enter into a lower Advanced Configuration and Power Interface (ACPI) D-state.

In the absence of a host-initiated power management transition, the device should attempt to transition the link into Partial after some appropriate short timeout. Also, if the host doesn't transition the PHY into Slumber after sending a D-state command, the device should do it.

After a longer period of timeout, it can be assumed that the link will remain idle for a longer period of time. Either the host, through the host software, or the device should at this point transition the link to Slumber.

Device Removal During Power Management

SATA Link Power Management disables the PHY between the host and device. Because the PHY is disabled, a device removal cannot be immediately detected without waiting for the next command to be sent to the drive. The AHCI specification provides the capability for a second electrical signal to be provided to the host controller to notify the host that the drive has been removed. Typically this will be implemented as an interlock switch on the mobile platform.

Enabling Link Power Management in Software

SATA Link Power Management can be implemented in a variety of ways, two of which are presented here.

The first of these is the Intel Application Accelerator driver. The Intel Application Accelerator driver, supported on Windows operating systems, replaces the native Microsoft storage driver. It is provided with the ICH6-M chipset, and it utilizes the new AHCI interface. It implements the following AHCI-specific software features: Native Command Queuing, Native Hot Plug, Staggered Spinup, and Aggressive Link Power Management.

The Intel Application Accelerator driver unlocks the full potential of the new SATA bus through AHCI. It provides best-of-breed performance through Native Command Queuing, while reducing power consumption to a minimum through a full implementation of both host-initiated and device-initiated SATA Link Power management.

The second is through BIOS or other software cooperating with and using an existing legacy software stack. In this case the host controller is left in legacy mode, and no host-initiated link power management features, or other AHCI-related features, are available. In this case the BIOS or other software queries the drive through the ATA IDENTIFY_DEVICE data structure to determine its capability of supporting device-initiated link power management, and it enables it via the ATA SET_FEATURES command.

In BIOS this can be implemented using the _GTF ACPI object. For further details see the Advanced Configuration and Power Interface (ACPI) specification [4].

Power Measurement Results

In Figure 10 we show the chipset power consumption of an ICH6-M-based system at Windows idle. Here we see that a chipset with a hard disk on the SATA interface with Link Power Management active consumes nearly equivalent power to the chipset when connected to a PATA hard drive.

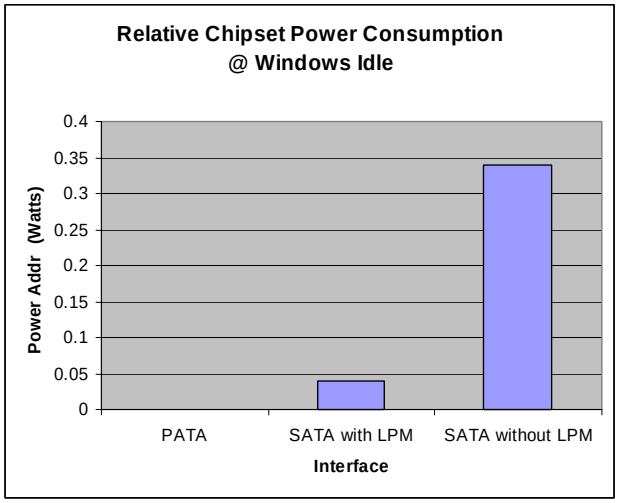


Figure 10: Power savings due to SATA link power management

Notice that Link Power Management provides nearly a 300 mW savings in the ICH6-M over a SATA drive without Link Power Management.

On the drive side, similar power savings can be achieved with a typical SATA hard disk drive consuming approximately 150 mW less power when Link Power Management is active. Total system-wide power savings attributable to Link Power Management in a single drive system reaches 450 mW.

INTEL HD AUDIO DISCUSSION

Intel HD Audio implementation in ICH6-M is the first instantiation of an Intel HD Audio host controller, along with the first generation of software. As the first instantiation, all of the features enabled by the Intel HD Audio architecture may not be available with a given codec or a given driver. Early instantiations have focused on higher quality audio (word size and bit rate), and have focused less on advanced features, such as array microphones and jack retasking.

Likewise, the power savings seen in a given instantiation of the platform will be very dependent on the codec and driver used. The power savings quoted in this paper were measured using an Intel mobile reference design.

The significance of the audio results is that Intel HD Audio provides meaningful increases in audio quality and functionality while reducing system power.

INTEL HD AUDIO CONCLUSION

Intel HD Audio provides significant improvement over prior integrated audio solutions. The ICH6-M implementation of the host controller supports improvements in audio quality and functionality. By carefully designing the architecture to minimize CPU interaction, Intel HD Audio was shown to save power relative to AC'97 in both DVD playback and CD playback. These savings were measured as 730 mW (DVD) and 850 mW (CD), which should result in prolonging battery life by 8 to 17 minutes for DVD playback, and by 35 to 42 minutes for CD playback. This shows that Intel HD Audio provides advanced capability while reducing system power.

SATA DISCUSSION

SATA Link Power Management was first implemented in the ICH6-M.

The Link Power Management results shown here were taken at Windows idle when no data were being transferred to the device. The system-level power savings realized by the end user will depend on the amount of disk traffic sent to the device. It is expected that with well-designed disk drives, and typical user loads, a typical user will experience nearly all of the 450 mW saved at Windows idle.

Furthermore, Link Power Management can be implemented on SATA optical drives as well as SATA hard drives. In typical mobile systems, the optical drive is rarely used. As the time of publication, SATA optical drives are not mature enough to measure the effect of Link Power Management, but by extrapolating from existing hard disk data we expect the optical drive to realize the same 150 mW savings that we see in the hard disk drive.

Total expected power savings in a typical 2 SATA drive system approach 600 mW, compared to a system without Link Power Management.

SATA CONCLUSIONS

SATA is an important technology to the Industry in that it provides a storage interconnect technology with room to grow over the next ten years. SATA provides significant improvements in storage performance and features for both desktop and mobile platforms. A significant challenge to the adoption of any new technology in the mobile space is power consumption. With a potential power savings of 600 mW, Link Power Management becomes an important component to the adoption rate of SATA in the mobile platform by providing the additional value of SATA with minimal power impact.

The Intel Application Accelerator driver provides the best implementation of Link Power Management, coupled with best-of-class performance through Native Command Queuing.

ACKNOWLEDGMENTS

Our thanks to Ed Pole and Les Cline for their valuable review feedback.

REFERENCES

[1] Serial ATA International Organization, Serial ATA 1.0a Specification, http://www.sata-io.org.* 7 January 2003.

[2] Serial ATA International Organization, Serial ATA II: Extensions to Serial ATA 1.0a, Revision 1.2, http://www.sata-io.org.* 27 August 2004.

[3] Intel Corporation "Advanced Host Controller Interface (AHCI) Revision 1.0," www.intel.com.

[4] Intel Corporation, "Advanced Configuration and Power Interface Specification (AHCI) Revision 3.0," www.acpi.info.*

[5] "ANSI, Information Technology–AT Attachment with Packet Interface–6," (ATA/ATAPI-6), www.ansi.org.*

[6] Intel® High Definition Audio Specification Revision 1.0, http://www.intel.com/standards/hdaudio/, April 15, 2004.

AUTHORS' BIOGRAPHIES

Tom Clark is an I/O program manager for the Client Platform Division in Intel's Mobility Group. Tom is interested in integrating I/O technologies into core chipsets. He received his B.Sc. degree in Electrical Engineering from the University of Florida. His e-mail is thomas.m.clark at intel.com.

Feng (Frank) Yang is a software architect with the Mobility Group at Intel. He holds a Ph.D degree in Computer Science from Syracuse University. His current focus is on mobile platform technology. His e-mail is frank.yang at intel.com.

Greg Kaine received his B.S. degree in Computer Engineering from Purdue University at West Lafayette, IN and joined the Intel Rotation Engineering Program in 2001. Currently he is a platform architect in Intel's Mobility Group and is responsible for analysis of platform power and performance for mobile work loads. His current interests include analysis of mobile applications and multimedia workloads. His e-mail is greg.d.kaine at intel.com.

Brian Leete received his Bachelors degree in Electrical Engineering from Brigham Young University and is a software architect in Intel's Mobility Group, with background and interests in Serial Interconnect technologies, including SATA and USB, and driver- and OS-level software design and development. His e-mail is brian.a.leete at intel.com.

Sriram Ranganathan received his Bachelors degree in Mathematics in 1994, his Masters degree in Computer Application from Bharathidasan University in 1997, and his Masters degree in System Software from Birla Institute of Technology and Science in 1999. He works as a senior software engineer with Intel's Mobility Group. His current interest is in SATA and storage subsystems, and he has a background in developing system software and drivers for Windows and Linux* operating systems. His e-mail is sriram.ranganathan at intel.com.

* Other brands and names are the property of their respective owners.

THIS PAGE INTENTIONALLY LEFT BLANK

Performance and Power Consumption for Mobile Platform Components Under Common Usage Models

Eunice Chang, Mobility Group, Intel Corporation
Jim Fung, Mobility Group, Intel Corporation
Greg Kaine, Mobility Group, Intel Corporation
Ali Saeed, Mobility Group, Intel Corporation
Randy P. Stanley, Mobility Group, Intel Corporation

Index words: processor, GMCH, ICH, DDR2, WLAN, Audio, Power, Performance, Mobile, Notebook, Laptop, Intel Centrino mobile technology

ABSTRACT

The second-generation PC platform built on Intel® Centrino™ mobile technology is the second generation of Intel's mobile technology, designed to address the four vectors of mobility: breakthrough mobile performance, integrated wireless LAN (WLAN) capability, great battery life, and thinner, lighter designs. Many architectural features have been added to improve the performance/watt behavior of the platform as compared to the first generation of the Intel Centrino mobile technology. Some of the new features provide significant performance benefits to the platform with a minimal power increase. Other features have been added to provide platform scalability headroom when running demanding applications under future usage models. Several new power-management techniques have been added to deliver on the Intel Centrino mobile technology goal of great battery life.

This new platform is comprised of the Intel Pentium® M processor, the Intel® 915 Express Chipset Family (GMCH), the Intel® 82801FBM I/O Controller Hub (ICH6-M), and the Intel® PRO/Wireless Connection, which combine to support new technologies such as Intel® High Definition (HD) Audio, Double Data Rate II (DDR2) memory, PCI Express* (PCIe), and Serial Advanced Technology Attachment (SATA). Various platform power-savings techniques such as Dynamic Row Power Management for memory and Intel Display Power Savings Technology 2.0 (DPST2.0) reduce the power consumption of the platform to account for the increased power demands the new features bring with them.

In this paper we discuss some of these new features and the impact they have on platform performance and power as observed while executing industry benchmarks. The benchmarks cover a wide range of usage models including typical office productivity, scientific programs, and DVD playback. The intent of the analysis is to give the reader a broad perspective on the next-generation Intel Centrino mobile technology's capabilities across a wide range of usages.

INTRODUCTION

Intel Centrino mobile technology was introduced in March 2003 in order to address the four vectors of mobility: breakthrough mobile performance, integrated wireless LAN (WLAN) capability, great battery life, and thinner, lighter, designs. Similar to the first Intel Centrino mobile technology, this generation continues with a design emphasis on mobile vectors while introducing many architectural features not present in a mobile platform until now. New interfaces are supported including Serial ATA (SATA) and PCI Express (PCIe) in addition to power-savings techniques, both old and new, including Dynamic Row Power Management for memory and Intel Display Power Savings Technology 2.0 (DPST2.0). Some of the enhancements are made possible in part by advances in process technologies increasing the density of transistors in silicon. Both

generations of Intel Centrino mobile technology share a similar topology as illustrated in Figure 1.

This platform is comprised of the Intel Pentium M processor, the i915 GMCH, the ICH6-M, and the Intel PRO/Wireless Connection. This paper provides an analysis of the performance benefits and power consumption (which determines the battery life) that are realized through the new features of each component. In each section of the paper we focus on a given feature and discuss the power and/or performance impact of that feature. Our analysis consists of workloads commonly used in the industry for platform evaluation, among which are the following:

- SPEC CPU2000* (http://www.spec.org)
  - SPECint 2000* (http://www.spec.org)
  - SPECfp 2000* (http://www.spec.org)
- SYSmark 2004* (http://www.bapco.com)
- MobileMark 2002* (http://www.bapco.com)
- 3DMark2001* SE (http://www.futuremark.com)
- 3DMark03* (http://www.futuremark.com)
- DVD Playback
- CD Audio Playback
- Windows* XP* Idle

Many of these workloads are industry benchmarks frequently used to evaluate platforms. This paper assumes the reader is familiar with these workloads. For further information regarding any of the benchmarks mentioned above, please visit the publishing Web sites listed by their name.

Many acronyms are used in this paper. For the purposes of clarity, we have included an appendix of acronyms at the end of this paper.

*Other brands and names are the property of their resepctive owners.

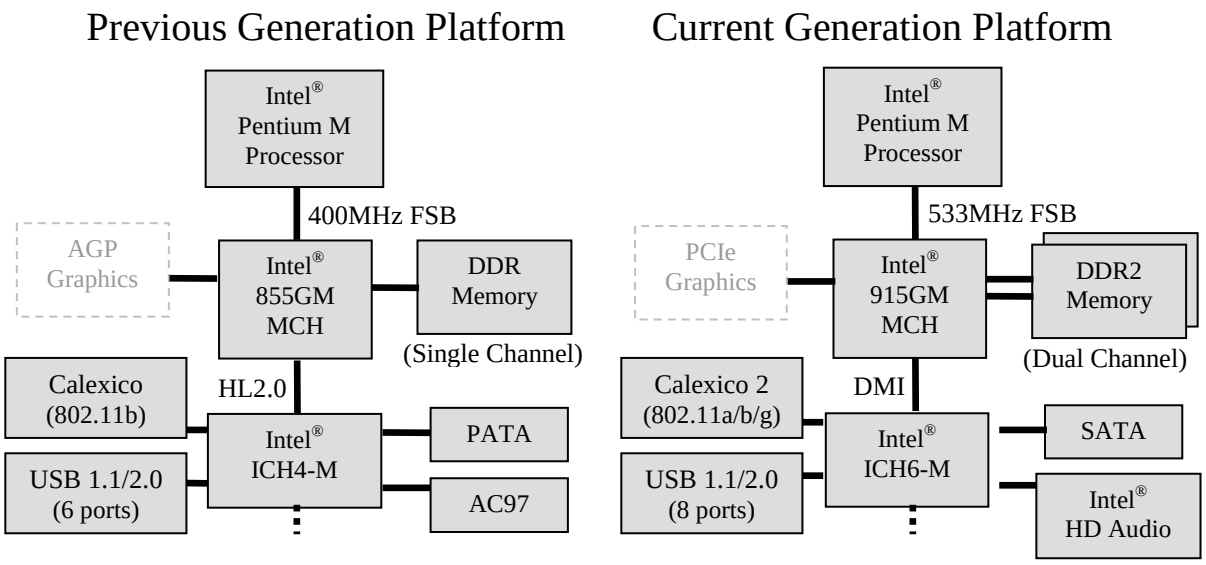


Figure 1: Comparison of topology for Intel Centrino mobile technology

In order to determine the power and performance impact of these new features, workloads were executed on an Intel Centrino mobile technology reference design instrumented for power measurements. Instrumented sense resistors allow for measurement of signals with a NetDaq* data logger. This reference platform follows the Intel specifications that Original Equipment Manufacturers (OEMs) are recommended to follow for notebook designs. Processor C-state residency and performance metrics generated by benchmarks have been included where appropriate. C-state residency is measured using a utility with a negligible impact on system behavior when running.

All data presented in this paper have been generated on a specific platform configuration that may vary for each set of results. All numbers presented are representative of the platform configuration for only that given set of data. Results will vary based on hardware and software configuration.

PROCESSOR

The processor for the second-generation Intel Centrino mobile technology is the second-generation Intel Pentium M processor and is manufactured using 90nm process technology. This processor brings with it a number of new features and improvements over its predecessor. The L2 cache has been doubled to 2 MB and is optimized for reduced power consumption. The Front-Side Bus (FSB) has increased from 400 MHz to 533 MHz and the available processor frequencies have been increased. The Enhanced Register Access Manager provides increased performance for legacy scenarios that would normally stall under register renaming procedures. Intelligent branch prediction is made more efficient through the Advanced Tight Loop Execution feature, and security for the platform is increased with the introduction of the Execute Disable Bit (XD). These features and the impact they have on power and performance are discussed in detail.

90nm Process Technology

The second-generation Intel Pentium M processor is manufactured with Intel's advanced 90nm process technology entirely on 300 mm wafers. This enables Intel to provide the mobile community with significant performance headroom for tomorrow's usage models within today's thermal envelope. These 300 mm wafers provide twice the capacity of their 200 mm counterparts while the process dimensions double the transistor density. This allows Intel to double its already generous on-die cache to 2 MB. This provides a generous workspace for cache-sensitive workloads, such as imaging applications, while at the same time reducing the average power and maintaining the ability to extract the heat dissipated from periods of peak power. The latter is accomplished by maintaining approximately the same surface area as its predecessor.

This second-generation 90nm lithography produces 50nm gate dimensions (half the diameter of an influenza virus) with 50% faster transistors. Intel also strains N-type and P-type devices, increasing their electrical conduction by enhancing carrier mobility through the Si lattice by 10-20%. This significantly lowers the device voltages, a key ingredient to lower-power operation.

Ultimately the goal of the processor is to accomplish more work per unit time for a given level of power than its predecessor. This requires greater activity levels not only for the switching devices already detailed, but for the transistor interconnect structures as well. Intel has addressed this by providing seven layers of Cu interconnect with low-K dielectric to reduce overall capacitance.

The collective result of these advances supports a high-volume commercial device with 140 million switching devices whose gate oxide is less than five atomic layers thick. This enables the device to draw less than 1 W average power at greater than 10% less Thermal Design Power (TDP), which is at 1.8 GHz as compared to the original Intel Pentium M processor.

2 MB Power-Optimized L2 Cache

The L2 cache is one of the most effective business productivity per watt enhancements offered by mobile processors today. The second-generation Intel Pentium M processor offers an unprecedented 2 MB, low-power, low-latency L2 cache to efficiently tackle the growing working sets of present and future applications. The effect of on-die cache is to reduce the number of cycles wasted while waiting on memory. Not surprisingly, the percent performance impact of this grows with increasing processor frequency. When operating at the maximum performance level tasks get done sooner allowing the entire system to return to a low-power state. Power is also conserved while in the adaptive state, a Windows XP Power Management Policy, which allows the processor to switch voltage and frequency dynamically to address processor demand. Maximizing the potential of each intermediate-level performance point reduces the likelihood of crossing the utilization threshold that initiates a jump to the next frequency. This has a linear effect on power, due to frequency, and a square effect due to voltage ($P \propto cfv^2$) and corresponding voltage level, which further reduces power demands.

The performance/power benefits of a larger L2 cache vary depending on the workload. Office productivity usage models (MS Word*, Excel*, Virus Scan*, etc.) typically have large data sets that force lots of off-chip accesses to system memory; a larger L2 cache is definitely a plus in such an application environment.

* Other brands and names are the property of their respective owners.

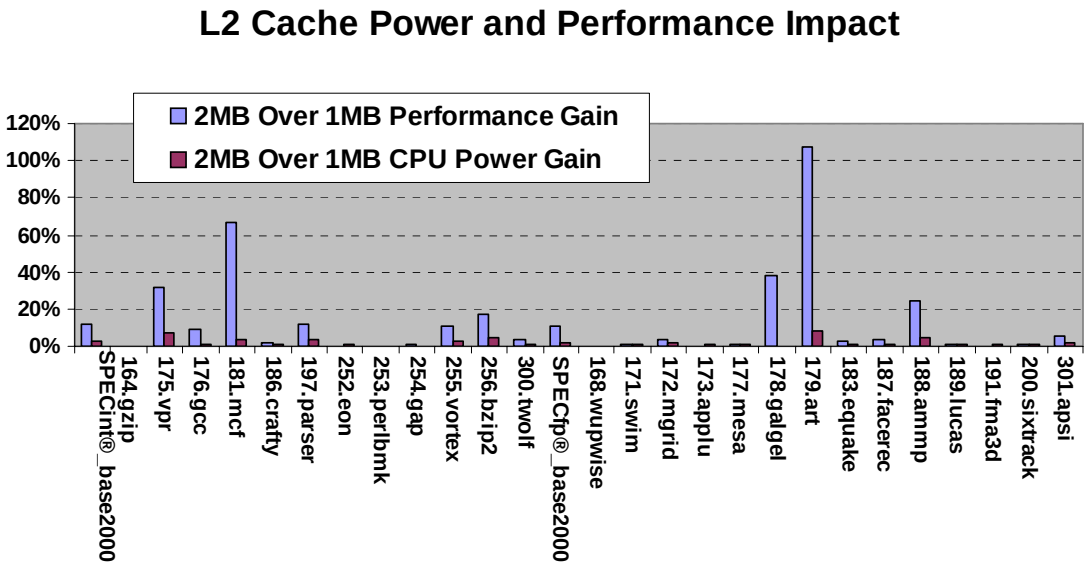


Figure 2: Performance and power comparison of 2 MB vs. 1 MB L2 cache

In a comparison of L2 cache sizes of 2 MB compared to 1 MB when running application-based benchmarks such as MobileMark 2002 and SYSmark 2004, we see ~5-7% performance gains for a 2 MB L2 cache over a 1 MB L2 cache with very little increase in processor power. For scientific-application candidates, we can use SPEC CPU2000 as a good proxy that shows that 2 MB outperforms 1 MB by ~11%, and processor power only increases by ~2-3%. Certain SPEC CPU2000 sub-tests show even larger benefits with a larger 2 MB L2 cache: for e.g., 175.vpr shows a ~30% and 7% performance and processor power increase, respectively with a 2 MB L2 cache; 181.mcf shows a ~65% and 3.5% performance and processor power increase, and a 179.art performance more than doubles for a ~8% processor power increase. These are excellent examples of mobile design tradeoffs that increase performance practically for free.[1] Figure 2 illustrates the performance improvements of the 2 MB cache compared to the increase in power.

An important concept to point out is that any time savings from getting the job done quicker, due to the 2 MB L2 cache, translates into additional platform power savings. This is accomplished by allowing various platform power-management policies to execute sooner, which forces the entire platform to go into its lowest power state. The net result is a much more energy efficient platform for workloads that benefit from this behavior.

Faster Front-Side Bus

With the previous platform, both the processor and memory components worked in near synchronous bandwidth levels. For example, traffic between the processor and GMCH runs at ~3.2 GB/s whereas traffic between GMCH and Memory is at ~2.7 GB/s. Note that the previous platform design was bottlenecked by system memory. One of the key bottlenecks that needed to be addressed with the new generation was the increasing gap between the available memory bandwidth and how it was available to the processor.

Two major improvements took place in the memory sub-system design. First, the second-generation DDR DRAM, DDR2, enabled much faster speeds of 533 MT/s (millions of transfers per second) compared to 333 MT/s for DDR. Secondly, the Intel 915GM Express Chipset memory controller provides two channels to simultaneously access system memory. Figure 3 compares this difference in topology. The net result is a significant increase in available bandwidth (533 MT/s* 8 bits *2 ch = ~8.4 GB/s) for the platform. This allows the FSB to operate at the maximum FSB throughput of 4.2 GB/s (533 MHz * 8 bits = ~4.2 GB/s). Figure 4 compares a second-generation Intel Pentium M processor at a clock frequency of 2.0 GHz with 400 MHz and 533 MHz FSB under SPEC CPU2000. During these experiments the 533FSB was run with DDR2-533 MT/s memory, and the 400FSB was run with DDR2-400 MT/s memory.

[1] Source: Intel Pentium M Processor 2.26 GHz– Intel 915 Express Chipset, Intel Customer Reference Board, 2x512MB DDR2 533MHz S0-DIMMs, 60GB 7200RPM Hitachi* Travelstar HTS7260609AT00 HDD

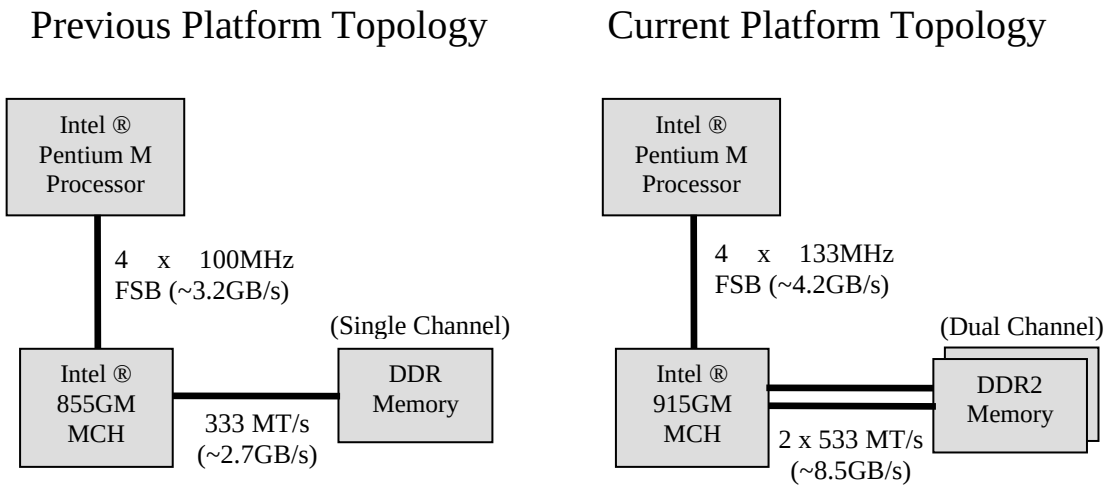


Figure 3: Comparison of memory topology for the previous platform vs. the new platform

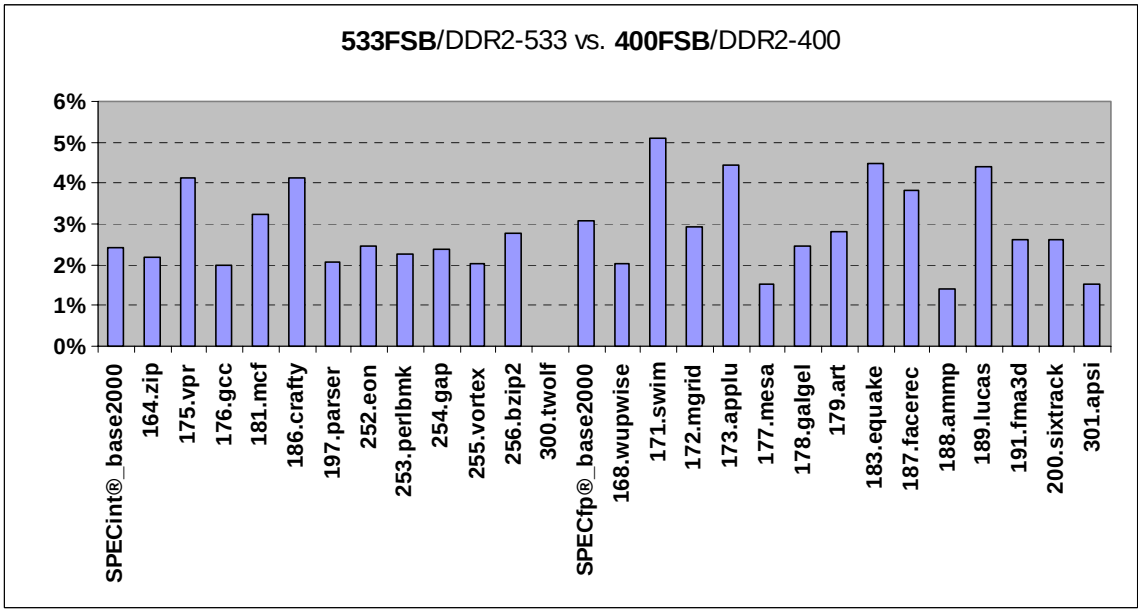


Figure 4: Front-side bus comparisons (533 MHz vs. 400 MHz)[2]

Enhanced Register Access Manager

The register renaming support provided by the Pentium Pro processor sought to relieve the limited number of eight architectural registers provided by the iA instruction set. It did this by recognizing "false dependencies" within the instruction flow and providing virtually unlimited distinct physical locations for results when said cases were identified. Previous to register renaming the practice of partial register writes provided relief to certain code that produced changes to a restricted 8- or 16-bit portion of the register.

This older technique adversely impacts the register renaming logic as it forces the renaming of the referenced register to stall until the partial register is written. Thus the throughput of the complementary one or two instructions that reference the alternate portion, or the entire register, is reduced.

[2] Source: Intel Pentium M Processor 2.0 GHz– Intel 915 Express Chipset, Intel Customer Reference Board, 2x256 MB DDR2 533 MHz S0-DIMMs, 60 GB 7200 RPM Hitachi Travelstar* HTS7260609AT00 HDD

To minimize the impact of such legacy code, the processor defines a "calculate and merge" fused μOP and ALU support that provides for the partial references to be calculated in parallel. This μOP provision removes the stall condition that would otherwise impact the register renaming mechanism, regaining anywhere from 0 to 20% performance on legacy application scenarios.

Advanced Tight Loop Execution

The processor recognizes even the smallest opportunity to apply clever solutions to increase performance while reducing power: for example, when a small well-behaved loop is found to be contained within the Instruction Length Decode Queue, power is conserved by stalling the Instruction Fetch Unit and running code from the Prefetch Buffer. As the loop characteristics are well understood, the Branch Prediction Unit can be powered down saving even more power.

Execute Disable Bit Support

The predominant vulnerability to system attack is from malicious code that is stored as data and later run via an intentional corruption of a return address. The second-generation Intel Pentium M processor provides protection from such mechanisms through support for the Execute Disable (XD) bit functionality. This functionality is enabled when operating in the Physical Address Extension (PAE) mode, and the XD bit is set at the Page Directory and/or Table Entry level. Once set, attempts to execute "instructions" within the sequentially referenced pages of memory will result in a page fault to the OS. This added protection is enabled in Windows XP Service Pack 2 without cost to performance or battery life.

Processor Summary

The second-generation Intel Pentium M processor continues its success in responding to the increasing performance demands of existing and future workloads and continues to effectively reduce power consumption that would inevitably increase if not addressed. Traditional techniques such as increased capacity (2MB L2 cache) and increases in frequency (processor clock and FSB) have been coupled with other novel techniques described to address these concerns. Workload analysis has shown these techniques to be effective in making the new processor a worthy successor to its predecessor.

GMCH

The Intel 915GM Express Chipset is the integrated graphics memory controller hub (GMCH) for the second-generation platform built on Intel Centrino mobile technology. The chipset supports several new/faster interfaces such as a 533 MHz FSB, Double Data Rate II (DDR2) Dual Channel, a PCI Express-based (PCIe-based) Direct Media Interface (DMI) link to the ICH6-M, and an x16 PCIe bus. There are also several power-management enhancements including C2 Pop-up and Rapid Memory Power Management (RMPM). The GMCH includes support for Pixel Shader 2.0 and DX9, integrated TVout, and improved power and performance techniques such as Intel Display Power Savings Technology 2.0 (DPST2).

Frequency Enhancements for GMCH

The GMCH supports an increase in frequency on the FSB from 400 MHz to 533 MHz. The performance results of this change are covered in detail in the processor section of this paper. In addition to an increase in the FSB frequency, the Intel 915GM Express Chipset now adds support for DDR2 dual-channel memory extending the memory transfer rate to 533 MT/s. Further details of the impact of these new features are discussed in the ICH6-M section.

Improved Interfaces for GMCH

Direct Media Interface

The Direct Media Interface is discussed in the ICH6-M section.

PCIe Graphics/IO

The GMCH contains a new universal scalable serial interface with a maximum number of 16 lanes of PCIe for the graphics interface to allow for maximum performance when needed. Different PCIe interfaces are used on other components of the platform as well. The serial interface in ICH uses this interconnect for devices such as Gigabit Ethernet and WLAN. For more details regarding PCIe please refer to the ICH6-M section of this paper.

GMCH Power-Savings Techniques

PCIe Active State Power Management

PCIe Active State Power Management (ASPM) is provided to reduce the power of the PCIe serial bus links. Although serial bus links provide great performance and bandwidth, they come at a high power cost and must be actively power managed to reduce the power of the controller (GMCH) and the graphics component. Three link states (L0, L0s, L1) are available. Power savings of approximately 1.2 W for the GMCH have been measured when the L0s power-management state is utilized on a x16 PCIe lane for graphics. Further power savings of an additional 0.3 W have been measured if the lower power L1 state is enabled.

C2 Pop-up

The C2 Pop-up feature was originally implemented to solve the problem when the UHCI USB master prevented entry of the processor to the C3/C4 state. This feature not only effectively addresses this problem but also allows the processor to enter a lower power C3/C4 state for I/O data transfers such as PCIe bus master traffic. Previous platforms only allowed the processor to enter the C3/C4 states when certain periods of inactivity were encountered. The new platform allows entry into C3/C4 and utilizes C2 Pop-up to place the processor into the higher power state only when necessary.

Evidence of power savings is demonstrated with Windows XP Idle. Without C2 Pop-up support C2 residency is ~95% in Idle. With C2 Pop-up support all 95% of the C2 residency in Idle is shifted to C4 residency. This accounts for significant power savings due to the difference in power consumption for the processor at these states. Processor power in C2 Pop-up can be 1.5-6 W higher than processor power in C4. Greater power consumption in C2 Pop-up is due to the increased processor voltage levels when it is entered at higher clock frequencies. This power savings may already be realized with some USB devices that already support selective suspend, which gives a similar result. This feature, however, provides the same benefit to all UHCI-attached USB devices.

Rapid Memory Power Management (RMPM)

RMPM is a feature in the chipset that saves platform power by checking the processor utilization. When the processor is in C2-C4 states it does not need to access memory, thus allowing memory to enter Self Refresh mode to save power. This also allows the MCH to turn off logic related to reading/writing memory to save power. This feature provides significant power savings due to the degree of clock gating and DLL shutdown on GMCH allowed during this state.

With MobileMark 2002, we see significant power savings due to the high percentage of C3/C4 processor C-state residency during extended periods of Idle built into the workload. Over 90 minutes of MobileMark 2002 GMCH power savings from this feature is ~1 W. Figure 5 illustrates this power savings. Memory power decreases by ~0.3 W as well. This feature along with other power-savings techniques working together saves additional power, as described later. Workloads such as SPEC CPU2000 and 3DMark03 are more processor and memory intense so they spend almost 100% of the time in C0; this means that the RMPM feature would not save power in these other workloads since no time is spent in the C2-C4 states.

Dynamic Row Power Management

Dynamic Row Power Management (DRPM) is one of the more useful power-savings features of GMCH since it generally saves power under any workload without impacting performance for the investigated workloads. This feature existed in the previous-generation GMCH, but it is particularly highlighted here for its ability to save power on the GMCH and Main Memory regardless of workload. Under MobileMark 2002, this feature alone saves ~0.4 W on GMCH and another ~0.4 W on the memory. Figure 5 illustrates this power savings. It works well with other power-savings features to save more power than each feature independently. Under SPEC CPU2000 and 3DMark03, DRPM saves at least 0.1 W on GMCH and 0.2 W on the memory, without performance degradation. These savings help to increase the battery life of the notebook and do so with no performance impact.

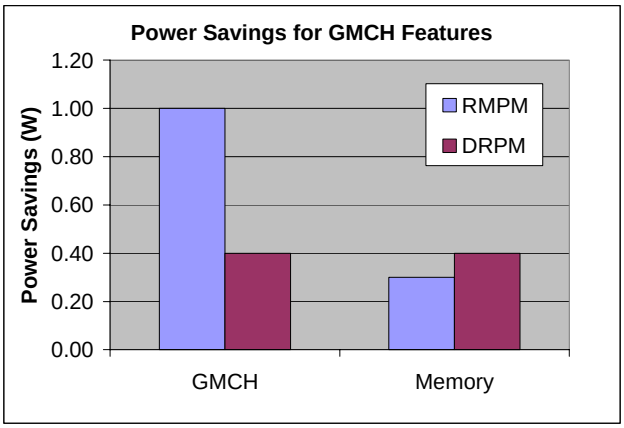


Figure 5: MobileMark 2002 power savings per feature

Graphics Core Features

So far in this paper, the MCH portion of the Intel 915GM Express Chipset has been our focus; we now briefly describe the graphics portion of the chipset [12]. The feature set for the chipset integrated graphics includes increased core frequency of 200 MHz to 333 MHz, support for DirectX9 over the previous platform support for DirectX7.1, and four pixel pipes over the two in the previous platform. Also new is hardware acceleration for Pixel Shader 2.0 (Shadow Maps, Volumetric Textures, Slope Scale Depth Bias, and Two-Sided Stencils) and software accelerated vertex shaders. These graphics features coupled with DDR2 Dual-Channel support result in improved 3D performance. This is seen in 3D benchmarks including 3DMark2001 SE and 3DMark03. 3DMark2001 SE scores of the GMCH are more than double that of previous-generation GMCHs.

Intel Display Power Savings Technology 2.0

DPST2.0 has enhancements over the first implementation including a more accurate picture representation. Power savings as before are primarily in the LCD backlight, and levels of savings are dependent on the original brightness and power of the panel, the image displayed, and the aggressiveness level that DPST2.0 is set to. There is no performance impact and no platform power increase due to additional computations. Additional circuitry for this feature is marginal, and 15-25% power savings for the LCD backlight may be achieved. Power savings for MobileMark 2002 with DPST2.0 has been measured at ~0.5 W for low power 14.1" SXGA+ panel at ~60 nits. DVD Playback[3] power savings for DPST2 has been measured at ~1.1 W as well as ~0.8 W savings during 3DMark03 with the same panel and brightness. Figure 6 illustrates these power savings. Power savings depend on the panel power, the panel type, and the LCD brightness so the absolute power savings may be higher even though the percentage savings should be similar.

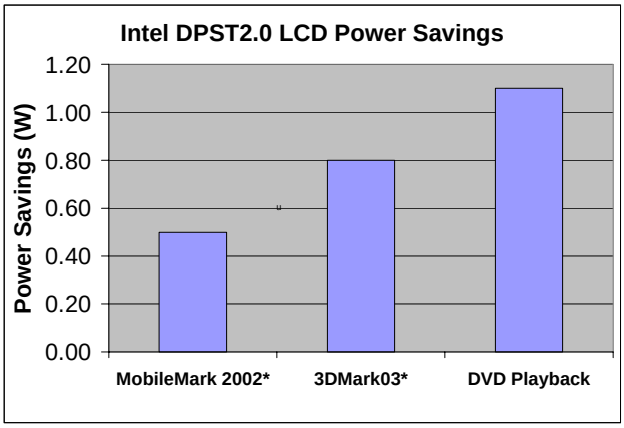


Figure 6: Power savings from DPST2.0

GMCH Summary

All together the GMCH has enabled several performance features for the platform. These include faster buses to the processor, memory, external graphics, and ICH. They also include an increase in frequencies for the FSB and DDR memory interface. The new performance features are paired with platform power-reduction techniques that not only reduce GMCH power consumption but also impact components such as the processor, memory, and LCD backlight.

[3] DVD playback software used for testing is Intervideo WinDVD Platinum 5

SYSTEM MEMORY

The new platform supports next-generation memory technology by implementing the DDR2 specification [10], an evolutionary technology that extends DDR [11] (supported on the previous platform). The new platform also introduces a second memory channel to the system memory, effectively doubling the total available memory bandwidth. With dual-channel DDR2 memory support, new platforms based on next-generation Intel Centrino mobile technology have increased peak bandwidth and lower Small Outline Dual Inline Memory Modules (SO-DIMM) power consumption over the previous generation.

New Feature Support

DDR2 Support

The DDR2 specification allows increased clock rates over DDR. The new platform supports a peak bandwidth of 4.3 GB/s, a 60% increase over the previous platform having a peak bandwidth of 2.7 GB/s (DDR 333 MT/s). The data path width remains the same so the increased peak bandwidth comes from the increased transfer rate from 333 MT/s to 533 MT/s. The platform also supports a 400 MT/s transfer rate. The DDR2 specification has a maximum transfer rate of 667 MT/s, which is not supported on the second-generation Intel Centrino mobile technology, but may be supported on future platforms.

The other major change in the DDR2 specification is the change to a 1.8 V operating voltage. In the DDR specification, the operating voltage is 2.5 V. This means that memory operations now consume less power on platforms based on next-generation Intel Centrino mobile technology, enabling increased battery life.

Single-Channel and Dual-Channel Modes

The next-generation platform built on Intel Centrino mobile technology introduces support for dual-channel mode when both SO-DIMM slots are populated (otherwise the platform operates in single-channel mode). This technique is commonly used on desktop systems. In this configuration, the memory controller may access both SO-DIMMs, simultaneously resulting in 8.5 GB/s maximum theoretical aggregate bandwidth.

Feature-Specific Workload Analysis Data

System Memory Performance and Power

With the increased memory frequency and dual-channel capability, the new platforms are capable of increased performance over the previous-generation Intel Centrino mobile technology. The amount of performance increase varies depending on the workload. System memory power consumption depends mainly on the memory vendor and to a lesser degree, on frequency. Certain memory vendors

design memory components with lower power consumption than others.

Under a typical business productivity usage model, system memory enters the low-power Self Refresh mode when there is no activity. Under this usage model, a platform based on next-generation Intel Centrino mobile technology consumes approximately 40% less power than a platform based on previous-generation technology[4].

When the new platform is executing a processor-to-memory intensive workload, increased memory frequency and dual-channel mode result in higher performance for a modest increase in memory power. In this configuration, the processor cannot take full advantage of the second memory channel because of the FSB limitation. Under this workload, a platform based on the next-generation Intel Centrino mobile technology scores approximately 40% higher than a previous generation system with a modest 6% power increase.

ICH6-M

The Intel 82801FBM I/O Controller Hub (ICH6-M) [6] enables next-generation Intel Centrino mobile technology by introducing support for current and future peripherals:

- PCIe devices
- SATA devices with Advanced Host Controller Interface support.
- Intel HD audio devices.

In addition to these new interfaces, ICH6-M provides support for USB2.0 as on ICH4-M [1] but adds two additional ports. The ICH6-M also provides the interface for the Intel PRO/Wireless 2915ABG wireless module and ExpressCard* devices.

ICH6-M provides these devices high-speed access to system memory via the DMI, a proprietary PCIe-based high-speed link to the GMCH.

New Feature Support

PCI Express

The ICH6-M supports the PCIe Base Specification, Revision 1.0a [2]. PCIe is a third-generation high-performance general input/output interconnect. Leveraging existing PCI architecture, PCIe introduces a scalable, point-to-point, power-managed, serial interface. Each basic connection is constructed with low-voltage, differential signal pairs.

The ICH6-M implements four PCIe root ports and each x1 port is capable of supporting 250 MB/s bandwidth in a single direction (500 MB/s concurrent). A connection between two PCIe ports is created with a link. An xN Link is built from N Lanes. Each lane is a set of differential signal pairs, one to transmit and the other to receive. For example, the ICH6-M can create a maximum of four x1 Links between itself and other PCIe-compliant devices.

Legacy bus power management states are replaced by new link power management states to account for the new interface. Link states are defined by the power management D-states of connected components or by Active State Power Management (ASPM) protocols. The ICH6-M implements L0, L0s, and L1 Link power management states for PCIe devices.

Serial ATA and AHCI

The ICH6-M supports the following specifications:

- SATA Specification, Revision 1.0a [3].
- Several optional sections of the SATA II: Extensions to SATA 1.0a Specification, Revision 1.2 [4].

SATA is the next-generation replacement for Parallel ATA (PATA). It is a serial interconnect for mass storage devices utilizing high-speed differential signals. The ICH6-M supports the Intel SATA AHCI Specification, Revision 1.0 [5]. AHCI enables better performance and additional power-management capability over the base SATA specification.

The ICH6-M provides two SATA ports each with a maximum bandwidth of 150 MB/s with independent DMA operation. Coupled with AHCI, functionality is extended to incorporate hardware-assisted native command queuing and aggressive host-initiated device/bus power management. Native command queuing allows multiple commands issued to the device to be internally queued and serviced for better performance [7]. Combined operation with SATA and legacy PATA devices is supported, but AHCI operation must be disabled under certain platform configurations.

[4] The next-generation Intel Centrino mobile technology system is configured with 512 MB of 533 MT/s DDR2 memory in dual-channel mode. The system based on previous-generation Intel Centrino mobile technology is configured with 512 MB of 333 MT/s DDR. Though the previous-generation Intel Centrino mobile technology-based platform has both SO-DIMM slots populated, it is not capable of operating in dual-channel mode.

* Other brands and names are the property of their respective owners.

Intel High Definition Audio

The ICH6-M supports the Intel HD Audio, Revision 1.0 [8] and is Universal Audio Architecture compliant. Intel HD Audio can support 192 kHz multi-channel output vs. 96 kHz stereo-channel output with AC'97 [9]. The two architectures are incompatible, and concurrent operation of Intel HD Audio and AC'97 functionality on ICH6-M is not supported.

The ICH6-M Intel HD Audio controller supports up to three external codecs, each of which is configured by software. Codecs are connected to the controller via a serial link, which is shared with the AC'97 controller. Codecs render audio streams when DMA requests are initiated by the controller. Each audio stream may contain one or more channels. Each channel can be dynamically assigned to a single codec converter.

Because Intel HD Audio data stream traffic does not require snooping of the processor cache, it enables processor entry into the C3/C4 state when the system is otherwise idling. This allows for significant power savings from the processor of at least 600 mW for DVD playback and CD audio playback [7].

Direct Media Interface

To meet the device-to-memory bandwidth requirements of PCIe, SATA, USB 2.0, Intel HD Audio and others, a proprietary serial interface, based on PCIe, was developed, the DMI link. This link connects the ICH6-M and the GMCH. It offers 2 GB/s maximum bandwidth compared to 266 MB/s available with the ICH4-M hub interface. The DMI integrates priority-based servicing to allow concurrent traffic and isochronous data transfer capabilities for Intel HD Audio.

ICH6-M Summary

The ICH6-M device introduces support for many new interfaces providing enhanced bandwidth for high-bandwidth applications and devices. SATA with AHCI support enables increased disk bandwidth, performance benefits from native command queuing, and more aggressive power management. Furthermore, Intel HD Audio provides better audio support over AC'97 in terms of audio quality, device features, and power management. Finally, the DMI provides high-bandwidth access to system memory.

WIRELESS LAN

Next-generation Intel Centrino mobile technology carries with it a full range of WLAN support with the Intel PRO/Wireless 2915ABG WLAN solution. This product not only supports 802.11a, 802.11b, and 802.11g, but it also provides the Intel Wireless Coexistence System (WCS) v.2 that mitigates wireless interference between Bluetooth[Δ] and WLAN. This solution also supports industry-standard wireless security features including WPA2 and Cisco* Compatible Extensions v.3.

Intel PRO/Wireless 2915ABG is capable of providing high-quality streaming video playback. An Intel Centrino mobile technology client containing this card is capable of >99% packet arrival rate[5] with a 5Mbps MPEG2 stream thus producing a high-quality image. Figure 7 demonstrates the difference in video quality that may be observed for a data stream arriving at less than 99% packet arrival.

[Δ] Bluetooth is a trademark owned by its proprietor and used by Intel Corporation under license.

* Other brands and names are the property of their respective owners.

[5] Source: Intel PRO/Wireless 2915ABG Network Connection; Driver: 9.0.0.60; Software TIC:87406; 802.11A/B/G Access Point: Cisco Aironet 1200 ABG*, System Firmware v12.2(15)JA3; All Platforms–Intel Pentium M Processor 1.40 GHz–Intel 855PM chipset, IBM* X31, BIOS: 1.01, 256 MB PC2100 DDR266, ATI* Mobility* 9000 w/ graphics driver, ATI* Mobility* 9000 6.13.10.6269, LCD Screen: 12.1" TFT 1024x768x32-bit color; IBM* IC25N04ATCS05-0 TravelStar* 40 GB hard drive 5400 rpm, Laptop running on AC power, Windows* XP* Professional SP1 Build 2600. Chariot* Console version 5.0 with Chariot* Endpoint version 4.5. 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.

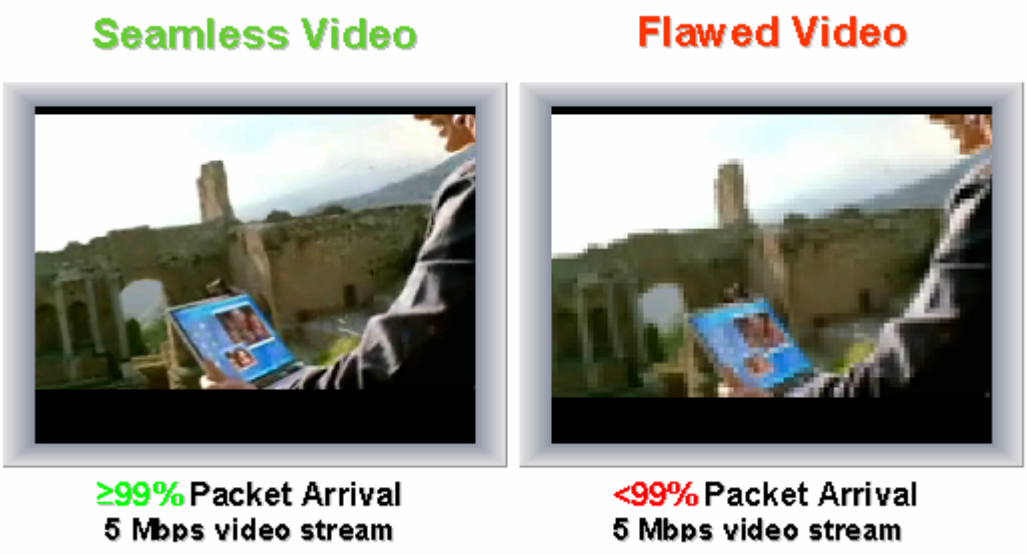


Figure 7: Example of video quality

In addition to quality video streaming, the Intel PRO/Wireless 2915ABG provides support for low-power states when idle. Power consumption is ~100 mW in idle for both associated and unassociated states to the AP. With the radio disabled the device consumes even less power, around 20 mW. Usage models such as Web browsing and e-mail can take advantage of these low-power-consuming states as typical usage in these cases is idle to a large degree.

The WCS v.2 feature allows for coexistence of both Bluetooth and 802.11b/g transmitter/receivers on Intel Centrino mobile technology while mitigating interference problems that result when these two protocols are exercised concurrently. Both 802.11b/g and Bluetooth transmit and receive on the 2.4 GHz base band. Without this feature wireless throughput rates may be significantly reduced.

CONCLUSION

Collectively, the enhancements in the second-generation PC platform built on Intel Centrino mobile technology deliver generous processing power to the mobile business platform while maintaining robust battery-life levels. When a given performance level is accomplished at less utilization of system resources less platform power is consumed. This power can be applied to additional work or more demanding tasks. While today's benchmarks may not take advantage of the full potential of this platform, this platform can handle the work levels sure to be present in tomorrow's applications and work environment. This paper has illustrated the wide range of usages that the second-generation platform built on Intel Centrino mobile technology is suited for and the impact of these usage models on system performance and power consumption.

TABLE OF ACRONYMS

AHCI	Advanced Host Controller Interface
ALU	Arithmetic Logic Unit
ASPM	Active State Power Management
DDR	Double Data Rate
DLL	Delay Lock Loop
DMA	Direct Memory Access
DMI	Direct Media Interface
DPST2	Display Power Savings Technology 2
DRAM	Dynamic Random Access Memory
DRPM	Dynamic Row Power Management
DVD	Digital Versatile Disk
EHCI	Enhanced Host Controller Interface
GMCH	Graphics Memory Controller Hub
ICH	I/O Controller Hub
HD	High Definition
LCD	Liquid Crystal Display
MCH	Memory Controller Hub
OEM	Original Equipment Manufacturer

PAE	Physical Address Extension
PATA	Parallel Advanced Technology Attachment
PCI	Peripheral Component Interconnect
PCIe	Peripheral Component Interconnect Express
FSB	Front-Side Bus
SATA	Serial Advanced Technology Attachment
SO-DIMM	Small Outline Dual Inline Memory Modules
SXGA+	Super eXtended Graphics Array Plus
μOP	Micro-Op
USB	Universal Serial Bus
WCS	Wireless Coexistence Solution
WLAN	Wireless Local Area Network
XD	Execute Disable

ACKNOWLEDGMENTS

Thanks to Les Cline and Nirav Shah for contributing to the ICH6-M section of this paper. Thanks also to Paul Cannan for providing information regarding Intel PRO/Wireless 2915 features and benefits.

REFERENCES

[1] ICH-4M product documentation.

[2] PCI Express Base Specification, Revision 1.0a, PCI-SIG Web site. http://www.pcisig.com/specifications/pciexpress/base/*.

[3] Serial ATA Specification, Revision 1.0a, Serial ATA International Organization Web site. http://www.serialata.org/specifications.asp.*

[4] Serial ATA II: Extensions to Serial ATA 1.0a Specification, Revision 1.2. Serial ATA International Organization Web site. http://www.serialata.org/specifications.asp.*

[5] Advanced Host Controller Interface Specification for Serial ATA, Revision 1.0, Intel Corporation Web site. http://www.intel.com/technology/serialata/ahci.htm.

[6] ICH6-M product specification.

[7] Tom Clark et al., "Low-Power Audio and Storage Input/Output Technologies for the Second-Generation Intel® Centrino™ mobile technology Platform," *Intel Technology Journal*, Volume 9, Issue 1, 2005.

[8] Intel High Definition Audio Specification, Revision 1.0. Intel Corporation Web site. http://www.intel.com/standards/hdaudio/.

[9] AC'97 Specification.

[10] JEDEC Specification JESD79-2A (DDR2). JEDEC Web site. http://www.jedec.org.*

[11] JEDEC Specification JESD79D (DDR). JEDEC Web site. http://www.jedec.org.*

[12] Kam Leung, et al., "High-Performance Graphics and TV Output Comes to the Second-Generation Intel® Centrino™ mobile technology Platform," *Intel Technology Journal*, Volume 9, Issue 1, 2005.

AUTHORS' BIOGRAPHIES

Eunice Chang received her B.S. degree in EECS from the University of California at Berkeley and joined the Intel Rotation Engineering Program in 1999. She is currently a system architect in Intel's Mobility Group responsible for platform power and performance analysis of notebook systems. Her current interests include analysis of chipset power management features and mobile platform average power. Her e-mail is eunice.y.chang at intel.com.

Jim Fung received his B.S. and M.S. degrees in Computer Engineering from the University of Illinois at Urbana-Champaign and joined the Intel Rotation Engineering Program in 2001. Currently he is a platform architect in Intel's Mobility Group. Since joining MPG in 2003 he has been responsible for platform power analysis and processor performance. His research interests include processor and systems architecture. His e-mail is jim.t.fung at intel.com.

Greg Kaine received his B.S. degree in Computer Engineering from Purdue University at West Lafayette, IN and joined the Intel Rotation Engineering Program in 2001. Currently he is a platform architect in Intel's Mobility Group and is responsible for analysis of platform power and performance for mobile workloads. His current interests include analysis of mobile applications and multimedia workloads. His e-mail is greg.d.kaine at intel.com.

Ali Saeed received his B.S. degree in Electrical Engineering from the University of Houston. Currently he is a senior systems architect in Intel's Mobility Group. His primary research area includes performance and power profiling of processors and chipsets with an emphasis on performance, power, and battery-life estimations. Prior to joining Intel, Ali was manager of performance analysis in

Compaq's Commercial PC Group. His e-mail is ali.saeed at intel.com.

Randy P. Stanley received his B.S. degree from the University of California at Berkeley, CA. Since joining Intel in 1981 he has been responsible for the memory management implementations in the Pentium and Pentium II processors and built Intel's current mobile power and performance analysis lab. His current interests include the development of power and performance forecasting models. His e-mail is randy.p.stanley at intel.com.

THIS PAGE INTENTIONALLY LEFT BLANK

Interface Material Selection and a Thermal Management Technique in Second-Generation Platforms Built on Intel® Centrino™ Mobile Technology

Eric C. Samson, Mobility Group, Intel Corporation
Sridhar V. Machiroutu, Mobility Group, Intel Corporation
Je-Young Chang, Technology and Manufacturing Group, Intel Corporation
Ishmael Santos, Mobility Group, Intel Corporation
Jim Hermerding, Mobility Group, Intel Corporation
Ashay Dani, Technology and Manufacturing Group, Intel Corporation
Ravi Prasher, Technology and Manufacturing Group, Intel Corporation
David W. Song, Technology and Manufacturing Group, Intel Corporation

Index words: cooling, TIM, thermal management, Intel® 915, Pentium® M processor

ABSTRACT

The mobile thin and light platform has a limited cooling capability, in part due to a form factor that limits the volume available for the thermal solution. The high performance of the Pentium® M processor on 90nm process technology and the Intel® 915 Chipset in the second-generation platform built on Intel® Centrino™ mobile technology demands a high electrical power and generates substantial heat, presenting a challenge to the thin and light notebook system designer. In this paper, we addresses two methods of dealing with the thermal challenge.

First, we discuss the path-finding effort to improve the thermal interface materials (TIMs) that allow a good thermal contact between processor and thermal solution, minimizing the transistor temperature of the bare-die Pentium M processors. Two tester methodologies are described, and the need to test TIMs under mobile usage conditions is emphasized. We also discuss the reliability test methodology for TIMs with a focus on the effect of mobile usage conditions affecting long-term reliability of the TIM.

We then focus on power-based thermal state estimation as a platform thermal management technique. This technique is used to detect and limit the thermal impact of power virus[1] workloads. In the second-generation platforms built on Intel Centrino mobile technology, the Intel 915 Chipset Graphics and Memory Controller Hub (GMCH) is uniquely positioned to understand much of the workload for the platform. The Intel 915 GMCH has implemented filter-based thermal management. Several key usage models for filter-based thermal management are explored in detail: detecting and limiting the impact of power viruses on system memory and detecting and limiting the impact of power viruses on chipset memory controller hubs.

INTRODUCTION

Notebook system designs vary significantly from designer to designer; however, they are all densely packed with components and devices, which leave the system with little room for cooling. The problem is compounded by the limited room inside a thin and light notebook product, which typically has a one-inch total thickness when folded and a 17 mm inner vertical space in the lower half of the notebook computer. Figure 1 shows a schematic electrical layout of the major components in the second-generation

[1] A power virus is an unusually intensive workload that maximizes power consumption. Most useful applications draw only a fraction of the power a power virus consumes.

platforms built on Intel Centrino mobile technology. Figure 2 shows a layout of platform-based notebook system that is roughly representative of performance thin and light notebook designs in the industry. In general, the low-profile thin and light form factor limits the flexibility of thermal solution choices for the system components that must be cooled in order to get any appreciable performance.

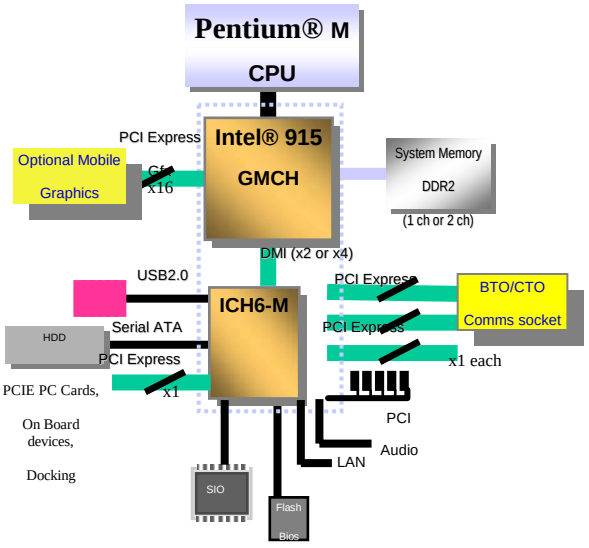


Figure 1: Electrical schematic of the second-generation platforms built on Intel Centrino mobile technology

Figure 3 shows the use of the remote heat exchange, the predominant method of cooling of high-power components that require dedicated active cooling. In remote heat exchange, the thermal energy is transported to a location, typically via a heat pipe, where a larger fan and heat exchanger can be used. Also shown in Figure 3 are the silicon portion of the hot component (bare die assumed), the attached hardware for coupling the thermal solution to the hot component, and the key temperature monitor points typically used to characterize the performance of the solution. In the first section, we discuss the ability to transfer the thermal energy from the processor to the thermal solution by using thermal interface materials (TIMs).

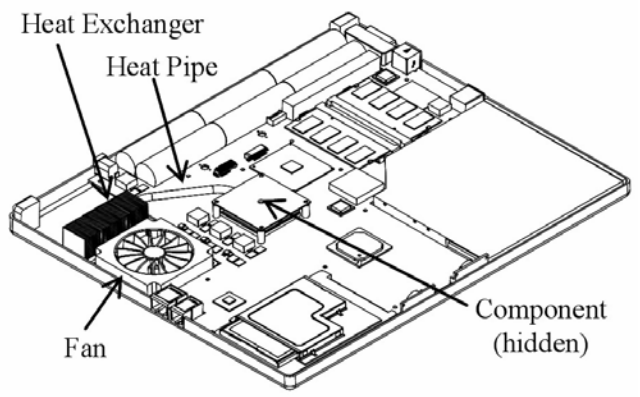


Figure 2: Layout of typical thin and light notebook (base only with top surface removed)

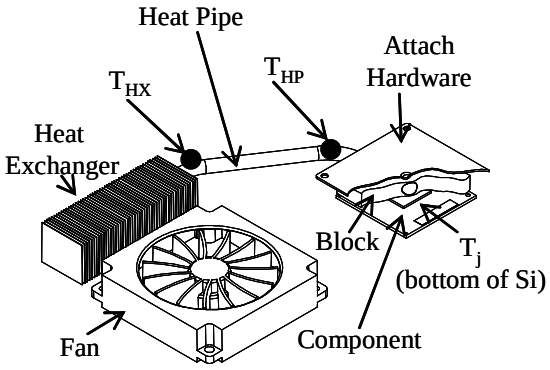


Figure 3: Remote heat exchanger using a heat pipe embedded in a block

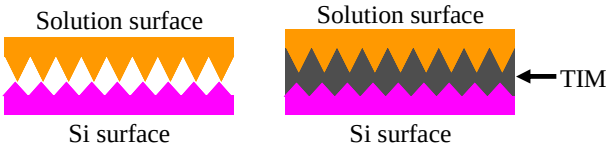
The thermal solution shown in Figure 3 is used to cool the processor, and the fan allows for the cooling of the platform by pulling air flow over various components in the platform. Under thermally high-stress applications such as a power virus or due to improper design of the thermal solution, the platform components could generate a severe thermal environment. To protect the notebook computer and the component functionality, Intel has built multiple thermal management elements into the second-generation platform built on Intel Centrino mobile technology.

THERMAL INTERFACE MATERIALS

For the bare-die Pentium M processor, a successful thermal solution design would allow a minimal temperature drop from the silicon transistor temperature to the ambient temperature. A remote heat exchanger (RHE) is used as the thermal solution for cooling the Pentium M processors in the second-generation platforms built on Intel Centrino mobile technology for thin and light systems. Figure 3 shows a schematic of an RHE. The efficiency of heat draw by the heat pipe (in the attach

block) from the processor depends on the quality of thermal contact between the attach block and the processor. The lower the thermal contact resistance, the lower the temperature drop from the silicon transistor to the ambient.

Even in a direct contact, the processor and the attach block do not transfer heat efficiently, because the quality of contact between two non-conforming solid surfaces is poor, as shown in Figure 4 (a). To enhance the thermal contact between the processor and the attach block, TIMs are inserted into the interface, as shown in Figure 4 (b). Under mechanical pressure, the soft TIMs conform to the microscopic surface contours of the adjacent solid surfaces and increase the (microscopic) area of contact between the thermal solution surface (block) and the silicon die surface (processor) and therefore reduce the temperature drop across this contact.



a) Dry Contact b) Contact using a TIM

Figure 4: Magnified schematic of contact between Si surface and Solution surface

The quality of contact between the processor and the attach block, or TIM performance, depends on the quality of the thermal conduction through the TIMs and the quality of contact between the TIMs and the two surfaces. Mathematically this can be represented as follows [1]:

$$\theta_{TIM} = \frac{BLT}{k_{TIM}} + R_c \qquad \text{Equation 1}$$

where θ_{TIM} is the effective performance of the TIM, k_{TIM} is the bulk thermal conductivity of the TIM, BLT (Bond Line Thickness) is the thickness of the TIM under usage, and R_c is the contact resistance between the TIM and the mating surfaces. The contact resistance is mainly due to the irregularities or roughness of the surfaces of processor and attach block, so the resistance is negligible, if the surfaces are perfectly smooth.

Based on Equation 1, we can consider three approaches (or strategies) for reducing θ_{TIM}. One approach is to increase the conductivity of TIM, k_{TIM}. This is generally done by using a high thermal conductivity material (like metal or liquid metal) or a low thermal conductivitiy base material loaded with highly conductive particles. Another approach to reduce θ_{TIM} is to reduce the BLT of the TIM. This is done by reducing the bulk modulus of elasticity of the TIM. The final approach is to reduce R_c by filling the crevices of the processor and attach block surfaces. Wetting materials allow for low R_c. TIM developers are trying to achieve good performance by optimizing or improving one or more of the three parameters. Table 1 lists various types of TIM, their properties, advantages and issues.

Table 1: Properties and applicability of TIMs

Material Type	Typical composition	Advantages	Disadvantages	+BLT (mil)	Thermal conductivity (W/m-K)
Grease	AIN1, Ag, ZnO, Silicon oil	High bulk thermal conductivity, conforms to surface irregularities, no cure needed, re-usable	Pump out and phase separation (post rel test), migration	2.0	3 to 5
Gel	Al, Ag, Silicone oil, Olefin	Gook bulk thermal conductivity, conforms to surface irregularities before cure, no pump out or migration, re-usable	Cure needed, lower thermal conductivity than grease	1-1.5	3 to 4
*PCM	Polyolefins, epoxies, polyesters, acrylics, BN, Alumina, Al, Carbon nanotubes	Conforms to surface irregularities, no cure needed, no delamination, easy handling, re-usable	lower thermal conductivity than grease, no uniform BLT	1.5-2	0.5 to 5
**PCMA	Pure In, In/AG, Sn/Ag/Cu, In/Sn/Bi	High thermal conductivity, easy handling, re-usale	Complete melt possible, voiding	2-5	30 to 50
Solder	Pure In, In/AG, Sn/Ag/Cu, In/Sn/Bi	High thermal conductivity, easy handling, no pump out	Reflow needed, stress crack/delamination (post rel), voiding possible, Not re-usable	2-5	30 to 50

*PCM: Phase Change Material
**PCMA: Phase Change Metallic Alloy
+BLT: Bold Line Thickness of TIM

In addition to improving the thermal performance of contact between two surfaces, TIM is also required to be reworkable; i.e., allow the separation of the processor and thermal solution. Processors are expensive (> $100) in comparison to the thermal solution (~ $10) or the TIM (~$1). If the thermal solution fails or the TIM is incorrectly applied, it should be possible to salvage the processor.

Table 2 lists the performance of best-in-class measurements of TIM, by material type. It also indicates

the potential performance of that material class in the near future.

Table 2: Representative θ_{TIM} performance range of best-in-class TIM (Unit in $^{o}C\text{-}cm^{2}/W$)

Material Type	1995	1998	2001	2004	Beyond
Elastomer	1.0	0.7	0.6	0.5	0.4 ?
Grease	0.3	0.2	0.2	0.1	0.05 ?
*PCM	-	0.25	0.2	0.1	0.05 ?
+PCMA	-	-	0.1	0.05	0.03 ?
Gel	-	0.3	0.2	0.08	0.05 ?

*PCM: Phase Change Material
+PCMA: Phase Change Metallic Alloy

Gels (since 2000) and phase change metal alloys (since 2001) are newer to the market. A review of Table 2 also indicates that the potential for improved TIM performance exists based on past trends. Realizing this potential improvement requires the updated formulations of TIMs, the optimization of current each ingredient material of TIMs, the selection of base polymers, and the improvement of filler properties. Since the potential performance improvement diminishes with material maturity (the law of diminishing returns), the tailored optimization of TIMs to a specific application becomes important. The first step in the development of application-specific materials is the ability to quantify the TIM performance under the required application (in this case, notebook systems).

TIM Characterization Methodology

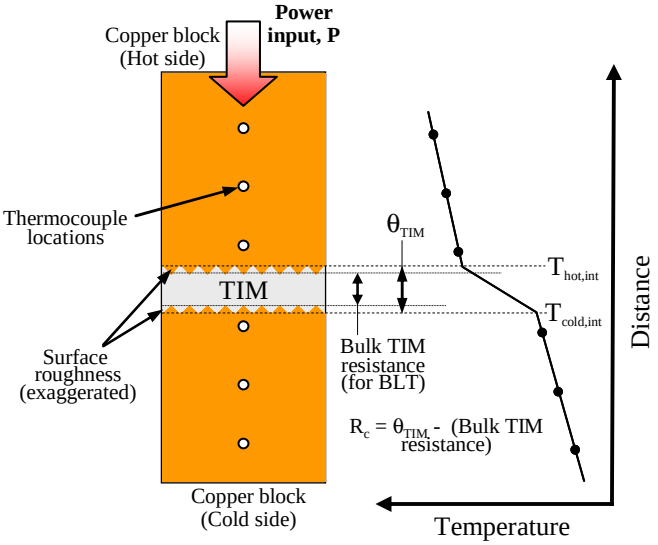
Two methods to quantify the performance of TIMs are discussed next. The first method, often used by material developers, is to quantify the material characteristics of the TIM. This method involves the use of a material tester, and is described first.

Material Tester

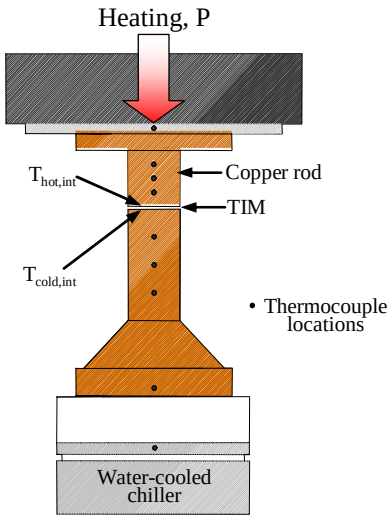
In simple terms, a material tester (ASTM D-5470 based [2]) consists of a TIM filled between two blocks with co-planar surfaces. One block is heated and the other block is cooled. The temperature at the interface of each block is measured. The heat flux through the TIM is measured and the TIM resistance (a measure of TIM performance) is calculated as follows:

$$\theta_{TIM} = \frac{T_{hot,\text{int}} - T_{cold,\text{int}}}{P} \qquad \text{Equation 2}$$

Where P is the power (heat flux) conducted through the TIM, $T_{hot,int}$ is the temperature of the TIM interface on the hot side and $T_{cold,int}$ is the temperature of the TIM interface at the cold side. Figure 5(a) shows a schematic representation using the temperature gradient approach, where θ_{TIM} can be calculated from Equation 2.



(a) Schematic showing temperature gradient method to measure θ_{TIM}



(b) Schematic of Material Tester

Figure 5: ASTM-based Material Tester

Figure 5(b) shows a schematic of an ASTM-based Material Tester. It consists of two copper rods, one of which is heated by electric resistive heaters and the other is cooled by a water-cooled thermoelectric chiller; and a facility to apply a mechanical pressure and adjust planarity at the interface between the copper rods. The TIM to be tested is placed between the copper rods. Each copper rod has three thermocouples embedded along its axis to measure the temperature gradient along the copper rod. The BLT is measured by reflecting two laser beams off the hot and cold copper bars, respectively, and measuring the distance of the reflected laser beams both with and without a TIM sample in the material tester. The

increase in the measured distance with the TIM is taken as its BLT.

Table 3 shows the TIM performance and BLT measured in an ASTM-based material tester. The test data range from 0.02 ^{o}C-cm^{2}/W to 0.2 ^{o}C-cm^{2}/W. The above method is used by many TIM manufacturers to estimate the performance of TIM and to develop new formulations of TIMs. Some advantages of the material tester include the controlled co-planarity between surface of interest, the ability to measure the TIM thickness, the ability to apply an accurate pressure on the TIM, and a controlled uniform temperature at the TIM-surface interface.

Table 3: TIM performance data from Material Tester measured at 90 psi loading on TIM

Material	θ_{TIM} (^{o}C-cm^{2}/W)	$^{+}$BLT (mil)
Grease 1	0.02	0.3
Grease 2	0.06	0.7
Grease 3	0.16	0.3
Gel 1	0.05	0.3
*PCM 1	0.08	1.1
*PCM 2	0.11	1.9
*PCM 3	0.11	0.3
*PCM 4	0.18	4
Foil 1	0.16	2
Foil 2	0.11	0.4
Solder 1	0.11	5
Solder 2	0.18	5

*PCM: Phase Change Material
$^{+}$BLT: Bond Line Thickness of TIM

A caveat with the material tester is that it does not reflect a realistic use-condition environment for the TIM. In a notebook system, such ideal conditions and controls don't exist, so a Mobile TIM Tester was developed to measure the TIM performance under more realistic application conditions. This is discussed below.

Mobile TIM Tester

A Mobile TIM Tester was developed to ensure a realistic characterization of TIM performance in a notebook environment. As shown in Figure 6, it consists of a Thermal Test Vehicle (that simulates a real processor), a wide, thin, flat copper plank with symmetric fan-heat sinks (to simulate a mobile thermal solution), and a real mobile mechanical attach to load the copper plate to the test vehicle. A Thermal Test Vehicle is made from the same technology as a real processor and is a thermal "replica" of the actual processor. It consists of heating elements within the silicon to heat the die surface (in a manner similar to that expected in a real processor). In notebook systems, a heat pipe is used to move energy from the processor to the fan-heat exchanger. However, no standard heat pipe is available on the market with calibrated performance. Hence, a flat, wide copper plank with the symmetric placement of over-sized fan-heat sinks is used to replicate a heat pipe-based thermal solution. The properties of the copper plank and the fan-heat sink unit are controlled.

A Mobile Mechanical Attach consists of a dimple plate that applies a point center load on the top of the copper plank to ensure minimal tilt and uniform pressure on the TIM and die surface.

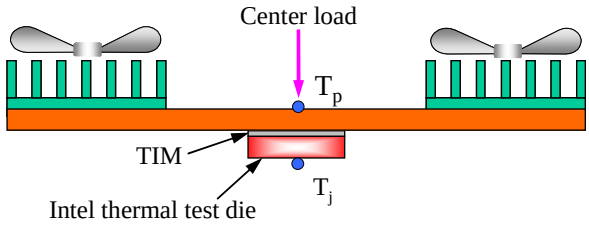


Figure 6: Schematic of Mobile TIM Tester

The temperature of the Thermal Test Vehicle, T_j, is measured based on a thermal sensor on the Thermal Test Vehicle. The temperature on the copper plank, T_p, is read by a thermocouple attached to the center of the copper plank on the top. The power input to the processor, P, is read using a combination of a voltage and current meter, or a power meter.

The performance of the TIM is captured as a resistance from the Test Vehicle to the copper plank, thus:

$$\theta_{j-p} = \frac{(T_j - T_p)}{P} \qquad \text{Equation 3}$$

The performance of the TIM in a Mobile TIM Tester measured as $\theta_{j\text{-}p}$ includes the resistances from the processor and copper plank in addition to the TIM itself.

To obtain the true TIM performance, θ_{TIM}, from a Mobile TIM Tester data, $\theta_{j\text{-}p}$, a decoder (translator) is needed that removes the contribution from the processor and copper plank in the $\theta_{j\text{-}p}$ measurement. For the development of translator ($\theta_{j\text{-}p}$ to θ_{TIM}), a thermal model (simulation) of the Mobile TIM Tester was developed with two fan-heat sink units, copper plate, TIM, and the processor with PCB. Variation in the TIM performance was correlated to the measured values of $\theta_{j\text{-}p}$ in the thermal model. This led to the translator below:

$$\theta_{TIM} = 1.026\ \theta_{j\text{-}p} - 0.095 \qquad \text{Equation 4}$$

The above correlation is useful to calculate the TIM performance from measured $\theta_{j\text{-}p}$ values from Intel's Mobile TIM Tester. This equation depends on the specific Mobile Test hardware. A TIM user is expected to develop an independent correlation to translate measurement to TIM performance. The measured TIM performance values in Intel's Mobile TIM Tester are presented in Table 4 (the

same materials were tested in an ASTM-based Material Tester and are presented in Table 3).

Table 4: TIM performance measurements in Intel's Mobile TIM Tester (Unit in ^{o}C-cm^{2}/W)

Material	θ_{j-p}	θ_{TIM}
Grease 1	0.17	0.08
Grease 2	0.19	0.10
Grease 3	0.36	0.26
Gel 1	0.19	0.10
*PCM 1	0.21	0.12
*PCM 2	0.21	0.12
*PCM 3	0.21	0.12
*PCM 4	0.25	0.16
Foil 1	0.22	0.13
Foil 2	0.23	0.14
Solder 1	0.17	0.08
Solder 2	0.23	0.14

*PCM: Phase Change Material

Comparisons of TIM Measurements

For seeking the best-in-class TIM for Pentium M processors, Intel explored various materials. Intel procured the samples of TIMs from more than 10 TIM suppliers. All test samples were tested at least four times for repeatability and an average of the TIM performance data for each TIM was calculated. This test data are presented in Tables 3 and 4 and plotted for comparisons in Figure 7.

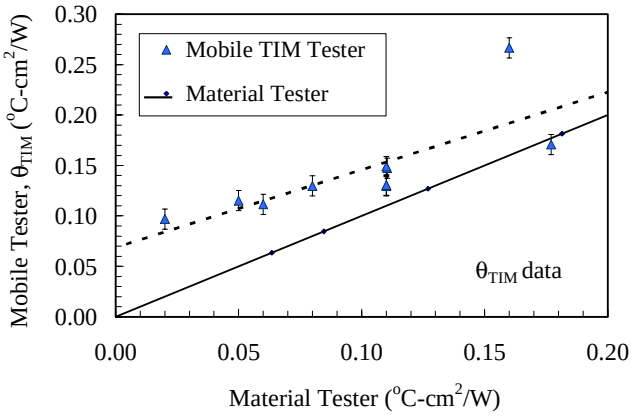


Figure 7: Test data comparing Material Tester and Mobile TIM Tester results

The test data in Figure 7 do not indicate a good match between the Material Tester and Mobile TIM Tester (a Mobile TIM Tester duplicates a notebook application environment). The TIM performance in a Material Tester shows a lower thermal resistance generally in comparison to a Mobile TIM Tester, which suggests that the Material Tester generates an optimistic view of TIM performance, a view that is not achieved in a real notebook environment.

TIM performance data collected in a Material Tester are unlikely to indicate a performance that is representative of a real application. Intel emphasizes the need to test TIMs in real notebook environments to characterize the TIM performance and to make an accurate assessment of the performance of the thermal solution. The method described in the previous section enables the accurate measurement of TIM performance in a notebook system environment without requiring the assembly or testing of an actual notebook.

TIM Degradation

The performance of TIMs can degrade with usage. Degradation in TIM performance depends on usage temperature, the time of usage, mechanical loading, and material properties. Since all factors are not well understood, TIM degradation is characterized using empirical methods. TIM degradation can vary with application and test conditions and is measured in a Mobile TIM Tester environment. The idea, again, is to replicate a true notebook environment and characterize degradation therein (via accelerated testing).

TIM Degradation Estimation

The degradation of TIM performance is measured as a change in the measured value of θ_{j-p} (since θ_{j-p} directly relates to TIM performance, as shown previously in Equation 3) after each stress cycle (e.g., high-temperature bake or temperature-humidity) compared to pre-stress state.

Since testing for long-term (one to five years) degradation requires a long period of time that is not realistic in a lab environment, the reliability characterization process uses accelerated testing. In accelerated testing, the severity of stress generated by critical parameters affecting TIM performance is increased. TIM performance data are collected as a time series under a severe stress environment, and these data are empirically modeled with the stress applied and the period of application of the stress.

For example, a TIM in a notebook system might experience temperatures in the 50-90 ^{o}C range over its life of three or more years. Under accelerated testing, the thermal stress is enhanced and the TIM is tested in the 125 to 150 ^{o}C range for shorter durations. The Mobile TIM Tester is put in a bake chamber and heated at a temperature of 125 ^{o}C or 150 ^{o}C for about 2000 hours. The TIM performance is measured at every interval as needed (e.g., 100 hours, 500 hours). In Figure 8, the measured TIM performance values of a sample TIM at three different bake temperatures are plotted as a function of bake time. An empirical model is developed that best

fits the data plotted in the same figure. An example of empirical correlation is a modified Arrhenius model shown below [3]:

$$\theta_{j-p,t} = \theta_{j-p,t=0} + B\, e^{-\left(\frac{E_a}{KT}\right)} \sqrt{t} \qquad \text{Equation 5}$$

where $\theta_{j\text{-}p,t}$ is the degraded TIM performance at time, t; $\theta_{j\text{-}p,t=0}$ is the TIM performance (before application of stress at time = 0), B is an acceleration coefficient, E_a is the activation energy for the TIM in Mobile Attach; K is Boltzmann constant, and T is the temperature (in Kelvin scale).

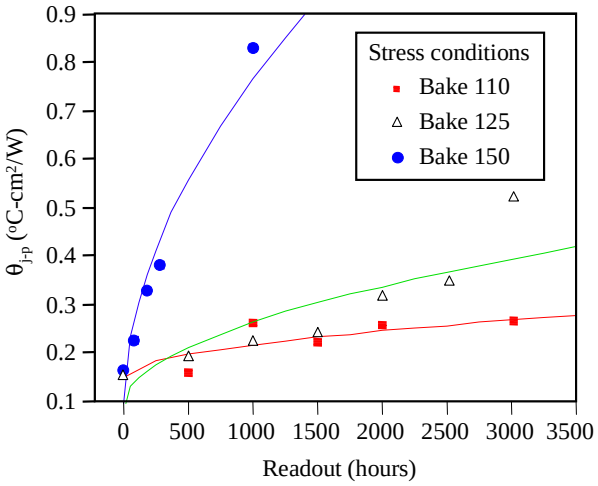


Figure 8: Plot showing change in a particular TIM performance at different stress levels

The rate of degradation of TIM depends on the TIM and its interaction with the Mobile TIM Tester (real notebook environment). Different materials have different degradation rates under similar stress conditions. Once a model is developed, the predictions of TIM performance are extrapolated to the end-of-life condition, based on the stress condition in real mobile usage conditions. The end-of-life represents the Intel-specified service life of a notebook product during which no field failure is allowed.

Different materials have different degradation rates under similar stress conditions. Figure 9 shows an example of TIM performance degradation for three materials: PCM-a, PCM-b, and Gel-a, which are selected as the best choices for Pentium M processors through extensive path-finding efforts with TIM manufacturers. The rate of degradation of Gel-a and PCM-b are very similar but substantially lower than PCM-a. The figure also indicates the impact of long-term stress on TIM performance degradation. After five years of usage, the TIM performance should depend on the long-term usage temperature in a notebook system. The long-term usage temperature depends on the notebook thermal solution design and the usage of notebook. At lower temperatures, the rate of degradation must be lower.

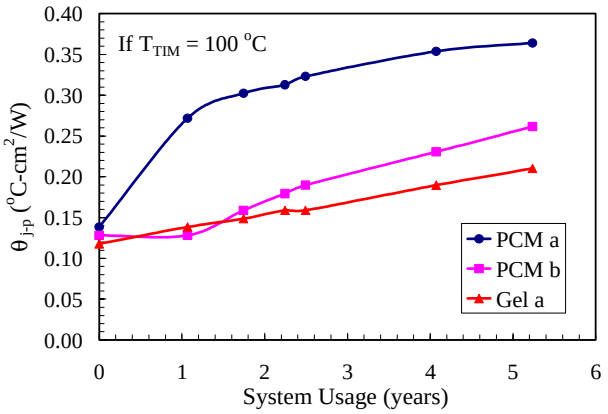


Figure 9: Projection of TIM performance degradation

The degradation in TIM performance is attributed to a prolonged usage under thermal stresses. However, how degradation occurs is not clearly understood. The analysis of TIM samples in a Mobile TIM Tester subjected to accelerated testing has been performed. Some samples have been sheared and laser cut to view the TIM-to-processor and TIM-to-copper plank interfaces. Delamination of the TIM from either the processor surface or copper plank was observed to start at the corners of the processor and progress toward its center with prolonged accelerated testing (or prolonged usage). So delamination of TIM is speculated to cause the degradation of TIM performance. Changes in the material properties with a prolonged use, evaporation/loss of volatile components, the interaction between the thermo-mechanical elements of the Mobile Attach, and oxidation of TIM are possible reasons.

System designers would like to measure the degradation rate of different TIMs and select a material that has a lower rate of degradation. It is also expected that they would estimate the long-term temperature of the processor and the TIM in their system to capture the degradation and TIM performance accurately.

FILTER-BASED THERMAL STATE ESTIMATION AND MANAGEMENT

This section of the paper is devoted to a method for managing high power consumption.

Figure 1 shows a high-level view of the main silicon components on a second-generation platform built on Intel Centrino mobile technology. All data transfers to and from the system memory are managed by the GMCH. If an application (running on the CPU) wants to get data from the hard disk or network, these data will make their way

from the appropriate port on the ICH to system memory and then from system memory to the CPU. If an application is operating on a large data set, the system memory will also get accessed to service the CPU cache misses. In integrated graphics mode, graphics and display also have high bandwidth data streams to the system memory.

In summary, if the GMCH/system memory is idle, the platform itself is generally in an idle state. If the GMCH/system memory is not idle, it is consuming more power, and the GMCH and memory components are heating up.

However, the GMCH and memory components react differently to the same activity. If the GMCH is performing a memory read, the data input buffers on the GMCH will be consuming power, but the GMCH data output buffers will not be active. From the memory's viewpoint, the GMCH memory read will activate the memory's output buffers, as well as the memory's core. So while the GMCH and memory power consumption and temperatures will be somewhat correlated, they will not be identical. Therefore independent mechanisms are required to detect GMCH and memory overheating.

The CPU has a thermal sensor, and there may be skin temperature monitoring thermal sensors, but usually the rest of the platform components do not have thermal sensors.

Thermal design guidelines are provided to customers to ensure that component over-temperature conditions do not occur for normal workloads. It may be possible, however, to design atypical (power virus) workloads that may cause overheating in some components.

A lumped model of component power is as follows:

Power = Dynamic Power + Static Power + Leakage

Equation 6

where

Dynamic Power = C × V^2 × f × AF
Static Power = I × V
Leakage Power = Leakage × stacking factor × V
C is the total capacitance toggling at rate AF × f
f is the effective operating frequency
AF is activity factor
V is voltage change or applied
I is effective constant current sources
Leakage represents all the component transistor leakage sources
Stacking factor takes into account the data-dependent component of Leakage

The expansion of the lumped model of dynamic power is as follows:

$$C \times V^2 \times f \times AF = \sum_i C_i \times f_i \times AF_i \times V_i^2 = \sum_i w_i \times AF_i$$

Equation 7

where each component of die functionality contributes its specific fraction w_i (weight) to the total dynamic power.

The relationship between component die junction temperature and steady state power is as follows:

Tdie_junction = Tambient + Θja*Power Equation 8

where Tambient is the component ambient temperature, Θja is the thermal resistivity (°C/Watt), and Power is the component steady state power.

Equation 8 applies to static conditions only. Workloads are time variant, and so die temperature is also time variant.

Lumped thermal analysis [4] (i.e., assume that thermal energy is leaving the die from all elements on its surface and the temperature of the die is uniform) leads to the time behavior response of temperature to a step response ΔT in temperature as follows:

$$T(t) = \Delta T(1 - e^{-t/\tau}) + \text{constant}$$ Equation 9

This equation corresponds to that of a first order low pass filter, and can be written recursively as

$$T(n) = (1-\alpha)T(n-1) + \alpha\,\Delta T + \text{constant}$$ Equation 10

where n is the time index, normalized to a sampling frequency of 1 and α is the filter time constant.

We can then combine Equations 8, 9, and 11 as

$$T(n) = (1-\alpha)T(n-1) + \alpha\,\Theta_{ja}\sum w_i AF_i + \text{constant}$$

Equation 11

This equation is that of a weighted input power filter.

It has been found (empirically) that Equation 12

$$T(n) = \sum \Delta T_i (1 - e^{-t/\tau_i}) + \text{constant}$$ Equation 12

gives a good fit to the time variant behavior of the die junction temperature to a step response in power (the constant term is a function of I, stacking factor, and ambient temperature). For the current generation silicon fabrication process, the leakage and constant current source of GMCH, ICH, and memory power will tend to be relatively constant at the higher power levels. So measurements of component AF_i can be converted into a reasonable indicator of component power, which can in turn be converted into an indication of die junction temperature.

The maximum possible power will correspond to states that toggle on every clock (AF_i = 1). But this rate of

toggling is unlikely to happen for the total capacitance of a component. For example, the very highest activity factors are found in clock trees.

So not only is the activity factor proportional to power, high AF_i serves as an indicator of workloads that are of relatively low utility[2]. It makes sense then to combine information about AF_i to slow down the GMCH core and/or memory activity when over-temperature conditions are suspected:

While T(n) > Threshold_Temp[3], throttle Equation 13

The next section details the implementation of GMCH and memory throttling mechanisms in the Intel 915 GMCH.

SYSTEM MEMORY AND GMCH THROTTLING USAGE MODELS

The Intel 915 GMCH has two independent mechanisms that cause system memory bandwidth throttling: GMCH thermal management, and DRAM thermal management.

GMCH Thermal Management

GMCH thermal management ensures the GMCH chipset is operating within thermal limits. The underlying theory is that GMCH heating is caused by the activity required to access system memory. The implementation provides a mechanism that controls the amount of system memory activity (i.e., Double-Data-Rate-II (DDR2) IO transactions) to a programmable threshold limit as per Equation 13. Memory activity throttling blocks all transactions or a selected set of transactions to the system memory.

System Memory

System DRAM thermal management ensures that the DRAM modules are operating within thermal limits. The DRAM modules are not equipped with thermal sensors, so their temperatures must be tracked via indirect means. The underlying theory is that a DRAM device heats up by different amounts based on the type of activity it is subject to. For example, the amount of heat contributed by a read command is different to that contributed by a write command. The implementation accounts for this variation by using the appropriate values for w_i for each memory transaction type. Throttling can be initiated by a DRAM activity measurement exceeding a programmed threshold.

[2] There are exceptions such as test patterns for silicon validation, for example.

[3] Threshold_Temp is set taking into account product reliability limits, and sufficient guardband to compensate for the estimation error in T(n).

GMCH Thermal Throttling

If the weighted power transaction filter output exceeds a programmable activity threshold (Threshold_Temp_GMCH), then the GMCH starts throttling GMCH activity. As per Equation 13, the throttling lasts for as long as the throttling threshold is exceeded.

Since GMCH thermal throttling is specific to the GMCH, there are three types of transactions (AF_i) that can be assigned different weights (w_i):

1. *Write*: This value is loaded in the filter when a write command is issued on the bus.
2. *Idle*: This value is loaded in the filter when there is no command being issued on the memory bus.
3. *Non-Write*: This value is loaded in the filter when a command other than write (read, precharge, activate, etc.) is issued on the bus.

Figure 10 shows an example of the relationship between the threshold value chosen for throttling and the GMCH power. Figure 11 shows that 3D graphics performance is unaffected for throttling levels above 38 h (programmable activity threshold). Since this behavior is typical of other workloads, a throttling level of >38 h will filter out undesirable workloads while leaving desirable applications unaffected.

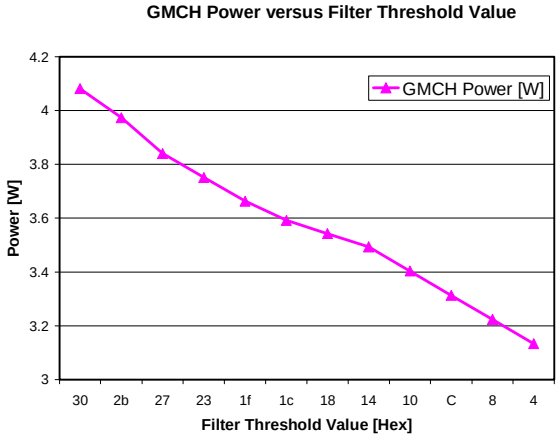


Figure 10: GMCH power as a function of throttling threshold

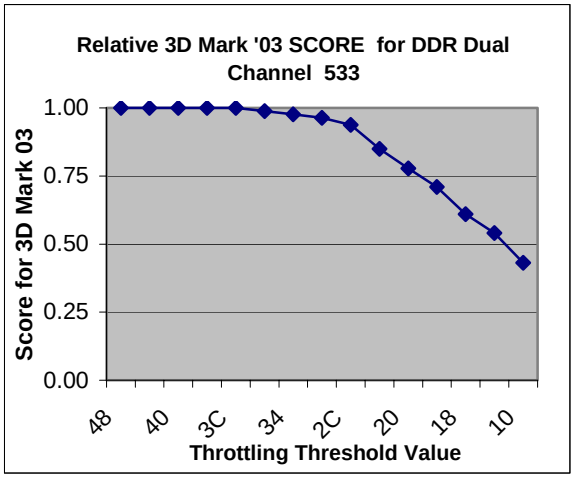


Figure 11: Relative graphics benchmark as a function of throttling threshold

DRAM Rank-Based Throttling

System DRAM devices are organized in ranks. Typically, the memory devices on a side of a memory module form a rank. Each rank heats up independently based on the activity it is subject to by the GMCH. Hence each rank requires an independent power filter.

If any of the rank weighted power transaction filter outputs exceed a programmable activity threshold (Threshold_Temp_DRAM), then the GMCH starts throttling memory. As per Equation 13, the throttling lasts for as long as the throttling threshold is exceeded.

DRAM devices have many pertinent power states:

1. Active State: Reads with Autoprecharge
2. Writes with Autoprecharge
3. Bust Reads
4. Burst Writes
5. Precharge All
6. Precharge
7. Activate
8. Page Open Idle
9. Page Close Idle (all banks precharge)
10. Page Open Power down
11. Page Close Power down (all banks precharge and power down)
12. Rank in Self refresh

The power values for these states are DRAM vendor specific. The Joint Electron Device Engineering Council (JEDEC) has a Serial Presence Device (SPD) specification[4] that allows the memory module vendors to fill in the register fields for the equivalent temperature delta (= $\Theta_{ja} \times w_i$) that can be mapped into most of these power states. These values can be read by the system bios from the memory modules during boot time and placed into the GMCH registers that store the state weights. After this operation is completed, the DRAM filters will look up and use the appropriate stored values as the particular memory states occurs.

For example, the memory module's SPD has the DT4R register field reserved for the temperature rise from ambient due to continual burst read operation. So each time a new memory read burst is started on a particular rank of memory, the DT4R value is used as an input into the filter for that rank.

Figure 12 shows memory power varying as a function of throttling level (expressed as a function of equivalent read bandwidth) for two different memory vendors.

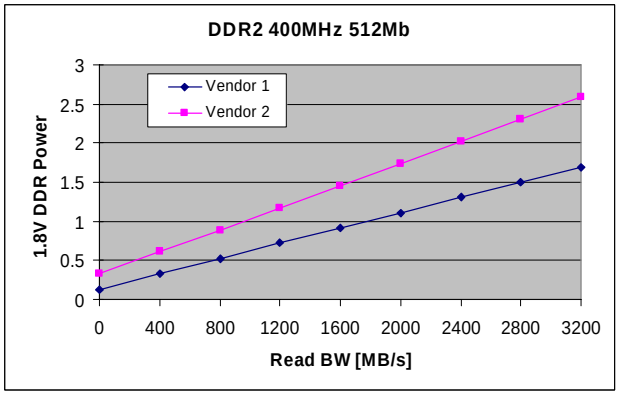


Figure 12: DRAM power

Since one vendor's power consumption per unit bandwidth is less than another's, the filter output for the lower power vendor will be smaller, and hence less likely to cause an overheating indication.

Thermal Power Filter Usage In Conjunction with Other Methods

While this method is able to monitor power variations due to GMCH activity, there are several factors that need to be encompassed in the guardband applied to the thresholds used to decide throttling.

For example, this method is not aware of the actual ambient temperature, so must assume the worst-case specification. Voltage variation must also assume worst-case specifications.

Since an approximation to the thermal diffusion equation is used, the actual thermal transient behavior modeled will be conservatively set so that it is slower for temperature increases, and faster for temperature decreases. This

[4] Joint Electron Device Engineering Council DDR2 SPD Revision 1.1

mismatch also contributes to the guardbanding requirements.

On-die thermal sensors, when available with sufficient die coverage and accuracy, are a better way to deal with these uncertainties. But since an over-temperature condition may last an indefinite time, it is desirable to allow some work to progress rather than hang the platform (due to halting all activity) until the temperature has fallen back to reasonable levels. The filter-based method is a good way to control the amount of throttling happening during over-temperature conditions since it can allow work to proceed at a reduced rate:

```
While (thermal sensor is sensing overheating) {
        While T(n) > threshold {
                throttle.
        }
}                                              Equation 14
```

While circuitry for direct power measurements can also be used to obtain instantaneous power measurements, these power measurements would still need to be filtered to correlate with the thermally relevant workload.

SUMMARY

Improving the platform component performance/power efficiency and improving the box cooling capability are the primary vectors to maximizing the platform performance given finite heat budgets. By using better material for component packages, component cooling is improved. Improvements in TIM performance are expected, and better characterization of the performance and reliability of these materials will aid this process.

Data were presented to demonstrate the improved accuracy of a Mobile TIM tester or testing in real-use conditions rather than just ASTM D5470-based Material testing.

A simple, accurate, and accelerated reliability test method based on the Arrhenius reliability model was presented.

Notebook system designers and TIM vendors are encouraged to study the degradation of TIM in real systems using the methodology outlined here.

Once a platform's cooling solution and components have been chosen, there is still some added scope to maximize performance by minimizing the guardband between the operating states allowed and the physical thermal limits.

We have shown how filter-based thermal state estimation can be used to help control overheating in platform components, even when these platform components do not contain internal thermal sensors. We have shown how these filters can be performance-transparent in many usage scenarios, and yet still be able to detect and limit power viruses. We have also shown how filters can be useful even if there is a thermal sensor available. Filter-based thermal estimation enables the second-generation platform built on Intel Centrino mobile technology to set a higher level of performance than it might otherwise allow, while still providing some protection from power viruses.

ACKNOWLEDGMENTS

We thank Lloyd Pollard and Yuchen Huang for supporting much of the work related to JEDEC standardization of memory event weights, and Terry Fletcher for the GMCH power weight event analysis. We thank Pachi Chen for providing the initial data on thermal transient characterization of GMCH heating, and Abdul Momem for providing the platform performance sensitivity to throttling.

REFERENCES

[1] R. S. Prasher, J. Shipley, S. Prstic, P. Koning and J. L. Wang, "Thermal Resistance of Particle Laden Polymeric Thermal Interface Materials," in *Proceedings of International Mechanical Engineering Congress and Exposition*, Nov. 15-21, 2003, Washington, D.C.

[2] ASTM, "Standard Test Method for Thermal Transmission Properties of Thin Thermally Conductive Solid Electrical Insulation Materials," ASTM D-5470-01, 1999.

[3] F. W. Billmeyer, Jr., Textbook of Polymer Science, John Wiley Publication, 3rd Edition, pp. 72-73, 1984.

[4] D. Pitts and L. Sissom, *Schaum's Outlines Heat Transfer*, 2nd Edition, pp. 88-89, 1998, McGraw-Hill, New York.

AUTHORS' BIOGRAPHIES

Eric C. Samson joined Intel in 1983 to design communication chipsets. He currently focuses on power management architecture for chipsets. His e-mail is eric.c.samson at intel.com.

Sridhar V. Machiroutu joined Intel in 2000 with a Ph.D. in Mechanical Engineering. He works on developing cooling technologies for mobile platforms including TIMs and heat pipes. His e-mail is sridhar.v.machiroutu at intel.com.

Je-Young Chang joined Intel in 2000 with a Ph.D degree in Mechanical Engineering. He worked in the Mobility Group developing thermal solutions for mobile processors and platforms for four years, and then in Assembly Technology Development developing advanced thermal packaging technologies. His e-mail is je-young.chang at intel.com.

Ishmael Santos works on Mechanical/Thermal Enabling in the Mobility Group. His interests include computer architecture with an emphasis on platform thermals, power consumption, and performance. He received his B.S. degree in Electrical Engineering and Computer Science from the University of California, Los Angeles. His e-mail is ishmael.f.santos at intel.com.

Jim Hermerding joined Intel in 1993 with a B.S. degree in Computer Engineering. He has worked in the Mobility Group for 11 years in various roles including architecture, engineering, product technical marketing, and thermal and mechanical applications. His current role is a manager within the Thermal Mechanical Enabling team. This involves managing a team to help define, test, and enable thermal control and protection mechanisms in Intel mobile products and to deliver component thermal testing software. He also manages a team of thermal simulation engineers to analyze cooling limits of mobile platforms. His e-mail is jim.g.hermerding at intel.com.

Ashay Dani obtained his Ph.D degree in Chemical Engineering in 1995. He worked for five years in various roles supporting plastics molding materials and process research at KeyTronic Corp (ex-Honeywell), before joining the Intel Assembly Technology Development Group in 2000. Within Intel, he has supported development of various polymer packaging materials and suppliers (thermal interface materials, thermal heat spreaders, sealants, adhesives, die attach, encapsulants, molding compounds, etc.) for several internal and external packaging platforms and products. He has published several papers in referenced journals and conferences within Intel. His e-mail is ashay.a.dani at intel.com.

Ravi Prasher joined Intel in 1999 with a Ph.D degree in Mechanical Engineering. He has worked on many aspects of TIMs and heat pipes for mobile applications and is currently involved in evaluating microchannel cooling technology for electronics cooling. He currently holds eight US patents with approximately 20 patent applications pending with the US patent office. He has published more than 40 articles in journals and conference proceedings. He is also an adjunct professor in the Department of Mech. and Aerospace Engineering at Arizona State University. His e-mail is ravi.s.prasher at intel.com.

David W. Song joined Intel in 2003 with a Ph.D degree in Mechanical Engineering from UCLA. He is a thermal engineer in the Design Process Development Group in Assembly Technology Development. He is currently working on the batch thermal property characterization of mobile test packages for product certification, as well as the current carrying capacity characterization of socket contacts. His graduate study and continuing research interests are in the microscale heat transfer in semiconductor and nano/microcomposites. His e-mail is david.w.song at intel.com.

Next-Generation PC Platform Built on Intel® Centrino™ Mobile Technology New Usage Models

Camille Chen, Mobility Group, Intel Corporation
Penny Chen, Mobility Group, Intel Corporation
Sherry Chen, Corporate Technology Group, Intel Corporation
Gordon Chinn, Mobility Group, Intel Corporation
Hani Elgebaly, Mobility Group, Intel Corporation
Uma Gadamsetty, Mobility Group, Intel Corporation
Ranjit Narjala, Corporate Technology Group, Intel Corporation
Paul K. Reddy, Mobility Group, Intel Corporation
Jeremy Rover, Corporate Technology Group, Intel Corporation
Paul Schmitz, Corporate Technology Group, Intel Corporation

Index words: EMA, VoIP, SNS, OBR, IRAP, WCS, VoIP, Bluetooth, 802.11, SIM, Roaming

ABSTRACT

Notebook computers based on the next-generation PC platform built on Intel® Centrino™ mobile technology are designed to meet and enhance the mobility requirements of a business user in the office as well as on the road. In this paper, we explore the capabilities of this platform, as well as the interfaces and the wireless ecosystem used to enable new usage models: Extended Mobile Access (EMA), Voice over Internet Protocol (VoIP), Simplified Network Selection (SNS), and One Bill Roaming (OBR). This next-generation platform is enabled to work with the EMA capability, which includes a small integrated display on the back of the lid that provides the user with up-to-the-minute calendar, daily tasks lists, and e-mail information while the user is in transit between meetings on the office campus while the notebook lid is closed. The EMA technology reduces the notebook power consumption by turning off the main LCD; hence saving battery life. The closed-lid notebook can still maintain network connectivity to the campus network via the Wireless Local Area Network (WLAN). The platform capabilities and the VoIP software will allow a user's desk phone to follow the user anywhere wireless or wired connectivity is available using VoIP. The latest platform incorporates the Wireless Coexistence System (WCS) Phase II solution that mitigates the potential interference between WLAN and Bluetooth[1]. Hence, the next-generation platform built on Intel Centrino mobile technology delivers a better audio experience for the user when used in the WLAN and Bluetooth environments. The platform includes SNS technology that enables end users to roam between WLAN and Wireless Wide Area Networks (WWAN) across multiple locations using a single set of Subscriber Identity Module (SIM) credentials and OBR across these heterogeneous wireless networks. Moreover, Intel Corporation is leading the industry to define International Roaming Access Protocols (IRAPs) [1], and in promoting adoption of IRAP. In essence, IRAP is a set of standards that are adopted by this next-generation platform.

INTRODUCTION

Notebook business users are looking for greater mobility in the office and on the road to engage in continuous business activities. The business user is looking for new powerful usage scenarios that enable continuous voice and data communications for access to e-mail, calendars, Personal Information Management (PIM), and desk phone

® Intel is a registered trademark of Intel Corporation or its subsidiaries in the Untied States and other countries.

™ Centrino is a trademark of Intel Corporation or its subsidiaries in the Untied States and other countries.

[1] Bluetooth is a trademark owned by its proprietor and used by Intel under license.

facilities whenever a network connection is available. The next-generation platform built on Intel Centrino mobile technology makes this mobile digital office a reality and provides a better user experience. The following key capabilities are either enabled or provided:

- Extended Mobile Access (EMA) technology that provides users the ability to view new e-mail messages or view their calendar to find out where their next meeting is, through an integrated auxiliary display on the notebook lid when the notebook lid is closed. This technology improves users' productivity as the user roams inside the enterprise with the notebook lid closed.
- Intel's Wireless Coexistence System (WCS) Phase II technology provides a better user experience while receiving or making voice calls utilizing Voice over Internet Protocols (VoIPs) over WLAN with Bluetooth headsets. This technology enables VoIP soft phones on notebooks to provide a virtual office experience for the business user where the user's desk phone can be routed to the notebook as long as it is connected to the enterprise network.
- Simplified Network Selection (SNS) technology enables users to utilize WWAN SIM to roam to WLANs. This technology enables selection of the most cost-effective network among available networks without the user's intervention. The technology also enables a single bill model using One Bill Roaming (OBR).

To enable this on-the-go lifestyle of an enterprise mobile user, the wireless ecosystem around the platform must be enabled to provide a seamless connectivity experience to the user. In this paper we introduce the International Roaming Access Protocol (IRAP), which outlines this proposed wireless ecosystem.

EXTENDED MOBILE ACCESS FOR INCREASED PRODUCTIVITY

The next-generation platform built on Intel Centrino mobile technology enables Extended Mobile Access (EMA) hardware and software that has been developed to meet the requirements of the "Office Warrior" usage model, which describes the behavior of a user within a corporate office (enterprise) environment where 802.11 wireless connectivity is available. To support this usage model the EMA capability introduces a user interface on a small LCD with buttons to support navigation. The following are the key features available to an end user:

- *Closed-Lid Synchronization* that supports synchronization of Personal Information Management (PIM) data from a server to a local data store during closed-lid operation. The PIM data are current and available for viewing even if network connectivity is lost during transit.
- *PIM Mode* consists of continuous access to a server via a Wireless Local Area Network (WLAN) interface. The PIM data are updated continuously while the lid is closed and made available for viewing.
- *Network Detection and Status* indicates the status of the WLAN connectivity including Service Set Identifier (SSID) and signal strength.
- *Battery-Level Indication* provides information on the current battery strength of the notebook.

Figure 1 illustrates the EMA LCD on the mobile platform and the usage model.

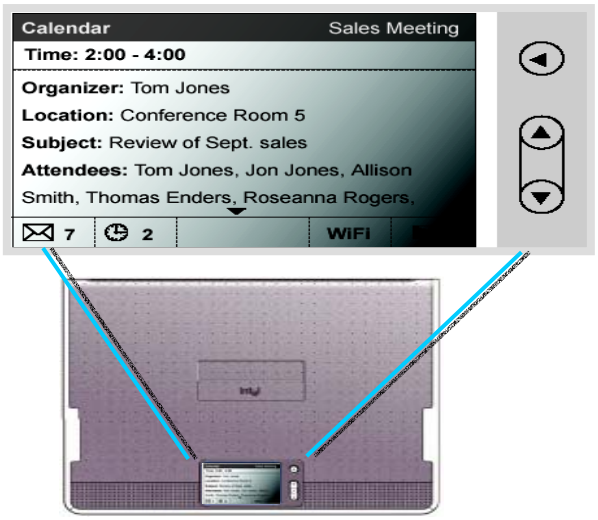


Figure 1: EMA usage model

Architecture

EMA technology consists of an operating mode that performs a number of functions such as powering down non-essential devices and synchronizing PIM information upon the closure of the notebook by the mobile user. Figure 2 provides an overview of the EMA architecture.

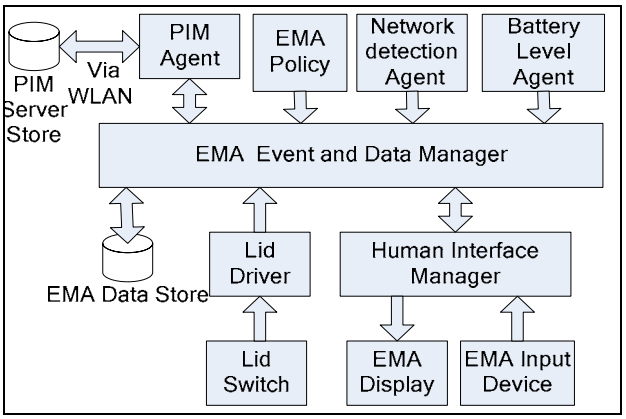


Figure 2: EMA architecture overview

This section provides an overview of the following key EMA subsystems and concepts:

- EMA Closed-Lid operating mode
- WLAN communication subsystem
- EMA display and input devices
- EMA agents: PIM, network, and battery

EMA Closed-Lid Operating Mode

The EMA policy manager is responsible for the policy settings and execution of closed-lid operations. The end user will be able to set policies to have the notebook enter EMA Closed-Lid operating mode and perform functions within this mode when the lid is closed (lid event). In this mode, subsystems that are not used are powered down to save energy. Table 1 lists the state of various subsystems under one possible policy configuration.

Table 1: Subsystems state

WLAN Card	On
Processor and Chipset (including memory)	On
HDD	On
Small EMA Display	On
Main LCD	Off
Wired Ethernet	Off
CD-ROM/DVD	Off
COM and LPT Ports	Off
Mouse and Keyboard	Off

The EMA Closed-Lid operating mode also employs processor performance control utilizing Intel SpeedStep® technology to implement a combination of processor P-state management and CPU throttling to reduce platform power and cap thermal dissipation of the notebook.

WLAN Communication Subsystem

EMA technology employs the WLAN IEEE 802.11 a/b/g communication framework for enabling closed-lid usage models. The WLAN interface provides connection to the PIM servers on the network to get continuous updates. One of challenges is to address Virtual Private Network (VPN) connectivity issues. In the enterprise, in some deployment models, WLAN is deployed outside the firewall. VPN connectivity is essential to maintain security of enterprise user data over WLAN. To initiate VPN connection, one has to enter a Personal Identification Number (PIN) into the notebook. In closed-lid mode this could be inconvenient. Intel Corporation worked with VPN vendors and enabled hands-free VPN connectivity for EMA closed-lid operation. The VPN connection persists as the user moves across Access Points (APs).

EMA Display and Input Devices

The fundamental feature of EMA is instant access to PIM data and status. A simple glance at the EMA display yields some useful information about unread e-mail messages, current calendar alerts, network connection, and battery status. The technology also will have buttons to allow the user to navigate through various menus and screens to view critical PIM data and status information about the WLAN and battery.

EMA Agents

When the lid is closed, the PIM agent clears any stale data in the EMA data store and establishes a connection with the PIM server using the WLAN interface; it then registers for new PIM updates and caches the data for viewing.

The network agent detects the WLAN environment including SSID, AP-friendly name, and signal strength and updates these details periodically in the EMA data store for viewing.

The battery agent continuously updates the battery-level indication in the EMA data store for viewing.

EMA Usage Model Extensions

The EMA technology can be extended to include new usage models in the future. One possible extension is to receive or make voice calls while users move about the office campus when the notebook lid is closed. The user

® SpeedStep is a registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.

can interact with the lid display and navigation buttons to accept or reject voice calls. The user also initiates calls using PIM data on the integrated display. The user experience can further be enhanced by using a Bluetooth headset in conjunction with closed-lid notebook for Voice over Internet Protocol (VoIP) applications.

Another possible extension to EMA technology could be a closed-lid usage model that utilizes speech recognition technology for speaker commands and control of notebook applications. While the closed-lid notebook user is mobile and wearing a Bluetooth headset, the user can access his/her voice mail, e-mail, and task and calendar items. The user can also get speech-based calendar and e-mail alerts through the Bluetooth-based headset while the user is on the go. Adding MP3 player features to EMA technology is desirable as users can listen to music while the lid is closed using Bluetooth headsets.

EMA Key Benefits

EMA features allow the notebook to run in a lower power profile without a full power drain on the battery and enable users to be more productive as they roam the office campus freely, while keeping up-to-the-minute with e-mails, tasks, and appointments. EMA enables the notebook to continue to provide access to critical information even while it is closed. The increased convenience of EMA will result in a new set of user needs and requirements. The next-generation platform built on Intel Centrino mobile technology establishes a foundation for these usage models, and the user experience will keep getting better with each new-generation platform.

ENABLING VOIP APPLICATIONS FOR BETTER USER EXPERIENCE

The "voice on the go" usage scenario is about supporting VoIP soft-phones on Intel Centrino notebooks for enterprise users. The soft-phones enable enterprise users to make calls to other soft-phones or Public Switched Telephone Network (PSTN) phones using VoIP technology as long as the user is connected to the enterprise network. This usage model allows users to replicate a virtual office environment on their notebooks, and as long as they are connected to the enterprise they will have access to their desk phone and all their telephone services including call transfer, speed dialing, conferencing, hold, access to voice-mail, etc. The value proposition of this usage model includes convenience for enterprise workers and also brings tremendous cost savings for IT.

This usage scenario allows the user to make and receive phone calls in any location with IP network connectivity as long as they are connected to the enterprise network. Users in most cases will be using WLAN to connect to their IP network and will use Bluetooth wireless headsets. VoIP over WLAN is a key capability that enables this "voice on the go" scenario. Due to the simultaneous use of WLAN and Bluetooth, the next-generation platform built on Intel Centrino mobile technology offers a Bluetooth/WLAN coexistence solution to support this usage.

When Intel launched PC platforms enabled with Intel Centrino mobile technology in 2003, the Intel Wireless Coexistence System (WCS) Phase I was implemented between Intel's PRO/Wireless 2100/2100A Network Connection solution and a third-party Bluetooth module [2] that mitigated Bluetooth interference and restored WLAN 802.11b data throughput almost completely. The following sections detail current problems, audio quality issues, and how WCS Phase II further mitigates RF interference.

Wireless Coexistence Problem Statement

Figure 3 shows a typical usage scenario for a VoIP application and illustrates how RF interference causes audio quality degradation. In this example, a "speaker" is talking to the "listener" using the left mobile PC via a Bluetooth headset. At the same time, the "speaker" is also sending some files or sharing some document with the "listener." In this scenario, the left mobile PC is receiving Bluetooth audio packets, and at the same time, it is transmitting files via WLAN. Since WLAN (802.11b/g) and Bluetooth use the same frequency (ISM band 2.4 GHz) at the same time, RF interference occurs in the left mobile PC. As a result, the listener at the right mobile PC will hear the corrupted sound.

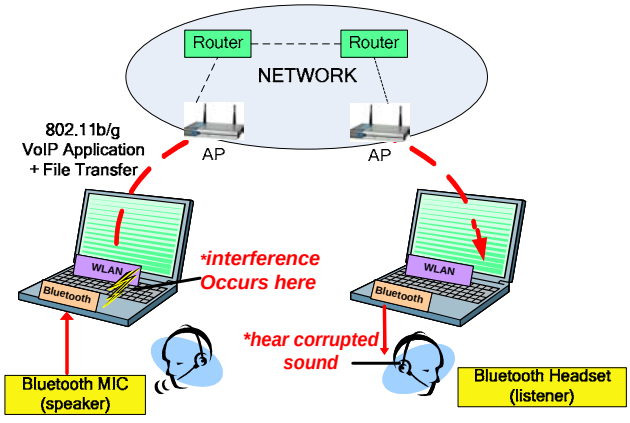


Figure 3: VoIP usage model

To further illustrate the problem, Figure 4 shows WLAN 802.11b/g and Bluetooth audio High quality Voice (HV) 3 Synchronous Connection-Oriented (SCO) packet activity over the time and frequency. The blue boxes represent Bluetooth HV3 packets, hopping between 2.40-2.48 GHz.

The red boxes represent 802.11b/g packets; in this case, it is operating in Channel 5. Since WLAN and Bluetooth operation are not coordinated, a collision happens when Bluetooth packets hop into the WLAN channel while WLAN is transmitting/receiving packets. Furthermore, since HV3 packets have no Cyclic Redundancy Checksum (CRC) in Bluetooth Version 1.1 [3], the corrupted packets will not be retransmitted.

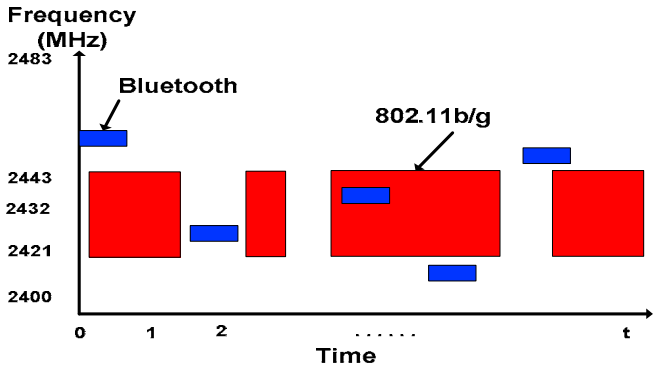


Figure 4: Uncoordinated WLAN and Bluetooth activities

Intel WCS Phase II Solution

In 2004, Intel developed PRO/Wireless 2200BG and 2915ABG Network Connection modules supporting 802.11b/g and 802.11a/b/g WLAN standards, respectively. Phase II of Intel WCS was also developed and implemented between these modules and third-party Bluetooth modules for the next-generation platform built on Intel Centrino mobile technology. The third-party Bluetooth modules include Cambridge Silicon Radio (CSR) BC02 and BC4, and RFMD SiW3000.

Figure 5 shows the Intel WCS Phase II interface diagram. There are two pins between the WLAN and Bluetooth modules: the CH_Data and the CH_CLK, which are multiplexed with Bluetooth priority data named BT_Priority. The general idea is that the WLAN module will send channel information to the Bluetooth module whenever it is active. Similarly, if the Bluetooth module is transmitting and/or receiving high-priority packets for audio, Human Interface Device (HID: mouse or keyboard), or link establishment, it will assert BT_Priority if the transmit/receive frequency is within the WLAN channel. The WLAN module then receives this signal and decides whether or not to hold off its transmission. This coordination is implemented in the Bluetooth and WLAN hardware and firmware; it is independent of the Bluetooth software stack, which is a key advantage of Intel WCS.

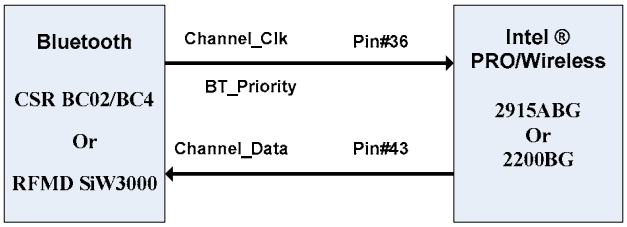


Figure 5: Intel WCS Phase II block diagram

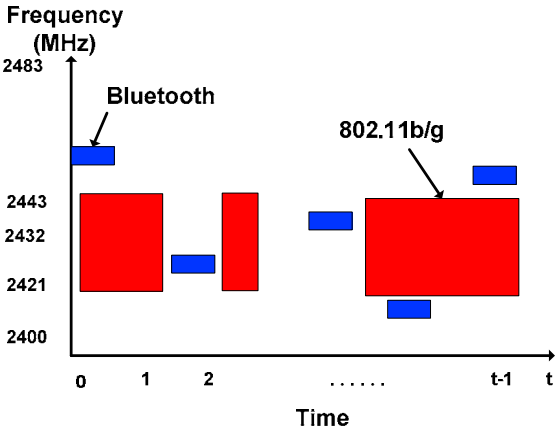


Figure 6: Coordinated WLAN and Bluetooth activities

With Intel WCS Phase II implemented, Figure 6 shows the modified WLAN 802.11b/g and Bluetooth audio HV3 SCO packet activity over the time and frequency. As shown in the figure, when a Bluetooth high-priority packet hops into the WLAN channel, the WLAN module will decide whether to hold off the packet transmission or not. By doing this, it reduces collision significantly, and Bluetooth audio quality is recovered.

Intel WCS Phase II Benefits

Now that the problem and solution have been described, let's look at what kind of benchmark is used to measure audio quality and to quantify Phase II's benefits. After investigating several audio benchmark tools available on the market, Opticom's OPERA* system was selected for several reasons. First, it adopts the latest international standard Perceptual Evaluation of Speech Quality (PESQ) for VoIP applications. Second, this tool is easy to use, and third, it is widely adopted by the cellphone industry. The OPERA system will generate Mean Opinion Scores (MOS), ranging from 0 to 4.5. A score of 4.0 or above signifies that the audio quality is good, 3.0-4.0 signifies that it is fair, and finally, 2 and below signifies that the audio quality is poor.

* Other brands and names are the property of their respective owners.

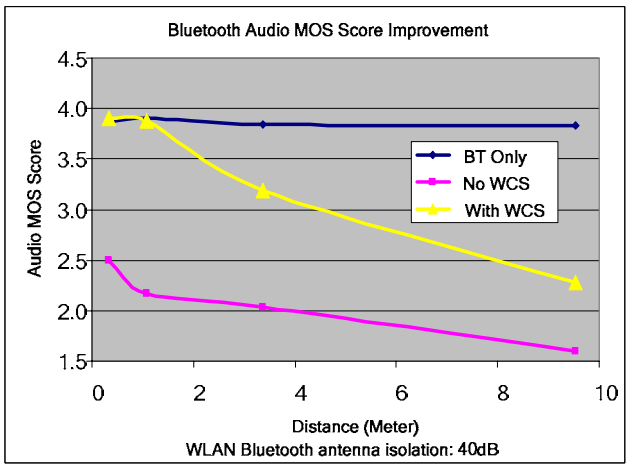


Figure 7: Bluetooth audio MOS

Figure 7 shows the audio-quality performance for a simulated VoIP scenario where the test system is cabled up. Indicated distances are calculated using path loss models. The key reason in utilizing a cabled-up system is to generate repeatable results. "Over the air" testing is subject to sometimes unpredictable RF interference, which could result in erratic test results. In Figure 7, the horizontal scale shows the calculated distance between the Bluetooth headset and the notebook with Bluetooth radio.

The Bluetooth audio streams are recorded using Windows* audio recorder and when compared with reference audio clips, MOSs are generated. For example, if there is only Bluetooth audio traffic in the system (i.e., no WLAN interference), audio quality is reasonable with the MOSs close to 4.0, up to about 10 meters. Then, with simultaneous WLAN traffic and Intel WCS disabled, the magenta line shows that even when the Bluetooth signal is strong, the MOS drops dramatically, starting from 2.5 and going down to 1.5. When Intel WCS Phase II is turned on, however, the yellow line shows that the Bluetooth audio quality improves significantly, especially at shorter distances where VoIP is anticipated.

Intel WCS Phase II's goal is to minimize the impact to WLAN performance while improving the quality of Bluetooth audio. The impact to WLAN performance due to Intel WCS Phase II is shown in Figure 8. The horizontal scale shows the calculated distance between the AP and the test notebook. The distance is calculated from the receiver signal strength and a path loss model for the office environment. The vertical line shows the WLAN throughput. Different color lines denote these three cases:

1. WLAN only–blue
2. WLAN with Bluetooth audio, but with Intel WCS Phase II disabled–magenta.
3. WLAN with Bluetooth audio, but with Intel WCS enabled–yellow line.

It is clear that the delta between the magenta and yellow lines is relatively small. In other words, the impact of the Intel WCS Phase II to WLAN performance is relatively minimal.

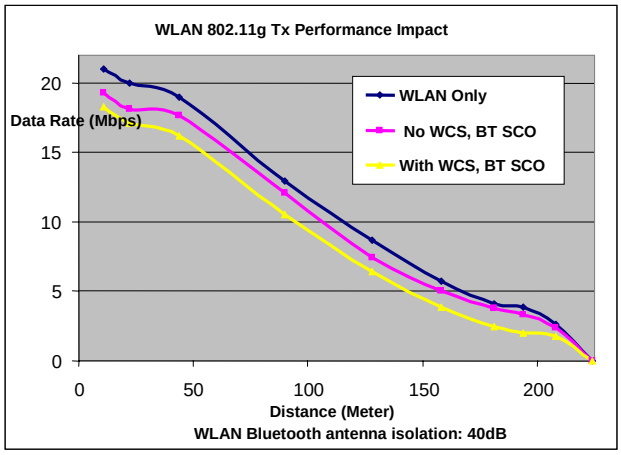


Figure 8: 802.11g throughput impact from WCS Phase II

Intel WCS Phase II Summary

Intel WCS Phase II helps to mitigate interference between WLAN and Bluetooth critical packets, such as packets for audio, HIDs, and link establishment. As a result, it helps to improve VoIP audio quality significantly when a user uses a Bluetooth headset and WLAN as the network connection. Furthermore, it has minimal impact on WLAN performance. Intel WCS Phase II works with legacy Bluetooth Version 1.1 devices, and it is also forward-compatible with Bluetooth Version 1.2 devices. In conclusion, the next-generation platform built on Intel Centrino mobile technology with Intel WCS Phase II provides a better user experience for VoIP applications using Bluetooth and WLAN technologies.

SIMPLIFIED NETWORK SELECTION

Simplified Network Selection (SNS) is an Intel initiative based on the International Roaming Access Protocols (IRAPs), Global Systems for Mobile Communications (GSM) Association, Internet Engineering Task Force (IETF) and 3rd Generation Partnership Project (3GPP) standards and forums under which Intel product teams developed SNS technologies such as the Extensible Authentication Protocol with Subscriber Identity Module (EAP-SIM) authentication method and the SIM Reuse client software, which is in compliance with the GSM

* Other brands and names are the property of their respective owners.

SIM AT command and Smart Card interface specifications to provide end users with the ability to roam between different wireless network types (WLAN and WWAN) across multiple locations using a single set of SIM credentials. In addition to a common authentication model, this technology also enables a One Bill Roaming (OBR) mechanism across heterogeneous wireless networks. The next-generation version of the Intel Centrino mobile technology platform contains the Intel PROSet/Wireless Software v.9.0.0, dual-mode Wi-Fi 2200BG, or tri-mode Wi-Fi 2915ABG hardware modules. EAP-SIM/SIM Reuse features are part of the Intel PROSet/Wireless software v.9.0.0 that use SIM AT command or PC/SC interfaces to enable access to the user's credentials from a SIM contained on a WWAN card, a Smart Card, or a USB SIM reader.

Problem Statement

With many wireless technologies available today, wireless data users experience interrupted wireless service when roaming between wireless networks. The wireless users have to negotiate the different access mechanisms, manage multiple accounts, and receive multiple bills for wireless services. Improving wireless users experience by providing simpler, safer methods of access via a single set of credentials that can be used over different wireless network types has become an essential fact for wireless technology adaptation, and SNS addresses these issues to provide a compelling solution.

SNS Solution

The aim of the SNS solution is to give notebook users automatic, constant, and economical access to the Internet and corporate Intranet, whenever their notebooks are turned on. Constant and automatic access means that users do not have to go through a manual process of connecting to the Internet/Intranet and entering authentication credentials each time they need to communicate with a remote server. Economical access means that constant Internet connectivity will not impose a significant cost premium and will furthermore allow users to specify network selection based on cost. EAP-SIM/SIM reuse and OBR are two of the key features available in the next-generation version of Intel Centrino mobile technology notebooks.

EAP-SIM is an implementation of an authentication method for Wi-Fi roaming and network access, which is based on GSM technologies. It provides mutual authentication of the client device to the network, and the network to the client device, to ensure that only valid user devices gain access to the mobile telephony network. It features the use of a SIM card, a type of Smart Card, containing user information that can be used in accounting/billing procedures, as well as data that are used in the encryption of transmitted voice and data. SIM cards, though most commonly used in mobile phones, are emerging for use with laptops, notebooks, PDA handhelds, and other devices to integrate the WLAN and GSM-capable intelligent networks.

OBR is an added benefit when user notebooks have built-in WLAN modules and add-on/built-in WWAN modules with SIM reuse technology. With the same set of SIM credentials, a carrier is able to generate one single bill for users regardless of the network types.

OBR Architecture

Figure 9 illustrates the overall WLAN and WWAN inter-working Authentication Authorization and Accounting (AAA) protocol architecture for WLAN and WWAN roaming. The interface between the WLAN hotspot and the GSM home network is called the inter-working roaming AAA interface. This is based on RADIUS as shown in Figure 9.

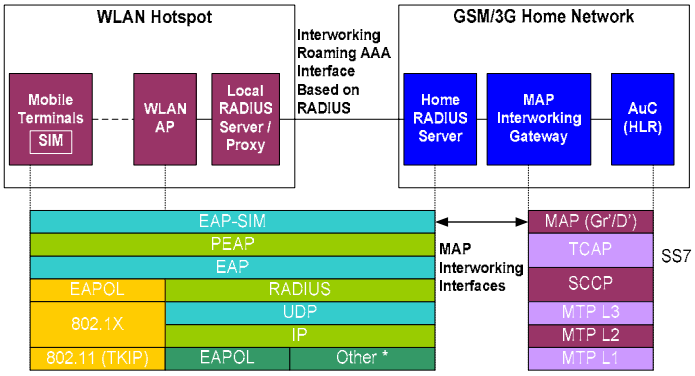


Figure 9: SIM based WLAN authentication protocol stack overview

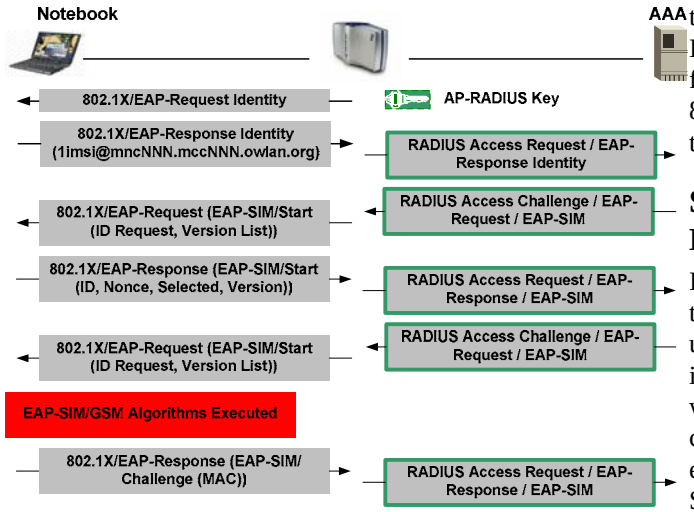


Figure 10: EAP-SIM authentication message flows

Figure 10 captures the high-level EAP-SIM authentication message flows between the mobile notebook, AP, and RADIUS server. The IEEE 802.1X protocol runs between the mobile notebook and the AP for the purposes of authentication. The Port Access Entity (PAE) implemented on the AP is responsible for blocking all user traffic except EAP packets until authentication completes. The AP also has a RADIUS client function that is responsible for initialing the RADIUS protocol that finally terminates on the home network RADIUS server, where the subscriber information resides. The RADIUS client is implemented based on the "IEEE 802.1X RADIUS usage guidelines" and is responsible for interpreting and appropriately forwarding the EAP packets between the AP and the home network.

Once the notebook associates with the AP, the EAP Over LAN (EAPOL) Start message is optionally sent by the notebook to trigger the 802.1X authentication process based on EAP. The EAP Identity Request is then sent to the notebook which responds with the EAP Identity Response that carries the NAI (username@realm). The "username" can be an International Mobile Subscriber Identity (IMSI) number or a temporary pseudonym, and the "realm" will be a fully qualified domain name that identifies the home network. The home network RADIUS server recognizing the "username" will attempt to start the EAP-SIM exchanges using the RADIUS Access Challenges/Requests. Following the successful authentication using EAP-SIM, the RADIUS Access Accept message sent from the server will result in an EAP Success message to the Client by the AP. This RADIUS message also carries EAP-SIM derived keying material for the session, which needs to be provided to the AP. The notebook also derives keying material as part of EAP-SIM authentication. Ciphering keys are set up using this keying material on the notebook and the AP as part of completing the 802.1X key-setting procedures. Please refer to the IEEE 802.1X, 802.11i and IETF EAP-SIM specifications for details on the key setting procedures [9,10]. The 802.1X PAE now opens the WLAN for user data traffic that is ciphered using the keys set up earlier.

SNS Real-Life Network Trial and Key Benefits

In 2004, Intel, selected carriers, and OEMs worked together and conducted successful end-user pilots. Pilot users have tested OBR usages for about four to six months in real-life inter-operator networks. One of the key things we learned from these pilots is that a user-friendly common connection manager is needed. Therefore, Intel enabled third-party Connection Manager Independent Software Vendors (ISVs) on next-generation platforms to meet users' requirements.

The key benefits of the Intel SNS technologies are outlined below:

- They have the ability to automatically detect the currently available networks and optimize the connection for the best available network.
- They can authenticate/authorize and generate a single-bill model when roaming between WWAN and WLAN hotspots.

Intel SNS technology is also designed to address the mobile market segments and to provide enterprise-class features, security, manageability, seamless roaming across the enterprise networks and easy-to-use, zero-click enterprise connectivity at the office, home, or hotpots using EAP-SIM methods. Intel's vision is to make notebook wireless data connection as easy as cell phone voice connection. The next-generation platform built on Intel Centrino mobile technology is the first step in realizing this vision.

PUBLIC WLAN ECOSYSTEM

Although the next-generation platform built on Intel Centrino mobile technology itself adds capabilities to enable new mobile usage models, it is essential that the platform operates seamlessly with the public WLAN ecosystem, which has grown immensely to support the mobile enterprise user on the road. IRAP addresses how the public WLAN (PWLAN) system is being enabled to work seamlessly in conjunction with this new platform.

Problem Statement in the Current Ecosystem

The analyst firm Gartner predicts by the year 2008 there will be more than 167 thousand hotspots in the world and over 75 million hotspot users. While the outlook for

PWLANs is promising, there are a number of obstacles that the industry must address for PWLANs to reach their full potential.

Current WLAN standards do not adequately address all of the system-wide issues that affect end-to-end roaming. Even vendors that are committed to a standards-based approach have developed implementations with incompatibilities with other vendors due to the ongoing evolution of standards and different choices of optional features. The rapid deployment of PWLAN currently results in a large installed base of fragmented and incompatible implementations. Fragmentation could increase the cost and complexity of supporting worldwide WLAN roaming to such an extent that it becomes difficult, expensive, and ultimately impractical to deploy. End users cannot count on their particular service experience being available at any particular hotspot and they face the possible dilemma of needing to re-configure their network devices to make use of these hotspots, while also negotiating varied login methodologies and authentication processes to be permitted to obtain access, and manage multiple bills for service. In addition, most of the current PWLAN security methods are insufficient for protecting data sent within a wireless hotspot. The challenges lie in the complexity of issues in the legacy network infrastructure, WLAN coverage, operators' business models, and billing and settlement requirements in a roaming context.

Today, service providers, equipment manufacturers, aggregators, and client manufacturers look at pieces of the solution focusing only on their individual product or offering without considering the entire PWLAN ecosystem. The solution lies in the adoption of a standards-based end-to-end architecture that addresses the technical challenges and facilitates ecosystem growth while improving the end-user experience (e.g., the next-generation platform built on Intel Centrino mobile technology). The solution to this problem is to follow the example of the cellular phone industry and enable PWLAN roaming. To solve these issues, Intel is working with wireless service providers, operators, and network equipment vendors from around the world [4, 5] to drive the definition and adoption of IRAP.

International Roaming Access Protocols Solution

IRAP [6] is an open framework to unify a global architecture. It is a set of core protocols that, when employed, make it easier for providers to build, test, and employ effective networks. The premise and goals of IRAP are to support secure worldwide seamless roaming by encouraging early alignment with and between existing and emerging standards. Users will be best served when there is a consistent service experience across all networks, logon is consistent worldwide, and there is one bill to pay and one method to access their network services and applications.

The IRAP framework defines a set of standards-based interfaces that establish a common baseline of features to facilitate seamless network interoperability between clients and different WLAN service providers. This framework is based on several architectural principles, co-existence between legacy browser-based authentication methods, and more secure authentication methods based on Wi-Fi Protected Access (WPA), end-to-end authentication and authorization, and a common accounting framework.

These interfaces are depicted in Figure 11:

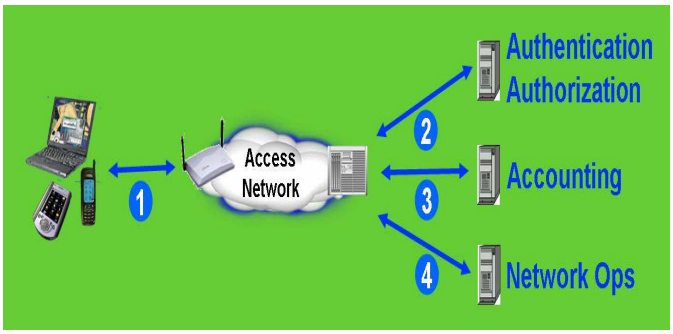


Figure 11: IRAP interfaces

Interface 1: Wireless Station to Access Network. This supports legacy browser clients and also enables the more secure and convenient next-generation authentication methods based on WPA/802.1X.

Interface 2: Access Network to Home Service Provider Authentication System. This supports AAA connections for authentication and service authorization between the visited hotspot network operator and the roaming user's service provider.

Interface 3: Access Network to Home Service Provider Billing System. This provides the accounting data to enable scalable and consistent OBR (a.k.a. unified billing) for multiple home operator types and billing models.

Interface 4: Access Network to Home Service Provider Operations Subsystem. This enables troubleshooting problems as they arise, via remote diagnostics. If customers are unable to connect in a visited network and contact their service provider, Interface 4 helps the service provider diagnose and resolve the problem.

IRAP Key Benefits

Adoption of these standardized interface profiles provides the following benefits:

- **Usability**. These profiles provide a common login process and allow the client to use a single set of credentials when roaming across PWLAN hotspots operated by different service providers.
- **Security**. The IRAP profiles improve security by encouraging adoption of (and providing a framework for) robust, over-air security and encryption methods. These methods include migration to WPA, which supports mutual authentication to protect both the user and the network, and higher security than the traditional browser-hijack methodology.
- **Interoperability**. The IRAP profiles prescribe a minimum set of accounting information that must be exchanged between different service providers. This allows the user to obtain a single bill from the home service provider regardless of the ownership of the visited networks to which they are connected.

These profiles facilitate easier integration between roaming service provider networks, and they enable end-to-end authentication between the client and its home service provider, regardless of the network being used.

IRAP Testing and Validation

As well as defining the technical requirements, Intel is driving the testing and validation of IRAP. By establishing interfaces within an ecosystem, IRAP facilitates points of verification where compliance can be confirmed. IRAP Conformance Tests [7] offer standalone tests for interface endpoints whereas IRAP Interoperability Tests validate the interoperability of two devices sharing an IRAP interface (e.g., interaction between a client and an access network).

In addition to conformance and interoperability tests, end-to-end testing [8] becomes increasingly important. With separate vendor networks interoperating with one another, the interfaces within vendors' networks must now interact with other roaming partner networks. Field trials [5], pilots, and deployment troubleshooting can be aided by running the IRAP end-to-end tests. The concept of end-to-end tests relates mainly to round-trip authentication and authorization (Interfaces 1 and 2) and billing support services (Interfaces 1 and 3). By creating and validating the plug-and-play vision of IRAP-based networks, deployment becomes more affordable, faster, and simpler.

SUMMARY

In this paper, we described new usage models for business users that are enabled by the next-generation platform built on Intel Centrino mobile technology. We have also demonstrated how this platform with the new capabilities increases users' productivity and efficiency while the user is in the office campus or away from the office. The key benefits are summarized as follows:

- EMA enables the notebook to continue to provide access to critical information even while it is closed.
- WCS Phase II Solution provides a better user experience for VoIP applications using Bluetooth and WLAN technologies.
- SNS is all about making it easier to connect to different wireless networks and providing single billing.
- IRAP adds a unified global infrastructure to the ecosystem to resolve seamless roaming obstacles and paves the way for new compelling usage models.

Some of the features that are discussed in this paper are pilot technology features and may not be supported on all the next-generation platforms built on Intel Centrino mobile technology. The platform with EMA, WCS Phase II solution, and SNS, in conjunction with the use of IRAP networks' secure authentication, authorization, roaming and seamless connectivity, allow the typical business user to be more productive by staying connected and synchronized.

ACKNOWLEDGMENTS

We acknowledge Duncan Glenndinning, Jeff Huckins, Glenn Olson, Asha Keddy, and Francis Truntzer for their help in writing this paper.

REFERENCES

[1] Paul Schmitz, Tom Van Voy, et al., "IRAP Public WLAN Roaming Market Requirements, version 0.8," *IRAP Fall Summit*, September 2004.

[2] G. Chinn, et al., "Mobile PC Platforms Enabled with Intel® Centrino™ Mobile Technology," *Intel Technology Journal*, vol. 7, Issue 2, May 2003.

[3] C. Chen, et al., "Bluetooth Wireless Technology: Delivering the Best Experience in Next Generation Notebooks," *Intel® Developer Forum*, September 2004.

[4] Ranjit Narjala, Infocomm Development Authority of Singapore, et al., "Public WLAN Interworking Study (IWS)," version 1.0, *IRAP Fall Summit*, September 2004.

[5] Sherry Chen, IDA, et al., "Public WLAN Interworking Study–Validation Report" version 1.0, *IRAP Fall Summit*, September 2004.

[6] Ranjit Narjala, et al., "IRAP Public WLAN Roaming Interfaces," version 0.85, IRAP Fall Summit, September 2004.

[7] Jeremy Rover, et al., "IRAP Interface Conformance Test Specification [s2], version 0.5c."

[8] Amber Sistla, Sherry Chen, Infocomm Development Authority of Singapore, et al., "Public WLAN Interworking Study: Test Case Specifications," version 0.8, *IRAP Fall Summit*, September 2004.

[9] P. Congdon, et al., "RFC3580 IEEE 802.1X Remote Authentication Dial In User Service (RADIUS) Usage Guidelines," IETF Networking Group, September 2003.

[10] H. Haverienen, et al., "Extensible Authentication Protocol Method for GSM Subscriber Identity Modules (EAP-SIM)," IETF Networking Group Internet Draft, available at http://www.potaroo.net/ietf/idref/draft-haverinen-pppext-eap-sim/.*

AUTHORS' BIOGRAPHIES

Camille Chen received B.S. and M.S. degrees from Zhejiang University, PRC and M.S. and Ph.D. degrees from Osaka University, Japan, all in Computer Science. She joined Intel in 1999. Currently she is a staff system architect in Intel's Mobility Group. Her professional interests include multi-radio coexistence, networking, and mobile wireless architectures. Her e-mail is camille.c.chen at intel.com.

Penny Chen received a B.S. degree in Electrical Engineering from Central South University, PRC, an M.S. degree in Electrical Engineering from Northern Illinois University, and a Ph.D. degree in Electrical Engineering from Northwestern University. She is a senior architecture design engineer in Intel's Mobility Group and focuses on wireless technology and integration solution software architecture design. Her e-mail is penny.p.chen at intel.com.

Sherry Chen earned an M.S. degree in Biomedical Engineering from National Yangming University, Taiwan, and an M.S. degree in Computer Science from the University of Denver. She joined Intel in 2003, is currently a senior network software engineer in the Intel Communication Technology Lab (CTL) within the Corporate Technology Group (CTG). Her main focus is software architecture and innovative test automation on network integration and end-to-end PWLAN validation. Her e-mail is sherry.s.chen at intel.com.

Gordon Chinn is the Radio Frequency (RF) Systems architecture manager in Intel's Mobility Group and is responsible for enabling mobile platform PCs with multiple, concurrent, embedded radio technologies. He has over 30 years of experience in various RF communications development positions in the commercial and aerospace industry. Gordon holds a B.S. and M.S. degree in Electrical Engineering and Computer Science from the University of California at Berkeley. His e-mail is gordon.chinn at intel.com.

Hani Elgebaly is the digital office architecture manager in Intel's Mobility Group. Hani's team is responsible for wireless and IP networking software technologies as they apply to Digital Office initiatives. He has been with Intel for more than nine years and has held multiple engineering management and lead architect positions. Hani has over 14 pending US patents and 15 journal and conference publications. He received a B.S. degree in Electrical and Computer Engineering with high honors from Cairo University, an M.S. degree in Computer Science from the University of Saskatchewan, and a Ph.D. degree in Computer Science from the University of Victoria, Canada. His e-mail is hani.elegebaly at intel.com.

Uma Gadamsetty is a staff software engineer in Intel's Mobility Group. He has been with Intel for more than eight years and his current focus is on various mobile PC wireless communication projects. Uma holds an M.S. degree in Electrical Engineering. His e-mail is uma.m.gadamsetty at intel.com.

Ranjit Narjala is a network software engineer in Intel's Corporate Technology Group at Portland, Oregon. His professional interests include wireless technologies, networking, and mobility protocols. Ranjit received his Masters degree in Information Networking from Carnegie Mellon University in 2000. He also holds a Bachelors degree in Computer Science and Engineering. His e-mail is ranjit.s.narjala at intel.com.

Paul K. Reddy is a wireless architecture manager in the Mobility Group, Intel Corporation. Paul has 15 years of experience in wireless and telecommunication development and deployment. He is currently working on the mobile wireless architectures, Mobile Digital Office, Digital On-the-Go Lifestyle usage models, prototype development, and user trials of next-generation Mobile Internet PCs. Paul holds two masters degrees, one in Electrical Engineering, which he received in 1985, and another in Computer Science, which he received in 1988. His e-mail is paul.k.reddy at intel.com.

Jeremy Rover is a network software engineer in the Communications Technology Lab in the Corporate Technology Group at Intel Corporation. His current interests include wireless networking, dynamic network reconfiguration, industry test specifications and WiMAX.

He received his B.A. degree in Computer Science from Linfield College. His e-mail is jeremy.rover at intel.com.

Paul Schmitz is a technical marketing engineer in Intel Corporation's Corporate Technology Group, where he is focusing on development and marketing of mobility technologies and network architectures. In his 17+ years at Intel, his areas of focus have included wired and wireless networking, firewall technologies, mobile IP and roaming technologies, and end-to-end security. He was a member of the team that verified Intel certified hotspots for the Centrino launch worldwide in early 2003. He received his Associate of Applied Science degree from ITT Technical Institute. His e-mail is paul.schmitz at intel.com.

For further information visit:

developer.intel.com/technology/itj/index.htm