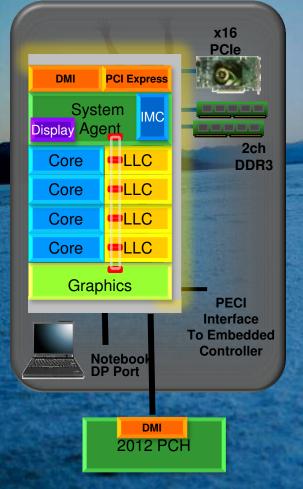
Power Management of the Third Generation Intel Core Micro Architecture formerly codenamed Ivy Bridge



Sanjeev Jahagirdar Varghese George, Inder Sodhi, Ryan Wells

Contents

> Ivy Bridge Overview > Power Scaling & Efficiency > Idle power Management Configurable TDP > Clocking > Additional Information



Contents

Ivy Bridge Overview > Power Scaling & Efficiency > Idle power Management > Configurable TDP > Clocking > Additional Information



Intel's Tick-Tock Philosophy

Tock Processors

- Provide substantial microarchitecture improvement...
- ...on existing manufacturing process

Tick Processors

- Retain existing microarchitecture, ...
- ...but utilize next generation fabrication technology to drive high volume and low product cost

□ The Tock: *Sandy Bridge*

- Brought new ring/LLC microarchitecture
- Integrated Graphics on ring
- Integrated North Bridge ("System Agent"), including memory controller

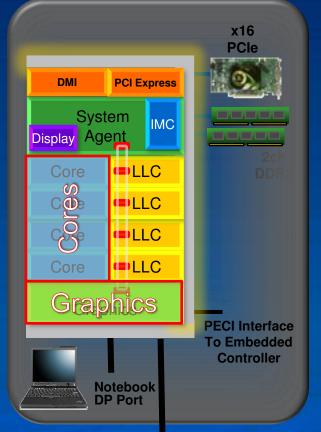
□ The Tick: Ivy Bridge

- Process lead vehicle: Intel's 22nm process node
- The Caveat:
 - Some Ivy Bridge areas have substantial (tock-like) change (Graphics)



Ivy Bridge – the 1st 22 nm Core Product

Ivy Bridge Microarchitecture



<u>омі</u> 2012 РСН

eap ahead

Leveraged from Sandy Bridge:

- Continue the 2-chip platform partition (CPU + PCH)
- Fully integrated on silicon:
 - 2-4 IA Cores
 - Processor Graphics, Media, Display Engine
 - Integrated Memory Controller
 - PCIe Controllers
 - Modular On-Die Ring Interconnect
 - Shared LLC between IA Cores and Graphics
- Same socket, similar packages
 - Similar SKUs (TDP, die configurations)
- IVB backwards compatible with SNB

Ivy Bridge – Key New Things

Entire chip moves to 22nm

- Higher performance/Lower power
- Instruction Set Architecture Enhancements
 - Float16 / Fast FS/GS support / REP MOVSB / RDRAND
- Security Enhancements
 - DRNG / SMEP
- Power Improvements
 - Scalability features: ConfigTDP
 - Average Power features: DDR power gates / PAIR
- □ IO/Memory
 - DDR3L support
 - Improved overclocking support

Performance Improvements (Instructions/clock)



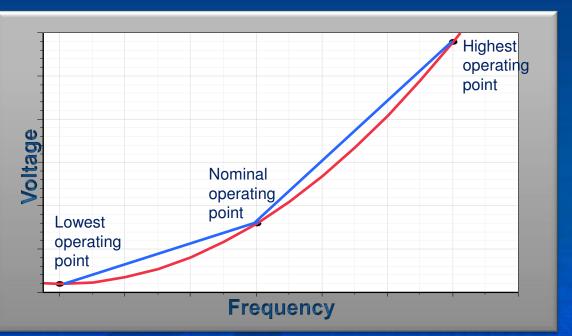
Contents

> Ivy Bridge Overview > Power Scaling & Efficiency > Idle power Management > Configurable TDP > Clocking > Additional Information



Power efficiency via scaling & testing

- Power Scaling in 22nm process extracted in two ways
 - Higher performance in IA & Graphics within a power envelope
 - Lower operating Voltage in System Agent and Memory controller



- Power loss from discrete test points and interpolation (blue line)
- Ivy Bridge builds a quadratic model of the VF based on enhanced testing (red line)
- Optimal voltage at all operating points



Power efficiency via interrupt routing

PAIR algorithm lowers power or performance impact of re-routable interrupts

- Compares power-state of all cores eligible to service interrupt
- Chooses "best core" based on optimization mode (Power vs. Performance)
- "Best Core" based on the following
 - Core C-states
 - P-state request (turbo vs. non-turbo)

Example: 1 core in C6 & 1 in C0

- Power bias will direct the interrupt to core in C0
- Performance bias will wake the C6 core



Temperature effects

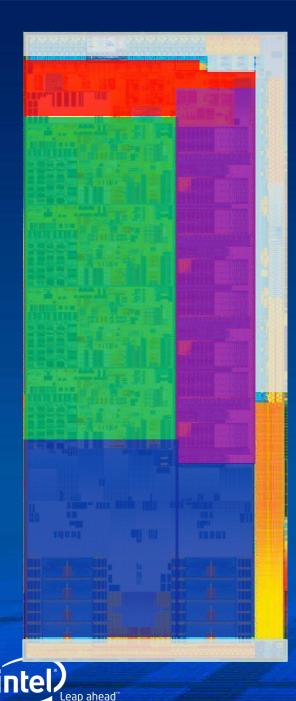
- Thermal sensors are located in the hot spots in the IA core and GPU core
- Inverse temperature dependence (ITD) effects more pronounced in the 22nm node
 - No sensors at the cold spots
 - IVB estimates the coldest point on the die to based on thermal sensors compensate for the effect
- Manufacturing test voltages at hot and cold temperatures
 - PCU interpolates linearly at run time to determine the voltage
 - Temperature moves slowly enough for the PCU and voltage regulator to keep up



Contents

> Ivy Bridge Overview > Power Scaling & Efficiency Idle power Management > Configurable TDP > Clocking > Additional Information

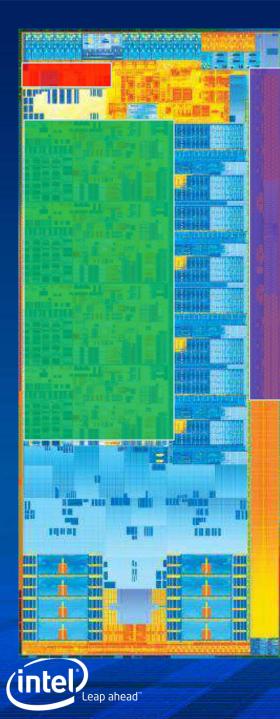




Ivy Bridge Power Planes

Key Power planes

- Core (Gated Green)
- LLC (Ungated Purple)
- SA/Display Red
- GT Blue
- Others (like IO, PLL etc) Gray



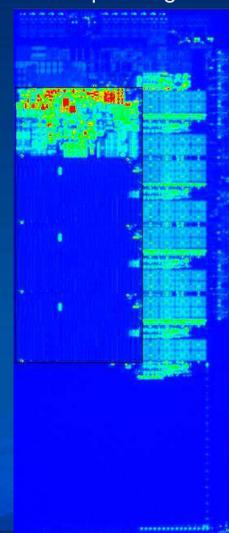
IVB Embedded Power Gate

Ivy Bridge has 3 on-die power gating areas

- Cores (Green)
 - Independent Gating per Core
 - Unified Cache
- PCIE controller(Red)
 - Gating static only when no connection
- DDR (Purple)
 - Gating of digital logic in the buffer applied during self-refresh mode

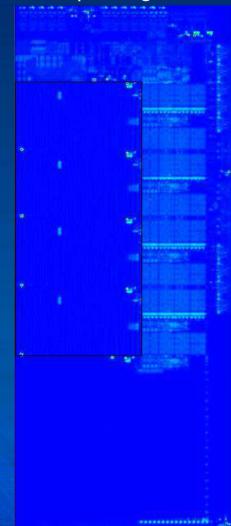
IREM images

1 core in turbo, other 3 cores power gated



_eap ahead"

Typical Usage of Cores and Graphics Cores and Graphics gated



DDR I/O Power Gating

Ivy Bridge implements on-die Embedded Power Gating (EPG) on DDR I/O Latency & Tradeoffs

Latency considerations

- Enabled on entry into Package C3 and deeper (memory in Self Refresh) to deal with latency of power gate
- Additional latency of <5uS for device access to memory during exit
- Conditional enabling only if devices can tolerate the latency
- No Impact to exit latency for interrupts
- Design tradeoffs
 - To get around saving and restoring context, the DDR state is put on an ungated power island

 For Idle/MM07-OP, Intel expects DDR IO to be gated ~90% of the time



Low Voltage optimizations

- Small Signal arrays and register files limit the lowest operating voltage and retention voltage
- 1. Dynamic cache sizing to achieve a lower cache Vmin
 - Cache Vmin is limited by 'bad cells' or defects distributed across the cache
 - A smaller size cache has a lower Vmin due to fewer defects
- 2. PCU Firmware based register file reinitialization on exit from standby states
 - Allows reduction of retention voltage below the retention level of the register file

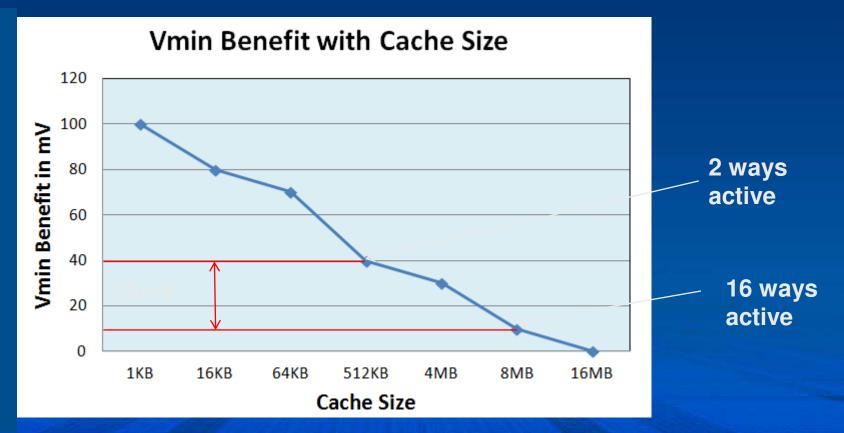


Max(Vmin1,2,3)



LLC - Dynamic Cache Shrink Feature

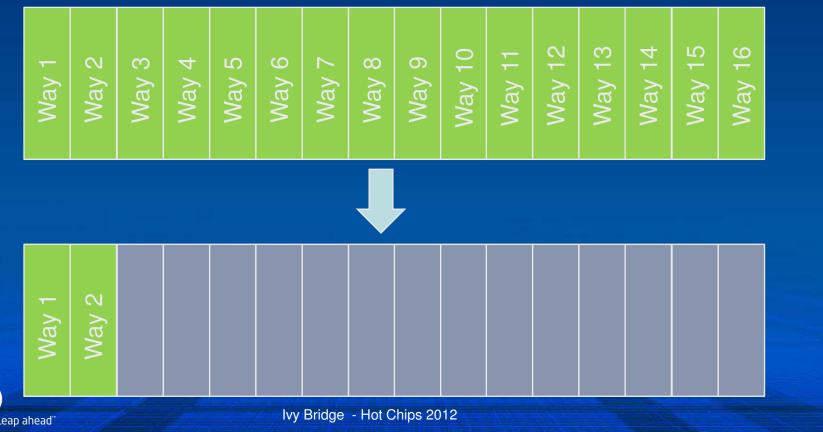
- Reduce LLC cache size dynamically from 8MB to 512KB to gain 30mV Vmin benefit
- LLC Expand/Shrink algorithm is developed for this purpose Entry/exit points were defined based on the work loads &
- perfórmance





LLC - Dynamic Cache Shrink Feature

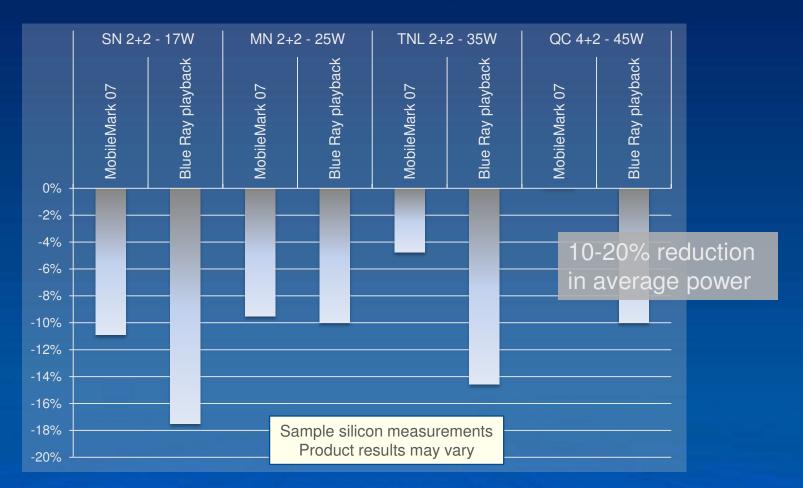
- LLC organized in 16 ways.
 When PCU detects low activity workload
 Flushes 14 ways of the cache and puts ways to sleep
 Shrinks active ways from 16 to 2 to improve VccMin
 When PCU detects high activity
- - Expands active ways back to 16 to improve cache hit rate.



Active

Sleep

Ivy Bridge average power reduction (relative to SNB)



Power reduction via new PM features and process scaling benefits Benefits on other SKUs varies

Ivy Bridge - Hot Chips 2012

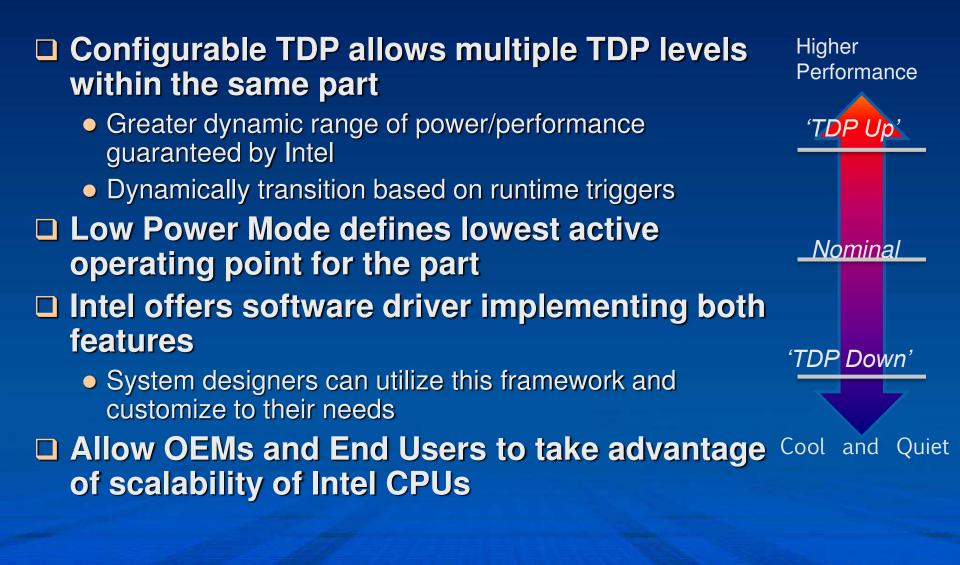
eap ahead

Contents

> Ivy Bridge Overview > Power Scaling & Efficiency > Idle power Management Configurable TDP > Clocking > Additional Information

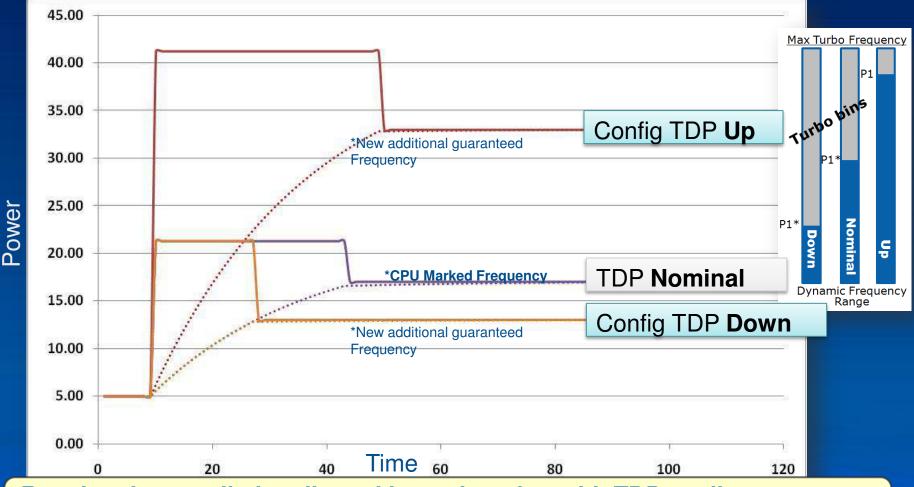


Configurable TDP & Low Power Mode





cTDP Power Control



Regulated power limit adjusted in conjunction with TDP to allow guaranteed frequency (performance) at a specific power level

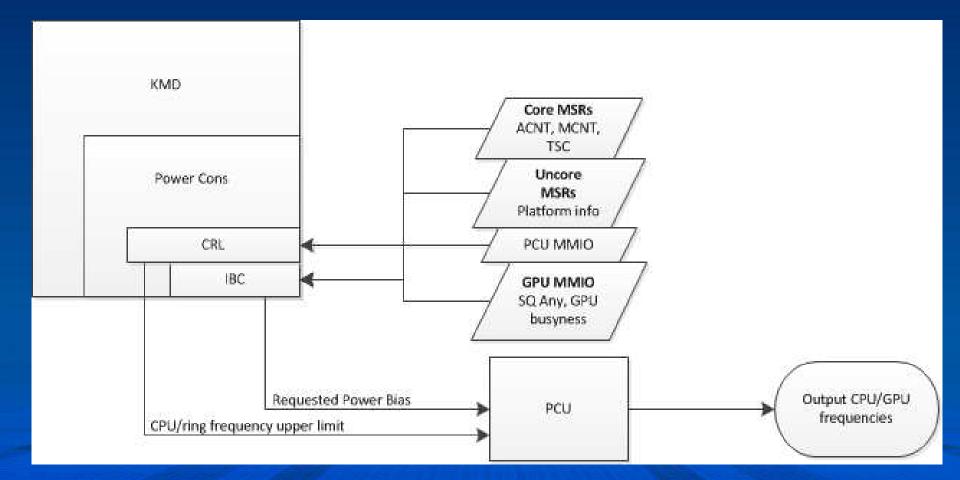
IA/GPU Power sharing

OEMs can configure the cooling limits to <17W
 Static biasing (X% to GPU and 100-X% to IA) results in sub optimal performance
 Solution: Distribute power based on workload demand

- Determine target CPU/ring frequency based on workload
 - If actual CPU/ring freq < (target frequency guard band)
 Move bias toward CPU Else Move bias toward GPU With hysteresis



Intelligent Bias Control Architecture





Platform Power management

Power delivery management

- How do we deal with the platform need to divert current from the CPU to other components dynamically?
 - IVB PCU will manage the current draw and will honor dynamic max current updates
- Platform debug and tuning hooks
 - IVB provides feedback to platform designers if power delivery, & cooling is limiting performance

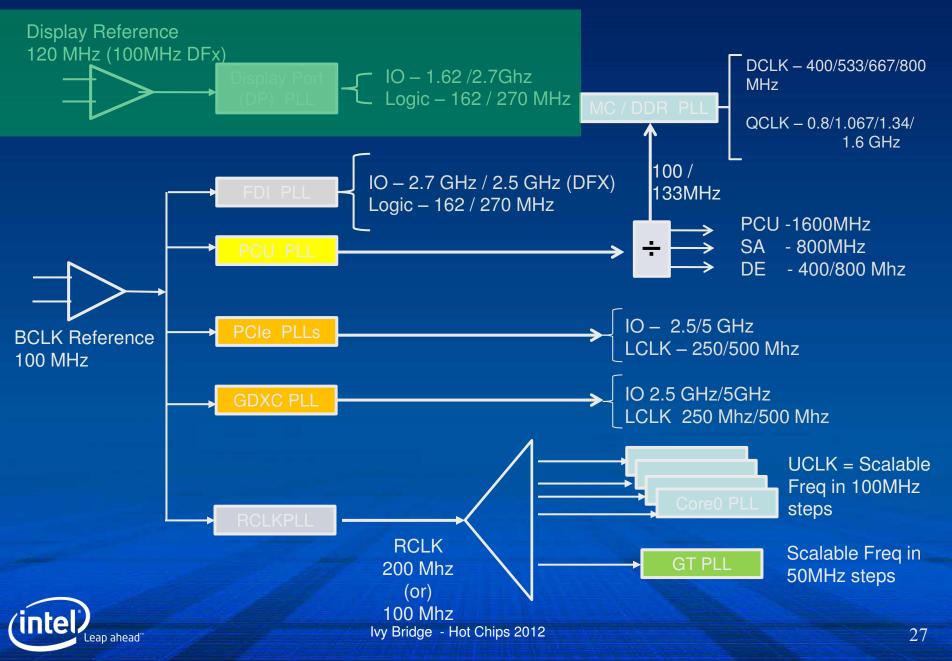


Contents

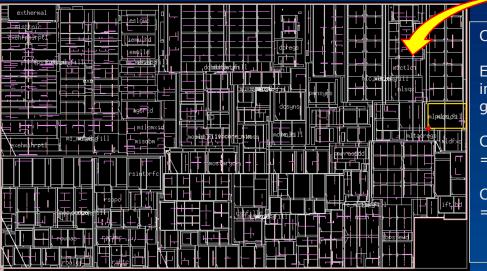
> Ivy Bridge Overview > Power Scaling & Efficiency > Idle power Management Configurable TDP > Clocking > Additional Information



IVB Clock Domains



PLL/Clocking



Clock	Isla	ands	in	Core
-------	------	------	----	------

Each Island can be independently clock gated.

Clock Islands in Core =180

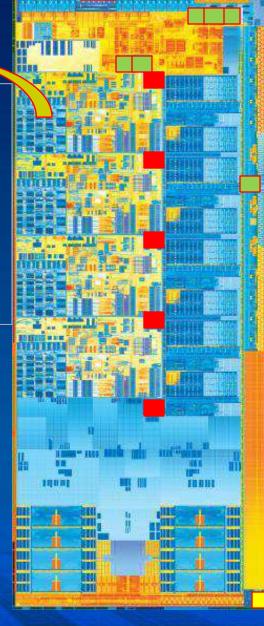
Clock Islands in LLC = 48

Wide Range SB PLL

Single Ratio SB PLL

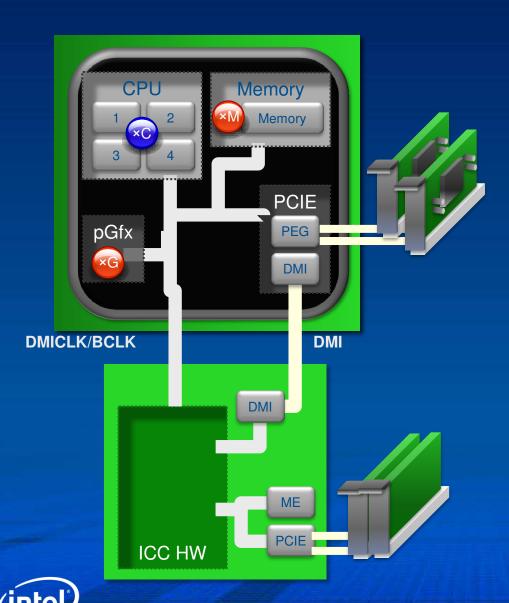
PCIE LC PLL

1				
	Slice Clocking Cdyn	CORE (pF)	L3(pF)	TOTAL(pF)
	Total (including RT)	179	77	256
	Global Drivers + Islands	109	46	155
	Clock Source + Spines	70	31	101
	Silicon Measurement	181	81	262
1				





Overclocking Enhancements



eap ahead

Core Frequency



- Unlocked turbo limits
- Unlocked core ratios up to 63 in 100MHz increments[†]
- Programmable voltage offset
- Graphics Frequency



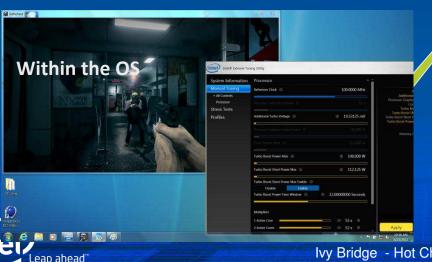
- Unlocked graphics turbo limits
- Unlocked graphics ratios up to 60 in 50MHz increments
- Programmable voltage offset
- Memory Ratio
 - 0
- Unlocked memory controller
- Granularity options for 200 and 266MHz
- Logical support up to 2666MHz
- DMICLK (aka BCLK)
- Unlocked PCH clock controller (1MHz increments)
- PEG and DMI
- Fixed ratios

Real-Time Overclocking

intel

Profiles

- PCU samples OC parameters continuously and updates power limits
- OC without reboot:
 - Maximum Core Ratio
 - Processor Graphics Ratio
 - BCLK (small increments)
 - Power Limits: PL1, PL2, Tau
 - Additional Turbo Voltage for CPU and pGfx



Intel® Extreme Tuning Utility System Information Processor Manual Tuning Reference Clock ① 100.0000 MHz All Controls Processor Stress Tests Additional Turbo Voltage 🕧 19.53125 mV Turbo Boost Power Max 0 100.000 W Turbo Boost Short Power Max 0 8 112.125 W Turbo Boost Short Power Max Enable ① Disable Fnable Turbo Boost Power Time Window 🕕 32.00000000 Seconds Multipliers 53 x 🛞 1 Active Core 2 Active Cores 52 x ⊗ 3 Active Cores 51 x ⊗ 50 x ⊗ **4** Active Cores

Changes effective immediately

Acknowledgements

Authors would like to thank the entire lvy Bridge team for their dedicated work.



Contents

> My Bridge Overview > Power Scaling & Efficiency > Idle power Management > Configurable TDP > Clocking >Additional Information



Ivy Bridge ISA & Security enhancements



Float16 Data Conversion Instructions

New instructions for supporting conversion between a 16-bit floating point memory format and 32-bit single precision

- VCVTPH2PS, VCVTPS2PH
- Both 128 (SSE) and 256 bit (AVX) wide vector flavors supported
- Only supported in the VEX prefix context
- Facilitates use of single-precision floating point computations from a more compressed memory format
 - 1-bit sign, 5-bit exponent, 10-bit significand (+ implicit integer bit)
- Enables higher dynamic range compared to fixed point within the same storage footprint
 - Image processing, video decode, audio processing
 - 50% reduction in storage v. single-precision FP (w/ loss of fidelity)
- Enumerated via new CPUID feature flag
 - CPUID.1.ECX[29]

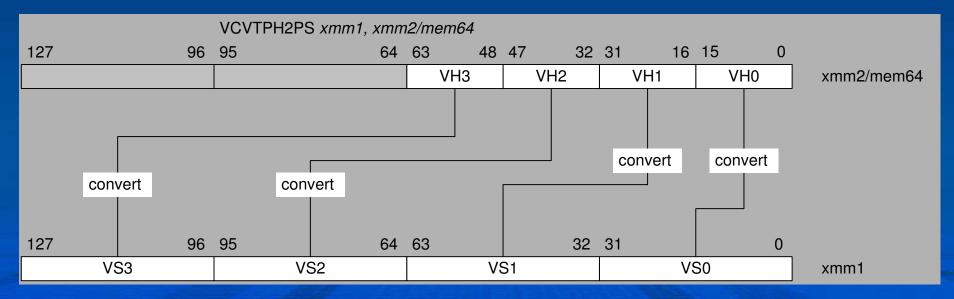


VCVTPH2PS – Convert 16-bit float to SP

VCVTPH2PS *ymm1*, *xmm2/mem128* VCVTPH2PS *xmm1*, *xmm2/mem64*

- 256 bit vector - 128 bit vector

Converts four packed 16-bit floating-point values in the low 64 bits of XMM2 or 64-bit memory location to four single-precision floating-point values and writes the results in the destination (XMM1 register).



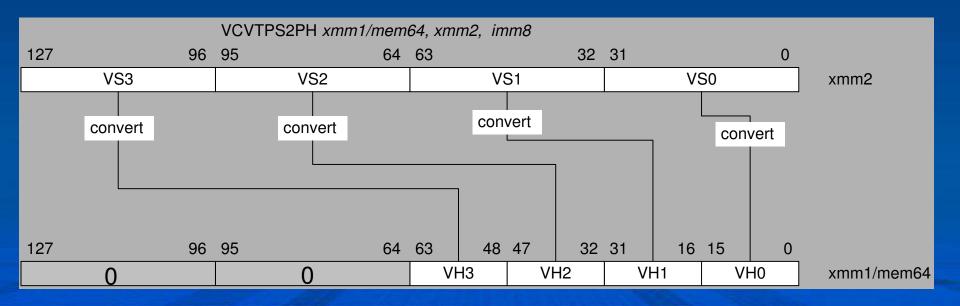


VCVTPS2PH – Convert SP to 16-bit float

VCVTPS2PH xmm1/mem64, xmm2, imm8 VCVTPS2PH xmm1/mem128, ymm2, imm8 - 256 bit vector

- 128 bit vector

Converts four packed single-precision floating-point values in XMM2 to four 16bit floating-point values and writes the results in the destination (XMM1 register or memory location).





Write/Read FS/GS Base Instructions

- New ring-3 instructions for read/write of the FS & GS segment base registers
 - To be used by user level code for thread local storage
 - Enumerated via new CPUID feature flag
 - CPUID.7.0.EBX[0] indicates availability (leaf 7, subleaf 0)
 - Requires enabling by OS to permit FS/GS segment base access
 - CR4.RDWRGSFS (bit 16) = 0 (default)
- Motivation:
 - Improve scalability and programming ease for user threads

REP MOVSB/STOSB improvements

- Historically optimizing block copy/fill operations tends to be microarchitecture specific
 - Lack of a "one size fits all" solution implies CPU model specific algorithms for best performance

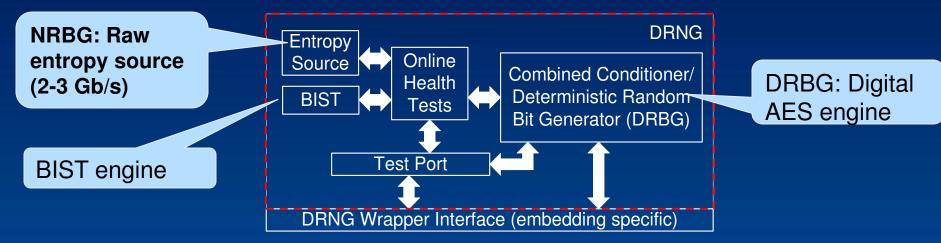
IVB address this through more optimized REP MOVSB and REP STOSB instructions

- Expect this to replace the need for manual tuning solutions
- Limitation: If block size is known at compile time and size <=64 bytes, then scalar loads & stores are still considered faster

Enhancement availability indicated by CPUID.7.0.EBX[9] (ENFSTRG)

This bit can be used by run time SW (Libraries, JIT) for tuning to a specific implementation
 Ivy Bridge - Hot Chips 2012

Digital Random Number Generator (DRNG)



Background:

- Entropy is valuable in a variety of uses Example: "keying material" in cryptography
- Historically, computing platforms did not have a good source of a high quality/high performance "entropy source"
- Typical sources used today are slow (bit rate in Kb/s) (key strokes, mouse clicks etc)
- IVB introduces high quality/high performance DRNG
- The DRNG is designed to be Standards compliant
 - ANSI X9.82,NIST SP 800-90 and NIST FIPS 140-2/3 Level 2 certifiable entropy source
- New instruction: RDRAND Available at all privilege levels/operating modes
 - Instruction will return a random number (16, 32 or 64-bit) to the destination register
- New CPUID feature flag for RDRAND enumeration

CPUID.1.ECX[30]

eap ahead

Supervisory Mode Execute Protection (SMEP)

Background:

0xF..FFF

System call

0x0..000

eap ahead

Kernel

Memory

NULL

pointer vulnerability

App

memory

Attack code Jump-or-call

- Privilege Escalation Attack causes CPL 0 access to user mode pages
- Example:
 - Step 1: Compromise user mode app or trick user into installing attack app
 - Step 2: Exploit OS vulnerability to force control transfer to user mode attack code while CPU remains in supervisory mode => privilege escalation

IVB introduces SMEP to help prevent such attacks

- Prevents execution of user mode pages while in supervisor mode
- If CR4.SMEP set to 1 and in supervisor mode (CPL<3), instructions may not be executed from a linear address for which the user mode flag is 1

• Available in both 32- and 64-bit operating modes

• SMEP is enumerated via CPUID.7.0.EBX[7] Ivy Bridge - Hot Chips 2012

PCI Express Gen 3



Ivy Bridge PCI Express Gen 3

□ Third generation of the PCI Express I/O interface

- Delivers nearly twice the I/O bandwidth v. Gen 2
- Improves performance for applications sensitive to I/O bandwidth
 - Enables smaller form factors via narrower, faster physical links

Bandwidth realized through:

eap ahead

- Faster signaling speed: 8 GT/s
- More efficient lane encoding: 128/130

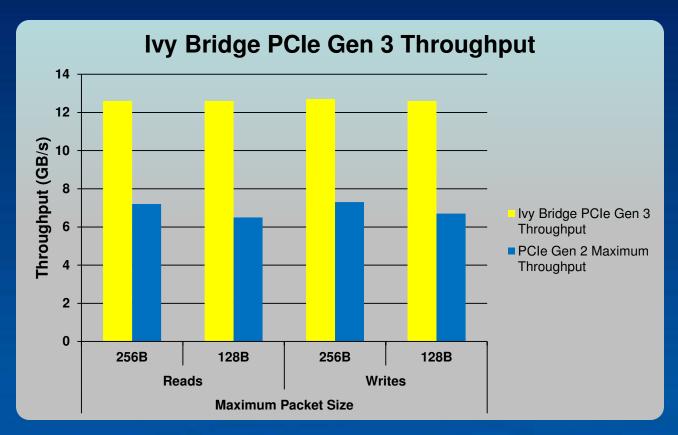
Utilizes Gen 2 I/O channel characteristics

- Enables compatibility with previous Gen components
- Enables drop-in upgrade for Sandy Bridge-based platforms

Supports PCIe bandwidth management & ASPM states

Dynamic Link Width Configuration, L0s (Rx & Tx), L1

Ivy Bridge PCIe Performance*



Ivy Bridge delivers nearly 2x Gen 2 bandwidth

At similar latencies

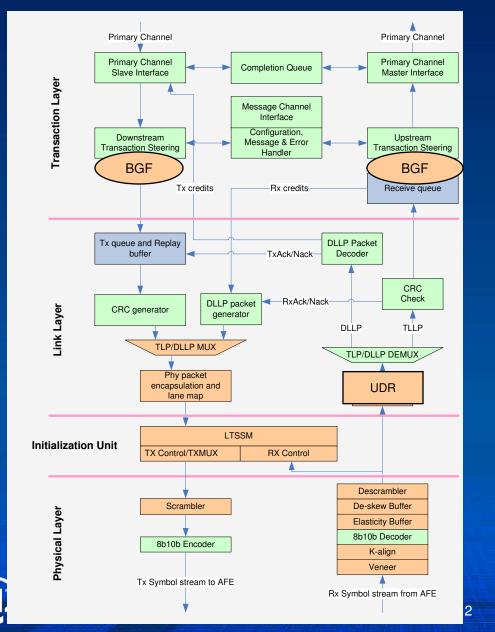


~300ns typical for upstream read request

Ivy Bridge - Hot Chips 2012

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

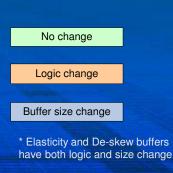
Ivy Bridge PCIe Logic Changes



- Sandy Bridge PCIe uArch unchanged
 - No change to primary channel hub/TL interface

 No change to controller/PHY lanes interface

Gen 3 changes layered on top of Gen 2 functionality (additional states, arcs)
Parallel flows implemented where feasible



43

IPC Improvements



Most Significant IVB IPC Improvements

Pipeline MOV elimination

 Eliminates Move related micro-operations from the processor execution pipeline

Pipelined divider

• Improves throughput of divide related computations

Next page prefetcher

• Enables prefetching to span across a 4K page boundary

Shift/Rotate performance

- Addresses glass jaw concern with crypto and hashing algorithms
- Addresses clumsiness of partial flag handling
- 6 additional split load registers
 - Improves performance for loads splitting cache lines
 - Especially critical for AVX or SSE



Uncore IPC Features

□ AFP – Adaptive Fill Policy

Cache heuristics to identify and segregate streaming applications

QLRU – Quad-Age LRU algorithm

- Allows fine-grain "age assignment" on cache allocation
- E.g.: prefetched requests are allocated at "middle age"

DPT – Dynamic Prefetch Throttling

- Real-time memory bandwidth monitor
- Directs core prefetchers to reduce prefetch aggressiveness during high memory load scenarios

Channel Hashing -- DRAM channel selection mechanism

- Allows channel selection to be made based on multiple address bits
- Historically, it had been "A[6]"
- Allows more even distribution of memory accesses across channels



Legal Notices and Disclaimers

- INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL® PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. INTEL PRODUCTS ARE NOT INTENDED FOR USE IN MEDICAL, LIFE SAVING, OR LIFE SUSTAINING APPLICATIONS.
- Intel may make changes to specifications and product descriptions at any time, without notice.
- All products, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice.
- Intel, processors, chipsets, and desktop boards may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available on request.
- Any code names featured are used internally within Intel to identify products that are in development and not yet publicly announced for release. Customers, licensees and other third parties are not authorized by Intel to use code names in advertising, promotion or marketing of any product or services and any such use of Intel's internal code names is at the sole risk of the user.
- Intel product plans in this presentation do not constitute Intel plan of record product roadmaps. Please contact your Intel representative to obtain Intel's current plan of record product roadmaps.
- Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to <u>http://www.intel.com/performance</u>
- Intel, Intel Inside, the Intel logo, Centrino, Intel Core, Intel Atom, Pentium and UltraBook are trademarks of Intel Corporation in the United States and other countries.
- <u>Material in this presentation is intended as product positioning and not approved end user messaging.</u>
- This document contains information on products in the design phase of development.
- *Other names and brands may be claimed as the property of others.
- Copyright © 2012 Intel Corporation, All Rights Reserved

Legal Notices and Disclaimers, cont.

- WiMAX connectivity requires a WiMAX enabled device and subscription to a WiMAX broadband service. WiMAX connectivity may require you to purchase additional software or hardware at extra cost. Availability of WiMAX is limited, check with your service provider for details on availability and network limitations. Broadband performance and results may vary due to environment factors and other variables. See www.intel.com/go/wimax for more information.
- Intel® My WiFi Technology is an optional feature and requires additional software and a Centrino® wireless adapter. Wi-Fi devices must be certified by the Wi-Fi Alliance for 802.11b/g/a in order to connect. See mywifi.intel.com for more details.
- Hyper-Threading Technology requires a computer system with a processor supporting HT Technology and an HT Technology-enabled chipset, BIOS and operating system. Performance will vary depending on the specific hardware and software you use. For more information including details on which processors support HT Technology, see <u>here</u>
- Intel® Turbo Boost Technology requires a PC with a processor with Intel Turbo Boost Technology capability. Intel Turbo Boost Technology performance varies depending on hardware, software and overall system configuration. Check with your PC manufacturer on whether your system delivers Intel Turbo Boost Technology. For more information, see <u>http://www.intel.com/technology/turboboost</u>
- Requires an Intel® Wireless Display enabled PC, TV Adapter, and compatible television. Available on select Intel® Core processors. Does not support Blu-Ray or other protected content playback. Consult your PC manufacturer. For more information, see <u>www.intel.com/go/wirelessdisplay</u>
- (Built-in Visuals) Available on the 2nd gen Intel® Core[™] processor family. Includes Intel® HD Graphics, Intel® Quick Sync Video, Intel® Clear Video HD Technology, Intel® InTru[™] 3D Technology, and Intel® Advanced Vector Extensions. Also optionally includes Intel® Wireless Display depending on whether enabled on a given system or not. Whether you will receive the benefits of built-in visuals depends upon the particular design of the PC you choose. Consult your PC manufacturer whether built-in visuals are enabled on your system. Learn more about built-in visuals at http://www.intel.com/technology/visualtechnology/index.htm.
- Intel® Insider™ is a hardware-based content protection mechanism. Requires a 2nd generation Intel® Core™ processor-based PC with built-in visuals enabled, an Internet connection, and content purchase or rental from qualified providers. Consult your PC manufacturer. For more information, visit <u>www.intel.com/go/intelinsider</u>.
- Viewing Stereo 3D content requires 3D glasses and a 3D capable display. Physical risk factors may be present when viewing 3D material

Legal Notices and Disclaimers, cont.

- Security features enabled by Intel® AMT require an enabled chipset, network hardware and software and a corporate network connection. Intel AMT may not be available or certain capabilities may be limited over a host OS-based VPN or when connecting wirelessly, on battery power, sleeping, hibernating or powered off. Setup requires configuration and may require scripting with the management console or further integration into existing security frameworks, and modifications or implementation of new business processes. For more information, see http://www.intel.com/technology/manage/iamt.
- No system can provide absolute security under all conditions. Requires an enabled chipset, BIOS, firmware and software and a subscription with a capable Service Provider. Consult your system manufacturer and Service Provider for availability and functionality. Intel assumes no liability for lost or stolen data and/or systems or any other damages resulting thereof. For more information, visit <u>http://www.intel.com/go/anti-theft</u>
- Requires an Execute Disable Bit enabled system. Check with your PC manufacturer to determine whether your system delivers this functionality. For more information, visit <u>http://www.intel.com/technology/xdbit/index.htm</u>
- Intel® vPro[™] Technology is sophisticated and requires setup and activation. Availability of features and results will depend upon the setup and configuration of your hardware, software and IT environment. To learn more visit: <u>http://www.intel.com/technology/vpro</u>
- The original equipment manufacturer must provide TPM functionality, which requires a TPM-supported BIOS. TPM functionality must be initialized and may not be available in all countries.
- Intel® AES-NI requires a computer system with an AES-NI enabled processor, as well as non-Intel software to execute the instructions in the correct sequence. AES-NI is available on select Intel® processors. For availability, consult your reseller or system manufacturer. For more information, see http://software.intel.com/enus/articles/intel-advanced-encryption-standard-instructions-aes-ni/
- No system can provide absolute security under all conditions. Requires an Intel IPT enabled system, including a 2nd generation Intel Core processor, enabled chipset, firmware, and software. Available only on participating websites. Consult your system manufacturer. Intel assumes no liability for lost or stolen data and/or systems or any resulting damages. For more information, visit <u>http://www.ipt.intel.com</u>

