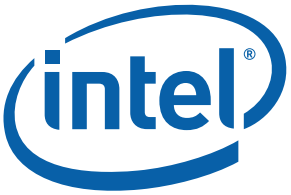


Intel® Xeon® Processor E5 v2 Product Family

Specification Update

June 2014

Reference Number: 329189-005

Contents

Tables

Revision History

Version	Description	Date
001	Initial release.	September 2013
002	Added Errata CA110 - CA116	November 2013
003	Added Errata CA117 - CA120 Added Intel® Xeon® E5-4600 v2 product family	March 2014
004	Updated Erratum CA21 Updated Erratum CA93 Updated Erratum CA120 Updated Related Documents Updated Table 4. Documentation Changes Added Errata CA121 - CA135	May 2014
005	Added Errata CA136 through CA143	June 2014

Preface

This document is an update to the specifications contained in the Affected Documents table below. This document is a compilation of device and documentation errata, specification clarifications and changes. It is intended for hardware system manufacturers and software developers of applications, operating systems, or tools.

Information types defined in Nomenclature are consolidated into the specification update and are no longer published in other documents.

This document may also contain information that was not previously published.

Affected Documents

Document title	Document number / location
Intel® Xeon® Processor E5-1600 v2/E5-2600 v2 Product Families Datasheet - Volume One of Two	329187
Intel® Xeon® Processor E5 v2 Product Family Datasheet- Volume Two: Registers	329188

Related Documents

Nomenclature

Errata are design defects or errors. These may cause the Intel® Xeon® Processor E5 v2 Product Family's behavior to deviate from published specifications. Hardware and software designed to be used with any given stepping must assume that all errata documented for that stepping are present on all devices.

Specification changes are modifications to the current published specifications. These changes will be incorporated in any new release of the specification.

Specification clarifications describe a specification in greater detail or further highlight a specification's impact to a complex design situation. These clarifications will be incorporated in any new release of the specification.

Documentation changes include typos, errors, or omissions from the current published specifications. These will be incorporated in any new release of the specification.

Note: Errata remain in the specification update throughout the product's lifecycle, or until a particular stepping is no longer commercially available. Under these circumstances, errata removed from the specification update are archived and available upon request. Specification changes, specification clarifications and documentation changes are removed from the specification update when the appropriate changes are made to the appropriate product specification or user documentation (datasheets, manuals, and so forth).

Summary Tables of Changes

The following tables indicate the errata, specification changes, specification clarifications, or documentation changes which apply to the Intel® Xeon® Processor E5 v2 Product Family. Intel may fix some of the errata in a future stepping of the component, and account for the other outstanding issues through documentation or specification changes as noted. These tables uses the following notations:

Codes used in summary tables

Stepping

X:	Errata exists in the stepping indicated. Specification Change or Clarification that applies to this stepping.
(No mark) or (Blank box):	This erratum is fixed in listed stepping or specification change does not apply to listed stepping.

Page

(Page):	Page location of item in this document.

Status

Doc:	Document change or update will be implemented.
Plan Fix:	This erratum may be fixed in a future stepping of the product.
Fixed:	This erratum has been previously fixed.
No Fix:	There are no plans to fix this erratum.

Row

	Change bar to left of table row indicates this erratum is either new or modified from the previous version of the document.

Table 1. Errata (Sheet 1 of 5)

Number	Stepping	Status	Errata
	C-1 / M-1 / S-1		
CA1	X	No Fix	Core Frequencies at or Below the DRAM DDR Frequency May Result in Unpredictable System Behavior.
CA2	X	No Fix	DWORD Aligned XOR DMA Sources May Prevent Further DMA XOR Progress.
CA3	X	No Fix	Rank Sparing May Cause an Extended System Stall.
CA4	X	No Fix	Intel® QuickData Technology DMA Lock Quiescent Flow Causes DMA State Machine to Hang.
CA5	X	No Fix	Suspending/Resetting a DMA XOR Channel May Cause an Incorrect Data Transfer on Other Active Channels.
CA6	X	No Fix	Quad Rank DIMMs May Not be Properly Refreshed During IBT_OFF Mode.
CA7	X	No Fix	Intel® QuickData Technology Continues to Issue Requests After Detecting 64-bit Addressing Errors.
CA8	X	No Fix	PCIe* TPH Attributes May Result in Unpredictable System Behavior.
CA9	X	No Fix	PCIe* Rx Common Mode Return Loss is Not Meeting the Specification.
CA10	X	No Fix	Intel® QPI Tx AC Common Mode Fails Specification.
CA11	X	No Fix	PCIe* Rx DC Common Mode Impedance is Not Meeting the Specification.
CA12	X	No Fix	QPILS Reports the VNA/VN0 Credits Available for the Processor Rx Rather Than Tx.
CA13	X	No Fix	A PECI RdPciConfigLocal Command Referencing a Non-Existent Device May Return an Unexpected Value.
CA14	X	No Fix	The Vswing of the PCIe* Transmitter Exceeds the Specification.
CA15	X	No Fix	PECI Write Requests That Require a Retry Will Always Time Out.
CA16	X	No Fix	The Intel® QPI Link Status Register LinkInitStatus Field Incorrectly Reports "Internal Stall Link Initialization" for Certain Stall Conditions.
CA17	X	No Fix	The Processor Does Not Detect Intel® QPI RSVD_CHK Field Violations.
CA18	X	No Fix	Intel® QuickData Technology DMA Non-Page-Aligned Next Source/Destination Addresses May Result in Unpredictable System Behavior.
CA19	X	No Fix	Intel® QPI May Report a Reserved Value in The Link Initialization Status Field During Link Training.
CA20	X	No Fix	Enabling Opportunistic Self-Refresh and Pkg C2 State Can Severely Degrade PCIe* Bandwidth.
CA21	X	No Fix	Functionally Benign PCIe* Electrical Specification Violation Compendium.
CA22	X	No Fix	Warm Resets May be Converted To Power-On Resets When Recovering From an IERR.
CA23	X	No Fix	Patrol Scrubbing During Memory Mirroring May Improperly Signal Uncorrectable Machine Checks.
CA24	X	No Fix	A Modification To The Multiple Message Enable Field Does Not Affect The AER Interrupt Message Number Field.
CA25	X	No Fix	Long latency Transactions Can Cause I/O Devices On The Same Link to Time Out.
CA26	X	No Fix	Intel® QPI Link Layer Does Not Drop Unsupported or Undefined Packets.
CA27	X	No Fix	Coherent Interface Write Cache May Report False Correctable ECC Errors During Cold Reset.
CA28	X	No Fix	Combining ROL Transactions With Non-ROL Transactions or Marker Skipping Operations May Result in a System Hang.
CA29	X	No Fix	Excessive DRAM RAPL Power Throttling May Lead to a System Hang or USB Device Offlining.
CA30	X	No Fix	TSOD-Related SMBus Transactions may not Complete When Package C-States are Enabled.
CA31	X	No Fix	The Integrated Memory Controller does not Enforce CKE High For tXSDLL DCLKs After Self-Refresh.

Table 1. Errata (Sheet 2 of 5)

Number	Stepping	Status	Errata
	C-1 / M-1 / S-1		
CA32	X	No Fix	Intel® QuickData Technology DMA Suspend does not Transition From ARMED to HALT State.
CA33	X	No Fix	Routing Intel® High Definition Audio Traffic Through VC1 May Result in System Hang.
CA34	X	No Fix	NTB Operating in NTB/RP Mode with MSI/MSI-X Interrupts May Cause System Hang.
CA35	X	No Fix	Patrol Scrubbing does not Skip Ranks Disabled After DDR Training.
CA36	X	No Fix	Writes to SDOORBELL or B2BDOORBELL in Conjunction With Inbound Access to NTB MMIO Space May Hang System.
CA37	X	No Fix	DR6.B0-B3 May Not Report All Breakpoints Matched When a MOV/POP SS is Followed by a REP MOVSB or STOSB
CA38	X	No Fix	64-bit REP MOVSB/STOSB May Clear The Upper 32-bits of RCX, RDI And RSI Before Any Data is Transferred
CA39	X	No Fix	An Interrupt Recognized Prior to First Iteration of REP MOVSB/STOSB May Result EFLAGS.RF Being Incorrectly Set
CA40	X	No Fix	Instructions Retired Event May Over Count Execution of IRET Instructions
CA41	X	No Fix	An Event May Intervene Before a System Management Interrupt That Results from IN or INS
CA42	X	No Fix	Execution of VAESIMC or VAESKEYGENASSIST With An Illegal Value for VEX.vvvv May Produce a #NM Exception
CA43	X	No Fix	Unexpected #UD on VZEROALL/VZEROUPPER
CA44	X	No Fix	Successive Fixed Counter Overflows May be Discarded
CA45	X	No Fix	Execution of FXSAVE or FXRSTOR With the VEX Prefix May Produce a #NM Exception
CA46	X	No Fix	VM Exits Due to "NMI-Window Exiting" May Not Occur Following a VM Entry to the Shutdown State
CA47	X	No Fix	Execution of INVVPID Outside 64-Bit Mode Cannot Invalidate Translations For 64-Bit Linear Addresses
CA48	X	No Fix	REP MOVSB May Incorrectly Update ECX, ESI, and EDI
CA49	X	No Fix	Performance-Counter Overflow Indication May Cause Undesired Behavior
CA50	X	No Fix	VEX.L is not Ignored with VCVT*2SI Instructions
CA51	X	No Fix	Concurrently Changing the Memory Type and Page Size May Lead to a System Hang
CA52	X	No Fix	MCI_ADDR May be Incorrect For Cache Parity Errors
CA53	X	No Fix	Instruction Fetches Page-Table Walks May be Made Speculatively to Uncacheable Memory
CA54	X	No Fix	REP MOVS/STOS Executing with Fast Strings Enabled and Crossing Page Boundaries with Inconsistent Memory Types may use an Incorrect Data Size or Lead to Memory-Ordering Violations
CA55	X	No Fix	The Processor May Not Properly Execute Code Modified Using A Floating-Point Store
CA56	X	No Fix	VM Exits Due to GETSEC May Save an Incorrect Value for "Blocking by STI" in the Context of Probe-Mode Redirection
CA57	X	No Fix	IA32_MC5_CTL2 is Not Cleared by a Warm Reset
CA58	X	No Fix	The Processor May Report a #TS Instead of a #GP Fault
CA59	X	No Fix	IO_SMI Indication in SMRAM State Save Area May be Set Incorrectly
CA60	X	No Fix	Performance Monitor SSE Retired Instructions May Return Incorrect Values
CA61	X	No Fix	IRET under Certain Conditions May Cause an Unexpected Alignment Check Exception
CA62	X	No Fix	Performance Monitoring Event FP_MMX_TRANS_TO_MMX May Not Count Some Transitions
CA63	X	No Fix	General Protection Fault (#GP) for Instructions Greater than 15 Bytes May be Preempted
CA64	X	No Fix	LBR, BTS, BTM May Report a Wrong Address When an Exception/Interrupt Occurs in 64-bit Mode

Table 1. Errata (Sheet 3 of 5)

Number	Stepping	Status	Errata
	C-1 / M-1 / S-1		
CA65	X	No Fix	Incorrect Address Computed For Last Byte of FXSAVE/FXRSTOR or XSAVE/XRSTOR Image Leads to Partial Memory Update
CA66	X	No Fix	Values for LBR/BTS/BTM Will be Incorrect after an Exit from SMM
CA67	X	No Fix	EFLAGS Discrepancy on Page Faults and on EPT-Induced VM Exits after a Translation Change
CA68	X	No Fix	B0-B3 Bits in DR6 For Non-Enabled Breakpoints May Be Incorrectly Set
CA69	X	No Fix	MCi_Status Overflow Bit May Be Incorrectly Set on a Single Instance of a DTLB Error
CA70	X	No Fix	Debug Exception Flags DR6.B0-B3 Flags May Be Incorrect for Disabled Breakpoints
CA71	X	No Fix	LER MSRs May Be Unreliable
CA72	X	No Fix	Storage of PEBS Record Delayed Following Execution of MOV SS or STI
CA73	X	No Fix	PEBS Record Not Updated When in Probe Mode
CA74	X	No Fix	Faulting MMX Instruction May Incorrectly Update x87 FPU Tag Word
CA75	X	No Fix	#GP on Segment Selector Descriptor That Straddles Canonical Boundary May Not Provide Correct Exception Error Code
CA76	X	No Fix	APIC Error "Received Illegal Vector" May Be Lost
CA77	X	No Fix	Changing the Memory Type for an In-Use Page Translation May Lead to Memory-Ordering Violations
CA78	X	No Fix	Reported Memory Type May Not be Used to Access the VMCS and Referenced Data Structures
CA79	X	No Fix	LBR, BTM or BTS Records May have Incorrect Branch From Information After an EIST/T-state/S-state/C1E Transition or Adaptive Thermal Throttling
CA80	X	No Fix	FP Data Operand Pointer May Be Incorrectly Calculated After an FP Access Which Wraps a 4-Gbyte Boundary in Code That Uses 32-Bit Address Size in 64-bit Mode
CA81	X	No Fix	VMREAD/VMWRITE Instruction May Not Fail When Accessing an Unsupported Field in VMCS
CA82	X	No Fix	An Unexpected PMI May Occur After Writing a Large Value to IA32_FIXED_CTR2
CA83	X	No Fix	A Write to the IA32_FIXED_CTR1 MSR May Result in Incorrect Value in Certain Conditions
CA84	X	No Fix	#GP May be Signaled When Invalid VEX Prefix Precedes Conditional Branch Instructions
CA85	X	No Fix	Interrupt From Local APIC Timer May Not Be Detectable While Being Delivered
CA86	X	No Fix	PCMPESTRI, PCMPESTRM, VPCMPESTRI and VPCMPESTRM Always Operate with 32-bit Length Registers
CA87	X	No Fix	During Package Power States Repeated PCIe* and/or DMI L1 Transitions May Cause a System Hang
CA87	X	No Fix	During Package Power States Repeated PCIe* and/or DMI L1 Transitions May Cause a System Hang
CA88	X	No Fix	RDMSR of IA32_PERFEVTSEL4-7 May Return an Incorrect Result
CA89	X	No Fix	MONITOR or CLFLUSH on the Local XAPIC's Address Space Results in Hang
CA90	X	No Fix	PCMPESTRI, PCMPESTRM, VPCMPESTRI and VPCMPESTRM Always Operate With 32-bit Length Registers
CA91	X	No Fix	Clock Modulation Duty Cycle Cannot Be Programmed to 6.25%
CA92	X	No Fix	Processor May Livelock During On Demand Clock Modulation
CA93	X	No Fix	Performance Monitor Counters May Produce Incorrect Results
CA94	X	No Fix	Virtual-APIC Page Accesses With 32-Bit PAE Paging May Cause a System Crash
CA95	X	No Fix	IA32_FEATURE_CONTROL MSR May be Un-Initialized on a Cold Reset
CA96	X	No Fix	PEBS May Unexpectedly Signal a PMI After the PEBS Buffer is Full
CA97	X	No Fix	Execution of GETSEC[SEXIT] May Cause a Debug Exception to Be Lost

Table 1. Errata (Sheet 4 of 5)

Number	Stepping	Status	Errata
	C-1 / M-1 / S-1		
CA98	X	No Fix	An Uncorrectable Error Logged in IA32_CR_MC2_STATUS May also Result in a System Hang
CA99	X	No Fix	The Corrected Error Count Overflow Bit in IA32_ MC0_STATUS is Not Updated After a UC Error is Logged
CA100	X	No Fix	IA32_VMX_VMCS_ENUM MSR (48AH) Does Not Properly Report the Highest Index Value Used for VMCS Encoding
CA101	X	No Fix	The Upper 32 Bits of CR3 May be Incorrectly Used With 32-Bit Paging
CA102	X	No Fix	EPT Violations May Report Bits 11:0 of Guest Linear Address Incorrectly
CA103	X	No Fix	Intel® QuickData Technology DMA Access to Invalid Memory Address May Cause System Hang
CA104	X	No Fix	CPUID Faulting is Not Enumerated Properly
CA105	X	No Fix	TSC is Not Affected by Warm Reset
CA106	X	No Fix	Incorrect Size Reported by PCIe* NTB BAR Registers
CA107	X	No Fix	The Vswing of the PCIe* Transmitter Exceeds The Specification
CA108	X	No Fix	PECI_WAKE_MODE is Always Reported as Disabled
CA109	X	No Fix	Poisoned PCIe* AtomicOp Completions May Return an Incorrect Byte Count
CA110	X	No Fix	Incorrect Speed and De-emphasis Level Selection During DMI Compliance Testing
CA111	X	No Fix	PCIe* Device 3 Does Not Log an Error in UNCERRSTS When an Invalid Sequence Number in an Ack DLLP is Received
CA112	X	No Fix	Programmable Ratio Limits For Turbo Mode is Reported as Disabled
CA113	X	No Fix	PCIe* TLPs in Disabled VC Are Not Reported as Malformed
CA114	X	No Fix	PCIe* Link May Fail to Train to 8.0 GT/s
CA115	X	No Fix	PCIe* Header of a Malformed TLP is Logged Incorrectly
CA116	X	No Fix	PCIe* May Associate Lanes That Are Not Part of Initial Link Training to L0 During Upconfiguration
CA117	X	No Fix	Single PCIe* ACS Violation or UR Response May Result in Multiple Correctable Errors Logged
CA118	X	No Fix	PCIe* Extended Tag Field May be Improperly Set
CA119	X	No Fix	Power Meter May Under-Estimate Package Power
CA120	X	No Fix	DTS2.0 May Report Inaccurate Temperature Margin
CA121	X	No Fix	PMON Counters Overflow May Not Trigger PMON Global Freeze
CA122	X	No Fix	Processor May Log a Machine Check when MSI is signaled by a device
CA123	X	No Fix	Spurious Patrol Scrub Errors Observed During a Warm Reset
CA124	X	No Fix	PECI May Not be Able to Access IIO CSRs
CA125	X	No Fix	A DMI UR May Unexpectedly Cause a CATERR# After a Warm Reset
CA126	X	No Fix	Duplicate. Erratum Removed
CA127	X	No Fix	VM Exit May Set IA32_EFER.NXE When IA32_MISC_ENABLE Bit 34 is Set to 1
CA128	X	No Fix	Receiver Termination Impedance On PCIe* 3.0 Does Not Comply With The Specification
CA129	X	No Fix	Platform Recovery After a Machine Check May Fail
CA130	X	No Fix	PECI May be Non-responsive When System is in BMC Init Mode
CA131	X	No Fix	Processor May Issue Unexpected NAK DLLP Upon PCIe* L1 Exit
CA132	X	No Fix	Surprise Down Error Status is Not Set Correctly on DMI Port
CA133	X	No Fix	Core C-state Residency Counters May Return Stale Data

Table 1. Errata (Sheet 5 of 5)

Number	Stepping	Status	Errata
	C-1 / M-1 / S-1		
CA134	X	No Fix	PCIe* Ports Operating at 8 GT/s May Issue an Additional Packet After Stop and Scream Occurs
CA135	X	No Fix	Incorrect Page Translation when EPT is enabled
CA136	X	No Fix	A Machine Check Exception Due to Instruction Fetch May Be Delivered Before an Instruction Breakpoint
CA137	X	No Fix	Accessing Physical Memory Space 0-640K through the Graphics Aperture May Cause Unpredictable System Behavior
CA138	X	No Fix	The RDRAND Instruction Will Not Execute as Expected
CA139	X	No Fix	Writes to B2BSPAD[15:0] Registers May Transfer Corrupt Data Between NTB Connected Systems
CA140	X	No Fix	Reading DDRIO Broadcast CSRs Via PECI May Return Incorrect Data
CA141	X	No Fix	Corrected Filtering Indication May be Incorrect in LLC Machine Check Bank Status Register
CA142	X	No Fix	RTID_POOL_CONFIG Registers Incorrectly Behave as a Read-Write Registers
CA143	X	No Fix	Catastrophic Trip Triggered at Lower Than Expected Temperatures

Table 2. Specification Clarifications

No.	Specification clarifications
1	None

Table 3. Specification Changes

No.	Specification changes
1	None

Table 4. Documentation Changes

No.	Documentation changes
1	SDM, Volume 3B: On-Demand Clock Modulation Feature Clarification
2	Intel® Xeon® Processor E5 v2 Product Family Datasheet- Volume Two: Device Mapping Addition

Identification Information

Component identification via programming interface

The Intel® Xeon® Processor E5 v2 Product Family stepping can be identified by the following register contents.

Table 5. Intel® Xeon® Processor E5 v2 Product Family Signature/Version

Reserved	Extended family[1]	Extended model[2]	Reserved	Processor type[3]	Family code[4]	Model number[5]	Stepping ID
31:28	27:20	19:16	15:14	13:12	11:8	7:4	3:0
	00000000b	0011b		00b	0110b	1110b	C1 =0100
	00000000b	0011b		00b	0110b	1110b	M1 =0100
	00000000b	0011b		00b	0110b	1110b	S1 =0100

Notes:

1. The Extended Family, Bits [27:20] are used in conjunction with the Family Code, specified in Bits [11:8], to indicate whether the processor belongs to the Intel386™, Intel486™, Pentium®, Pentium 4, or Intel® Core™ processor family.
2. The Extended Model, Bits [19:16] in conjunction with the Model Number, specified in Bits [7:4], are used to identify the model of the processor within the processor's family.
3. The Family Code corresponds to Bits [11:8] of the EDX register after RESET, Bits [11:8] of the EAX register after the CPUID instruction is executed with a 1 in the EAX register, and the generation field of the Device ID register accessible through Boundary Scan.
4. The Model Number corresponds to Bits [7:4] of the EDX register after RESET, Bits [7:4] of the EAX register after the CPUID instruction is executed with a 1 in the EAX register, and the model field of the Device ID register accessible through Boundary Scan.
5. The Stepping ID in Bits [3:0] indicates the revision number of that model.

The following table provides the CPUID, CAPID0 and CAPID4 values for each stepping used for across -EP and -EN packages.

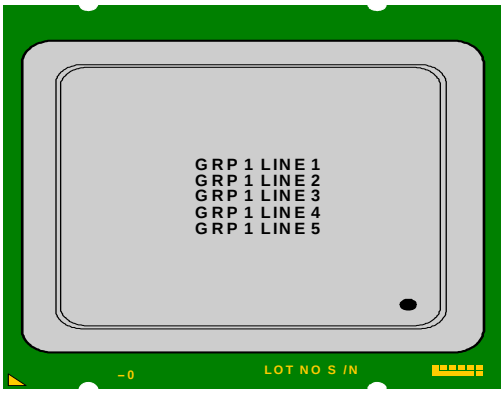
Table 6. Intel® Xeon® Processor E5 v2 Product Family Stepping Identification

Core Count	Stepping	Package	CPUID	CAPID4 (CPUBUSNO(1): 0n10:3:0x94) sv.socket0.uncore0.capid4						CAPID0 (CPUBUSNO(1): 0n10:3:0x84) sv.socket0.uncore0.capid0
			Bit	19	20	21	22	23	25	1
HCC 12-core	C1	EP/EP 4S	0x306E4	x	x	0	1	1	1	0
MCC 10/8-core	M1	EP/EP 4S	0x306E4	0	0	0	0	1	1	0
	M1	EN	0x306E4	0	0	0	0	1	1	1
LCC 6/4-core	S1	EP/EP 4S	0x306E4	0	1	0	0	1	1	0
	S1	EN	0x306E4	0	1	0	0	1	1	1

Component marking

The Intel® Xeon® Processor E5 v2 Product Family stepping can be identified by the following component markings:

Figure 1. Intel® Xeon® Processor E5 v2 Product Family top-side markings (example)



Legend	Mark Text (Production Mark):
GRP1 LINE1:	i{M}{C}YY
GRP1 LINE2:	SUB- BRAND PROC #
GRP1 LINE3:	SSPEC SPEED
GRP1 LINE4:	XXXXX
GRP1 LINE5:	{FPO} {e4}

Table 7. Intel® Xeon® Processor E5-1600/2600 v2 Product Family Identification (Sheet 1 of 2)

S-Spec Number	Stepping	Model Number	CPUID	Core frequency (GHz)/DDR3 (MHz)/ Intel® QPI (GHz)	Available bins of Intel® Turbo Boost Technology	TDP (W)	# Cores	Cache size (MB)	Notes
SR1BA	C-1	E5-2695v2	0x306E4	2.40/1866/8	4/4/4/4/4/4/ 4/4/5/6/7/8	115	12	30	1,2,3,7
SR19H	C-1	E5-2697v2	0x306E4	2.70/1866/8	3/3/3/3/3/3/ 3/4/5/6/7/8	130	12	30	1,2,3,7
SR19V	M-1	E5-2687Wv2	0x306E4	3.40/1866/8	2/2/2/2/3/4/ 5/6	150	8	25	1,2,3,6, 7
SR19W	M-1	E5-2667v2	0x306E4	3.30/1866/8	3/3/3/3/4/5/ 6/7	130	8	25	1,2,3,7
SR19X	M-1	E5-2643v2	0x306E4	3.50/1866/8	1/1/1/1/2/3	130	6	25	1,2,3,7
SR19Y	M-1	E5-2650Lv2	0x306E4	1.70/1600/8	2/2/2/2/2/2/ 2/2/3/4	70	10	25	1,2,3,7
SR19Z	M-1	E5-2640v2	0x306E4	2/1600/7.2	3/3/3/3/3/3/ 4/5	95	8	20	1,2,3,7
SR1A0	M-1	E5-2658v2	0x306E4	2.40/1866/8	2/2/3/3/4/4/ 5/5/6/6	95	10	25	1,2,3,7
SR1A2	M-1	E5-2648Lv2	0x306E4	1.90/1866/8	2/2/3/3/4/4/ 5/5/6/6	70	10	25	1,2,3,7
SR1A5	M-1	E5-2690v2	0x306E4	3/1866/8	3/3/3/3/3/3/ 3/4/5/6	130	10	25	1,2,3,7
SR1A6	M-1	E5-2680v2	0x306E4	2.80/1866/8	3/3/3/3/3/4/ 5/6/7/8	115	10	25	1,2,3,7
SR1A7	M-1	E5-2670v2	0x306E4	2.50/1866/8	4/4/4/4/4/4/ 5/6/7/8	115	10	25	1,2,3,7
SR1A8	M-1	E5-2650v2	0x306E4	2.60/1866/8	5/5/5/5/5/6/ 7/8	95	8	20	1,2,3,7
SR1AB	M-1	E5-2660v2	0x306E4	2.20/1866/8	4/4/4/4/4/4/ 5/6/7/8	95	10	25	1,2,3,7

Table 7. Intel® Xeon® Processor E5-1600/2600 v2 Product Family Identification (Sheet 2 of 2)

S-Spec Number	Stepping	Model Number	CPUID	Core frequency (GHz)/DDR3 (MHz)/ Intel® QPI (GHz)	Available bins of Intel® Turbo Boost Technology	TDP (W)	# Cores	Cache size (MB)	Notes
SR1AF	M-1	E5-2628Lv2	0x306E4	1.90/1600/7.2	2/2/3/3/4/4/5/5	70	8	20	1,2,3,7
SR1MJ	M-1	E5-1680v2	0x306E4	3/1866/NA	4/4/4/4/5/7/8/9	130	8	25	1,2,3,6,7
SR1AM	S-1	E5-2630v2	0x306E4	2.60/1600/7.2	3/3/3/3/4/5	80	6	15	1,2,3,7
SR1AN	S-1	E5-2620v2	0x306E4	2.10/1600/7.2	3/3/3/3/4/5	80	6	15	1,2,3,7
SR1AP	S-1	E5-1660v2	0x306E4	3.70/1866/NA	1/1/1/1/2/3	130	6	15	1,2,3,6,7
SR1AQ	S-1	E5-1650v2	0x306E4	3.50/1866/NA	1/1/2/2/2/4	130	6	12	1,2,3,6,7
SR1AR	S-1	E5-1620v2	0x306E4	3.70/1866/NA	0/0/0/2	130	4	10	1,2,3,6,7
SR1AX	S-1	E5-2609v2	0x306E4	2.50/1333/6.4	NA	80	4	10	1,2,3,4,5
SR1AY	S-1	E5-2603v2	0x306E4	1.80/1333/6.4	NA	80	4	10	1,2,3,7
SR1AZ	S-1	E5-2630Lv2	0x306E4	2.40/1600/7.2	2/2/2/2/3/4	60	6	15	1,2,3,7
SR1B7	S-1	E5-2637v2	0x306E4	3.50/1866/8	1/1/2/3	130	4	15	1,2,3,7
SR1B8	S-1	E5-2618Lv2	0x306E4	2/1333/6.4	NA	50	6	15	1,2,3,5

Notes:
1. Intel® Xeon® Processor E5-1600 v2 and E5-2600 v2 Product Families VID codes will change due to temperature and/or current load changes in order to minimize the power of the part. For specific voltages, refer to the latest Intel® Xeon® Processor E5-1600 v2/E5-2600 v2 Product Families Datasheet - Volume One of Two, #329187.
2. Refer to the latest revision of the following documents for information on processor specifications and features: Intel® Xeon® Processor E5-1600 v2/E5-2600 v2 Product Families Datasheet - Volume One of Two #329187, Intel® Xeon® Processor E5-1 v2 Product Families Datasheet - Volume Two #329188-001.
3. Please refer to the latest Intel® Xeon® Processor E5-1600 v2/E5-2600 v2 Product Families Datasheet - Volume One of Two, #329187 for information on processor operating temperature and thermal specifications.
4. This SKU does not support Intel® Hyper Threading Technology.
5. This SKU does not support Intel® Turbo Boost Technology.
6. This SKU is intended for workstations only and uses workstation specific use conditions for reliability assumptions.
7. Intel® Turbo Boost Technology performance varies depending on hardware, software and overall system configuration.
8.

Table 8. Intel® Xeon® Processor E5-4600 v2 Product Family Identification (Sheet 1 of 2)

S-Spec Number	Stepping	Model Number	CPUID	Core frequency (GHz)/DDR3 (MHz)/ Intel® QPI (GHz)	Available bins of Intel® Turbo Boost Technology	TDP (W)	# Cores	Cache size (MB)	Notes
QF53	C-1	E5-4657Lv2	0x306E4	2.4/1866/8	3/3/3/3/3/3/3/3/3/4/5	115	12	30	1,2,3,6,7
QF96	C-1	E5-4610v2	0x306E4	2.3/1600/7.2	2/2/2/2/2/2/3/4	95	8	16	1,2,3,6,7
QF5X	M-1	E5-4640v2	0x306E4	2.2/1866/8	3/3/3/3/3/3/3/3/4/5	95	10	20	1,2,3,6,7
QF6H	M-1	E5-4624Lv2	0x306E4	1.9/1866/8	2/2/3/3/4/4/5/5/6/6	70	10	25	1,2,3,6,7
QF71	M-1	E5-4620v2	0x306E4	2.6/1600/7.2	2/2/2/2/2/2/3/4	95	8	20	1,2,3,6,7

Table 8. Intel® Xeon® Processor E5-4600 v2 Product Family Identification (Sheet 2 of 2)

S-Spec Number	Stepping	Model Number	CPUID	Core frequency (GHz)/DDR3 (MHz)/ Intel® QPI (GHz)	Available bins of Intel® Turbo Boost Technology	TDP (W)	# Cores	Cache size (MB)	Notes
QF77	M-1	E5-4627v2	0x306E4	3.3/1866/7.2	2/2/2/2/2/2/2/3	130	8	16	1,2,3,4,6,7
QF7D	M-1	E5-4650v2	0x306E4	2.4/1866/8	3/3/3/3/3/3/3/3/4/5	95	10	25	1,2,3,6,7
QF8P	S-1	E5-4607v2	0x306E4	2.6/1333/6.4	NA	95	6	15	1,2,3,5,6,7
QF8T	S-1	E5-4603v2	0x306E4	2.2/1333/6.4	NA	95	4	10	1,2,3,5,7
SR19F	C-1	E5-4657Lv2	0x306E4	2.4/1866/8	3/3/3/3/3/3/3/3/3/3/4/5	115	12	30	1,2,3,6,7
SR19L	C-1	E5-4610v2	0x306E4	2.3/1600/7.2	2/2/2/2/2/2/3/4	95	8	16	1,2,3,6,7
SR19R	M-1	E5-4640v2	0x306E4	2.2/1866/8	3/3/3/3/3/3/3/3/4/5	95	10	20	1,2,3,6,7
SR1A1	M-1	E5-4624Lv2	0x306E4	1.9/1866/8	2/2/3/3/4/4/5/5/6/6	70	10	25	1,2,3,6,7
SR1AA	M-1	E5-4620v2	0x306E4	2.6/1600/7.2	2/2/2/2/2/2/3/4	95	8	20	1,2,3,6,7
SR1AD	M-1	E5-4627v2	0x306E4	3.3/1866/7.2	2/2/2/2/2/2/2/3	130	8	16	1,2,3,4,6,7
SR1AG	M-1	E5-4650v2	0x306E4	2.4/1866/8	3/3/3/3/3/3/3/3/4/5	95	10	25	1,2,3,6,7
SR1B4	S-1	E5-4607v2	0x306E4	2.6/1333/6.4	NA	95	6	15	1,2,3,4,5,6,7
SR1B6	S-1	E5-4603v2	0x306E4	2.2/1333/6.5	NA	95	4	10	1,2,3,4,5,6,7

Notes:

1. Intel® Xeon® Processor E5-4600 v2 Product Families VID codes will change due to temperature and/or current load changes in order to minimize the power of the part. For specific voltages refer to the latest Intel® Xeon® Processor E5-V2 Product Families Datasheets volume 1, #329187-001.
2. Please refer to the latest rev of the following documents for information on processor specifications and features: Intel® Xeon® Processor E5 V2 Product Families Datasheet - Volume One #329187-001, Intel® Xeon® Processor E5-1 V2 Product Families Datasheet - Volume Two #329188-001.
3. Please refer to the latest Intel® Xeon® Processor E5-1 V2 Product Families Datasheet - Volume One, #329187-001 for information on processor operating temperature and thermal specifications.
4. This SKU does not support Intel® Hyper Threading Technology.
5. This SKU does not support Intel® Turbo Boost Technology.
6. Intel® Turbo Boost Technology performance varies depending on hardware, software and overall system configuration.

Errata

CA1 Core Frequencies at or Below the DRAM DDR Frequency May Result in Unpredictable System Behavior.

Problem: The Enhanced Intel SpeedStep® Technology can dynamically adjust the core operating frequency to as low as 1200 MHz. Due to this erratum, under complex conditions and when the cores are operating at or below the DRAM DDR frequency, unpredictable system behavior may result.

Implication: Systems using Enhanced Intel SpeedStep Technology with DDR3-1333 or DDR3-1600 memory devices are subject to unpredictable system behavior.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA2 DWORD Aligned XOR DMA Sources May Prevent Further DMA XOR Progress.

Problem: XOR DMA channels may stop further progress in the presence of Locks/PHOLDs if the source pointed to by a DMA XOR descriptor is not cacheline aligned.

Implication: Non-cacheline aligned DMA XOR sources may hang both channels 0 and 1. A reset is required in order to recover from the hang. Legacy DMA descriptors on any channel have no source alignment restrictions.

Workaround: Software must either:

- Ensure XOR DMA descriptors only point to cacheline aligned sources (best performance) OR
- A legacy DMA copy must be used prior to non-cacheline aligned DMA operations to guarantee that the source misalignment is on DWORD15 of the cacheline. The required source that must be misaligned to DWORD15, depends on the following desired subsequent DMA XOR operations:
  - — DMA XOR Validate (RAID5/ P-Only): The P-source must be misaligned to DWORD15 (last DWORD).
  - — DMA XOR Validate (RAID6/P+Q): The Q-source must be misaligned to DWORD15 (last DWORD).
  - — DMA XOR Generate or Update: The last source (which will be different based on numblk) must be misaligned to DWORD15 (last DWORD).

Status: For the affected steppings, see the Summary Tables of Changes.

CA3 Rank Sparing May Cause an Extended System Stall.

Problem: The Integrated Memory Controller sequencing during a rank sparing copy operation blocks all writes to the memory region associated with the rank being taken out of service. Due to this erratum, this block can result in a system stall that persists until the sparing copy operation completes.

Implication: The system can stall at unpredictable times which may be observed as one time instance of system unavailability.

Workaround: A BIOS workaround has been identified. Refer to Intel® Xeon® Processor E5 Product Family-based Platform CPU/Intel QPI/Memory Reference Code version 1.0.006 or later and release notes.

Status: For the affected steppings, see the Summary Tables of Changes.

CA4 Intel® QuickData Technology DMA Lock Quiescent Flow Causes DMA State Machine to Hang.

Problem: The lock quiescent flow is a means for an agent to gain sole ownership of another agent's resources by preventing other devices from sending transactions. Due to this erratum, during the lock quiescent flow, the Intel® QuickData Technology DMA read and write queues are throttled simultaneously. This prevents subsequent read completions from draining into the write queue, hanging the DMA lock state machine.

Implication: The DMA lock state machine may hang during a lock quiescent flow.

Workaround: Fix was provided in Reference Code version 1.0.000 or later.

Status: For the affected steppings, see the Summary Tables of Changes.

CA5 Suspending/Resetting a DMA XOR Channel May Cause an Incorrect Data Transfer on Other Active Channels.

Problem: Suspending an active DMA XOR channel by setting CHANCMD.Suspend DMA bit (Offset 84; Bit 2) while XOR type DMA channels are active may cause incorrect data transfer on the other active legacy channels. This erratum may also occur while resetting an active DMA XOR channel CHANCMD.Reset DMA bit (Offset 84; Bit 5). CHANCMD is in the region described by CB_BAR (Bus 0; Device 4; function 0-7; Offset 10H).

Implication: Due to this erratum, an incorrect data transfer may occur on the active legacy DMA channels.

Workaround: Software must suspend all legacy DMA channels before suspending an active DMA XOR channel (channel 0 or 1).

Status: For the affected steppings, see the Summary Tables of Changes.

CA6 Quad Rank DIMMs May Not be Properly Refreshed During IBT_OFF Mode.

Problem: The Integrated Memory Controller incorporates a power savings mode known as IBT_OFF (Input Buffer Termination disabled). Due to this erratum, Quad Rank DIMMs may not be properly refreshed during IBT_OFF mode.

Implication: Use of IBT_OFF mode with Quad Rank DIMMs may result in unpredictable system behavior.

Workaround: A BIOS workaround has been identified. Refer to Intel® Xeon® Processor E5 Product Family-based Platform CPU/Intel® QPI/Memory Reference Code version 1.0.006 or later and release notes.

Status: For the affected steppings, see the Summary Tables of Changes.

CA7 Intel® QuickData Technology Continues to Issue Requests After Detecting 64-bit Addressing Errors.

Problem: Intel® QuickData Technology uses the lower 48 address bits of a 64-bit address field. Detection of accesses to source address, destination address, descriptor address, chain address, or completion address outside of this 48-bit range are flagged as "64-bit addressing errors" and should halt DMA processing. Due to this erratum, the Intel® QuickData Technology DMA continues to issue requests after detecting certain 64-bit addressing errors involving RAID operations. The failing condition occurs for 64-bit addressing errors in either a Channel Completion Upper Address Register (CHANCMP_0, CHANCMP_1) (Bus 0; MMIO BAR CB_BAR [0:7]; Offset 98H, 9CH), or in the source or destination addresses of a RAID descriptor.

Implication: Programming out of range DMA address values may result in unpredictable system behavior.

Workaround: Ensure all RAID descriptors, CHANCMP_0, and CHANCMP_1 addresses are within the 48-bit range before starting the DMA engine.

Status: For the affected steppings, see the Summary Tables of Changes.

CA8 PCIe* TPH Attributes May Result in Unpredictable System Behavior.

Problem: TPH (Transactions Processing Hints) are optional aids to optimize internal processing of PCIe* transactions. Due to this erratum, certain transactions with TPH attributes may be misdirected, resulting in unpredictable system behavior.

Implication: Use of the TPH feature may affect system stability.

Workaround: A BIOS workaround has been identified. Refer to Intel® Xeon® Processor E5 Family-based Platform CPU/Intel® QPI/Memory Reference Code version 1.0.006 or later and release notes.

Status: For the affected steppings, see the Summary Tables of Changes.

CA9 PCIe* Rx Common Mode Return Loss is Not Meeting the Specification.

Problem: The PCIe* specification requires that the Rx Common Mode Return Loss in the range of 0.05 to 2.5 GHz must be limited to -6 dB. The processor's PCIe* Rx do not meet this requirement. The PCIe* Rx Common Mode Return at 500 MHz has been found to be between -3.5 and -4 dB on a limited number of samples.

Implication: Intel has not observed any functional failures due to this erratum with any commercially available PCIe* devices.

Workaround: None identified.

Problem: For the affected steppings, see the Summary Tables of Changes.

CA10 Intel® QPI Tx AC Common Mode Fails Specification.

Problem: The Intel® QPI specification requires Tx AC Common Mode (ACCM) to be between -50 mV and 50 mV at 8.0 GT/s. Testing across process, voltage, and temperature showed that the ACCM exceeded the upper end of the specification on several lanes.

Implication: Those performing an electrical characterization of the Intel® QPI interface may notice a violation of the upper end of the ACCM specification by no more than 5 mV.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA11 PCIe* Rx DC Common Mode Impedance is Not Meeting the Specification.

Problem: When the PCIe* Rx termination is not powered, the DC Common Mode impedance has the following requirement: ≥10 kohm over 0 to 200 mV range with respect to ground and ≥20 kohm for voltages ≥200 mV with respect to ground. The processor's PCIe* Rx do not meet this requirement at 85°C or greater. In a limited number of samples Intel has measured an impedance as low as 9.85 kohm at 50 mV.

Implication: Intel has not observed any functional impact due to this violation with any commercially available system.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA12 QPILS Reports the VNA/VN0 Credits Available for the Processor Rx Rather Than Tx.

Problem: The QPILS register (CPUBUS(1); Devices 8,9; Function 0; Offset 0x48), according to the Intel® QuickPath Interconnect Specification at revision 1.1 and later, should report the VNA/VN0 credits available for the processor Tx (Transmit port). Due to this erratum, the QPILS register reports the VNA/VN0 credits available for the processor Rx (Receive port).

Implication: This is a violation of the specification but no functional failures have been observed due to this erratum.

Workaround: None

Status: For the affected steppings, see the Summary Tables of Changes.

CA13 A PECI RdPciConfigLocal Command Referencing a Non-Existent Device May Return an Unexpected Value.

Problem: Configuration reads to nonexistent PCI configuration registers should return 0FFFF_FFFFH. Due to this erratum, when the PECI RdPciConfigLocal command references a nonexistent PCI configuration register, the value 0000_0000H may be returned instead of the expected 0FFFF_FFFFH.

Implication: A PECI RdPciConfigLocal command referencing a nonexistent device may observe a return value of 0000_0000H. Software expecting a return value of 0FFFF_FFFFH to identify nonexistent devices may not work as expected.

Workaround: Software that performs enumeration via the PECI "RdPciConfigLocal" command should interpret 0FFFF_FFFFH and 0000_0000H values for the Vendor Identification and Device Identification Register as indicating a nonexistent device.

Status: For the affected steppings, see the Summary Tables of Changes.

CA14 The Vswing of the PCIe* Transmitter Exceeds the Specification.

Problem: The PCIe* specification defines a limit for the Vswing (voltage swing) of the differential lines that make up a lane to be 1200 mV peak-to-peak when operating at 2.5 GT/s and 5 GT/s. Intel has found that the processor's PCIe* transmitter may exceed this specification. Peak-to-peak swings on a limited number of samples have been observed up to 1450 mV.

Implication: For those taking direct measurements of the PCIe* transmit traffic coming from the processor may detect that the Vswing exceeds the PCIe* specification. Intel has not observed any functional failures due to this erratum.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA15 PECI Write Requests That Require a Retry Will Always Time Out.

Problem: PECI 3.0 introduces a 'Host Identification' field as a way for the PECI host device to identify itself to the PECI client. This is intended for use in future PECI systems that may support more than one PECI originator. Since PECI 3.0 systems do not support the use of multiple originators, PECI 3.0 host devices should zero out the unused Host ID field. PECI 3.0 also introduces a 'retry' bit as a way for the PECI host to indicate to the client that the current request is a 'retry' of a previous read or write operation. Unless the PECI 3.0 host device zeroes out the byte containing the 'Host ID & Retry bit' information, PECI write requests that require a retry will never complete successfully.

Implication: PECI write requests that require a retry may never complete successfully. Instead, they will return a timeout completion code of 81H for a period ranging from 1 ms to 30 ms if the 'RETRY' bit is asserted.

Workaround: PECI 3.0 host devices should zero out the byte that contains the Host ID and Retry bit information for all PECI requests at all times including retries.

Status: For the affected steppings, see the Summary Tables of Changes.

CA16 The Intel® QPI Link Status Register LinkInitStatus Field Incorrectly Reports "Internal Stall Link Initialization" for Certain Stall Conditions.

Problem: The Intel® QPI Link Control register (CPUBUS(1), Devices 8, 9; Function 0; Offset 0x44) bits 17 and 16 allow for the control of the Link Layer Initialization by forcing the link to stall the initialization process until cleared. The Intel® QPI Link Status register (CPUBUS(1), Device 8, 9; Function 0; Offset 0x48) bits 27:24 report the Link Initialization Status (LinkInitStatus). The LinkInitStatus incorrectly reports "Internal Stall Link Initialization" (0001b) for non-Intel® QPI link control register, bit[17,16] stall conditions. The Intel® QPI specification does not intend for internal stall conditions to report that status, but rather report the normal "Waiting for Physical Layer Ready" (0000b).

Implication: There is no known problem with this behavior since there is no usage model that relies on polling of the LinkInitStatus state in the "Waiting for Physical Layer Ready" versus

"Internal Stall Link Initialization" state, and it only advertises the "Internal Stall Link Initialization" state for a brief period of time during Link Layer Initialization.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA17 The Processor Does Not Detect Intel® QPI RSVD_CHK Field Violations.

Problem: According to the Intel® QPI specification, if a target agent receives a packet with a nonzero RSVD_CHK field, it should flag it as an "Intel QPI Link Layer detected unsupported/undefined" packet. Due to this erratum, the processor does not check the RSVD_CHK field nor report the expected error.

Implication: The processor will not flag the "Intel QPI Link Layer detected unsupported/undefined" packet error in the case that the RSVD_CHK field is nonzero.

Workaround: None identified.

Problem: For the affected steppings, see the Summary Tables of Changes.

CA18 Intel® QuickData Technology DMA Non-Page-Aligned Next Source/Destination Addresses May Result in Unpredictable System Behavior.

Problem: Non-page aligned Intel® QuickData Technology DMA next source/destination addresses may cause memory read-write collisions.

Implication: Due to this erratum, using non-page aligned next source/destination addresses may result in unpredictable system behavior.

Workaround: Next source/destination addresses must be page aligned. The Intel-provided Intel® QuickData Technology DMA driver abides by this alignment rule.

Status: For the affected steppings, see the Summary Tables of Changes.

CA19 Intel® QPI May Report a Reserved Value in The Link Initialization Status Field During Link Training.

Problem: An Intel® QPI (Intel® QuickPath Interconnect) link reports its Link Training progress in the Intel® QPI Link Status register. Due to this erratum, the Link Initialization Status (QPILS Bus 1;Device 8,9; Function 0; Offset 48H; bits [27:24]) incorrectly reports a reserved encoding of 1101b while in the "Initial Credit return (initializing credits)" state. The correct encoding for the "Initial Credit return (initializing credits)" state is 0101b.

Implication: Software that monitors the Link Initialization Status field during Link Training may see a reserved encoding reported.

Workaround: None identified. Software may ignore or reinterpret the incorrect encoding for this processor.

Status: For the affected steppings, see the Summary Tables of Changes.

CA20 Enabling Opportunistic Self-Refresh and Pkg C2 State Can Severely Degrade PCIe* Bandwidth.

Problem: Due to this erratum, enabling opportunistic self-refresh can lead to the memory controller over-aggressively transitioning DRAM to self-refresh mode when the processor is in Pkg C2 state.

Implication: The PCIe* interface peak bandwidth can be degraded by as much as 90%

Workaround: A BIOS workaround has been identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA21 Functionally Benign PCIe* Electrical Specification Violation Compendium.

Problem: Violations of PCIe* electrical specifications listed in the table below have been observed.

Specification	Violation Description
Deemphasis ratio limit: -3.5±0.5 dB	Ave: -3.8 dB, Min: -4.09 dB
At 5 GT/s operation, the receiver must tolerate AC common mode voltage of 300 mV (peak-to-peak) and must tolerate 78.1 ps jitter.	Simultaneous worst case AC common mode voltage and worst case jitter during 5 GT/s operation may result in intermittent failures leading to subsequent recovery events.
TTX-UPW-TJ (uncorrelated total pulse width jitter) maximum of 24 ps.	Samples have measured as high as 25 ps.
The Transmitter PLL bandwidth and peaking for PCIe* at 5 GT/s is either 8 to 16 MHz with 3 dB of peaking or 5 to 16 MHz with 1 dB of peaking.	Samples have measured 7.8-16 MHz with 1.3 dB of peaking.
During the LTSSM Receiver Detect State, common-mode resistance to ground is 40 to 60 ohms.	Samples have measured up to 100 ohms.
8 GT/s Receiver Stressed Eye	Samples marginally pass or fail the 10-12 BER target under stressed eye conditions.
8 GT/s PLL Bandwidth: 2 to 4 MHz with 2 dB peaking.	Samples have a measured bandwidth of up to 4.1 MHz

Implication: Occasional RxDetect failure has been observed on a PCIe* Gen2 device caused by the Intel® Xeon® processor E5 v2 product families and Intel® Xeon® processor E7 2800/4800/8800 v2 product families violation of the PCIe specification max limit of 60 ohms common mode resistance to ground.

Workaround: Change the setting of the following register bits from "0x1" to "0x2":

Bus: 0x0
Offset: 0x648
Device: 0x6
Function: 0x7
Bits: 10:9

Status: For the affected steppings, see the Summary Tables of Changes.

CA22 Warm Resets May be Converted To Power-On Resets When Recovering From an IERR.

Problem: When a warm reset is attempted and an IERR (Internal Error) happens as indicated by the IA32_MCi_STATUS.MCOD of 0000_0100_0000_0000, a power-on reset occurs instead.

Implication: The values in the machine check bank will be lost as a result of the power-on reset. This prevents a OS, BIOS, or the BMC (Baseboard Management Controller) from logging the content of the error registers or taking any post-reset actions that are dependent on the machine check information.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA23 Patrol Scrubbing During Memory Mirroring May Improperly Signal Uncorrectable Machine Checks.

Problem: With memory mirroring enabled, Patrol Scrub detection of an uncorrectable error on one channel of the mirror should be downgraded to a correctable error when valid data is present on the other channel of the mirror. Due to this erratum, patrol Scrub detection of an uncorrectable error always signals an uncorrectable Machine Check.

Implication: This erratum may cause reduced availability of systems with mirrored memory.

Workaround: It is possible for BIOS to contain processor configuration data and code changes as a workaround for this erratum. Refer to Intel® Xeon® Processor E5 Product Family-based Platform CPU/Intel® QPI/Memory Reference Code version 0.8.312 or later and release notes.

Status: For the affected steppings, see the Summary Tables of Changes.

CA24 A Modification To The Multiple Message Enable Field Does Not Affect The AER Interrupt Message Number Field.

Problem: The (Advanced Error Interrupt) Message Number field (RPERRSTS Devices 0-3; Functions 0-3; Offset 178H; bits[31:27]) should be updated when the number of messages allocated to the root port is changed by writing the Multiple Message Enable field (MSIMSGCTL Device 3; Function 0; Offset 62H; bits[6:4]). However, writing the Multiple Message Enable in the root port does not update the Advanced Error Interrupt Message Number field.

Implication: Due to this erratum, software can allocate only one MSI (Message Signaled Interrupt) to the root port.

Workaround: None identified

Status: For the affected steppings, see the Summary Tables of Changes.

CA25 Long latency Transactions Can Cause I/O Devices On The Same Link to Time Out.

Problem: Certain long latency transactions - for example, master aborts on inbound traffic, locked transactions, peer-to-peer transactions, or vendor defined messages - conveyed over the PCIe* and DMI2 interfaces can block the progress of subsequent transactions for extended periods. In certain cases, these delays may lead to I/O device timeout that can result in device error reports and/or device off-lining.

Implication: Due to this erratum, devices that generate PCIe* or DMI2 traffic characterized by long latencies can interfere with other traffic types on the same link. This may result in reduced I/O performance and device timeout errors. USB traffic can be particularly sensitive to these delays.

Workaround: Avoid the contributing conditions. This can be accomplished by separating traffic types to be conveyed on different links and/or reducing or eliminating long latency transactions.

Status: For the affected steppings, see the Summary Tables of Changes.

CA26 Intel® QPI Link Layer Does Not Drop Unsupported or Undefined Packets.

Problem: The Intel® QPI should detect an unsupported or undefined packet, drop the offending packet, and log a correctable error with an IA32_MCi_STATUS.MCOD of 0000_1100_0000_1111. When the Intel QPI detects an unsupported or undefined packet it does not drop the offending packet but it does log the error.

Implication: Due to this erratum, Intel QPI does not drop unsupported packets. Intel has not observed any functional failure on commercially available systems due to this erratum.

Workaround: None

Status: For the affected steppings, see the Summary Tables of Changes.

CA27 Coherent Interface Write Cache May Report False Correctable ECC Errors During Cold Reset.

Problem: The Integrated I/O's coherent interface write cache includes ECC logic to detect errors. Due to this erratum, the write cache can report false ECC errors. This error is signaled by asserting bit 1 (Write Cache Corrected ECC) in the IRPP0ERRST CSR (Bus 0; Device 5; Function 2; Offset 230H) or the IRPP1ERRST CSR (Bus 0; Device 5; Function 2; Offset 2B0H).

Implication: If the coherent interface write cache ECC is enabled, the processor may incorrectly indicate correctable ECC errors in the write cache.

Workaround: A BIOS workaround has been identified. Refer to Intel® Xeon® Processor E5 Product Family-based platform CPU/Intel® QPI/Memory Reference Code version 1.0.006 or later and release notes.

Status: For the affected steppings, see the Summary Tables of Changes.

CA28 Combining ROL Transactions With Non-ROL Transactions or Marker Skipping Operations May Result in a System Hang.

Problem: When Intel® QuickData Technology DMA ROL (Raid On Load) transactions and non-ROL transactions are simultaneously active, and the non-ROL address offsets are not cacheline boundary aligned, the non-ROL transaction's last partial cacheline data write may be lost leading to a system hang. In addition, when Intel® QuickData Technology DMA ROL transactions are active, marker skipping operations may lead to a system hang.

Implication: When this erratum occurs, the processor may live lock resulting in a system hang.

Workaround: None identified. When ROL transactions and non-ROL transactions are simultaneously active, all non-ROL address offsets must be aligned on cacheline boundaries. Further, marker skipping operations may not be used on any DMA channel when ROL transactions are active.

Status: For the affected steppings, see the Summary Tables of Changes.

CA29 Excessive DRAM RAPL Power Throttling May Lead to a System Hang or USB Device Offlining.

Problem: DRAM RAPL (Running Average Power Limit) is a facility for limiting the maximum power consumption of the memory subsystem. DRAM RAPL's control mechanism constrains the number of memory transactions during a particular time period. Due to this erratum, a very low power limit can throttle certain memory subsystem configurations to an extent that system failure, ranging from permanent loss of USB devices to system hangs, may result.

Implication: Using DRAM RAPL to regulate the memory subsystem power to a very low level may cause platform instability.

Workaround: It is possible for the BIOS to contain processor configuration data and code changes as a workaround for this erratum. Refer to Intel® Xeon® Processor E5 Product Family-based Platform CPU/Intel® QPI/Memory Reference Code version 1.0.013 or later and release notes.

Status: For the affected steppings, see the Summary Tables of Changes.

CA30 TSOD-Related SMBus Transactions may not Complete When Package C-States are Enabled.

Problem: The processor may not complete SMBus (System Management Bus) transactions targeting the TSOD (Temperature Sensor On DIMM) when Package C-States are enabled. Due to this erratum, if the processor transitions into a Package C-State while an SMBus transaction with the TSOD is in process, the processor will suspend receipt of the transaction. The transaction completes while the processor is in a Package C-State. Upon exiting Package C-State, the processor will attempt to resume the SMBus transaction, detect a protocol violation, and log an error.

Implication: When Package C-States are enabled, the SMBus communication error rate between the processor and the TSOD may be higher than expected.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA31 The Integrated Memory Controller does not Enforce CKE High For tXSDLL DCLKs After Self-Refresh.

Problem: The JEDEC STANDARD DDR3 SDRAM Specification (No. 79-3E) requires that the CKE signal be held high for tXSDLL DCLKs after exiting self-refresh before issuing commands that require a locked DLL (Delay-Locked Loop). Due to this erratum, the Integrated Memory Controller may not meet this requirement with 512 Mb, 1 Gb, and 2 Gb devices in single rank per channel configurations.

Implication: Violating tXSDLL may result in DIMM clocking issues and may lead to unpredictable system behavior.

Workaround: A BIOS workaround has been identified. Refer to the Intel® Xeon® Processor E5 Product Family-based platform CPU/Intel® QPI/Memory Reference Code (RC), version 0.8.0 or later.

Status: For the affected steppings, see the Summary Tables of Changes.

CA32 Intel® QuickData Technology DMA Suspend does not Transition From ARMED to HALT State.

Problem: Suspending an Intel® QuickData Technology DMA channel while in the ARMED state should transition the channel to the HALT state. Due to this erratum, suspending a DMA channel while in the ARMED state does not change the state to HALT and will cause the DMA engine, when subsequently activated, to ignore the first descriptor's fence control bit and may cause the DMA engine to prematurely discard the first descriptor during the copy stage.

Implication: Suspending a DMA channel while in the ARMED state will cause the DMA engine to ignore descriptor fencing, possibly issue completion status without actually completing all descriptors, and may be subject to unexpected activation of DMA transfers.

Workaround: Check the DMA_trans_state (CHANSTS_0; Bus 0; MMIO BAR: CB_BAR [0:7]; Offset 88H; bits[2:0]) to ensure the channel state is either IDLE (001b) or ACTIVE (000b) before setting Susp_DMA (CHANCMD; Bus 0; MMIO BAR: CB_BAR [0:7]; Offset 84H;bit 2).

Status: For the affected steppings, see the Summary Tables of Changes.

CA33 Routing Intel® High Definition Audio Traffic Through VC1 May Result in System Hang.

Problem: When bit 9 in the IIOMISCCTRL CSR (Bus 0; Device 5; Function 0; Offset 1C0H) is set, VCp inbound traffic (Intel® HD Audio) is routed through VC1 to optimize isochronous traffic performance. Due to this erratum, VC1 may not have sufficient bandwidth for all traffic routed through it; overflows may occur.

Implication: This erratum can result in lost completions that may cause a system hang.

Workaround: A BIOS workaround has been identified. Refer to the Intel® Xeon® Processor E5 Product Family-based Platform CPU/QPI/Memory Reference Code version 1.0.006 or later and release notes.

Status: For the affected steppings, see the Summary Tables of Changes.

CA34 NTB Operating in NTB/RP Mode with MSI/MSI-X Interrupts May Cause System Hang.

Problem: The NTB (nontransparent bridge) operating in NTB/RP (NTB to Root Port mode) using Message Signaled Interrupts (MSI or MSI-X) in the presence of locks may result in a system hang.

Implication: Due to this erratum, system may hang under the condition described above.

Workaround: A BIOS workaround has been identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA35 Patrol Scrubbing does not Skip Ranks Disabled After DDR Training.

Problem: If a rank is detected as failed after completing DDR training then BIOS will mark it as disabled. Disabled ranks are omitted from the OS memory map. Due to this erratum, a rank disabled after DDR training completes is not skipped by the Patrol Scrubber. Patrol Scrubbing of the disabled ranks may result in superfluous correctable and uncorrectable memory error reports.

Implication: Disabling ranks after DDR training may result in the over-reporting of memory errors.

Workaround: A BIOS workaround has been identified. Refer to the Intel® Xeon® Processor E5 Product Family-based platform CPU/Intel QPI/Memory Reference Code version 1.0.013 or later and release notes.

Status: For the affected steppings, see the Summary Tables of Changes.

CA36 Writes to SDOORBELL or B2BDOORBELL in Conjunction With Inbound Access to NTB MMIO Space May Hang System.

Problem: A posted write targeting the SDOORBELL (Offset 64H) or B2BDOORBELL (Offset 140H) MMIO registers in the region define by Base Address Register PB01BASE (Bus 0; Device 3; Function 0: Offset 10H) or SB01BASE (Bus M; Device 0; Function 0; Offset 10H) may hang the system. This system hang may occur if the NTB (Non-Transparent Bridge) is processing a transaction from the secondary side of the NTB that is targeting the NTB shared MMIO registers or targeting the secondary side configuration registers when the write arrives.

Implication: The system may hang if the processor writes to the local SDOORBELL or B2BDOORBELL register at the same time that the NTB is processing an inbound transaction.

Workaround: In NTB/NTB (back-to-back) mode, do not use the B2BDOORBELL to send interrupts from the local to remote host. Instead, configure one of the following local register pairs to point to the remote SB01BASE region:

- PB23BASE (Device: 3; Function: 0; Offset: 18H) and PBAR2XLAT (Offset 10H) from PB01BASE or SB01BASE regions;
- PB45BASE (Device: 3; Function: 0; Offset: 20H) and PBAR4XLAT (Offset 18H) from PB01BASE, or SB01BASE regions;

The local host may then write directly to the PDOORBELL (Offset 60H) from the PB23BASE/PB45BASE region defined above.

In NTB/RP (bridge to root port) mode, the SDOORBELL register cannot be used by the processor on the primary side of the NTB to interrupt the processor on the secondary side. Instead, dedicate a BAR and XLAT pair, either PB23BASE/PBAR2XLAT or PB45BASE/PBAR4XLAT, to generate an interrupt directed directly into the MSI/MSIx (Message Signaled Interrupt) interrupt range on the remote processor. The device driver or client on the remote host must point the appropriate PBARnXLAT register to its MSI/MSIx interrupt range. The processor on the primary side can then write the MSI/MSIx interrupt to the dedicated BAR which will be translated by the NTB to the MSI/MSIx region of the secondary side's processor.

Status: For the affected steppings, see the Summary Tables of Changes.

CA37 DR6.B0-B3 May Not Report All Breakpoints Matched When a MOV/POP SS is Followed by a REP MOVSB or STOSB

Problem: Normally, data breakpoints matches that occur on a MOV SS, r/m or POP SS will not cause a debug exception immediately after MOV/POP SS but will be delayed until the instruction boundary following the next instruction is reached. After the debug exception occurs, DR6.B0-B3 bits will contain information about data breakpoints matched during the MOV/POP SS as well as breakpoints detected by the following instruction. Due to this erratum, DR6.B0-B3 bits may not contain information about data breakpoints matched during the MOV/POP SS when the following instruction is either an REP MOVSB or REP STOSB.

Implication: When this erratum occurs, DR6 may not contain information about all breakpoints matched. This erratum will not be observed under the recommended usage of the MOV SS,r/m or POP SS instructions (that is, following them only with an instruction that writes (E/R)SP).

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA38 64-bit REP MOVSB/STOSB May Clear The Upper 32-bits of RCX, RDI And RSI Before Any Data is Transferred

Problem: If a REP MOVSB/STOSB is executed in 64-bit mode with an address size of 32 bits, and if an interrupt is being recognized at the start of the instruction operation, the upper 32-bits of RCX, RDI and RSI may be cleared, even though no data has yet been copied or written.

Implication: Due to this erratum, the upper 32-bits of RCX, RDI and RSI may be prematurely cleared.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA39 An Interrupt Recognized Prior to First Iteration of REP MOVSB/STOSB May Result EFLAGS.RF Being Incorrectly Set

Problem: If a REP MOVSB/STOSB is executed and an interrupt is recognized prior to completion of the first iteration of the string operation, EFLAGS may be saved with RF=1 even though no data has been copied or stored. The Software Developer's Manual states that RF will be set to 1 for such interrupt conditions only after the first iteration is complete.

Implication: Software may not operate correctly if it relies on the value saved for EFLAGS.RF when an interrupt is recognized prior to the first iteration of a string instruction. Debug exceptions due to instruction breakpoints are delivered correctly despite this erratum; this is because the erratum occurs only after the processor has evaluated instruction-breakpoint conditions.

Workaround: Software whose correctness depends on value saved for EFLAGS.RF by delivery of the affected interrupts can disable fast-string operation by clearing Fast-String Enable in bit 0 in the IA32_MISC_ENABLE MSR (1A0H).

Status: For the affected steppings, see the Summary Tables of Changes.

CA40 Instructions Retired Event May Over Count Execution of IRET Instructions

Problem: Under certain conditions, the performance monitoring event Instructions Retired (Event C0H, Unmask 00H) may over count the execution of IRET instruction.

Implication: Due to this erratum, performance monitoring event Instructions Retired may over count.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA41 An Event May Intervene Before a System Management Interrupt That Results from IN or INS

Problem: If an I/O instruction (IN, INS, OUT, or OUTS) results in an SMI (system-management interrupt), the processor will set the IO_SMI bit at offset 7FA4H in SMRAM. This interrupt should be delivered immediately after execution of the I/O instruction so that the software handling the SMI can cause the I/O instruction to be re-executed. Due to this erratum, it is possible for another event (for example. a non maskable interrupt) to be delivered before the SMI that follows the execution of an IN or INS instruction.

Implication: If software handling an affected SMI uses I/O instruction restart, the handler for the intervening event will not be executed.

Workaround: The SMM handler has to evaluate the saved context to determine if the SMI was triggered by an instruction that read from an I/O port. The SMM handler must not restart an I/O instruction if the platform has not been configured to generate a synchronous SMI for the recorded I/O port address.

Status: For the affected steppings, see the Summary Tables of Changes.

CA42 Execution of VAESIMC or VAESKEYGENASSIST With An Illegal Value for VEX.vvvv May Produce a #NM Exception

Problem: The VAESIMC and VAESKEYGENASSIST instructions should produce a #UD (Invalid-Opcode) exception if the value of the vvvv field in the VEX prefix is not 1111b. Due to this erratum, if CR0.TS is "1", the processor may instead produce a #NM (Device-Not-Available) exception.

Implication: Due to this erratum, some undefined instruction encodings may produce a #NM instead of a #UD exception.

Workaround: Software should always set the vvvv field of the VEX prefix to 1111b for instances of the VAESIMC and VAESKEYGENASSIST instructions.

Status: For the affected steppings, see the Summary Tables of Changes.

CA43 Unexpected #UD on VZEROALL/VZEROUPPER

Problem: Execution of the VZEROALL or VZEROUPPER instructions in 64-bit mode with VEX.W set to 1 may erroneously cause a #UD (invalid-opcode exception).

Implication: The affected instructions may produce unexpected invalid-opcode exceptions in 64-bit mode.

Workaround: Compilers should encode VEX.W = 0 for the VZEROALL and VZEROUPPER instructions.

Status: For the affected steppings, see the Summary Tables of Changes.

CA44 Successive Fixed Counter Overflows May be Discarded

Problem: Under specific internal conditions, when using Freeze PerfMon on PMI feature (bit 12 in IA32_DEBUGCTL.Freeze_PerfMon_on_PMI, MSR 1D9H), if two or more PerfMon Fixed Counters overflow very closely to each other, the overflow may be mishandled for some of them. This means that the counter's overflow status bit (in MSR_PERF_GLOBAL_STATUS, MSR 38EH) may not be updated properly; additionally, PMI interrupt may be missed if software programs a counter in Sampling-Mode (PMI bit is set on counter configuration).

Implication: Successive Fixed Counter overflows may be discarded when Freeze PerfMon on PMI is used.

Workaround: Software can avoid this by:

1. Avoid using Freeze PerfMon on PMI bit
2. Enable only one fixed counter at a time when using Freeze PerfMon on PMI

Status: For the affected steppings, see the Summary Tables of Changes.

CA45 Execution of FXSAVE or FXRSTOR With the VEX Prefix May Produce a #NM Exception

Problem: Attempt to use FXSAVE or FXRSTOR with a VEX prefix should produce a #UD (Invalid-Opcode) exception. If either the TS or EM flag bits in CR0 are set, a #NM (device-not-available) exception will be raised instead of #UD exception.

Implication: Due to this erratum a #NM exception may be signaled instead of a #UD exception on an FXSAVE or an FXRSTOR with a VEX prefix.

Workaround: Software should not use FXSAVE or FXRSTOR with the VEX prefix.

Status: For the affected steppings, see the Summary Tables of Changes.

CA46 VM Exits Due to "NMI-Window Exiting" May Not Occur Following a VM Entry to the Shutdown State

Problem: If VM entry is made with the "virtual NMIs" and "NMI-window exiting," VM-execution controls set to 1, and if there is no virtual-NMI blocking after VM entry, a VM exit with exit reason "NMI window" should occur immediately after VM entry unless the VM entry put the logical processor in the wait-for SIPI state. Due to this erratum, such VM exits do not occur if the VM entry put the processor in the shutdown state.

Implication: A VMM may fail to deliver a virtual NMI to a virtual machine in the shutdown state.

Workaround: Before performing a VM entry to the shutdown state, software should check whether the "virtual NMIs" and "NMI-window exiting" VM-execution controls are both 1. If they are, software should clear "NMI-window exiting" and inject an NMI as part of VM entry.

Status: For the affected steppings, see the Summary Tables of Changes.

CA47 Execution of INVVPID Outside 64-Bit Mode Cannot Invalidate Translations For 64-Bit Linear Addresses

Problem: Executions of the INVVPID instruction outside 64-bit mode with the INVVPID type "individual-address invalidation" ignore bits 63:32 of the linear address in the INVVPID descriptor and invalidate translations for bits 31:0 of the linear address.

Implication: The INVVPID instruction may fail to invalidate translations for linear addresses that set bits in the range 63:32. Because this erratum applies only to executions outside 64-bit mode, it applies only to attempts by a 32-bit virtual-machine monitor (VMM) to invalidate translations for a 64-bit guest. Intel has not observed this erratum with any commercially available software.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA48 REP MOVSB May Incorrectly Update ECX, ESI, and EDI

Problem: Under certain conditions, if the execution of a REP MOVSB instruction is interrupted, the values of ECX, ESI and EDI may contain values that represent a later point in the execution of the instruction than the actual interruption point.

Implication: Due to this erratum ECX, ESI, and EDI may be incorrectly advanced, resulting in unpredictable system behavior.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA49 Performance-Counter Overflow Indication May Cause Undesired Behavior

Problem: Under certain conditions (listed below) when a performance counter overflows, its overflow indication may remain set indefinitely. This erratum affects the general-purpose performance counters IA32_PMC{0-7} and the fixed-function performance counters IA32_FIXED_CTR{0-2}. The erratum may occur if any of the following conditions are applied concurrent to when an actual counter overflow condition is reached:

1. Software disables the counter either globally through the IA32_PERF_GLOBAL_CTRL MSR (38FH), or locally through the IA32_PERFEVTSEL{0-7} MSRs (186H-18DH), or the IA32_FIXED_CTR_CTRL MSR (38DH).
2. Software sets the IA32_DEBUGCTL MSR (1D9H) FREEZE_PERFMON_ON_PMI bit [12].
3. The processor attempts to disable the counters by updating the state of the IA32_PERF_GLOBAL_CTRL MSR (38FH) as part of transitions such as VM exit, VM entry, SMI, RSM, or processor C-state.

Implication: Due to this erratum, the corresponding overflow status bit in IA32_PERF_GLOBAL_STATUS MSR (38DH) for an affected counter may not get cleared when expected. If a corresponding counter is configured to issue a PMI (performance monitor interrupt), multiple PMIs may be signaled from the same overflow condition. Likewise, if a corresponding counter is configured in PEBS mode (applies to only the general purpose counters), multiple PEBS events may be signaled.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA50 VEX.L is not Ignored with VCVT*2SI Instructions

Problem: The VEX.L bit should be ignored for the VCVTSS2SI, VCVTSD2SI, VCVTTSS2SI, and VCVTTSD2SI instructions, however due to this erratum the VEX.L bit is not ignored and will cause a #UD.

Implication: Unexpected #UDs will be seen when the VEX.L bit is set to 1 with VCVTSS2SI, VCVTSD2SI, VCVTTSS2SI, and VCVTTSD2SI instructions.

Workaround: Software should ensure that the VEX.L bit is set to 0 for all scalar instructions.

Status: For the affected steppings, see the Summary Tables of Changes.

CA51 Concurrently Changing the Memory Type and Page Size May Lead to a System Hang

Problem: Under a complex set of microarchitectural conditions, the system may hang if software changes the memory type and page size used to translate a linear address while a TLB (Translation Lookaside Buffer) holds a valid translation for that linear address.

Implication: Due to this erratum, the system may hang. Intel has not observed this erratum with any commercially available software.

Workaround: None identified. Please refer to Software Developer's Manual, volume 3, section "Recommended Invalidation" for the proper procedure for concurrently changing page attributes and page size.

Status: For the affected steppings, see the Summary Tables of Changes.

CA52 MCI_ADDR May be Incorrect For Cache Parity Errors

Problem: In cases when a WBINVD instruction evicts a line containing an address or data parity error (MCOD of 0x124, and MSCOD of 0x10), the address of this error should be logged in the MCi_ADDR register. Due to this erratum, the logged address may be incorrect, even though MCi_Status.ADDRV (bit 63) is set.

Implication: The address reported in MCi_ADDR may not be correct for cases of a parity error found during WBINVD execution.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA53 Instruction Fetches Page-Table Walks May be Made Speculatively to Uncacheable Memory

Problem: Page-table walks on behalf of instruction fetches may be made speculatively to uncacheable (UC) memory.

Implication: If any paging structures are located at addresses in uncacheable memory that are used for memory-mapped I/O, such I/O operations may be invoked as a result of speculative execution that would never actually occur in the executed code path. Intel has not observed this erratum with any commercially available software.

Workaround: Software should avoid locating paging structures at addresses in uncacheable memory that are used for memory-mapped I/O.

Status: For the affected steppings, see the Summary Tables of Changes.

CA54 REP MOVS/STOS Executing with Fast Strings Enabled and Crossing Page Boundaries with Inconsistent Memory Types may use an Incorrect Data Size or Lead to Memory-Ordering Violations

Problem: Under certain conditions as described in the Software Developers Manual section "Out-of-Order Stores For String Operations in Pentium 4, Intel Xeon, and P6 Family Processors" the processor performs REP MOVS or REP STOS as fast strings. Due to this erratum fast string REP MOVS/REP STOS instructions that cross page boundaries from WB/WC memory types to UC/WP/WT memory types, may start using an incorrect data size or may observe memory ordering violations.

Implication: Upon crossing the page boundary the following may occur, dependent on the new page memory type:

- UC the data size of each write will now always be 8 bytes, as opposed to the original data size.
- WP the data size of each write will now always be 8 bytes, as opposed to the original data size and there may be a memory ordering violation.
- WT there may be a memory ordering violation.

Workaround: Software should avoid crossing page boundaries from WB or WC memory type to UC, WP or WT memory type within a single REP MOVS or REP STOS instruction that will execute with fast strings enabled.

Status: For the affected steppings, see the Summary Tables of Changes.

CA55 The Processor May Not Properly Execute Code Modified Using A Floating-Point Store

Problem: Under complex internal conditions, a floating-point store used to modify the next sequential instruction may result in the old instruction being executed instead of the new instruction.

Implication: Self- or cross-modifying code may not execute as expected. Intel has not observed this erratum with any commercially available software.

Workaround: None identified. Do not use floating-point stores to modify code.

Status: For the affected steppings, see the Summary Tables of Changes.

CA56 VM Exits Due to GETSEC May Save an Incorrect Value for "Blocking by STI" in the Context of Probe-Mode Redirection

Problem: The GETSEC instruction causes a VM exit when executed in VMX non-root operation. Such a VM exit should set bit 0 in the interruptibility-state field in the virtual-machine control structure (VMCS) if the STI instruction was blocking interrupts at the time GETSEC commenced execution. Due to this erratum, a VM exit executed in VMX non-root operation may erroneously clear bit 0 if redirection to probe mode occurs on the GETSEC instruction.

Implication: After returning from probe mode, a virtual interrupt may be incorrectly delivered prior to GETSEC instruction. Intel has not observed this erratum with any commercially available software.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA57 IA32_MC5_CTL2 is Not Cleared by a Warm Reset

Problem: IA32_MC5_CTL2 MSR (285H) is documented to be cleared on any reset. Due to this erratum this MSR is only cleared upon a cold reset.

Implication: The algorithm documented in Software Developer's Manual, Volume 3, section titled "CMCI Initialization" or any other algorithm that counts the IA32_MC5_CTL2 MSR being cleared on reset will not function as expected after a warm reset.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA58 The Processor May Report a #TS Instead of a #GP Fault

Problem: A jump to a busy TSS (Task-State Segment) may cause a #TS (invalid TSS exception) instead of a #GP fault (general protection exception).

Implication: Operation systems that access a busy TSS may get invalid TSS fault instead of a #GP fault. Intel has not observed this erratum with any commercially available software.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA59 IO_SMI Indication in SMRAM State Save Area May be Set Incorrectly

Problem: he IO_SMI bit in SMRAM's location 7FA4H is set to "1" by the CPU to indicate a System Management Interrupt (SMI) occurred as the result of executing an instruction that reads from an I/O port. Due to this erratum, the IO_SMI bit may be incorrectly set by:

- A non-I/O instruction
- SMI is pending while a lower priority event interrupts

- A REP I/O read
- A I/O read that redirects to MWAIT

Implication: SMM handlers may get false IO_SMI indication.

Workaround: The SMM handler has to evaluate the saved context to determine if the SMI was triggered by an instruction that read from an I/O port. The SMM handler must not restart an I/O instruction if the platform has not been configured to generate a synchronous SMI for the recorded I/O port address.

Status: or the affected steppings, see the Summary Tables of Changes.

CA60 Performance Monitor SSE Retired Instructions May Return Incorrect Values

Problem: Performance Monitoring counter SIMD_INST_RETIRED (Event: C7H) is used to track retired SSE instructions. Due to this erratum, the processor may also count other types of instructions resulting in higher than expected values.

Implication: Performance Monitoring counter SIMD_INST_RETIRED may report count higher than expected.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA61 IRET under Certain Conditions May Cause an Unexpected Alignment Check Exception

Problem: In IA-32e mode, it is possible to get an Alignment Check Exception (#AC) on the IRET instruction even though alignment checks were disabled at the start of the IRET. This can only occur if the IRET instruction is returning from CPL3 code to CPL3 code. IRETs from CPL0/1/2 are not affected. This erratum can occur if the EFLAGS value on the stack has the AC flag set, and the interrupt handler's stack is misaligned. In IA-32e mode, RSP is aligned to a 16-byte boundary before pushing the stack frame.

Implication: In IA-32e mode, under the conditions given above, an IRET can get a #AC even if alignment checks are disabled at the start of the IRET. This erratum can only be observed with a software generated stack frame.

Workaround: Software should not generate misaligned stack frames for use with IRET.

Status: For the affected steppings, see the Summary Tables of Changes.

CA62 Performance Monitoring Event FP_MMX_TRANS_TO_MMX May Not Count Some Transitions

Problem: Performance Monitor Event FP_MMX_TRANS_TO_MMX (Event CCH, Umask 01H) counts transitions from x87 Floating Point (FP) to MMX™ instructions. Due to this erratum, if only a small number of MMX instructions (including EMMS) are executed immediately after the last FP instruction, a FP to MMX transition may not be counted.

Implication: The count value for Performance Monitoring Event FP_MMX_TRANS_TO_MMX may be lower than expected. The degree of undercounting is dependent on the occurrences of the erratum condition while the counter is active. Intel has not observed this erratum with any commercially available software.

Workaround: None Identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA63 General Protection Fault (#GP) for Instructions Greater than 15 Bytes May be Preempted

Problem: When the processor encounters an instruction that is greater than 15 bytes in length, a #GP is signaled when the instruction is decoded. Under some circumstances, the #GP fault may be preempted by another lower priority fault (for example, Page Fault (#PF)). However, if the preempting lower priority faults are resolved by the operating system and the instruction retried, a #GP fault will occur.

Implication: Software may observe a lower-priority fault occurring before or in lieu of a #GP fault. Instructions of greater than 15 bytes in length can only occur if redundant prefixes are placed before the instruction.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA64 LBR, BTS, BTM May Report a Wrong Address When an Exception/Interrupt Occurs in 64-bit Mode

Problem: An exception/interrupt event should be transparent to the LBR (Last Branch Record), BTS (Branch Trace Store) and BTM (Branch Trace Message) mechanisms. However, during a specific boundary condition where the exception/interrupt occurs right after the execution of an instruction at the lower canonical boundary (0x00007FFFFFFFFFFF) in 64-bit mode, the LBR return registers will save a wrong return address with bits 63 to 48 incorrectly sign extended to all 1's. Subsequent BTS and BTM operations which report the LBR will also be incorrect.

Implication: LBR, BTS and BTM may report incorrect information in the event of an exception/interrupt.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA65 Incorrect Address Computed For Last Byte of FXSAVE/FXRSTOR or XSAVE/XRSTOR Image Leads to Partial Memory Update

Problem: A partial memory state save of the FXSAVE or XSAVE image or a partial memory state restore of the FXRSTOR or XRSTOR image may occur if a memory address exceeds the 64KB limit while the processor is operating in 16-bit mode or if a memory address exceeds the 4 GB limit while the processor is operating in 32-bit mode.

Implication: FXSAVE/FXRSTOR or XSAVE/XRSTOR will incur a #GP fault due to the memory limit violation as expected but the memory state may be only partially saved or restored.

Workaround: Software should avoid memory accesses that wrap around the respective 16-bit and 32-bit mode memory limits.

Status: For the affected steppings, see the Summary Tables of Changes.

CA66 Values for LBR/BTS/BTM Will be Incorrect after an Exit from SMM

Problem: After a return from SMM (System Management Mode), the CPU will incorrectly update the LBR (Last Branch Record) and the BTS (Branch Trace Store), hence rendering their data invalid. The corresponding data if sent out as a BTM on the system bus will also be incorrect.

Note: This issue would only occur when one of the 3 above mentioned debug support facilities are used.

Implication: The value of the LBR, BTS, and BTM immediately after an RSM operation should not be used.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA67 EFLAGS Discrepancy on Page Faults and on EPT-Induced VM Exits after a Translation Change

Problem: This erratum is regarding the case where paging structures are modified to change a linear address from writable to non-writable without software performing an appropriate TLB invalidation. When a subsequent access to that address by a specific instruction (ADD, AND, BTC, BTR, BTS, CMPXCHG, DEC, INC, NEG, NOT, OR, ROL/ROR, SAL/SAR/SHL/SHR, SHLD, SHRD, SUB, XOR, and XADD) causes a page fault or an EPT-induced VM exit, the value saved for EFLAGS may incorrectly contain the arithmetic flag values that the EFLAGS register would have held had the instruction completed without

fault or VM exit. For page faults, this can occur even if the fault causes a VM exit or if its delivery causes a nested fault.

Implication: None identified. Although the EFLAGS value saved by an affected event (a page fault or an EPT-induced VM exit) may contain incorrect arithmetic flag values, Intel has not identified software that is affected by this erratum. This erratum will have no further effects once the original instruction is restarted because the instruction will produce the same results as if it had initially completed without fault or VM exit.

Workaround: If the handler of the affected events inspects the arithmetic portion of the saved EFLAGS value, then system software should perform a synchronized paging structure modification and TLB invalidation.

Status: For the affected steppings, see the Summary Tables of Changes.

CA68 B0-B3 Bits in DR6 For Non-Enabled Breakpoints May Be Incorrectly Set

Problem: Some of the B0-B3 bits (breakpoint conditions detect flags, bits [3:0]) in DR6 may be incorrectly set for non-enabled breakpoints when the following sequence happens:

1. MOV or POP instruction to SS (Stack Segment) selector;
2. Next instruction is FP (Floating Point) that gets FP assist
3. Another instruction after the FP instruction completes successfully
4. A breakpoint occurs due to either a data breakpoint on the preceding instruction or a code breakpoint on the next instruction.

Due to this erratum a non-enabled breakpoint triggered on step 1 or step 2 may be reported in B0-B3 after the breakpoint occurs in step 4.

Implication: Due to this erratum, B0-B3 bits in DR6 may be incorrectly set for non-enabled breakpoints.

Workaround: Software should not execute a floating point instruction directly after a MOV SS or POP SS instruction.

Status: For the affected steppings, see the Summary Tables of Changes.

CA69 MCi_Status Overflow Bit May Be Incorrectly Set on a Single Instance of a DTLB Error

Problem: A single Data Translation Look Aside Buffer (DTLB) error can incorrectly set the Overflow (bit [62]) in the MCi_Status register. A DTLB error is indicated by M error code (bits [15:0]) appearing as binary value, 000x 0000 0001 0100, in the MCi_Status register.

Implication: Due to this erratum, the Overflow bit in the MCi_Status register may not be an accurate indication of multiple occurrences of DTLB errors. There is no other impact to normal processor functionality.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA70 Debug Exception Flags DR6.B0-B3 Flags May Be Incorrect for Disabled Breakpoints

Problem: When a debug exception is signaled on a load that crosses cache lines with data forwarded from a store and whose corresponding breakpoint enable flags are disabled (DR7.G0-G3 and DR7.L0-L3), the DR6.B0-B3 flags may be incorrect.

Implication: The debug exception DR6.B0-B3 flags may be incorrect for the load if the corresponding breakpoint enable flag in DR7 is disabled.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA71 LER MSRs May Be Unreliable

Problem: Due to certain internal processor events, updates to the LER (Last Exception Record) MSRs, MSR_LER_FROM_LIP (1DDH) and MSR_LER_TO_LIP (1DEH), may happen when no update was expected.

Implication: The values of the LER MSRs may be unreliable.

Workaround: None Identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA72 Storage of PEBS Record Delayed Following Execution of MOV SS or STI

Problem: When a performance monitoring counter is configured for PEBS (Precise Event Based Sampling), overflow of the counter results in storage of a PEBS record in the PEBS buffer. The information in the PEBS record represents the state of the next instruction to be executed following the counter overflow. Due to this erratum, if the counter overflow occurs after execution of either MOV SS or STI, storage of the PEBS record is delayed by one instruction.

Implication: When this erratum occurs, software may observe storage of the PEBS record being delayed by one instruction following execution of MOV SS or STI. The state information in the PEBS record will also reflect the one instruction delay.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA73 PEBS Record Not Updated When in Probe Mode

Problem: When a performance monitoring counter is configured for PEBS (Precise Event Based Sampling), overflows of the counter can result in storage of a PEBS record in the PEBS buffer. Due to this erratum, if the overflow occurs during probe mode, it may be ignored and a new PEBS record may not be added to the PEBS buffer.

Implication: Due to this erratum, the PEBS buffer may not be updated by overflows that occur during probe mode.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA74 Faulting MMX Instruction May Incorrectly Update x87 FPU Tag Word

Problem: Under a specific set of conditions, MMX stores (MOVD, MOVQ, MOVNTQ, MASKMOVQ) which cause memory access faults (#GP, #SS, #PF, or #AC), may incorrectly update the x87 FPU tag word register.

Problem: This erratum will occur when the following additional conditions are also met.

- The MMX store instruction must be the first MMX instruction to operate on x87 FPU state (that is, the x87 FP tag word is not already set to 0x0000).
- For MOVD, MOVQ, MOVNTQ stores, the instruction must use an addressing mode that uses an index register (this condition does not apply to MASKMOVQ).

Implication: If the erratum conditions are met, the x87 FPU tag word register may be incorrectly set to a 0x0000 value when it should not have been modified.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA75 #GP on Segment Selector Descriptor That Straddles Canonical Boundary May Not Provide Correct Exception Error Code

Problem: During a #GP (General Protection Exception), the processor pushes an error code on to the exception handler's stack. If the segment selector descriptor straddles the canonical boundary, the error code pushed onto the stack may be incorrect.

Implication: An incorrect error code may be pushed onto the stack. Intel has not observed this erratum with any commercially available software.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA76 APIC Error "Received Illegal Vector" May Be Lost

Problem: APIC (Advanced Programmable Interrupt Controller) may not update the ESR (Error Status Register) flag Received Illegal Vector bit [6] properly when an illegal vector error is received on the same internal clock that the ESR is being written (as part of the write-read ESR access flow). The corresponding error interrupt will also not be generated for this case.

Implication: Due to this erratum, an incoming illegal vector error may not be logged into ESR properly and may not generate an error interrupt.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA77 Changing the Memory Type for an In-Use Page Translation May Lead to Memory-Ordering Violations

Problem: Under complex microarchitectural conditions, if software changes the memory type for data being actively used and shared by multiple threads without the use of semaphores or barriers, software may see load operations execute out of order.

Implication: Memory ordering may be violated. Intel has not observed this erratum with any commercially available software.

Workaround: Software should ensure pages are not being actively used before requesting their memory type be changed.

Status: For the affected steppings, see the Summary Tables of Changes.

CA78 Reported Memory Type May Not be Used to Access the VMCS and Referenced Data Structures

Problem: Bits 53:50 of the IA32_VMX_BASIC MSR report the memory type that the processor uses to access the VMCS and data structures referenced by pointers in the VMCS. Due to this erratum, a VMX access to the VMCS or referenced data structures will instead use the memory type that the MTRRs (memory-type range registers) specify for the physical address of the access.

Implication: Bits 53:50 of the IA32_VMX_BASIC MSR report that the WB (write-back) memory type will be used but the processor may use a different memory type.

Workaround: Software should ensure that the VMCS and referenced data structures are located at physical addresses that are mapped to WB memory type by the MTRRs.

Status: For the affected steppings, see the Summary Tables of Changes.

CA79 LBR, BTM or BTS Records May have Incorrect Branch From Information After an EIST / T-state / S-state / C1E Transition or Adaptive Thermal Throttling

Problem: The "From" address associated with the LBR (Last Branch Record), BTM (Branch Trace Message) or BTS (Branch Trace Store) may be incorrect for the first branch after a transition of:

- Enhanced Intel® SpeedStep Technology
- T-state (Thermal Monitor states)
- S1-state (ACPI package sleep state)
- C1E (Enhanced C1 Low Power state)
- Adaptive Thermal Throttling

Implication: When the LBRs, BTM or BTS are enabled, some records may have incorrect branch "From" addresses for the first branch after a transition of Enhanced Intel SpeedStep Technology, T-states, S-states, C1E, or Adaptive Thermal Throttling.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA80 FP Data Operand Pointer May Be Incorrectly Calculated After an FP Access Which Wraps a 4-Gbyte Boundary in Code That Uses 32-Bit Address Size in 64-bit Mode

Problem: The FP (Floating Point) Data Operand Pointer is the effective address of the operand associated with the last non-control FP instruction executed by the processor. If an 80-bit FP access (load or store) uses a 32-bit address size in 64-bit mode and the memory access wraps a 4-Gbyte boundary and the FP environment is subsequently saved, the value contained in the FP Data Operand Pointer may be incorrect.

Implication: Due to this erratum, the FP Data Operand Pointer may be incorrect. Wrapping an 80-bit FP load around a 4-Gbyte boundary in this way is not a normal programming practice. Intel has not observed this erratum with any commercially available software.

Workaround: If the FP Data Operand Pointer is used in a 64-bit operating system which may run code accessing 32-bit addresses, care must be taken to ensure that no 80-bit FP accesses are wrapped around a 4-Gbyte boundary

Status: For the affected steppings, see the Summary Tables of Changes.

CA81 VMREAD/VMWRITE Instruction May Not Fail When Accessing an Unsupported Field in VMCS

Problem: The Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 2B states that execution of VMREAD or VMWRITE should fail if the value of the instruction's register source operand corresponds to an unsupported field in the VMCS (Virtual Machine Control Structure). The correct operation is that the logical processor will set the ZF (Zero Flag), write 0CH into the VM-instruction error field and for VMREAD leave the instruction's destination operand unmodified. Due to this erratum, the instruction may instead clear the ZF, leave the VM-instruction error field unmodified and for VMREAD modify the contents of its destination operand.

Implication: Accessing an unsupported field in VMCS will fail to properly report an error. In addition, VMREAD from an unsupported VMCS field may unexpectedly change its destination operand. Intel has not observed this erratum with any commercially available software.

Workaround: Software should avoid accessing unsupported fields in a VMCS.

Status: For the affected steppings, see the Summary Tables of Changes.

CA82 An Unexpected PMI May Occur After Writing a Large Value to IA32_FIXED_CTR2

Problem: If the fixed-function performance counter IA32_FIXED_CTR2 MSR (30BH) is configured to generate a performance-monitor interrupt (PMI) on overflow and the counter's value is greater than FFFFFFFFFFC0H, then this erratum may incorrectly cause a PMI if software performs a write to this counter.

Implication: A PMI may be generated unexpectedly when programming IA32_FIXED_CTR2. Other than the PMI, the counter programming is not affected by this erratum as the attempted write operation does succeed.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA83 A Write to the IA32_FIXED_CTR1 MSR May Result in Incorrect Value in Certain Conditions

Problem: Under specific internal conditions, if software tries to write the IA32_FIXED_CTR1 MSR (30AH) a value that has all bits [31:1] set while the counter was just about to overflow when the write is attempted (i.e. its value was 0xFFFF FFFF FFFF), then due to this erratum the new value in the MSR may be corrupted.

Implication: Due to this erratum, IA32_FIXED_CTR1 MSR may be written with a corrupted value.

Workaround: Software may avoid this erratum by writing zeros to the IA32_FIXED_CTR1 MSR, before the desired write operation.

Status: For the affected steppings, see the Summary Tables of Changes.

CA84 #GP May be Signaled When Invalid VEX Prefix Precedes Conditional Branch Instructions

Problem: When a 2-byte opcode of a conditional branch (opcodes 0F8xH, for any value of x) instruction resides in 16-bit code-segment and is associated with invalid VEX prefix, it may sometimes signal a #GP fault (illegal instruction length > 15-bytes) instead of a #UD (illegal opcode) fault.

Implication: Due to this erratum, #GP fault instead of a #UD may be signaled on an illegal instruction.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA85 Interrupt From Local APIC Timer May Not Be Detectable While Being Delivered

Problem: If the local-APIC timer's CCR (current-count register) is 0, software should be able to determine whether a previously generated timer interrupt is being delivered by first reading the delivery-status bit in the LVT timer register and then reading the bit in the IRR (interrupt-request register) corresponding to the vector in the LVT timer register. If both values are read as 0, no timer interrupt should be in the process of being delivered. Due to this erratum, a timer interrupt may be delivered even if the CCR is 0 and the LVT and IRR bits are read as 0. This can occur only if the DCR (Divide Configuration Register) is greater than or equal to 4. The erratum does not occur if software writes zero to the Initial Count Register before reading the LVT and IRR bits.

Implication: Software that relies on reads of the LVT and IRR bits to determine whether a timer interrupt is being delivered may not operate properly.

Workaround: Software that uses the local-APIC timer must be prepared to handle the timer interrupts, even those that would not be expected based on reading CCR and the LVT and IRR bits; alternatively, software can avoid the problem by writing zero to the Initial Count Register before reading the LVT and IRR bits.

Status: For the affected steppings, see the Summary Tables of Changes.

CA86 PCMPESTRI, PCMPESTRM, VPCMPESTRI and VPCMPESTRM Always Operate with 32-bit Length Registers

Problem: In 64-bit mode, using REX.W=1 with PCMPESTRI and PCMPESTRM or VEX.W=1 with VPCMPESTRI and VPCMPESTRM should support a 64-bit length operation with RAX/RDX. Due to this erratum, the length registers are incorrectly interpreted as 32-bit values.

Implication: Due to this erratum, using REX.W=1 with PCMPESTRI and PCMPESTRM as well as VEX.W=1 with VPCMPESTRI and VPCMPESTRM do not result in promotion to 64-bit length registers.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA87 During Package Power States Repeated PCIe* and/or DMI L1 Transitions May Cause a System Hang

Problem: Under a complex set of internal conditions and operating temperature, when the processor is in a deep power state (package C3, C6 or C7) and the PCIe and/or DMI links are toggling in and out of L1 state, the system may hang.

Implication: Due to this erratum, the system may hang.

Workaround: None Identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA88 RDMSR of IA32_PERFEVTSEL4-7 May Return an Incorrect Result

Problem: When CPUID.A.EAX[15:8] reports 8 general-purpose performance monitoring counters per logical processor, RDMSR of IA32_PERFEVTSEL4-7 (MSR 18AH:18DH) may not return the same value as previously written.

Implication: Software should not rely on the value read from these MSRs. Writing these MSRs functions as expected.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA89 MONITOR or CLFLUSH on the Local XAPIC's Address Space Results in Hang

Problem: If the target linear address range for a MONITOR or CLFLUSH is mapped to the local xAPIC's address space, the processor will hang.

Implication: When this erratum occurs, the processor will hang. The local xAPIC's address space must be uncached. The MONITOR instruction only functions correctly if the specified linear address range is of the type write-back. CLFLUSH flushes data from the cache. Intel has not observed this erratum with any commercially available software.

Workaround: Do not execute MONITOR or CLFLUSH instructions on the local xAPIC address space.

Status: For the affected steppings, see the Summary Tables of Changes.

CA90 PCMPESTRI, PCMPESTRM, VPCMPESTRI and VPCMPESTRM Always Operate With 32-bit Length Registers

Problem: In 64-bit mode, using REX.W=1 with PCMPESTRI and PCMPESTRM or VEX.W=1 with VPCMPESTRI and VPCMPESTRM should support a 64-bit length operation with RAX/RDX. Due to this erratum, the length registers are incorrectly interpreted as 32-bit values.

Implication: Due to this erratum, using REX.W=1 with PCMPESTRI and PCMPESTRM as well as VEX.W=1 with VPCMPESTRI and VPCMPESTRM do not result in promotion to 64-bit length registers.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA91 Clock Modulation Duty Cycle Cannot Be Programmed to 6.25%

Problem: When programming field T_STATE_REQ of the IA32_CLOCK_MODULATION MSR (19AH) bits [3:0] to '0001, the actual clock modulation duty cycle will be 12.5% instead of the expected 6.25% ratio.

Implication: Due to this erratum, it is not possible to program the clock modulation to a 6.25% duty cycle.

Workaround: None Identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA92 Processor May Livelock During On Demand Clock Modulation

Problem: The processor may livelock when (1) a processor thread has enabled on demand clock modulation via bit 4 of the IA32_CLOCK_MODULATION MSR (19AH) and the clock modulation duty cycle is set to 12.5 % (02H in bits 3:0 of the same MSR), and (2) the other processor thread does not have on demand clock modulation enabled and that thread is executing a stream of instructions with the lock prefix that either split a cache line or access UC memory.

Implication: Program execution may stall on both threads of the core subject to this erratum.

Workaround: This erratum will not occur if clock modulation is enabled on all threads when using on demand clock modulation or if the duty cycle programmed in the IA32_CLOCK_MODULATION MSR is 18.75% or higher.

Status: For the affected steppings, see the Summary Tables of Changes.

CA93 Performance Monitor Counters May Produce Incorrect Results

Problem: When operating with SMT enabled, a memory at-retirement performance monitoring event (from the list below) may be dropped or may increment an enabled event on the corresponding counter with the same number on the physical core's other thread rather than the thread experiencing the event. Processors with SMT disabled in BIOS are not affected by this erratum.

The list of affected memory at-retirement events is as follows:

- MEM_UOP_RETIRED.LOADS
- MEM_UOP_RETIRED.STORES
- MEM_UOP_RETIRED.LOCK
- MEM_UOP_RETIRED.SPLIT
- MEM_UOP_RETIRED.STLB_MISS
- MEM_LOAD_UOPS_RETIRED.HIT_LFB
- MEM_LOAD_UOPS_RETIRED.L1_HIT
- MEM_LOAD_UOPS_RETIRED.L2_HIT
- MEM_LOAD_UOPS_RETIRED.LLC_HIT
- MEM_LOAD_UOPS_LLC_HIT_RETIRED.XSNP_HIT
- MEM_LOAD_UOPS_LLC_HIT_RETIRED.XSNP_HITM
- MEM_LOAD_UOPS_LLC_HIT_RETIRED.XSNP_MISS
- MEM_LOAD_UOPS_LLC_HIT_RETIRED.XSNP_NONE
- MEM_LOAD_UOPS_RETIRED.LLC_MISS
- MEM_LOAD_UOPS_LLC_MISS_RETIRED.LOCAL_DRAM
- MEM_LOAD_UOPS_LLC_MISS_RETIRED.REMOTE_DRAM
- MEM_LOAD_UOPS_RETIRED.L2_MISS

Implication: Due to this erratum, certain performance monitoring event may produce unreliable results when SMT is enabled.

Workaround: None Identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA94 Virtual-APIC Page Accesses With 32-Bit PAE Paging May Cause a System Crash

Problem: If a logical processor has EPT (Extended Page Tables) enabled, is using 32-bit PAE paging, and accesses the virtual-APIC page then a complex sequence of internal processor micro-architectural events may cause an incorrect address translation or machine check on either logical processor.

Implication: This erratum may result in unexpected faults, an uncorrectable TLB error logged in IA32_MCi_STATUS.MCACOD (bits [15:0]) with a value of 0000_0000_0001_xxxxb (where x stands for 0 or 1), a guest or hypervisor crash, or other unpredictable system behavior.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA95 IA32_FEATURE_CONTROL MSR May be Un-Initialized on a Cold Reset

Problem: IA32_FEATURE_CONTROL MSR (3Ah) may have random values after RESET (including the reserved and Lock bits), and the read-modify-write of the reserved bits and/or the Lock bit being incorrectly set may cause an unexpected GP fault.

Implication: Due to this erratum, an unexpected GP fault may occur and BIOS may not complete initialization.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA96 PEBS May Unexpectedly Signal a PMI After the PEBS Buffer is Full

Problem: The Software Developer's Manual states that no PMI should be generated when PEBS index reaches PEBS Absolute Maximum. Due to this erratum a PMI may be generated even though the PEBS buffer is full.

Implication: PEBS may trigger a PMI even though the PEBS index has reached the PEBS Absolute Maximum.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA97 Execution of GETSEC[SEXIT] May Cause a Debug Exception to Be Lost

Problem: A debug exception occurring at the same time that GETSEC[SEXIT] is executed or when an SEXIT doorbell event is serviced may be lost.

Implication: Due to this erratum, there may be a loss of a debug exception when it happens concurrently with the execution of GETSEC[SEXIT]. Intel has not observed this erratum with any commercially available software.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA98 An Uncorrectable Error Logged in IA32_CR_MC2_STATUS May also Result in a System Hang

Problem: Uncorrectable errors logged in IA32_CR_MC2_STATUS MSR (409H) may also result in a system hang causing an Internal Timer Error (MCACOD = 0x0400h) to be logged in another machine check bank (IA32_MCi_STATUS).

Implication: Uncorrectable errors logged in IA32_CR_MC2_STATUS can further cause a system hang and an Internal Timer Error to be logged.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA99 The Corrected Error Count Overflow Bit in IA32_ MC0_STATUS is Not Updated After a UC Error is Logged

Problem: When a UC (uncorrected) error is logged in the IA32_MC0_STATUS MSR (401H), corrected errors will continue to update the lower 14 bits (bits 51:38) of the Corrected Error Count. Due to this erratum, the sticky count overflow bit (bit 52) of the Corrected Error Count will not get updated after a UC error is logged.

Implication: The Corrected Error Count Overflow indication will be lost if the overflow occurs after an uncorrectable error has been logged.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA100 IA32_VMX_VMCS_ENUM MSR (48AH) Does Not Properly Report the Highest Index Value Used for VMCS Encoding

Problem: IA32_VMX_VMCS_ENUM MSR (48AH) bits 9:1 report the highest index value used for any VMCS encoding. Due to this erratum, the value 21 is returned in bits 9:1 although there is a VMCS field whose encoding uses the index value 23.

Implication: Software that uses the value reported in IA32_VMX_VMCS_ENUM[9:1] to read and write all VMCS fields may omit one field.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA101 The Upper 32 Bits of CR3 May be Incorrectly Used With 32-Bit Paging

Problem: When 32-bit paging is in use, the processor should use a page directory located at the 32-bit physical address specified in bits 31:12 of CR3; the upper 32 bits of CR3 should be ignored. Due to this erratum, the processor will use a page directory located at the 64-bit physical address specified in bits 63:12 of CR3.

Implication: The processor may use an unexpected page directory or, if EPT (Extended Page Tables) is in use, cause an unexpected EPT violation. This erratum applies only if software enters 64-bit mode, loads CR3 with a 64-bit value, and then returns to 32-bit paging without changing CR3. Intel has not observed this erratum with any commercially available software.

Workaround: Software that has executed in 64-bit mode should reload CR3 with a 32-bit value before returning to 32-bit paging.

Status: For the affected steppings, see the Summary Tables of Changes.

CA102 EPT Violations May Report Bits 11:0 of Guest Linear Address Incorrectly

Problem: If a memory access to a linear address requires the processor to update an accessed or dirty flag in a paging-structure entry and if that update causes an EPT violation, the processor should store the linear address into the "guest linear address" field in the VMCS. Due to this erratum, the processor may store an incorrect value into bits 11:0 of this field. (The processor correctly stores the guest-physical address of the paging-structure entry into the "guest-physical address" field in the VMCS.)

Implication: Software may not be easily able to determine the page offset of the original memory access that caused the EPT violation. Intel has not observed this erratum to impact the operation of any commercially available software.

Workaround: Software requiring the page offset of the original memory access address can derive it by simulating the effective address computation of the instruction that caused the EPT violation.

Status: For the affected steppings, see the Summary Tables of Changes.

CA103 Intel® QuickData Technology DMA Access to Invalid Memory Address May Cause System Hang

Problem: When an Intel QuickData Technology DMA access request references an invalid memory address, the channel generating the request may fail to abort the invalid address access and cause all channels to hang.

Implication: An Intel QuickData Technology DMA access to an invalid memory address may cause all channels to hang.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA104 CPUID Faulting is Not Enumerated Properly

Problem: A processor that supports the CPUID-faulting feature enumerates this capability by setting PLATFORM_INFO MSR (CEH) bit 31. Due to this erratum, the processor erroneously clears this bit.

Implication: Software that depends upon CPUID faulting will incorrectly determine that the processor does not support the feature.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA105 TSC is Not Affected by Warm Reset

Problem: The TSC (Time Stamp Counter MSR 10H) should be cleared on reset. Due to this erratum the TSC is not affected by warm reset.

Implication: The TSC is not cleared by a warm reset. The TSC is cleared by power-on reset as expected. Intel has not observed any functional failures due to this erratum.

Workaround: None Identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA106 Incorrect Size Reported by PCIe* NTB BAR Registers

Problem: The PCIe NTB (Non-Transparent Bridge) BARs (Base Address Register) at PB23BASE (Bus 0; Device 3; Function 0; Offset 0x18) and PB45BASE ((Bus 0; Device 3; Function 0; Offset 0x20) are used to determine the size of the requested memory region. Due to this erratum, these registers return incorrect values.

Implication: When this erratum occurs, incorrect memory region size(s) can lead to overlapping BAR regions.

Workaround: A BIOS workaround has been identified. Please refer to Intel® Xeon® Processor E5-1600/2400/2600/4600 Product Families and Intel® Xeon® Processor E5-1600/2400/2600/4600 v2 Product Families BIOS Specification Update - NDA version 2.0.2 or later and release notes.

Workaround: A BIOS code change has been identified and may be implemented as a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA107 The Vswing of the PCIe* Transmitter Exceeds The Specification

Problem: The PCIe Specification defines a limit for the Vswing (Voltage Swing) of the differential lines that make up a lane to be 1200 mV peak-to-peak when operating at 2.5 GT/s and 5 GT/s. Intel has found that the processor's PCIe transmitter may exceed this specification. Peak-to-peak swings on a limited number of samples have been observed up to 1450 mV.

Implication: For those taking direct measurements of the PCIe transmit traffic coming from the processor may detect that the Vswing exceeds the PCIe Specification. Intel has not observed any functional failures due to this erratum.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA108 PECI_WAKE_MODE is Always Reported as Disabled

Problem: Due to this erratum, the state of PECI_WAKE_MODE is always reported as disabled. The PECI (Platform Environment Control Interface) PCS (Package Configuration Service) WRITE_PECI_WAKE_MODE (0x5) command correctly updates the state of PECI_WAKE_MODE, but the PECI PCS READ_PECI_WAKE_MODE (0x5) always reports the PECI_WAKE_MODE as disabled.

Implication: Software depending on the reported value for PECI_WAKE_MODE may not behave as expected.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA109 Poisoned PCIe* AtomicOp Completions May Return an Incorrect Byte Count

Problem: A poisoned PCIe AtomicOp request completion may have an incorrect byte count.

Implication: When this erratum occurs, PCIe devices which enable byte count checking will log an unexpected completion and issue a CTO (Completion Time Out).

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA110 Incorrect Speed and De-emphasis Level Selection During DMI Compliance Testing

Problem: When the DMI port is operating as a PCIe* port, it supports only 2.5GT/s and 5GT/s data rates. According to the PCIe specification, the data rate and de-emphasis level for the compliance patterns should be based on the maximum data rate supported. Due to

this erratum, the port may select an 8 GT/s data rate and associated de-emphasis level during compliance testing mode.

Implication: When doing PCIe load board compliance testing, the DMI port may transmit using 8GT/s data rate and de-emphasis levels.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA111 PCIe* Device 3 Does Not Log an Error in UNCERRSTS When an Invalid Sequence Number in an Ack DLLP is Received

Problem: If the processor's PCIe device 3 controller receives an invalid sequence number in an Ack DLLP (Data Link Layer Packet), it is expected to log an uncorrectable error for the affected port in bit [4] of the UNCERRSTS register (Bus 0; Device 3; Function 3:0; Offset 14CH). Due to this erratum, no data link protocol error is logged when an invalid sequence number in an Ack DLLP occurs on PCIe device 3.

Implication: Software that uses this register upon an uncorrectable PCIe error will not be able to identify this specific error type.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA112 Programmable Ratio Limits For Turbo Mode is Reported as Disabled

Problem: The Programmable Ratio Limits for Turbo Mode in bit 28 of the MSR_PLATFORM_INFO MSR (CEH) should be 1 but, due to this erratum, it is reported as 0.

Implication: Due to this erratum, software will incorrectly assume it cannot dynamically vary the factory configured ratio limit values specified in MSR_TURBO_RATIO_LIMIT MSR (1ADH) and MSR_TURBO_RATIO_LIMIT1 MSR (1AEH).

Workaround: Software should treat the Programmable Ratio Limits for Turbo Mode bit as set.

Status: For the affected steppings, see the Summary Tables of Changes.

CA113 PCIe* TLPs in Disabled VC Are Not Reported as Malformed

Problem: The PCIe Base Specification requires processors to report a TLP (Transaction Layer Packet) with a TC (Traffic Class) that is not mapped to any enabled VC (Virtual Channel) in an Ingress Port as a Malformed TLP. Due to this erratum, a TLP received on the DMI port that not is mapped to an enabled VC is not reported as a Malformed TLP.

Implication: Receipt of a TLP with an unmapped PCIe TC may lead to completion time out events or other unexpected system behavior. Intel has not observed this erratum with any commercially available software or platform.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA114 PCIe* Link May Fail to Train to 8.0 GT/s

Problem: Due to this erratum, with certain 8.0 GT/s-capable link partners, the PCIe link may fail to train to 8.0 GT/s as requested.

Implication: When this erratum occurs, the PCIe link will enter an infinite speed-change request loop.

Workaround: A BIOS code change has been identified and may be implemented as a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes

CA115 PCIe* Header of a Malformed TLP is Logged Incorrectly

Problem: If a PCIe port receives a malformed TLP (Transaction Layer Packet), an error is logged in the UNCERRSTS register (Device 0; Function 0; Offset 14CH and Device 2-3; Function 0-3; Offset 14CH). Due to this erratum, the header of the malformed TLP is

logged incorrectly in the HDRLOG register (Device 0; Function 0; Offset 164H and Device 2-3; Function 0-3; Offset 164H).

Implication: The PCIe header of a malformed TLP is not logged correctly.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA116 PCIe* May Associate Lanes That Are Not Part of Initial Link Training to L0 During Upconfiguration

Problem: The processor should not associate any lanes that were not part of the initial link training in subsequent upconfiguration requests from an endpoint. Due to this erratum, the processor may associate any Lane that has exited Electrical Idle, even if it is beyond the width of the initial Link training.

Implication: Upconfiguration requests may result in a Link wider than the initially-trained Link.

Workaround: Endpoints must ensure that upconfiguration requests do not request a Link width wider than that negotiated during initial Link training.

Status: For the affected steppings, see the Summary Tables of Changes.

CA117 Single PCIe* ACS Violation or UR Response May Result in Multiple Correctable Errors Logged

Problem: An ACS (Access Control Services) error or UR (Unsupported Request) PCIe completion status can trigger a LER (Live Error Recovery) if they are unmasked in the LER_UNCERRMSK (Bus 0; Device 0/1/2/3; Function 0/0-1/0-3/0-3; Offset 28CH) and LER_XPUNCERRMSK (Bus 0; Device 0/1/2/3; Function 0/0-1/0-3/0-3; Offset 290H) CSRs, respectively. Due to this erratum, the Root Port Error Status "multiple_correctable_error_received" bit (RPERRSTS[1], CPUBUSNO(0), Device 0:3, Functions 0/0-1/0-3/0-3, Offset 0x178) may be set upon on a single ACS or UR error.

Implication: PCIe error handling software may not behave as expected after an ACS error or a UR completion status. Intel has not observed this erratum with any commercially available software or system.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA118 PCIe* Extended Tag Field May be Improperly Set

Problem: The Extended Tag field in the TLP Header will not be zero for TLPs issued by PCIe ports 1a, 1b, 2c, 2d, 3c, and 3d even when the Extended Tag Field Enable bit in the Device Control Register (Offset 08H, bit 8) is 0.

Implication: This erratum does not affect ports 0, 2a, 2b, 3a and 3b. This erratum will not result in any functional issues when using device that properly track and return the full 8 bit Extended Tag value with the affected ports. However, if the Extended Tag field is not returned by a device connected to an affected port then this erratum may result in unexpected completions and completion timeouts.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA119 Power Meter May Under-Estimate Package Power

Problem: Power Meter provides a real-time power consumption estimate for the processor. Depending on operating conditions and variations in certain component-specific characteristics, the reported power may be below the actual power consumption.

Implication: Due to Intel® Turbo Boost Technology using the Power Meter to compare instantaneous power consumption to the rated TDP, the core frequency in P0 may set to a ratio where the processor exceeds its rated TDP. Further, using the average power limit facility (RAPL) may cause the processor to run at a power consumption level that is higher than expected.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA120 DTS2.0 May Report Inaccurate Temperature Margin

Problem: When DTS (Digital Thermal Sensor) 2.0 is enabled on the E5-4600 v2 product family, the thermal margin reported by the PACKAGE_THERM_MARGIN MSR (1A1H) THERMAL_MARGIN bits [15:0] may be inaccurate.

Implication: Due to this erratum, fan speed control algorithms may set the fan speed incorrectly.

Workaround: It is possible for BIOS to contain processor configuration data and code changes as a workaround for this erratum. Please refer to the latest version of the Intel® Xeon® Processor E5-1600/2400/2600/4600 Product Families and Intel® Xeon® Processor E5-1600/2400/ 2600/4600 v2 Product Families BIOS spec update version 2.0.5 or later and release notes.

Status: For the affected steppings, see the Summary Tables of Changes.

CA121 PMON Counters Overflow May Not Trigger PMON Global Freeze

Problem: In the event of a R2PCIe / R3QPI PMON register PMONCTRSTATUS0 (Device 19; Function 1; Offset 0xF8) and PMONCTRSTATUS{0:1} (Device 18,19; Function 5,6; Offset 0xF8) overflow, the PMON unit sends an overflow message to a global PMON manager. The global PMON manager then asserts the global freeze signal and disables all of its counters. Due to this erratum, the global PMON manager will not assert the global freeze signal and disable all of its counters.

Implication: PMON counters may fail to freeze when either of R2PCIe /R3QPI PMON Counters Overflow. The counts on PMON unit may not be accurate.

Workaround: Upon receipt of an overflow message, software should set bit 8 of the PMONUNITCTRL0 (Device 19; Function 1; Offset 0xF4) and PMONUNITCTRL{0:1} (Device 18,19; Function 5,6; Offset 0xF4) to freeze the counter.

Status: For the affected steppings, see the Summary Tables of Changes.

CA122 Processor May Log a Machine Check when MSI is signaled by a device

Problem: In certain cases, when MSI (Message Signaled Interrupt) is signaled by a device, a machine check error may be logged in the IA32_MC4_STATUS MSR (411H) with MSCOD field bits [31:24] equal to values of 73H or 74H.

Implication: A machine check may be observed when MSI is signaled by a device.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA123 Spurious Patrol Scrub Errors Observed During a Warm Reset

Problem: The patrol scrub engine continues to run during a warm reset; this can lead to spurious errors being reported by the Memory Controller while memory is in Self Refresh.

Implication: Due to this erratum, erroneous patrol scrub errors may be observed during a warm reset.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA124 PECI May Not be Able to Access IIO CSRs

Problem: Due to this erratum, when the processor has viral enabled and an uncorrectable error occurs in the core, PECI (Platform Environment Control Interface) may not be able to access IIO (Integrated I/O) CSRs.

Implication: When this erratum occurs, IIO CSR access using a PECI RdPCIConfigLocal() or WrPCIConfigLocal() command will return a status of 91H, indicating that the request could not be processed.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA125 A DMI UR May Unexpectedly Cause a CATERR# After a Warm Reset

Problem: Reset disables certain error detection facilities to prevent error signaling from interfering with system initialization. Due to this erratum, DMI UR (Unsupported Request) error reporting, if previously enabled by BIOS, is not disabled by a warm reset.

Implication: A platform event shortly after a warm reset that produces a DMI UR is subject to this erratum. When this erratum occurs, a CATERR# is signaled with IA32_MCi_STATUS.MCACOD = 0xE0B. Some platforms automatically reset after a CATERR# so this erratum may be seen as an unexpected re-boot.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA126 Duplicate. Erratum Removed

CA127 VM Exit May Set IA32_EFER.NXE When IA32_MISC_ENABLE Bit 34 is Set to 1

Problem: When "XD Bit Disable" in the IA32_MISC_ENABLE MSR (1A0H) bit 34 is set to 1, it should not be possible to enable the "execute disable" feature by setting IA32_EFER.NXE. Due to this erratum, a VM exit that occurs with the 1-setting of the "load IA32_EFER" VM-exit control may set IA32_EFER.NXE even if IA32_MISC_ENABLE bit 34 is set to 1. This erratum can occur only if IA32_MISC_ENABLE bit 34 was set by guest software in VMX non-root operation.

Implication: Software in VMX root operation may execute with the "execute disable" feature enabled despite the fact that the feature should be disabled by the IA32_MISC_ENABLE MSR. Intel has not observed this erratum with any commercially available software.

Workaround: A virtual-machine monitor should not allow guest software to write to the IA32_MISC_ENABLE MSR.

Status: For the affected steppings, see the Summary Tables of Changes.

CA128 Receiver Termination Impedance On PCIe* 3.0 Does Not Comply With The Specification

Problem: The PCIe* Base Specification revision 3.0 defines ZRX-HIGH-IMP-DC-NEG and ZRX-HIGH-IMP-DC-POS for termination impedance of the receiver. The specified impedance for a negative voltage (-150 mV to 0V) is expected to be greater than 1 Kohm. Sampled measurements of this impedance as low as 400 ohms have been seen. The specified impedance for a positive voltage (> 200 mV) is greater than 20 Kohms. Sampled measurements of this impedance as low as 14.6 Kohms have been seen.

Implication: Intel has not observed functional failures from this erratum on any commercially available platforms using any commercially available PCIe device.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA129 Platform Recovery After a Machine Check May Fail

Problem: While attempting platform recovery after a machine check (as indicated by CATERR# signaled from the legacy socket), the original error condition may prevent normal platform recovery which can lead to a second machine check. A remote processor detecting a second Machine Check Event will hang immediately.

Implication: Due to this erratum, it is possible a system hang may be observed during a warm reset caused by a CATERR#.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA130 PECI May be Non-responsive When System is in BMC Init Mode

Problem: The allow_peci_pcode_error_rsp field in the DYNAMIC_PERF_POWER_CTL CSR (Device 10; Function 2; Offset 0x64H; bit 16) does not retain its value after a warm reset. When the system is in BMC Init mode, this erratum can cause PECI (Platform Environment Control Interface) access to be non-responsive after a warm reset caused by a Machine Check Event.

Implication: When this erratum occurs, PECI requests will return a status of 91H, indicating that the request could not be processed.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA131 Processor May Issue Unexpected NAK DLLP Upon PCIe* L1 Exit

Problem: Upon exiting the L1 link power state, the processor's PCIe port may unexpectedly issue a NAK DLLP (Data Link Layer Packet).

Implication: PCIe endpoints may unexpectedly receive and log a NAK DLLP.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA132 Surprise Down Error Status is Not Set Correctly on DMI Port

Problem: Due to this erratum, the Surprise_down_error_status (UNCERRSTS Device 0; Function 0; Offset 0x14C; bit 5) is not set to 1 when DMI port detects a surprise down error.

Implication: Surprise down errors will not be logged for the DMI port. This violates the PCIe* Base Specification. Software that relies on this status bit may not behave as expected. It is likely that conditions resulting in surprise down error would lead to other errors being logged; a surprise down on DMI is not likely to be a silent event.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA133 Core C-state Residency Counters May Return Stale Data

Problem: Updates to the Core C-state residency counter MSRs occur after a core wakeup has completed. Reads of these MSRs which occur between a core wakeup and the subsequent Core C-state residency counter updates will return the pre-sleep values. The affected MSRs are MSR_CORE_C3_RESIDENCY (3FCH), MSR_CORE_C6_RESIDENCY (3FDH), and MSR_CORE_C7_RESIDENCY (3FEH).

Implication: Due to this erratum, reads of the Core C-state may return stale data. As a result, software may interpret the data incorrectly.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA134 PCIe* Ports Operating at 8 GT/s May Issue an Additional Packet After Stop and Scream Occurs

Problem: The processor's Stop and Scream mode requires that an attempt to send a poisoned packet results in that poisoned packet and all subsequent packets being dropped. When operating a PCIe link at 8 GT/s, Stop and Scream will drop the attempted poisoned packet but, due to this erratum, one additional good packet may be sent.

Implication: When this erratum occurs, the end-device may detect a sequencing issue because of the additional packet.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA135 Incorrect Page Translation when EPT is enabled

Problem: If EPT (Extended Page Tables) is enabled, then a complex sequence of internal processor events may result in unexpected page faults or use of incorrect page translations.

Implication: Due to this erratum a guest may crash or experience unpredictable system behavior.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes

CA136 A Machine Check Exception Due to Instruction Fetch May Be Delivered Before an Instruction Breakpoint

Problem: Debug exceptions due to instruction breakpoints take priority over exceptions resulting from fetching an instruction. Due to this erratum, a machine check exception resulting from the fetch of an instruction may take priority over an instruction breakpoint if the instruction crosses a 32-byte boundary and the second part of the instruction is in a 32-byte poisoned instruction fetch block.

Implication: Instruction breakpoints may not operate as expected in the presence of a poisoned instruction fetch block.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA137 Accessing Physical Memory Space 0-640K through the Graphics Aperture May Cause Unpredictable System Behavior

Problem: The physical memory space 0-640K when accessed through the graphics aperture may result in a failure for writes to complete or reads to return incorrect results.

Implication: A hang or functional failure may occur during graphics operation such as OGL or OCL conformance tests, 2D/3D games and graphics intensive application.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

CA138 The RDRAND Instruction Will Not Execute as Expected

Problem: On processors that support the RDRAND instruction, that capability should be reported via the setting of CPUID.01H:ECX.RDRAND[bit 30]. Due to this erratum, that bit will not be set, and the execution of the RDRAND instruction will result in a #UD exception.

Implication: Software will not be able to utilize the RDRAND instruction.

Workaround: It is possible for the BIOS to contain a workaround for this erratum to report RDRAND as present via CPUID and allow proper execution of RDRAND.

Status: For the affected steppings, see the Summary Tables of Changes.

CA139 Writes to B2BSPAD[15:0] Registers May Transfer Corrupt Data Between NTB Connected Systems

Problem: Writes to the NTB (Non-Transparent Bridge) B2BSPAD[15:0] registers (BAR PB01BASE, SB01BASE; Offsets 100H - 13FH) may result in corrupted data transfer between systems.

Implication: Using B2BSPAD[15:0] registers to transfer data may not work as expected.

Workaround: Do not use the B2BSPAD[15:0] to send data from the local to remote host. Instead, configure one of the following local register pairs to point to the remote SB01BASE region:

- PB23BASE (Device 3; Function 0; Offset 18H) and PBAR2XLAT (Offset 10H) from PB01BASE or SB01BASE regions
- PB45BASE (Device 3; Function 0; Offset 20H) and PBAR4XLAT(Offset 18H) from PB01BASE, or SB01BASE regions

The local host may then write directly to the SPAD[15:0] registers (Offsets 80H - 0BFH) of the remote system from the PB23BASE/PB45BASE region defined above.

Status: For the affected steppings, see the Summary Tables of Changes.

CA140 Reading DDRIO Broadcast CSRs Via PECI May Return Incorrect Data

Problem: Device 15, Functions 6 and 7 are DDRIO broadcast CSRs; writing one broadcast CSR writes the corresponding CSR in each of the four memory channels (Individual memory channel CSRs are located at Device 17, Functions 0, 1, 4 and 5). Due to this erratum, using the PECI (Platform Environment Control Interface) RdConfig() command to read a DDRIO broadcast CSR may return incorrect data.

Implication: When this erratum occurs, software that reads broadcast CSRs may behave unexpectedly.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA141 Corrected Filtering Indication May be Incorrect in LLC Machine Check Bank Status Register

Problem: The Corrected Filtering indication in IA32_MC{17-28}_STATUS.MCACOD bit 12 may be calculated incorrectly under some conditions.

Implication: Software using the Corrected Filtering indication in IA32_MC{17-28}_STATUS.MCACOD may not function as expected.

Workaround: None identified.

Status: For the affected steppings, see the Summary Tables of Changes.

CA142 RTID_POOL_CONFIG Registers Incorrectly Behave as a Read-Write Registers

Problem: The RTID_POOL_CONFIG CSRs (Device 12; Function 0-5; Offset ACH and Device 13, Function 0-5; Offset ACH) were intended to be Read-Only. Due to this erratum, these registers behave incorrectly as Read-Write.

Implication: Writes to the RTID_POOL_CONFIG CSRs may lead to unexpected results.

Workaround: None identified. Software should write to RTID_POOL_CONFIG_SHADOW CSRs (Device 12; Function 0-5; Offset B0H and Device 13; Function 0-5; Offset B0H) rather than RTID_POOL_CONFIG CSRs.

Status: For the affected steppings, see the Summary Tables of Changes.

CA143 Catastrophic Trip Triggered at Lower Than Expected Temperatures

Problem: Catastrophic Trip is intended to provide protection when the temperature of the processor exceeds a critical threshold by immediately shutting down the processor. Due to this erratum, the Catastrophic Trip may be triggered well below the critical threshold.

Implication: When this erratum occurs, the processor improperly issues a catastrophic shutdown causing a system failure.

Workaround: It is possible for the BIOS to contain a workaround for this erratum.

Status: For the affected steppings, see the Summary Tables of Changes.

§

Specification Changes

1. There are no specification changes in this specification update revision.

Specification Clarifications

1. There are no specification clarifications in this specification update revision.

§

Documentation Changes

The Documentation Changes listed in this section apply to the following documents:

- Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 1: Basic Architecture
- Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 2A: Instruction Set Reference Manual A-M
- Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 2B: Instruction Set Reference Manual N-Z
- Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 3A: System Programming Guide
- Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 3B: System Programming Guide

All Documentation Changes will be incorporated into a future version of the appropriate Processor documentation.

Note: Documentation changes for Intel® 64 and IA-32 Architecture Software Developer's Manual volumes 1, 2A, 2B, 3A, and 3B will be posted in a separate document, Intel® 64 and IA-32 Architecture Software Developer's Manual Documentation Changes. Follow the link below to become familiar with this file.

http://developer.intel.com/products/processor/manuals/index.htm

1. SDM, Volume 3B: On-Demand Clock Modulation Feature Clarification

Software Controlled Clock Modulation section of the Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 3B: System Programming Guide will be modified to differentiate On-demand clock modulation feature on different processors. The clarification will state:

For Intel® Hyper-Threading Technology enabled processors, the IA32_CLOCK_MODULATION register is duplicated for each logical processor. In order for the On-demand clock modulation feature to work properly, the feature must be enabled on all the logical processors within a physical processor. If the programmed duty cycle is not identical for all the logical processors, the processor clock will modulate to the highest duty cycle programmed for processors if the CPUID DisplayFamily_DisplayModel signatures is listed in Appendix 14-2. For all other processors, if the programmed duty cycle is not identical for all logical processors in the same core, the processor will modulate at the lowest programmed duty cycle.

For multiple processor cores in a physical package, each core can modulate to a programmed duty cycle independently.

For the P6 family processors, on-demand clock modulation was implemented through the chipset, which controlled clock modulation through the processor's STPCLK# pin.

Table 14-2. CPUID Signatures for Legacy Processors That Resolve to Higher Performance Setting of Conflicting Duty Cycle Requests

DisplayFamily_DisplayModel	DisplayFamily_DisplayModel	DisplayFamily_DisplayModel	DisplayFamily_DisplayModel
0F_xx	06_1C	06_1A	06_1E
06_1F	06_25	06_26	06_27
06_2C	06_2E	06_2F	06_35
06_36			

2. Intel® Xeon® Processor E5 v2 Product Family Datasheet- Volume Two: Device Mapping Addition

The following register group will be added to the table 1-1 titled "Functions Specifically Handled by the Processor" in section 1.2.1.4 in the document Intel® Xeon® Processor E5 v2 Product Family Datasheet- Volume Two.

Register Group	DID	Device	Function	Comment
PCI Express Root Port 1	0xE02, 0xE03	1	0-1	x8 or x4 max link width

Mixed Processors Within DP Platforms

Mixed Processor Consistency Requirements

Intel supports dual processor (DP) configurations consisting of processors:

1. From the same power rating.
2. That support the same maximum Intel® QuickPath Interconnect (Intel® QPI) and DDR3 memory speeds.
3. That share symmetry across physical packages with respect to the number of logical processors per package, number of cores per package, number of Intel QPI Interfaces, and cache topology.
4. That have identical Extended Family, Extended Model, Processor Type, Family Code and Model Number as indicated by the function 1 of the CPUID instruction.

Note: Processors must operate with the same Intel QPI, DDR3 memory and core frequency.

While Intel does nothing to prevent processors from operating together, some combinations may not be supported due to limited validation, which may result in uncharacterized errata. Coupling this fact with the large number of Intel Xeon processor E5 v2 series attributes, the following population rules and stepping matrix have been developed to clearly define supported configurations.

1. Processors must be of the same power rating. For example, mixing of 95W Thermal Design Power (TDP) processors is supported. Mixing of dissimilar TDPs in the same platform is not supported (for example, 95W with 130W, and so forth).
2. Processors must operate at the same core frequency. Note: Processors within the same power-optimization segment supporting different maximum core frequencies (for example, a 2.93 GHz / 95 W and 2.66 GHz / 95W) can be operated within a system. However, both must operated at the highest frequency rating commonly supported. Mixing components operating at different internal clock frequencies is not supported and will not be validated by Intel.
3. Processors must share symmetry across physical packages with respect to the number of logical processors per package, number of Intel QPI Interfaces, and cache topology.
4. Mixing dissimilar steppings is only supported with processors that have identical Extended Family, Extended Model, Processor type, Family Code and Model Number as indicated by the function 1 of the CPUID instruction. Mixing processors of different steppings but the same model (as per CPUID instruction) is supported. Details regarding the CPUID instruction are provided in the *AP-487, Intel® Processor Identification and the CPUID Instruction* application note and *Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 2A*.
5. After AND'ing the feature flag and extended feature flag from the installed processors, any processor whose set of feature flags exactly matches the AND'ed feature flags can be selected by the BIOS as the BSP. If no processor exactly matches the AND'ed feature flag values, then the processors with the numerically lower CPUID should be selected as the BSP.
6. Intel requires that the processor microcode update be loaded on each processor operating within the system. Any processor that does not have the proper microcode update loaded is considered by Intel to be operating out of specification.
7. The workarounds identified in this, and subsequent specification updates, must properly applied to each processor in the system. Certain errata are specific to the

multiprocessor environment. Errata for all processor steppings will affect system performance if not properly worked around.

8. Customers are fully responsible for the validation of their system configurations.

Mixed Steppings

Mixing processors of different steppings but the same model (as per CPUID instruction) is supported provided there is no more than one stepping delta between the processors, for example, S and S+1.

S and S+1 is defined as mixing of two CPU steppings in the same platform where one CPU is S (stepping) = CPUID.(EAX=01h):EAX[3:0], and the other is S+1 = CPUID.(EAX=01h):EAX[3:0]+1. The stepping ID is found in EAX[3:0] after executing the CPUID instruction with Function 01h.

§