



4th Gen Intel® Xeon® Processor Scalable Family, Codename Sapphire Rapids

Data Sheet Vol. 2 Registers

Rev. 001

July 2025



Notice: This document contains information on products in the design phase of development. The information here is subject to change without notice. Do not finalize a design with this information.

Intel technologies may require enabled hardware, software or service activation.

You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein. You agree to grant Intel a non-exclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Performance varies by use, configuration, and other factors. Learn more on the [Performance Index site](#).

Your costs and results may vary.

"Conflict-free" refers to products, suppliers, supply chains, smelters, and refiners that, based on our due diligence, do not contain or source tantalum, tin, tungsten or gold (referred to as "conflict minerals" by the U.S. Securities and Exchange Commission) that directly or indirectly finance or benefit armed groups in the Democratic Republic of the Congo or adjoining countries.

All product plans and roadmaps are subject to change without notice.

Code names are used by Intel to identify products, technologies, or services that are in development and not publicly available. These are not "commercial" names and not intended to function as trademarks.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

Copies of documents which have an order number and are referenced in this document may be obtained by calling 1-800-548-4725 or visiting the [Resource and Documentation Center](#).

Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

Copyright © 2025, Intel Corporation. All rights reserved.

Contents

Revision History.....	6
1.0 Introduction.....	7
1.1 Related Documents.....	7
1.2 Terminology.....	8
2.0 Registers Overview.....	12
2.1 Register Spaces and Structure.....	12
2.2 Access Types.....	14
2.3 Configuration Rules.....	16
2.4 Notational Conventions.....	16
3.0 Intel® Ultra Path Interconnect (Intel® UPI) Registers.....	17
3.1 Bus: B(30), Device: 1-2-3-4, Function: 0 (UPI Misc).....	17
3.1.1 Vendor ID (VID) — Offset 0h.....	17
3.1.2 Device ID (DID) — Offset 2h.....	18
3.1.3 PCI Status (PCISTS) — Offset 6h.....	18
3.1.4 Revision ID and Class Code Register (RID_CCR) — Offset 8h.....	19
3.1.5 Cache Line Size Register (CLSR) — Offset Ch.....	19
3.1.6 Header Type (HDR) — Offset Eh.....	20
3.1.7 Subsystem Vendor ID (SVID) — Offset 2Ch.....	20
3.1.8 Subsystem ID (SDID) — Offset 2Eh.....	20
3.1.9 Capability Pointer (CAPPTR) — Offset 34h.....	21
3.1.10 PCI Express Capability (PXPCAP) — Offset 40h.....	21
3.1.11 Miscellaneous Clock Control (KTIMISCCLKCTRL) — Offset Ech.....	22
3.1.12 PCI Express Enhanced Capability (XPENHCAP) — Offset 100h.....	22
3.1.13 kti misc privc0 (KTIMISCPRIVC0) — Offset 1D0h.....	23
3.1.14 kti misc scratch cfg32 (KTIMISCSCRATCHCFG32) — Offset 304h.....	23
4.0 Integrated Memory Controller (IMC) Registers.....	25
4.1 Bus: B(30), Device: 12, Function: 0 (IMC0 Mesh to Mem Registers).....	25
4.1.1 Vendor ID (VID_0_12_0_CFG) — Offset 0h.....	25
4.1.2 Device ID (DID_0_12_0_CFG) — Offset 2h.....	26
4.1.3 Header Type (HDR_0_12_0_CFG) — Offset Eh.....	26
4.1.4 PCI Express Capability (PXPCAP_0_12_0_CFG) — Offset 40h.....	26
4.1.5 PCI Express Enhanced Capability (XPENHCAP_0_12_0_CFG) — Offset 100h....	27
4.1.6 TAD Read (TAD_RD) — Offset 238h.....	27
4.1.7 TAD Write (TAD_WR) — Offset 240h.....	30
4.1.8 Mode register (MODE) — Offset 318h.....	32
4.1.9 UPI Routing Config (UPIROUTINGCONFIG) — Offset 320h.....	34
4.1.10 UPI Routing Table (UPIROUTINGTABLE) — Offset 328h.....	35
5.0 Configuration Agent (Ubox) Registers.....	36
5.1 Bus: B(30), Device: 0, Function: 0 (EVNTS).....	36
5.1.1 Vendor ID (VID_0_0_0_CFG) — Offset 0h.....	36
5.1.2 Device ID (DID_0_0_0_CFG) — Offset 2h.....	36
5.1.3 The register type IntControl_CFG (INTCONTROL_CFG) — Offset 94h.....	37
5.1.4 The register type CPUNODEID_CFG (CPUNODEID_CFG) — Offset C0h.....	37
5.1.5 The register type GIDNIDMAP_CFG (GIDNIDMAP_CFG) — Offset D4h.....	38

5.2 Bus: B(30), Device: 0, Function: 2 (DECS).....	39
5.2.1 Vendor ID (VID_0_0_2_CFG) — Offset 0h.....	39
5.2.2 Device ID (DID_0_0_2_CFG) — Offset 2h.....	39
5.2.3 The register type CPUBUSNO_CFG (CPUBUSNO_CFG) — Offset D0h.....	40
5.2.4 SMI generation control (UBOX_SMICTRL_CFG) — Offset D8h.....	40
6.0 Power Control Unit (PCU) Registers.....	42
6.1 Bus: B(31), Device: 30, Function: 0 (PCU CR0 Registers).....	42
6.1.1 Vendor ID (VID_1_30_0_CFG) — Offset 0h.....	42
6.1.2 Device ID (DID_1_30_0_CFG) — Offset 2h.....	43
6.1.3 PACKAGE ENERGY STATUS (PACKAGE_ENERGY_STATUS_CFG) — Offset 90h.....	43
6.1.4 PACKAGE TEMPERATURE (PACKAGE_TEMPERATURE_CFG) — Offset C8h.....	43
6.1.5 TEMPERATURE TARGET (TEMPERATURE_TARGET_CFG) — Offset E4h.....	43
6.2 Bus: B(31), Device: 30, Function: 2 (PCU CR2 Registers).....	44
6.2.1 Vendor ID (VID_1_30_2_CFG) — Offset 0h.....	44
6.2.2 Device ID (DID_1_30_2_CFG) — Offset 2h.....	45
6.2.3 PCI Command (PCICMD_1_30_2_CFG) — Offset 4h.....	45
6.2.4 PCI Status (PCISTS_1_30_2_CFG) — Offset 6h.....	46
6.2.5 Revision ID and Class Code Register (RID_CCR_1_30_2_CFG) — Offset 8h.....	47
6.2.6 Cache Line Size Register (CLSR_1_30_2_CFG) — Offset Ch.....	47
6.2.7 PCI Latency Timer (PLAT_1_30_2_CFG) — Offset Dh.....	48
6.2.8 Header Type (HDR_1_30_2_CFG) — Offset Eh.....	48
6.2.9 Built In Self Test (BIST_1_30_2_CFG) — Offset Fh.....	48
6.2.10 Subsystem Vendor ID (SVID_1_30_2_CFG) — Offset 2Ch.....	49
6.2.11 Subsystem ID (SDID_1_30_2_CFG) — Offset 2Eh.....	49
6.2.12 Capability Pointer (CAPPTR_1_30_2_CFG) — Offset 34h.....	49
6.2.13 Interrupt Line (INTL_1_30_2_CFG) — Offset 3Ch.....	50
6.2.14 Interrupt Pin (INTPIN_1_30_2_CFG) — Offset 3Dh.....	50
6.2.15 Minimum Grant (MINGNT_1_30_2_CFG) — Offset 3Eh.....	50
6.2.16 Maximum Latency (MAXLAT_1_30_2_CFG) — Offset 3Fh.....	51
6.2.17 DRAM ENERGY STATUS (DRAM_ENERGY_STATUS_CFG) — Offset 80h.....	51
6.2.18 PACKAGE OVERFLOW STATUS (PACKAGE_RAPL_PERF_STATUS) — Offset 88h.....	52
6.2.19 DRAM POWER INFORMATION (DRAM_POWER_INFO_CFG) — Offset A8h.....	52
6.2.20 MEM TRML TEMPERATURE REPORT (MEM_TRML_TEMPERATURE_REPORT[0]) — Offset C8h.....	53
6.2.21 MEM TRML TEMPERATURE REPORT (MEM_TRML_TEMPERATURE_REPORT[1]) — Offset CCh.....	54
6.2.22 MEM TRML TEMPERATURE REPORT (MEM_TRML_TEMPERATURE_REPORT[2]) — Offset D0h.....	54
6.2.23 MEM TRML TEMPERATURE REPORT (MEM_TRML_TEMPERATURE_REPORT[3]) — Offset D4h.....	54
6.2.24 DRAM PLANE OVERFLOW STATUS (DRAM_RAPL_PERF_STATUS_CFG) — Offset D8h.....	54
6.2.25 THERMTRIP CONFIGURATION (THERMTRIP_CONFIG_CFG) — Offset F8h.....	55



Tables

1	Public Publications.....	7
2	Base Address Register (BAR) for Functional Blocks.....	13
3	Field Access Types.....	14
4	Summary of Power Control Unit (PCU) Registers.....	42



Revision History

Revision	Description	Date
001	Initial release of the document	July 2025

1.0 Introduction

NOTE

All information in this document is preliminary and subject to change.

The 4th Gen Intel® Xeon® Processor Scalable Family, Codename Sapphire Rapids server is the leadership wave 3 10 nm+++ Intel® Xeon® multi-core server processor. The processor supports up to 52 bits of physical address space and 57 bits of virtual address space. The processor is designed for a platform consisting of at least one 4th Gen Intel® Xeon® Processor Scalable Family, Codename Sapphire Rapids and the Intel® C740 Series Chipset Platform Controller Hub (PCH). Included in this family of processors are Integrated Memory Controller (IMC) and an Integrated I/O (IIO) on a single silicon die.

All processor types support up to 64 lanes of PCI Express* 5.0 links capable of 32 GT/s, and eight lanes of DMI. It features four Integrated Memory Controllers (IMC), each IMC supporting up to two channels of DDR5 DIMMs with up to two DIMMs per channel.

1.1 Related Documents

Materials and concepts available in the following documents may be beneficial when reading this document.

The following table lists public references that provide context, details, and further clarifications related to the topics of this document.

Table 1. Public Publications

Document	Document Number/Location
<i>Advanced Configuration and Power Interface Specification 4.0</i>	www.acpi.info
<i>PCI Local Bus Specification 3.0</i>	http://www.pcisig.com/
<i>PCI Express Base Specification, Revision 5.0, version 1.0</i>	http://www.pcisig.com/
<i>DDR5 SDRAM Specification and Register Specification</i>	http://www.jedec.org/
Intel® 64 and IA-32 Architectures Software Developer Manuals, which consists of: • Volume 1: Basic Architecture • Volume 2A: Instruction Set Reference, A-M • Volume 2B: Instruction Set Reference, N-Z • Volume 3A: System Programming Guide • Volume 3B: System Programming Guide • Intel® 64 and IA-32 Architectures Optimization Reference Manual	767375 Intel® 64 and IA-32 Architectures Software Developer Manuals
<i>Intel® Virtualization Technology Specification for Directed I/O Architecture Specification</i>	Intel® Virtualization Technology for Directed I/O Architecture Specification
<i>Intel® Trusted Execution Technology (Intel® TXT): Software Development Guide</i>	315168
continued...	

Document	Document Number/Location
	https://www.intel.com/content/www/us/en/content-details/315168/intel-trusted-execution-technology-intel-txt-software-development-guide.html?wapkw=315168
4th Gen Intel® Xeon® Processor Scalable Family (codename Sapphire Rapids) Data Sheet Vol 1 - Architecture	814093
4th Gen Intel® Xeon® Processor Scalable Family (codename Sapphire Rapids) Data Sheet Vol 3 - Electrical	TBD

NOTE

The most up-to-date collateral list is found on rdc.intel.com.

1.2 Terminology

The following acronyms and terms are used in this document and related documentation.

Term	Description
AC	Read and Write Access Control
ASPM	Active State Power Management
BL	Burst Length
BMC	Baseboard Management Controller
CA	Caching Agent
CHA	Caching and Home Agent
CP	Control Policy
CSR	Configuration Space Register
DDR5	Fifth generation Double Data Rate SDRAM memory technology.
DMA	Direct Memory Access
DTLB	Data Translation Look-aside Buffer. Part of the processor core architecture.
DTS	Digital Thermal Sensor
ECC	Error Correction Code
EDAF	eSPI direct attached flash
EDP	Electrical Design Point
Enhanced Intel SpeedStep® Technology	Allows the operating system to reduce power consumption when performance is not needed.
Execute Disable Bit	The Execute Disable bit allows memory to be marked as executable or non-executable, when combined with a supporting operating system. If code attempts to run in non-executable memory the processor raises an error to the operating system. This feature can prevent some classes of viruses or worms that exploit buffer overrun vulnerabilities and can thus help improve the overall security of the system. Refer to the <i>Intel® 64 and IA-32 Architectures Software Developer's Manual</i> for detailed information.
continued...	

Term	Description
FIVR	Fully Integrated Voltage Regulator
FLIT	Flow Control Unit. The Intel® UPI Link layer's unit of transfer. A FLIT is made of multiple PHITS. A Flit is always a fixed amount of information (192 bits).
Functional Operation	Refers to the normal operating conditions in which all processor specifications, including DC, AC, system bus, signal quality, mechanical, and thermal, are satisfied.
GSSE	Extension of the SSE/SSE2 (Streaming SIMD Extensions) floating point instruction set to 256b operands.
HA	A Home Agent (HA) orders read and write requests to a piece of coherent memory. The HA is implemented in the CHA logic.
HDCC	High Density Core Count
ICU	Instruction Cache Unit. Part of the processor core architecture.
IFU	Instruction Fetch Unit. Part of the processor core.
IIO	Integrated I/O Controller. An I/O controller that is integrated in the processor die.
IMC	Integrated Memory Controller. A Memory Controller that is integrated in the processor die.
Intel® 64 Technology	64-bit memory extensions to the IA-32 architecture. Further details on Intel® 64 architecture and programming model can be found at http://developer.intel.com/technology/intel64/ .
Intel® AVX	Intel® Advanced Vector Extensions promotes legacy 128-bit SIMD instruction sets that operate on XMM register set to use a "vector extension" (VEX) prefix and operates on 256-bit vector registers (YMM).
Intel® DLB	Intel® Dynamic Load Balancer (Intel® DLB)
Intel® DSA	Intel® Data Streaming Accelerator
Intel® IAA	Intel® In-Memory Analytics Accelerator
Intel Omni-Path	This is provided through a primary side connector on the package.
Intel® QAT	Intel® QuickAssist Technology is a platform solution designed to maximize the throughput of server data traffic across a broader range of configurations and server environments to achieve faster, scalable, and more reliable I/O.
Intel® TXT	Intel® Trusted Execution Technology
Intel® Turbo Boost Technology	A feature that opportunistically enables the processor to run a faster frequency. This results in increased performance of both single and multi-threaded applications.
Intel® UPI	Intel® Ultra Path Interconnect. A cache-coherent, link-based Interconnect specification for Intel® processors.
Intel® VT	Intel® Virtualization Technology. Processor Virtualization which, when used in conjunction with Virtual Machine Monitor software, enables multiple, robust independent software environments inside a single platform.
Intel® VT-d	Intel® Virtualization Technology (Intel® VT) for Directed I/O. Intel® VT-d is a hardware assist, under system software (Virtual Machine Manager or OS) control, for enabling I/O device Virtualization. Intel® VT-d also brings robust security by providing protection from errant DMAs by using DMA remapping, a key feature of Intel® VT-d.
Integrated Heat Spreader (IHS)	A component of the processor package used to enhance the thermal performance of the package. Component thermal solutions interface with the processor at the IHS surface.
continued...	

Term	Description
IOV	I/O Virtualization
IRP	I/O Request Packet (IRP) is the cache tracker for the IO M2 SSD
IVR	Integrated Voltage Regulation (IVR): The processor supports several integrated voltage regulators.
LLC	Last Level Cache
LRDIMM	Load Reduced Dual In-line Memory Module
LRU	Least Recently Used. A term used in conjunction with cache allocation policy.
M2M	Mesh to Memory. Logic in the IMC which interfaces the IMC to the mesh.
M2PCIE	The logic in the IIO modules which interface the modules to the mesh.
MCP	Multi-Chip Package
MESH	The on die interconnect that connects modules in the processor.
MESI	Modified/Exclusive/Shared/Invalid. States used in conjunction with cache coherency
MLC	Mid Level Cache
MSR	Model-Specific Register
NCTF	Non-Critical to Function: NCTF locations are typically redundant ground or non-critical reserved, so the loss of the solder joint continuity at end of life conditions will not affect the overall product functionality.
NID \ NodeID	Node ID (NID) or NodeID (NID). The processor implements up to 4- bits of NodeID (NID).
PBM	Power Budget Management
PCIe*	PCI Express*
Pcode	Pcode is microcode which is run on the dedicated microcontroller within the PCU.
PCU	Power Control Unit
PECI	Platform Environment Control Interface
Phit	The data transfer unit on Intel® UPI at the Physical layer is called a phit (physical unit). A Phit will be either 20 bits, or 8 bits depending on the number of active lanes.
Processor	Includes the 64-bit cores, uncore, I/Os and package
Processor Core	The term "processor core" refers to Si die itself which can contain multiple execution cores. Each execution core has an instruction cache and data cache and Mid-Level Cache (MLC) cache. All execution cores share the L3 cache.
RAC	Read Access Control
Rank	A unit of DRAM corresponding four to eight devices in parallel, ignoring ECC. These devices are usually, but not always, mounted on a single side of a DDR5 DIMM.
RAPL	Running Average Power Limit
RDIMM	Registered Dual In-line Memory Module
RTID	Request Transaction IDs are credits issued by the CHA to track outstanding transaction, and the RTIDs allocated to a CHA are topology dependent.
SCF	Scalable Coherent Fabric
continued...	

Term	Description
SCI	System Control Interrupt. Used in ACPI protocol.
SKU	Stock Keeping Unit (SKU) is a subset of a processor type with specific features, electrical, power and thermal specifications. Not all features are supported on all SKUs. A SKU is based on specific use condition assumption.
RsvdP	Reserved and Preserved
RsvdZ	Reserved and Zero
SMBus	System Management Bus. A two-wire interface through which simple system and power management related devices can communicate with the rest of the system.
Storage Conditions	A non-operational state. The processor may be installed in a platform, in a tray, or loose. Processors may be sealed in packaging or exposed to free air. Under these conditions, processor landings should not be connected to any supply voltages, have any I/Os biased or receive any clocks. Upon exposure to "free air" (that is, unsealed packaging or a device removed from packaging material) the processor must be handled in accordance with moisture sensitivity labeling (MSL) as indicated on the packaging material.
TAC	Thermal Averaging Constant
TDP	Thermal Design Power
TSOD	Temperature Sensor On DIMM
UDIMM	Unbuffered Dual In-line Memory Module
Unit Interval	Signaling convention that is binary and unidirectional. In this binary signaling, one bit is sent for every edge of the forwarded clock, whether it be a rising edge or a falling edge. If a number of edges are collected at instances $t_1, t_2, t_n, \dots, t_k$ then the UI at instance "n" is defined as: $UI_n = t_n - t_{n-1}$.
VCCIN	Primary voltage input to the voltage regulators integrated into the processor.
VSS	Processor ground
VSSA	System agent supply for Intel® UPI and PCIe*
VCCIO	I/O voltage supply input
VCCD	DDR power rail
WAC	Write Access Control
x1, x4, x8, x16	Refers to a link or port with one, four, eight, and sixteen physical lane(s).

2.0 Registers Overview

This document provides the register information for the processor.

NOTE

This datasheet covers multiple product types and SKUs. Some register and field descriptions apply only to the specific product types and SKUs. Not all features specific for each processor type have been explicitly identified in this specification, and not all features documented are available for all SKUs. See the specific product specifications for product type feature support, and refer to marketing collateral for SKU specific attributes.

Default values may vary based on processor type and SKU. In many cases these are read-only register fields that provide feature support visibility to firmware. Firmware should not rely on default values found in this document and must verify values by reading register values from the installed processor.

There are two bus ranges supported for the uncore [1-0]. The bus number is configurable in the Ubox, CSR CPUBUSNO_CFG (Bus: 30: Device: 0: Function: 2, Offset: D0h).

This document uses bus notation: Bus: 30 is Uncore Bus 0, and Bus: 31 is Uncore Bus 1. The bus number for CPUBUSNO0 is 0 and CPUBUSNO1 is 1.

2.1 Register Spaces and Structure

Configuration and Status Registers (CSR)

CSRs are accessed via the configuration transaction mechanism defined in the PCI specification using the Bus, Device, Function number concept to address a specific device's configuration space. All configuration register accesses are accessed over Message Channels through the UBox, but can come from processor cores or via PECl or JTAG interfaces.

Configuration registers can be read or written in byte, WORD (16-bit), or DWORD (32-bit) quantities. Accesses larger than a DWORD to PCI Express configuration space result in unexpected behavior. All multi-byte numeric fields use "little-endian" ordering (lower addresses contain the least significant parts of the field). For a listing of PCI devices with device ID, refer to the Device List Map in the preceding section.

Memory-Mapped I/O (MMIO) Registers

The PCI standard provides CSRs as well as registers which reside in memory-mapped space. For PCI devices, this is typically where the majority of the driver programming occurs and the specific register definitions and characteristics are provided by the device manufacturer. Access to these registers is typically accomplished via CPU reads and writes to non-coherent (UC) or write combining (WC) space. Reads and writes to memory-mapped registers can be accomplished with 1, 2, 4, or 8 byte transactions.

Devices support MMIO space within their IP blocks. This adds support to the Ubox to allow IP blocks to contain MMIO registers that are accessible via the GPSB interface. This permits IP blocks containing MMIO registers to be directly connected (without change) to the internal fabric accessed by proxy through the Ubox.

MMIO Address Space Allocation

Even though BIOS allocates a contiguous 8 MB region to each socket via SAD programming, the Ubox sub-divides this address range into separate BARs. BARs are not relocatable by the operating system.

The Ubox hosted BARs do not follow standard PCIe BAR semantics or address offsets. Having the BARs located outside of the PCIe standard location adds the requirement that BIOS be aware of the location of these registers and will not be determined through device enumeration. BIOS requires access to the MMIO registers across the chip prior to device enumeration.

NOTE

The terms "MMIO" and "MEM" are interchangeable.

The following table summarizes the Base Address Register (BAR) for most IP blocks.

Table 2. Base Address Register (BAR) for Functional Blocks

Chapter	Functional Block Registers	BAR
Accelerators Registers	Intel® DSA	BAR0
Accelerators Registers	Intel® DLB Physical Function	FUNC_BAR
Accelerators Registers	Intel® DLB Virtual Function	SRIOV_CAP_FUNC_BAR
Host I/O Processor (HIOP) Registers	Intel® VMD	MEMBAR2
Host I/O Processor (HIOP) Registers	Intel® VT-d	VTDBAR
Integrated Error Handler (IEH) Registers	Error Logger	SCF BASE
I3C* Registers	I3C	SCF BASE
Compute Express Link* (CXL*) Cache Memory Controller (CXL.CM) Registers	CXLCM Port Base	EXPPTMBAR
Compute Express Link* (CXL*) Cache Memory Controller (CXL.CM) Registers	Port Encryption, RDT	CXLRBBAR
Compute Express Link* (CXL*) Cache Memory Controller (CXL.CM) Registers	Bridge to Compute Express Link (B2CXL)	SCF BASE
PCI Express* (PCIe*) and CXL IO (CXL.io) Registers	Express Port (EXPPTMBAR)	EXPPTMBAR
PCI Express* (PCIe*) and CXL IO (CXL.io) Registers	Integrity Data Encryption (IDE)	KEY_CONFIG_BAR
PCI Express* (PCIe*) and CXL IO (CXL.io) Registers	Bridge to High Bandwidth Order Traffic (B2HOT)	SCF BASE
Out-Of-Band Management Service Module (OOBMSM) Registers	Management Service Module	PCI_CFG_BAR0
Intel® Trace Hub (Intel® TH) Registers	Intel® Trace Hub	MTB_BAR
continued...		

Chapter	Functional Block Registers	BAR
Secured Startup Services Module (S3M) Registers	ACPI	PWRMBASE
Secured Startup Services Module (S3M) Registers	eSPI	SBREG_BAR + CS0/1#
Secured Startup Services Module (S3M) Registers	ILMI	SBREG_BAR
Secured Startup Services Module (S3M) Registers	SMBus	SMBMBAR
Secured Startup Services Module (S3M) Registers	SPI	BIOS_SPI_BAR0
Secured Startup Services Module (S3M) Registers	UART	MEMBAR2
Memory Controller (MC) Registers	Memory Channel (MCCHAN)	SCF BASE
DDR Physical Layer (DDRIO) Registers	DDRIO	SCF BASE
Memory Security Engine (MSE) Registers	MSE	SCF BASE
Caching and Home Agent (CHA) Registers	Bridge to Intra-Die Interconnect (B2IDI)	SCF BASE
Configuration Agent (UBOX) Registers	B2UBOX	SCF BASE

Model Specific Registers (MSRs)

All MSRs are 64-bit and accessed as a Qword quantity through the RDMSR and WRMSR instructions. See the Model Specific Registers (MSRs) chapter in this document for further details.

See also the "Model-Specific Registers" section in the [Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 3A](#) for more information.

2.2 Access Types

Access Types are constructed from a base attribute and any number of additional modifier attributes. The full list of access type values is provided in the following table.

Table 3. Field Access Types

Access Type	Behavior	Description
RO	Read-Only	Read returns constant value.
RO/V	Read-Only, Variant	Variant indicates hardware updates the value.
RO/P	Read-Only, Sticky	Reset on PowerGood reset only (not functional reset).
RO/V/P	Read-Only, Variant with Sticky	Same as RO/V, but also Sticky (reset on PowerGood only).
RO/C/V	Read-Only, Cleared by Read, Variant	Same as RO/C, but also Variant (hardware sets value).
RW	Read-Write	Readable and writable.
RW/1C	Read-Write, Write 1 to Clear	Write 1 to clear.
RW/1C/P	Read-Write, Write 1 to Clear, Sticky	Same as RW/1C, but also Sticky (reset on PowerGood only).
RW/1C/V	Read-Write, Write 1 to Clear, Variant	Same as RW/1C, but also Variant (hardware sets value).
continued...		

Access Type	Behavior	Description
RW/1C/V/P	Read-Write, Write 1 to Clear, Variant with Sticky	Same as RW/1C/V, but also Sticky (reset on PowerGood only).
RW/1L	Read-Write, Write 1 to Lock (bitwise)	Writing 1 to any bit of the field will lock that respective bit only (other bits that are still 0 can still be written).
RW/1S	Read-Write, Write 1 to Set	Writing 1 sets bits. Writing 0 has no effect.
RW/1S/V	Read-Write, Write 1 to Set, Variant	Same as RW/1S, but also Variant (hardware clears value). Example: SW written "go" bit that HW clears upon completion.
RW/1S/V/P	Read-Write, Write 1 to Set, Variant with Sticky	Same as RW/1S/V, but also Sticky (reset on PowerGood only).
RW/C	Read-Write, Cleared by Write	Writing clears the value.
RW/L	Read-Write, Lock	Behaves as RW if the corresponding LockKeyField is "off", and RO if the LockKeyField is "on".
RW/O	Read-Write, Write-Once	Only writable once. After first write, behaves as RO.
RW/O/L	Read-Write, Write-Once, Lock	Same as RW/O, but also Lockable.
RW/O/P	Read-Write, Write-Once, Sticky	Same as RW/O, but also Sticky (reset on PowerGood only).
RW/P	Read-Write, Sticky	Same as RW, but also Sticky (reset on PowerGood only).
RW/P/L	Read-Write, Lock with Sticky	Same as RW/P, but also Lockable.
RW/V	Read-Write, Variant	Same as RW, but also Variant (hardware updates the value).
RW/V/L	Read-Write, Variant with Lock	Same as RW/V, but also Lockable.
RW/V/P	Read-Write, Variant with Sticky	Same as RW/V, but also Sticky (reset on PowerGood only).
RW/V/P/L	Read-Write, Variant with Lock, Sticky	Same as RW/V/P, but also Lockable.
WO	Write-Only	Writes update hardware value. Reads return "reset" value of the field.
WO/P	Write-Only, Sticky	Same as WO, but also Sticky (reset on PowerGood only).
WOC	Write-Only, Write will invert	Same as WO, but write causes field value to invert (written value is of no importance).

2.3 Configuration Rules

Unimplemented Devices, Functions, and Registers

Configuration reads to unimplemented functions and devices will return all ones emulating a controller abort response. There is no asynchronous error reporting that happens when a configuration read controller aborts. Configuration writes to unimplemented functions and devices will return a normal response.

Software must not attempt or rely on reads or writes to unimplemented registers or register bits. Unimplemented registers return all zeroes when read. Writes to unimplemented registers are ignored. For configuration writes to these registers (require a completion), the completion is returned with a normal completion status (not controller-aborted).

2.4 Notational Conventions

Hexadecimal and Binary Numbers

Base 16 numbers are represented by a string of hexadecimal digits followed by the character h (for example, F82Eh). A hexadecimal digit is a character from the following set: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, and F. Hexadecimal numbers can also be shown with standard leading characters "0x" (for example, 0x2A).

Base 2 (binary) numbers are represented by a string of 1s and 0s, sometimes followed by the character b (for example, 101b). The "b" designation is only used in situations where confusion as to the type of the number might arise.

Base 10 numbers are represented by a string of decimal digits followed by the character d (for example, 23d). The "d" designation is only used in situations where confusion as to the type of the number might arise.

3.0 Intel® Ultra Path Interconnect (Intel® UPI) Registers

The Intel® UPI module is the coherent communication interface between processors. 4th Gen Intel® Xeon® Processor Scalable Family, Codename Sapphire Rapids implements either two, three, or four Intel® UPI links based on processor type.

The Intel® UPI Registers are implemented in the following Bus, Device, Functions:

- Bus: B(30), Device: 1-2-3-4, Function: 0 (UPI Misc)

NOTE

For additional details on the Intel® UPI, see *4th Gen Intel® Xeon® Processor Scalable Family, Codename Sapphire Rapids, Architecture Specification*, chapter "Intel® Ultra Path Interconnect (Intel® UPI) Functional Description."

3.1 Bus: B(30), Device: 1-2-3-4, Function: 0 (UPI Misc)

This chapter documents the Intel® UPI Misc registers in Bus (30), Device 1-2-3-4, Function 0.

- Device 1 applies to Intel® UPI Port 0
- Device 2 applies to Intel® UPI Port 1
- Device 3 applies to Intel® UPI Port 2
- Device 4 applies to Intel® UPI Port 3

3.1.1 Vendor ID (VID) — Offset 0h

PCI Vendor ID Register

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:1-4, F:0] + 0h	8086h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	1

Bit Range	Default & Access	Field Name (ID): Description
15:0	8086h RO	VENDOR_IDENTIFICATION_NUMBER: The value is assigned by PCI-SIG to Intel.

3.1.2 Device ID (DID) — Offset 2h

PCI Device Identification Number

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:1-4, F:0] + 2h	3240h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	1

Bit Range	Default & Access	Field Name (ID): Description
15:0	3240h RO	DEVICE_IDENTIFICATION_NUMBER: The value is assigned by each IP/function owner as a unique identifier.

3.1.3 PCI Status (PCISTS) — Offset 6h

PCI Status Register

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:1-4, F:0] + 6h	0010h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	1

Bit Range	Default & Access	Field Name (ID): Description
15	0h RO	DETECTED_PARITY_ERROR: This bit is set when the device receives a packet on the primary side with an uncorrectable data error (including a packet with poison bit set) or an uncorrectable address/control parity error. The setting of this bit is regardless of the Parity Error Response bit (PERRE) in the PCICMD register.
14	0h RO	SIGNALED_SYSTEM_ERROR: Hardwired to 0
13	0h RO	RECEIVED_MASTER_ABORT: Hardwired to 0
12	0h RO	RECEIVED_TARGET_ABORT: Hardwired to 0
11	0h RO	SIGNALED_TARGET_ABORT: Hardwired to 0
10:9	0h RO	DEVSEL_TIMING: Not applicable to PCI Express. Hardwired to 0.
8	0h RO	MASTER_DATA_PARITY_ERROR: Hardwired to 0
7	0h RO	FAST_BACK_TO_BACK: Not applicable to PCI Express. Hardwired to 0.
continued...		

Bit Range	Default & Access	Field Name (ID): Description
6	0h RO	Reserved
5	0h RO	X66MHZ_CAPABLE: Not applicable to PCI Express. Hardwired to 0.
4	1h RO	CAPABILITIES_LIST: This bit indicates the presence of a capabilities list structure. When set to 1, indicates the register at 34h provides an offset into the function.
3	0h RO	INTX_STATUS: Reflects the state of the INTA# signal at the input of the enable/disable circuit. This bit is set by HW to 1 when the INTA# is asserted. This bit is reset by HW to 0 after the interrupt is cleared (independent of the state of the Interrupt Disable bit in the PCICMD register). Hardwired to 0 on the processor
2:0	0h RO	Reserved

3.1.4 Revision ID and Class Code Register (RID_CCR) — Offset 8h

PCIe header Revision ID register and Class Code register

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:1-4, F:0] + 8h	08800000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	1

Bit Range	Default & Access	Field Name (ID): Description
31:24	08h RO	BASE_CLASS: Base class code
23:16	80h RO	SUB_CLASS: Sub class code
15:8	00h RO	REGISTER_LEVEL_PROGRAMMING_INTERFACE: Register Level Programming Interface
7:0	00h RO/V	REVISION_ID: Reflects the Uncore Revision ID after reset. Reflects the Compatibility Revision ID if BIOS updates.

3.1.5 Cache Line Size Register (CLSR) — Offset Ch

PCI Cache Line Size Register

Type	Size	Offset	Default
PCI	8 bit	[B:30, D:1-4, F:0] + Ch	00h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	1

Bit Range	Default & Access	Field Name (ID): Description
7:0	00h RO	CACHELINE_SIZE: Size of Cacheline

3.1.6 Header Type (HDR) — Offset Eh

PCI Header Type

Type	Size	Offset	Default
PCI	8 bit	[B:30, D:1-4, F:0] + Eh	80h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	1

Bit Range	Default & Access	Field Name (ID): Description
7	1h RO	MULTI_FUNCTION_DEVICE: This bit defaults to 1b since all these devices are multi-function
6:0	00h RO	CONFIGURATION_LAYOUT: Type 0 header

3.1.7 Subsystem Vendor ID (SVID) — Offset 2Ch

Subsystem Vendor ID Register

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:1-4, F:0] + 2Ch	8086h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
RW	RW	RW	1

Bit Range	Default & Access	Field Name (ID): Description
15:0	8086h RW/O	SUBSYSTEM_VENDOR_IDENTIFICATION_NUMBER: The default value specifies Intel® but can be set to any value once after reset.

3.1.8 Subsystem ID (SDID) — Offset 2Eh

Subsystem ID Register

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:1-4, F:0] + 2Eh	0000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
RW	RW	RW	1

Bit Range	Default & Access	Field Name (ID): Description
15:0	0000h RW/O	SUBSYSTEM_IDENTIFICATION_NUMBER: Assigned by the subsystem vendor to uniquely identify the subsystem

3.1.9 Capability Pointer (CAPPTR) — Offset 34h

PCI Capability Pointer Register

Type	Size	Offset	Default
PCI	8 bit	[B:30, D:1-4, F:0] + 34h	40h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	1

Bit Range	Default & Access	Field Name (ID): Description
7:0	40h RO	CAPABILITY_POINTER: Points to the first capability structure for the device which is the PCIe capability.

3.1.10 PCI Express Capability (PXPCAP) — Offset 40h

PCI Express Capability

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:1-4, F:0] + 40h	00920010h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	1

Bit Range	Default & Access	Field Name (ID): Description
31:30	0h RO	Reserved
29:25	00h RO	INTERRUPT_MESSAGE_NUMBER: Not valid for this device, since the device does not generate interrupts
24	0h RO	SLOT_IMPLEMENTED: Not valid for PCIe integrated Endpoints
23:20	9h RO	DEVICE_PORT_TYPE: Device type is Root Complex Integrated Endpoint
continued...		

Bit Range	Default & Access	Field Name (ID): Description
19:16	2h RO	CAPABILITY_VERSION: PCI Express Capability is Compliant with Version 2.0 of the PCI Express Spec.
15:8	00h RO	NEXT_PTR: Pointer to the next capability. Set to 0 to indicate there are no more capability structures.
7:0	10h RO	CAPABILITY_ID: Identifies the PCI Express capability assigned by PCI-SIG.

3.1.11 Miscellaneous Clock Control (KTIMISCCLKCTRL) — Offset ECh

Intel® UPI Miscellaneous Clock Control Register

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:1-4, F:0] + ECh	00000020h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	RW	R

Bit Range	Default & Access	Field Name (ID): Description
31:6	0h RO	Reserved
5	1h RW/P	KTI_MISC_CSR_CLKGAT_EN: clock gate enable for UPI misc CSR and CSR chain
4:2	0h RO	Reserved
1	0h RW/P	DISABLE_IOV_CHAIN: Obsolete, HSX doesn't have 8-bit IOV csr critter chain. Recommend remove this field in future project.
0	0h RW/P	DISABLE_AGENT_CHAIN: Obsolete, SKX doesn't have an agent csr critter chain. Recommend remove this field in future project.

3.1.12 PCI Express Enhanced Capability (PXPENHCAP) — Offset 100h

PCI Express Enhanced Capability

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:1-4, F:0] + 100h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	1

Bit Range	Default & Access	Field Name (ID): Description
31:20	000h RO	NEXT_CAPABILITY_OFFSET: Pointer to the next capability in the enhanced configuration space. Set to 0 to indicate there are no more capability structures.
19:16	0h RO	CAPABILITY_VERSION: There is no capability at this location.
15:0	0000h RO	CAPABILITY_ID: There is no capability at this location

3.1.13 kti misc privc0 (KTIMISCPRIVC0) — Offset 1D0h

ktimiscprivc0

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:1-4, F:0] + 1D0h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	R	R

Bit Range	Default & Access	Field Name (ID): Description
31	0h RW/P/L	FORCE_CFGWRTLCK_NONPOSTED: Force all CfgWrTLck requests targeting UPI agent to be treated as non-posted. The purpose is to revert to JKT/IVT/HSX QPI behavior (SKX HSD 298055). Locked by: KTIMISCFUSE.PRIVDBGLCK
30:4	0h RO	Reserved
3	0h RW/P/L	PN_TERM_DISABLE: Disable PN termination feature and go back to JKT. (Need to set bit[12] to override the fuse) Locked by: KTIMISCFUSE.PRIVDBGLCK
2	0h RW/P/L	OVERRIDE_FUSE_PN_TERM_DISABLE: Override fuse setting and enable Bit13 programming to disable PN termination. Locked by: KTIMISCFUSE.PRIVDBGLCK
1	0h RW/P/L	TAP_CSR_OVERRIDE_EN: To enable some PLL CSR override function from minitap. Locked by: KTIMISCFUSE.PRIVDBGLCK
0	0h RW/P/L	MSGCH_GATECLK_DIS: If set disables clock gating on message channel endpoint and router Locked by: KTIMISCFUSE.PRIVDBGLCK

3.1.14 kti misc scratch cfg32 (KTIMISCSCRATCHCFG32) — Offset 304h

ktimiscscratchcfg32

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:1-4, F:0] + 304h	00000000h

Register Level Access:



BIOS Access	SMM Access	OS Access	Policy Group ID
RW	RW	RW	1

Bit Range	Default & Access	Field Name (ID): Description
31:16	0000h RW	SCRATCH_UP: Scratch CFG Upper 16-bits (non-sticky) register field.
15:0	0000h RW/P	SCRATCH_DN: Scratch CFG Lower 16-bits (sticky) register field.

4.0 Integrated Memory Controller (IMC) Registers

The 4th Gen Intel® Xeon® Processor Scalable Family, codename Sapphire Rapids implements four Integrated Memory Controllers (IMC). Each IMC is capable of controlling two DDR5 memory channels, and two DIMMs per channel.

The IMC Host Configuration Space Registers are implemented in Bus: B(30), Device: 12-13-14-15, Function: 0

- Device 12 applies to IMC 0
- Device 13 applies to IMC 1
- Device 14 applies to IMC 2
- Device 15 applies to IMC 3

NOTE

For additional details on the IMC, see 4th Gen Intel® Xeon® Processor Scalable Family, codename Sapphire Rapids Data Sheet Volume 1, chapter "Integrated Memory Controller (IMC) Functional Description."

4.1 Bus: B(30), Device: 12, Function: 0 (IMC0 Mesh to Mem Registers)

This chapter documents the IMC0 Mesh to Mem (MS2MEM) registers in Bus (30), Device 12, Function 0.

- IMC 0: Bus(30) Device 12 Function 0
- IMC 1: Bus(30) Device 13 Function 0
- IMC 2: Bus(30) Device 14 Function 0
- IMC 3: Bus(30) Device 15 Function 0

4.1.1 Vendor ID (VID_0_12_0_CFG) — Offset 0h

PCI Vendor ID Register

Type	Size	Offset	Default
PCI	16 bit	MEM0_BAR + 0h	8086h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	2

Bit Range	Default & Access	Field Name (ID): Description
15:0	8086h RO	VENDOR_IDENTIFICATION_NUMBER: The value is assigned by PCI-SIG to Intel.

4.1.2 Device ID (DID_0_12_0_CFG) — Offset 2h

PCI Device Identification Number

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:12, F:0] + 2h	324Ah

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	2

Bit Range	Default & Access	Field Name (ID): Description
15:0	324Ah RO	DEVICE_IDENTIFICATION_NUMBER: The value is assigned by each IP/function owner as a unique identifier.

4.1.3 Header Type (HDR_0_12_0_CFG) — Offset Eh

PCI Header Type

Type	Size	Offset	Default
PCI	8 bit	[B:30, D:12, F:0] + Eh	80h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	2

Bit Range	Default & Access	Field Name (ID): Description
7	1h RO	MULTI_FUNCTION_DEVICE: This bit defaults to 1b since all these devices are multi-function
6:0	00h RO	CONFIGURATION_LAYOUT: Type 0 header

4.1.4 PCI Express Capability (PXPCAP_0_12_0_CFG) — Offset 40h

PCI Express Capability

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:12, F:0] + 40h	00920010h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	2

Bit Range	Default & Access	Field Name (ID): Description
31:30	0h RO	Reserved
29:25	00h RO	INTERRUPT_MESSAGE_NUMBER: Not valid for this device, since the device does not generate interrupts
24	0h RO	SLOT_IMPLEMENTED: Not valid for PCIe integrated Endpoints
23:20	9h RO	DEVICE_PORT_TYPE: Device type is Root Complex Integrated Endpoint
19:16	2h RO	CAPABILITY_VERSION: PCI Express Capability is Compliant with Version 2.0 of the PCI Express Spec.
15:8	00h RO	NEXT_PTR: Pointer to the next capability. Set to 0 to indicate there are no more capability structures.
7:0	10h RO	CAPABILITY_ID: Identifies the PCI Express capability assigned by PCI-SIG.

4.1.5 PCI Express Enhanced Capability (PXPENHCAP_0_12_0_CFG) — Offset 100h

PCI Express Enhanced Capability

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:12, F:0] + 100h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	2

Bit Range	Default & Access	Field Name (ID): Description
31:20	000h RO	NEXT_CAPABILITY_OFFSET: Pointer to the next capability in the enhanced configuration space. Set to 0 to indicate there are no more capability structures.
19:16	0h RO	CAPABILITY_VERSION: There is no capability at this location.
15:0	0000h RO	CAPABILITY_ID: There is no capability at this location

4.1.6 TAD Read (TAD_RD) — Offset 238h

TAD mapping:

This register allows indirect access to read each of the 20 internal TAD table entries.

The reset value for all the fields in the internal TAD table is 0, except for AddressLimit, which resets to all-1

This internal TAD table gets reset on a cold reset (pwr pulse) only.

To read a TAD entry:

1. Configure the TAD entry/index field and at the same time set the read-enable bit to 1.

2. Now read this register which will have its fields populated with the information of the particular TAD entry.

- When reading this register, check that TadId matches the value which was written in step 2.

If the value does not match, then another software agent wrote this register between steps 1 and 2, and the read results are not valid. Steps 1-2 should be repeated.

Type	Size	Offset	Default
PCI	64 bit	[B:30, D:12, F:0] + 238h	0000000000000000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
RW	RW	RW	2

Bit Range	Default & Access	Field Name (ID): Description
63:52	0h RO	Reserved
51:26	0000000h RW/V/P	ADDRESSLIMIT: Limit address Specifies the upper limit (inclusive) for address[51:26] (64MB aligned) to be part of this memory region. An address hits a particular TAD entry if all of the following conditions are true (where number): TadVld[N] == 1 address[51:26] <= AddressLimit[N] address[51:26] > AddressLimit[N-1] (only for N>0) System firmware must program the TAD AddressLimit fields such that AddressLimit[N] > AddressLimit[N-1] for all N>0 with TadVld[N]==1.
25	0h RO	Reserved
24	0h RW/V/P	DEDUPVLD: DedupRegion (DedupVld) This TAD entry represents a region for which CHA uses Dedup IP to access memory Setting this bit would mean m2mem would use a dedicated VC to issue transactions to the memory controller which will be routed via the Dedup IP to MC
23	0h RW/V/P	LOWBW: Low Bandwidth Memory (LowBW): This TAD entry represents a region for which CHA uses a low bandwidth credit.
continued...		

Bit Range	Default & Access	Field Name (ID): Description
		Setting this bit causes m2mem to set the ddrt_cdt bit on BL credit returns for writes targeting this region.
22	0h RW/V/P	FRCNPWR: Force Non-Posted Writes (FrcNPWr): This TAD entry represents region that requires NP write semantics on CMI. Only applies to the far-memory region if this is a 2LM region.
21	0h RW/V/P	SECONDARY1ST: Read secondary mirror channel first (Secondary1st): For this memory region, read secondary mirror channel first. Used in mirroring to prevent (non error flow) accesses to a primary mirror channel for a bad primary region/address. Note, if an error flow does get invoked because of an error on secondary for an access to this region then the primary storage will still be accessed (when channel not failed over).
20	0h RW/V/P	MIRROR: Mirror region: This TAD entry represents a mirroring region.
19	0h RW/V/P	NMCACHEABLEVLD: Near-Memory Cacheable Valid (NmCacheableVld): This TAD entry represents a near-memory cached region, i.e. not a 1LM region.
18	0h RW/V/P	DDR4: DDR4: This TAD entry represents near-memory at NMC or 1LM DDR4. In non-Xtile config: this just represents the 1LM DDR4 region. In Xtile config: this represents NM access at the near-memory controller (or DDR4 access at the FMC). BIOS to calculate as follows:
17	0h RW/V/P	BLKVLD: Block Region Valid (BlkVld): This TAD entry represents a block region.
16	0h RW/V/P	PMEMVLD: PMem Region Valid (PmemVld): This TAD entry represents a persistent memory region.
15	0h RW/V/P	NONPERSISTENTFM: Non-Persistent Far Memory (NonPersistentFm): This TAD entry represents a near-memory cached region in which the far-memory is non-persistent. This bit only takes effect when NmCacheableVld for this TAD entry is set, and should only be set if the far-memory being cached by near-memory is not persistent (i.e. is not DDRT). NOTE: This bit only works if the far-memory is on a different MC tile than the near-memory. It is currently not supported where a single MC has both near-memory AND far-memory and the far-memory being cached is not persistent/DDRT
14:12	0h RO	Reserved
11:8	0h RW/V/P	DDRTADID: DDR TAD entry (DDRTadId): DDR TAD entry associated with this TAD table entry. M2mem puts no restrictions on the values here, but this field should be programmed consistent with the channel

continued...

Bit Range	Default & Access	Field Name (ID): Description
		the channel, hence the current legal range is 0-11. DDR4 can only use the first 8 TAD entries, so the legal range is 0-7 for regions with the DDR4 bit set. HBM should only use TAD entry 0.
7	0h RW/V/P	TADVLD: TAD Table Entry Valid (TadVld): Specifies whether this TAD entry in the TAD table is valid. It is illegal to program TadVld[N]==1 if TadVld[N-1]==0 (for N>0).
6	0h RW/V	TADRDN: TAD Read (TadRdEn): Read the TAD table entry. The read information will get stored in this register. Hardware will read the TAD table entry when seeing a zero-to-one transition on this bitfield. Hardware will reset this bit to 0 the cycle after it is written to 1.
5	0h RO	Reserved
4:0	00h RW/P	TADID: TAD Table SAD Entry/Index (TadId): Specifies what TAD entry in the TAD table to read indirectly through this register. There are 20 TAD entries so the legal range is 0-19.

4.1.7 TAD Write (TAD_WR) — Offset 240h

TAD mapping:

This register allows indirect access to write each of the 20 internal TAD table entries.

The reset value for all the fields in the internal TAD table is 0, except for AddressLimit, which resets to all-1

This internal TAD table gets reset on a cold reset (pwr pulse) only.

To write a TAD entry:

1. Configure all fields and at the same time set the write-enable bit to 1.

Type	Size	Offset	Default
PCI	64 bit	[B:30, D:12, F:0] + 240h	0000000000000000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
RW	RW	R	1

Bit Range	Default & Access	Field Name (ID): Description
63:52	0h RO	Reserved
51:26	0000000h RW/P	ADDRESSLIMIT: Limit address Specifies the upper limit (inclusive) for address[51:26] (64MB aligned) to be part of this memory region.
continued...		

Bit Range	Default & Access	Field Name (ID): Description
		An address hits a particular TAD entry if all of the following conditions are true (where number): TadVld[N] == 1 address[51:26] <= AddressLimit[N] address[51:26] > AddressLimit[N-1] (only for N>0) System firmware must program the TAD AddressLimit fields such that AddressLimit[N] > AddressLimit[N-1] for all N>0 with TadVld[N]==1.
25	0h RO	Reserved
24	0h RW/P	DEDUPVLD: DedupRegion (DedupVld) This TAD entry represents a region for which CHA uses Dedup IP to access memory Setting this bit would mean m2mem would use a dedicated VC to issue transactions to the memory controller which will be routed via the Dedup IP to MC
23	0h RW/P	LOWBW: Low Bandwidth Memory (LowBW): This TAD entry represents a region for which CHA uses a low bandwidth credit. Setting this bit causes m2mem to set the ddr_t_cdt bit on BL credit returns for writes targeting this region.
22	0h RW/P	FRCNPWR: Force Non-Posted Writes (FrcNPWr): This TAD entry represents region that requires NP write semantics on CMI. Only applies to the far-memory region if this is a 2LM region.
21	0h RW/P	SECONDARY1ST: Read secondary mirror channel first (Secondary1st): For this memory region, read secondary mirror channel first. Used in mirroring to prevent (non error flow) accesses to a primary mirror channel for a bad primary region/address. Note, if an error flow does get invoked because of an error on secondary for an access to this region then the primary storage will still be accessed (when channel not failed over).
20	0h RW/P	MIRROR: Mirror region: This TAD entry represents a mirroring region.
19	0h RW/P	NMCACHEABLEVLD: Near-Memory Cacheable Valid (NmCacheableVld): This TAD entry represents a near-memory cached region, that is, not a 1LM region.
18	0h RW/P	DDR4: DDR4: This TAD entry represents near-memory at NMC or 1LM DDR4. In non-Xtile config: this just represents the 1LM DDR4 region. In Xtile config: this represents NM access at the near-memory controller (or DDR4 access at the FMC). BIOS to calculate as follows:
17	0h RW/P	BLKVLD: Block Region Valid (BlkVld): This TAD entry represents a block region.
16	0h RW/P	PMEMVLD: PMem Region Valid (PmemVld): This TAD entry represents a persistent memory region.
continued...		

Bit Range	Default & Access	Field Name (ID): Description
15	0h RW/P	NONPERSISTENTFM: Non-Persistent Far Memory (NonPersistentFm): This TAD entry represents a near-memory cached region in which the far-memory is non-persistent. This bit only takes effect when NmCacheableVld for this TAD entry is set, and should only be set if the far-memory being cached by near-memory is not persistent (that is, is not DDRT). <i>Note:</i> This bit only works if the far-memory is on a different MC tile than the near-memory. It is currently not supported where a single MC has both near-memory AND far-memory and the far-memory being cached is not persistent/DDRT
14:12	0h RO	Reserved
11:8	0h RW/P	DDRTADID: DDR TAD entry (DDRTadId): DDR TAD entry associated with this TAD table entry. M2mem puts no restrictions on the values here, but this field should be programmed consistent with the channel the channel, hence the current legal range is 0-11. DDR4 can only use the first 8 TAD entries, so the legal range is 0-7 for regions with the DDR4 bit set. HBM should only use TAD entry 0.
7	0h RW/P	TADVLD: TAD Table Entry Valid (TadVld): Specifies whether this TAD entry in the TAD table is valid. It is illegal to program TadVld[N]=1 if TadVld[N-1]=0 (for N>0).
6	0h RW/V	TADWREN: TAD Write (TadWrEn): Write the TAD table entry with the information contained in this register. Hardware will write the TAD table entry when seeing a zero-to-one transition on this bitfield. Hardware will reset this bit to 0 the cycle after it is written to 1.
5	0h RO	Reserved
4:0	00h RW/P	TADID: TAD Table SAD Entry/Index (TadId): Specifies what TAD entry in the TAD table to write indirectly through this register. There are 20 TAD entries so the legal range is 0-19.

4.1.8 Mode register (MODE) — Offset 318h

Configuration register for the main m2m modes.

Important configuration restrictions:

- Isoch is not supported by the m2m if DDRT is present on a channel.
- Isoch is not supported if any channel gets serviced by FNV.
- Mirroring and Isoch mode are mutually exclusive.

Whenever mirroring is enabled on a mesh2mem, then also following bits need to get set on that mesh2mem:

- Defeatures0[IngBypDis]=1 (functional requirement)

- Sysfeatures0[PrefDisable]=1 (functional requirement)

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:12, F:0] + 318h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	R	R

Bit Range	Default & Access	Field Name (ID): Description
31:14	0h RO	Reserved
13	0h RW	NMCMODE: Near Memory Controller mode: When set, this indicates that the M2M is in an HBM configuration where the HBMtile is configured to handle route through or xtile 2lm NM transactions This should only be set on the M2M in the HBM tile. Do not set this if not in 2LM mode or not an HBM product where cross-tile 2LM or route through flows are applicable. 0: This M2M is not a HBM controller in cross-tile 2LM or route through 1: This M2M is a HBM controller in cross-tile 2LM or route through.
12	0h RW	FMCMODE: Far Memory Controller mode: When set, this indicates that the M2M is in an HBM configuration where 2LM is enabled and this Memory controller handles the Far Memory(DDR). This should only be set on the M2M in the DDR tile, and not on the M2M in the HBM tile. Do not set this if not in 2LM mode or not an HBM product where cross-tile 2LM flows are applicable. 0: This M2M is not a Far Memory controller in cross-tile 2LM. 1: This M2M is a Far Memory controller in cross-tile 2LM.
11	0h RW/P	ADDDC: Unused: available for ECO if necessary.
10	0h RW/P	Reserved
9	0h RW/P	MIRRORDDR4: DDR4 Mirroring (MirrorDDR4): Enable DDR4 mirroring. It is illegal to set this bit when Near Memory Caching is enabled. MmCapabilities[CapMirrorEn] needs to be 1 for mirroring.
8	0h RW/P	NMCACHING: Near Memory Caching (NMCaching): Enable Near Memory caching, that is, two level memory with the first level (that is, the near memory) having the capability to cache second level memory (that is, far memory). Setting this bit also implies there is a second memory level present. Furthermore, in this implementation, near memory accesses get scheduled by the DDR4 scheduler while the far memory gets serviced by the DDRT scheduler. <i>Note:</i> MmCapabilities[CapNMCachingEn] also needs to be 1 for Near Memory Caching.
7	0h RO	Reserved
6	0h RW/P	PMEM: Persistent Memory (PMem):

continued...

Bit Range	Default & Access	Field Name (ID): Description
		Enable persistent memory flows. Furthermore, MmCapabilities[CapPMemEn] needs to be 1 for PMem support.
5	0h RW/P	BLOCKREGION: Block Region (BlockRegion): Enable block region memory flows. Furthermore, MmCapabilities[CapBlockRegionEn], Mode[PMem] and MmCapabilities[CapPMemEn] need to be 1 for block region support.
4	0h RO	Reserved
3	0h RW/P	MEE: Memory Encryption Engine (MEE): Memory Encryption Engine enable. Set this bit to 1 if transactions destined for the Memory Encryption Engine (MEE) might come in.
2	0h RW/P	DEDUP: Memory Deduplication Engine (DEDUP) Memory Deduplication Engine enable. Set this bit to 1 if transactions destined for the Memory Deduplication Engine (Dedup) might come in.
1:0	0h RO	Reserved

4.1.9 UPI Routing Config (UPIROUTINGCONFIG) — Offset 320h

Selects the index bits used for indexing into the UpiRoutingTable.LogicalUpiTarget. The register supports defining up to 16 different indexing/interleave modes to UPI. Currently only 3 modes are used, the remaining are reserved for future.

Type	Size	Offset	Default
PCI	32 bit	MEM0_BAR + 320h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
RW	RW	R	1

Bit Range	Default & Access	Field Name (ID): Description
31:4	0h RO	Reserved
3:0	0h RW	UPIINTERLEAVEMODE: 0 : Index = {NID[2:0], PA[12]^PA[6]} (interleaves 1 to 2 UPI links between 8 nodes) 1 : Index = {NID[1:0], MOD3[1:0]} (interleaves 1 to 3 UPI links between 4 nodes) 2 : Index = {NID[1:0], PA[13:12]^PA[7:6]} (interleaves 1 to 4 UPI links between 4 nodes) 3-15 : RSVD

4.1.10 UPI Routing Table (UPIROUTINGTABLE) — Offset 328h

The 48 bit vector defined in this register captures 16 UPI ID values (supporting up to 8 UPI links). The index bit scheme selected by UpiRoutingConfig.UpiInterleaveMode is used to index to this UPI Routing Table.

NOTE

If the number of UPI links supported between the nodes is less than the max number of UPI links supported by UpiRoutingConfig.UpiInterleaveMode, then the LogicalUpiTarget entries need to be replicated appropriately (that is, aliasing).

Type	Size	Offset	Default
PCI	64 bit	MEM0_BAR + 328h	0000000000000000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
RW	RW	R	1

Bit Range	Default & Access	Field Name (ID): Description
63:48	0h RO	Reserved
47:0	0000000000 00h RW	LOGICALUPITARGET: This 48 bit vector captures 16 UPI ID values, each ID being 3 bits (supporting up to 8 UPI links). The index bit scheme selected by UpiRoutingConfig.UpiInterleaveMode is used to index to this UPI Routing Table. Bits [2:0] correspond to the UPI ID selected by Index==0, bits[5:3] correspond to the UPI ID selected by Index==1, etc.

5.0 Configuration Agent (Ubox) Registers

The Ubox handles transactions such as register accesses, interrupt flows, lock flows, and events. The Ubox houses coordination for the performance architecture, and scratchpad and semaphore registers.

The Ubox Registers are implemented in the following Bus, Device, Functions:

- Bus: B(30), Device: 0, Function: 0 (EVNTS)
- Bus: B(30), Device: 0, Function: 2 (DECS)

5.1 Bus: B(30), Device: 0, Function: 0 (EVNTS)

This chapter documents the registers in Bus: 30, Device 0, Function 0.

5.1.1 Vendor ID (VID_0_0_0_CFG) — Offset 0h

PCI Vendor ID Register

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:0, F:0] + 0h	8086h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	2

Bit Range	Default & Access	Field Name (ID): Description
15:0	8086h RO	VENDOR_IDENTIFICATION_NUMBER: The value is assigned by PCI-SIG to Intel.

5.1.2 Device ID (DID_0_0_0_CFG) — Offset 2h

PCI Device Identification Number

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:0, F:0] + 2h	3250h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
R	R	R	2

Bit Range	Default & Access	Field Name (ID): Description
15:0	3250h	DEVICE_IDENTIFICATION_NUMBER:

Bit Range	Default & Access	Field Name (ID): Description
	RO/V	The value is assigned by each IP/function owner as a unique identifier.

5.1.3 The register type IntControl_CFG (INTCONTROL_CFG) — Offset 94h

Interrupt Configuration Register

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:0, F:0] + 94h	00000001h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	RW	R

Bit Range	Default & Access	Field Name (ID): Description
31:19	0h RO	Reserved
18	0h RW	LOGFLATCLUSTOVREN: 0 : IA32 Logical Flat or Cluster Mode bit is locked as Read only bit. 1 : IA32 Logical Flat or Cluster Mode bit may be written by SW, values written by xTPR update are ignored. For one time override of the IA32 Logical Flat or Cluster Mode value, return this bit to its default state after the bit is changed. Leaving this bit as 1 will prevent automatic update of the filter.
17	0h RW/V	LOGFLTCLUSTMOD: Set by bios to indicate if the OS is running logical flat or logical cluster mode. This bit can also be updated by IntPrioUpd messages. This bit reflects the setup of the filter at any given time. 0 - flat, 1 - cluster.
16	0h RW	CLASTCHKSMPPMOD: 0: Disable checking for Logical_APICID[31:0] being non-zero when sampling flat/ cluster mode bit in the IntPrioUpd message as part of setting bit 1 in this register 1: Enable the above checking
15:2	0h RO	Reserved
1	0h RW	FORCEX2APIC: Write: 1: Forces the system to move into X2APIC Mode. 0: No affect
0	1h RW	XAPICEN: Set this bit if you would like extended XAPIC configuration to be used. This bit can be written directly, and can also be updated using XTPR messages

5.1.4 The register type CPUNODEID_CFG (CPUNODEID_CFG) — Offset C0h

Node ID Configuration Register

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:0, F:0] + C0h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
RW	RW	R	1

Bit Range	Default & Access	Field Name (ID): Description
31:16	0h RO	Reserved
15:13	0h RW	NODECTRLID: Node ID of the Node Controller. Set by the BIOS. (EX processor only)
12:10	0h RW	LGCNODEID: NodeID of the legacy socket
9:8	0h RO	Reserved
7:5	0h RW	LOCKNODEID: NodeID of the lock master
4:3	0h RO	Reserved
2:0	0h RW	LCLNODEID: Node Id of the local Socket

5.1.5 The register type GIDNIDMAP_CFG (GIDNIDMAP_CFG) — Offset D4h

Mapping between group id and nodeid

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:0, F:0] + D4h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access	Policy Group ID
RW	RW	R	1

Bit Range	Default & Access	Field Name (ID): Description
31:24	0h RO	Reserved
23:21	0h RW	NODEID7: NodeID for group id 7
20:18	0h RW	NODEID6: Node Id for group 6
17:15	0h	NODEID5:
continued...		

Bit Range	Default & Access	Field Name (ID): Description
	RW	Node Id for group 5
14:12	0h RW	NODEID4: Node Id for group id 4
11:9	0h RW	NODEID3: Node Id for group 3
8:6	0h RW	NODEID2: Node Id for group Id 2
5:3	0h RW	NODEID1: Node Id for group Id 1
2:0	0h RW	NODEID0: Node Id for group 0

5.2 Bus: B(30), Device: 0, Function: 2 (DECS)

This chapter documents the registers in Bus: 30, Device 0, Function 2.

5.2.1 Vendor ID (VID_0_0_2_CFG) – Offset 0h

PCI Vendor ID Register

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:0, F:2] + 0h	8086h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
15:0	8086h RO	VENDOR_IDENTIFICATION_NUMBER: The value is assigned by PCI-SIG to Intel.

5.2.2 Device ID (DID_0_0_2_CFG) – Offset 2h

PCI Device Identification Number

Type	Size	Offset	Default
PCI	16 bit	[B:30, D:0, F:2] + 2h	3252h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
15:0	3252h RO	DEVICE_IDENTIFICATION_NUMBER: The value is assigned by each IP/function owner as a unique identifier.

5.2.3 The register type CPUBUSNO_CFG (CPUBUSNO_CFG) — Offset D0h

Bus Number Configuration

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:0, F:2] + D0h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	RW	R

Bit Range	Default & Access	Field Name (ID): Description
31:24	0h RO	Reserved
23:16	00h RW	SEGMENT: Segment number of Uncore Bus A & B
15:8	00h RW	CPUBUSNO1: Uncore Bus B
7:0	00h RW	CPUBUSNO0: Uncore Bus A

5.2.4 SMI generation control (UBOX_SMICTRL_CFG) — Offset D8h

SMI generation control

Type	Size	Offset	Default
PCI	32 bit	[B:30, D:0, F:2] + D8h	04000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	RW	R

Bit Range	Default & Access	Field Name (ID): Description
31:30	0h RO	Reserved
29	0h RW/P	UPIDISABLE: If set, only local SMIs will be logged
28	0h RW/P	SMIDIS4: Disable Generation of SMI from CSMI from MsgCh
continued...		

Bit Range	Default & Access	Field Name (ID): Description
27	0h RW/P	SMIDIS3: Disable Generation of SMI from message channel
26	1h RW/P	SMIDIS2: Disable generation of SMI for lock timeout, cfg write mis-align access, and cfg read mis-align access.
25	0h RW/P	SMIDIS: Disable generation of SMI
24	0h RW/P	UMCSMIEN: This is the enable bit that enables SMI generation due to a UMC 1 -> Generate SMI after the threshold counter expires. 0 -> Disable generation of SMI
23:20	0h RO	Reserved
19:0	00000h RW/P	THRESHOLD: This is the countdown that happens in the hardware before an SMI is generated due to a UMC

6.0 Power Control Unit (PCU) Registers

The Power Control Unit (PCU) is a dedicated controller that provides power and thermal management for the processor. The PCU implements a PECI interface for out-of-band management. The PCU consists of a dedicated micro-controller, ROM and RAM for Pcode (PCU microcode), HW state machines, I/O registers for interfacing to the microcontroller and interfaces to the hardware units in the processor.

The PCU Registers are implemented in the following Bus, Device, Functions:

- Bus: B(31), Device: 30, Function: 1 (PCU CR1)
- Bus: B(31), Device: 30, Function: 2 (PCU CR2)

NOTE

For additional details on the PCU, see *4th Gen Intel® Xeon® Processor Scalable Family, Codename Sapphire Rapids, Architecture Specification*, chapter “Power Controller Unit (PCU) Functional Description.”

Table 4. Summary of Power Control Unit (PCU) Registers

Bus: B(31), Device: 30, Function: 0 (PCU CR0 Registers) on page 42
Bus: B(31), Device: 30, Function: 2 (PCU CR2 Registers) on page 44

6.1 Bus: B(31), Device: 30, Function: 0 (PCU CR0 Registers)

This chapter documents the Power Control Unit CR0 registers in Bus (31), Device 30, Function 0.

6.1.1 Vendor ID (VID_1_30_0_CFG) — Offset 0h

PCI Vendor ID Register

Type	Size	Offset	Default
PCI	16 bit	[B:31, D:30, F:0] + 0h	8086h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
15:0	8086h RO	VENDOR_IDENTIFICATION_NUMBER: The value is assigned by PCI-SIG to Intel.

6.1.2 Device ID (DID_1_30_0_CFG) — Offset 2h

PCI Device Identification Number

Type	Size	Offset	Default
PCI	16 bit	[B:31, D:30, F:0] + 2h	3258h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
15:0	3258h RO	DEVICE_IDENTIFICATION_NUMBER: The value is assigned by each IP/function owner as a unique identifier.

6.1.3 PACKAGE ENERGY STATUS (PACKAGE_ENERGY_STATUS_CFG) — Offset 90h

Package energy consumed by the entire CPU (including IA, Uncore). The counter will wrap around and continue counting when it reaches its limit.

NOTE

note: Bit definitions are the same as PRIP_NRG_STTS_CFG, offset 88h.

6.1.4 PACKAGE TEMPERATURE (PACKAGE_TEMPERATURE_CFG) — Offset C8h

Package temperature in degrees (C). This field is updated by FW.

Type	Size	Offset	Default
PCI	32 bit	[B:31, D:30, F:0] + C8h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
31:8	0h RO	Reserved
7:0	00h RO/V	DATA: Package temperature in degrees (C) PLUS 64C to prevent underflow.

6.1.5 TEMPERATURE TARGET (TEMPERATURE_TARGET_CFG) — Offset E4h

Legacy register holding temperature related constants for Platform use.

Type	Size	Offset	Default
PCI	32 bit	[B:31, D:30, F:0] + E4h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	R	R

Bit Range	Default & Access	Field Name (ID): Description
31	0h RW/1S/V	LOCKED: When set, this entire register becomes read-only. Locked by: TEMPERATURE_TARGET.LOCKED
30	0h RO	Reserved
29:24	00h RW/L	TJ_MAX_TCC_OFFSET: Temperature offset in degrees (C) from the TJ Max. Used for throttling temperature. Will not impact temperature reading. If offset is allowed and set - the throttle will occur and reported at lower than Tj_max. Locked by: TEMPERATURE_TARGET.LOCKED
23:16	00h RO/V	REF_TEMP: This field indicates the maximum junction temperature, also referred to as the Throttle Temperature, TCC Activation Temperature or Prochot Temperature. This is the temperature at which the Adaptive Thermal Monitor is activated. MSR alias previously with access type RO_FW changed to RW_L with lock bit 1
15:8	00h RO/V	FAN_TEMP_TARGET_OFST: Fan Temperature Target Offset (a.k.a. T-Control) indicates the relative offset from the Thermal Monitor Trip Temperature at which fans should be engaged. MSR alias previously with access type RO_FW changed to RW_L with lock bit 1
7	0h RW/L	TCC_OFFSET_CLAMPING_BIT: When enabled will allow RATL throttling below P1 Locked by: TEMPERATURE_TARGET.LOCKED
6:0	00h RW/L	TCC_OFFSET_TIME_WINDOW: Describes the RATL averaging time window Locked by: TEMPERATURE_TARGET.LOCKED

6.2 Bus: B(31), Device: 30, Function: 2 (PCU CR2 Registers)

This chapter documents the Power Control Unit CR2 registers in Bus (31), Device 30, Function 2.

6.2.1 Vendor ID (VID_1_30_2_CFG) — Offset 0h

PCI Vendor ID Register

Type	Size	Offset	Default
PCI	16 bit	[B:31, D:30, F:2] + 0h	8086h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
15:0	8086h RO	VENDOR_IDENTIFICATION_NUMBER: The value is assigned by PCI-SIG to Intel.

6.2.2 Device ID (DID_1_30_2_CFG) — Offset 2h

PCI Device Identification Number

Type	Size	Offset	Default
PCI	16 bit	[B:31, D:30, F:2] + 2h	325Ah

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
15:0	325Ah RO	DEVICE_IDENTIFICATION_NUMBER: The value is assigned by each IP/function owner as a unique identifier.

6.2.3 PCI Command (PCICMD_1_30_2_CFG) — Offset 4h

PCI Command Register

Type	Size	Offset	Default
PCI	16 bit	[B:31, D:30, F:2] + 4h	0000h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
15:11	0h RO	Reserved
10	0h RO	INTX_DISABLE: N/A for these devices
9	0h RO	FAST_BACK_TO_BACK_ENABLE: Not applicable to PCI Express and is hardwired to 0
8	0h RO	SERR_ENABLE: This bit has no impact on error reporting from these devices
7	0h RO	IDSEL_STEPPING_WAIT_CYCLE_CONTROL: Not applicable to internal devices. Hardwired to 0.
continued...		

Bit Range	Default & Access	Field Name (ID): Description
6	0h RO	PARITY_ERROR_RESPONSE: This bit has no impact on error reporting from these devices
5	0h RO	VGA_PALETTE_SNOOP_ENABLE: Not applicable to internal devices. Hardwired to 0.
4	0h RO	MEMORY_WRITE_AND_INVALIDATE_ENABLE: Not applicable to internal devices. Hardwired to 0.
3	0h RO	SPECIAL_CYCLE_ENABLE: Not applicable. Hardwired to 0.
2	0h RO	BUS_MASTER_ENABLE: Hardwired to 0 since these devices do not generate any transactions
1	0h RO	MEMORY_SPACE_ENABLE: Hardwired to 0 since these devices do not decode any memory BARs
0	0h RO	IO_SPACE_ENABLE: Hardwired to 0 since these devices do not decode any IO BARs

6.2.4 PCI Status (PCISTS_1_30_2_CFG) – Offset 6h

PCI Status Register

Type	Size	Offset	Default
PCI	16 bit	[B:31, D:30, F:2] + 6h	0000h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
15	0h RO	DETECTED_PARITY_ERROR: This bit is set when the device receives a packet on the primary side with an uncorrectable data error (including a packet with poison bit set) or an uncorrectable address/control parity error. The setting of this bit is regardless of the Parity Error Response bit (PERRE) in the PCICMD register.
14	0h RO	SIGNALED_SYSTEM_ERROR: Hardwired to 0
13	0h RO	RECEIVED_MASTER_ABORT: Hardwired to 0
12	0h RO	RECEIVED_TARGET_ABORT: Hardwired to 0
11	0h RO	SIGNALED_TARGET_ABORT: Hardwired to 0
10:9	0h RO	DEVSEL_TIMING: Not applicable to PCI Express. Hardwired to 0.
8	0h RO	MASTER_DATA_PARITY_ERROR: Hardwired to 0
continued...		

Bit Range	Default & Access	Field Name (ID): Description
7	0h RO	FAST_BACK_TO_BACK: Not applicable to PCI Express. Hardwired to 0.
6	0h RO	Reserved
5	0h RO	X66MHZ_CAPABLE: Not applicable to PCI Express. Hardwired to 0.
4	0h RO	CAPABILITIES_LIST: This bit indicates the presence of a capabilities list structure. When set to 1, indicates the register at 34h provides an offset into the function.
3	0h RO	INTX_STATUS: Reflects the state of the INTA# signal at the input of the enable/disable circuit. This bit is set by HW to 1 when the INTA# is asserted. This bit is reset by HW to 0 after the interrupt is cleared (independent of the state of the Interrupt Disable bit in the PCICMD register). Hardwired to 0 on the processor
2:0	0h RO	Reserved

6.2.5 Revision ID and Class Code Register (RID_CCR_1_30_2_CFG) — Offset 8h

PCIe header Revision ID register and Class Code register

Type	Size	Offset	Default
PCI	32 bit	[B:31, D:30, F:2] + 8h	08800000h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
31:24	08h RO	BASE_CLASS: Base class code
23:16	80h RO	SUB_CLASS: Sub class code
15:8	00h RO	REGISTER_LEVEL_PROGRAMMING_INTERFACE: Register Level Programming Interface
7:0	00h RO/V	REVISION_ID: Reflects the Uncore Revision ID after reset. Reflects the Compatibility Revision ID if BIOS updates.

6.2.6 Cache Line Size Register (CLSR_1_30_2_CFG) — Offset Ch

PCI Cache Line Size Register

Type	Size	Offset	Default
PCI	8 bit	[B:31, D:30, F:2] + Ch	00h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
7:0	00h RO	CACHELINE_SIZE: Size of Cacheline

6.2.7 PCI Latency Timer (PLAT_1_30_2_CFG) — Offset Dh

PCI Latency Timer

Type	Size	Offset	Default
PCI	8 bit	[B:31, D:30, F:2] + Dh	00h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
7:0	00h RO	PRIMARY_LATENCY_TIMER: Not applicable to PCI-Express. Hardwired to 00h.

6.2.8 Header Type (HDR_1_30_2_CFG) — Offset Eh

PCI Header Type

Type	Size	Offset	Default
PCI	8 bit	[B:31, D:30, F:2] + Eh	80h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
7	1h RO	MULTI_FUNCTION_DEVICE: This bit defaults to 1b since all these devices are multi-function
6:0	00h RO	CONFIGURATION_LAYOUT: Type 0 header

6.2.9 Built In Self Test (BIST_1_30_2_CFG) — Offset Fh

PCI BIST Register

Type	Size	Offset	Default
PCI	8 bit	[B:31, D:30, F:2] + Fh	00h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
7:0	00h RO	BIST_TESTS: Not supported. Hardwired to 00h

6.2.10 Subsystem Vendor ID (SVID_1_30_2_CFG) — Offset 2Ch

Subsystem Vendor ID Register

Type	Size	Offset	Default
PCI	16 bit	[B:31, D:30, F:2] + 2Ch	8086h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	RW	RW

Bit Range	Default & Access	Field Name (ID): Description
15:0	8086h RW/O	SUBSYSTEM_VENDOR_IDENTIFICATION_NUMBER: The default value specifies Intel® but can be set to any value once after reset.

6.2.11 Subsystem ID (SDID_1_30_2_CFG) — Offset 2Eh

Subsystem ID Register

Type	Size	Offset	Default
PCI	16 bit	[B:31, D:30, F:2] + 2Eh	0000h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	RW	RW

Bit Range	Default & Access	Field Name (ID): Description
15:0	0000h RW/O	SUBSYSTEM_IDENTIFICATION_NUMBER: Assigned by the subsystem vendor to uniquely identify the subsystem

6.2.12 Capability Pointer (CAPPTR_1_30_2_CFG) — Offset 34h

PCI Capability Pointer Register

Type	Size	Offset	Default
PCI	8 bit	[B:31, D:30, F:2] + 34h	00h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
7:0	00h RO	CAPABILITY_POINTER: Points to the first capability structure for the device which is the PCIe capability.

6.2.13 Interrupt Line (INTL_1_30_2_CFG) — Offset 3Ch

PCI Interrupt Line Register

Type	Size	Offset	Default
PCI	8 bit	[B:31, D:30, F:2] + 3Ch	00h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
7:0	00h RO	INTERRUPT_LINE: N/A for these devices

6.2.14 Interrupt Pin (INTPIN_1_30_2_CFG) — Offset 3Dh

PCI Interrupt Pin Register

Type	Size	Offset	Default
PCI	8 bit	[B:31, D:30, F:2] + 3Dh	00h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
7:0	00h RO	INTERRUPT_PIN: N/A since these devices do not generate any interrupt on their own

6.2.15 Minimum Grant (MINGNT_1_30_2_CFG) — Offset 3Eh

PCI Min Grant Register

Type	Size	Offset	Default
PCI	8 bit	[B:31, D:30, F:2] + 3Eh	00h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
7:0	00h RO	MGV: The device does not burst as a PCI compliant master.

6.2.16 Maximum Latency (MAXLAT_1_30_2_CFG) — Offset 3Fh

PCI Max Latency Register

Type	Size	Offset	Default
PCI	8 bit	[B:31, D:30, F:2] + 3Fh	00h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
7:0	00h RO	MLV: The device has no specific requirements for how often it needs to access the PCI bus.

6.2.17 DRAM ENERGY STATUS (DRAM_ENERGY_STATUS_CFG) — Offset 80h

DRAM energy consumed by all the DIMMS in all the Channels. The counter will wrap around and continue counting when it reaches its limit.

ENERGY_UNIT for DRAM domain is 15.3uJ.

The data is updated by PCODE and is Read Only for all SW.

Type	Size	Offset	Default
PCI	32 bit	[B:31, D:30, F:2] + 80h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
31:0	00000000h	DATA:

Bit Range	Default & Access	Field Name (ID): Description
	RO/V	Energy Value

6.2.18 PACKAGE OVERFLOW STATUS (PACKAGE_RAPL_PERF_STATUS) — Offset 88h

This register is used by Pcode to report Package Power limit violations in the Platform PBM.

Provides information on the performance impact of the RAPL power limit.

Provides the total time for which PACKAGE was throttled because of the RAPL power limit. Throttling here is defined as going below O.S requested P-state.

Total time = Reg value * 1s * 1 / 2(time unit)

Usage model: Enables the O.S/Driver to learn about PACKAGE throttling as a result of RAPL limit.

Can be used by other S/W components that control the PACKAGE power.

This register is mapped as an MSR with the name Package_Perf_Status

Type	Size	Offset	Default
PCI	32 bit	[B:31, D:30, F:2] + 88h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
31:0	00000000h RO/V	PWR_LIMIT_THROTTLE_CTR: Reports the number of times the Power limiting algorithm had to clip the power limit due to hitting the lowest power state available. Accumulated PACKAGE throttled time

6.2.19 DRAM POWER INFORMATION (DRAM_POWER_INFO_CFG) — Offset A8h

Power allowed for DRAM

Type	Size	Offset	Default
PCI	64 bit	[B:31, D:30, F:2] + A8h	0028025800780118h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	R	R

Bit Range	Default & Access	Field Name (ID): Description
63	0h RW/L	LOCK: Lock bit to lock the Register Locked by: DRAM_POWER_INFO_CFG.LOCK
62:55	0h RO	Reserved
54:48	28h RW/L	DRAM_MAX_WIN: The maximal time window allowed for the DRAM. Higher values will be clamped to this value. $x = \text{PKG_MAX_WIN}[54:53]$ $y = \text{PKG_MAX_WIN}[52:48]$ The timing interval window is Floating Point number given by $1.x * \text{power}(2,y)$. The unit of measurement is defined in DRAM_POWER_INFO_UNIT_MSR[TIME_UNIT]. Locked by: DRAM_POWER_INFO_CFG.LOCK
47	0h RO	Reserved
46:32	0258h RW/L	DRAM_MAX_PWR: The maximal power setting allowed for DRAM. Higher values will be clamped to this value. The maximum setting is typical (not guaranteed). The units for this value are defined in DRAM_POWER_INFO_UNIT_MSR[PWR_UNIT]. Locked by: DRAM_POWER_INFO_CFG.LOCK
31	0h RO	Reserved
30:16	0078h RW/L	DRAM_MIN_PWR: Intel® Reserved. This field is deprecated. User should ignore any values in this field.
15	0h RO	Reserved
14:0	0118h RW/L	DRAM_TDP: The Spec power allowed for DRAM. The TDP setting is typical (not guaranteed). The units for this value are defined in DRAM_POWER_INFO_UNIT_MSR[PWR_UNIT]. Locked by: DRAM_POWER_INFO_CFG.LOCK

6.2.20 MEM TRML TEMPERATURE REPORT (MEM_TRML_TEMPERATURE_REPORT[0]) — Offset C8h

This register is used to report the thermal status of the memory. The channel max temperature field is used to report the maximal temperature of all ranks.

Type	Size	Offset	Default
PCI	32 bit	[B:31, D:30, F:2] + C8h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
31:24	0h RO	Reserved
23:16	00h RO/V	CHANNEL2_MAX_TEMPERATURE: Temperature in Degrees (C).
15:8	00h RO/V	CHANNEL1_MAX_TEMPERATURE: Temperature in Degrees (C).
7:0	00h RO/V	CHANNELO_MAX_TEMPERATURE: Temperature in Degrees (C).

6.2.21 MEM TRML TEMPERATURE REPORT (MEM_TRML_TEMPERATURE_REPORT[1]) — Offset CCh

This register is used to report the thermal status of the memory. The channel max temperature field is used to report the maximal temperature of all ranks.

NOTE

note:Bit definitions are the same as MEM_TRML_TEMPERATURE_REPORT[0], offset C8h.

6.2.22 MEM TRML TEMPERATURE REPORT (MEM_TRML_TEMPERATURE_REPORT[2]) — Offset D0h

This register is used to report the thermal status of the memory. The channel max temperature field is used to report the maximal temperature of all ranks.

NOTE

note:Bit definitions are the same as MEM_TRML_TEMPERATURE_REPORT[0], offset C8h.

6.2.23 MEM TRML TEMPERATURE REPORT (MEM_TRML_TEMPERATURE_REPORT[3]) — Offset D4h

This register is used to report the thermal status of the memory. The channel max temperature field is used to report the maximal temperature of all ranks.

NOTE

note:Bit definitions are the same as MEM_TRML_TEMPERATURE_REPORT[0], offset C8h.

6.2.24 DRAM PLANE OVERFLOW STATUS (DRAM_RAPL_PERF_STATUS_CFG) — Offset D8h

This register is used by Pcode to report DRAM Plane Power limit violations in the Platform PBM.

Provides information on the performance impact of the RAPL power limit.

Provides the total time for which DRAM was throttled because of the RAPL power limit.
Total time = Reg Value * 1 s * 1/ (2time_unit)

Usage model: Enables the OS/Driver to learn about DRAM throttling as a result of RAPL limit.

Can be used by other S/W components that control the DRAM power.

This register is mapped as an MSR with the name DRAM_RAPL_PERF_STATUS.

Type	Size	Offset	Default
PCI	32 bit	[B:31, D:30, F:2] + D8h	00000000h

Register Level Access:

BIOS Access	SMM Access	OS Access
R	R	R

Bit Range	Default & Access	Field Name (ID): Description
31:0	00000000h RO/V	PWR_LIMIT_THROTTLE_CTR: Reports the number of times the Power limiting algorithm had to clip the power limit due to hitting the lowest power state available. Accumulated DRAM throttled time

6.2.25 THERMTRIP CONFIGURATION (THERMTRIP_CONFIG_CFG) – Offset F8h

This register is used to configure whether the Thermtrip signal only carries the processor Trip information, or does it carry the Mem trip information as well. The register will be used by HW to enable ORing of the memtrip info into the thermtrip OR tree.

Type	Size	Offset	Default
PCI	32 bit	[B:31, D:30, F:2] + F8h	000000C0h

Register Level Access:

BIOS Access	SMM Access	OS Access
RW	RW	R

Bit Range	Default & Access	Field Name (ID): Description
31:8	0h RO	Reserved
7	1h RW/P	MEMTRIP1_TO_XXMEMTRIP_N_EN: When set, Punit memtrip network 1 is enabled to cause xxMEMTRIP_N pin assertion. The specific meaning and usage model of Memtrip network 1 is defined by the SoC architecture and integration.
6	1h RW/P	MEMTRIP0_TO_XXMEMTRIP_N_EN:
continued...		

Bit Range	Default & Access	Field Name (ID): Description
		When set, Punit memtrip network 0 is enabled to cause xxMEMTRIP_N pin assertion. The specific meaning and usage model of Memtrip network 0 is defined by the SoC architecture and integration.
5	0h RW/P	MEMTRIP1_TO_XXTHERMTRIP_N_EN: When set, Punit memtrip network 1 is enabled to cause xxTHERMTRIP_N pin assertion. The specific meaning and usage model of Memtrip network 1 is defined by the SoC architecture and integration.
4	0h RW/P	MEMTRIP0_TO_XXTHERMTRIP_N_EN: When set, Punit memtrip network 0 is enabled to cause xxTHERMTRIP_N pin assertion. The specific meaning and usage model of Memtrip network 0 is defined by the SoC architecture and integration.
3:0	0h RO	Reserved