Arctic Sound-P Design Overview

Intel Corporation
Data Center Platform Application Engineering
March 2021
Reference Number: 616579
Revision Number: 1.2
Intel Confidential

Legal Disclaimer

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' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at Intel.com, or from the OEM or retailer.

No computer system can be absolutely secure. Intel does not assume any liability for lost or stolen data or systems or any damages resulting from such losses.

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.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps.

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.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Copies of documents which have an order number and are referenced in this document may be obtained by calling 1-800-548-4725 or by visiting www.intel.com/design/literature.htm.

Intel, VTune, and the Intel logo are trademarks of Intel Corporation in the U. S. or its subsidiaries.

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

Objective

- To provide an overview of Arctic Sound-P PCI Express* (PCIe*) product hardware design considerations.

Target Audience

- System Design Engineers (Electrical, Thermal, Mechanical)
- System Software Engineers (including Manageability and Reliability, Availability, and Serviceability [RAS])

Scope of This Training

- Product Overview
- Electrical, Thermal, Mechanical Specification Summary
- Power and Thermal Management
- Reliability, Availability, and Serviceability (RAS)
- Card Management Firmware

Other Available Trainings

- oneAPI
- Media Analytics Software
- OpenCL*
- Deep Learning (DL) Frameworks
- Virtualization
- Open Visual Inference and Neural network Optimization* (OpenVINO*)

Introduction

Overview Data Center GPU Use Cases and Swimlanes

Artificial Intelligence (AI)

High Performance Computing (HPC)

Media analytics

Media processing and delivery

Cloud graphics and gaming

Immersive media

Compute accelerator

Multi-purpose Graphic Processing Unit (GPU)

Note: Intel may make changes to specifications and product descriptions at any time, without notice. The information here is subject to change without notice. Do not finalize a design with this information. Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration.

Arctic Sound-P

Multi-tile, multi-purpose Graphic Processing Unit (GPU) for data center workloads

Industry-leading media engine

Up to 32 GB High Bandwidth Memory (HBM2E)

Architected for scalable performance

Flexible Thermal Design Power (TDP) and form factor design points

Multi-use for many markets

Note: Intel may make changes to specifications and product descriptions at any time, without notice. The information here is subject to change without notice. Do not finalize a design with this information. Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance scores are estimates using internal Intel analysis or architecture simulation or modeling and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

Arctic Sound-P Form Factors

	ARCTIC SOUND-P – 1T	ARCTIC SOUND-P – 2T
Segments	Media, Media analytics, Artificial Intelligence (AI)	AI (training and inference), High Performance Computing (HPC), Media, Media analytics
Form factor	PCIe* card: 1/2L, FH, Software (SW) - passively cooled Software Development Vehicle (SDV) only in 2H'20	PCIe* card: 3/4L, FH, DW - passively cooled SDV in 2H'20-1H'21 and Product Release Qualification (PRQ) in Q1'22
Thermal Design Power (TDP)	150W	300W
Memory	16 GB High Bandwidth Memory (HBM2E)	32 GB HBM2E
Memory bandwidth	716 GB/s peak	2x 716 GB/s peak
Input/Output (I/O)	PCI Express* (PCIe*) 4.0 x16	
Media support	Codecs: JPEG, MPEG2, Advanced Visual Computing (AVC), High Efficiency Video Coding (HEVC), VP9d, AV1d, Programmable Media	
Numeric support	Non-systolic: FP64, FP32; Systolic: FP16, Int8, Int4	

Note: Intel may make changes to specifications and product descriptions at any time, without notice. The information here is subject to change without notice. Do not finalize a design with this information. Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance scores are estimates using internal Intel analysis or architecture simulation or modeling and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

Multi-Purpose GPU 2021 Runtime Stack

Arctic Sound Graphic Processing Unit (GPU) Family Customer Solution

Orchestration: Kubernetes*

Container: Docker*, Singularity	Container: Docker*	Container: Docker*	Container: Docker*	Container: Docker*
HPC WL 1 · WL 2 · ● ● ●	**AI Training** WL1, 2 … ● ● ●	**AI Inference** WL 1, 2 … ● ● ●	**Media/Media Analytics** WL 1 · WL 2 · ● ● ●	**Graphics** WL 1 · WL 2 · ● ● ●
Intel® MPI Library, MPICH	TensorFlow* PyTorch*; Python* programming	TensorFlow*, PyTorch*, OpenVINO*	OpenVINO*, Gstreamer*, FFMPEG*	Virtual Display
oneMKL, oneTBB, oneDAL, OneAPI Rendering Toolkit , oneDPC++	oneDNN, oneCCL, oneTBB, oneDPC++	oneDNN, oneDPC++	oneDNN, oneVPL, oneDPC++	Capture and Stream SDK*
Compiler Runtimes, (DPC++, Python* OpenCL*, Fortran* with OpenMP*, C++ with OpenMP*)	DPC++ Compiler Runtime	DPC++ Compiler Runtime, Python*	DPC++ Compiler Runtime	
Driver, Xe Level Zero	Driver, Xe Level Zero	Driver, Xe Level Zero	Driver, Xe Level Zero	Driver, DirectX*, OGL*, Vulkan*
Ubuntu*	Ubuntu*	Ubuntu*	Ubuntu*	Ubuntu*, Windows* 19H1

Drivers: i915, MDAPI

Virtualization/Native OS: Ubuntu* 20.04 – RHEL* 8.3 – CentOS* 8.3 - SLES* 15 SP2 – Windows* Server (Arctic Sound-M ONLY)

Host OS Hypervisor: KVM*, Xen, VMware*, Hyper-V*

Notes:
- Not an exhaustive list of WLs supported for Multi-Purpose GPUs.
- Arctic Sound-P SW support for Linux* only.
- SG1 Software stack is focused on Android* Cloud Gaming and Media Transcode. See the *SG1 Snapshot* for more details.

Hardware (Arctic Sound-P, Arctic Sound-M)

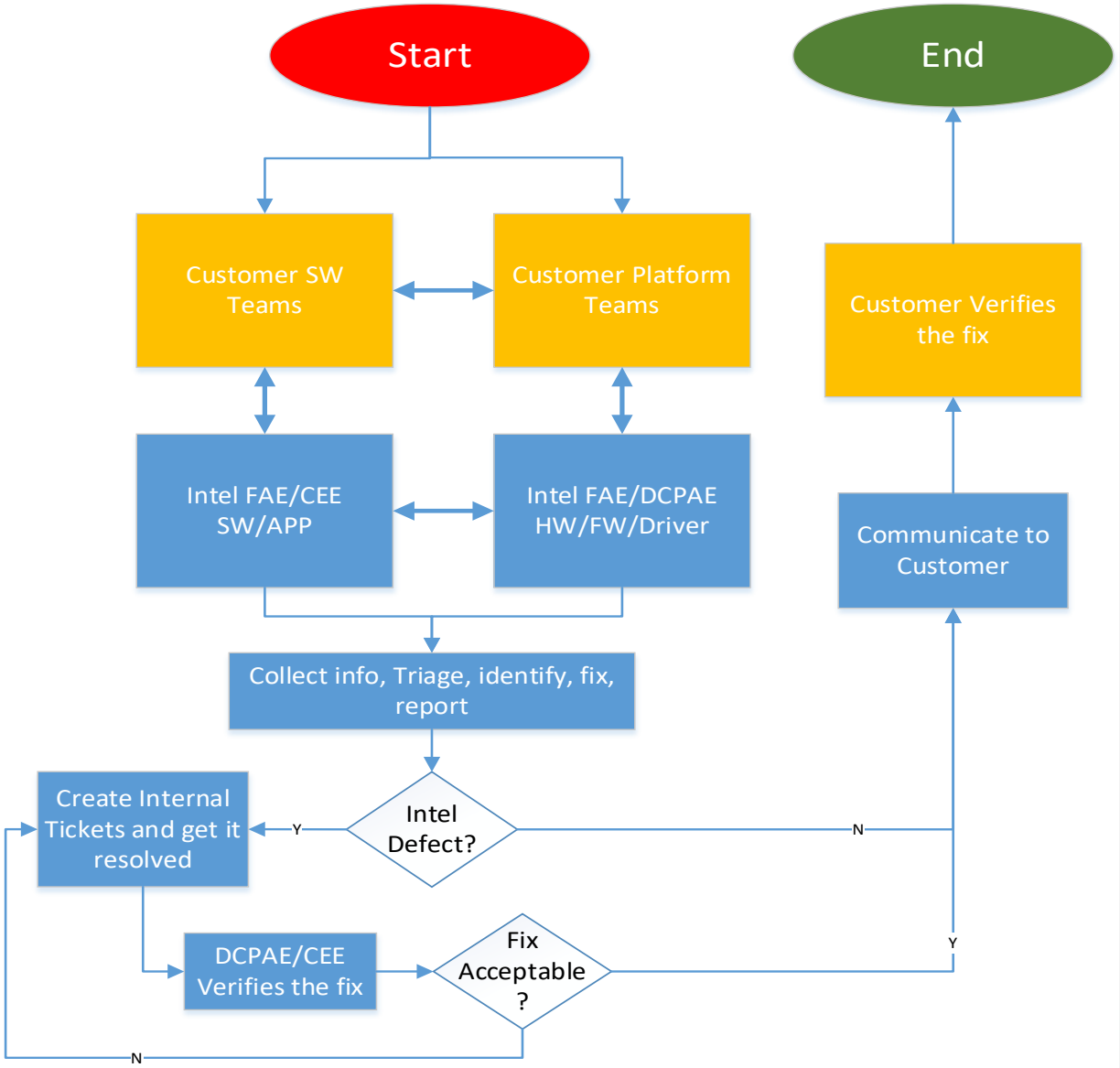
Intel Processes

Intel Processes (1/2)

- Intel Contacts:
  - **Marketing**: Nicholas (Nick) Della Cioppa (nicholas.della.cioppa@intel.com)
  - **AE (Hardware [HW]/Firmware [FW]/Driver)**: Saimanohara (Sai) Alapati (saimanohara.alapati@intel.com)
  - **AE (Software [SW]/APP)** Tao Zhiqi (zhiqi.tao@intel.com)
- RDC – where all Arctic Sound-P Non-Disclosure Agreement (NDA) collaterals (datasheets, models, tools, SW releases, among others) are posted:
  - https://www.intel.com/content/www/us/en/secure/design/confidential/products-and-solutions/processors-and-chipsets/arctic-sound/technical-library.html

Intel Processes (2/2)

- Intel® Premier Support:
  - Place to file and track bugs
- https://premiersupport.intel.com/
- File Intel® Premier Support portal tickets under “Arctic Sound” Project.



Collaterals Planned (1/2)

Collateral Name	Description
Dear Customer Letter (DCL)	Provides details (silicon stepping, frequencies, device IDs, firmware versions, and so on) of the physical samples at each sampling phase.
Message of the Week (MOW)	Provides timely updates (for example: collateral releases, future changes, important issues discovered, and their workarounds, and so on) to the program.
BKC Release	Firmware (FW)/Driver/Software (SW) packages validated internally will be released periodically and include release notes, known issues, and errata.
Arctic Sound Product Family Datasheet, document number 618616	Discusses the architecture, thermal boundary conditions, mechanical dimensions, electrical interface, power requirements, and system management of the PCI Express* (PCIe*) card SKUs.
Arctic Sound-P Product Family Compact Thermal and Mechanical Models, document number 619422	Provides compact thermal models (in FloTHERM* and Icepak* formats), Computer-Aided Design (CAD) mechanical models (in Standard for the Exchange of Product Data [STEP] format), and detailed mechanical drawings (in PDF format).

Collaterals Planned (2/2)

Collateral Name	Description
Arctic Sound-P Product Family Software Guide, document number 631277	Describes setup instructions for Hots/Hardware (HW)/Firmware (FW)/Drivers, quick start instructions on different segment applications, Virtualization and Software (SW) distribution (containers).
Arctic Sound-P Product Family GPU Management Guide, document number 631276	Describes the In-Band (INB) Management and the Out-of-Band (OOB) Management commands and how to use them.
Performance Guide	Describes how to run benchmarks and how to tune to get the best performance.
OpenStack Guide	Documents all the openstack components and how to obtain and use them.
OEM Certification Guide	Provides details of different tests (Thermal, Performance, Longevity, Power Cycling, Shock and Vibe, and so on) that OEMs need to perform before they declare that Arctic Sound-P works in their server.
Arctic Sound-P Product Family Tools Guide, document number 633850	Provides details of different tools available for Arctic Sound-P and how to use them.
Arctic Sound-P Product Family Sightings Report, document number 627380	Describes all known issues of the Arctic Sound-P product and their disposition.

Collaterals Planned Timelines

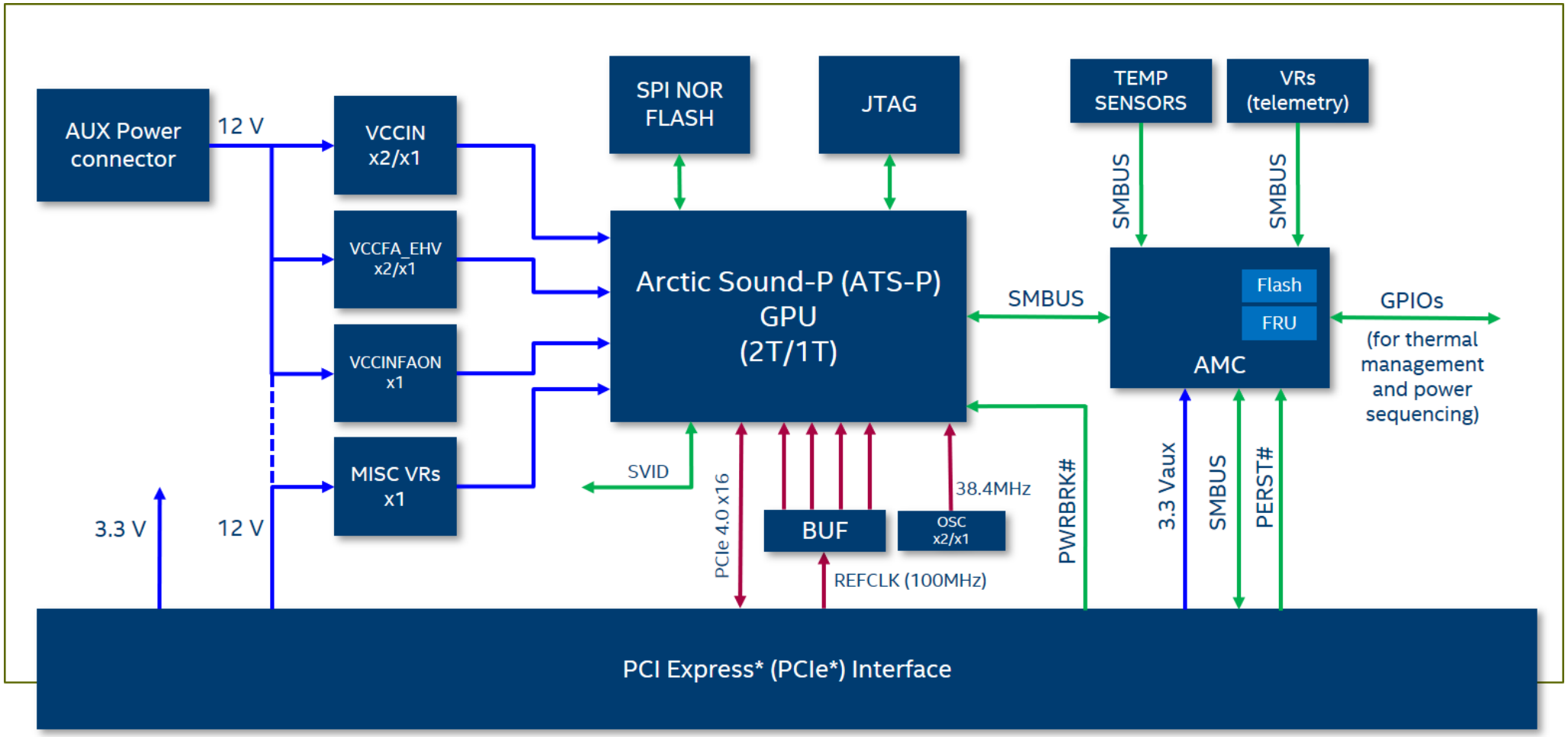
Collateral Name	Document Number	0.5	1.0	1.5	2.0
Arctic Sound Product Family Datasheet	618616	Q1'20	Q2'20		
Arctic Sound-P Product Family Compact Thermal and Mechanical Models	619422	Q1'20	Q2'20		
Message of the Week (MOW)			Q1'21 – Q2'21		
BKC Release			Q1'21 – Q2'21		
Arctic Sound-P Product Family Software Guide	631277	Q1'21	Q1'21 – Q2'21		
Arctic Sound-P Product Family GPU Management Guide	631276	Q1'21	Q2'21		
Performance Guide					Q2'21 – Q4'21
OpenStack Guide					Q2'21 – Q4'21
Arctic Sound-P Product Family Tools Guide	633850		Q1'21 – Q2'21		
OEM Certification Guide			Q1'21 – Q2'21		Q2'21 – Q4'21
Arctic Sound-P Product Family Sightings Report	627380	Q4'20	Q1'21 – Q2'21		

Tools Planned

Tool Name/Functionality	Description	Document Number
Intel® Power Thermal Utility (Intel® PTU) – GPU + HBM	Run workloads that consume maximum Graphic Processing Unit (GPU) and maximum High Bandwidth Memory (HBM) power simultaneously to check the platform power supply and thermal solution	631587
FW Update Tool – IFWI	Integrated Firmware Image (IFWI) Firmware (FW) update tool	
FW Update Tool – AMC	Add-in-card Management Controller (AMC) FW update tool	
Memory Eraser	Erase HBM memory to protect the user data	
Debug Tools	Debug tools to support PAE debug of Arctic Sound-P, including capturing crash dumps, serial dumps, other internal logs (Hardware [HW]/Software [SW]), and performing at-scale-debug	
Management Tool	Tool providing card, Field Replaceable Unit (FRU), monitoring, configuration, both In-Band (INB) and Out-of-Band (OOB)	
Microbenchmark tool	To benchmark underlying HW features (cores, memory, PCI Express* [PCIe*])	
BKCChecker	Validates system configuration for SW	
OEM/ODM Manufacturing Test Suite	Set of test tools for OxM to verify the card functionality after installing in a server system (functional test, burn-in, outgoing quality audit, and so on)	

Hardware

Block Diagram for Arctic Sound-P Cards



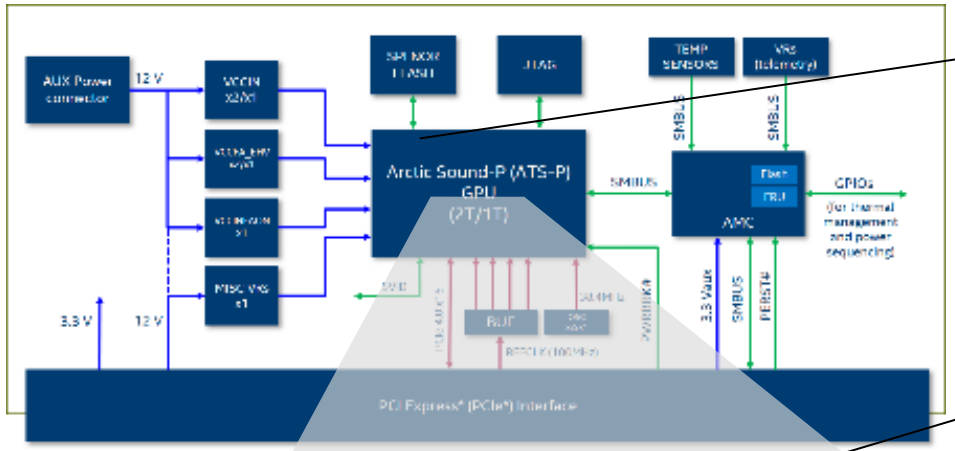
Notes:

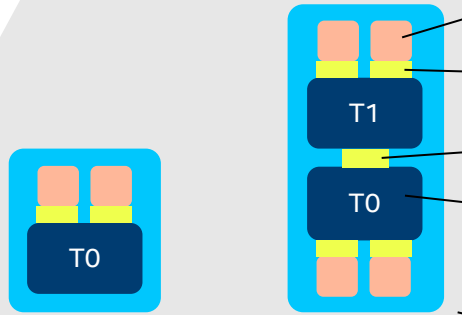
- † All names are preliminary and can change.
- Diagram is not to scale.

Card Block Diagram Details

- The cards are PCI Express* (PCIe*) Card Electric Mechanical (CEM) 4.0 compliant.
- For faster power-throttling response, PWRBRK# from the PCIe* slot goes directly to the Graphic Processing Unit (GPU).
- Full card manageability can be obtained in systems which provide both +3.3 Vaux and Serial Management Bus (SMBus) at the PCIe* slot; however, the cards function properly in systems which do not provide this.
- Add-in-card Management Controller (AMC) performs card power sequencing and Out of Band (OOB) management. It allows the host to set card configuration, get status (such as: temperatures, power), read/write the Field Replaceable Unit (FRU), perform debugging, and others.

GPU Diagram





T0 is the master tile

A A Graphic Processing Unit (GPU) is the heart of the card

F Two High Bandwidth Memory (HBM2E) stacked-die per tile, providing up to 32 GB capacity and 716 GB/s bandwidth

E Embedded Multi-die Interconnect Bridge (EMIB) connects a System on Chip (SoC) with an HBM2E or an adjacent SoC

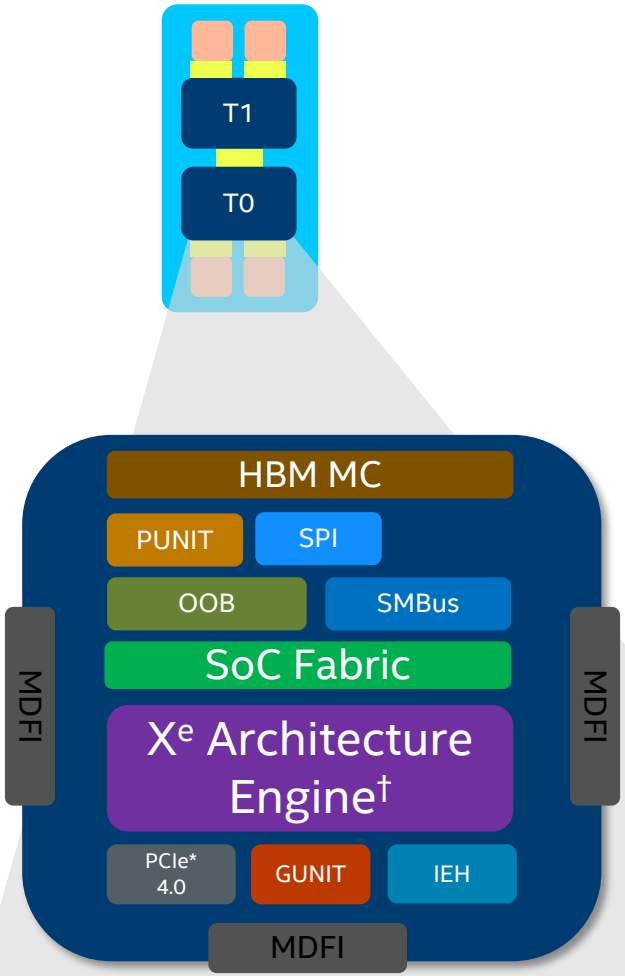
D SoC

C A 2-tile (2T) package has 11 pieces of silicon (6 active, 5 EMIB)

B A Tile (T) comprises an SoC, two HBM2e, and the connecting EMIBs

Notes:
- † All names are preliminary and can change.
- Diagram is not to scale.

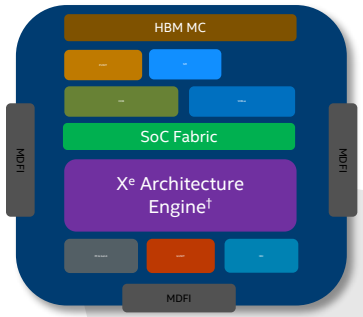
SoC Diagram

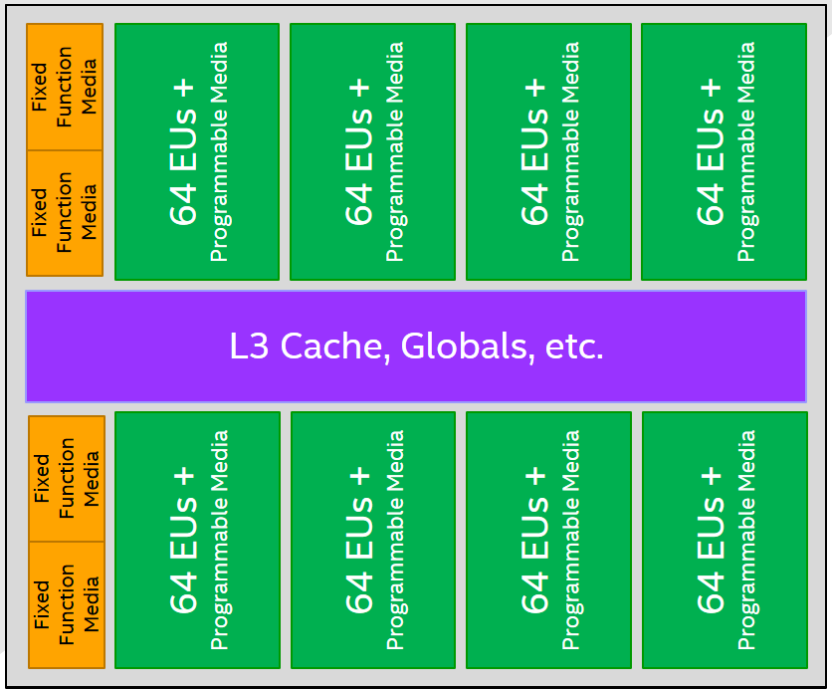


- MDFI: The Modular Die Fabric Interface connects to the adjacent System on Chip (SoC) via the Embedded Multi-die Interconnect Bridge (EMIB), with a bandwidth of up to 260 GB/s in each direction.
- HBM MC: A Memory Controller which interfaces to two High Bandwidth Memory stacked dies.
- PUNIT: Provides reset and power management functionality.
- SPI: A Serial Peripheral Interface controller which interfaces to the flash device on the card.
- OOB: Provides infrastructure support for Out-of-Band access to the resources on the SoC.
- SMBus: A Serial Management Bus is an interface to the Add-in-card Management Controller (AMC) for OOB management.
- Xe Architecture Engine†: See the following slide for further details.
- PCI Express* (PCIe*) 4.0: Interface to the host platform, x16 width (maximum bandwidth of 32 GB/s in each direction).
- GUNIT: Represents the graphics device to the host.
- IEH: The Integrated Error Handler accumulates errors from different functional blocks and reports them to the host.

Notes:
- † All names are preliminary and can change.
- Diagram is not to scale.

Xe Architecture Engine†

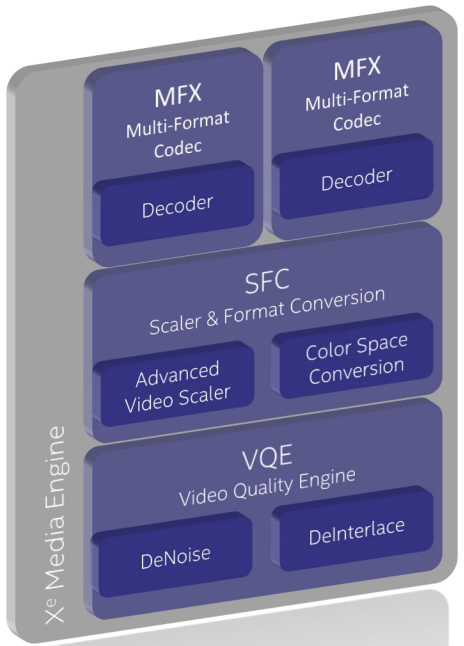


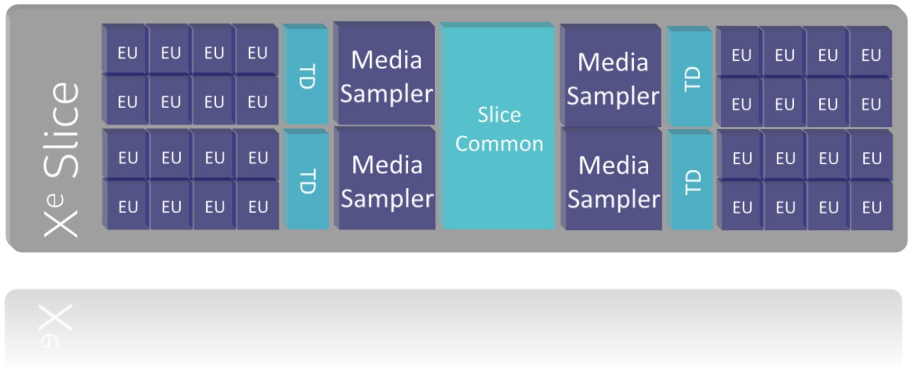


- Execution Unit (EU):
  - Programmable Single Instruction Multiple Data (SIMD) Arithmetic and Logic Units (ALUs) for integer and floating-point compute with shared local memory. Systolic array for accelerating Machine Learning (ML) workloads. 384-512 EUs per tile
- Programmable Media
  - Video Motion Estimation (VME)
- Fixed Function Media:
  - Codecs for video decode (AV1, VP9, High Efficiency Video Coding [HEVC], H.264, and so on)
  - Multi-Format Codex (MFX), Visual Quality Enhancement (VQE), Scaler and Format Conversion (SFC)
  - Seven or eight VD boxes, four VQE boxes, and four SFC per tile
- L3 Cache:
  - Local storage of data from this or other tile's memory space 10 MB/Tile

Notes:
- † All names are preliminary and can change.
- Diagram is not to scale.

Arctic Sound-P Xe Advanced Media Engine



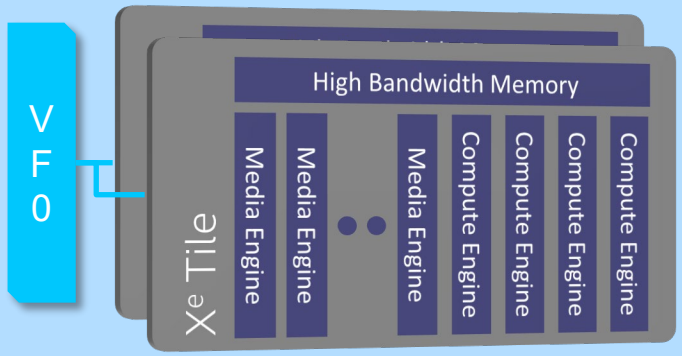


High Density Transcoding
6-10x 4K60 HEVC transcodes per Tile
24-44x 1080p60 AVC transcodes per Tile
6-11x 1:8 1080i60 AVC transcodes per Tile
5040 images/s of 1080p JPEG decode per Tile

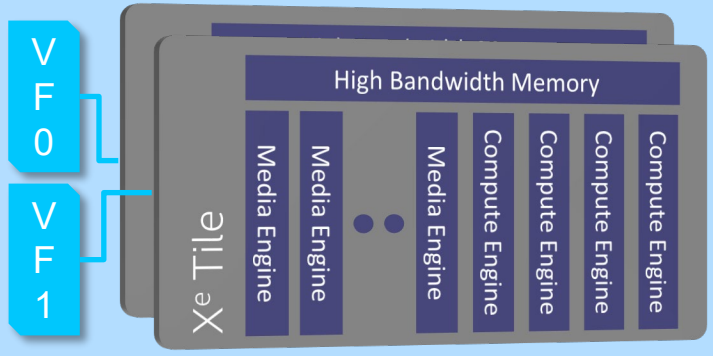
- Full decode of MPEG2, Advanced Video Coding (AVC), High Efficiency Video Coding (HEVC), VP9, and AV1
- Legacy MPEG2 and AVC encode support
  - Interlaced, 4:2:0, 8 bits, up to 4K60
- Enhanced HEVC encode
  - 4x Media Sampler (total 8x Video Motion Estimation [VME]) per slice
  - Video quality enhancements HM + 15% (quality) to HM + 45% (speed)

Arctic Sound-P Xe Virtualization

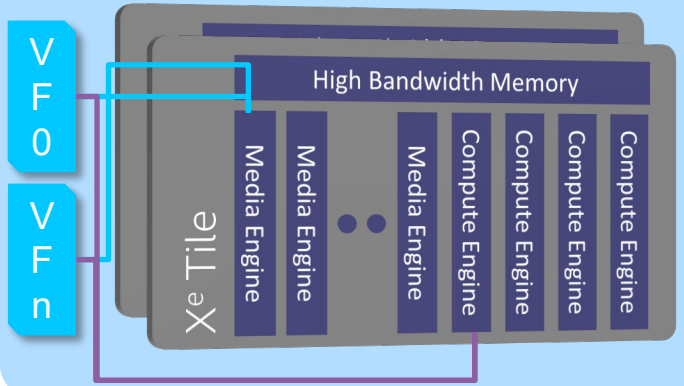
Option 1: Full GPU virtualized to single Virtual Function



Option 2: Individual tile virtualized to separate Virtual Functions



Option 3: Individual Engine executes across all VMs[1]



Single-Root Input/Output Virtualization (SR-IOV) Support:

- Ensures the hypervisor full management of the device
- Multiple modes for flexibility
- Full memory isolation between virtual functions
- Supports up to 63x virtual functions

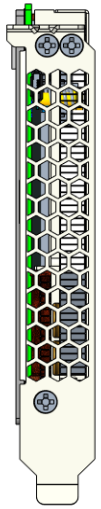
Considerations:

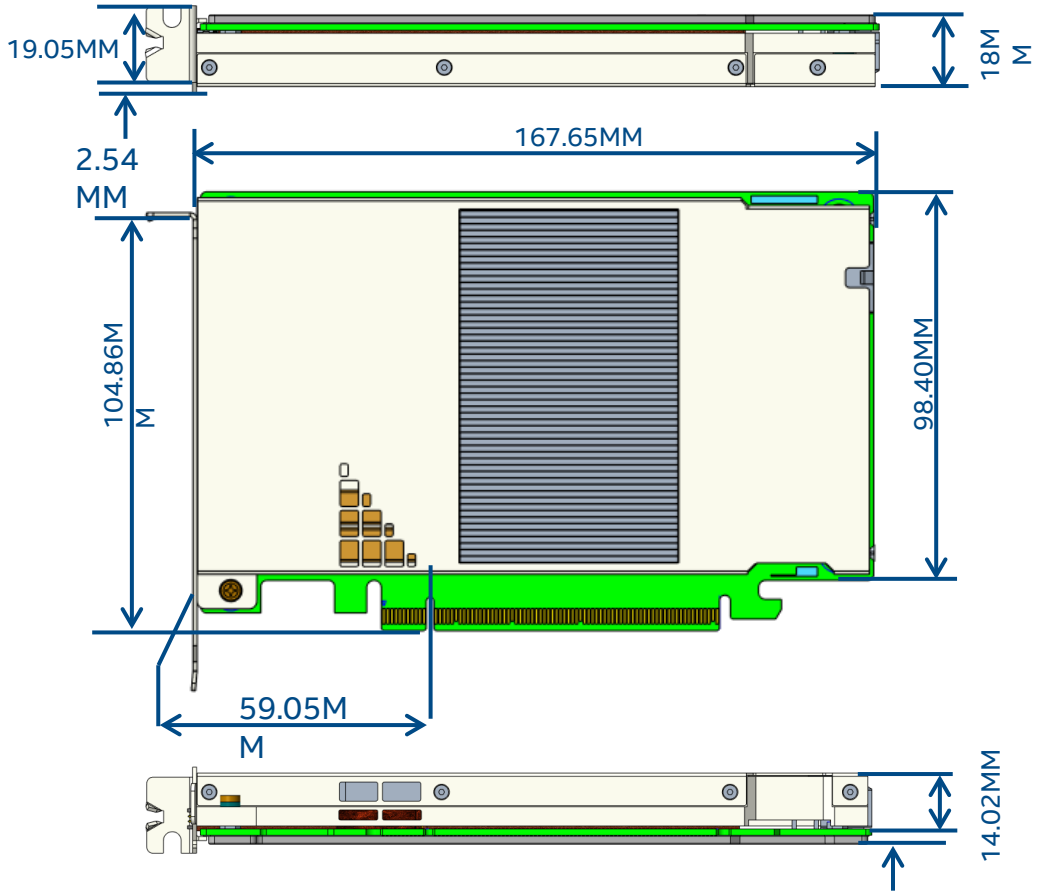
- Option 3 to be used with caution.
- Limited in-Virtual Machine (VM) Telemetry (Full capability provided to the hypervisor)
- Migration time depends on virtual function memory sizing

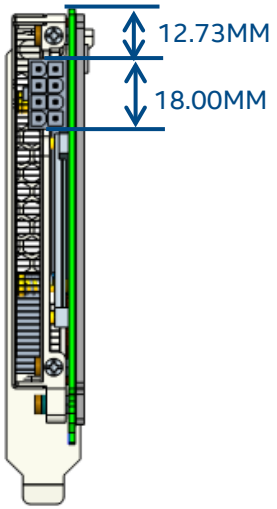
Notes: [1]Use with caution: Engines run to completion and can cause denial of service to VMs, or depending on the application, cause execution or reset issues between VMs.

Mechanical Specifications – 1T/150W Card

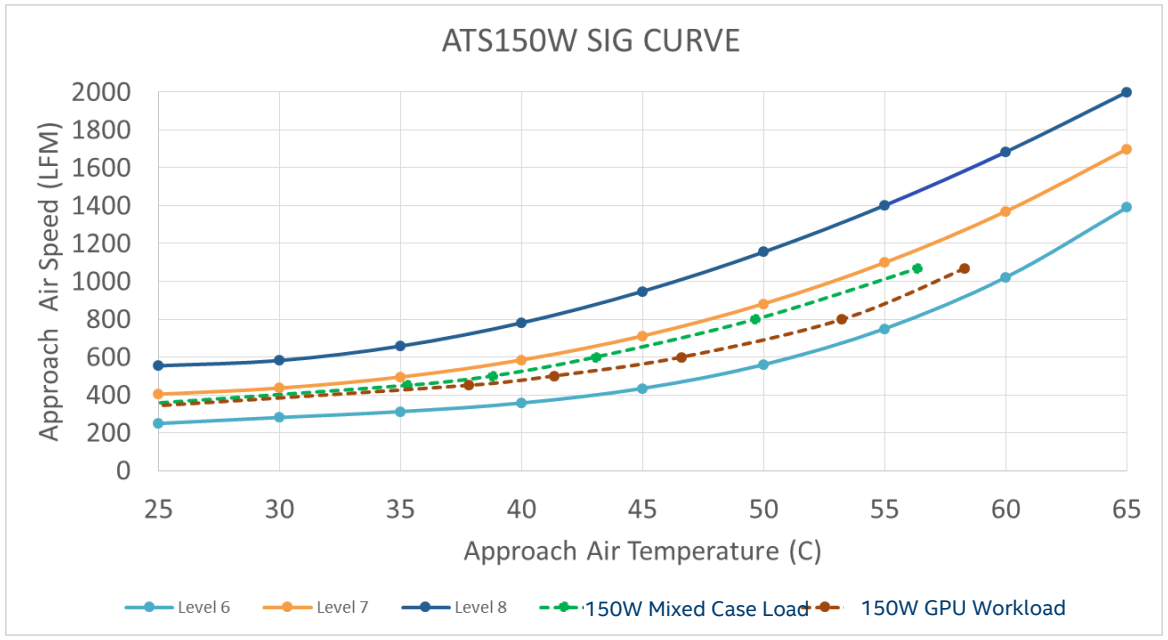
- The full height, half length 1T/150 W card will fully comply with the Card Electric Mechanical (CEM) 4.0 specification (dimensions in mm).

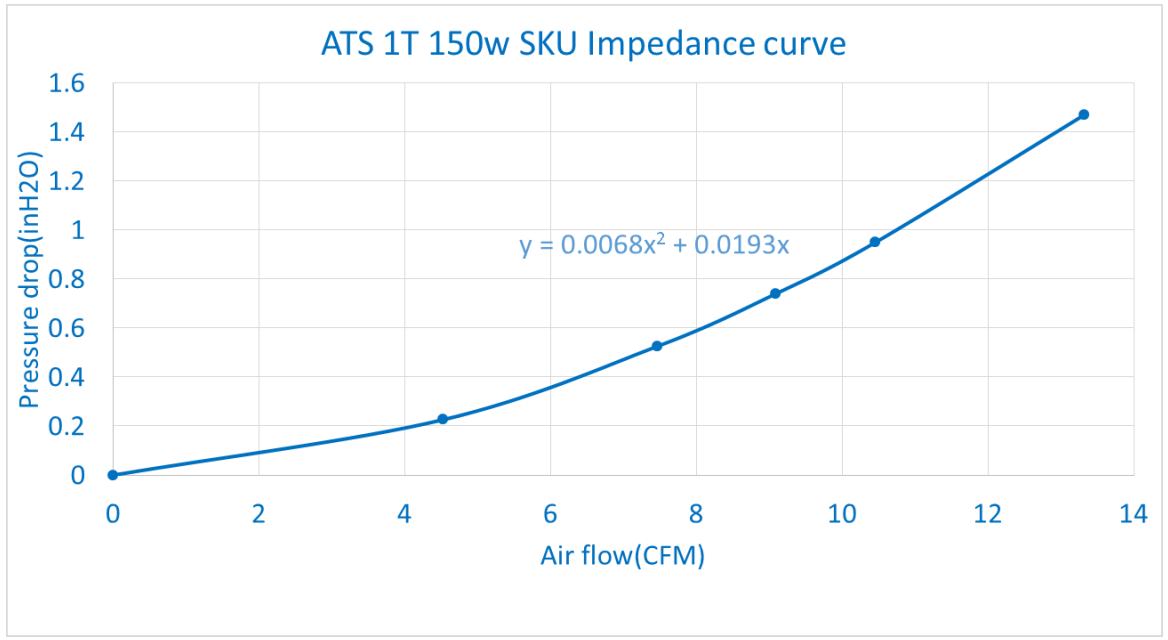






Thermal Specifications – 1T/150 W Card

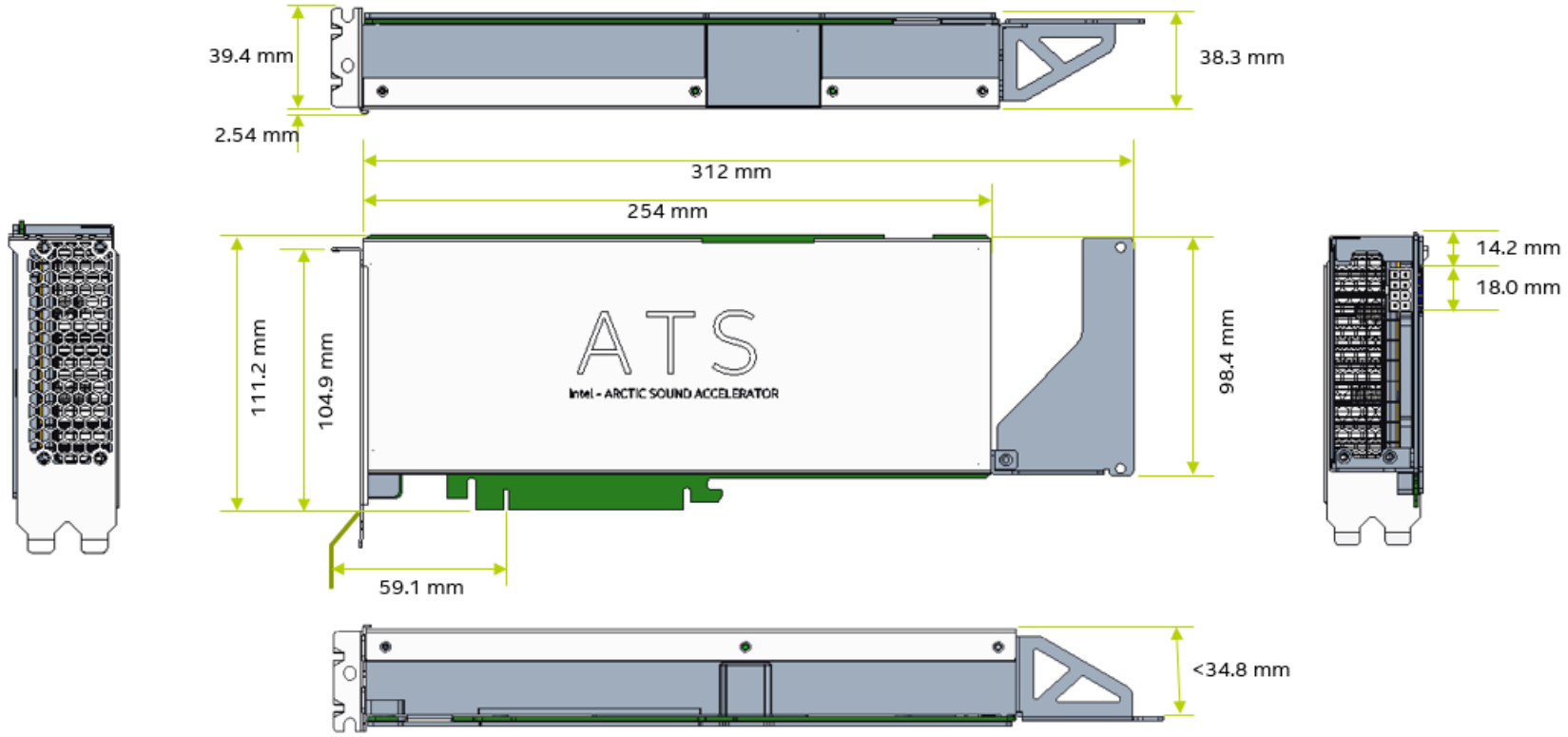




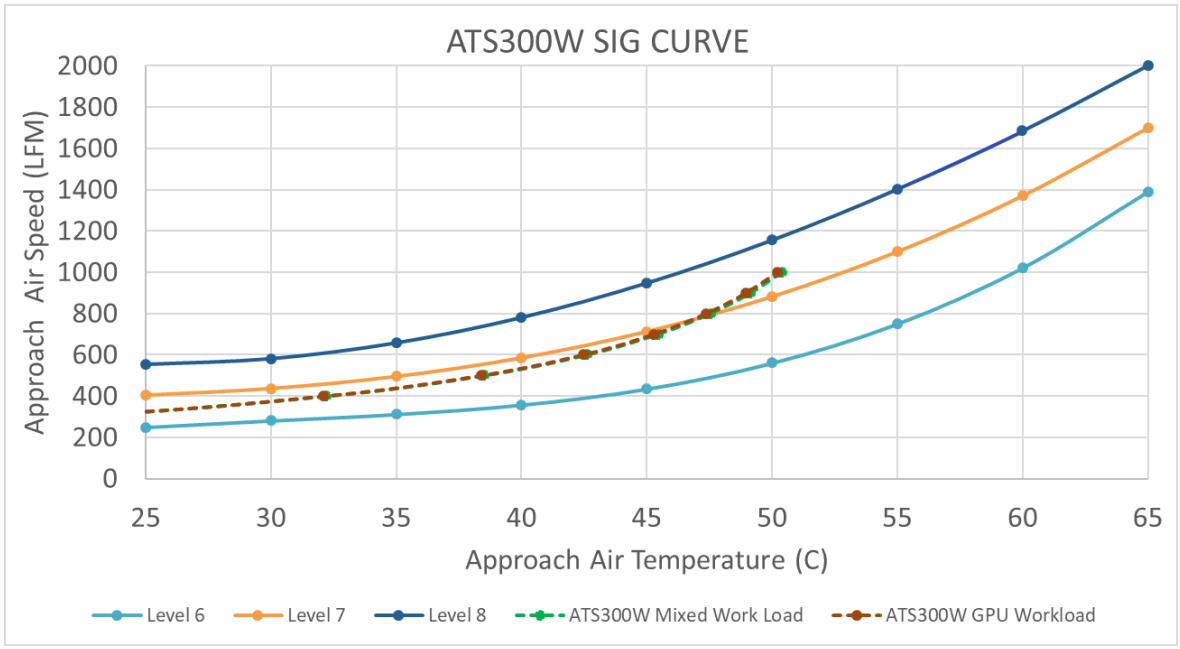
- **Graphic Processing Unit (GPU)-Centric Workloads**: Workloads are compute bound and performance scales up with more compute availability. Example: SGEMM
- **Memory-Centric Workloads**: These workloads are memory Bandwidth (BW) bound (High Bandwidth Memory [HBM]) and performance scales up with more BW availability. Example: High Performance Computing (HPC) Title (stream_triad)
- **Mixed Case**: Combination of GPU-centric and memory-centric workloads (HBM) running at the same time to produce worst-case thermals.

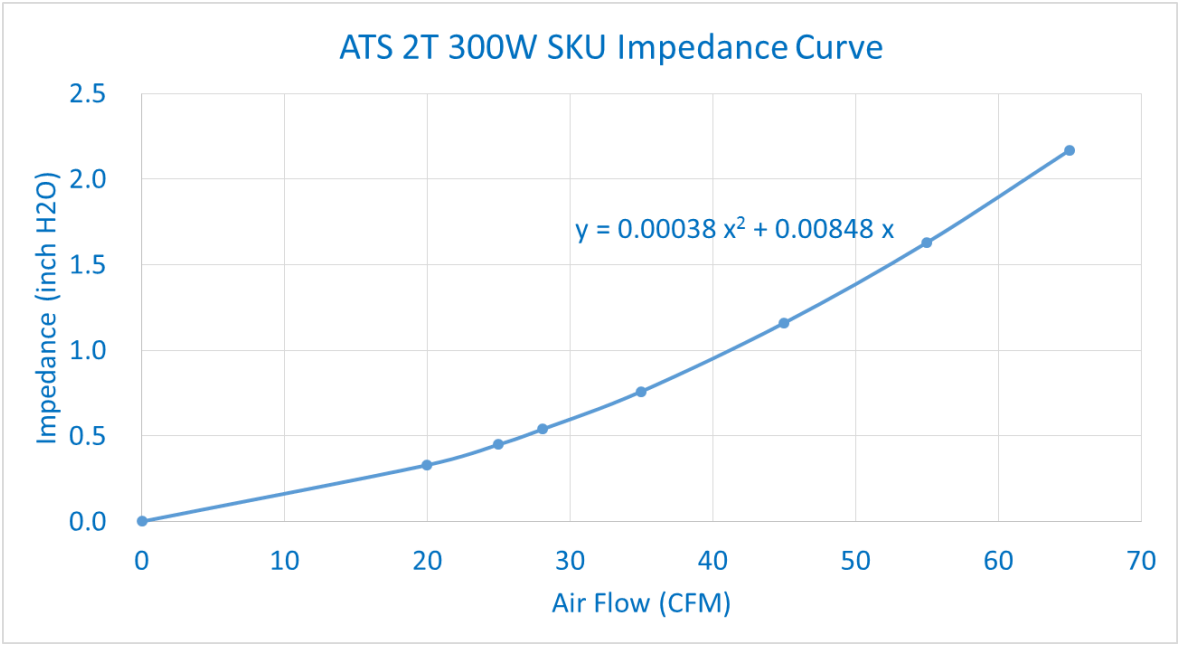
Mechanical Specifications – 2T/300W Card

- The three-quarter-length 2T/300 W card will fully comply with the Card Electric Mechanical (CEM) 4.0 specification (dimensions in mm).



Thermal Specifications – 2T/300W Card





- **Graphic Processing Unit (GPU)-Centric Workloads**: Workloads are compute bound and performance scales up with more compute availability. Example: SGEMM
- **Memory-Centric Workloads**: These workloads are memory Bandwidth (BW) bound (High Bandwidth Memory [HBM]) and performance scales up with more BW availability. Example: High Performance Computing (HPC) Title (stream_triad)
- **Mixed Case**: Combination of GPU-centric and memory-centric workloads (HBM) running at the same time to produce worst-case thermals.

Electrical Specifications

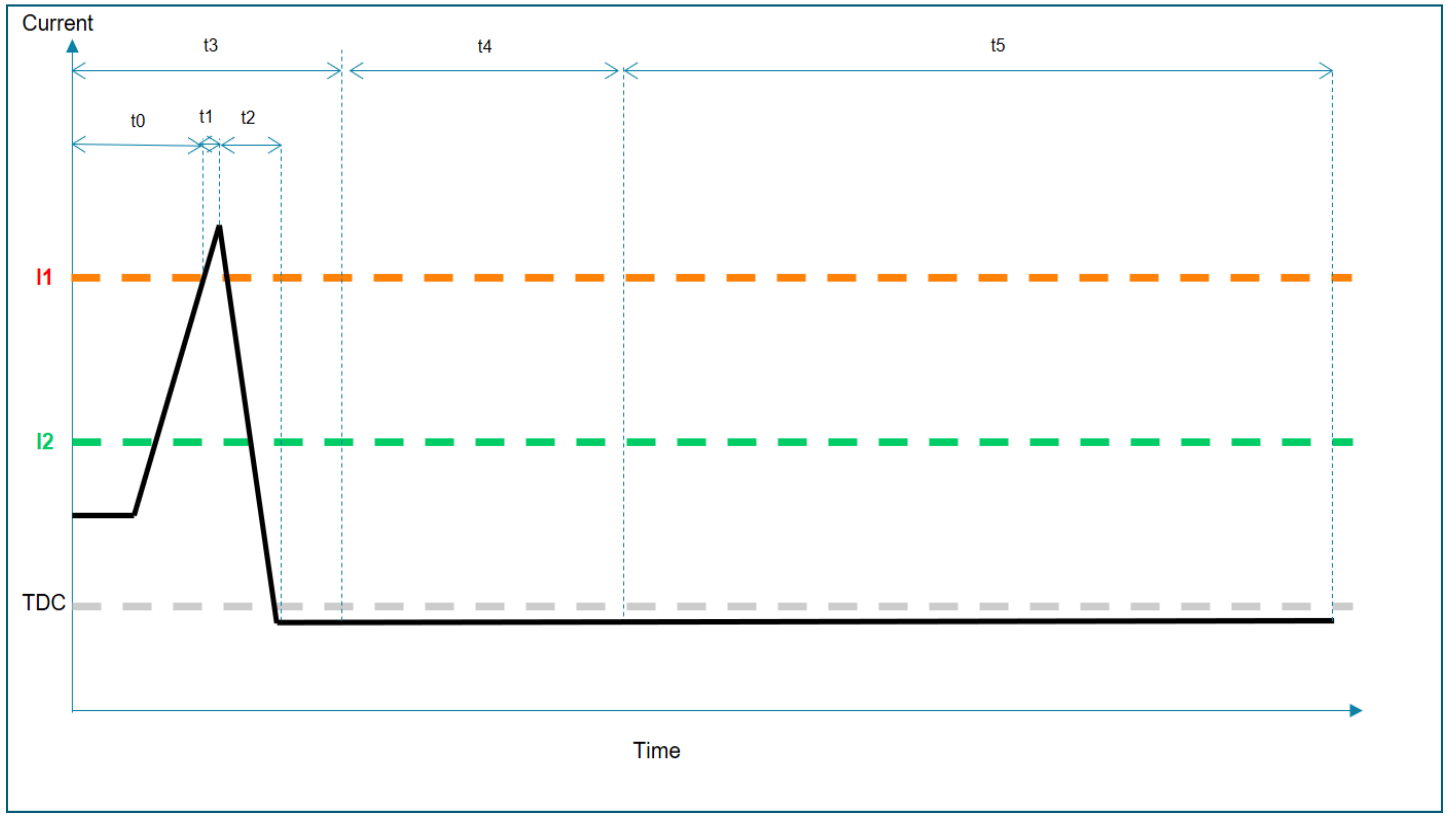
Card SKU	Power Connector
150 watts	EPS-12V 2x4
300 watts	EPS-12V 2x4

- PCI Express* (PCIe*) x16 signaling complies with the *Base 4.0 Specification*
- Optional +3.3 Vaux, Serial Management Bus (SMBus), and PWRBRK# are used for full functionality (but are not required)

Power Management

- The Psys sensor measures the total card current (from 12V edge connector and 12V auxiliary power connector) and is continually read over the Serial Voltage ID (SVID) by the master tile's PUNIT, which will throttle all tiles if set thresholds are exceeded:
  - A Thermal Design Power (TDP), which is a Running Average Power Limit (RAPL), is fused for each SKU. This can be programmed to a lower value if desired.
  - Short bursts of operation at two higher power levels are allowed. The highest level can be programmed to a lower value if desired.
  - Two power-capping loops run in parallel: a current limiter and a power limiter, as shown on the following two slides.
- The host platform can throttle all tiles by asserting the PWRBRK# signal:
  - PWRBRK# response time is 2 µs.

Current Limiter in Action

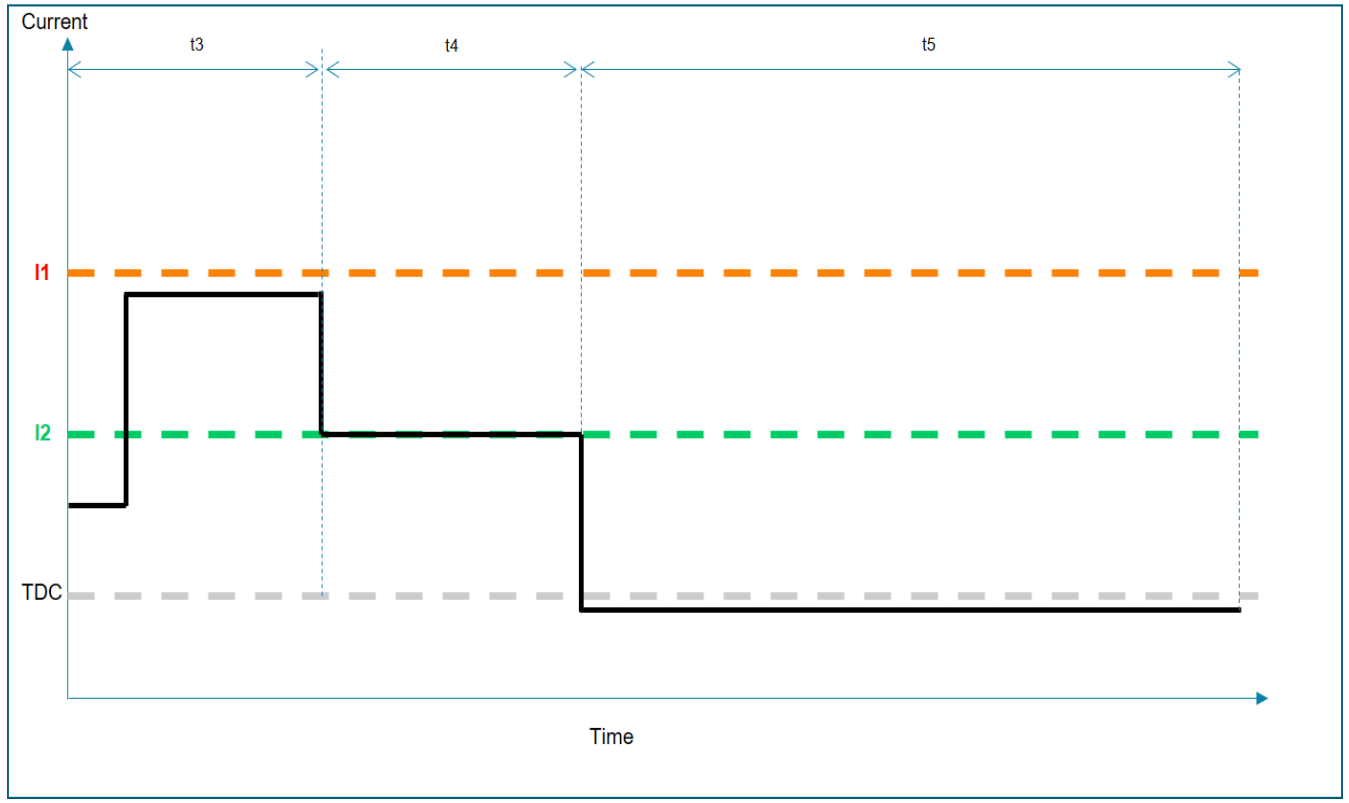


Parameter	Description
TDC/TDP[1]	Programmable sustainable current/power ≤ SKU announced power
I1	Programmable threshold ≤ 2 * TDC
I2	Intermediate level between I1 and TDP
t1	~ 2 µs
t2	~ 2 µs
t3	~ 500 µs worst case
t4	~ 500 µs worst case
t5	t3 + t4 + t5 = ~ 8 ms minimum[2]

Notes:

1. Average power will not be more than Thermal Design Power (TDP) within a 3 * Tau window (1s default, 16 ms minimum), where Tau is programmable
2. This is the cycle time for a spike/pulse above Thermal Design Current (TDC)/TDP, that is, one such is allowed every t3 + t4 + t5 ms

Power Limiter in Action



Parameter	Description
TDC/TDP[1]	Programmable sustainable current/power ≤ SKU announced power
I1	Programmable threshold ≤ 2 * TDC
I2	Intermediate level between I1 and TDP
t1	~ 2 µs
t2	~ 2 µs
t3	~ 500 µs worst case
t4	~ 500 µs worst case
t5	t3 + t4 + t5 = ~ 8 ms minimum[2]

Notes:
1. Average power will not be more than Thermal Design Power (TDP) within a 3 * Tau window (1s default, 16 ms minimum), where Tau is programmable
2. This is the cycle time for a spike/pulse above Thermal Design Current (TDC)/TDP, that is, one such is allowed every t3 + t4 + t5 ms

Thermal Management

- The PUNIT of each tile reads the on-die Digital Temperature Sensors (DTS) and manages the tile temperature accordingly by throttling Xe Architecture Engine† or the High Bandwidth Memories (HBMs).
  - The Arctic Sound-P package's THERMTRIP# signal is asserted if any System on Chip (SoC) or HBM reaches a catastrophic temperature. The PUNIT shuts down all tiles, and the Add-in-card Management Controller (AMC) shuts down all Arctic Sound-P on-card power rails and logs this event.
- Each SoC Voltage Regulator (VR)'s VR_HOT signal can assert the tile's PROCHOT# signal, which throttles the Xe Architecture Engine† on just that tile.
- Each HBM VR's VR_HOT signal can assert the tile's MEMHOT# signal, which throttles the HBMs on just that tile.
- The AMC reads the DTS sensors on each tile at a regular interval and reports it to the Baseboard Management Controller (BMC) for platform control (for example, to increase or decrease the system's fan speed).
  - It also reports the card's inlet and outlet air temperatures.

Notes: † All names are preliminary and can change.

Arctic Sound-P Reliability, Availability, and Serviceability (RAS) Features

High Bandwidth Memory (HBM) Memory Controller (MC) Protection:

- Command, control, and data parity

HBM MC ⇔ HBM Protection:

- Command / address parity
- Error Correction Code (ECC): Single-bit correction, double-bit detection
- Transient retry: Multiple reads on bit errors

HBM Error Scrubbing and Testing:

- Built-In Self Test (BIST)
- Patrol scrubbing
- Demand scrubbing

HBM Error Logging and Reporting:

- Correctable failure log
- Poison bit encoding on uncorrectable errors
- Hard errors reported to the software for sparing options
- Error messages to host via PCI Express* (PCIe*) Advanced Error Reporting (AER)
- Errors on command-and-control trigger fatal error on the PCIe* port
- Errors on data trigger uncorrectable non-fatal error on the PCIe* port

PCIe* RAS Features:

- AER
- Parity on command and data
- End-to-end Cyclic Redundancy Check (ECRC)
- Link recovery
- Link retraining

Host Platform

Hardware Considerations

- The Whitley platform (CPU: Ice Lake) will be used for the power on and validation of Arctic Sound-P:
  - Coyote Pass 2U system (accommodates two Arctic Sound-P cards)
  - PCI Express* (PCIe*) breakout box (accommodates sixteen Arctic Sound-P cards; pairs with Coyote Pass 2U system)
- The host platform should follow the PCIe* Card Electric Mechanical (CEM) 4.0 specification; in addition:
  - It should provide an EPS-12 V 2x4 auxiliary power cable
  - It should provide the thermal boundary conditions (that is, inlet air temperature and airflow) required by the specific Arctic Sound-P card SKU

BIOS Considerations

BAR Size (Physical Function)	
BAR0	16 MB/32 MB (based on SKU)
BAR1	16 GB/32 GB (based on SKU)
BAR2	3 MB

BAR Size (Virtual Function)	
BAR0	1 GB (16 MB*63)
BAR1	504 GB (8 GB*63)

- The host system BIOS must be configured to enable large Memory-Mapped Input/Output (MMIO) and allow for large per-device Base Address Register (BAR) allocations. BAR must have 64-bit address enabled.

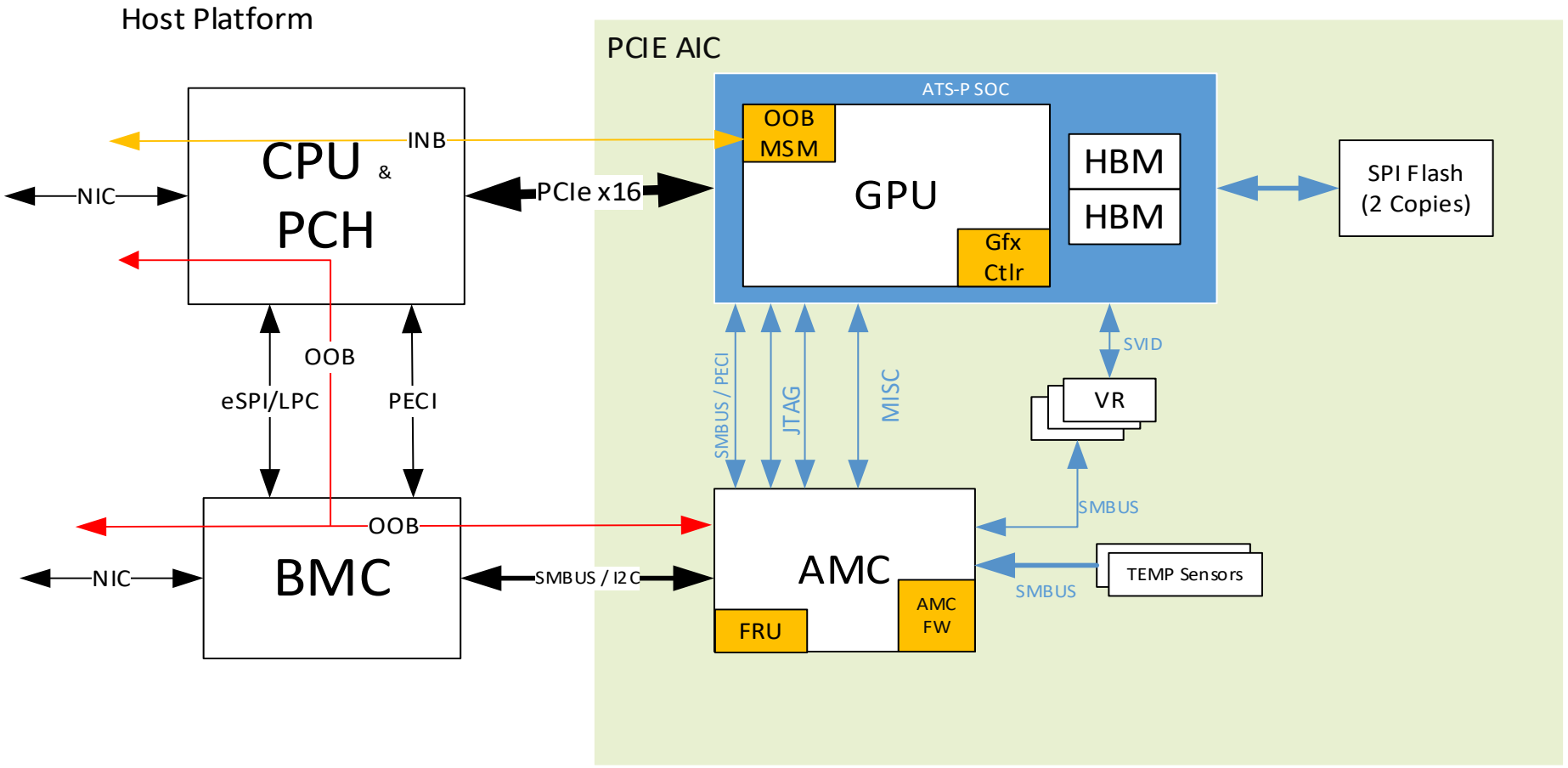
- The minimum requirements for BAR and MMIO are:
  - MMIO mapping above 4 GB is enabled
  - MMIO size is ~538 GB with Single-Root Input/Output Virtualization (SR-IOV) enabled and 63 Virtual Functions (VFs)

- For example, on Intel® Server Board S2600 WT based systems, this can be enabled in the BIOS setup by configuring the following two options on the Peripheral Configuration Interface (PCI) configuration screen:
  - Set **Memory Mapped Above 4 GB** to **Enabled**
  - Set **Memory Mapped IO Size** to **1 TB or higher**

- The specific wording in the BIOS setting of non-Intel produced motherboards may vary, but the ability to enable a large MMIO range is required for proper operation of Arctic Sound-P.

Firmware and Manageability

Arctic Sound-P Firmware and Manageability



OOB: Out of Band Management

INB: In Band Management

PCI Express* (PCIe*)

Firmware (FW)

- Arctic Sound-P FW (Serial Peripheral Interface [SPI] flash, graphics controller FW):
  - Fault tolerant (redundant image)
  - Initializes and configures Arctic Sound-P System on Chip (SoC):
    - Chassis security control
    - Memory controller
    - Out-of-Band (OOB)-MSM
    - PCI Express* (PCIe*)
    - Others
- Add-in-card Management Controller (AMC) FW
  - Fault tolerant (redundant image)
  - Powerup sequencing of SoC
  - Manageability
    - SoC and card temperature monitoring
    - Power monitoring or throttling
    - Card inventory through the Field Replaceable Unit (FRU)
    - Reading crash dump data
    - Remote flash update
    - Status and error reporting
    - Remote debug
- FRU
  - Basic inventory information: model, SKU, manufacturer, and others

Out-of-Band Management (OOB) Manageability and Debug Features under Evaluation

Features in Priority Order	Arctic Sound-P Direction
Manageability protocol	The Baseboard Management Controller (BMC) can communicate to the Add-in-card Management Controller (AMC) via Management Component Transport Protocol (MCTP) over the System Management Bus (SMBus)/I2C*, Platform Level Data Model (PLDM) over MCTP
Thermal management	The BMC can read thermal sensors (System on Chip [SoC], High Bandwidth Memory [HBM], card inlet, card outlet) The AMC drives SoC PROCHOT and monitors THERMTRIP
Power management	The BMC can read current power sensors The BMC can set power limits and emergency throttling via POWERBRK#
Field Replaceable Unit (FRU)	The BMC can read basic inventory information (model, SKU, manufacturer, and others)
AMC Firmware (FW) update	The BMC can update the AMC FW over I2C*. AMC FW is signed by Intel. Fault tolerant updates
Crashdump	The BMC can extract error status in case of boot or runtime failure
Intel® At-Scale Debug (Intel® ASD)	The BMC can forward JTAG* commands to Arctic Sound-P over the SMBus

Out-of-Band Management (OOB) Manageability

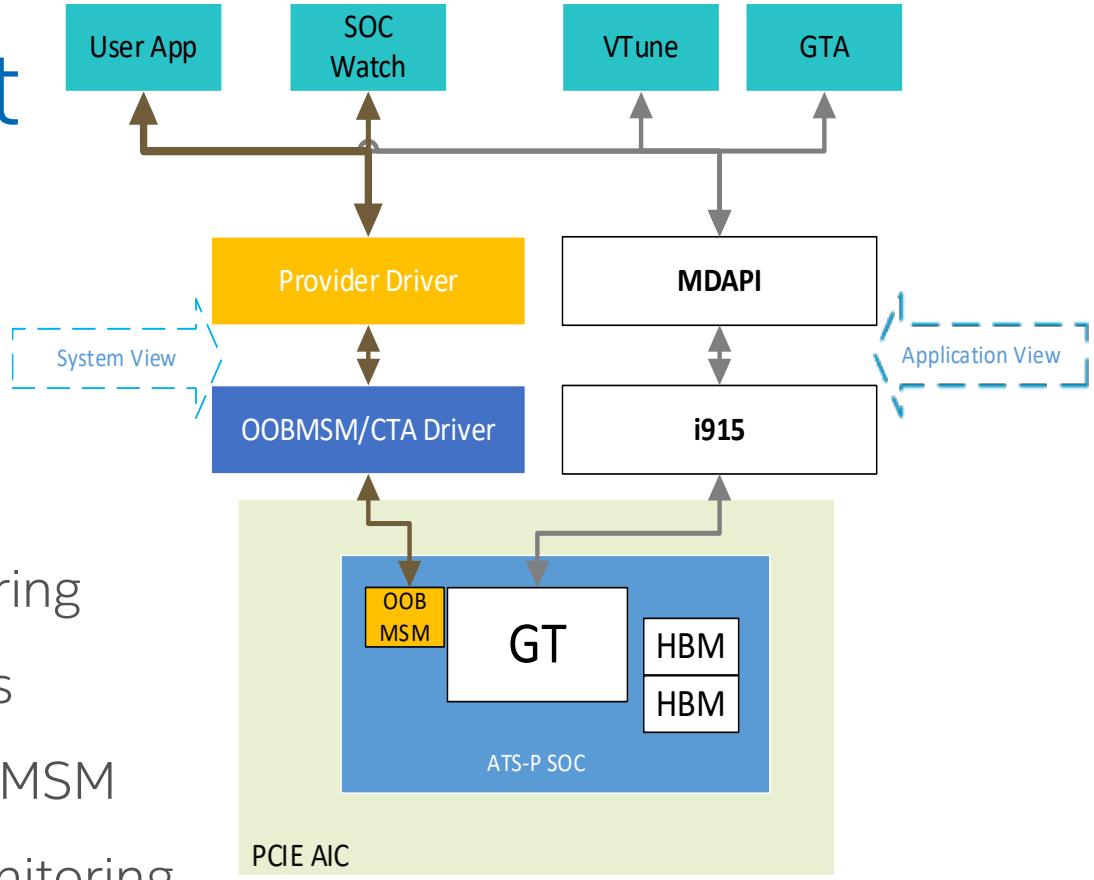
Feature	Desktop Management Task Force* (DMTF*) Specifications	Description
Baseboard Management Controller (BMC) External Interfaces	-	The BMC is to provide web interface/Redfish* interface to provide the features mentioned in this table.
Sensor Monitoring (Thermal, Energy, Thermtrip and device initialization) and user configuration of PL1/PL2, ICCmax	Management Component Transport Protocol (MCTP) Base/Serial Management Bus (SMBus) Binding + Platform Level Data Model (PLDM) Base + PLDM Control and Monitoring. Numeric Sensors – 5 to 14 thermal sensor PDRs. State sensors – Two PDRs. Effecters – Three PDRs. No sensor initialization PDRs.	The AMC exposes sensors and provides effectors to configure the System on Chip (SoC) Running Average Power Limit (RAPL) registers as defined in DSP0248 – *PLDM Monitoring and Control Specification.* No PLDM event logging. BMC to scan sensors at ~2s frequency.
Inventory: Field Replaceable Unit (FRU)	MCTP Base/SMBus Binding + PLDM Base + PLDM FRU.	Host BMC will get the card capability from FRU through SMBus (PLDM FRU over MCTP over SMBus).
Add-in-card Management Controller (AMC) Firmware update	MCTP Base/SMBus Binding + PLDM Base + PLDM Firmware Update.	BMC to provide UA capabilities. AMC will implement FD capabilities.
AMC Remote Debug	MCTP Base/SMBus Binding + Vendor-defined Intel® At-Scale Debug (Intel® ASD) messaging (*DoPCIe Specification*)	BMC to facilitate remote debug on Arctic Sound-P using AMC's JTAG* bit bang driver.
AMC Crash dump	MCTP Base/SMBus Binding + Vendor defined crash dump	AMC to provide crash dump on BMC's request.

Desktop Management Task Force* (DMTF*) Specifications

DMTF* Specification	Description	Arctic Sound-P Context
DSP0236	*MCTP Base Specification* 1.3.0	Defines the Management Component Transport Protocol (MCTP) control messages and endpoint device functionalities
DSP0237	*MCTP SMBus/I2C Transport Binding Specification* 1.1.0	Defines how MCTP messages are carried over the I2C* bus
DSP0239	*MCTP IDs and Codes* 1.6.0	Defines the MCTP message types
DSP0240	*Platform Level Data Model (PLDM) Base Specification* 1.0.0	Defines the base PLDM architecture, packet format, and mandatory commands
DSP0241	*Platform Level Data Model (PLDM) Over MCTP Binding Specification* 1.0.0	Describes how the PLDM traffic is carried over the MCTP transport
DSP0245	*Platform Level Data Model (PLDM) IDs and Codes Specification* 1.3.0	PLDM message types
DSP0248	*Platform Level Data Model (PLDM) Platform Monitoring and Control Specification* 1.1.1	Defines how the Add-in-card Management Controller (AMC) can export Arctic Sound-P sensors (entities that provide status, for example: temperature) and effectors (control knobs) to the Baseboard Management Controller (BMC)
DSP0249	*Platform Level Data Model (PLDM) State Set Specification* 1.0.0	State set definitions
DSP0257	*Platform Level Data Model (PLDM) for FRU Data Specification* 1.0.0	Defines how the AMC can export the Arctic Sound-P card's Field Replaceable Unit (FRU) information to the BMC. See section PLDM FRU implementation
DSP0267	*Platform Level Data Model (PLDM) for Firmware Update Specification* 1.0.1	Defines the firmware update methodology
Vendor Defined	*DoPCIe for Remote Debug*	Card Intel® At-Scale Debug (Intel® ASD) enablement
Vendor Defined	*Crash Dump*	Crash dump collection

In-Band (INB) Management (1/2)

- Over PCI Express* (PCIe*) x16
- INB application view through MDAPI
  - MDAPI: Metrics Discovery API
  - GFX: Abstraction layer for performance monitoring
  - VTune™, GTA, System on Chip (SoC) watch tools
- INB local system view through Out-of-Band (OOB)-MSM
  - Used for load balancing, orchestration, and monitoring
  - Part of Converged Telemetry Architecture (CTA) across multiple IPS
  - Intel provides OOB-MSM library functions, and customers can write their own provider driver and API for applications.



In-Band (INB) Management (2/2)

- INB local system view through Out-of-Band (OOB)-MSM
  - Per High Bandwidth Memory (HBM) stack, temperature, reads and writes
  - Per PCI Express* (PCIe*) link, reads and writes
  - Per tile, tile maximum, GT maximum
  - Per tile, GT throttle cycles (Thermal, Thermal Design Power [TDP], ICCmax)
  - Per tile, margin to PROCHOT, TCONTROL
  - Per tile, total energy (GT + HBM + Miscellaneous)

Software

Targeted OS for Arctic Sound-P

OS	Version	Bare Metal and Guest OS Support
Red Hat* Enterprise Linux* (RHEL*)	8.x	Yes
SLES*	SLE15.SPx	Yes
Ubuntu*	20.10 LTS and 21.04 or newer	Yes
CentOS*	8.x	Yes

Hypervisor	Version
KVM*	TBD
Xen*	TBD
VMware*	TBD

Note: OS targets may subject to change.

Arctic Sound-P Target Deployment Models

- Bare Metal
  - Host/Native OS
- Machine Images
  - Binary image for launching in virtual machine
  - Contains a guest OS with system configuration
    - Vendor specific images
    - Generic images
- Containers
  - Package software into standardized units for development, shipment, and deployment

Terminology

Term	Definition	Term	Definition
ALU	Arithmetic and Logic Unit	GPU	Graphics Processing Unit
AMC	Add-In-Card Management Controller	GUNIT	Graphics Unit
ATS-P	Arctic Sound-P (Intel Graphics SOC)	HBM	High Bandwidth Memory
BMC	Baseboard Management Controller	HPC	High Performance Computing
BW	Bandwidth	IEH	Integrated Error Handler
CEM	Card Electromechanical (the standard PCIe card form-factor)	JTAG*	Joint Test Action Group*
CFM	Cubic Feet per Minute	KOZ	Keep Out Zone
DTS	Digital Temperature Sensors	MC	Memory Controller
EU	Execution Unit	MCTP	Management Component Transport Protocol
EVAC	Extended Volume Air Cooling	MDFI	Modular Die Fabric Interface
FF	Form Factor	MFX	Multi-Format Codex
FRU	Field Replaceable Unit	ML	Machine Learning

Terminology continued

Term	Definition	Term	Definition
OOB	Out Of Band	SVID	Serial VID
PCIe*	PCI Express*	TDC	Thermal Design Current
PSU	Power Supply Unit	TDP	Thermal Design Power
PUNIT	Power Unit	Tla	Temperature of the local ambient air
RAPL	Running Average Power Limit	VD	Video Decode
SFC	Scaler and Format Conversion	VME	Video Motion Estimation
SGEMM	Single precision floating General Matrix Multiply	VQE	Visual Quality Enhancement
SIMD	Single Instruction Multiple Data	VR	Voltage Regulator
SKU	Stock Keeping Unit	TDC	Thermal Design Current
SMBUS	System Management Bus	TDP	Thermal Design Power
SoC	System on Chip	Tla	Temperature of the local ambient air
SPI	Serial Peripheral Interface	VD	Video Decode


intel