NVIDIA Jetson AGX Orin Series

Ampere GPU + Arm Cortex-A78AE CPU + LPDDR5 + 64GB eMMC5.1

Data Sheet

NVIDIA Jetson AGX Orin Modules

Module	Description
Jetson AGX Orin 64GB (JAO 64GB)	Ampere GPU + Arm Cortex-A78AE CPU + 64GB LPDDR5 + 64GB eMMC 5.1
Jetson AGX Orin 32GB (JAO 32GB)	Ampere GPU + Arm Cortex-A78AE CPU + 32GB LPDDR5 + 64GB eMMC 5.1

References to JAO and Jetson AGX Orin include and can be read as Jetson AGX Orin 64GB and Jetson AGX Orin 32GB except where explicitly noted.

AI Performance

JAO 64GB: Up to 275 Sparse TOPS (INT8)

JAO 32GB: Up to 200 Sparse TOPs (INT8)

Ampere GPU

JAO 64GB: 2 GPC | 8 TPC | 2048 NVIDIA® CUDA® cores | 64 Tensor cores Ray-Tracing cores | 170 Sparse TOPS | Maximum Operating Frequency: 1.3 GHz

JAO 32GB: 2 GPC | 7 TPC | | 1792 NVIDIA® CUDA® cores | 56 Tensor cores Ray-Tracing cores | 108 Sparse TOPS Maximum Operating Frequency: 939 MHz

JAO: End-to-end lossless compression | Tiled Caching | OpenGL® 4.6+ | OpenGL ES 3.2 | Vulkan™ 1.2+◇ | CUDA 10.2+ | Maximum Operating Frequency: 1.3 GHz

Arm Cortex-A78AE CPU

Arm v8.2 (64-bit) heterogeneous multi-processing (HMP) CPU architecture |

JAO 64GB: 12x cores | 3 CPU clusters (4 cores/cluster) | 259 SPECint_rate2006

JAO 32GB: 8x cores | 2 CPU clusters (4 cores/cluster) | 177 SPECint_rate2006

JAO: L1 Cache: 64 KB L1 instruction cache (I-cache) + 64 KB L1 data cache (D-cache) per CPU core | L2 Cache: 256 KB per CPU core | L3 Cache: 2MB per CPU cluster | Maximum Operating Frequency: 2.2 GHz

DL Accelerator

JAO: 2x NVDLA 2.0 Engines |

JAO 64GB: Maximum Operating Frequency: 1.6 GHz | 52.5 TOPS each(Sparse INT8)

JAO 32GB: Maximum Operating Frequency: 1.4 GHz | 46 TOPs each (Sparse INT8)

Memory

JAO 64GB: 64GB 256-bit LPDDR5 DRAM

JAO 32GB: 32GB 256-bit LPDDR5 DRAM

JAO: Secure External Memory Access Using TrustZone® Technology | System MMU | Maximum Operating Frequency: 3200 MHz

Storage

64GB eMMC 5.1 Flash Storage | Bus Width: 8-bit | Maximum Bus Frequency: 200 MHz (HS400 or HS533)

64MB NOR Boot Flash | 8MB NOR Secure Key Flash

Display Controller

1x shared HDMI 2.1, eDP1.4, VESA DisplayPort 1.4a HBR3

Maximum Resolution (eDP/DP/HDMI): (up to) 8K60 (up to 36 bpp) | Multiple displays can be supported over DP interface with MST

Multi-Stream HD Video and JPEG

Video Decode: H.265 (HEVC), H.264, AV1, VP9, VP8, MPEG-4, MPEG-2, VC-1

Video Encode: H.265 (HEVC), H.264, AV1

JPEG (Decode & Encode)

Optical Flow Accelerator
- o   Optical Flow
- o   Stereo Disparity Estimation

Audio

Dedicated programmable audio processor | ARM Cortex A9 with NEON | PDM in/out | Industry-standard High-Definition Audio (HDA) controller provides a multi-channel audio path to the HDMI® interface

Imaging

16x lanes total | D-PHY v2.1 (40 Gbps)

16x trio links total | C-PHY v2.0 (164 Gbps)

Networking

1x GbE | 4x 10GbE

Peripheral Interfaces

USB: xHCI host controller with integrated PHY (up to) 3x USB 3.2 Gen2 (10Gbps), 4x USB2.0 | PCIe Gen4: 1 x8, 2 x4, 2 x2, 2 x1 | SD/MMC controller (supporting eMMC 5.1, SD 4.0, SDHOST 4.0 and SDIO 3.0) | 4x UART | 3x SPI | 8x I²C | 2x CAN | 4x I²S | 2x DMIC | 1x DSPK | GPIOs

Mechanical

Module Size: 100.0 mm x 87.0 mm x 16.0 mm | 699 pin B2B Connector | Integrated Thermal Transfer Plate (TTP) with Heatpipe

Operating Requirements

TTP Surface Temperature*: -25 °C to 80 °C | TTP Surface: 80 °C max Power Input: 5V (MV) and 12V (HV) | Operating Lifetime (24x7): 5 years

JAO 64GB Maximum Module Power: Up to 60 W

JAO 32GB Maximum Module Power: Up to 40W

Note: The Jetson AGX Orin Developer Kit can be used to develop the Jetson AGX Orin Series. The Developer Kit has the full GPU, CPU, DLA, NVENC, and NVDEC performance of Jetson AGX Orin 64GB, but with 32 GB memory. Refer to the "Software Features" section of the latest *L4T Developer Guide* for a list of supported features; all features may not be available.

◇ Product is based on a published Khronos Specification and is expected to pass the Khronos Conformance Process. Current conformance status can be found at www.khronos.org/conformance.

*   See the *NVIDIA Jetson AGX Orin Thermal Design Guide* for details

Document History

DS-10662-001_v1.0

Version	Date	Description of Change
1.0	March 18, 2022	Initial release

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

The NVIDIA® Jetson® AGX Orin System-on-Module (SOM) blends industry-leading performance, power efficiency, integrated deep learning capabilities, and rich I/O to enable emerging technologies with compute-intensive requirements. The Jetson AGX Orin SOM is designed for a wide variety of applications requiring varying performance metrics.

Table 1-1.   Jetson AGX Orin SOM Product Summary

Specification	Description
Total module power	JAO 64GB: 15 W | 30 W | 50 W, and up to 60 W
	JAO 32GB: 15 W | 30 W | 40 W
CPU	JAO 64GB: Arm® v8.2 (64-bit) | 12× (up to 6× lock step) Arm Cortex-A78AE cores | 3 CPU clusters (4 cores/cluster) | 259 SPECint_rate2006
	JAO 32GB: Arm® v8.2 (64-bit) | 8× Arm Cortex-A78AE cores | 2 CPU clusters (4 cores/cluster) | 177 SPECint_rate2006
GPU	JAO 64GB: Ampere GPU 2 GPC | 8 TPC | Up to 170 INT8 Sparse TOPS or 85 FP16 TFLOPS (Tensor Cores) | Up to 5.32 FP32 TFLOPS or 10.649 FP16 TFLOPS (CUDA cores)
	JAO 32GB: Ampere GPU 2 GPC | 7 TPC | Up to 108 INT8 Sparse TOPS or 54 FP16 TFLOPS (Tensor Cores)
	Up to 3.365 FP32 TFLOPS or 6.73 FP16 TFLOPS (CUDA cores)
Vision and DNN accelerators	Deep Learning Accelerator (DLA)
	JAO 64GB: Up to 105 INT8 TOPS (Sparse, Deep Learning Inference)
	JAO 32GB: Up to 92 INT8 TOPS (Sparse, Deep Learning Inference)
	JAO: 2 MB dedicated SRAM
	Programmable Vision Accelerator (PVA) | Up to 512 INT16 GMACS or 2048 INT8 GMACS | 2 MB dedicated SRAM
Platform security controller	RISC-V subsystem, PKC crypto (RSA3K)
Memory	JAO 64GB: 64 GB LPDDR5
	JAO 32GB: 32 GB LPDDR5
Display	1×: shared HDMI™ 2.1, eDP1.4, VESA® DisplayPort™ (DP) HBR3
Storage	64 GB eMMC 5.1
	64 MB NOR Boot Flash
	8 MB NOR Secure Key Flash

Specification	Description
Encoder/Decoder	JAO 64GB: 2×: 4K60 (H.265) /1x: 8K30 (H.265)
	JAO 32GB: 1x: 4K60 (H.265) /2x: 4K60 (H.265)
CSI	16×: CSI Lanes D-PHY v2.1	C-PHY v2.0
Ethernet	RGMII
	4×: MGBE
PCIe	PCI Express 4.0 x1, x2, x4 and x8
	Up to 2x Endpoint supported
USB	3x: USB 3.2
	4x: USB 2.0
Other	UART, SPI, CAN, I2C, I2S, GPIOs
Module dimensions	100.0 mm × 87.0 mm × 16.0 mm
Temperature range (at Thermal Transfer Plate (TTP) surface)	-25 °C to 80 °C
Operating humidity	5% to 85% RH
Storage temperature (Ambient)[1]	-25 °C to 80 °C
Storage humidity	30% to 70% RH

Table 1-2.   NVIDIA Orin SoC Features on Jetson AGX Orin SOM

Description	Orin SoC Features on Jetson AGX Orin SOM
NVIDIA Ampere GPU	
Advanced GPU including CUDA cores, Ray-Tracing (RT) cores and 3rd Generation Tensor cores	Enhanced compute capability	
JAO 64GB: 2 GPC	8 TPC	Up to 170 INT8 Sparse TOPS or 85 FP16 TFLOPS	Up to 5.32 FP32 TFLOPS or 10.649 FP16 TFLOPS (CUDA cores)	
JAO 32GB: 2 GPC	7 TPC	Up to 108 INT8 Sparse TOPS or 54 FP16 TFLOPS	Up to 3.365 FP32 TFLOPS or 6.73 FP16 TFLOPS (CUDA cores)	
Arm Cortex-A78AE CPU	
JAO 64GB:12× Arm Cortex-A78AE cores	3 CPU clusters (4 cores/cluster)	259 SPECint_rate2006	
JAO 32GB: 8× Arm Cortex-A78AE cores	2 CPU clusters (4 cores/cluster)	177 SPECint_rate2006	
Arm® v8.2 (64-bit)	Symmetric multi-processing (SMP)	NEON SIMD	High-performance coherent interconnect fabric	
L1 Cache: 64 KB Instruction Cache (I) + 64 KB Data Cache (D) per CPU core	
L2 Cache: 256 KB per CPU core	
L3 Cache: 2 MB per CPU cluster	
Vision and DNN Accelerators	
2x Deep Learning Accelerator (DLA)	2MB dedicated SRAM JAO 64GB: Up to 105 INT8 TOPS (Deep Learning Inference) JAO 32GB: Up to 98 INT8 TOPS (Deep Learning Inference)	
Programmable Vision Accelerator (PVA)	Up to 512 INT16 GMACS or 2048 INT8 GMACS	2MB dedicated SRAM	

Description	Orin SoC Features on Jetson AGX Orin SOM
Platform Security Controller (PSC)	
RISC-V subsystem, PKC crypto (RSA3K)	
Sensor Processing Engine (SPE)	
Arm Cortex R5 processor	Always-on (AON) cluster
Boot and Power Management Processor (BPMP) complex	
Cold and Warm Boot | System power state transitions | Voltage / Frequency management	
On-Chip Memory	
System Cache: 4 MB	
Memory Subsystem	
Memory Type	256-bit LPDDR5
Maximum Memory Bus Bandwidth (up to)	204.8 GB/s
Maximum Capacity	JAO 64GB: 64 GB
	JAO 32GB: 32 GB
Memory encryption support | 2-stage CPU MMU and SMMU, GPU MMU | Lossless memory bandwidth compression for GPU R/W	
Multi-Steam HD Video Decode	
H.264, H.265 (HEVC), VP8, VP9, VC1, AV1, MPEG2, MPEG4	
Multi-Stream HD Video Encode	
H.264, H.265 (HEVC), AV1	
JPEG Image Encode/Decode	
Supported	
Audio Subsystem (APE)	
ADSP: Arm Cortex A9 processor with NEON	
APE subsystem contains Audio HUB | Audio DMA | 4× DAP | 16 slot TDM | PDM (multiple TX/RX)	
2× DAP additional ports available outside APE for functional safety audio or other purposes	
Display Controller Subsystem	
2× display heads (pipelines) sharing 4 surfaces	1× shared HDMI v2.1 / DP v1.4a / eDP v1.4
DP supports MST, HBR3, VESA DSC per head	HDCP v2.2 and v1.4
Imaging System	
ISP (up to 1.85 GPix/s) with HDR support	6 MIPI CSI-2 v3.0 links
D-PHY	16× lanes total | D-PHY v2.1 up to 2.5 Gbps/lane | 40 Gbps total
C-PHY	16× trio links total | C-PHY v2.0 up to 4.5GSps (10.25 Gbps)/trio | 164 Gbps total
Video Imaging Compositor (VIC) 2D Engine	
Gen 4.2 VIC | 2 GPix/s | 16-surface blending | Lens distortion correction, HiQ scaling, HiQ deinterlacing, blending, rotation, cadence detection, temporal noise filtering, pixel/memory format conversions	
Boot Sources	

Description	Orin SoC Features on Jetson AGX Orin SOM
QSPI serial flash, USB (Recovery Mode)	
Security	
Security boot, Arm TrustZone TEE, secure memory (TZ SRAM, DRAM protection with memory encryption), Platform Security Controller, hardware symmetric/asymmetric crypto acceleration, hardware root-of-trust, physical attack protection, secure debug (DFD) and test (DFT), life cycle management	
Storage Interfaces	
2 × SD/MMC controllers (1 supporting eMMC 5.1 (used on-module), both supporting SD 4.2 and SDIO 4.1) | 1 × QSPI (used on module)	
Peripheral Interfaces	
XHCI USB host controller with integrated PHY: up to 3× USB 3.2 Gen 2 (10 Gbps), up to 4× USB 2.0 | USB device controller for 1× USB 3.2 Gen 1 SS (5 Gbps) and 1× USB 2.0 |	
PCIe (7× controllers, 22 shared lanes, up to Gen 4 (16 Gbps/lane))	
4× UART | 3x SPI | 8x I2C | 4x DAP ports: support I2S, RJM, LJM, PCM, TDM (multi-slot mode) | 2× PDM (DMIC) | 1x DSPK | 2× CAN (LS, FD) | 1× ETHER_QOS (RGMII) with AVB support | 4× MGBE (XFI) with AVB support | 4× PWM	

Notes:

1. Storage and peripheral interfaces are subject to pin-muxing. Not all interfaces are available in the same system design or available simultaneously.

2. Simultaneous support of various functions depends on the application use case and is subject to availability of memory bandwidth.

Chapter 2. Functional Description

The NVIDIA Jetson AGX Orin SOM is a high performance, small-form factor (SFF) device. It enables modular system design by mechanically isolating integrated components from external mechanical forces, standardizing thermal and mechanical interfaces, and exposing a comprehensive set of system and peripheral interfaces at the 699-pin board-to-board connector. The NVIDIA Jetson AGX Orin SOM can be used in a wide variety of applications requiring varying performance metrics. To accommodate these varying conditions, NVIDIA Jetson AGX Orin SOM implement a multitiered solution that focuses on the efficient application of performance to manage a complex environment:

▶ **Power Management Controller (PMC):** The PMC primarily controls voltage transitions for the NVIDIA Orin™ SoC as it transitions to and from different low-power modes. It also acts as a target receiving dedicated power and clock request signals as well as wake event from dedicated GPIO, which can wake the module from a deep sleep state.

▶ **Power Gating:** NVIDIA Jetson AGX Orin SOM aggressively employ power-gating (controlled by the PMC) to power-off blocks that are idle. CPU cores are on a separate power rail to allow complete removal of power and eliminate leakage. Each CPU can be power gated independently internally. Software provides context save and restore to and from DRAM.

▶ **Dynamic Voltage and Frequency Scaling (DVFS):** Raises voltages and clock frequencies when demand requires, lowers them when less is sufficient, and removes them when none is needed. DVFS is used to change the voltage and frequencies on the following rails:

- VDD_CPU
- VDD_GPU
- VDD_CV

▶ **Real Time Clock (RTC):** The RTC Always On partition logic of the CPU Complex is not power gated. It can wake the system based on either a timer event or an external trigger (for example, key press). Wake on RTC Alarm is NOT supported on PMIC while the PMIC is in the Global Shutdown power state.

NVIDIA Jetson AGX Orin SOM has three power inputs:

▶ SYS_VIN_HV (7V to 20V input)

▶ SYS_VIN_MV (A 5V regulated input)

▶ PMIC_BBATT (1.65V to 3.45V input) for RTC backup

Power is then supplied to the devices on board through a power sequencer IC and dedicated voltage regulators. All internal module voltages and I/O voltages are generated from these inputs. An optional back up battery can be attached to the **PMIC_BBATT** module input (this will maintain the on system RTC, when VIN is not present). **SYS_VIN_MV** and **SYS_VIN_HV** must be supplied by the carrier board that the NVIDIA Jetson AGX Orin SOM is designed to connect to.

Figure 2-1.   System Block Diagram
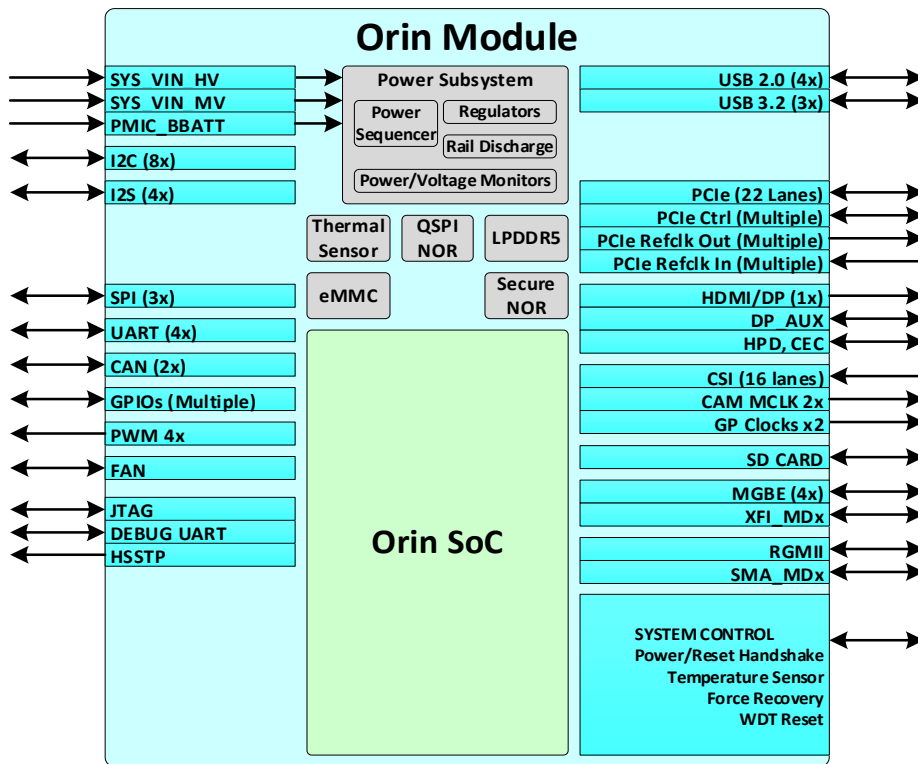
Chapter 3. Power and System Management

3.1 Input Power

The NVIDIA Jetson AGX Orin SOM has three power inputs:

▶ SYS_VIN_HV: 7V–20V

▶ SYS_VIN_MV: 5V

▶ PMIC_BBATT: (1.65V to 3.45V input) for RTC backup

Power is then supplied to the devices on board through power sequencer IC and dedicated voltage regulators. All internal module voltages and I/O voltages are generated from these inputs. Input powers must be supplied by the carrier board that the Orin Module is designed to connect to.

`PMIC_BBATT` provides power for RTC backup.

The input voltage measured at the module connector should never exceed the voltage range defined in Table 7-2.

3.2 Power Sequencing

NVIDIA Jetson AGX Orin SOM and the product carrier board must be power sequenced properly to avoid potential damage to components on either the module or the carrier board system. The module is powered before the main carrier board circuits. Refer to the *NVIDIA Jetson AGX Orin Design Guide* for system level details on the application of power, power-up sequencing, power-down sequencing, and power monitoring.

3.3 Power States

The NVIDIA Jetson AGX Orin SOM operates in three main power modes: ON, Deep Sleep State (SC7), and OFF.

3.3.1   ON State

The ON power state is entered from OFF state. In this state, NVIDIA Jetson AGX Orin SOM is fully functional and will operate normally. An ON event must occur for a transition between OFF and ON states. The **VDDIN_PWR_BAD_N** control is the carrier board indication to the NVIDIA Jetson AGX Orin SOM that the VIN power is good. The carrier board should assert this high only when VIN has reached its required voltage level and is stable. This prevents NVIDIA Jetson AGX Orin SOM from powering up until the VIN power is stable.

3.3.2   SC7 – Deep Sleep State

All CPU cores are powered off and software execution is suspended. System state is preserved in DRAM, which is put in self-refresh mode. Most I/Os and internal blocks are powered off.

Transitioning to a sleep state involves (among other things) the following:

▶ Freezing all running applications

▶ Synchronizing file system contents to storage devices

▶ Suspending individual device drivers and saving their state in DRAM

▶ Putting DRAM in self-refresh mode

▶ Powering off various Orin blocks

▶ Inactivity timeout, no CPU process needed, no devices are active

▶ OS is suspended

▶ CPU, SOC, GPU and CV power rails are OFF

▶ PMC and RTC still available

▶ Wake event is triggered through GPIO pins (Refer to the *Jetson AGX Orin Pinmux* for more details)

- PADs are powered off except for PADs which monitor wake events

3.3.3   OFF State

The OFF state is the default state when the system is not powered. It can only be entered from the ON state or through an OFF event.

Table 3-1.   OFF Events

Event	Details	Preconditions
Power OFF	VIN power is disconnected or MODULE_POWER_ON signal going low	In ON state
Software shutdown	Software will initiate	ON state, software operational
Thermal shutdown	If the internal temperature of the module reaches an unsafe temperature, the hardware is designed to initiate a shutdown	Any power state

Chapter 4. NVIDIA Orin SoC Overview

At the heart of the NVIDIA Jetson AGX Orin SOM, is the NVIDIA Orin system on chip (SoC). NVIDIA Orin™ is a versatile SoC appropriate for a wide variety of perception and general compute tasks. High-level architecture of the SoC, is organized into three main processing complexes: CPU, GPU, and hardware accelerators.

CPUs include the Arm Cortex-A78AE based main CPU complex, which provides the general-purpose high-speed computing capability.

The graphics processing unit (GPU) is an NVIDIA Ampere Architecture GPU. It provides advanced parallel-processing computing capability for the CUDA language. It supports rich range of tools from NVIDIA such as NVIDIA® TensorRT™, a deep learning inference optimizer and runtime that delivers low latency and high-throughput. Ampere also provides state-of-the-art graphics capabilities including real-time ray-tracing.

The domain-specific hardware accelerators (DSAs) are a set of special-purpose hardware engines. They intended to offload a variety of computing tasks from the computing engines, and to perform these with high throughput and power efficiency.

The premium performance and integrated capabilities of this purpose built SoC, coupled with its rich I/O, reduces complexity in system integration making the Orin SoC the ideal choice for variety of complex applications.

4.1 NVIDIA Ampere GPU

The NVIDIA Ampere GPU introduces a new design for the Streaming Multiprocessor (SM) that dramatically improves performance per watt and performance per area, along with supporting 3rd generation tensor cores and TensorRT cores. Ampere GPUs improve on the previous NVIDIA Turing™ generation; and are software compatible so that the same APIs are used.

The NVIDIA Ampere Architecture GPU has a number of enhancements for compute and graphics capability that include:

▶ Sparsity: fine grained structured sparsity doubles throughput and reduces memory usage.

▶ 2× CUDA floating-point performance: higher compute math speed.

▶ SM architecture improves bandwidth to the L1 cache and shared memory and reduces L1 miss latency.

▶ Improved async compute, and post-L2 cache compression compared to NVIDIA Turing.

4.1.1   Compute Features

Ampere introduces third-generation NVIDIA Tensor Cores which offer a wider range of precisions including TensorFloat-32 (TF32), bfloat16, FP16, and INT8 all of which provide unmatched versatility and performance.

TensorFloat-32 (TF32) is a new format that uses the same 10-bit mantissa as half-precision (FP16) math and is shown to have more than sufficient margin for the precision requirements of AI workloads. In addition, since the TF32 adopts the same 8-bit exponent as FP32 it can support the same numeric range.

Ampere adds support for structured sparsity. Not all the parameters of modern AI networks are needed for accurate predictions and inference, and some can be converted to zeros to make the models "sparse" without compromising accuracy. The Tensor Cores in Ampere can provide up to 2× higher performance for inference of sparse models.

Ampere supports Compute Data Compression which can accelerate unstructured sparsity and other compressible data patterns. Compression in L2 provides up to a 4× improvement in DRAM read/write bandwidth, up to 4× improvement in L2 read bandwidth, and up to a 2× improvement in L2 capacity.

Ampere also supports many other enhancements for higher compute throughput.

4.1.2   Graphics Features

Ampere graphics capabilities include:
- End-to-end lossless compression, including Post-L2 compression, enabling compression of SM stores.
- Tiled Caching
- OpenGL 4.6+, Vulkan 1.2+, CUDA 10.2+
- Adaptive Scalable Texture Compression (ASTC) LDR profile supported
- Modern Graphics features:
  - Ray Tracing
  - DL Inferencing
  - Mesh Shaders
  - Sampler Feedback
  - Variable Rate Shading
  - Texture LOD in compute programs
- Iterated blend, ROP OpenGL-ES blend modes
- 2D BLIT from 3D class avoids channel switch
- 2D color compression
- Constant color render SM bypass
- 2×, 4×, 8× MSAA with color and Z compression
- Non-power-of-2 and 3D textures, FP16 texture filtering

▶ FP16 shader support

▶ Geometry and Vertex attribute Instancing

▶ Parallel pixel processing

▶ Early-z reject: Fast rejection of occluded pixels acts as multiplier on pixel shader and texture performance while saving power and bandwidth

▶ Video protection region

4.1.3   GPU Architecture

There are multiple texture processing clusters (TPC) units within a graphics processing cluster (GPC), each TPC includes two SMs, a Polymorph Engine, two Texture Units, and a Ray Tracing core (RTcore). Each GPC includes a Raster Engine (ROP), which can access all of memory. Each SM is partitioned into four separate processing blocks, each with its own instruction buffer, scheduler and 128 CUDA cores.

The GPC is a dedicated hardware block for rasterization, shading, texturing, and compute. The GPU's core graphics functions are performed inside the GPC. Inside the GPC, the SM CUDA cores perform pixel/vertex/geometry shading and physics/compute calculations. Texture units perform texture filtering and load/store units fetch and save data to memory. Special Function Units (SFUs) handle transcendental and graphics interpolation instructions. Tensor cores perform matrix multiplies to greatly accelerate DL inferencing. The RTcore unit assists ray-tracing by accelerating Bounding Volume Hierarchy (BVH) traversal and intersection of scene geometry during ray tracing.

Finally, the PolyMorph engine handles vertex fetch, tessellation, viewport transform, attribute setup, and stream output. The SM geometry and pixel processing performance make it highly suitable for rendering advanced user interfaces and complex gaming applications. The power efficiency of the Ampere GPU enables this performance on devices with power-limited environments.

4.2   CPU Complex

The CPU cluster is comprised of 12-cores (JAO 64GB) or 8-cores (JAO 32GB) of Arm Cortex-A78AE Core processors organized as multiple quad-core clusters. Clusters contain private L1 and L2 caches per core, a Snoop Control Unit (SCU), and a cluster-level L3 cache (shared by the 4 cores), an interconnect fabric and debug support modules (CoreSight).

4.2.1   CPU

Features:

▶ Superscalar, variable-length, out-of-order pipeline

▶ Dynamic branch prediction with Branch Target Buffer (BTB) and a branch direction predictor using previous branch history, a return stack, a static predictor and an indirect predictor

▶ A 1.5K entry, 4-way skewed associative L0 Macro-OP (MOP) cache

▶ 32-entry fully-associative L1 instruction TLB with native support for 4KB, 16KB, 64KB, and 2MB page sizes.

▶ 32-entry fully-associative L1 data TLB with native support for 4KB, 16KB, 64KB, 2MB, and 512MB page sizes.

▶ 4-way set-associative unified 1024-entry Level 2 (L2) TLB in each processor

▶ L1 caches – separate 64 KB I-cache and 64 KB D-cache for each core.

▶ L2 cache – a unified, 8-way set associative, 256 KB L2 cache per core.

▶ 40-bit Physical Address (PA)

The Cortex-A78AE CPU supports:

▶ Full implementation of Armv8.2-A architecture instruction set and select instructions from Armv8.3-A, Armv8.4-A and Armv8.5-A extensions

▶ Embedded Trace Microcell (ETM) based on the ETMv4.2 architecture

▶ Performance Monitor Unit (PMU) based on the PMUv3 architecture

▶ CoreSight for debugging based on CoreSightv3 architecture

▶ Cross Trigger Interface (CTI) for multiprocessor debugging

▶ Generic Timer Interface based on Armv8-A architecture and 64-bit count input from external system counter

▶ Cryptographic Engine for crypto function support

▶ Interface to an external Generic Interrupt Controller based on GICv3 architecture

▶ Power management with multiple power domains

4.2.2   Supporting Features

The CPU clusters contain supporting features including:

▶ Debug, power-management

▶ Arm CoreLink GIC-600AE Generic Interrupt Controller

▶ Error detection and reporting

4.2.3   Performance Monitoring

A performance monitoring unit in each core (provided as part of the Arm Cortex-A78 core) provides six counters, each of which can count any of the events in the processor. The unit

gathers various statistics on the operation of the processor and memory system during runtime, based on Arm PMUv3 architecture. In addition, the DSU provides six counters to gather various statistics on the operation of the memory of the cluster during runtime.

4.3 Programmable Vision Accelerator and Deep Learning Accelerator Cluster

This cluster consists of two primary engines: Programmable Vision Accelerator (PVA) and Deep Learning Accelerator (DLA).

The Orin PVA is the second generation of NVIDIA's vision DSP architecture, which is an application-specific instruction vector processor that targets computer-vision along with virtual and mixed reality applications. These are some key areas where PVA capabilities are a good match for algorithmic domains that need to have a predictable processing capability, at low power and low latency.

A PVA cluster has the following components:

▶ Dual Vector Processing Units (VPU) with vector cores, instruction cache, and 3 vector data memories. Each unit has seven VLIW slots including both scalar and vector instructions.

▶ 384 KBytes of triple-port memory for each VPU

▶ Dual DMA engines with 5-dimensional addressing capability, each with 16 independent hardware channels, and sophisticated control to have both hardware and software events trigger the DMA channels.

▶ 1 MByte local L2 cache.

▶ Cortex-R5 subsystem for PVA control and task monitoring.

The DLA is a fixed function engine used to accelerate inference operations on convolutional neural networks (CNNs). Orin implements the second generation of NVIDIA's DLA architecture. The DLA supports accelerating CNN layers such as convolution, deconvolution, activation, pooling, local response normalization, and fully-connected layers.

Specific optimizations include:

▶ Structured Sparsity

▶ Depth-wise Convolution capability

▶ A dedicated Hardware Scheduler to maximize efficiency

4.4   Multi-Standard Video Decoder

The SOM incorporates a single instance of the NVIDIA Multi-Standard Video Decoder (NVDEC). This video decoder accelerates video decode, supporting low resolution mobile content, Standard Definition (SD), High Definition (HD) and UltraHD (8K, 4K, etc.) video profiles. The video decoder is designed to be extremely power efficient without sacrificing performance. The video decoder communicates with the memory controller through the video DMA which supports a variety of memory format output options. For low-power operations, the video decoder can operate at the lowest possible frequency while maintaining real-time decoding using dynamic frequency scaling techniques.

Video decode standards supported: H.265 (HEVC), H.264, VP9, VP8, AV1, MPEG-4, MPEG-2, VC-1.

Table 4-1.   Supported Video Decode Streams JAO 64GB

Standard	Profile(s)	Resolution (Maximum Number of Streams)	Throughput (Up to)
H.264	Baseline, Main, High	4K60 (1) | 4K30 (3) | 1080p60 (6) | 1080p30 (13)	850 MP/S
	High 444, High 444 Predictive, MVC (per view considering 2 views)[1]	4K30 (1) | 1080p60 (3) | 1080p30 (6)	425 MP/s
H.265 (HEVC)	Main, Main10	8K30 (1) | 4K60 (3) | 4K30 (7) | 1080p60 (11) | 1080p30 (22)	1300 MP/S
	Main444, Main444 10, MV (per view)	4K60 (1) | 4K30 (3) | 1080p60 (5) | 1080p30 (11)	650 MP/S
AV1	Main Profile	8K30 (1) | 4K60 (3) | 4K30 (6) | 1080p60 (12) | 1080p30 (24)	1200 MP/S
VP9	Profile 0, Profile 2	8K30 (1) | 4K60 (3) | 4K30 (6) | 1080p60 (9) | 1080p30 (18)	1200 MP/S

Notes:

[1]Max throughput half for YUV444 – as compared to YUV420

[2]AV1 Dec: Max aggregate Mbps should be <120

Table 4-2.   Supported Video Decode Streams JAO 32GB

Standard	Profile(s)	Resolution (Maximum Number of Streams)	Throughput (Up to)
H.264	Baseline, Main, High	4K60 (1) | 4K30 (2) | 1080p60 (5) | 1080p30 (11)	720 MP/S
	High 444, High 444 Predictive, MVC (per view considering 2 views)[1]	4K30 (1) | 1080p60 (2) | 1080p30 (5)	360 MP/s
H.265 (HEVC)	Main, Main10	8K30 (1) | 4K60 (2) | 4K30 (4) | 1080p60 (9) | 1080p30 (18)	1100 MP/S
	Main444, Main444 10, MV (per view)	4K60 (1) | 4K30 (2) | 1080p60 (4) | 1080p30 (9)	550 MP/S
AV1	Main Profile	8K30 (1) | 4K60 (2) | 4K30 (4) | 1080p60 (10) | 1080p30 (20)	1000 MP/S
VP9	Profile 0, Profile 2	4K60 (1) | 4K30 (3) | 1080p60 (7) | 1080p30 (15)	1000 MP/S

Notes:

[1]Max throughput half for YUV444 – as compared to YUV420

[2]AV1 Dec: Max aggregate Mbps should be <120

4.5   Multi-Standard Video Encoder

The SOM incorporates a single instance of the NVIDIA Multi-Standard Video Encoder (NVENC). This multi-standard video encoder enables full hardware acceleration of various encoding standards. It performs high quality video encoding operations for mobile applications such as video recording and video conferencing. The encode processor is designed to be extremely power efficient without sacrificing performance.

Video encode standards supported: H.265 (HEVC), H.264, AV1.

Table 4-3.   Supported Video Encode Streams JAO 64GB

Standard	Profile(s)	Resolution (Maximum Number of Streams)	Throughput (Up to)
H.264	UHP	4K60 (1) | 4K30 (3) | 1080p60 (7) | 1080p30 (14)	870 MP/s
	HP	4K60 (1) | 4K30 (2) | 1080p60 (4) | 1080p30 (9)	600 MP/S
	HQ	4K30 (1) | 1080p60 (2) | (4) 1080p30	280 MP/s
H.265 (HEVC)	UHP	4K60 (2) | 4K30 (4) | 1080p60 (8) | 1080p30 (16)	1000 MP/S
	HP	4K30 (1) | 1080p60 (3) | 1080p30 (7)	500 MP/S
	HQ	1080p60 (1) | 1080p30 (3)	200 MP/s
AV1	UHP	4K60 (2) | 4K30 (4) | 1080p60 (7) | 1080p30 (15)	1000 MP/S
	HQ	4K60 (1) | 4K30 (2) | 1080p60 (3) | 1080p30 (7)	480 MP/s

Table 4-4.   Supported Video Encode Streams JAO 32GB

Standard	Profile(s)	Resolution (Maximum Number of Streams)	Throughput (Up to)
H.264	UHP	4K60 (1) | 4K30 (2) | 1080p60 (5) | 1080p30 (11)	680 MP/s
	HP	4K30 (1) | 1080p60 (3) | 1080p30 (7)	470 MP/S
	HQ	1080p60 (1) | 1080p30 (3)	220 MP/s
H.265 (HEVC)	UHP	4K60 (1) | 4K30 (3) | 1080p60 (6) | 1080p30 (12)	800 MP/S
	HP	4K30 (1) | 1080p60 (3) | 1080p30 (6)	400 MP/S
	HQ	1080p60 (1) | 1080p30 (2)	140 MP/s
AV1	UHP	4K60 (1) | 4K30 (3) | 1080p60 (6) | 1080p30 (12)	750 MP/S
	HQ	4K30 (1) | 1080p60 (3) | 1080p30 (6)	380 MP/s

4.6   Optical Flow Accelerator

The Optical Flow Accelerator (OFA) is a hardware accelerator for computing optical flow and stereo disparity between the frames.

OFA can operate in Stereo Disparity Mode and Optical Flow Mode.

OFA generates disparity and flow vector block-wise, one output for each input block of 8x8, 4x4, 2x2, and 1x1 pixels (referred as output grid size). The generated output can be further post-processed to improve accuracy, up sampled to produce dense map.

▶ **Stereo Disparity Mode**

- OFA processes rectified left and right view of stereo captures and generates disparity values between them.

- The output stereo disparity format is fixed signed 10.5 (2 bytes per disparity output). We need to divide the output values by 32 to get a disparity value in terms of pixel units.

▶ **Optical Flow Mode**

- OFA generates optical flow between two given frames.

- The input to OFA in this mode is image pyramid of input and reference frames with fixed scale factor of 2. As search range of single layer is small, each pyramid level will search around output of previous pyramid level.

- OFA generates a flow vector has X and Y component that represent motion in X and Y direction. The output flow format is fixed signed 10.5 (4 bytes per flow vector). We need to divide the output values by 32 to get a disparity value in terms of pixel units.

Table 4-5.   Optical Flow Accelerator

OFA Parameter	Description
Input Image Size	Minimum size: 32 × 32
	Maximum size: 8192 × 8192
	No alignment requirement
Input Image format / bit depth	Luma / Single channel Input
	Supports bit depth of 8/10/12/16 bits
Disparity Output	Disparity Map in fixed S 10.5 format
Flow Output	Flow Map (mvx, mvy) in fixed S10.5 format
Hardware Cost Output	Hardware cost for winner disparity candidate
	8 bit per output / Range 0 - 255
Output Grid Size	1x1/2x2/4x4/8x8
Maximum Disparity Range	128 / 256
Search Direction	Left / Right Disparity Map
Region Of Interest Support	Supports maximum 32 ROI per stereo pair
Max Pyramid Levels for Flow	5

Table 4-6.   OFA Streams

Mode	Grid Size	Resolution (Maximum Number of Streams)	Throughput (Up to)
Optical Flow	8x8	4K60 (1) | 4K30 (3) | 1080p60 (7) | 1080p30 (15)	940 MP/S
	4x4	4K30 (1) | 1080p60 (2) | 1080p30 (5)	330 MP/S
Stereo	8x8	4K60 (3) | 4K30 (6) | 1080p60 (13) | 1080p30 (27)	1700 MP/S
	4x4	4K30 (1) | 1080p60 (3) | 1080p30 (7)	460 MP/S

4.7   NVJPEG

The JPEG processing block is responsible for JPEG (de)compression calculations (based on JPEG still image standard), image scaling, decoding (YUV420, YUV422H/V, YUV444, YUV400) and color space conversion (RGB to YUV).

It consists of hardware engine with two instances of NVJPEG HW:

▶   2x NVJPEG
▶   Perf: 2x 600Mpix/Sec

Table 4-7.   NVJPEG Streams per Instance

NVJPEG	Compression Ratio	Throughput (Up to)	Number of 1080p30 Streams	Number of 4K30 Streams
Decode	6:1	756 MP/S	12	3x
	10:1	952 MP/S	15	3x
Encode	6:1	922 MP/S	15	3x
	10:1	1253 MP/S	20	5x

Notes:

2x NVJPG engines are present in Orin. The data in this table is for single instance of NVJPG

Results at 880 MHz for 4:2:0 and aggregate across two NVJPEG blocks

Throughput for 4:4:4 will be roughly half of the above

Input (encode) formats:

- Pixel width: 8 bpc
- Subsample format: YUV420
- Resolution (up to): 16K x 16K
- Pixel pack format
  > Semi-planar/Planar for 420

Output (decode) formats:

- Pixel width 8 bpc
- Resolution (up to): 16K x 16K
- Pixel pack format
  > Semi-planar/Planar for YUV420
  > YUY2/Planar for 422H/422V
  > Planar for YUV444/YUV400
  > Interleaved RGBA

4.8   Sensor Processing Engine

The Cortex-R5 processor in the Always On (AON) block is also referred to as the Sensor Processing Engine (SPE). The AON cluster provides all the necessary hardware features to support low power sensor management and wake use cases. The cluster consists of an Arm Cortex-R5 processor core with a tightly coupled RAM, supporting peripherals (such as timers and an interrupt controller), various I/O controller peripherals, and routing logic.

AON Cortex-R5 implementation:

- ▶ Arm v7-R ISA
- ▶ Integrated instruction and data caches
- ▶ Tightly coupled memory (TCM) interface for local SRAM
- ▶ Vectored interrupt support
- ▶ 64-bit AXI Initiator interface for DRAM requests

▶ 32-bit AXI Initiator interface for MMIO requests

▶ 32-bit AHB Initiator interface for Arm Vectored Interrupt Controller (AVIC) access

▶ AXI Target interface for DMA access to the local SRAM

4.9 Security Subsystem

This subsystem is comprised of the following:

▶ Platform Security Controller (PSC)

▶ Security Engine (SE)

4.9.1 Platform Security Controller

The Platform Security Controller (PSC) is a highly secure subsystem to protect and manage assets (keys, fuses, functions, and features) within the SoC, provide trusted services, increase resilience against attacks on the SoC, and provide a greater level of protection against software and hardware attacks on the subsystem itself.

Key Management and Protection: The PSC will be the only mechanism with access to the most critical secrets in the chip. This subsystem represents the highest level of protection in Orin and the subsystem itself is highly resilient to a wide range of software and hardware attacks.

Trusted Services: The primary PSC services include secure authentication (for example, during SoC secure boot), provisioning of additional keys, ID, data, key access and management, random number generation, and trusted time reporting.

Security Monitor: The PSC will be responsible for periodic security housekeeping tasks, including continually assessing the security status of the SoC, actively monitor known or potential attack patterns (for example, such as voltage glitching or thermal attacks), mitigate hardware attack risks, and to take action in the case of a detected attack. The PSC will have the ability to accept updates as workarounds to improve the robustness of the system in the field.

4.9.2 Security Engine

The Security Engine (SE) provides hardware acceleration for cryptographic algorithms. There are two instances of SE available for software usage:

▶ TZ-SE: accessible only by TrustZone software

▶ NS/TZ-SE: configurable to be accessible only by TrustZone software or TrustZone and non-secure software

The SE provides hardware acceleration for various cryptographic operations and hardware-assisted Key protection. The crypto operations that the SE provides can be used by software to build crypto protocols and security features. All of these crypto operations are based on Crypto algorithms approved by the National Institute of Standards and Technology (NIST).

The SE supports the following:

▶ NIST-compliant asymmetric, symmetric cryptography and hashing

▶ Side channel countermeasures [AES/RSA/ECC]

▶ Independent channels for parallelization

▶ Hardware Key Access Controls (KAC): Rule-based, hardware-enforced access control for symmetric keys

▶ 16× AES, 4× RSA/ECC key slots

▶ Hardware key isolation (only AES keyslots)

▶ Read protection (only AES keyslots)

▶ Hardware keyslot functions

▶ Key wrap and unwrap functionality (AES -> AES keyslot)

▶ Key derivation into a keyslot (KDF -> AES keyslot)

▶ Random key generation (RNG -> AES keyslot)

4.10 Jetson AGX Orin SOM Memory

64 GB 256-bit LPDDR5 DRAM is used on the NVIDIA JAO 64GB. 32 GB 256-bit LPDDR5 DRAM is used on the NVIDIA JAO 32GB. JAO supports the following:

▶ Secure external memory access using TrustZone technology

▶ System MMU

▶ Maximum operating frequency: 3200 MHz

Other non-volatile memory used on the module is

▶ 64 GB eMMC 5.1

▶ 64 MB NOR Boot Flash (QSPI)

▶ 8 MB NOR Secure Key Flash

Chapter 5. Interfaces

5.1 USB Interfaces

The NVIDIA Jetson AGX Orin SOM provides 4x USB 2.0 and 3x USB3.2 Gen2 x1 port for communication to external peripheral devices. In host mode, the USB3.2 host controller supports up to Gen2 Super Speed+, 10 Gbps. In device mode, the USB3.2 controller supports up to Gen1 Super Speed. USB interfaces are compliant with the following USB specifications:

▶ Universal Serial Bus Specification Revision 3.2 Gen 1 and 2
▶ Universal Serial Bus Specification Revision 2.0, plus the following:
  • Modes: Host and Device (Only USB 2.0 port USB0 supports RCM, Host, Device Mode. All other ports are Host only)
  • Speeds: Low, Full, and High
  • *USB Battery Charging 1.2 Specification*
▶ *Enhanced Host Controller Interface Specification for Universal Serial Bus* Revision 1.0

5.2 SD and eMMC Controller

The Secure Digital (SD) and Embedded Multimedia Card (eMMC) controller is capable of interfacing to SD, SDIO, eSD, and eMMC cards. It has a direct memory interface and is capable of initiating data transfers between system memory and an external card or device. The SD and eMMC controller supports two different bus protocols: SD and eMMC bus protocol for eMMC cards.

Features of the controller are:

▶ Supports 4-bit data interface for SD cards
▶ Allows card to interrupt host in 1 bit, and 4-bit modes
▶ Supports Read wait Control, Suspend/Resume operation for SDIO cards
▶ Supports FIFO overrun and underrun condition by stopping SD clock
▶ Supports addressing larger capacity SD 3.0 or SD-XC cards up to 2 TB

5.3   Serial Peripheral Interface

There are 3x general-purpose serial peripheral interface (SPI) buses available on the NVIDIA Jetson AGX Orin SOM. The SPI controller allows a duplex, synchronous, serial communication between the controller and external peripheral devices. It consists of four signals:

▶ CS_N (Chip select)

▶ SCK (clock)

▶ MOSI (Initiator data out and Target data in)

▶ MISO (Initiator data in and Target data out)

The data is transferred on MISO or MOSI based on the data transfer direction on every SCK edge. The receiver always receives the data on the other edge of SCK.

Features of the SPI controller are:

▶ Initiator and target functionality

- Initiator: support all modes in the "SPI Mode Descriptions" table (Table 5-1)

- Target: support Mode 1 and Mode 3 in the "SPI Mode Descriptions" table (Table 5-1

▶ Independent RX FIFO and TX FIFO.

▶ Software-controlled bit-length supports packet sizes of 4-bits to 32-bits.

▶ Packed mode support for bit-length of three (4-bit packet size), seven (8-bit packet size) 15 (16-bit packet size), and 31 (32-bit packet size)

▶ CS_N can be selected to be controlled by software, or it can be generated automatically by the hardware on packet boundaries

▶ Simultaneous receive and transmit supported

▶ SPI1 and SPI3 support two chip-selects.

▶ SPI6 supports:

- Initiator Mode 0 only

- SDR mode only

- SPI ×1 and dual SPI ×2 modes (both half-duplex)

Table 5-1.   SPI Mode Descriptions

SPI Mode	Clock Polarity	Clock Phase	SCK Inactive State	Data Latch In	Data Latch Out
0	0	0	Low	Latched IN on the positive edge of clock	Latched OUT on the negative edge of clock
1	0	1	Low	Latched IN on the negative edge of clock	Latched OUT on the positive edge of clock
2	1	0	High	Latched IN on the negative edge of clock	Latched OUT on the positive edge of clock
3	1	1	High	Latched IN on the positive edge of clock	Latched OUT on the negative edge of clock

5.4   I2C Controller

8x general-purpose I2C controller allows system expansion for I2C-based devices, such as cameras, sensors, voltage monitor, thermal monitor, serial ADC/DAC, and serial EPROMs, as defined in the NXP inter-IC-bus (I2C) specification. I2C6 can be used either for DP AUX or I2C.

The I2C bus supports serial device communications to multiple devices. The I2C controller handles clock source negotiation, speed negotiation for standard and fast devices, and 7-bit target address support according to the I2C protocol and supports Initiator and Target mode of operation.

The I2C controller supports the following operating modes for both Initiator and Target: Standard-mode (Sm, up to 100 Kbit/s), Fast-mode (Fm, up to 371.585 Kbit/s), Fast-mode plus (Fm+, up to 985 Kbit/s). The I2C controller also supports Multi-Master operation.

5.5   UART

The NVIDIA Jetson AGX Orin SOM has 4x general-purpose UART ports. UART controller provides serial data synchronization and data conversion (parallel-to-serial and serial-to-parallel) for both receiver and transmitter sections. Synchronization for serial data stream is accomplished by adding start and stop bits to the transmit data to form a data character. Data integrity is accomplished by attaching a parity bit to the data character. The parity bit can be checked by the receiver for any transmission bit errors.

Features of UART are:

▶   Synchronization for the serial data stream with start and stop bits to transmit data and form a data character

▶   Supports both 16450- and 16550-compatible modes. Default mode is 16450

▶   Device clock up to 200 MHz, baud rate of 12.5 Mbits/second

▶   Data integrity by attaching parity bit to the data character

- ▶ Support for word lengths from five to eight bits, an optional parity bit and one or two stop bits
- ▶ Support for modem control inputs
- ▶ DMA capability for both TX and RX
- ▶ 8-bit × 36 deep TX FIFO
- ▶ 11-bit × 36 deep RX FIFO. 3 bits of 11 bits per entry will log the RX errors in FIFO mode (break, framing, and parity errors as bits 10,9,8 of FIFO entry)
- ▶ Auto sense baud detection
- ▶ Time out interrupts to indicate if the incoming stream stopped
- ▶ Priority interrupts mechanism
- ▶ Flow control support on RTS and CTS (hardware and software controlled)
- ▶ Internal loop-back
- ▶ SIR encoding/decoding (3/16 or 4/16 baud pulse widths to transmit bit zero)

5.6   RGMII

The NVIDIA Jetson AGX Orin SOM integrates an Ethernet controller/MAC with AVB support and provides a Reduced Gigabit Media Independent Interface (RGMII) to an external Ethernet PHY or switch. The transmit clock signal is provided by the MAC and is synchronous with the data signals. The timing of NVIDIA Orin SoC complies with the original RGMII mode of *Reduced Gigabit Media Independent Interface (RGMII) Specification*, Version 2.0

5.7   MGBE

The NVIDIA Jetson AGX Orin SOM has four integrated Multi-Gigabit Ethernet (MGBE) controllers that can support up to 40 Gbps of total bandwidth.

Each MGBE controller can independently operate in 2.5 Gbps, 5 Gbps, or 10 Gbps throughput mode, enabling NVIDIA Orin SoC to transmit and receive data over Ethernet in compliance with IEEE 802.3-2015 standard.

The NVIDIA Orin SoC MGBE controller can be connected to external devices like Ethernet PHY's and Switches through XFI differential lanes.

5.8   CAN

The Controller Area Network (CAN) is a vehicular bus standard for communication between microcontrollers and devices within the vehicle. The CAN bus is a multi-Initiator serial bus for connecting multiple nodes within a vehicle using a message-based protocol. The NVIDIA Jetson AGX Orin SOM supports connectivity to two CAN networks.

Features of CAN are:

▶ CAN protocol Version 2.0A, Version 2.0B, and ISO 11898-1:2006/11898-1:2015

▶ Support ISO11898-1:2006 FD format and BOSCH FD format

▶ Dual clock source, enabling FM-PLL designs

▶ 16, 32, 64 or 128 Message Objects (configurable)

▶ Each Message Object has its own Identifier mask

▶ Programmable FIFO mode

▶ Programmable loop-back mode for self-test

▶ Parity check for message RAM (optional)

▶ Maskable interrupt, two interrupt lines

▶ Power-down support

▶ Supports TT CAN

▶ TTCAN Level 0, 1, and 2

▶ Time Mark Interrupts

▶ Stopwatch

▶ Watchdog timer

▶ Synchronization to external events

5.9   Display Interfaces

The NVIDIA Jetson AGX Orin SOM provides 1x HDMI and DP port. The HDMI™ and VESA DisplayPort (DP) interfaces share the same set of interface pins.

HDMI provides a unified method of transferring both audio and video data. The HDMI block receives video from either display controller and audio from a separate high-definition audio (HDA) controller; it combines and transmits them as appropriate.

Supported HDMI features are:

▶ Compliant to the HDMI 2.0 (up to 594 MHz pixel clock rate) and HDMI 2.1 (Fixed Rate Link at 3 Gbps, 6 Gbps, 8 Gbps, 10 Gbps, or 12 Gbps)

- Support 8/10/12 bpc RGB, YUV444, YUV420, or YUV422 (HDMI 2.0 only)

▶ HDCP 2.2 and 1.4

▶ On-chip HDCP key storage, no external SecureROM required

▶ Multichannel audio from HDA controller, up to eight channels 192 KHz 24-bit

▶ 24-bit RGB and 24-bit YUV444 (HDMI) pixel formats

VESA DisplayPort (DP) is a digital display interface often used to connect a video source to a display device over a cable, in consumer or commercial applications. Embedded DisplayPort (eDP) is based on DP but intended for embedded applications where the display panel is integrated. For embedded use cases that require multiple display support using MST, and DP is intended to interface with SerDes devices that in turn could support multiple displays. Using SerDes can provide long-distance, low-EMI connection for multiple displays. DP or eDP is a mixed-signal interface consisting of four differential serial I/O lanes.

Supported DisplayPort features are:

▶ Compliant to the DisplayPort 1.4a Specification

- Support 16 bpp YUV422

- Support 18 bpp RGB

- Support 24 bpp RGB/YUV444

- Support 30 bpp RGB/YUV444

- Support 36 bpp RGB/YUV444

▶ Support up to 810 MHz pixel clock

▶ Support for 1/2/4 lanes

▶ Support for following bit rates:

- RBR (Reduced Bit Rate, 1.62 Gbps)

- HBR (High Bit Rate, 2.7 Gbps)

- HBR2 (High Bit Rate 2, 5.4 Gbps)

- HBR3 (High Bit Rate 3, 8.1 Gbps)

▶ Multi-Stream Transport (MST)

▶ Support for 2-8 channel audio streaming up to 96 KHz sample rate

▶ Support additional eDP 1.4 features:

- Additional link rates (2.16, 2.43, 3.24, 4.32 Gbps)

- Enhanced framing

- Power sequencing

- Reduced AUX timing

- Reduced main voltage swing

- Alternate Seed Scrambler Reset (ASSR) for internal eDP panels

5.10 Audio Interfaces

The Audio Controller transports streaming audio data between system memory and an audio codec. The controller supports I2S format, Left-justified Mode format, Right-justified Mode format, and DSP mode format, as defined in the Philips inter-IC-sound (I2S) bus specification. The timing in the following sections applies to any of these interfaces depending on whether they are configured for I2S or TDM mode.

The I2S controller supports point-to-point (P2P) serial interfaces for the I2S digital audio streams. I2S-compatible products, such as compact disc players, digital audio tape devices, digital sound processors, and those with digital TV sound may be directly connected to the I2S controller. The controller also supports the PCM and telephony mode of data-transfer. Pulse-Code-Modulation (PCM) is a standard method used to digitize audio (particularly voice) patterns for transmission over digital communication channels. The Telephony mode is used to transmit and receive data to and from an external mono codec in a slot-based scheme of time-division multiplexing. The I2S controller supports bidirectional audio streams and can operate in half-duplex or full-duplex mode.

When DAP port operates as I2S (Initiator and Target modes) interface, it supports clock rates up to 12.288 MHz and comply with I2S specification.

When DAP port operates as TDM/PCM interface, it supports clock rates up to 24.576 MHz.

Features for audio interfaces are:

▶ Basic I2S modes to be supported (I2S, RJM, LJM, and DSP) in both Initiator and Target modes.

▶ PCM mode with short (one-bit-clock wide) and long-fsync (two bit-clocks wide) in both Initiator and Target modes.

▶ NW-mode with independent slot-selection for both Tx and Rx

▶ TDM mode with flexibility in number of slots and slots selection.

▶ Capability to drive-out a High-z outside the prescribed slot for transmission.

▶ Flow control for the external input and output stream.

5.11 Pulse-Width Frequency Modulation Interface

There are 4x pulse-width frequency modulation (PWM) frequency divider with a varying pulse width available on the NVIDIA Jetson AGX Orin SOM. FAN PWM and PWM1 are assigned at the connector. Other two PWMs are available for customer use but may not be forward compatible.

The PWM runs off a device clock programmed in the Clock and Reset controller. The source can either be the OSC clock (38.4 MHz) or PLLP (408 MHz). The source is first divided by 256,

and then again by a 13-bit register value, to generate the PWM frequency. The duty cycle is a controlled by an 8-bit register value.

5.12 General Purpose I/O

The NVIDIA Jetson AGX Orin SOM offers several General Purpose I/O pins. Some GPIOs are dedicated, and others are alternative GPIOs. Each GPIO pin is configurable for input or output direction and can be read or written to individually. All GPIOs support interrupt capability. For GPIO voltage level and characteristics, refer to "Pin Type" in the SOM pin list which provides their voltage rail and pad type. Some limitations apply to alternative GPIOs.

> **Notes**: GPIOs are push-pull only. Configuring as open drain or bidirectional I/O is performed through software emulation.
>
> GPIOs supporting 3.3V tolerance can become true open drain if 3.3V tolerance is enabled.

5.13 JTAG

The NVIDIA Jetson AGX Orin SOM has a JTAG interface that can be used for boundary scan testing or for debugging. JTAG clock can be driven up to 15 MHz. However, during boundary scan, its frequency should not exceed 7.5 MHz.

5.14 System Control signals

The NVIDIA Jetson AGX Orin SOM provides a set of system control signals. They are used for the following:

▶ Power handshaking
▶ Temperature sensing of attached devices
▶ System forced recovery
▶ Voltage monitor interrupts
▶ In-System-Testing
▶ WDT reset

5.15   UPHY Configurations

The NVIDIA Jetson AGX Orin SOM supports up to PCIe Gen4 speed. The PCIe lanes can be configured as x1, x2, x4, and x8. The PCIe controllers can be used as root ports or endpoint devices modes. Refer to the *NVIDIA Jetson AGX Design Guide* for supported UPHY configurations for the NVIDIA Jetson AGX Orin SOM.

5.16   CSI Configurations

The NVIDIA Jetson AGX Orin SOM has four CSI x4 bricks (16 lanes and trios total) supporting a variety of device types and camera configurations. Data aggregated from physical lanes enters an asynchronous FIFO which interfaces to the NVCSI block. Both MIPI D-PHY v2.1 and C-PHY v2.0 modes are supported. In D-PHY mode, each data channel has peak bandwidth of up to 2.5 Gbps per lane. For C-PHY, each lane (Trio) supports up to 4.5 GSymb/s (10.25 Gbps).

Features of the CSI interface are as follows:

▶ MIPI CSI-2 3.0 receiver

▶ Supports up to 6 CSI-2 input ports operating concurrently - for up to 6 deserializers (six cameras) or more cameras through deserializer-aggregators (hubs):

- Up to six ×1 lane deserializer output port
- Up to six ×2 lane deserializer output port
- Up to four ×4 lane deserializer output port

▶ Supports 16 virtual channels (VC) per CSI link

▶ Supported input data formats:

- RGB
  - > RGB888
  - > RGB666
  - > RGB565
  - > RGB555
  - > RGB444
- YUV
  - > YUV422-8b
  - > YUV420-8b (legacy)
  - > YUV420-8b
  - > YUV420-10b
  - > YUV422-10b
- RAW
  - > RAW6
  - > RAW7
  - > RAW8

- > RAW10
- > RAW12
- > RAW14
- > RAW16
- > RAW20
- DPCM: user defined
- DPCM (predictor 1)
  - > 14-10-14
  - > 14-8-14
  - > 12-10-12
  - > 12-8-12
  - > 12-7-12
  - > 12-6-12
  - > 10-8-10
  - > 10-7-10
  - > 10-6-10
- Embedded control information
- ▶ MIPI D-PHY v2.1 Modes of Operation
  - High Speed Mode: High speed differential signaling up to 2.5 Gbps. Burst transmission for low power
  - Low Power Control: Single-ended 1.2V CMOS level. Low speed signaling for handshaking
  - Low Power Escape: Low speed (10 Mbps) signaling for data. Used for escape command entry only
- ▶ MIPI C-PHY v2.0 Modes of Operation
  - Shares same D-PHY low power receiver functionality
  - High Speed Mode: Requires one or more sets of three wires (referred to as a trio) for high-speed data communication

The following tables show CSI configurations for the NVIDIA Jetson AGX Orin SOM.

5.16.1   D-PHY Configurations

The following table contains the CSI configurations in D-PHY mode.

Table 5-2.   CSI Configurations D-PHY Mode

CSI Input	Module Connector Name	D-PHY x1	D-PHY x2	D-PHY x4
A	CSI0_CLK_N	Clock	Clock	Clock
	CSI0_CLK_P			
	CSI0_D0_N	Data	Data	Data
	CSI0_D0_P			
	CSI0_D1_N			
	CSI0_D1_P			
B	CSI1_CLK_N	Clock	Clock	
	CSI1_CLK_P			
	CSI1_D0_N	Data	Data	Data
	CSI1_D0_P			
	CSI1_D1_N			
	CSI1_D1_P			
C	CSI2_CLK_N	Clock	Clock	Clock
	CSI2_CLK_P			
	CSI2_D0_N	Data	Data	Data
	CSI2_D0_P			
	CSI2_D1_N			
	CSI2_D1_P			
D	CSI3_CLK_N	Clock	Clock	
	CSI3_CLK_P			
	CSI3_D0_N	Data	Data	Data
	CSI3_D0_P			
	CSI3_D1_N			
	CSI3_D1_P			
E	CSI4_CLK_N	Clock*	Clock*	Clock
	CSI4_CLK_P			
	CSI4_D0_N	Data*	Data*	Data
	CSI4_D0_P			
	CSI4_D1_N			
	CSI4_D1_P			
F	CSI5_CLK_N	Clock*	Clock*	
	CSI5_CLK_P			

		D-PHY		
CSI Input	**Module Connector Name**	**x1**	**x2**	**x4**
	CSI5_D0_N	Data*	Data*	Data
	CSI5_D0_P			
	CSI5_D1_N			
	CSI5_D1_P			
G	CSI6_CLK_N	Clock*	Clock*	Clock
	CSI6_CLK_P			
	CSI6_D0_N	Data*	Data*	Data
	CSI6_D0_P			
	CSI6_D1_N			
	CSI6_D1_P			
H	CSI7_CLK_N	Clock*	Clock*	
	CSI7_CLK_P			
	CSI7_D0_N	Data*	Data*	Data
	CSI7_D0_P			
	CSI7_D1_N			
	CSI7_D1_P			

📰 **Note**: For the E/F and G/H inputs, only one ×1 or ×2 interface can be used (either E or F, and either G or H).

5.16.2 Supported C-PHY Configurations

The following table contains the supported CSI configurations in C-PHY mode.

Table 5-3. CSI Configurations C-PHY Mode

					C-PHY		
CSI Input	**Module Connector Name**	**Trio**	**Mapping Option 1**	**Mapping Option 2**	**x1**	**x2**	**x4**
A	CSI0_D0_P	A0	A	A	■	■	■
	CSI0_D0_N		B	B			
	CSI0_CLK_P		C	C			
	CSI0_CLK_N	A1	C	A			
	CSI0_D1_P		A	B			

					C-PHY		
CSI Input	Module Connector Name	Trio	Mapping Option 1	Mapping Option 2	x1	x2	x4
	CSI0_D1_N		B	C			
B	CSI1_D0_P	B0	A	A			
	CSI1_D0_N	B0	B	B			
	CSI1_CLK_P	B0	C	C			
	CSI1_CLK_N	B1	C	A			
	CSI1_D1_P	B1	A	B			
	CSI1_D1_N	B1	B	C			
C	CSI2_D0_P	C0	A	A			
	CSI2_D0_N	C0	B	B			
	CSI2_CLK_P	C0	C	C			
	CSI2_CLK_N	C1	C	A			
	CSI2_D1_P	C1	A	B			
	CSI2_D1_N	C1	B	C			
D	CSI3_D0_P	D0	A	A			
	CSI3_D0_N	D0	B	B			
	CSI3_CLK_P	D0	C	C			
	CSI3_CLK_N	D1	C	A			
	CSI3_D1_P	D1	A	B			
	CSI3_D1_N	D1	B	C			
E	CSI4_D0_P	E0	A	A	See note	See note	
	CSI4_D0_N	E0	B	B			
	CSI4_CLK_P	E0	C	C			
	CSI4_CLK_N	E1	C	A			
	CSI4_D1_P	E1	A	B			
	CSI4_D1_N	E1	B	C			
F	CSI5_D0_P	F0	A	A	See note	See note	
	CSI5_D0_N	F0	B	B			
	CSI5_CLK_P	F0	C	C			
	CSI5_CLK_N	F1	C	A			
	CSI5_D1_P	F1	A	B			
	CSI5_D1_N	F1	B	C			
G	CSI6_D0_P	G0	A	A	See note		See note
	CSI6_D0_N	G0	B	B			
	CSI6_CLK_P	G0	C	C			
	CSI6_CLK_N	G1	C	A			
	CSI6_D1_P	G1	A	B			
	CSI6_D1_N	G1	B	C			

					C-PHY																	
CSI Input	Module Connector Name	Trio	Mapping Option 1	Mapping Option 2	x1						x2						x4					
H	CSI7_D0_P	H0	A	A					See note							See note						
	CSI7_D0_N		B	B																		
	CSI7_CLK_P		C	C																		
	CSI7_CLK_N	H1	C	A																		
	CSI7_D1_P		A	B																		
	CSI7_D1_N		B	C																		

Note: For the E/F and G/H inputs, only one ×1 or ×2 interface can be used (either E or F, and either G or H).

Chapter 6. Pin Definitions

The functions for each pin on the module are fixed to a single Special-Function I/O (SFIO) or software-controlled General Purpose I/O (GPIO). The NVIDIA Jetson AGX Orin SOM has multiple dedicated GPIOs. Each GPIO is individually configurable as Output/Input/Interrupt sources with level and edge controls. SFIO and GPIO functionality is configured using Multi-purpose I/O (MPIO) pads. Each MPIO pad consists of:

▶ An output driver with tristate capability, drive strength controls and push-pull mode, open-drain mode, or both

▶ An input receiver with either Schmitt mode, CMOS mode, or both

▶ A weak pull-up and a weak pull-down

MPIO pads are partitioned into multiple "pad control groups" with controls being configured for the group. During normal operation, these per-pad controls are driven by the pinmux controller registers.

Refer to the *NVIDIA Jetson AGX Orin Design Guide* for more information on pad behavior associated with different interfaces and the *Orin SoC Technical Reference Manual* for more information on modifying MPIO pad controls.

6.1   Power-On Reset Behavior

Each MPIO pad has a deterministic power-on reset (PoR) state. The reset state for each pad is chosen to minimize the need of additional on-board components; for example, on-chip weak pull-ups are enabled during PoR for pads which are usually used to drive active-low chip selects eliminating the need for additional pull-up resistors.

The following list is a simplified description of the NVIDIA Jetson AGX Orin SOM boot process focusing on those aspects which relate to the MPIO pins:

▶ System-level hardware executes the power-up sequence. This sequence ends when system-level hardware releases **SYS_RESET_N**.

▶ The boot ROM begins executing and programs the on-chip I/O controllers to access the secondary boot device.

▶ The boot ROM fetches the Boot Configuration Table (BCT) and boot loader from the secondary boot device.

▶ If the BCT and boot loader are fetched successfully, the boot ROM transfers control to the boot loader.

▶ Otherwise, the boot ROM enters USB recovery mode.

6.2   SOM B2B Connector Pinout

Simplified version of the NVIDIA Jetson AGX Orin SOM 699-pin B2B connector pinout is attached to this data sheet. For more details refer to the full pin description spreadsheet attached to the *Jetson AGX Orin Design Guide* (DG-10653-001).

To access the attached files, click the **Attachment** icon on the left-hand toolbar on this PDF (using Adobe Acrobat Reader or Adobe Acrobat). Select the file and use the Tool Bar options (**Open, Save**) to retrieve the documents. Excel files with the .nvxlsx extension will need to be renamed to .xlsx to open.

Chapter 7. Electrical and Mechanical Characteristics

7.1 Electrical Specifications

This section details the electrical specifications for the Jetson AGX Orin SOM.

7.1.1 Absolute Maximum Ratings

The absolute maximum ratings describe stress conditions. These parameters do not set minimum and maximum operating conditions that will be tolerated over extended periods of time. If the device is exposed to these parameters for extended periods of time, no guarantee is made, and device reliability may be affected. It is not recommended to operate a Jetson AGX Orin SOM under these conditions. Recommended operating conditions are provided in the following section.

Table 7-1.   Maximum Ratings

Symbol	Parameter	Min	Max	Unit	Notes
VIN	SYS_VIN_HV	-0.5	22.5	V	
	SYS_VIN_MV	-0.5	6.0	V	
	PMIC_BBATT	-0.3	5.4	V	
IDDMAX	VIN Imax (SYS_VIN_HV)		5.4	A	Software limited. IDDMAX (HV/MV current) reflects EDPp based on a 6 uS moving window.
	VIN Imax (SYS_VIN_MV)		6.0	A	5.4A is for VIN (20V) on SYS_VIN_HV. 6.0A is for VIN (5V) on SYS_VIN_MV. Actual IDDMAX is dependent on VIN (VINMIN)
VM_PIN	Voltage applied to any powered I/O pin	-0.5	VDD + 0.5	V	
	DD pads configured as open drain	-0.5	3.63		Pad's output driver must be set to open drain mode

Symbol	Parameter	Min	Max	Unit	Notes
Tttp	Operating Temperature: measured on Thermal Transfer Plate	-25	80	°C	
Tstg	Storage Temperature	-25	80	°C	

7.1.2 Recommended Operating Conditions

The parameters listed in following table are specific to a temperature range and operating voltage. Operating a NVIDIA Jetson AGX Orin SOM beyond these parameters is not recommended. Exceeding these conditions for extended periods may adversely affect device reliability.

Table 7-2.   Recommended Operating Conditions

Symbol	Parameter	Min	Typical	Max	Unit	Notes
VDDdc	SYS_VIN_HV	7	12	20	V	
	SYS_VIN_MV	4.75	5	5.25	V	
	PMIC_BBATT	1.65	3.3	3.45	V	Input Only

7.1.3 Digital Logic

Voltages less than the minimum stated value can be interpreted as an undefined state or logic level low which may result in unreliable operation. Voltages exceeding the maximum value can damage or adversely affect device reliability.

Table 7-3.   CMOS Pin Type DC Characteristics

Symbol	Description	Min	Max	Unit
VOL	Output Low Voltage (IOL = 1mA)	---	0.15 x VDD	V
VOH	Output High Voltage (IOH = -1mA)	0.75 x VDD	---	V
VIL	Input Low Voltage	-0.5	0.25 x VDD	V
VIH	Input High Voltage	0.70 x VDD	0.5 + VDD	V

Table 7-4.   Open Drain Pin Type DC Characteristics

Symbol	Description	Min	Max	Unit
VIL	Input Low Voltage	-0.5	0.25 x VDD	V
VIH	Input High Voltage	0.75 x VDD	3.63	V
VOL	Output Low Voltage (IOL = 1mA)	---	0.15 x VDD	V
	I2C Output Low Voltage (IOL = 3.3mA)	---	0.3 x VDD	V
VOH	Output High Voltage (IOH = -1mA)	0.85 x VDD	---	V

7.2   Mechanical Specifications

This section details the mechanical specifications for the Jetson AGX Orin SOM.

7.2.1   SOM Mechanical Drawing and Dimensions

The following are the module dimensions and weight for the module.

▶   Dimensions: 87.0 mm (width) × 100.0 mm (length) × 16.0 mm (height).

▶   Weight: TBD g (± 10 g)

Figure 7-1 shows the module outline.

Figure 7-1.   Module Outline Drawing 3D View
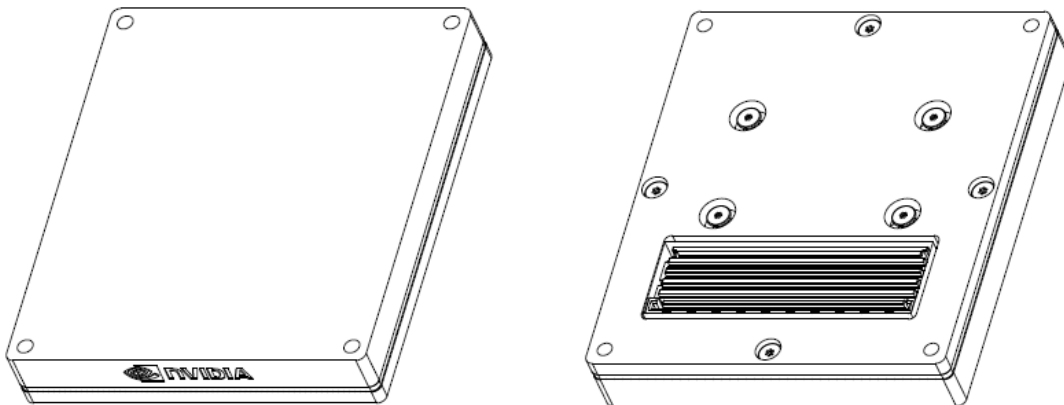
Figure 7-2. Module Mechanical Drawing Top View
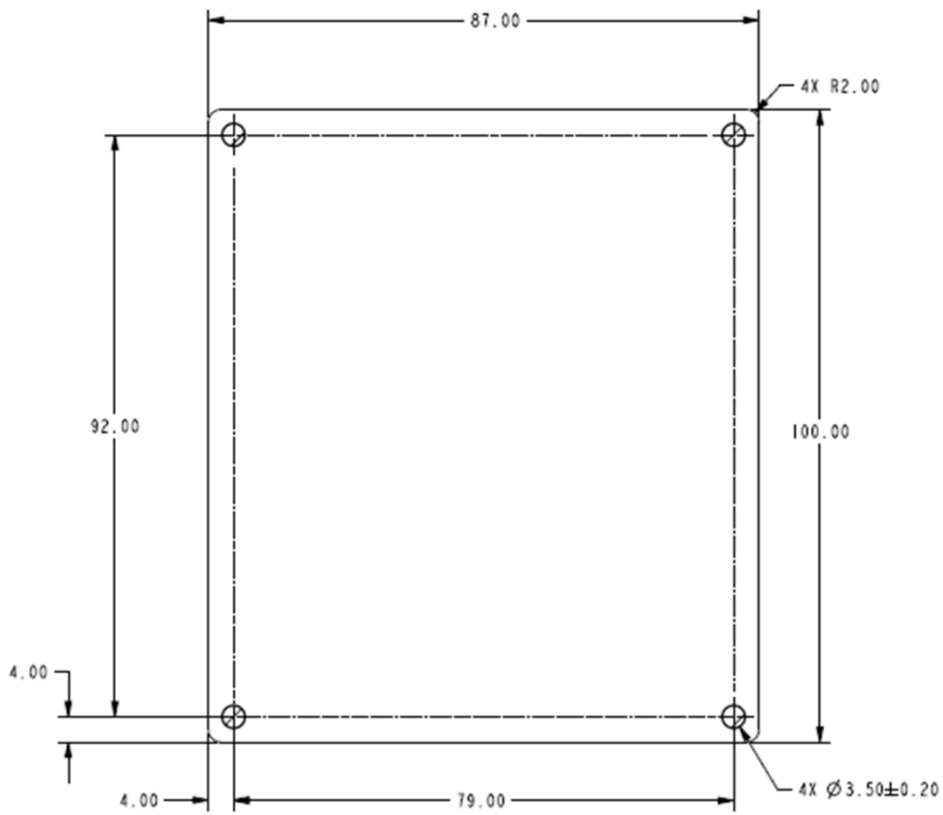
Figure 7-3. Module Mechanical Drawing Side View
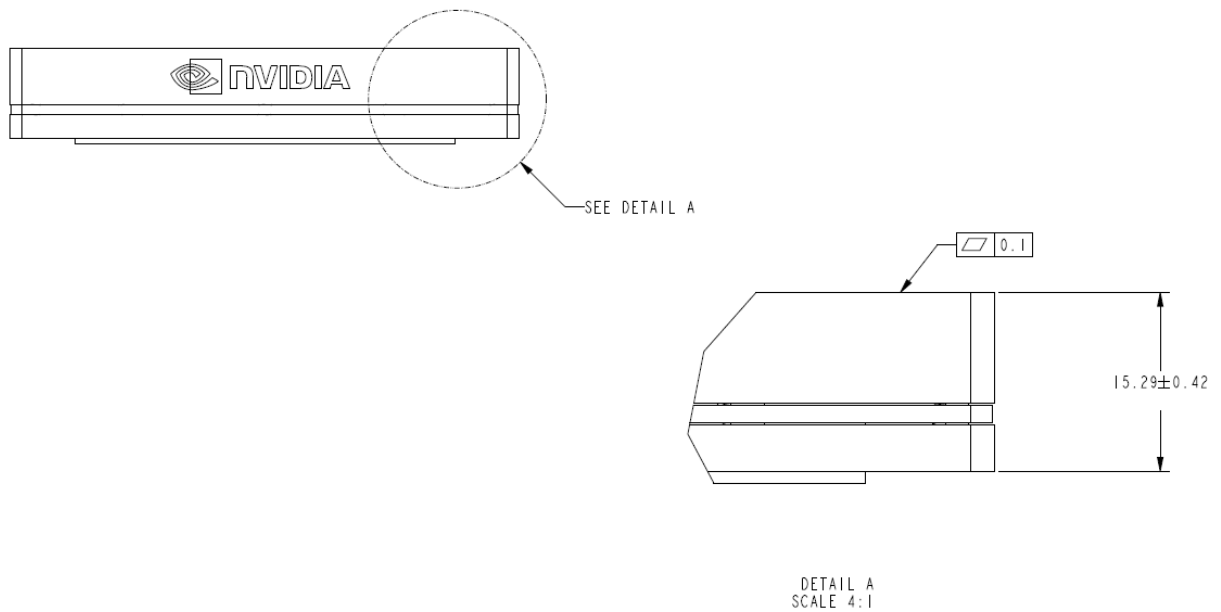
7.2.2　　Module Mounting Hole

The holes labeled "A" in the following figure are used for mounting purpose to mate the NVIDIA Jetson AGX Orin SOM, the system motherboard, and thermal solution.

Figure 7-4.　　Orin Module Mounting Hole