

Edge Computing AI Inference Module Thermal and Computing Benchmarks - Official Technical Overview & Hardware Datasheet

EXECUTIVE SUMMARY

The Edge Computing AI Inference Module Thermal and Computing Benchmarks represents a purpose-built, carrier-grade hardware platform engineered for deterministic, high-throughput artificial intelligence inference at the network edge. Designed to meet the rigorous demands of telecom operators, industrial IoT deployments, and latency-sensitive enterprise applications, this module integrates an advanced neural processing architecture with a ruggedized thermal management subsystem. The platform delivers sustained TOPS-class compute density within a compact, fanless or hybrid-cooled chassis, enabling real-time decision-making without backhaul dependency.

This datasheet provides comprehensive technical specifications, benchmark methodologies, thermal performance envelopes, and deployment guidance for system architects, pre-sales engineers, and procurement specialists evaluating edge AI infrastructure.



ARCHITECTURE & CHASSIS DESIGN

The Edge Computing AI Inference Module is built upon a heterogeneous system-on-module (SoM) architecture, combining a high-efficiency AI accelerator cluster with a general-purpose host processor and dedicated memory bandwidth. The chassis is constructed from extruded aluminum with integrated heat-pipe and vapor-chamber thermal coupling, enabling operation in ambient temperatures ranging from -40°C to +70°C without active forced-air cooling in standard configurations.

The module supports a modular mezzanine interface for optional 5G/LTE backhaul, Wi-Fi 6/6E, or fiber-optic uplink modules, ensuring seamless integration into existing telecom edge cabinets and street-side enclosures. All external interfaces are surge-protected and conformally coated for humidity,

salt-spray, and particulate resilience.

HARDWARE FEATURES

- AI Accelerator: Multi-core neural processing unit (NPU) delivering up to 32 INT8 TOPS and 8 FP16 TFLOPS
- Host CPU: 8-core ARM Cortex-A78AE or x86 equivalent, depending on SKU
- Memory: 16 GB LPDDR5 with ECC, expandable to 32 GB
- Storage: 128 GB industrial-grade eMMC 5.1 plus M.2 NVMe expansion slot
- Thermal Design: Passive heat-pipe cooling with optional 40 mm PWM fan kit for extreme environments
- Power Input: 9–36 VDC wide-range, with 1+1 redundant DC input on carrier board
- Networking: 2x 10GbE SFP+, 4x 1GbE RJ45, optional 5G sub-6 GHz module
- Management: Out-of-band BMC with IPMI 2.0, RESTful API, and SNMPv3
- Security: TPM 2.0, secure boot, hardware root of trust

COMPLIANCE & STANDARDS

The module is designed to meet or exceed the following standards:

- Safety: IEC 62368-1, UL 62368-1
- EMC: EN 55032 Class A, EN 55035, FCC Part 15 Subpart B

- Environmental: IEC 60068-2-1/2/14/30, IP40 rated enclosure
- Telecom: NEBS Level 3 (pending), ETSI EN 300 019
- RoHS: Directive 2011/65/EU, REACH compliant

TECHNICAL SPECIFICATIONS

Parameter	Specification
Form Factor	Compact module, 180 mm x 120 mm x 45 mm (W x D x H)
AI Compute	32 INT8 TOPS / 8 FP16 TFLOPS (peak)
Host CPU	8-core ARM Cortex-A78AE @ 2.2 GHz or x86 equivalent
Memory	16 GB LPDDR5 ECC (32 GB optional)
Storage	128 GB eMMC 5.1, M.2 NVMe expansion
Thermal Range	-40°C to +70°C ambient (passive); -40°C to +85°C with fan kit
Power Input	9–36 VDC, 1+1 redundant DC input
Power Consumption	15 W typical, 45 W maximum
Networking	2x 10GbE SFP+, 4x 1GbE RJ45, optional 5G/Wi-Fi
Management	BMC with IPMI 2.0, RESTful API,

	SNMPv3
Security	TPM 2.0, secure boot, hardware root of trust
Compliance	IEC 62368-1, EN 55032 Class A, NEBS Level 3 (pending)

ORDERING OPTIONS

- EC-AIM-32T-A: 32 TOPS NPU, 16 GB ECC, 128 GB eMMC, 2x 10GbE SFP+, passive cooling
- EC-AIM-32T-F: 32 TOPS NPU, 32 GB ECC, 256 GB NVMe, 2x 10GbE SFP+, fan kit included
- EC-AIM-16T-A: 16 TOPS NPU, 8 GB ECC, 64 GB eMMC, 2x 1GbE, passive cooling
- Optional Accessories: 5G sub-6 GHz module, Wi-Fi 6E module, DIN-rail mount kit, external antenna kit

