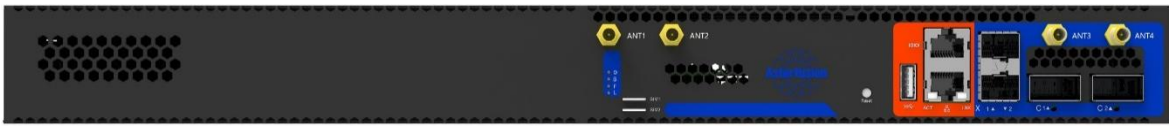
Overview

The Asterfusion ET3600 is an AI-ready edge network platform designed for modern deployments, including large-scale enterprise routing, security, BNG, MEC, and edge computing applications. Powered by up to two Marvell OCTEON 10 CN103XX processor featuring 8-core ARM 64-bit Neoverse N2 architecture, the platform combines high-performance computing, hardware-accelerated packet processing with up to 2x100Gbps networking throughput, inline security acceleration, and optional AI inference capabilities in a compact low-power appliance.

The platform supports major Linux distributions including Ubuntu, Debian, CentOS, and OpenWRT, enabling customers to deploy containerized edge applications and open-source software such as Nginx, iptables, UFW, and other edge services on a single device. Running Asterfusion's native AsterNOS-VPP router software, the ET3600 delivers enterprise-grade routing, secure VPN, firewall, and edge networking services optimized through VPP and DPDK acceleration.



ET3608-2P2S

ET3616-4P4S

An optional AI acceleration module delivering up to 160 TOPS of compute performance and 24 GB of onboard memory further extends the ET3600 with intelligent edge capabilities. The ET3600 supports deployment of modern AI frameworks and large language models (LLMs) in the 30B parameter class directly at the network edge, enabling use cases such as intelligent operations, automated troubleshooting, anomaly detection, and private AI assistants. By combining networking, security, edge compute, and AI acceleration within a single platform, the ET3600 enables tighter integration between AI workloads and network services while reducing hardware complexity, power consumption, and operational overhead.

Key Features and Benefits

- Compact 1RU edge network appliance
- 8 x 2.5GHz ARM64 Neoverse N2 Core
- 2 x 16GB pluggable DDR5 SO-DIMM, up to 48G
- Multiple interface combination:
  - 2x 2x100GE (QSFP28) + 2x 2x10GE (SFP+)
  - 2x100GE (QSFP28) + 2x10GE (SFP+)
- True inline crypto engine
- Optional AI accelerator module with 160TOPS INT8 computing performance
- Optional M.2 SSD up to 4TB
- Optional 5G/LTE extensible module
- Optional PTP/SyncE module with 20ns accuracy and OCXO holdover
- Up to 2 x 100Gbps intelligent data processing for routing, firewall, IPSec and SSL/TLS
- Power Consumption (full configuration & workload, w/o PoE):
  -ET3608: <100 W
  -ET3616: <200 W

Application Scenarios

Based on the open hardware-software decoupled architecture, the ET3600 combines a rich array of opensource software for control plane with hardware-optimized data plane. Here are some typical scenarios that can be used individually or in combination:

- **Border Router/VPN/Firewall:** AsterNOS-VPP (SONiC-VPP)
  - Hardware-optimized vector packet technology and DPDK accelerate data plane forwarding, delivering up to 2 x 100Gbps forwarding performance.
  - Multi-WAN Routing across Ethernet and 5G/LTE uplinks with traffic steering and failover.
  - Hierarchical QoS policies enable precise traffic for users, services, and applications.
  - Hardware-accelerated VPN with encryption/decryption engine supports up to 2 x 100Gbps throughput.
  - Advanced policy-based traffic control with GeoIP/GeoSite-aware filtering.

- **Broadband Network Gateway (BNG):** AsterNOS-VPP (SONiC-VPP)
  - AsterNOS-VPP supports subscriber management features such as PPPoE server and hierarchical QoS (HQoS), enabling deployment as a BNG/BRAS for broadband access networks.
  - The platform supports up to 50K concurrent PPPoE sessions, making it well-suited for small ISP deployments and low-density subscriber regions such as rural or remote access sites.

- **Edge ARM Server:** Standard Linux distribution with VPP/DPDK acceleration
  - Powered by an 8-core 2.5GHz ARM64 Neoverse N2 processor optimized for networking and edge infrastructure workloads.
  - Open Linux platform supporting standard software ecosystems, containers, and cloud-native applications.
  - Optimized for open software frameworks including DPDK, VPP, Nginx, featuring a hardware-based true inline crypto engine to boost cryptographic application performance.
  - Delivers up to 2 x 100Gbps accelerated networking performance for packet processing.
  - Maintains high energy efficiency, keeping total system power consumption below 100W even under full-workload operation.

- **AI Inference for Edge Networking:** AI Framework + Inference Engine + LLM
  - Combines edge networking, security acceleration, and AI inference capabilities in a compact and power-efficient platform.
  - Integrated AI module delivers up to 160 TOPS AI computing performance with 24GB LPDDR5 memory for edge AI workloads and local LLM inference.
  - Compatible with popular AI frameworks and runtimes including PyTorch, llama.cpp, Ollama, vLLM, and TensorFlow Lite.
  - Supports quantized AI models such as Llama 3 8B, GPT-OSS-20B-A3B, Qwen3 7B–30B, and other GGUF-based edge inference models.
  - Running Qwen3-30B with 2048-token input delivers approximately 37ms TTFT (Time to First Token) and 30 tokens/s generation throughput, with support for up to 32K context length.

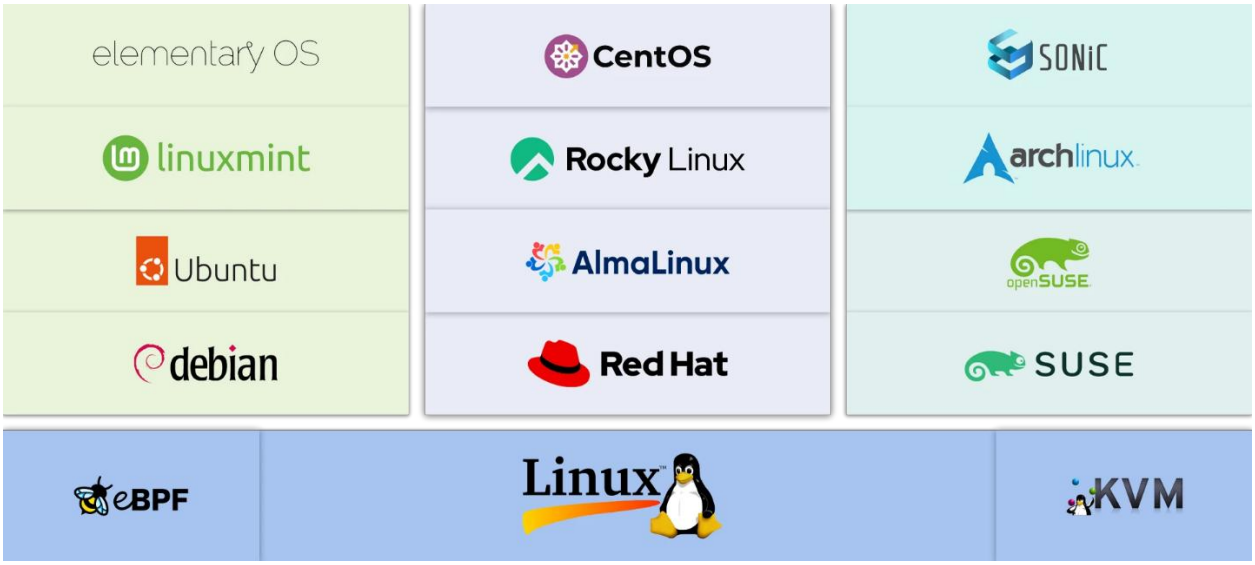
Additionally, users have the flexibility to install new software or develop their own software using the built-in toolchain as needed to address additional use cases.

Operating System

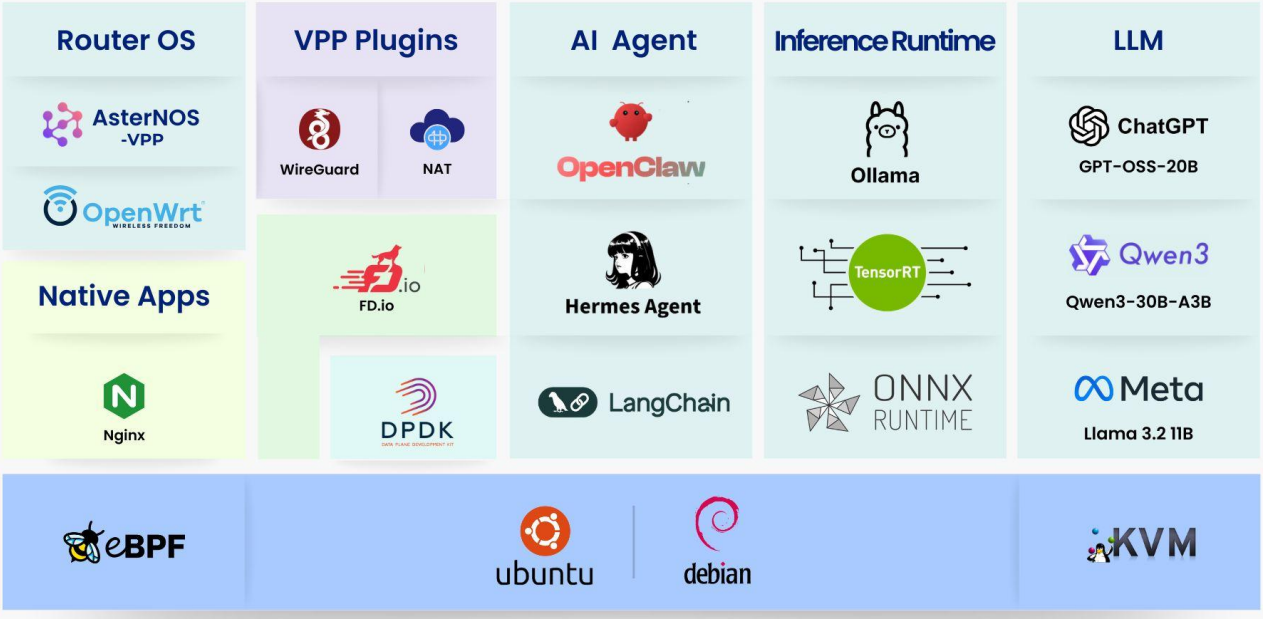
- Supports Ubuntu, Debian, SONiC and other Linux distribution, such as CentOS, OpenSUSE, Arch Linux, AlmaLinux, Rocky Linux, Linux Mint and Elementary OS
- Install and upgrade the OS using a USB disk with Arm Trusted Firmware and UEFI
- Embedded eBPF (extended Berkeley Packet Filter) in Linux kernel via XDP

Software Ecosystem

- Optimized DPDK (Data Plane Development Kit) tied to HW Acceleration
- Open-source routers, including VPP (Vector Packet Processing), AsterNOS-VPP, OpenWRT, etc.
- Open-source firewalls, including iptables, UFW, nftables, Firewalld, etc.
- Cloud-native deployment with Docker, Kubernetes, and containerized edge applications.
- Open-source load balances, including Nginx.
- Rich VPP plugins, including Marvel device plugin, QUIC, SRv6, LLDP, NAT64, SRTP etc.
- GCC, GDB, BinUtils, Buildroot and other tool chains.
- C/C++/Python/Go/Rust/Java/Lua and other programming languages.
- AI software ecosystem supporting PyTorch, TensorFlow Lite, Keras, ONNX Runtime, llama.cpp and Ollama.
- Applications from other Linux distributions on Ubuntu/Debian using Docker with direct access to the host network.
- Any software for ARM64 + Linux.

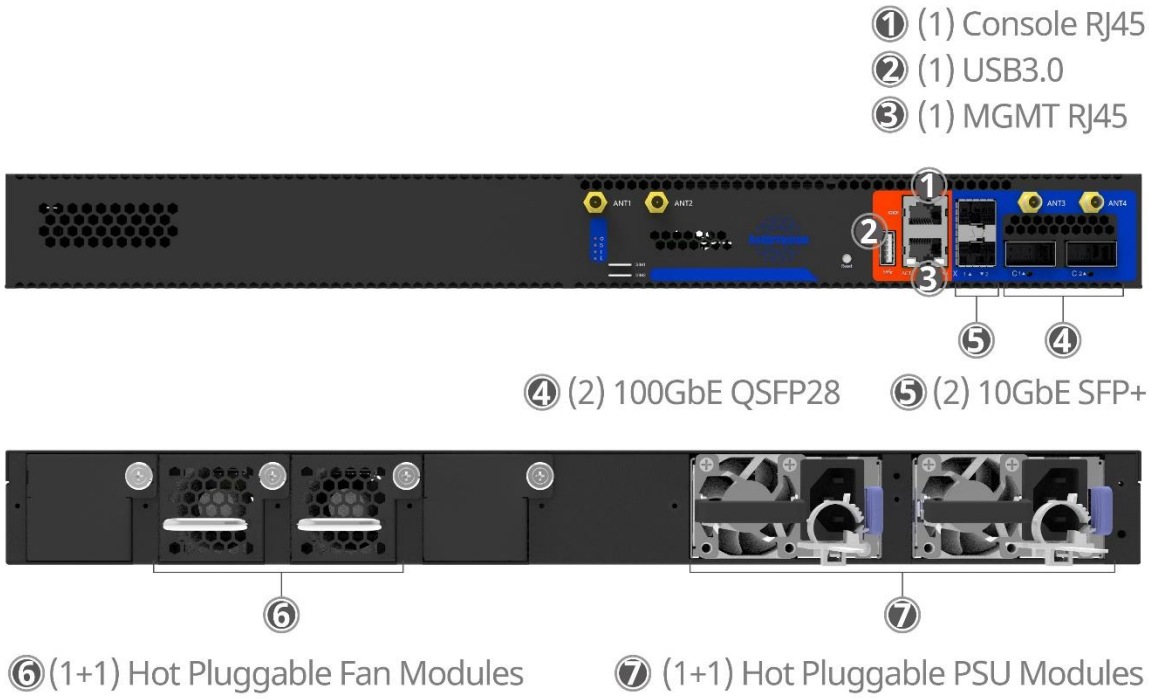


Operating System

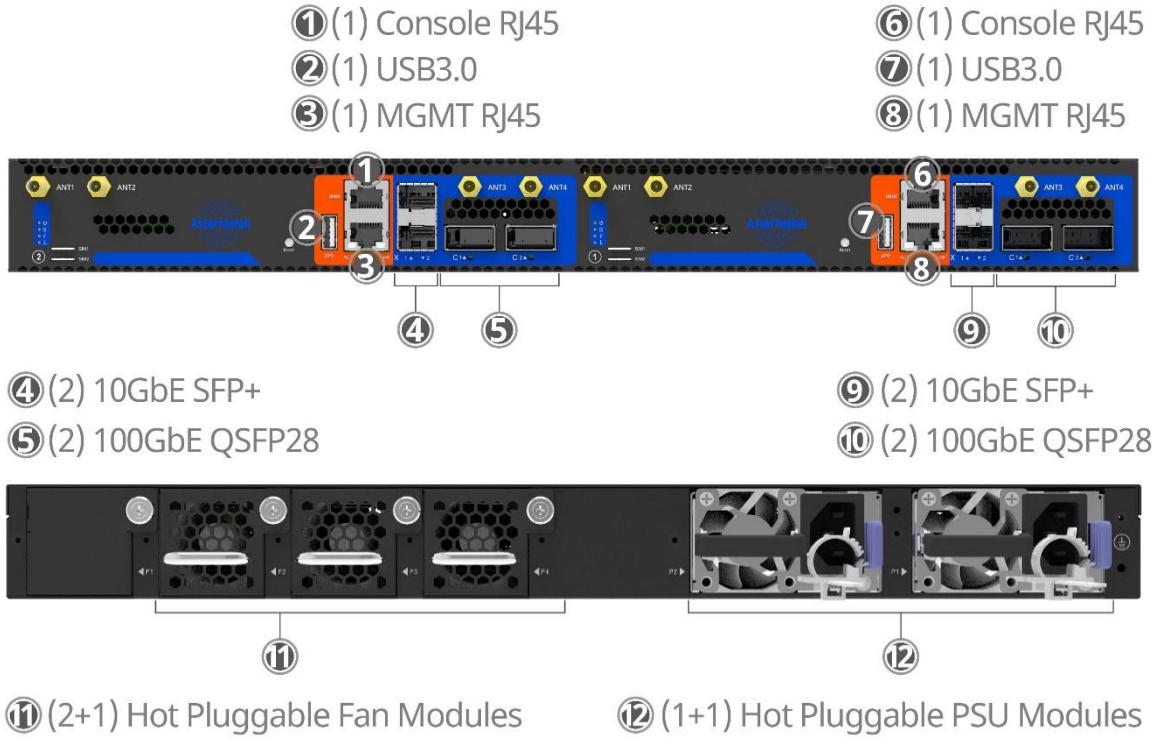


Software Ecosystem

Panel Illustration

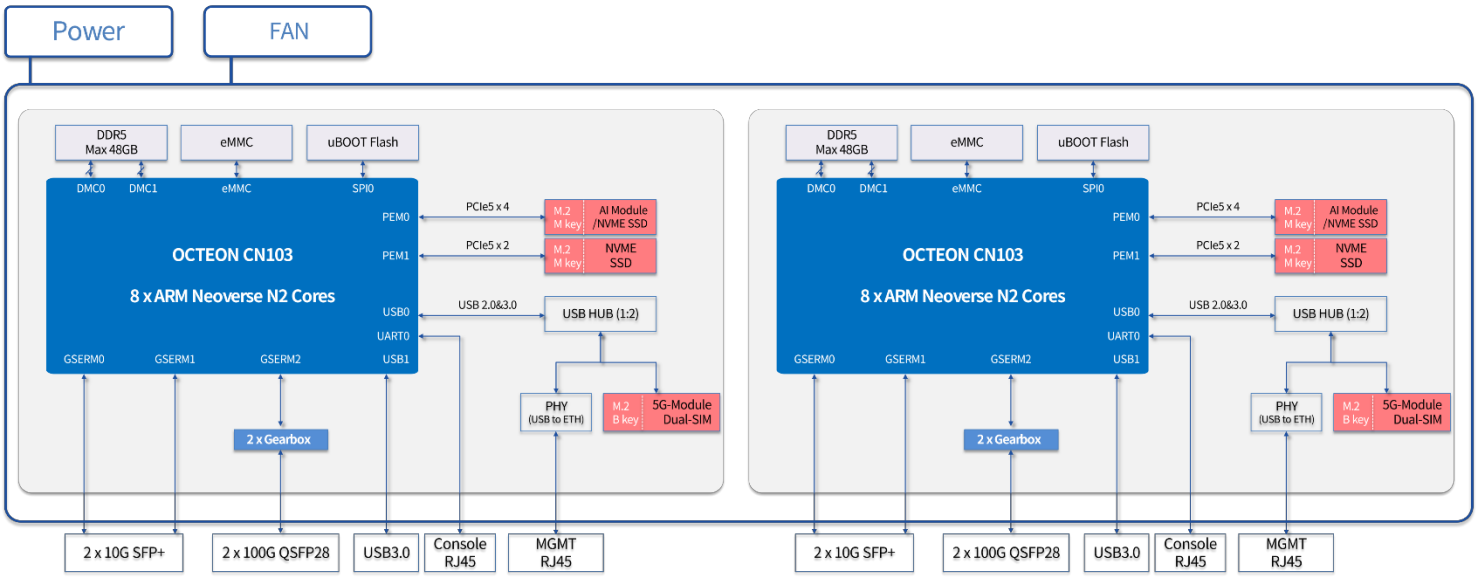


ET3608-2P2S



ET3616-4P4S

System Architecture



ET3600 series Edge Network Appliance system architecture

ET3600 Platform Specifications

Product Model		ET3608-2P2S	ET3616-4P4S
Network Interface	100GE (QSFP28)	2	2 x 2
	10GE (SFP+)	2	2 x 2
	5G/LTE (Optional)	2 SIM cards, M.2 B key	2x 2 SIM cards, M.2 B key
Management Interface	USB	1 x USB3.0	2 x USB3.0
	Console	1 x Console RJ45	2 x Console RJ45
	MGMT	1 x MGMT GE RJ45	2 x MGMT GE RJ45
Processor	DPU model	1 x Marvell CN103 8-core ARM64 2.5GHz	2 x Marvell CN103 8-core ARM64 2.5GHz
	Cache capacity	L2 8MB, L3 16MB	
Memory & Storage	Memory	16GB DDR5, maximum 48GB	2 x 16GB DDR5, maximum 2 x 48GB
	eMMC	1 x 64GB	2 x 64GB
	SSD (Optional)	1 x M.2 NVME	2 x M.2 NVME
Network Performance	Switching capacity	200Gbps	2 x 200Gbps
	Routing capacity	100Gbps	2 x 100Gbps
	Enc/Dec capacity	100Gbps	2 x 100Gbps
	PTP/SyncE accuracy	20ns	
	PTP/SyncE holdover	> 8hours	
Electrical Characteristics	Fan module	1 + 1	2 + 1
	Power module	1 + 1	
	Input voltage	100~240VAC	
	Maximum power consumption	100W (FULL configuration and workload)	200W (FULL configuration and workload)

Dimensions	Height	1U
	Dimensions (W x H x D, mm)	440 x 44 x 470
Operating Conditions	Operating temperature	0 - 45 °C
	Relative humidity	5% - 95%(non-condensing)

AI Accelerator Module Specification

Model	M2-AI-INFERENCE-24GB
AI Compute Perf.	• Up to 160 TOPS elastic AI computing performance • Up to 100 TFLOPS @ bFP16
Memory	• 24 GB LPDDR5 / LPDDR5X • Memory bandwidth: 153.6 GB/s
Storage	1 × 32 MB NOR Flash
Interface Type	M.2 Key M
PCIe Interface	1 × PCIe Gen4 interface Supports up to x4 lanes, compatible with x3 / x2 / x1 Maximum bandwidth up to 8 GB/s per interface
UART	Supports 2 × UART interfaces Standard baud rates supported, up to 921600 bps
I2C	1 × I²C controller Supports 100 kHz and 400 kHz communication rates
Dimensions (mm)	22 x 80 x 2.58
Weight	9 g (±10 g)
Operating Temperature	-10°C to +60°C
Operating Humidity	10% RH – 90% RH (non-condensing)
Supply Voltage	3.3 V ± 5%

Supported LLMs:

The AI module supports stable inference of large language models in the 7B–30B parameter range, using modern quantization (INT4/INT8).

- LLM: Qwen3-14B, Qwen2.5-7B, deepseek-8B, Qwen3-30B-A3B, GPT-OSS-30B-A3B, etc.
- MLLM: minicpm-o-2_6
- VLM: Qwen2.5-VL-7B, Qwen3-VL-2B, Qwen3-VL-4B, Qwen3-VL-8B, Qwen3-VL-30B-A3B

Ordering Information

Part Number	Description
ET3608-2P2S	Edge Network Appliance, 8 x ARM64 N2 CPU, 16GB DDR5, 64GB eMMC, 2 x 100GE + 2 x 10GE, 100W Power
ET3616-4P4S	Edge Network Appliance, 2 x 8 x ARM64 N2 CPU, 2 x 16GB DDR5, 2 x 64GB eMMC, 2 x 2 x 100GE + 2 x 2 x 10GE, 200W Power
Optional Components and Spares	
M2-AI-INFERENCE-24GB	AI Accelerator Module, 160 Tera-Operations Per Second, 32MB Flash, 24GB LPDDR5 with memory bandwidth 153.6GB, M.2 M Key 2280
SSD-M2-NVME-1TB	SSD,1TB, NVME, M.2 M Key 2280, PCIe5 x4
SSD-M2-NVME-2TB	SSD,2TB, NVME, M.2 M Key 2280, PCIe5 x4
SSD-M2-NVME-4TB	SSD,4TB, NVME, M.2 M Key 2280, PCIe5 x4
5G-M2-5Gbps	5G Module, USB2.0/3.0, M.2 B key, 5G NR, LTE, GNSS, 5.0Gbps (DL) / 650Mbps (UL)
PTP-20ns	PTP module, 20ns accuracy, BC support with holdover > 8hours
SVC-Basic-1Y-ET	Basic H/W Service and Warranty

Warranty and Service Support

Asterfusion ET3600 series appliance comes with 2-year Basic H/W service and warranty.

To acquire more info about company, products, and solutions: *www.cloudswit.ch*
Sales: *bd@cloudswit.ch*