

InfiniBand and RoCEv2 Gateway Modules for Compute Clusters - Official Technical Overview & Hardware Datasheet

EXECUTIVE SUMMARY

The InfiniBand and RoCEv2 Gateway Modules for Compute Clusters represent a purpose-built hardware solution that bridges the protocol gap between native InfiniBand fabrics and Ethernet-based RoCEv2 (RDMA over Converged Ethernet version 2) infrastructures. Designed for high-performance computing (HPC), artificial intelligence (AI) training clusters, and hyperscale compute environments, these gateway modules enable seamless, line-rate interoperability between InfiniBand and RoCEv2 domains without compromising latency, throughput, or congestion management capabilities.

As compute clusters grow in scale and heterogeneity, the need to interconnect disparate fabric technologies becomes paramount. Legacy InfiniBand deployments often coexist with newer RoCEv2-based Ethernet fabrics, creating islands of high-performance compute that require complex, software-based gateways or inefficient protocol translation layers. This hardware module eliminates those bottlenecks by providing a dedicated ASIC-based forwarding engine that natively parses, translates, and forwards both InfiniBand and RoCEv2 packets at wire speed. The result is a unified fabric that preserves the deterministic, low-latency characteristics of InfiniBand while extending the

reach and flexibility of RoCEv2 across standard Ethernet networks.

Each gateway module is hot-swappable and designed for integration into standard 1RU/2RU compute cluster chassis or dedicated gateway appliances. The module supports up to 400Gb/s per port, with aggregate throughput scaling to multiple terabits per second in a fully populated chassis. Advanced congestion control mechanisms, including InfiniBand credit-based flow control and RoCEv2 DCQCN (Data Center Quantized Congestion Notification), are implemented in hardware to ensure lossless operation under heavy load. The gateway also supports advanced routing, including InfiniBand Subnet Manager (SM) integration and Ethernet-based BGP/OSPF for RoCEv2 networks, enabling dynamic path selection and failover.



ARCHITECTURE & CHASSIS DESIGN

The gateway module is built around a custom, high-performance ASIC that integrates both InfiniBand and Ethernet MACs, a flexible packet parser, a translation engine, and a shared buffer architecture. The ASIC operates at line rate for all supported port speeds and provides deterministic forwarding latency of under 500 nanoseconds for both InfiniBand-to-RoCEv2 and RoCEv2-to-InfiniBand traffic. The module features a modular design with separate mezzanine cards for the InfiniBand interface (supporting HDR, NDR, and future XDR speeds) and the RoCEv2 interface (supporting 100GbE, 200GbE, and 400GbE). This modularity allows for field upgrades and customized port configurations to match specific cluster requirements.

The chassis design accommodates up to eight gateway modules in a 2RU form factor, with a passive backplane that provides redundant power and management connectivity. Each module slot supports hot-swap and is individually serviceable. The chassis includes dual redundant management modules that provide centralized configuration, monitoring, and firmware management via SNMP, REST API, and CLI. The backplane supports fabric-level telemetry, including per-port latency, throughput, and error counters, which are aggregated and exported to cluster management tools.

HARDWARE FEATURES

The InfiniBand and RoCEv2 Gateway Modules offer a comprehensive set of hardware features designed for demanding compute cluster environments:

- Protocol Translation: Native InfiniBand (IB) to RoCEv2 and RoCEv2 to IB translation at line rate, with support for both Unreliable Datagram (UD) and Reliable Connection (RC) transport services.
- Congestion Management: Hardware-based support for InfiniBand credit-based flow control, RoCEv2 DCQCN, and ECN (Explicit Congestion Notification) marking, ensuring lossless fabric operation.
- Quality of Service: Eight priority queues per port, with strict priority and weighted round-robin scheduling, plus support for InfiniBand Virtual Lanes (VLs) and RoCEv2 Priority Flow Control (PFC).
- Routing & Forwarding: Integrated InfiniBand Subnet Manager agent for SM-assisted routing, and Ethernet-based BGP/OSPF for RoCEv2 networks. Support for equal-cost multipath (ECMP) and dynamic failover.
- Telemetry: Per-port and per-flow telemetry, including latency histograms, throughput counters, and congestion events, exported via gRPC or REST API.
- Security: Secure boot, signed firmware, and optional MACsec encryption on RoCEv2 ports. InfiniBand partitions and RoCEv2 VLANs provide traffic isolation.
- Manageability: Dual redundant management modules, hot-swappable fans and power supplies, and support for in-band and out-of-band management.

COMPLIANCE & STANDARDS

The gateway modules comply with the following standards and specifications:

- InfiniBand Architecture Specification Volume 1 and Volume 2 (IBTA)
- RoCEv2 (RDMA over Converged Ethernet version 2) as defined in Annex A17 of the InfiniBand Architecture Specification
- IEEE 802.3 Ethernet standards for 100GbE, 200GbE, and 400GbE
- IEEE 802.1Q VLAN, IEEE 802.1p Priority, and IEEE 802.3x Flow Control
- IEEE 802.1AE MACsec (optional)
- PCIe 4.0 and 5.0 for host interface (management and telemetry)
- RoHS, REACH, and WEEE environmental directives
- FCC Part 15 Class A, CE, and VCCI emissions compliance
- UL/IEC 60950-1 and 62368-1 safety certifications

TECHNICAL SPECIFICATIONS

The following table summarizes the key technical specifications of the InfiniBand and RoCEv2 Gateway Modules.

Parameter	Specification
-----------	---------------

Form Factor	1RU / 2RU Chassis (module: half-width blade)
Switching Capacity	Up to 3.2 Tb/s per chassis (8 modules x 400Gb/s)
Port Speeds	InfiniBand: HDR/NDR (200/400Gb/s); RoCEv2: 100/200/400GbE
Protocol Support	InfiniBand UD/RC, RoCEv2, IPv4/IPv6, VLAN, PFC, ECN, DCQCN
Latency	< 500 ns port-to-port (IB to RoCEv2 or RoCEv2 to IB)
Congestion Control	Credit-based FC (IB), DCQCN, ECN, PFC (RoCEv2)
Management	Dual redundant management modules, SNMP, REST API, CLI, gRPC telemetry
Power Supply	1+1 Redundant AC/DC (chassis level), hot-swappable
Cooling	Hot-swappable fan trays, front-to-back airflow
Operating Temperature	0°C to 45°C (32°F to 113°F)
Dimensions (2RU chassis)	88.9 mm (H) x 440 mm (W) x 600 mm (D)

Weight (fully populated)	Approx. 18 kg (2RU), 10 kg (1RU)
Compliance	IBTA, IEEE 802.3, RoHS, REACH, FCC, CE, UL/IEC 62368-1

ORDERING OPTIONS

The gateway modules are available in multiple configurations to suit different cluster sizes and performance requirements. Ordering options include:

- IB-GW-400-2P: 2-port 400Gb/s InfiniBand (NDR) to 2-port 400GbE RoCEv2 gateway module
- IB-GW-200-4P: 4-port 200Gb/s InfiniBand (HDR) to 4-port 200GbE RoCEv2 gateway module
- IB-GW-100-8P: 8-port 100Gb/s InfiniBand (EDR) to 8-port 100GbE RoCEv2 gateway module
- CH-2RU-8S: 2RU chassis with 8 module slots, dual management modules, and redundant power supplies
- CH-1RU-4S: 1RU chassis with 4 module slots, single management module, and redundant power supplies

Each module ships with a comprehensive documentation kit, including

installation guide, configuration manual, and API reference. Optional support contracts provide 24/7 technical assistance and advanced replacement.

