QFX5240 SWITCHES DATASHEET

Product Overview

The QFX5240 Switches meet the advanced AI data center networking requirements of large-scale clusters. QFX5240 switches work with the automation in Juniper Apstra to assure daily operation in AI and ML workload training and access.

QFX5240 Switches:

— deliver high-density 800GbE ports on a fixed form factor with software to provide advanced network services tuned to the specific needs of AI/ML workloads

— are a foundation of AI networks, ensuring fast job completion time (JCT) to speed training through high GPU utilization

— help teams managing AI/ML environments realize improved economics

Product Description

Continuous evolution of AI/ML technology along with new applications are driving the next major shift in bandwidth requirement within the data center fabric. Juniper Networks® QFX5240 800GbE switch (64 port 800GbE) is a next-generation, fixed-configuration platform designed for spine, leaf, and border switch roles. The switch provides flexible, cost-effective, high-density 800GbE, 400GbE, 100GbE, and 50GbE interfaces for intra-IP fabric connectivity as well as higher density 200/400GbE NIC connectivity for AI/ML use cases. It's 51.2Tbps unidirectional throughput meets the bandwidth requirement of AI/ML workloads and storage systems with latency in the range of 700-750ns. Remote Direct Memory Access (RDMA) is the de-facto data transfer technology used in AI/ML workloads, and it uses Remote Direct Memory Access over Converged Ethernet v2 (ROCEv2) for transport at the network layer. QFX5240 supports ROCEv2 along with congestion management features like PFC, explicit congestion notification (ECN), and data center quantized congestion notification (DCQCN).

The QFX5240 helps reduce the number of network nodes deployed—decreasing the total power consumption of the data center fabric and improving the carbon footprint of the data center. These improvements are possible by having different breakout options like 128x400GbE and 256x200GbE/100GbE.

Table 1: QFX5240 Product Highlights

AI Data Center	• Leaf/spine in AI/ML cluster • ROCEv2 for AI/ML workloads • DCQCN-PFC, ECN for congestion management • Support for PFC watchdog for storm avoidance • Dynamic load balancing (DLB) for better load balancing • Configurable hash-bucket size to suit different flow scale
Cloud-ready Data Center	• Leaf/spine in IP fabric • Leaf/spine/super spine in EVPN-VXLAN fabric • Support for EVPN-VXLAN • 136K MAC scale • 860K IPv4 route scale
Port Options	• 64 ports of 800GbE • 128 ports of 400GbE (achieved with breakout cable) • 256 ports of 100GbE (achieved with breakout cable) • 256 ports of 50GbE (QFX5240-64OD) (achieved with breakout cable)
Platform Parameters	• Throughput: 51.2Tbps unidirectional • Buffer: 165MB • Tool less rack mount kit • Hot swappable power supplies and FAN trays • Power supply redundancy • Remote power cycling capability

Features and Benefits

AI/ML Design

Artificial intelligence puts new challenges on compute, network, and storage solutions with large models that run in parallel across many GPUs for training. These models require fast job completion time (JCT) with minimal delays for the last GPU to finish its calculations, that is, low tail latency. Architects optimize the cluster performance through rail-optimized design (Read this Juniper White Paper for more information about AI/ML cluster design). As model sizes and datasets continue to grow, designs must accommodate more GPUs in the cluster, requiring that the network seamlessly scale, without compromising performance, or introducing communication bottlenecks.

The QFX5240 meets the needs of these large-scale AI networks. The switch provides:

- 64 ports of 800GbE on a 2 U switch to reduce costs on both space and total power utilization
- Choice of connectivity with both OSFP and QSFP-DD variants of 800GbE for leaf-spine connectivity
- Availability of 2x400GbE per 800GbE port for increased connectivity to the GPUs
- Advanced telemetry capabilities to support ECN/PFC counters
- Fine-grained, load-balancing capability to handle reduced flow entropy
- Automation of rail-optimized design through Apstra

Automation

Automation tools, such as Apstra, ensure the reliable set up of expansive networks with ongoing verification of the deployment along with monitoring of operations. Apstra intent-based networking delivers full Day 0 through Day 2+ capabilities for IP/EVPN fabrics with closed-loop assurance in the data center. Apstra provides a broad set of operational capabilities, with multiple built-in intent-based analytics probes, flow visibility, and analysis to ensure that the AI network is running as designed. Apstra provides a simple UI workflow to create custom intent-based analytics to capture, enrich, and visualize data from the AI network.

Monitoring

The QFX5240 supports Junos® telemetry interface, a modern telemetry streaming tool that provides performance monitoring in complex, dynamic data centers. Streaming data to a performance management system lets network administrators measure trends in link and node utilization and troubleshoot issues such as network congestion in real time.

Junos telemetry interface provides:

- Application visibility and performance management by provisioning sensors to collect and stream data and analyze the application and workload flow path through the network
- Capacity planning and optimization by proactively detecting hotspots and monitoring latency and microbursts
- Troubleshooting and root cause analysis via high-frequency monitoring and correlating overlay and underlay networks

Additionally, the Junos Evolved operating system supports a robust API set to support automation through Terraform, Ansible, zero-touch provisioning (ZTP), operations and event scripts, automatic rollback, and Python scripts.

QFX5240-64OD

Specifications

Hardware Specifications

Table 2: QFX5240 System Capacity

Parameter	QFX5240-64OD	QFX5240-64QD
System throughput	51.2/102.4 Tbps uni/bidirectional	51.2/102.4 Tbps uni/bidirectional
Max Forwarding Rate	21.2Bpps unidirectional	21.2Bpps unidirectional
Port density	64 ports of OSFP 800GbE	64 ports of QSFP-DD 800GbE
Max ports with breakout	64 × 800GbE, 128 × 400GbE, 256 × 100GbE (Supports up to 320 interfaces in future options)	64 × 800GbE, 128 × 400GbE, 256 × 100GbE, 256x50GbE
Dimensions (W x H x D)	17.26 x 3.46 x 25.52 in (43.8 x 8.8 x 64.8 cm)	17.26 x 3.46 x 25.52 in (43.8 x 8.8 x 64.8 cm)
Rack units	2 U	2 U
Weight	22kgs (48.50lbs) fully loaded without optics	22kgs (48.50lbs) fully loaded without optics
Operating system	Junos OS Evolved	Junos OS Evolved
CPU	Intel Ice Lake (4 core)	Intel Ice Lake (4 core)
Memory	32GB (16GBx2) of DDR4	32GB (16GBx2) of DDR4
Storage	2x480GB	2x480GB
Power	Redundant (1+1) hot-pluggable 3000W AC (200 to 240V) power supplies	Redundant (1+1) hot-pluggable 3000W AC (200 to 240V) power supplies
Cooling	Ports-to-FRUs (AFO) 4 hot-pluggable fan modules	Ports-to-FRUs (AFO) 4 hot-pluggable fan modules
Total packet buffer	165 MB	165 MB
Warranty	Juniper standard one-year warranty	Juniper standard one-year warranty

Table 3: QFX5240 Feature Matrix

Features
Layer 2 Features
STP—IEEE 802.1D (802.1D-2004)
Rapid Spanning Tree Protocol (RSTP) (IEEE 802.1w); MSTP (IEEE 802.1s)
Bridge protocol data unit (BPDU) protect
Loop protect
Root protect
VLAN—IEEE 802.1Q VLAN trunking
Routed VLAN interface (RVI)
Static MAC address assignment for interface
Global MAC learning disable
Link Aggregation and Link Aggregation Control Protocol (LACP) (IEEE 802.3ad)
IEEE 802.1AB Link Layer Discovery Protocol (LLDP)
Layer 3 Features
Static routing
OSPF v2/v3
Filter-based forwarding
VRRP/VRRPv3
IPv6
Virtual routers
Loop-free alternate (LFA)
BGP
IS-IS
Dynamic Host Configuration Protocol (DHCP) v4/v6 relay (stateless)
VRF-aware DHCP
IPv4/IPv6 over GRE tunnels
Multicast
Internet Group Management Protocol (IGMP) v1/v2/v3
Multicast Listener Discovery (MLD) v2
IGMP proxy, querier
IGMP v1/v2/v3 snooping
Intersubnet multicast using IRB interface
MLD snooping
Protocol Independent Multicast PIM-SM, PIM-SSM, PIM-DM, PIM-Bidir Multicast Source Discovery Protocol (MSDP)
Quality of Service (QoS)
L2 and L3 QoS: Classification, rewrite, queuing
Rate limiting: - Ingress policing: 1 rate 2 color, 2 rate 3 color - Egress policing: Policer, policer mark down action - Egress shaping: Per queue, per port
10 hardware queues per port (8 unicast and 2 multicast)
Strict priority queuing (LLQ), shaped-deficit weighted round robin (SDWRR)
Layer 2 classification criteria: Interface, MAC address, Ether type, 802.1p, VLAN
Congestion avoidance capabilities: WRED, ECN
Trust IEEE 802.1p
Configurable shared buffer and buffer monitoring
Congestion Notification Profile
Priority-based flow control (PFC)—IEEE 802.1Qbb
High Availability
Bidirectional Forwarding Detection (BFD)
Visibility and Analytics
Switched Port Analyzer (SPAN)

Features
Remote SPAN (RSPAN)
Encapsulated Remote SPAN (ERSPAN)
sFlow v5
Junos Telemetry Interface Management and Operations
Role-based CLI management and access
Junos OS Evolved configuration rescue and rollback
Image rollback
SNMP v1/v2/v3
Junos OS Evolved XML management protocol
Automation and orchestration
Zero-touch provisioning (ZTP)
Python
Junos OS Evolved event, commit, and OP scripts
Network Services
ROCEv2
DCQCN, PFC, ECN
PFC watchdog
Dynamic load balancing (DLB)
Configurable hash-bucket size

Environmental Ranges

Table 4: QFX5240-64OD operating parameters

Parameter	QFX5240-64OD
Operating temperature	0° to 40°C @6000 ft
Storage temperature	-40° to 70°C
Operating altitude	6000 ft
Relative humidity operating	5 to 90% non-condensing
Relative humidity nonoperating	5 to 90% non-condensing
Seismic	Zone 4 earthquake rating (GR-63 EQ zone 4)
Typical power draw	932-Watt 100% traffic with DACs (without optics power) @ 25C)

Ordering Information

Product SKU	Description
QFX5240-64OD-AO	64x800GE OSFP800, AC, Front-to-back airflow
QFX5240-64OD-CHAS	64x800GE OSFP800, spare chassis
QFX5240-64QD-AO	64x800GE QSFP-DD 800, AC, Front-to-back airflow
QFX5240-64QD-CHAS	64x800GE QSFP-DD 800, spare chassis

License SKUs

S-QFX5K-C5-A1-X (X=3,5,P)	Advanced 1 Software License (X Years Subscription, X=3,5, or P for Perpetual) for QFX5240-64OD/QD
S-QFX5K-C5-A2-X (X=3,5,P)	Advanced 2 Software License (X Years Subscription, X=1,3,5, or P for Perpetual) for QFX5240-64OD/QD
S-QFX5K-C5-P1-X (X=3,5,P)	Premium Software License (X Years Subscription, X=1,3,5, or P for Perpetual) for QFX5240-64OD/QD

Note: The information provided is from early prototyping and may vary from the actual GA product.

About Juniper Networks

At Juniper Networks, we are dedicated to dramatically simplifying network operations and driving superior experiences for end users. Our solutions deliver industry-leading insight, automation, security, and AI to drive real business results. We believe that powering connections will bring us closer together while empowering us all to solve the world's greatest challenges of well-being, sustainability, and equality.

Corporate and Sales Headquarters

Juniper Networks, Inc.

1133 Innovation Way

Sunnyvale, CA 94089 USA

Phone: 888.JUNIPER (888.586.4737)

or +1.408.745.2000

www.juniper.net

APAC and EMEA Headquarters

Juniper Networks International B.V.

Boeing Avenue 240 1119 PZ Schiphol-Rijk

Amsterdam, The Netherlands

Phone: +31.207.125.700