

QFX5240 LINE OF SWITCHES DATASHEET



Product Overview

The [QFX5240 line of Switches](#) meet the advanced [AI data center networking](#) requirements of large-scale clusters. The QFX5240 line works with the automation—such as [Apstra Data Center Director](#) (formerly Juniper Apstra)—to assure daily operation in AI and ML workload training and access.

The QFX5240 line of 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, and their low latencies ensure 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 line of 800GbE switches 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 (store and forward). 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. The QFX5240 line of Switches supports ROCEv2 along with congestion management features like Priority Flow Control (PFC), explicit congestion notification (ECN), and data center quantized congestion notification (DCQCN).

The QFX5240 line of Switches 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, 256x200GbE, and 320x100GbE.

Table 1: QFX5240 line of Switches Product Highlights

| | | Line of Switches | |
|-------------------------|--|------------------|--|
| | | QFX5241-32 | QFX5241-64 |
| AI Data Center | <ul style="list-style-type: none"> 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 Selective-DLB - enable/disable DLB based on opcode/byte match | | |
| Cloud-ready Data Center | <ul style="list-style-type: none"> Leaf/spine in IP fabric Leaf/spine/super spine in EVPN-VXLAN fabric Support for EVPN-VXLAN 136K MAC scale IPv4 unicast routes: <ul style="list-style-type: none"> 2,000,000 routes (RIB), 880,000 LPM routes (FIB) and 880,000 host routes (FIB) IPv6 unicast routes: <ul style="list-style-type: none"> 2,000,000 routes (RIB), 875,000 LPM routes (FIB) and 875,000 host routes (FIB) | | |
| Port Options | <ul style="list-style-type: none"> 32 ports of 800GbE 64 ports of 400GbE (achieved with breakout cable) 128 ports of 200GbE (achieved with breakout cable) 160 ports of 100GbE (achieved with breakout cable) | | <ul style="list-style-type: none"> 64 ports of 800GbE 128 ports of 400GbE (achieved with breakout cable) 256 ports of 200GbE (achieved with breakout cable) 320 ports of 100GbE (achieved with breakout cable) |
| Platform Parameters | <ul style="list-style-type: none"> Throughput: 25.6Tbps unidirectional Buffer: 83MB Tool less rack mount kit Hot swappable power supplies and FAN trays Power supply redundancy Remote power cycling capability | | <ul style="list-style-type: none"> Throughput: 51.2Tbps unidirectional Buffer: 165MB |

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 line 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
- 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 Data Center Director

Automation Data Center Director

Automation tools, such as Data Center Director, ensure the reliable set up of expansive networks with ongoing verification of the deployment along with monitoring of operations. Data Center Director delivers full Day 0 through Day 2+ capabilities for IP/EVPN fabrics with closed-loop assurance in the data center. Data Center Director 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. Data Center Director provides a simple UI workflow to create custom intent-based analytics to capture, enrich, and visualize data from the AI network.

Monitoring

The QFX5240 line of Switches 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.

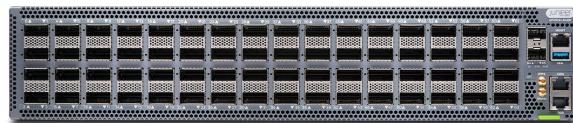
Junos telemetry interface provides:

- 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 Line of Switches



QFX5241-64QD Line of Switches



QFX5241-32OD Switch

Specifications

Hardware Specifications

Table 2: QFX5240 line of Switches System Capacity

| Parameter | QFX5240-64OD/QFX5241-64OD | QFX5240-64QD/QFX5241-64QD | QFX5241-32OD |
|-------------------------|---|--|---|
| System throughput | 51.2/102.4 Tbps uni/bidirectional | 51.2/102.4 Tbps uni/bidirectional | 25.6/51.2 Tbps uni/bidirectional |
| Max Forwarding Rate | 21.2Bpps | 21.2Bpps | 10.6 Bpps |
| Port density | 64 ports of OSFP 800GbE | 64 ports of QSFP-DD 800GbE | 32 ports of OSFP 800GbE |
| Max ports with breakout | 64 × 800GbE, 128 × 400GbE, 256 × 200GbE, 256 × 100GbE. Max port density with mixed speeds 320 (please use online port checker to validate mixed speed configurations) | 64 × 800GbE, 128 × 400GbE, 256 × 200GbE, 320 × 100GbE, 256 × 50GbE. Max port density with mixed speeds 320 (please use online port checker to validate mixed speed configurations) | 32 × 800GbE, 64 × 400GbE, 128 × 200GbE, 160 × 100GbE, 128 × 50GbE |
| 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) | 17.25 × 1.69 × 23.19 in (43.4 × 4.3 × 58.9 cm) |
| Rack units | 2 U | 2 U | 1 U |
| Weight | 22kgs (48.50lbs) fully loaded without optics | 22kgs (48.50lbs) fully loaded without optics | 15kgs (33.17lbs) |
| Operating system | Junos OS Evolved | Junos OS Evolved | Junos OS Evolved |
| Switch chip | Broadcom Tomahawk5 | Broadcom Tomahawk5 | Broadcom Tomahawk5 |
| CPU | Intel Ice Lake (4 core) | Intel Ice Lake (4 core) | Intel Ice Lake (4 core) |
| Memory | 32GB (16GBx2) of DDR4 | 32GB (16GBx2) of DDR4 | 32GB (16GBx2) of DDR |
| Storage | 2x480GB | 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 | Redundant (1+1) hot-pluggable 2400W power supplies |
| Cooling | Ports-to-PSU (AFO) 4 hot-pluggable fan modules | Ports-to-PSU (AFO) 4 hot-pluggable fan modules | Ports-to-PSU (AFO) 7 hot-pluggable fan modules |
| Total packet buffer | 165 MB | 165 MB | 83 MB |
| Warranty | Juniper standard one-year warranty | Juniper standard one-year warranty | Juniper standard one-year warranty |

Table 3: QFX5240 Line of Switches 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 |

| Features |
|---|
| Loop protect |
| Root protect |
| VLAN—IEEE 802.1Q VLAN trunking |
| Routed VLAN interface (RVI) |
| Static MAC address assignment for interface |

Features

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)

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

Features

Secure zero-touch provisioning (sZTP) (DevID)

Secure-BIOS

Secure-Boot

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

*Secure Boot is available on QFX5241 switches

Environmental Ranges

Table 4: QFX5240 line of switches operating parameters

| Parameter | QFX5240 line of switches |
|----------------------------------|--|
| Operating temperature | 0° to 40°C @ sea level |
| Storage temperature | -40° to 70°C |
| Operating altitude | Up to 6000 ft/3962 m |
| 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) |
| QFX5240-64OD/QFX5241-64OD | |
| Typical Power Consumption | 220-240 V: 613 W |
| Maximum Power Consumption | 220-240 V: 2198 W |
| QFX5240-64QD/QFX5241-QD | |
| Typical Power Consumption | 220-240 V: 616 W |
| Maximum Power Consumption | 220-240 V: 2258 W |
| QFX5241-32OD | |
| Typical Power Consumption | 220-240 V: 543 W |
| Maximum Power Consumption | 220-240 V: 1394 W |

Notes:

Max power consumption measured at 40°C ambient temperature with DR optics at 100% load with IMIX traffic.
 Typical power consumption measured at 25°C ambient temperature with DACs at 50% load with IMIX traffic, excluding transceivers.

Power consumption is subject to operating condition and unit-to-unit variations.

Electromagnetic compatibility

- FCC 47 CFR Part 15
- ICES-003 / ICES-GEN
- BS EN 55032
- BS EN 55035
- EN 300 386 V1.6.1
- EN 300 386 V2.2.1
- BS EN 300 386
- EN 55032
- CISPR 32
- EN 55035
- CISPR 35
- IEC/EN 61000 Series

- IEC/EN 61000-3-2
- IEC/EN 61000-3-3
- AS/NZS CISPR 32
- VCCI-CISPR 32
- BSMI CNS 15936
- KS C 9835 (Old KN 35)
- KS C 9832 (Old KN 32)
- KS C 9610
- BS EN 61000 Series

Environmental compliance

- Restriction of Hazardous Substances (RoHS)
- Toxic Substances Control Act (TSCA)
- Persistent Organic Pollutants (POPs)
- Recycled Material Waste Electronics and Electrical Equipment (WEEE)
- Registration, Evaluation, Authorization and Restriction of Chemicals (REACH)
- Substances of Concern in Products (SCIP)

Safety

- UL 60950-1:2007 R10.14 Information Technology Equipment
- CAN/CSA-C22.2 No. 60950-1-07, Amd 1:2011, Amd 2:2014 Information Technology Equipment
- IEC 62368-1:2014 (2nd Edition) Audio/Video, Information and Communication Technology Equipment (Include all country deviation)
- IEC 62368-1:2018 (3rd Edition) Audio/Video, Information and Communication Technology Equipment (Include all country deviation)
- EN 62368-1:2014+A11:2017 Audio/Video Information and Communication Technology Equipment
- UL/CSA 62368-1:2019 (3rd Edition) Audio/Video, Information and Communication Technology Equipment
- IEC/EN 60825-1 Safety of Laser Products – Part 1: Equipment classification and requirements

Ordering Information

| | |
|-------------------|--|
| QFX5241-64OD-AO | 64x800GbE OSFP, AC, AFO |
| QFX5241-64OD-DO | 64x800GbE OSFP, DC, Air-Out |
| QFX5241-64QD-AO | 64x800GbE QSFP-DD, AC, AFO |
| QFX5241-64QD-DO | 64x800GbE QSFP-DD, DC, Air-Out |
| QFX5240-2U-4PRMK | 4post Tool-less RMK for 2RU QFX5240 Series |
| QFX5240-2U-FANAO | Air Out FANs for QFX5240 Series-2RU |
| QFX5240-PWR-AC-AO | QFX5240 Series-2RU AC PSU, Air Out |

| | |
|-------------------|---|
| QFX5241-PWR-DC-AO | QFX5240 Series-2RU DC PSU, Air Out |
| QFX5240-64OD-AO | 64x800GE OSFP800, AC, Front-to-back airflow |
| QFX5241-64OD-CHAS | 64x800GE OSFP800, spare chassis |
| QFX5240-64QD-AO | 64x800GE QSFP-DD 800, AC, Front-to-back airflow |
| QFX5241-64QD-CHAS | 64x800GE QSFP-DD 800, spare chassis |
| QFX5240-64OD-CHAS | 64x800GE QSFP-DD 800, spare chassis |
| QFX5240-64QD-CHAS | 64x800GE QSFP-DD 800, spare chassis |
| QFX5241-32OD-AO | 32x800GbE OSFP, AC, AFO |
| QFX5241-32OD-DO | 32x800GbE OSFP, DC, Air-Out |
| QFX5241-1U-PWR-AO | QFX5241-1RU AC PSU, Air Out |
| QFX5241-1U-PWR-DO | QFX5241-1RU DC PSU, Air Out |
| QFX5241-1U-FAN-AO | Air Out FANs for QFX5241-1RU |
| QFX5241-1U-4PRMK | 4post Tool-less RMK for 1RU QFX5241 |

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-OD/QD line of switches |
| 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-OD/QD line of switches |
| 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-OD/QD line of switches |

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

Optics and Transceivers

Up-to-date information on supported optics can be found on the [Hardware Compatibility Tool](#). The QFX5240 line of Switches supports varying port speeds at 800GbE, 400GbE, and 100GbE, with different transceiver options.

Useful Links

[Feature Explorer](#)

[Hardware Compatibility Tool](#)

[Recommended Releases](#)

About Juniper Networks

Juniper Networks is leading the convergence of AI and networking. [Juniper's Mist™ AI-native networking platform](#) is purpose-built to run AI workloads and simplify IT operations assuring exceptional secure user and application experiences—from the edge, to the data center, to the cloud. Additional information can be found at www.juniper.net, [X](#), [LinkedIn](#), and [Facebook](#).

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

