# Cisco Nexus 9508 Switch Power and Performance

The Cisco Nexus 9508 brings together data center switching power efficiency and forwarding performance in a high-density 40 Gigabit Ethernet form factor.

The Cisco Nexus® 9500 Series is a family of modular switches that delivers industry-leading high-performance and high-density 1,10, 40, and future 100 Gigabit Ethernet with support for application infrastructure programmability. The 8-slot Cisco Nexus 9508 Switch (Figure 1) is the first available platform in the family.

Figure 1. Cisco Nexus 9508 Switch
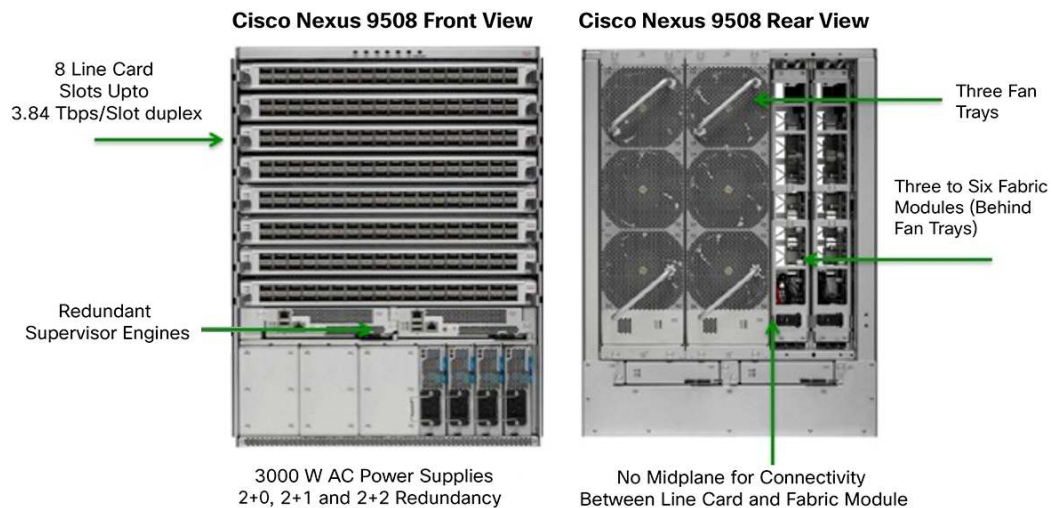
The Cisco Nexus 9508 Switch supports up to 1152 10-Gbps ports or 288 40-Gbps ports. Table 1 lists other specifications.

Cisco Nexus 9508 Architecture

The Cisco Nexus 9500 Series Switches share a common modular architecture that consists of switch chassis, supervisors, system controllers, fabric modules, line cards, power supplies, and fan trays.

Table 1.   Cisco Nexus 9508 Chassis and Forwarding Characteristics

Metric	Cisco Nexus 9508
Height	13 RU
Supervisor Slots	2
Fabric Module Slots	6
I/O Module Slots	8
Maximum Fabric Bandwidth per Slot	3.84 Tbps
Maximum Fabric Bandwidth per System	30 Tbps
Maximum Number of 10-Gbps Ports	1152
Maximum Number of 40-Gbps Ports	288
Maximum Forwarding Throughput per Line Card	2.88 Tbps
Maximum Forwarding Throughput per System	23.04 Tbps
Air Flow	Front-to-Back
Power Supplies	4 x 3 KW AC PSUs
Fan Trays	3

Chassis Design Optimized for Airflow

The chassis of the Cisco Nexus 9500 Series has an innovative midplane-free design (Figure 2). A midplane or passive midplane is commonly used in modular platform designs as a way to provide connectivity between line cards and fabric modules. The Cisco Nexus 9500 Series is the industry-first switch platform that eliminates the need for a chassis midplane. With a precise alignment mechanism, Cisco Nexus 9500 Series line cards and fabric modules directly attach to each other with connecting pins. Line cards and fabric modules have orthogonal orientations in the chassis so that each fabric module is connected to all line cards and vice versa. Without a midplane blocking the airflow path, the chassis delivers highly efficient cooling. It also allows a compact chassis design without the need for large cooling fans.

Figure 2.   Cisco Nexus 9500 Midplane-Free Chassis Design
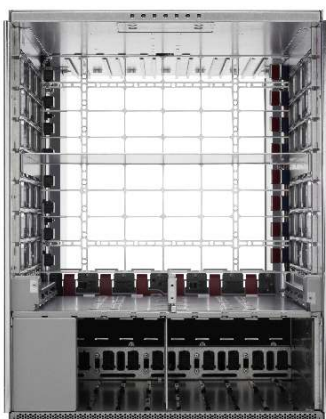
Cisco Nexus 9508 Power Efficiency

In addition to the high cooling efficiency, the Cisco Nexus 9500 Series also leads with high power efficiency. Its power supplies are certified for the industry highest efficiency standard, 80 Plus Platinum, with 90-94 percent power efficiency across all workloads.

Under the supervisor modules are eight power supply slots capable of supplying 3000 watts of AC power each; however, only two are needed to power a fully populated chassis. An additional two power supplies can be used for redundancy in a 2-plus-1 or 2-plus-2 configuration. The remaining four slots provide power headroom for later migration to support 100 Gbps and another generation of application-specific integrated circuits (ASICs).

Line cards and fabric modules of the Cisco Nexus 9500 Series are designed with minimal number of ASICs, a mix of 28-nanometer Cisco and 40-nanometer Broadcom custom and merchant ASICs, respectively, that reduces the number of heat pockets on a module.

The combined result of all these innovations is extremely low power consumption per port, as seen in Table 2:

Table 2.   Cisco Nexus 9508 Power Consumption

Port Speed	10 Gbps	40 Gbps
Watts per Port	~3.85	~15

Cisco Nexus 9508 Distributed Forwarding

In the Cisco Nexus 9500 Series Switches the packet lookup and forwarding functions in the data plane are conducted in a highly distributed fashion that involves line cards and fabric modules.

Both line cards and fabric modules of the Cisco Nexus 9500 Series are equipped with multiple network forwarding engines that perform packet lookup, processing, and forwarding. The fabric module for the Cisco Nexus 9508 has two network forwarding engines.

Each network forwarding engine is built with 32 40-Gbps links and a core switching capacity of 1.44 billion packets per second for even the smallest packet of 64 bytes.

The network forwarding engines use a combination of dedicated ternary content addressable memory (TCAM) table space and shared hash table memory to store Layer 2 and Layer 3 forwarding information (Figure 3). This shared hash table can be programmed in variable modes to provide different memory allocations for the three information tables.

Figure 3.   Dedicated and Shared Table Resources for Layer 2 and Layer 3 Forwarding

MAC Table	IP Host and IP Multicast Table	Longest Prefix Match Table
32k	16k	16k (TCAM)
256k Shared Memory Resource		

To further increase the system wide forwarding scalability, the Cisco Nexus 9500 Series Switches are designed to use tables on line cards and fabric modules for different forwarding lookup functions. The line cards store the Layer 2 MAC table and Layer 3 host table; therefore line cards are responsible for Layer 2 switching lookup and Layer 3 host routing lookup. The fabric modules host Layer 3 LPM tables, and fabric modules perform Layer 3 LPM routing lookup. Both line cards and fabric modules have multicast tables and take a part in distributed multicast lookup and packet replication. Figure 4 depicts the system wide forwarding scalability of Cisco Nexus 9500 Series Switches.

Figure 4.   Cisco Nexus 9500 Series Systemwide Forwarding Scalability

Fabric Module

Feature	Scale
Layer 3 Longest Prefix Match Table	128k
Layer 3 Host Table and Layer 3 Multicast	16k
Layer 2 MAC Table and Layer 2 Multicast	32k

Line Card

Feature	Scale
Layer 3 Host Table and Layer2/Layer3 Multicast	88k
Layer 2 MAC Table	160k

Cisco Nexus 9500 Series Fabric Module

A Cisco Nexus 9508 Switch can have up to six fabric modules that all function in active mode.

The fabric module of Cisco Nexus 9500 Series Switches performs the following functions in the modular chassis architecture:

- It provides high-speed nonblocking data forwarding connectivity for line cards. All links on network forwarding engines are active data paths. Each fabric module can provide up to 8 40-Gbps links to every line card slot. A Cisco Nexus 9500 chassis deployed with 6 fabric modules can potentially obtain 48 40-Gbps fabric paths for every line card slot. That is equivalent to 3.84 Tbps full duplex bandwidth per slot.
- It performs distributed longest-prefix match (LMP) routing lookup for IPv4 and IPv6 traffic. LPM forwarding information is stored on fabric modules on a Cisco Nexus 9500 Series Switch. It supports up to 128,000 IPv4 prefixes or 32,000 IPv6 prefixes.
- It performs distributed multicast lookup and packet replication to send copies of multicast packets to receiving egress line cards.

## Cisco Nexus 9500 Series Line Card Architecture

Network forwarding engines on a Cisco Nexus 9500 Series Switch line card perform ingress pipeline and egress pipeline packet processing.

Network forwarding engines on line cards conduct Layer 2 switching lookup and Layer 3 host routing lookup. Line cards are equipped with a sufficient number of network forwarding engines and fabric ports to support line-rate forwarding performance for all IP packet sizes.

36-Port 40-Gbps QSFP Line Card

The Cisco line card with product number N9K-X9636PQ (Figure 5) provides 36 40-Gbps Quad Small Form-Factor Pluggable (QSFP) front-panel ports. It has three network forwarding engines for packet forwarding; each supports 12 40-Gbps front-panel ports. The entire line card has a physical-layer-free (PHY-less) design. This reduces the data transport latency on the port by 100 nanoseconds, decreases the port power consumption, and improves reliability with a line card design of fewer active components.

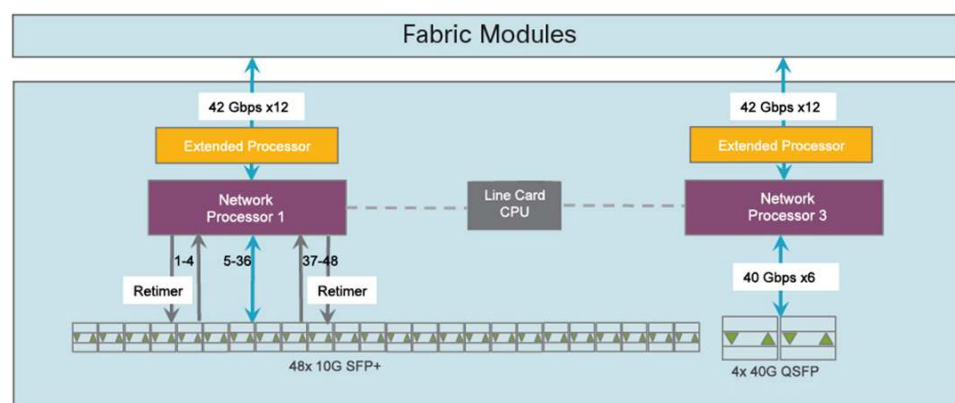**Figure 5.**   Cisco Nexus 9500 Series 36-Port 40-Gbps QSFP Line Card
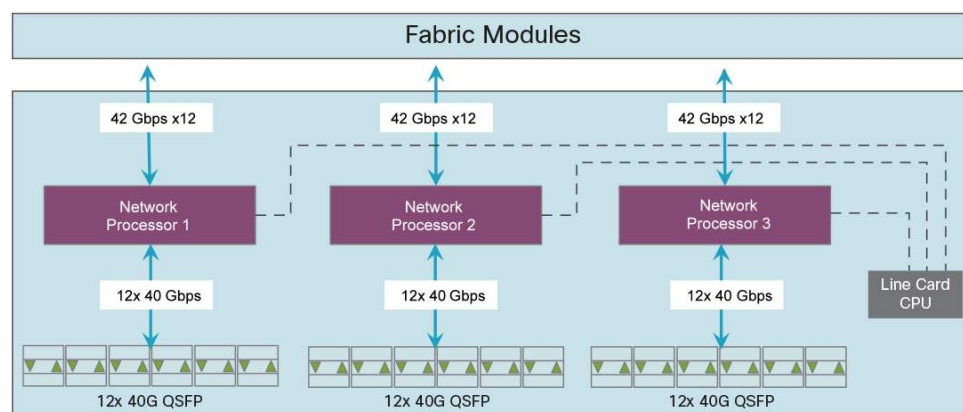
48-Port 1-Gbps or 10-Gbps SFP+ Line Card

The Cisco line card with product number N9K-X9564PX (Figure 6) provides 48 1-Gpbs or 10-Gbps Enhanced Small Form-Factor Pluggable (SPF+) ports and 4 40-Gbps QSFP ports.

There are two network forwarding engines on this line card. One supports the 48 1-Gbps or 10-Gbps front-panel ports and the other supports the 4 40-Gbps ports.

Because of the port speed flexibility, the front-panel ports can be operating at different speeds. Speed mismatch is one of the top reasons for port congestion and packet buffering. To accommodate this additional need for buffer, the line card is built with two extended processors, each offering an additional 40 MB packet buffer space while operating in a pass-through mode along the data path between a network forwarding engine and fabric modules. Transit traffic between front-panel ports and fabric modules, as well as locally switched traffic from a 10-Gbps port to a 1-Gbps port, can take advantage of the extended buffer space.
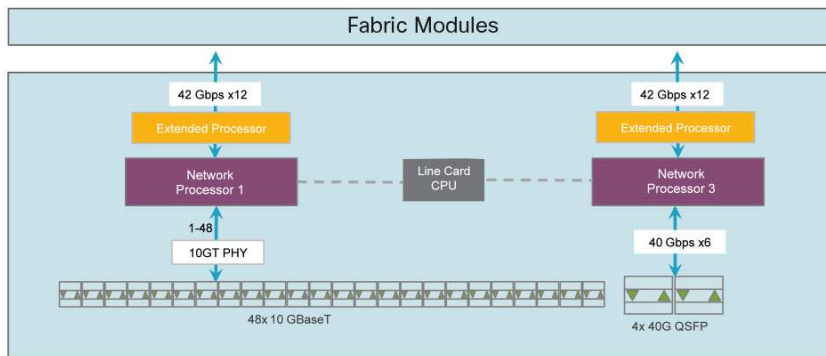
Figure 6.   Cisco Nexus 9500 Series 48-Port 1-Gbps or 10-Gbps SPF+ Line Card

48-Port 1-Gbps or 10BASE-T Line Card

The Cisco line card with product number N9K-X9564TX (Figure 7) provides 48 1-Gbps or 10BASE-T ports and 4 40-Gbps QSFP ports.

Figure 7.   Cisco Nexus 9500 Series 48-Port 1-Gbps or 10BASE-T Line Card

Cisco Nexus 9500 Series Table Sizes

Both line cards and fabric modules on the Cisco Nexus 9500 Series Switches have network forwarding engines that perform packet lookup and forwarding functions. The network forwarding engines have forwarding table resources, including TCAM tables and a programmable hash table called Unified Forwarding Table (UFT) that is shared for the Layer 2 MAC table, IP host table, and LPM table (Table 3). The shared memory resource can be flexibly allocated for Layer 2 MAC entries, IP host entries, or LPM entries. This sharing provides the ability to deploy the Cisco Nexus 9500 Series Switches in a broad range of data centers with a variety of application types.

**Table 3.**   Cisco Nexus 9500 Series Table Sizes

	Line Card	Fabric Module
Layer 2 MAC Table	160K	32K
Layer 3 Host Table	88K	16K
LPM Table	16K (TCAM)	128K

Cisco Nexus 9508 RFC Test Results

RFC 2544

RFC 2544 tests consist of connecting traffic generators to all 288 40-Gbps ports and sending bidirectional traffic between pairs of ports. In other words, half of the ports in the chassis talk to the other half and vice versa, simultaneously. The tests showed a 100 percent line rate throughput for any packet size (that is, no packet drops) with a fully loaded system (Figure 8).

Figure 8. Cisco Nexus 9508 RFC Test Results
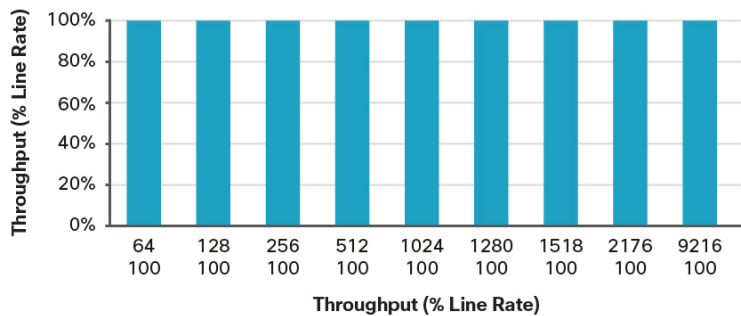
The measured latency for Layer 3 unicast traffic for this test ranged from a low of 1.6 microseconds (1617.2 nanoseconds) for 64-byte packets to a high of 3.5 microseconds (3562.6 nanoseconds) for 9216-byte packets.

RFC 2899

The RFC 2899 test consists of connecting all ports and sending traffic in a full mesh manner. That is, each port sends to every other port except itself. This test is designed to cause congestion in the system and to stress the scheduling capabilities of switches.

The Cisco Nexus 9508 Switch demonstrated a nearly 78 percent aggregate forwarding rate as percentage of line rate during congestion conditions for Layer 3 traffic flows. There was no head of line blocking (HOLB) observed. This means that congestion on a 40-Gbps port did not affect the performance of other switch ports.

RFC 3918

The RFC 3918 test consists of sending multicast traffic where one port transmits and all other ports in the system receive as the number of multicast groups increases.

The Cisco Nexus 9508 Switch provided 100 percent aggregated throughput for IP multicast traffic with latencies ranging from a low of 1.6 microseconds (1652 nanoseconds) for 64-byte packets to a high of 3.5 microseconds (3465.3 nanoseconds) for 9216-byte packets.

Cisco Nexus 9508 Power Consumption Analysis

This section illustrates the power consumption of the Cisco Nexus 9508 system with different configurations and under different types of loads.

A typical configuration may consist of the following components:

- 3 fan trays
- 6 fabric cards
- 2 supervisors
- 2 system controllers
- 2 x 36 x 40-Gbps line cards

The power consumption for such a system is shown in Table 4.

Table 4. Cisco Nexus 9508 Power Consumption (Typical) with 2x 40 GigE Line Cards

Component	Power Draw	Quantity	Extended
Fan	176 W	3	528 W

Fabric	176 W	6	1056 W
Supervisor	69 W	2	138 W
System Controller	13 W	2	26 W
36 x 40-Gbps Line Cards	260 W	2	520 W
Total			2268 W

A fully loaded configuration consists of the following components:

- 3 fan trays
- 6 fabric cards
- 2 supervisors
- 2 system controllers
- 8 x 36 x 40-Gbps line cards

The power consumption for such a system is shown in Table 5.

Table 5.   Cisco Nexus 9508 Power Consumption (Typical) with Fully Loaded System

Component	Power Draw	Quantity	Extended
Fan	176 W	3	528 W
Fabric	176 W	6	1056 W
Supervisors	69 W	2	138 W
System Controller	13 W	2	26 W
36 x 40-Gbps Line Cards	260 W	8	2080 W
Total			3828 W

At a 100 percent line rate, the power consumption for such a system goes to 5470 watts in the worst-case scenario of 64-byte packets and to 4746 watts for traffic with a mix of packet sizes.

Conclusion

The Cisco Nexus 9500 Series Switches are industry-leading data center class switches offering high port density for 1-, 10-, 40-, and future 100 Gbps connectivity and true line rate and low latency forwarding performance.

The midplane-free chassis design enables cooling efficiency. The combination of merchant and custom silicon allows the line cards to have few ASICs while providing record-setting performance, resulting in the lowest number of heat sinks in the system. With these innovations, the front-to-rear airflow, and its 80 Plus Platinum certified highly efficient power supplies, the Cisco Nexus 9500 Series Switches set a record for energy efficiency for data center - class switches.

The Cisco Nexus 9508 Switch provides high-density 40 Gigabit Ethernet in the data center with the best possible power consumption and latency figures even with a fully loaded system under stress, as demonstrated by the RFC test results.

For More Information

Cisco ACI strategy: http://www.cisco.com/go/aci