BROADCOM®

Product Brief

BCM78909
51.2-Tb/s Multilayer CPO Switch with 100G SerDes

Key Features

The BCM78909 has extensive features to address the rapidly increasing scale of data center network deployments and distributed computing applications:

• Large forwarding databases

• Powerful load balancing

• Extensive support for streaming and in-band telemetry

• Advanced congestion management

• Robust buffer performance, including many-to-one burst absorption capabilities that assist in TCP Incast scenarios

Applications

• Data center spine, leaf, ToR, blade, and aggregation switching

• Mobile core switches

• Cloud computing

• Large-scale enterprise campus backbone

• Service provider core switching

Overview

The Broadcom® BCM78909 51.2-Tb/s Co-Packaged Optics Switch is a high-radix, high-bandwidth network switching device that supports up to 64 × 800GbE, 128 × 400GbE, 256 × 200GbE, 320 × 100GbE, or 320 × 50GbE ports with integrated optical engines supporting 128xFR4 connectivity. The device family features a maximum of 64 integrated Peregrine SerDes cores, each with eight integrated 106-Gb/s PAM4 SerDes transceivers and associated physical coding sublayer (PCS). The BCM78909 delivers high-bandwidth, glueless network connectivity up to 51.2 Tb/s on a single chip.

The integrated optical engines enable direct high-speed I/O to the BCM78909 using 128 single-mode optical fibers to the front panel. Each optical engine is directly driven from the Peregrine SerDes, providing the lowest-power and lowest-cost optical interconnect solution. Each fiber is capable of supporting four individual 100G links using coarse wavelength division multiplexing (CWDM) with nominal wavelengths of 1271 nm, 1291 nm, 1311 nm, and 1331 nm. Co-packaging with optics eliminates the need to source and manage high-speed pluggable transceivers in deployment, dramatically lowers system power consumption, and simplifies system design leading to improved network reliability. The laser source is provided by 16 efficient pluggable laser modules (ARLM-96F8DMZ) which are Class 1 eye safe and field replaceable, enabling operator-level serviceability in the field.

The BCM78909 is a high-performance and high-capacity device designed to meet the requirements for next-generation data center and cloud computing environments. The BCM78909 architecture delivers complete L2 and L3 switching, routing, and tunneling capabilities at line rate and maximum port density, with low power and latency. Software compatibility is maintained across the StrataXGS® product portfolio to simplify customer designs and reduce customer time-to-market.

As server interfaces transition to higher Ethernet speeds and virtualization continues to increase link utilization, data center networks demand switches with dense 50GbE to 800GbE connectivity at the edge and aggregation layers. With up to 64 × [PM8x100] SerDes and flexibility in configuring 50GbE, 100GbE, 200GbE, 400GbE, and 800GbE ports, a single BCM78909 switching chip can be used to build scalable, cost-effective leaf, spine, top-of-rack (ToR), blade, and aggregation switches across the entire data center. The BCM78909 has extensive features to address the rapidly increasing scale of data center network deployments and distributed computing applications.

With the BCM78909, customers can build data centers with much higher server node counts while dramatically improving per-port power efficiency compared to pluggable optics solutions. The BCM78909 is built using Broadcom-proprietary silicon photonics (SiPh) technology and incorporates advanced power-management features to minimize power based on the features in use.

Features

General Features

- 512 × 100G PAM4 all optical I/O
- Flexible port configurations: 400GBASE-FR4, 200GBASE-FR4, 100GbE CWDM4
- Oversubscription to maximize I/O throughput
- Low pin-to-pin latency in cut-through and store-and-forward modes
- IEEE 802.3bj Clause 91 forward error correction (FEC)
- Eight integrated optical engines providing up to 128 FR4 optical links
- QSFP-DD Common Management Interface Specification (CMIS) compliant
- Light source is provided by 16 pluggable laser modules

Data Center Features

- Hardware-based encapsulation
- Data center bridging capability exchange protocol (DCBX) congestion management: priority-based flow control (PFC) and Enhanced Transmission Selection (ETS)
- IEEE 802.1Qbg Edge Virtual Bridging and IEEE 802.1Qbh Bridge Port Extension support
- Per-virtual machine traffic shaping
- Support for VXLAN overlays

L2 and L3 Packet Processing Features

- Full IPv4 and IPv6 routing support
- Hardware-based encapsulations, including Multiprotocol Label Switching (MPLS), and generic routing encapsulation (GRE)
- Three-stage field processing with ingress field processor (IFP) stage support for the exact match feature
- Unified forwarding table for flexible allocation of L3 host, longest prefix match (LPM) entry, IFP key compression entry, and IFP exact match entry
- Flexible ingress and egress counter pools

Buffering and Traffic Management Features

- Integrated high-performance SmartBuffer memory for maximum burst absorption and service guarantees
- Full Quality of Service (QoS) support:
  - PFC
  - Weighted random early discard (WRED)
  - Single-rate Three Color Marking (srTCM) and Two-rate Three Color Marking (trTCM) for color marking and metering
- Congestion management capabilities including destination module flow control and ECN
- Dynamic load balancing for equal cost multipath (ECMP) groups
- Packetized MMU statistics
- Network monitoring
- Network congestion detection

- sFlow redirect
- Visibility and packet tracing
- Transient capture buffer
- Enhanced load balancing
- Enhanced trunk hashing capabilities: RTAG7, symmetric hash, flex hash, and resilient hash
- Support for jumbo frames up to 9416 bytes
- Support for 12 queues per port and the following scheduling algorithms: strict priority (SP), round-robin (RR), weighted round-robin (WRR), and weighted deficit round-robin (WDRR)

Additional Features

- Comprehensive time synchronization support:
  - Integrated IEEE 1588v2 processor
  - IEEE 802.1AS support
  - Synchronous Ethernet (SyncE)
  - Ingress and egress per packet times stamping
- sFlow support including ingress, egress, and flex sampler, with an option to forward truncated sFlow packets to a remote agent
- Four dedicated 25G auxiliary Ethernet ports can be used for management
- x4 PCI Express (PCIe) Gen5 interface supports a local CPU
- Adjustable voltage scaling (AVS) for reduced average and peak power
- Scalable StrataXGS architecture supports high-performance fixed and modular switch designs