

NP-4™ Network Processor

100Gbps NPU for Carrier Ethernet Applications

Mellanox's NP-4 is a highly flexible network processor providing wire-speed packet processing with both an integrated traffic manager and a control CPU. NP-4 provides the silicon core of next-generation Carrier Ethernet Switches and Routers (CESR).

Through programming the NP-4 delivers a variety of applications such as L2 switching, QVLAN stacking, MPLS and VPLS, and IPv4/IPv6 routing coupled with QoS for providing flow-based service level agreements (SLA).

Mellanox's NP-4 is a highly flexible network processor providing wire-speed packet processing with both an integrated traffic manager and a control CPU. The NP-4 offers the speed of an ASIC combined with the flexibility of a programmable microprocessor. It provides the silicon core of next-generation Carrier Ethernet Switches and Routers (CESR). Through programming the NP-4 delivers a variety of applications such as L2 switching, QVLAN stacking, MPLS and VPLS, IPv4/IPv6 routing coupled with QoS for providing flow-based service level agreements (SLA).

The NP-4 integrates into a single chip several functions that would normally be found in separate chips:

- 50-Gigabit full-duplex processing
- Classification search engines
- 50/100-Gigabit traffic manager
- On-chip Control CPU
- On-chip Quality of Service CPU
- Fabric Interface Chip (FIC) functionality
- Integrated MACs:
 - 48 1-Gigabit MACs
 - Ten 10/20-Gigabit Ethernet MACs
 - Three Interlaken MACs
 - One 40-Gigabit MAC

FLEXIBLE PACKET PROCESSING

NP-4 provides exceptionally flexible packet processing enabling system designers to future proof their designs to support new protocols and features through software updates. Packet parsing is supported for any field anywhere in the packet. Various table lookup options are provided with support for long lookup keys and results. Flows are classified based on any combination of extracted packet information. Any packet header and content can be edited and packets can easily be replicated to support multicast applications. A 'run to completion' processing model guarantees support for processing scenarios of any complexity. Large code space is provided to support complex applications as well as true hitless code updates.

HIGHLIGHTS

- Single-chip, programmable, 100-Gigabit throughput (50-Gigabit full duplex) wire-speed network processor
- Line card, services card, pizza box and switch card applications
- Based on Mellanox's NP-3 with performance scaling and an enhanced feature set
- On-chip CPUs for control CPU offload
- On-chip Fabric Interface Controller for interfacing to Ethernet fabrics as well as third-party fabric solutions
- System-wide traffic management with hierarchical scheduling
- Flexible processing with programmable packet parsing, classifying, modifying and forwarding
- IP reassembly
- Enhanced support for video streams and IPTV
- Embedded search engine eliminating the need for external co-processors
- On-chip OAM protocol processing offload
- Serdes interfaces configurable to various network interfaces:
 - Ten XAUI/RXAUI interfaces
 - 24 quad-speed SGMII ports or 48 tri-speed QSGMII ports
 - Three Interlaken MACs
- Support for OC-768 framer, switch fabrics and external 100G Ethernet MAC
- Single 40G MAC
- Internal TCAM
- On-chip hardware time-stamping supporting IEEE1588v2
- Support for Synchronous Ethernet ITU-T G.8261 as required by Circuit Emulation Services
- PCI-Express external host interface
- Comprehensive on-chip diagnostic hardware support

FEATURES SUMMARY

Integrated Traffic Management

- 180Mpps throughput
- Dynamic hitless resource allocation
- Dynamic hitless reconfiguration
- LAG shaping
- Work conserving and non-work conserving schedulers
- Frame sizes from 1 byte to 11 KB
- Total frame memory up to 4 Gbytes
- Up to 8M frames
- Per Flow Queuing (PFQ) with 5-level hierarchical scheduling:
 - 32 interfaces
 - 256 ports
 - 4K subports
 - 32K classes/users
 - 256K flows
- Policing: Per-flow metering, marking and policing for millions of flows
- Configurable WRED profiles
- Per flow per color WRED statistics
- Shaping: Single and Dual leaky bucket on committed/peak rate/bursts (CIR, CBS, PIR, PBS), with IFG emulation for accurate rate control
- Scheduling: WFQ and priority scheduling at each hierarchy level
- Per frame timestamp and timeout drop
- Hardware flow control per port and TM Interface/Port/Subport and Class
- Link-level flow control generation management scheme based on flexible traffic aggregation per source/destination TM congestion
- Class-based flow control

Packet Manipulation & Reassembly

- TOPs control of TM buffered data
- Data reordering
- Data reassembly (e.g. IP reassembly)

Enhanced Video Transmission

- Caching video streams for retransmission
- Video data awareness, IPTV fast channel zapping
- Video de-multiplexing
- Video streams path redundancy

Integrated Search Engines

- Flexibly defined switching, routing, classification and policy lookup tables with millions of entries per table
- Programmable keys and results (associated information) per table
- Support for long keys and long results per table entry
- Table entries stored in DRAM to reduce power dissipation and cost and provide large lookup tables headroom

Stateful Classifying & Processing

- Access to all 7 layers for classify and modify
- Maintains state of millions of sessions simultaneously
- On-chip state updates and learning of millions of sessions per second

Programming

- Large code space memory for multiple and complex applications
- Hitless code upgrades
- Single-image programming model with no parallel programming or multi-threading
- Automatic ordering of frames
- Automatic allocation of frames to processing engines (TOPs)
- Automatic passing of messages among TOPs
- Microcode compatible with Mellanox's NP-2, NP-3 and NPA network processors

Interfaces

- Serdes interfaces configurable to various network interfaces
- 40 Gigabit Ethernet MAC compatible to 802.3ba standard over XLAUI Multi Lane Distribution over 8 physical lanes
- Ten XAU interfaces:
 - Ten on-chip 10G/20G MACs
 - 3.125Gbps; 6.25Gbps per lane
 - Channelized operation with up to 256 transmit channels
 - In band and out of band flow control
 - Connection to Ethernet and TDM framers
 - Support for SPAUI packet mode
 - Support for RXAU protocol
- 24 quad-speed SGMII/1000Base-X Ethernet interfaces or 48 tri-speed QSGMII Ethernet interfaces
- Three Interlaken MACs
- External Host interface:
 - 1-lane PCI-Express 2.5Gbps for control CPU interface
 - Additional 2xSGMII GE ports
 - MDC/MDIO master interface for external PHY control; continuous polling mode by HW
- LED interface for port status exporting
- External memory interfaces:
 - External TM memory interface (optional):
 - DDR3 SDRAM
 - 666 MHz DDR; 8x16 bit
 - ECC protected data
 - External lookup table memory interface:
 - DDR3 SDRAM
 - 666 MHz DDR; 8x16 bit or 16x8 bit
 - ECC protected structures

- External statistics memory interface:

- RDRAM2-SIO
- 533 MHz DDR; 2x18 bit, 1 or 2 devices
- ECC protected counters

- External TCAM interface:

- Especially useful for fast lookups through large tables with wildcards, such as Access Control Lists (ACL)

OAM Offload

- KeepAlive frame generation for precise and accurate session maintenance operations
- KeepAlive watchdog timers for fastest detection time
- 802.1ag compliant message generation/termination offload
- Per OAM session state tracking and reporting
- Flexible statistics and performance monitoring

Statistics and Counters

- Up to 16M 64-bit counters via external memory
- Per-flow statistics for programmable events, traffic metering, policing and shaping
- Programmable threshold settings and threshold exceeded notification
- Dynamic allocation and auto association between counters and flows. Counters are automatically recycled when a flow is deleted or aged.
- Auto implementation of token bucket per flow (srTCM, trTCM or MEF5):
 - Hardware implementation of token bucket calculations and coloring (green, yellow, red)

Power Management

- Per interface power-up/power-down
- Configurable number of active TOP engines at each stage, for best power optimization per application

Physical Specifications

- Package: HFCBGA, 1895 pins, 45x45 mm
- Process: 55nm
- Power supply: 1.0V core voltage
- Power dissipation typical: NP-4: 35W, NP-4L: 25W

Models

- NP-4 with 100-Gigabit throughput (50G full duplex)
- NP-4L with 50-Gigabit throughput (25G full duplex)
- Both devices have same package, pin out, interfaces and are software compatible

SAMPLE APPLICATIONS

NP-4's flexibility and integration allows system vendors to deliver cost effective solutions that can easily adapt to changing market requirements. Typical applications include:

- Line cards in modular chassis:
 - Metro Switches
 - Edge and Core Routers
 - Wireless Backhaul Aggregation Switch/Routers
 - Enterprise Backbone Switches
- Stand-alone box solutions:
 - Ethernet Aggregation Nodes
 - EPON/GPON OLTs and cable CMTS
 - Firewalls, VPN and Intrusion Detection Appliances
 - Server Load Balancing Switches
 - Network Monitoring and Analysis Services

Figure 1. Interfaces Diagram

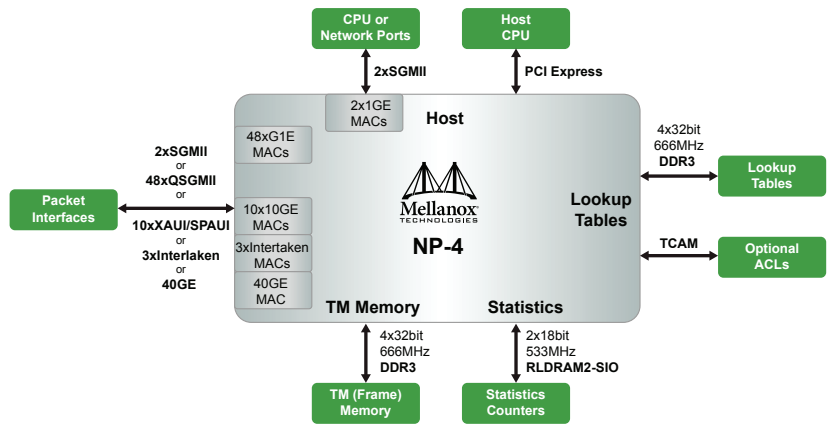


Figure 2. 20G Line Card: 24x1GE Ports Using NP-4 integrated 1G Ethernet MACs

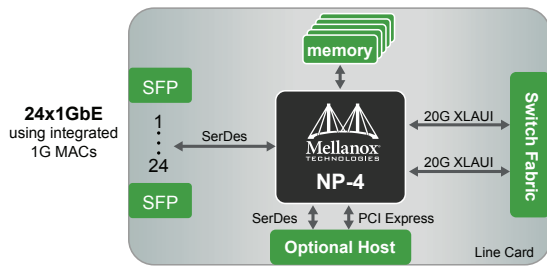


Figure 3. 100G Line Card: 100GE Pipe Application

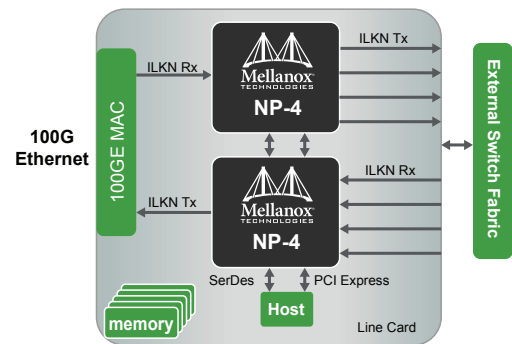


Figure 4. 40G Line Card: 4x10GE Ports

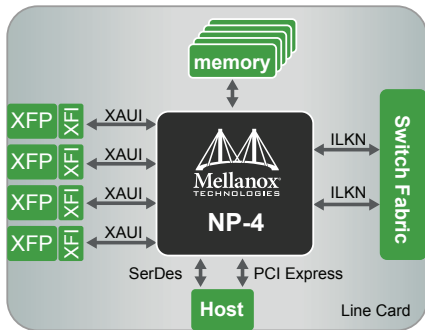


Figure 5. 24x1GE and 4x10GE Pizza Box

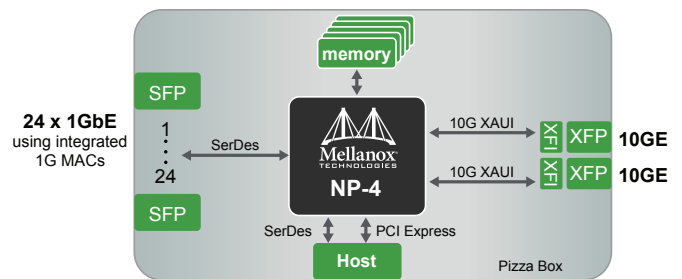


Table 1 - Part Numbers and Descriptions

OPN	Description
20779302, 20779301	NP-4 RoHS 100Gbps Network Processor
20780402, 20780401	NP-4L RoHS 50Gbps Network Processor