NP-4

## 100-Gigabit Network Processor for Carrier Ethernet Applications

Product Brief

Features

- 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 EZchip'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

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
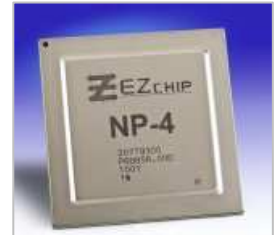
Packet Manipulation and Reassembly

- 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 and 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 EZchip's NP-2, NP-3 and NPA network processors

Interfaces ([Figure 1](#))

- 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 XAUI 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 RXAUI 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:
    - RLDRAM2-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 (i.e. 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

- The NP-4 is available in two models for diverse configurations and price points:
  - **NP-4** with 100-Gigabit throughput (50-Gigabit full duplex)
  - **NP-4L** with 50-Gigabit throughput (25-Gigabit full duplex)
- Both devices have same package, pin out, interfaces and are s/w compatible

Ordering Information

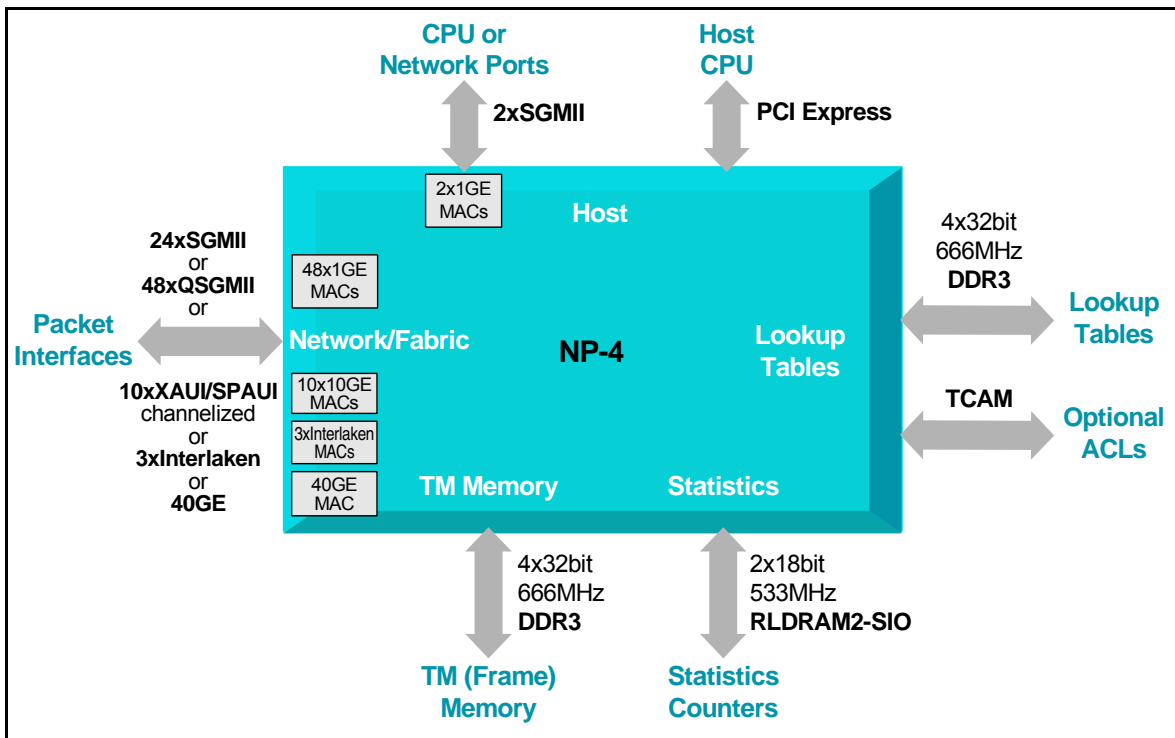| Device	Part No.	Device	Part No.
NP-4 RoHS	207793x1	NP-4L RoHS	207804x1

Interfaces Diagram

Figure 1. NP-4 interfaces

NPU with Integrated TM & CPUs

EZchip'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, and 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
- Traffic manager
- OAM processing offload
- On-chip Control CPU
- On-chip Quality of Service CPU
- Fabric Interface Chip (FIC) functionality
- Integrated MACs: 48 1-Gigabit, ten 10/20-Gigabit Ethernet MACs, one 40-Gigabit MAC, and three Interlaken MACs.

NP-4 provides exceptionally **flexible packet processing** enabling system designers to future proof their designs to support new protocols and features through s/w 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.

Main Functional Blocks

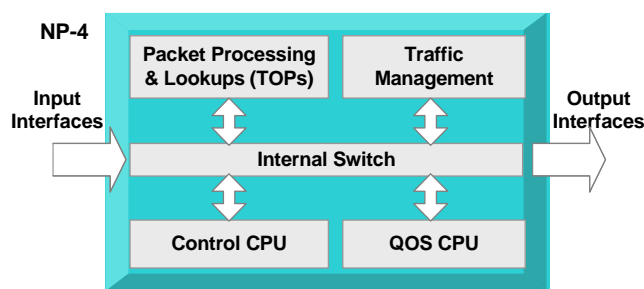These translate into the NP-4's main functional blocks:

Figure 2. Five main functional blocks

- Task Optimized Processors (TOPs) – fully programmable for packet processing and lookups
- Traffic Manager (TM) – configurable for advanced flow-based bandwidth control
- Control CPU
- Internal Switch – determines the data flow between the device's external interfaces (input and output) as well as the other main functional blocks

Task Optimized Processors (TOPs)

EZchip's innovative **TOPcore® technology** enables the NP-4 to deliver its exceptionally high performance. TOPcore technology integrates many high-speed processors; each optimized to perform a specific task. Four types of TOPs (Task Optimized Processors) – parse, search, resolve and modify – are employed to perform the main tasks of packet processing, i.e. classification, forwarding and modification. A programmable TOP corresponds to each of these tasks, and performs its respective task exceptionally fast.

Each TOP processor type employs a unique architecture with a customized, function-specific data path and instruction set. This minimizes the number of clock cycles required for complex packet manipulation and provides exceptionally fast packet processing. TOP performance is boosted by a super-scalar architecture in which multiple instances of the TOPs operate in parallel within each pipeline stage.

The NP-4 TOPs are fully programmable for a variety of applications and use a simple **single-image programming model** with no parallel programming or multi-threading. Allocation of the TOPs processing engines to incoming frames, passing messages between the TOPs as well as maintaining the ordering of frames is completely transparent to the programmer and performed in hardware. Large code space memory is available to support multiple and complex applications while providing headroom for adding new features. Full support for hitless code updates is provided to reduce system downtime for maintenance.

NP-4 features **integrated search engines** that perform lookups for implementing diverse applications in layer 2-4 switching/routing and layer 5-7 deep packet processing. These search engines deliver programmable lookups in a combination of tables with millions of entries per table. NP-4 stores its **lookup tables in DRAM** to reduce power dissipation and cost while supporting large tables and providing extensive classification headroom.

NP-4's leading-edge use of **embedded memory technology** provides aggregate bandwidth of hundreds of Gigabits per second, mandatory to sustain high throughput. Multiple embedded memory cores are utilized for queuing frame buffers that are being processed, while the rest of the frames are stored in external memory.

Lookup tables may be stored in the internal and/or in external memory for very large tables, via a fast DDR interface.

NP-4 supports three main types of **lookup tables** – direct access tables, hash tables and trees – each is flexibly defined and used for various applications. Tables may be used for forwarding and routing, flow classification, access control, etc. Numerous tables of each type can be defined, stored in internal memory and/or external memory and searched through per each packet. For maximum flexibility, the key size, result (i.e. associated data) size, and number of entries are all user-programmed per table. Longest prefix match and wildcards are supported in trees.

Patented **search algorithms** enable high-speed lookups in trees and hash tables, which are stored in the internal memory or external memory. All search algorithms for hash and trees are implemented in hardware for maximum performance and simplicity. Hash lookup performance is deterministic regardless of the hash table size. Trees, as well as hash tables, with millions of entries performing longest prefix match can be searched at sustained wire speed. Search keys can be up very long to support long table entries such as IPv6 5-tuple flows. Up to 96 bytes of result (associated information) can be stored per each entry and retrieved upon a lookup match.

Stateful auto learning and updating of table entries and the session's state is performed entirely by NP-4's TOPs with no intervention required by the host CPU. This provides for an extremely high rate of millions per second, for the addition of new flows or deletion of old flows to/from the flow table. Result information, statistics counters, and per-flow rate limiters can be automatically associated with new flows and recycled when deleting old flows. Session state can be updated per flow on-the-fly, and new packets can be generated to implement various stateful functions, such as TCP session 3-way handshake init/termination and dynamic TCP port tracking.

An optional interface is available to an **external TCAM** device(s). The TCAM option is especially useful for performing fast lookups through large tables with multiple wildcards such as Access Control Lists (ACL). TCAM lookups are performed in parallel to lookups performed by the integrated TOPsearch engines of NP-4.

Control CPU

The on-chip **Control CPU** provides extended flexibility for various applications such as OAM, statistics and performance monitoring, network management offload, interrupt monitoring and more. The on-chip Control CPU offloads the host CPU and may function as a smart Host Master for data transfers and status reporting.

System-wide protection and fast restoration features are enhanced utilizing the capabilities of the on-chip Control CPU for monitoring and reprogramming the NP-4 states. Communication channels established between several NP-4 on-chip CPU's on different line cards enable fast peer to peer synchronization and table upload in case of failures.

QOS CPU

The on-chip **QOS CPU** monitors the on-chip and system-wide Traffic Managers, providing end to end dynamic traffic control mechanisms. The QOS CPU is similar to the Control CPU, but includes hardware accelerations for accessing and reprogramming of the TM, which enables the execution of an optimized traffic engineering application.

The OQS CPU provides system-wide traffic management which enables use of standard Ethernet switches as backplane switches, while providing the required QoS features on top of the Ethernet switch features. The flexibility of the QOS CPU enables support for various congestion control protocols, such as IEEE802.1au Backwards Congestion Notification Messages.

Traffic Manager

NP-4 offers extensive **traffic management** capabilities on the ingress and egress paths. This enables frame queuing and traffic management for traffic on all NP-4 interfaces. Traffic transmitted to the network links and the system switching fabric as well to the host CPU interface can be assigned with specific QoS settings.

NP-4 supports DiffServ and IntServ services and a wide variety of QoS mechanisms, complying with the Metro Ethernet Services definitions as well as the DSL and PON networks QoS requirements. These include:

- Classification – assigning frames to specific flows with applicable QoS parameters.
- Metering – measuring per-flow traffic and determining compliance with traffic parameters. Two Rate Three Color Metering (trTCM) is used, compliant with the Metro Ethernet Forum specifications.
- Marking – individual frame's compliance.
- Congestion Avoidance – profile-based WRED early packet discards based on priority, available memory and metering results.

- Traffic Conditioning – enforcing rules on traffic flows, e.g. policing in which packet dropping is applied to non-compliant packets, or shaping which schedules a flow to conform to its assigned parameters (no packet dropping). Single and dual leaky bucket shaping is applied.
- Congestion Management – hierarchical scheduling of flows to the various interfaces using a priority scheme and Weighted Fair Queuing (WFQ).

IP Reassembly and Enhanced Video Support

NP-4 features extensive packet manipulation of the buffered data. This is enabled through special TOP instructions that control the TM to enable a variety of applications that require data reordering and reassembly, typically used in applications such as IP reassembly.

In addition, the TOPs control caching and retransmission of video streams and are video data-aware. This can be employed in line cards or dedicated video cards to enable a variety of video and IPTV related applications, including:

- Retransmission of lost video packets
- Seamless and fast IPTV channel zapping
- Video de-multiplexing of Multiple Program Transport Streams (MPTS) to Single Program Transport Stream (SPTS)
- Video stream path redundancy and Head-End redundancy by switching to alternative IPTV stream
- Switching to personalized commercial video clip

Internal Switch

The **Internal Switch** serves as a matrix between the chip's external interfaces (input and output) as well as the four main NP-4 functional blocks (TOPs, TM, Control CPU and QOS CPU). It determines the data flow of the frame through the network processor. Different types of frames may take different paths depending upon the level of processing required.

- Any queue can send to any destination
- Inherent loopback/bypass flows through the TM
- Flexible resource sharing between TM resources and interfaces
- Stream storage and control
- Flexible LAG shaping
- Redundancy support
- Line card scaling

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

Illustrated below are several sample solutions.

Line Card Solutions

20G Line Cards
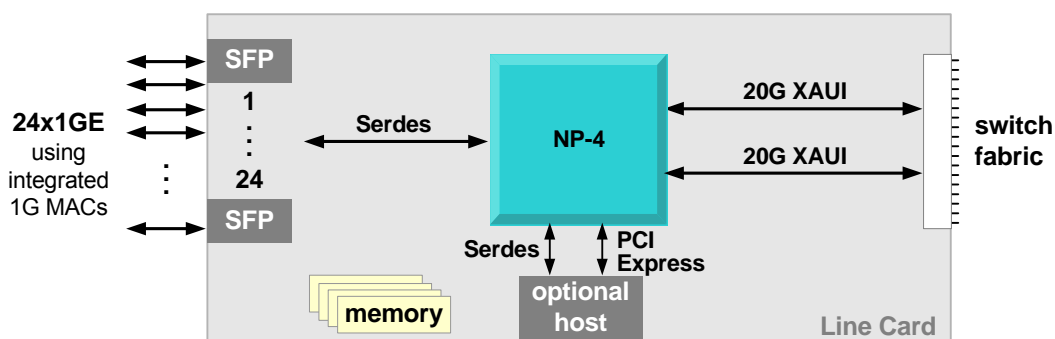
Figure 3. 24x1GE ports using NP-4 integrated 1G Ethernet MACs
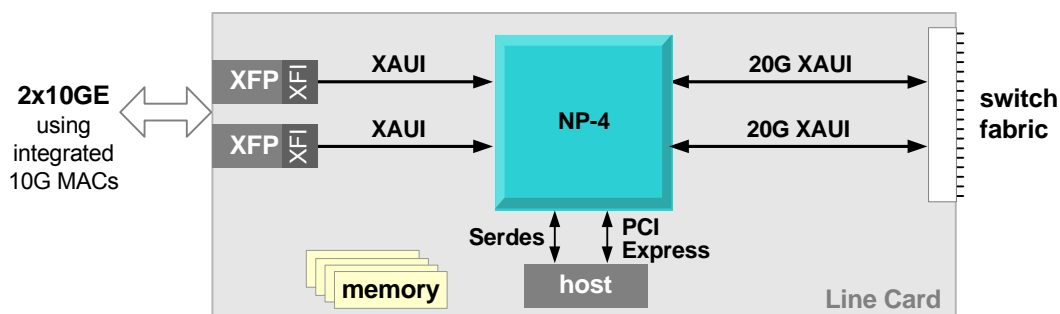
Figure 4. 2x10GE ports

40G Line Cards
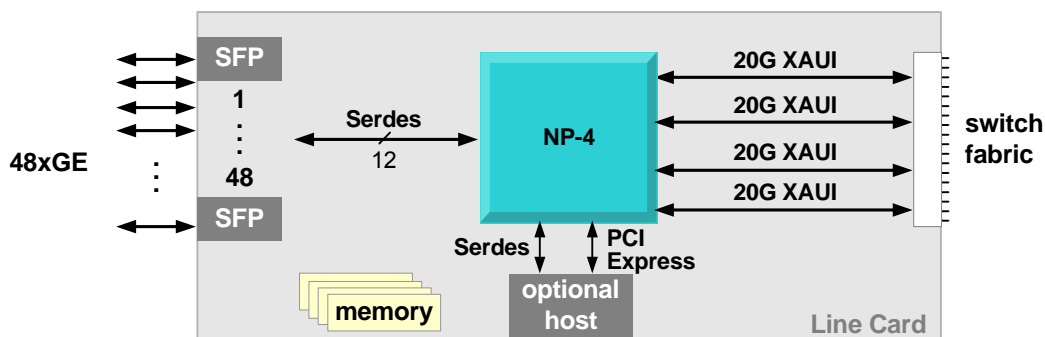
Figure 5. 48 GE ports using QGMII interfaces
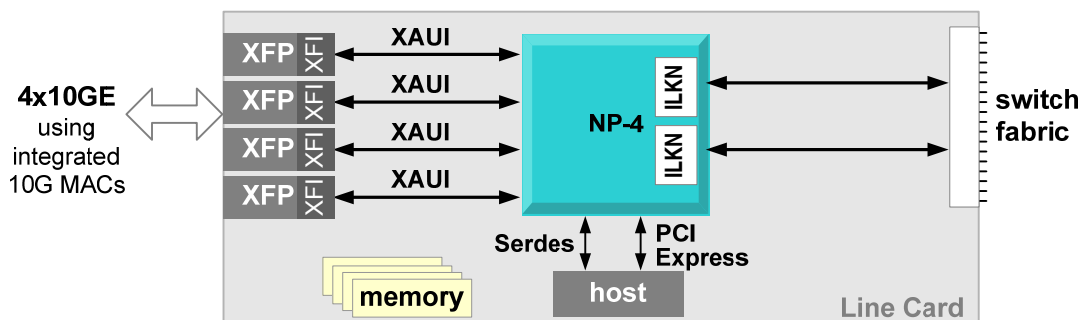
Figure 6. 4x10GE ports
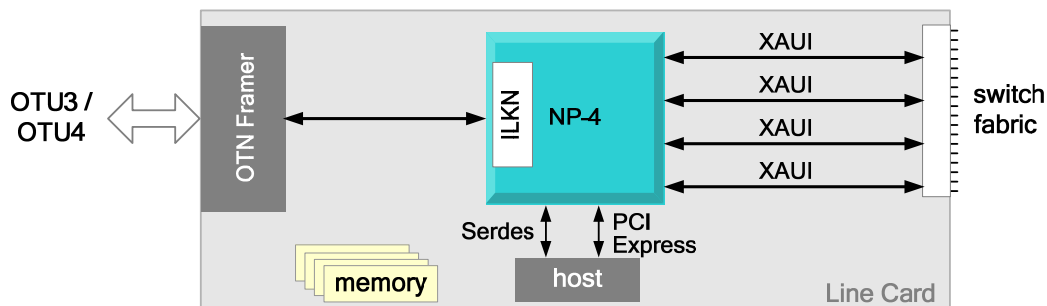
Figure 7. OTN application

100G Line Cards
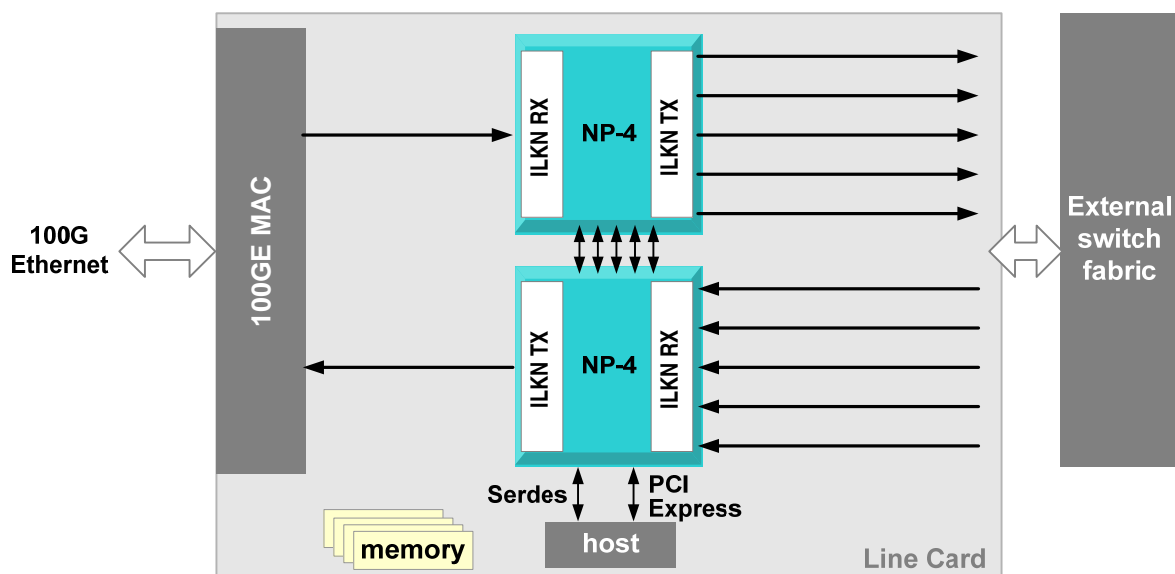
Figure 8. 100GE pipe application
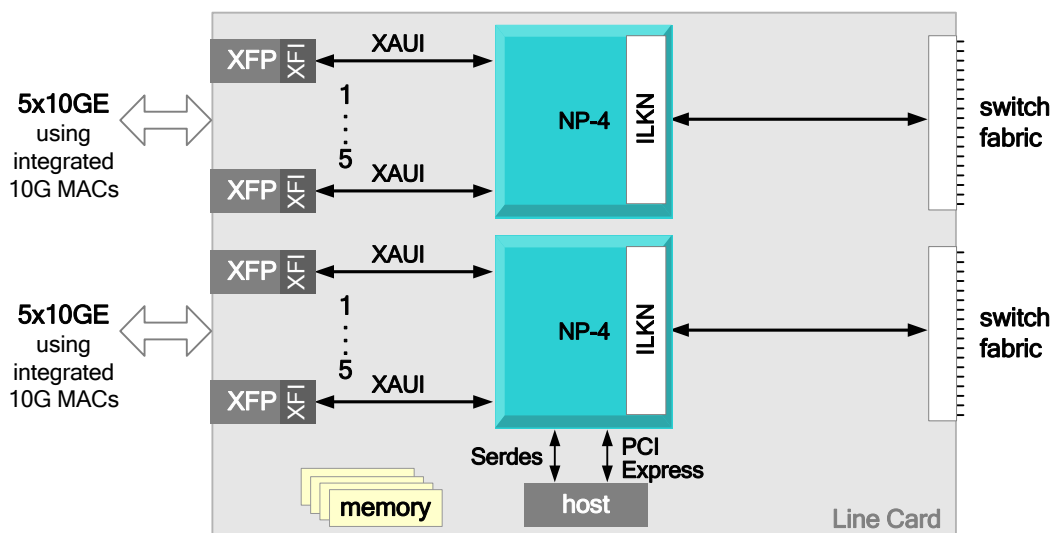
Figure 9. 10x10GE application

Stand-alone Solutions ("Pizza" Boxes)
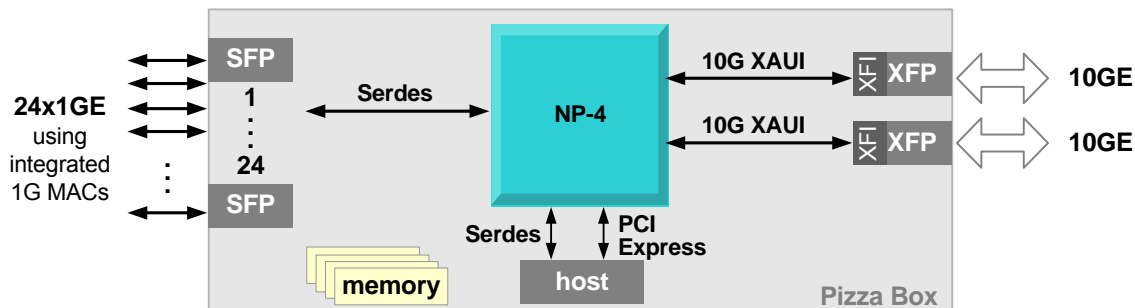
The following examples show pizza box applications.

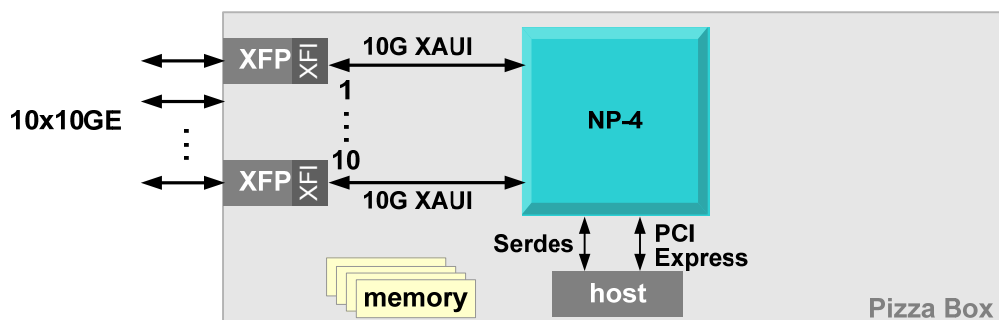

Figure 10. 24x1GE and 4x10GE pizza box
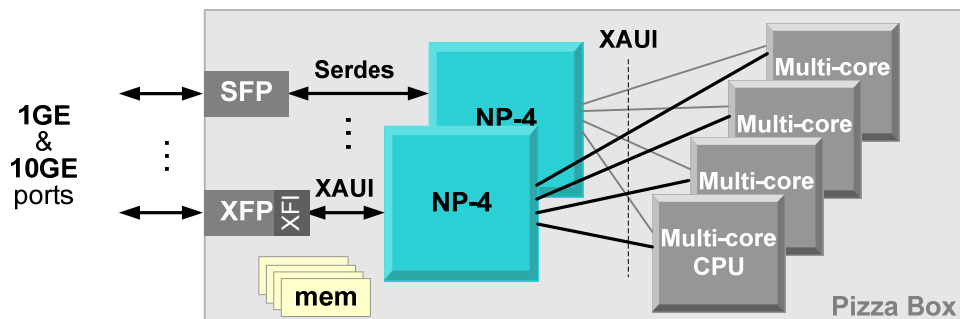

Figure 11. 10 x 10GE pizza box

Figure 12. Layer 4-7 board for DPI processing

EZdesign Software Toolset and Libraries

EZdesign is a comprehensive set of design and testing software tools for developers, enabling rapid delivery to production of new designs based on EZchip's network processors. EZdesign allows designers to create, verify and implement NP-4 applications to meet specific functionality and performance targets.

EZdesign components include:

- **Microcode Development Environment:** A unified GUI for editing and debugging code, including setting breakpoints, single-stepping program execution and access to internal resources. Features include a code editor, view of memory and register contents, performance charting, macro recording and script execution. The EZdesign MDE is used in development and debugging of code on both the simulator and the actual network processor.
- **Simulator:** Provides cycle accurate simulation of the EZchip network processor for code functionality testing and performance optimization.
- **Assembler and Preprocessor:** Generating optimized code for execution on EZchip's network processors. The assembly is interleaved with high-level macros.
- **Applications Library:** Sample code implementing high-level applications for reference when designing new networking platforms and services. Sample code is available for L2 switching, Metro Ethernet switch, MPLS LER and SER, VPLS and draft Martini, IPv4 and IPv6 routing, EPON/GPON OLT, Network Address Translation (NAT), Access Control Lists (ACL), firewall and more.

- **Frame Generator:** A GUI guiding the programmer through the process of creating frames, layer by layer. Allows for generation of frames of different types, protocols and user-defined fields.
- **EZconfig:** A GUI enabling configuration of the EZchip network processor and definition of data structures used by the network processor for forwarding and policy table lookups (e.g. hash, trees), their keys and associated result information.
- **Traffic Manager Configurator:** A high level GUI enabling you to configure both of the traffic managers embedded in the NP-4 network processor. This includes configuration of the hierarchical topology and QoS parameters for each level. A sanity check and creation of NPsl C source file are provided.

EZdriver Control Processor API Layer

EZdriver SDK is a toolset that facilitates the development of the control path software for EZchip network processor based systems. It enables applications that run on the control CPU to communicate with the EZchip network processor. EZdriver consists of routines that execute on the control CPU and provide an API for interfacing the network processor. It includes the chip configuration, microcode loading, creation and maintenance of lookup structures, sending and receiving frames to and from the network processor, as well as configuration and access to the statistics block.

EZdriver SDK along with EZdesign provide extensive debugging capabilities, and enable software-driven debugging features (e.g. breakpoints, single step, register and memory access) to be performed on both the simulator and the actual network processor.

About EZchip

EZchip Technologies is a fabless semiconductor company that provides Ethernet network processors. EZchip provides its customers with solutions that scale from 1-Gigabit to 200-Gigabits per second with a common architecture and software across all products. EZchip's network processors provide the flexibility and integration that enable triple-play data, voice and video services in systems that make up the new Carrier Ethernet networks. Flexibility and integration make EZchip's solutions ideal for building systems for a wide range of applications in telecom networks, enterprise backbones and data centers. Visit our web site at www.ezchip.com.

Email: ezsupport@ezchip.com ● Web: www.ezchip.com