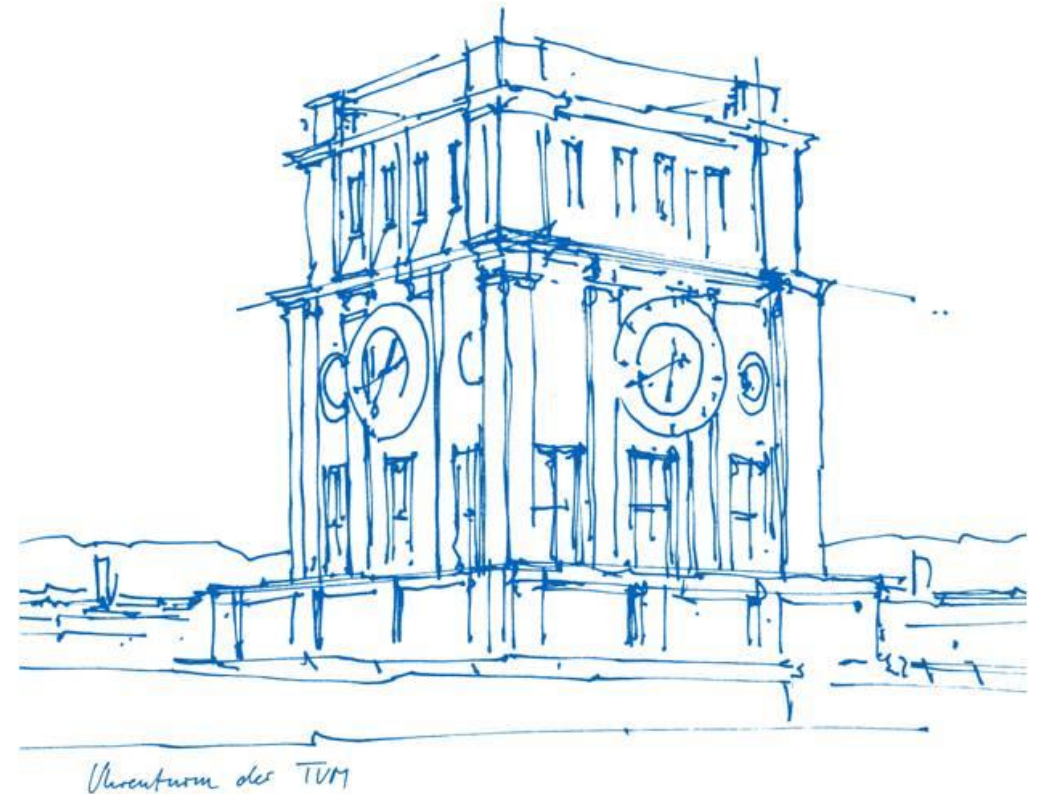# NIC Architectures and their Role in High-Speed Networking

Prof. Dr. Andreas Herkersdorf

School of Computation, Information and Technology

Technical University of Munich

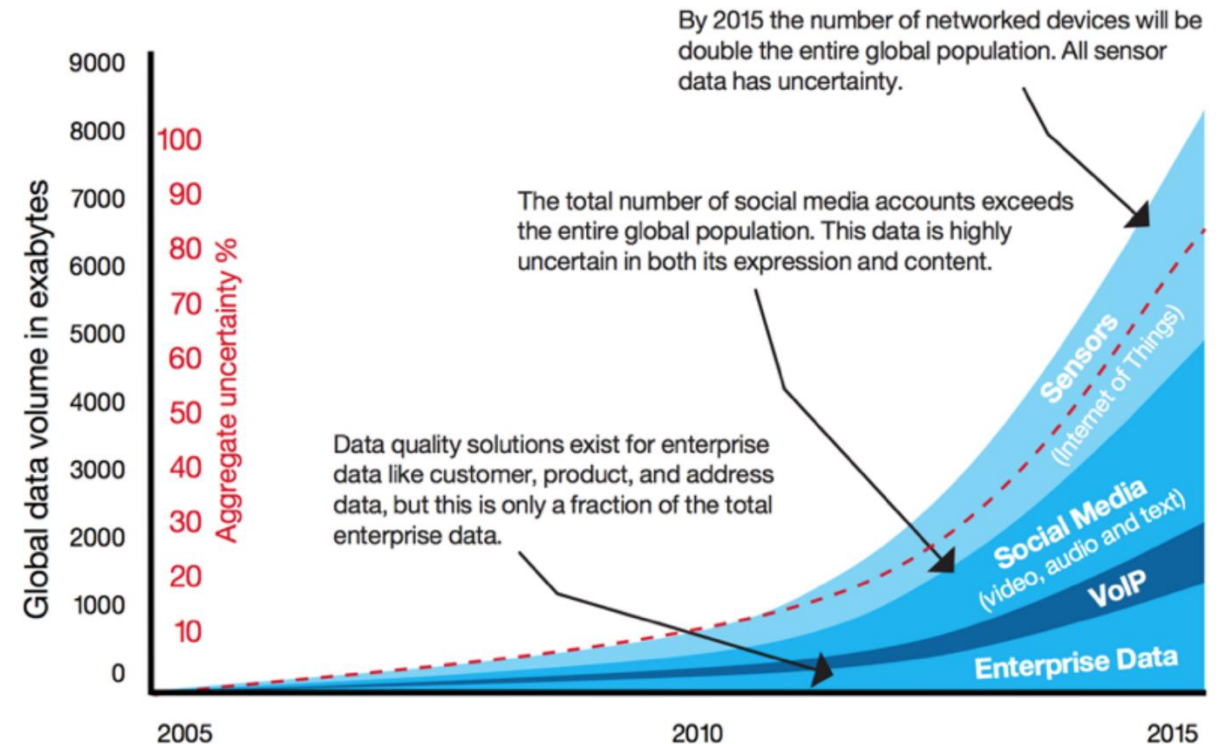Academic Salon on High-Performance Ethernet

March 12, 2025

Big Data in Motion

Feb 27, 2010

IBM estimates that 90% of all the world's data has been created in the last two years.

90%

2019

By 2015 the number of networked devices will be double the entire global population. All sensor data has uncertainty.

The total number of social media accounts exceeds the entire global population. This data is highly uncertain in both its expression and content.

Data quality solutions exist for enterprise data like customer, product, and address data, but this is only a fraction of the total enterprise data.

Sensors (Internet of Things)

Social Media (video, audio and text)

VoIP

Enterprise Data

Global data volume in exabytes: 0, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000

Aggregate uncertainty %: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100

2005   2010   2015

„The Promise and Prejudice of Big Data in Intelligence Community", K. Jani, Georgia Institute of Technology, October 26, 2016
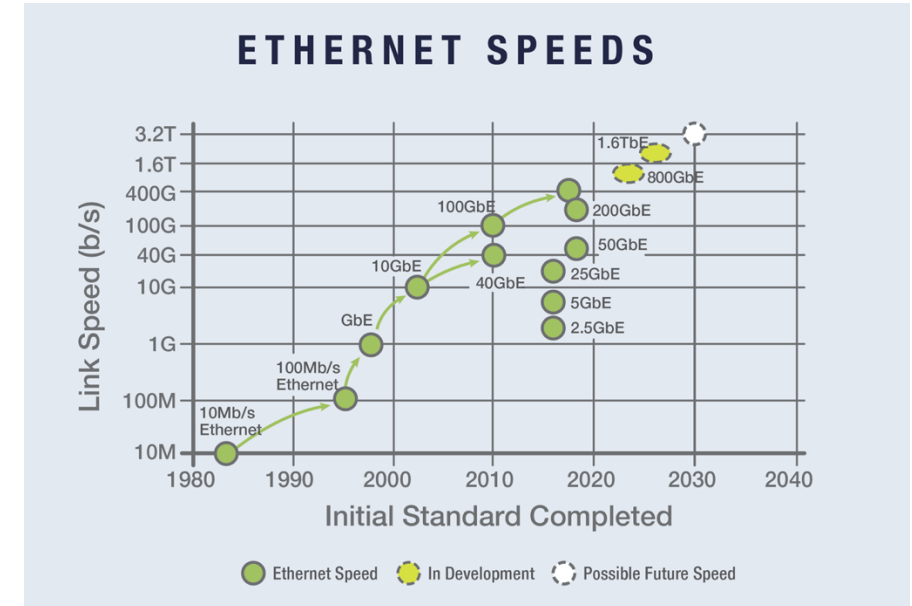
2

Challenges for high-speed Ethernet

The required processing capacity of a networking device depends on:

Packet rate / Packet interarrival time

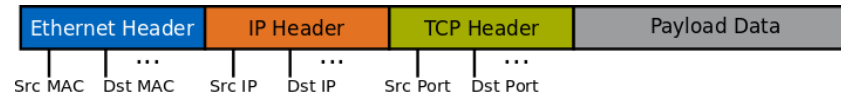→ *packets per second (PPS)* $PPS = \dfrac{link\_speed}{packet\_size}$

link_speed	40 Gbps		100 Gbps		800 Gbps	
packet_size	64B	512B	64B	512B	64B	512B
PPS	78M	9.8M	195M	24M	1.563G	195M
1 / PPS	12.8ns	102ns	5,1ns	41ns	640ps	5,1ns

Challenges for high-speed Ethernet

Header Processing

Ethernet Header	IP Header	TCP Header	Payload Data

Src MAC Dst MAC   Src IP   Dst IP   Src Port   Dst Port
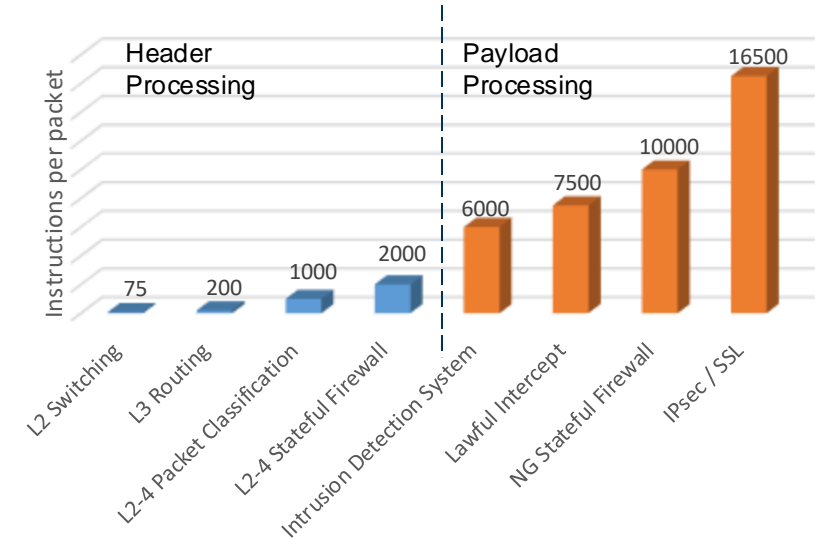
- Fixed header location allows simple parsing

- **Common use-cases:** routing, switching, IP/port-based firewalls, …

Payload Processing

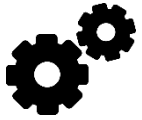| Eth.	IP	TCP	Payload	You Tube	…	Eth.	IP	TCP	Payload
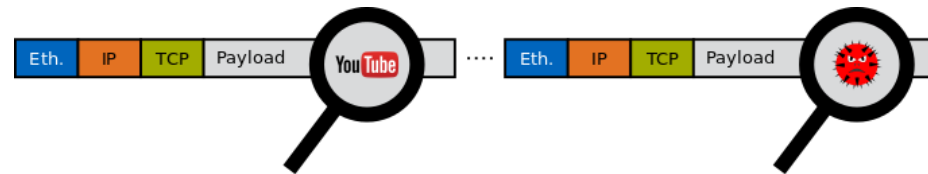
- **Common use-cases:** application/session/user identification for firewalls or bandwidth throttling, cryptology, intrusion detection, virus scanning, …

Instructions per packet

Header Processing | Payload Processing

L2 Switching	L3 Routing	L2-4 Packet Classification	L2-4 Stateful Firewall	Intrusion Detection System	Lawful Intercept	NG Stateful Firewall	IPsec / SSL
75	200	1000	2000	6000	7500	10000	16500

Source: Netronome

Challenges for high-speed Ethernet

The processing complexity of a networking function:
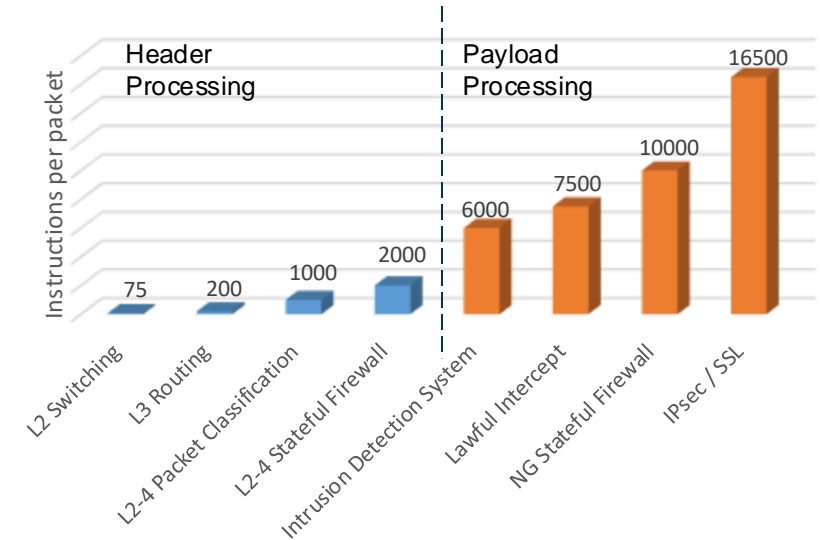
→ *instructions per packet (IPP)*

Required processing capacity = $PPS * IPP \left[\frac{instructions}{second}\right]$

Source: Netronome

@ 400Gbps

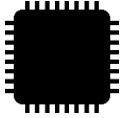L2 Switching (64B): $PPS * IPP = 781.250.000 \frac{packets}{second} * 75 \frac{instructions}{packet} = 58.6 * 10^9 \frac{instructions}{second}$

Intrusion Detection (512B): $588 * 10^9 \frac{instructions}{second}$ IPSec (512B): $1.6 * 10^{12} \frac{instructions}{second}$

Challenges for high-speed Ethernet

Intel Xeon E5-2660v3: 10 cores @ 2.6 GHz, 105 W TDP

$$IPS = 10\,cores * f_{clk} * IPC_{core} \approx 26 * 10^9\,\frac{instructions}{second}$$
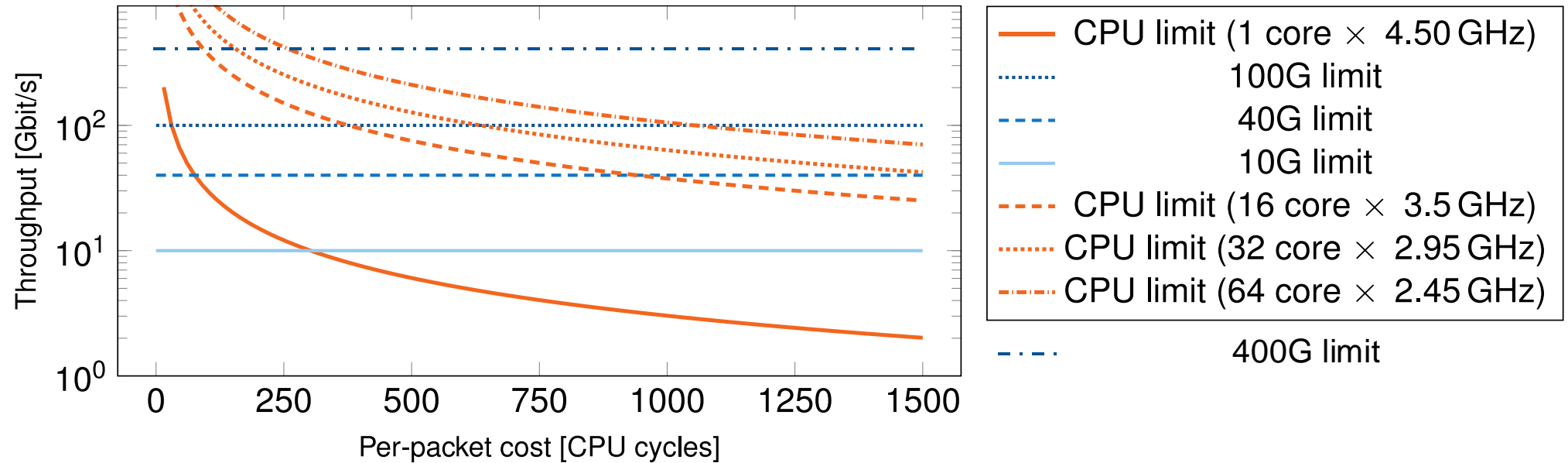
$$IPC_{core} \approx 1$$

Networking Function	400Gbps
L2 Switching	$58.6 * 10^9$
Intrusion Detection	$588 * 10^9$
IPSec	$1.6 * 10^{12}$

L2 Switching ✗   Intrusion Detection ✗   IPSec ✗

IPSec @ 400Gbps would require 70 CPUs with a TDP of 7 kW!

Challenges for high-speed Ethernet

- Multi- and manycore architectures help to achieve higher throughputs
- … but complexity of network services grows faster than processor performance

Ref.: S. Gallenmüller, P. Emmerich, F. Wohlfart, D. Raumer, and G. Carle. „Comparison of Frameworks for High-Performance Packet IO". In: ANCS '15

How to cope with these challenges?

Performance and Energy Efficiency

- cope with very high data rates (up to hundreds of Gbps)

- lowest packet delay as possible

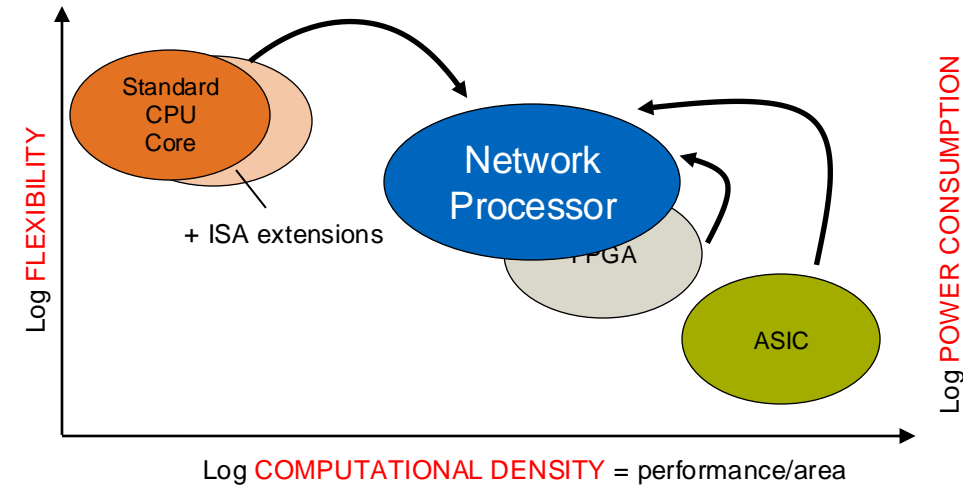- low power consumption → Customized **ASIC**

Flexibility

- adapt to evolving packet processing applications

- efficient resource sharing among network applications
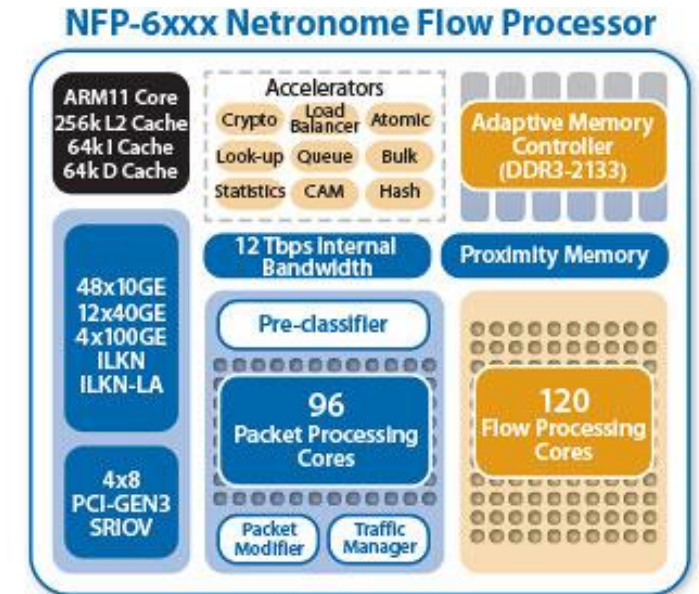
→ Programmable **CPU / ASIP**

Source: Blume et al., „Model-based exploration of the Design Space for Heterogeneous System on Chip", 2002

State-of-the-art Network Processors

Netronome NFP-6xxx Flow Processor

- 216 programmable cores to execute software
  - 96 *packet processing cores* for stateless processing
  - 120 *flow processing cores* for stateful processing

- More than $300 * 10^9 \frac{instructions}{second}$

- 100 hardware accelerators for
  - DPI, regular expression matching
  - Cryptography
  - Hash calculation
  - Packet I/O, Queue Management
  - …

- 50 Gbps bulk cryptography

- 720 Gbps I/O

Source: Netronome

State-of-the-art Network Processors

Netronome NFP-6xxx Flow Processor

- 216 programmable cores to execute software
  - 96 *packet processing cores* for stateless processing
  - 120 *flow processing cores* for stateful processing

- More than $300 * 10^9 \frac{\text{instructions}}{\text{second}}$

- 100 hardware accelerators for
  - DPI, regular expression matching
  - Cryptography
  - Hash calculation
  - Packet I/O, Queue Management
  - …

- 50 Gbps bulk cryptography

- 720 Gbps I/O
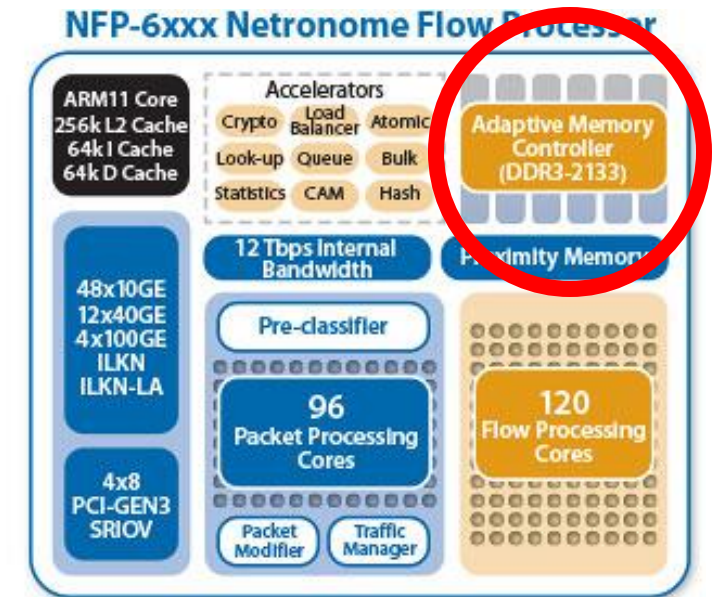
Source: Netronome

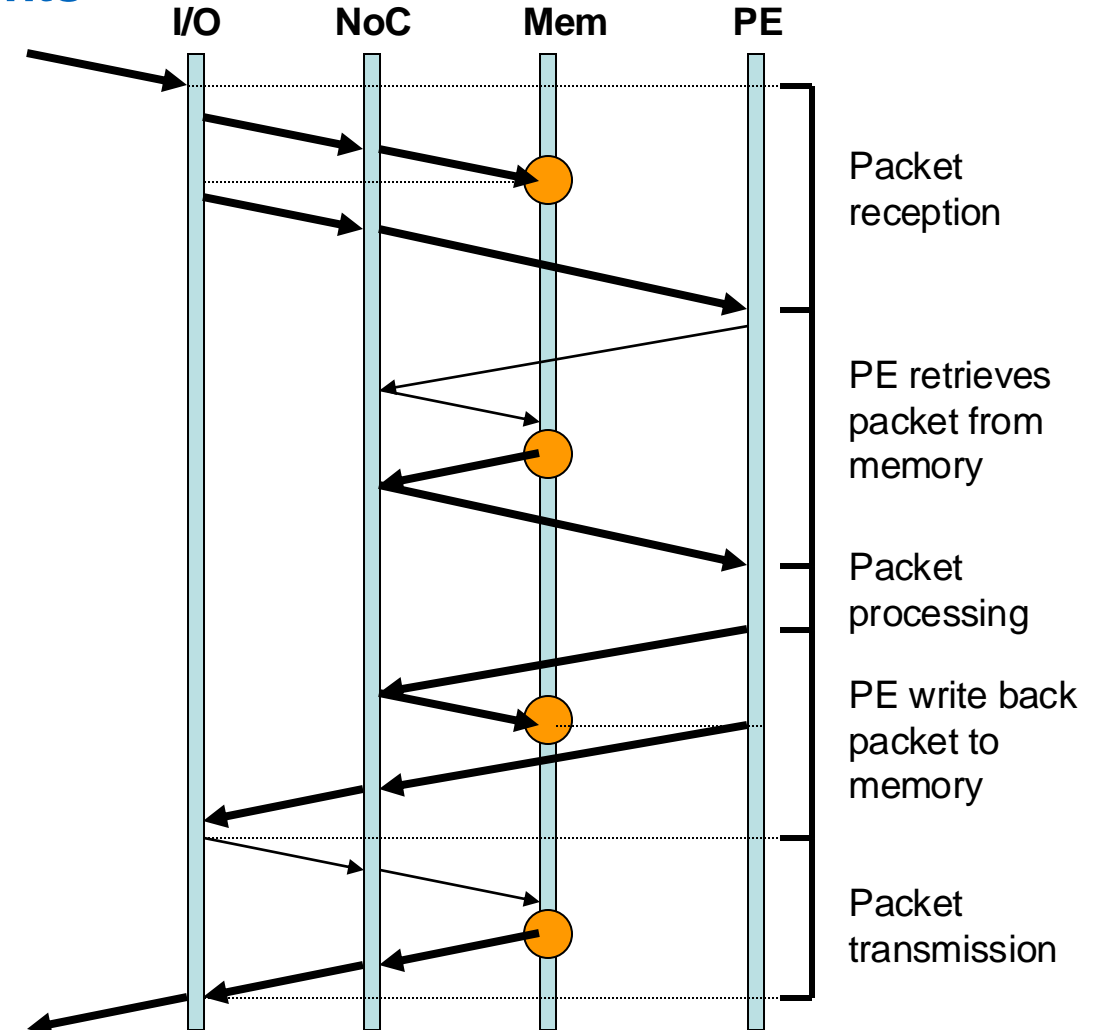| Memory	T_{acc} [ns]	$f_{I/O}$ [MHz]	Data width	BW_{max} [GByte/s]
DDR-400	28 - 35	200	64 bit	3.2
DDR3-2133	21 - 26	1066	64 bit	17.0
DDR4-3200	25 - 30	1600	64 bit	25.6

Network Processing Memory Bandwidth Requirements

IP packets …

- … traverse the memory interface at least 4 times! 🟠

- **Exceeds the peak streaming data rate of DDRx!**

$$BW_{store} \geq 4 \cdot 400 Gbps = 1600 \ Gbps = \mathbf{200 GByte/s}$$

Memory	T_{acc} [ns]	$f_{I/O}$ [MHz]	Data width	BW_{max} [GByte/s]
DDR-400	28 - 35	200	64 bit	3.2
DDR3-2133	21 - 26	1066	64 bit	17.0
DDR4-3200	25 - 30	1600	64 bit	25.6
HBM 3		3200	1024	820

I/O   NoC   Mem   PE

Packet reception

PE retrieves packet from memory

Packet processing

PE write back packet to memory

Packet transmission

Multi-Purpose SmartNICs

Offload compute node resources for ever increasing networking demands:

- Network and Node Resilience

  - Low-latency network coding / FEC

    - DFG SPP 2378 "Resilience in Connected Worlds" (joint project with G. Carle)

  - Reflex-based traffic steering

- Energy Efficiency & Power Management

  - ecoNIC-based workload pinning
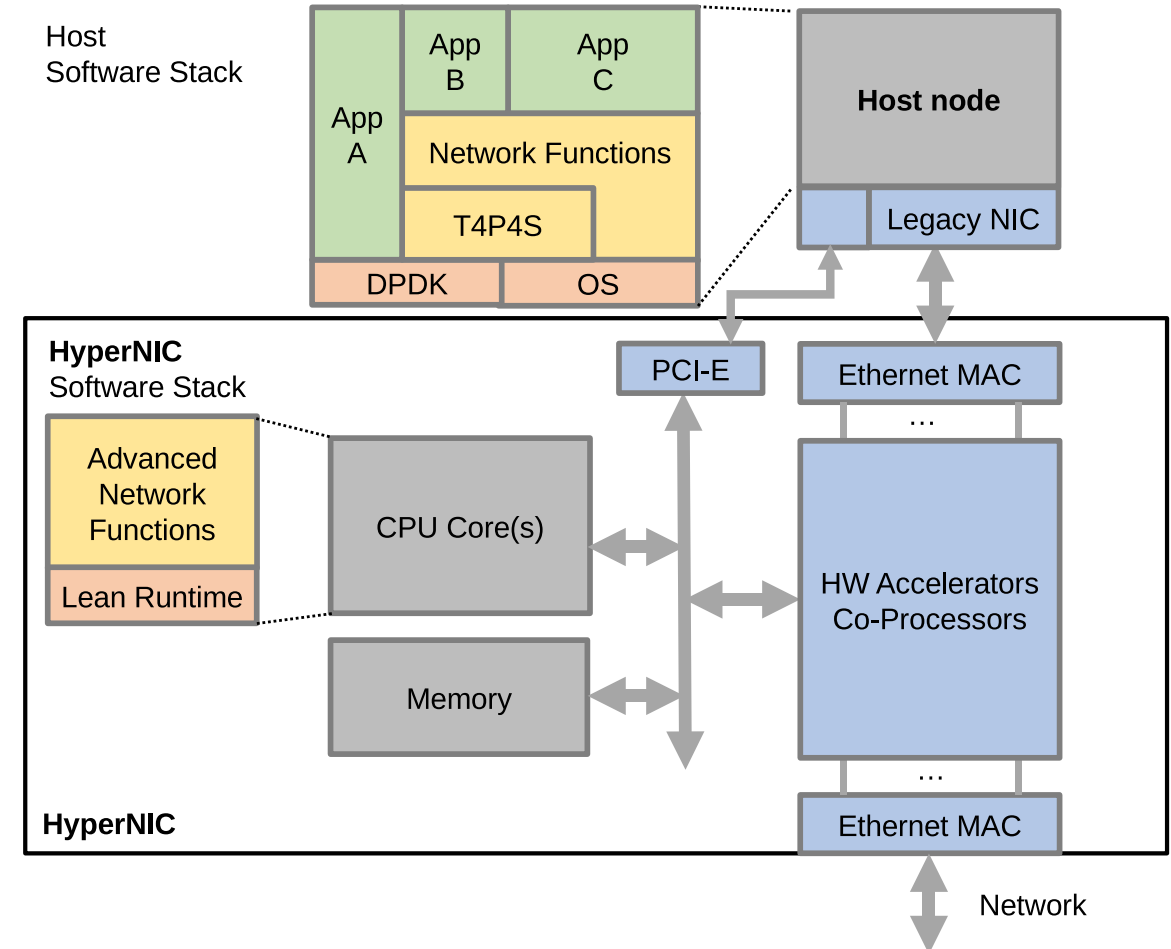
ecoNIC

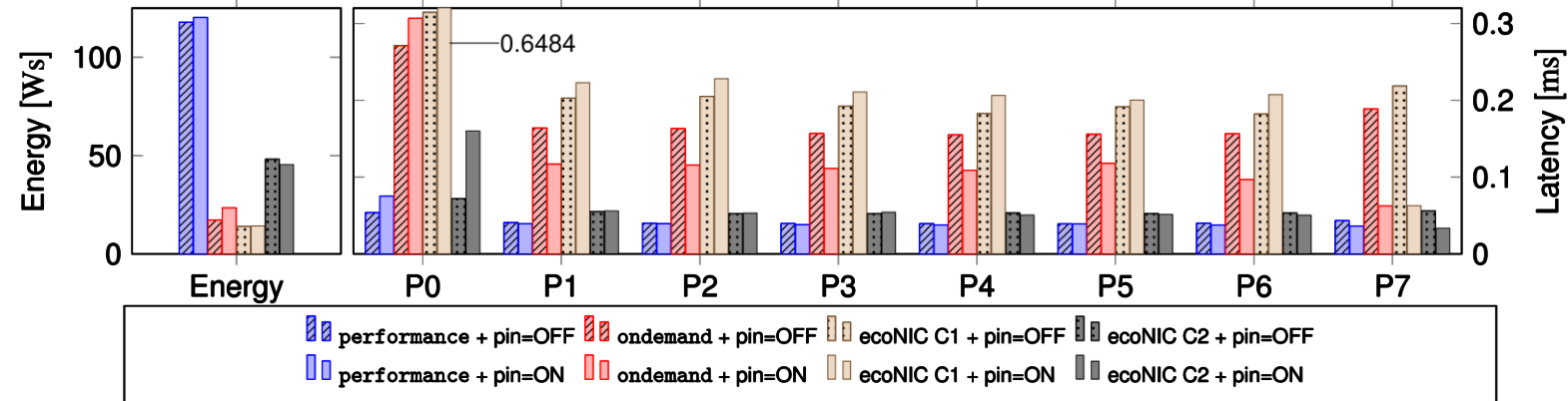- Combines workload power management with priority-based traffic steering / pinning
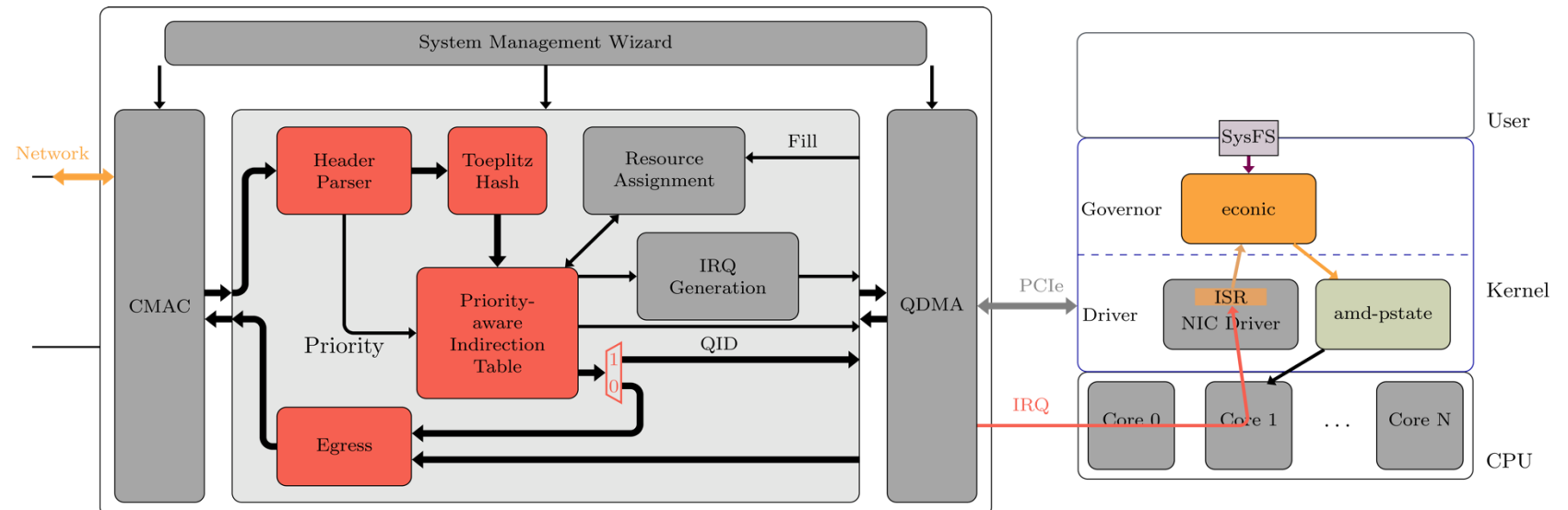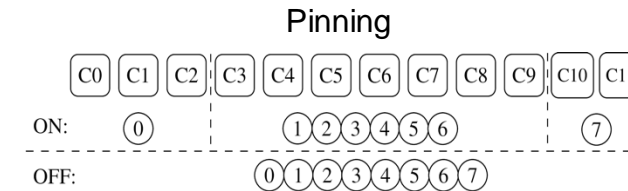
F. Biersack, M. Liess, M. Absmann, F. Lotter, T. Wild, A. Herkersdorf, "ecoNIC: Saving Energy through SmartNIC-based Load Balancing of Mixed-Critical Ethernet Traffic", 27th Euromicro Conference on Digital System Design (DSD), 2024.

Networking Testbed @ LIS
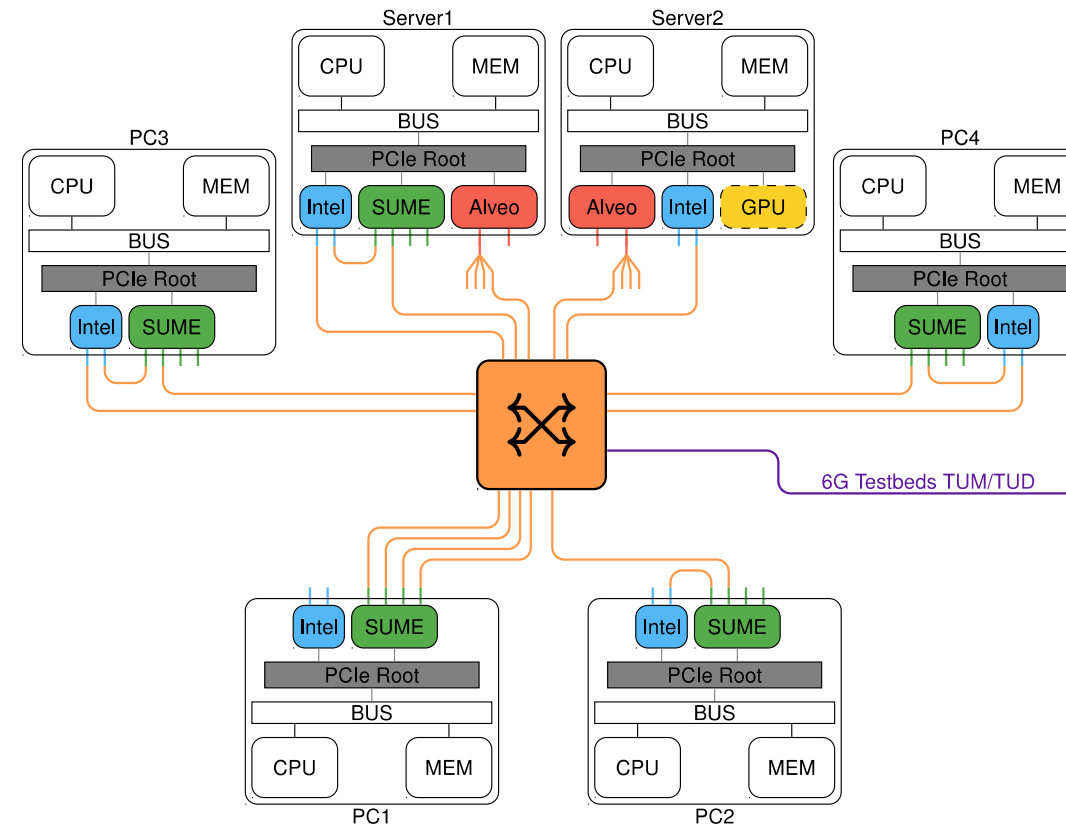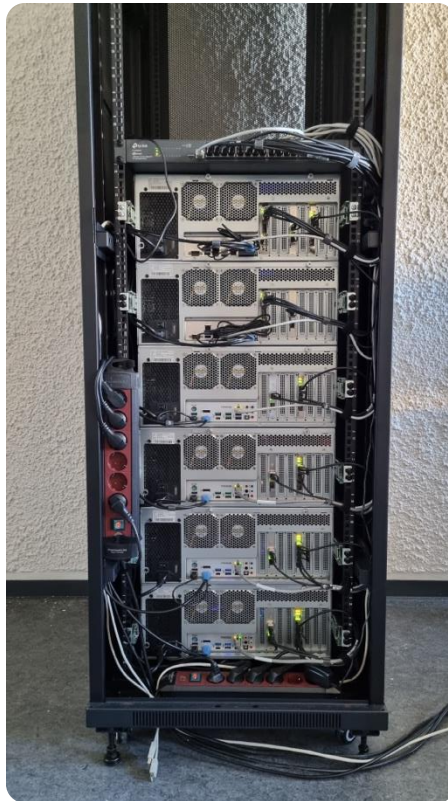
ecoNIC

- Comparison relative to Linux Power Governors (performance, ondemand)
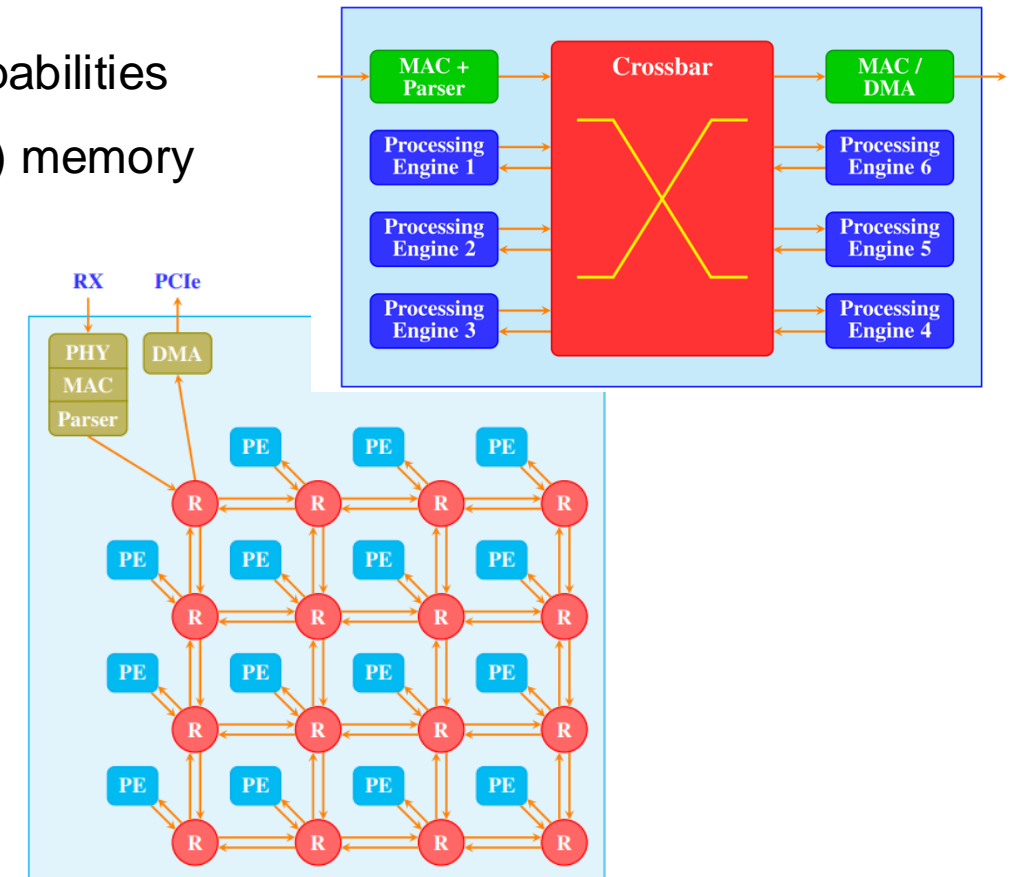- C1/C2: different parameter settings for switching between power states

Crossbar- / NoC-based SmartNIC Interconnect

High data rate SmartNICs demand server-like compute capabilities

- Mitigated by HW offloads (PEs) and pipelined (SRAM) memory buffers

FlexRoute / FlexCross / HiPerNoC

- Parse / Classify / Map packets to a PE traversal route

- Virtual cut-through (VCT) crossbar / router with two stage pipeline

- 512 bit data path; 102Gb/s

K. Zyla, M. Liess, T. Wild, A. Herkersdorf, „FlexCross: High-Speed and Flexible Packet Processing via a Crosspoint-Queued Crossbar", Euromicro DSD, 2024.
K. Zyla, M. Liess, T. Wild, A. Herkersdorf, "HiPerNoC: A High-Performance Network-on-Chip for Flexible and Scalable FPGA-Based SmartNICs", IEEE DATE, 2025.

Smart Network Interface for Predictable Services

Hardware and software for a new generation of architectures aiming for reliable, high data rate, low-latency, cut-through forwarding and processing

- **LML** Load Management Layer to balance resource utilization across available compute nodes
  → lower application latency and quick failure recovery
- **NHM** Network Health Monitoring for high precision flow-based traffic measurement,
  → quick detection of suboptimal network conditions
- **LML+NHM** incorporate inferred network state in load balancing decision
  → contribution to functional safety, overload mitigation and resource savings
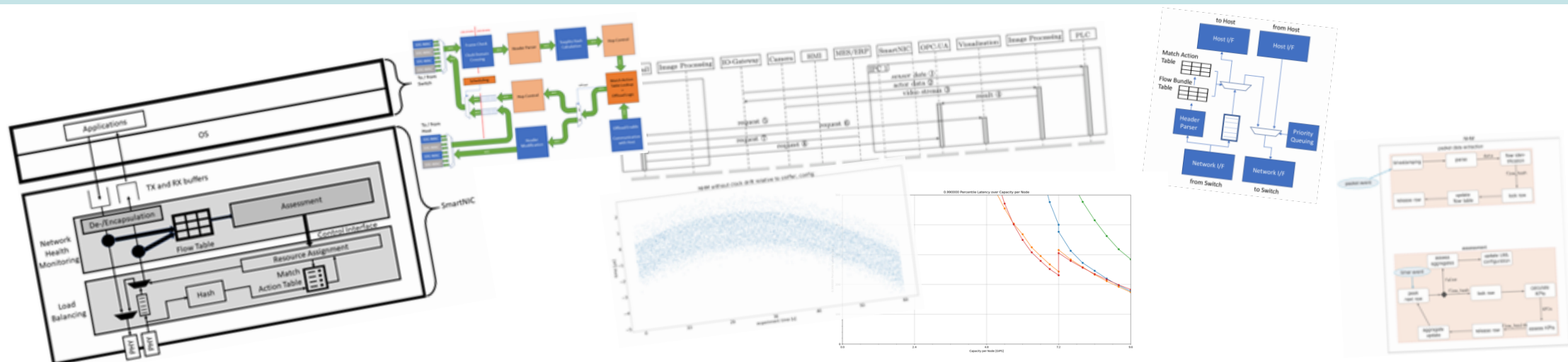
Take Aways …

- High-speed Ethernet poses technical challenges on the dataplane architecture
  - Not only on provisioning sufficient compute performance / accelerators, …
  - equally on data movement and storage

- Crucial relevance of ingress / egress wire-rate pre-/post-processing in NICs

  - Offloading heterogeneous host processing (function repartitioning)

  - Smart traffic steering and monitoring

  - Energy saving and low-latency priority services are not necessarily contradicting goals