



SmartNIC Design Considerations

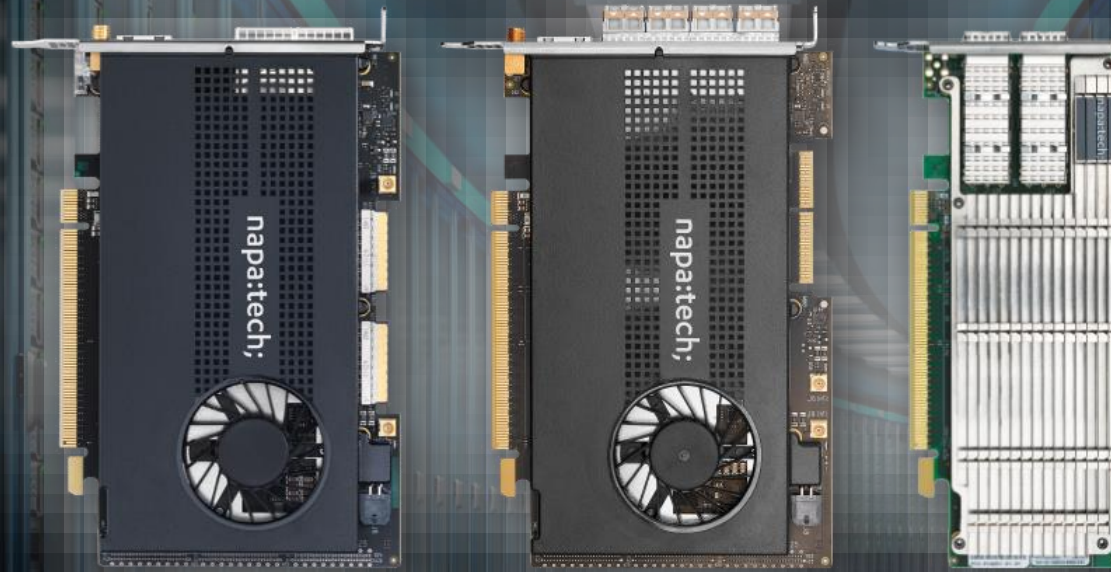
Latency Optimized vs No Loss Packet Capture

Global STAC Live
1 November 2021

Agenda

- Corporate overview
- Packet Capture Trade-offs
- Challenge: Microburst
- Challenge: Host Applications Cannot Keep Up
- How a Napatech SmartNIC Addresses these Challenges

Napatech Builds SmartNICs



- We design high-performance programmable SmartNICs to accelerate applications and capture data at line rate
- Napatech Link™ software supports Capture, Inline, Virtualization
- US Manufacturer with extremely high reliability

Napatech at a Glance

18+-year history delivering
FPGA-based technology to
customers globally

Unparalleled expertise
accelerating compute-
intensive applications on
standard open servers

Targeting rapidly expanding
\$2.7B Programmable NIC
market by 2024
Public company NAPA.OL



Some of the Finance Companies Working with Napatech



Morgan Stanley



Maven



The Dilemma

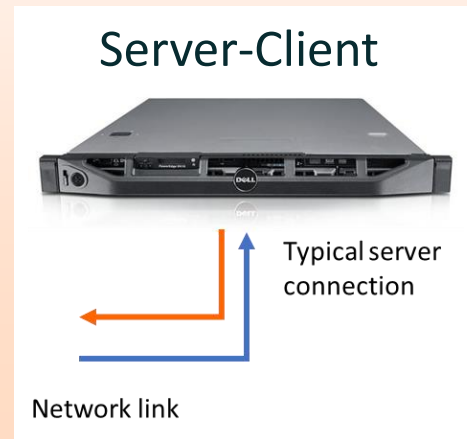


Network interface card architecture cannot be optimized for both ultra-low **latency client/server communication** and 100% packet capture at all speeds.

Packet Capture Trade-offs and Challenges

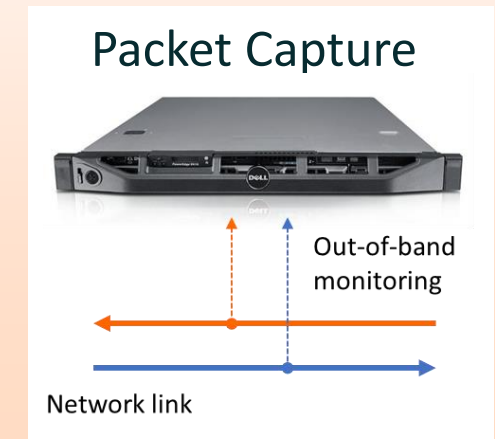
Server NICs are designed for client/server communication

- Ultra-low latency, packet by packet delivery to clients
- Packet delivery not guaranteed at high network traffic loads
- Not optimized for full throughput for any network traffic load
- When used as a capture device, packet loss can occur
- Packet ordering not guaranteed



Packet capture NICs are designed for 100% packet capture

- Designed for packet capture of market data for a complete market picture
- Optimized for capturing all packets no matter traffic conditions (bursts, congestion, packet size or overloaded application)
- Architected with large buffers to capture 100% of the network packets, including microbursts
- Nanosecond resolution time-stamping and synchronization



Challenge: Microburst

Problem

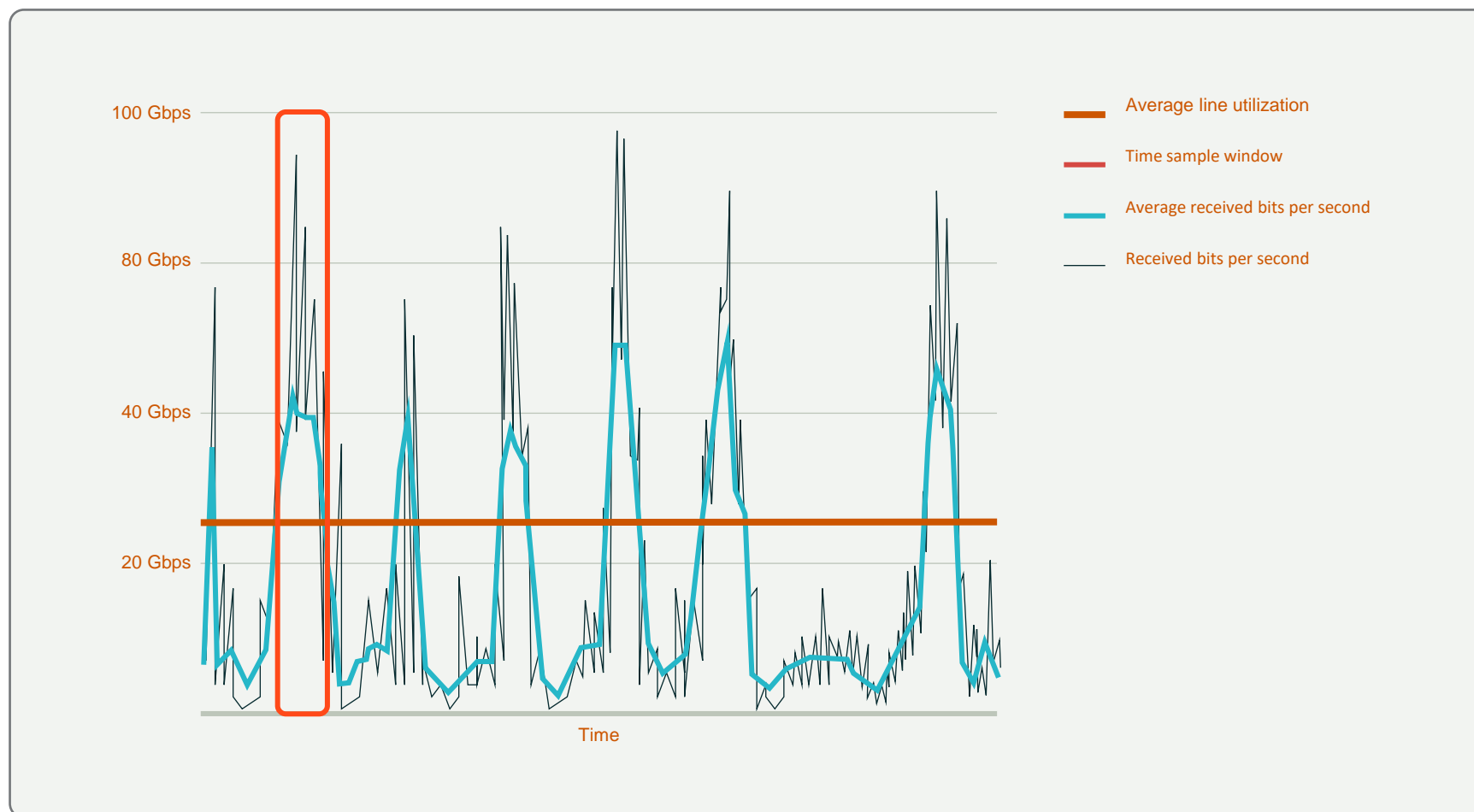
- Volatile stock markets put pressure on trading infrastructure
- Microsecond bursts are not always captured
- Packet queues are small in server NICs causing packets drops during a microburst event

On-board Packet Buffering to Absorb the Burst

- A SmartNIC that provides full theoretical throughput for all packet sizes and all network loads
- Advanced host memory buffer management guaranteeing 100% capture, even during a microburst event

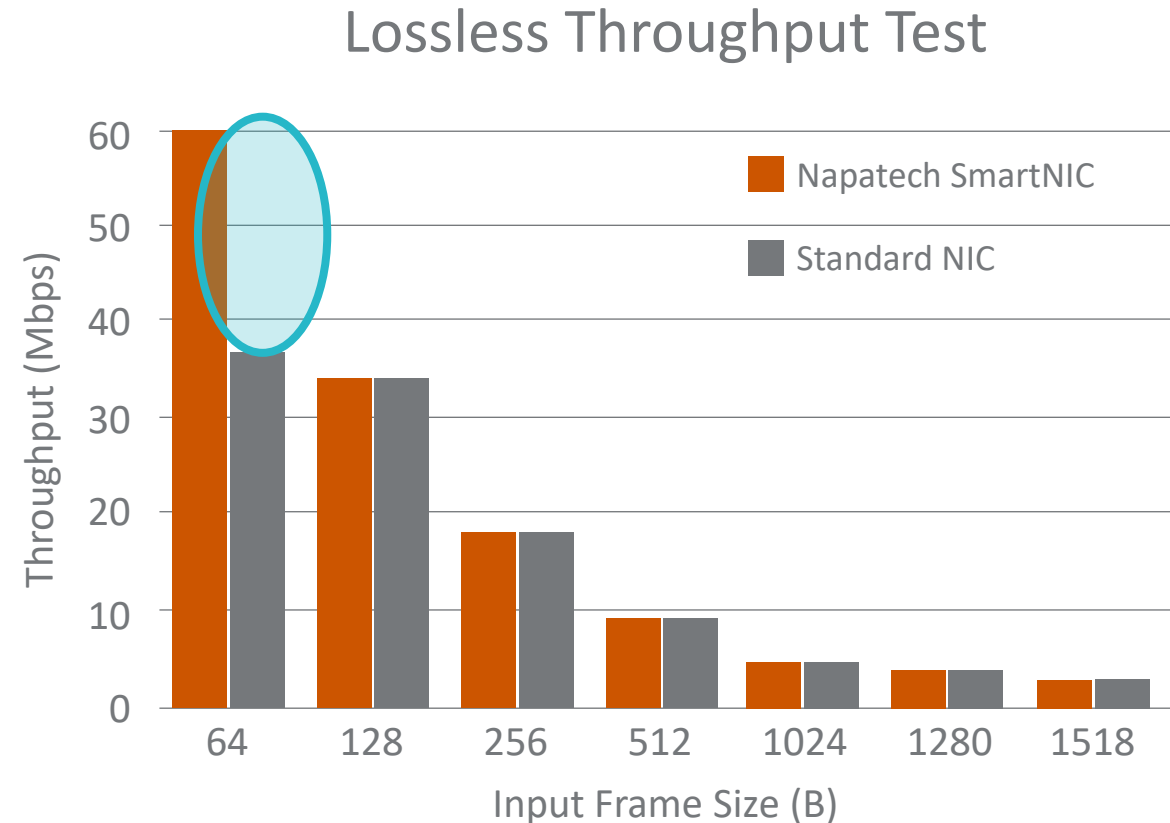


Microburst



Example: Handling Small Packets in Bursts

The standard NIC provides 61% line rate at 64 bytes compared to full theoretical throughput for the Napatech SmartNIC with Link™ Capture Software



See more: DN-1223 Solution Description Generic on Napatech SmartNICs

Challenge: Host Applications Cannot Keep Up

Problems

- Busy CPUs
- Applications cannot keep up with the network load
- Congestion in the delivery of data to the application

On-Board Memory

- Buffering of frames
- On-board memory for buffering assures guaranteed delivery of data, even when there is congestion in the delivery of data to the application

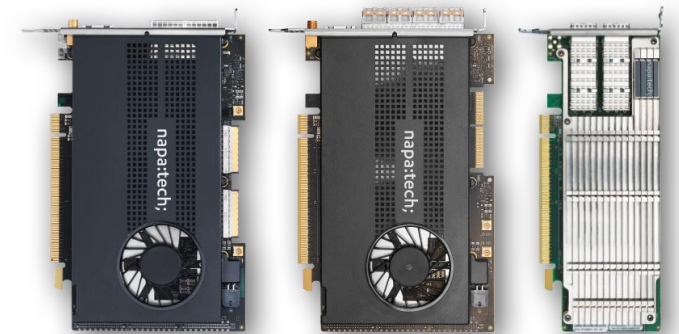
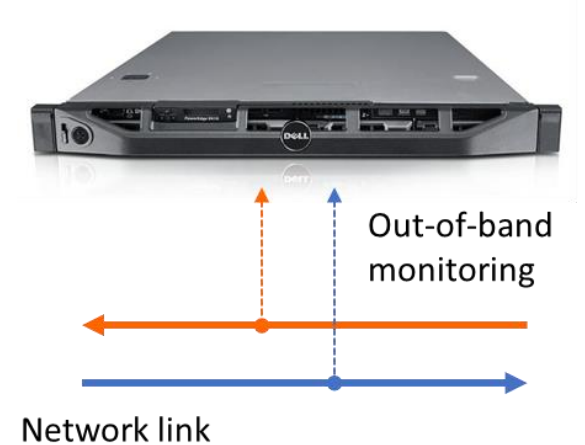
Large Host Buffers

- Advanced host memory buffer management enabling ultra-high CPU cache performance

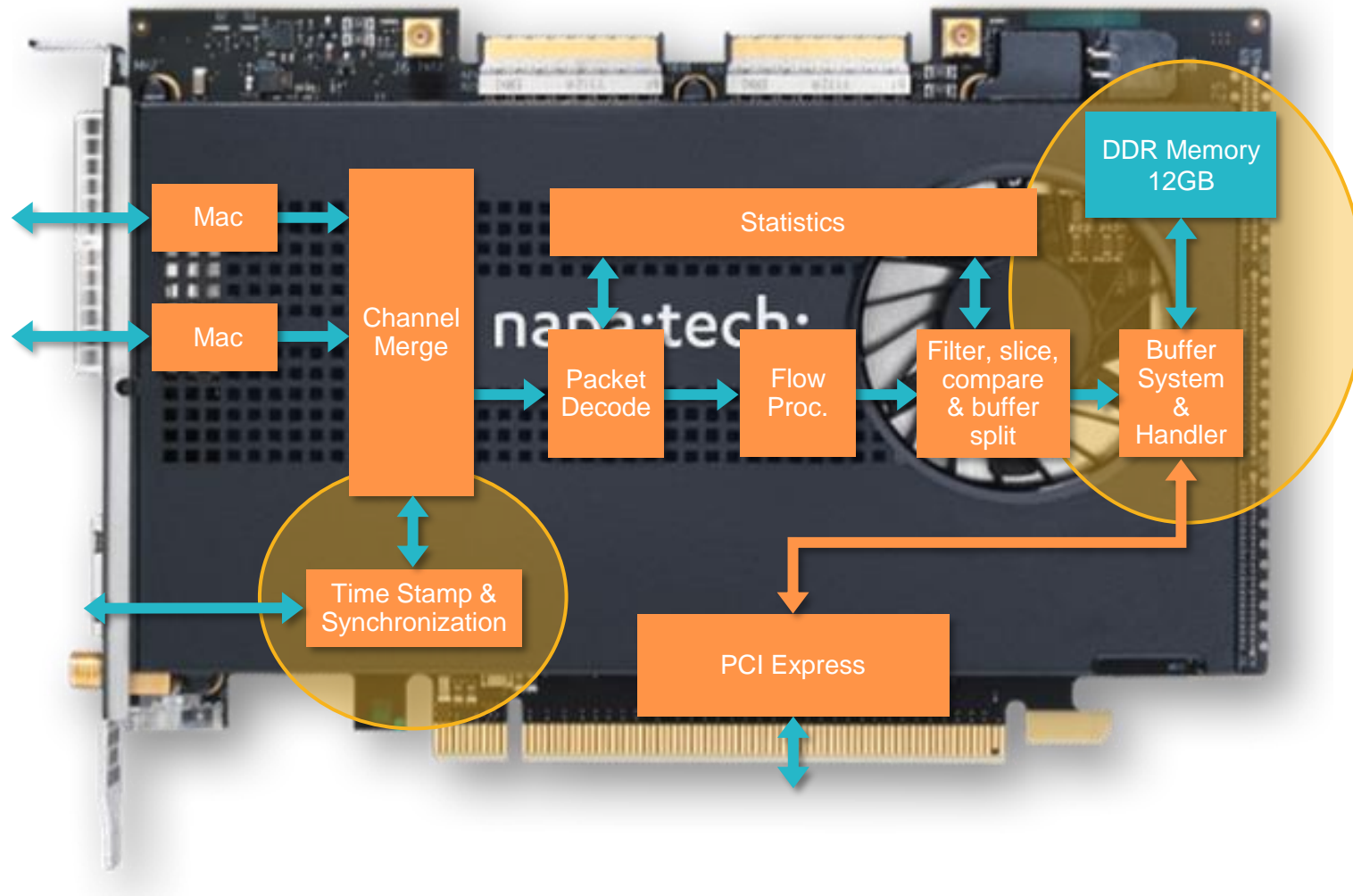
How a Napatech SmartNIC Addresses these Challenges

SmartNIC Hardware and FPGA software purposely designed for packet capture

- Architecture optimized to capture 100% of network packets regardless of network utilization or packet size
- Packets are timestamped in hardware for accurate event information and network performance analysis
- Physical ports are merged in FPGA, guaranteeing host applications see packets in exact order they traversed the network
- Large onboard packet buffer ensures:
 - Zero packet loss when network utilization is high (microbursts)
 - Zero packet loss when PCI express bus is busy
- Large host buffers ensure that host applications can keep up network load
- Zero-copy DMA kernel bypass
- Distribution to multiple host buffers based on flow (RSS)



SmartNIC FPGA Capture Architecture



Napatech Link™ Capture FPGA Software

Napatech NT200A02 running Napatech Link™ Capture Software

Guaranteed Delivery

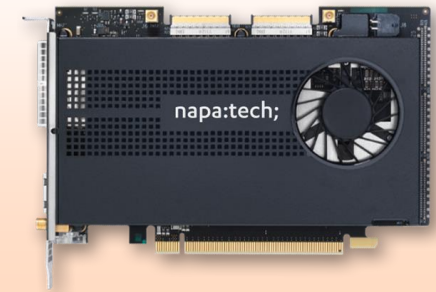
- Full throughput Rx/Tx for any packet size
- Zero packet loss
- Minimum 500 millisecond receive burst buffer @200Gbps (12GB)
- Optimized PCIe bandwidth utilization
- Guaranteed packet ordering

Low CPU Utilization

- Optimized for Intel/Xeon and AMD architectures
- Higher CPU cache performance, advanced host memory buffer system
- Higher CPU core utilization, advanced CPU load distribution



Link™ Capture Software



Advanced Features Examples

Zero Packet Loss

Guaranteed zero packet dropped under the most demanding conditions across all packet sizes, even at sustained line speeds.

Burst Buffering

Each Napatech SmartNIC has onboard memory to handle network microburst, ensuring that all packets are captured and delivered to the application over the PCIe bus.

Hardware Timestamping

Ensures exact timing information for when packets traversed network.

Traffic Replay

Allows high fidelity replay of a PCAP file with nanosecond precision to reproduce exact network behavior in test and measurement use cases.

Excellent PCIe Performance

PCIe throughput is maximized to so that the best possible network capture performance is achieved.

NUMA Balancing

Allows optimization of host processes and application threads to minimize NUMA to NUMA communication over the QPI interface.

Optimum Cache Utilization

Transfer of packets over the PCIe bus are optimized for the L3 cache in the processor, minimizing the frequency that the L3 cache is flushed, improving the overall performance of the host application.

Packet Sequencing

Ethernet port merging ensure that packets are delivered to application across multiple port in the correct order.

Traffic Forwarding

Built in packet broker functionality where incoming packets can be filtered and or load balanced out one or more interfaces, potentially eliminating the need for expensive load distribution solutions.

Intelligent Multi-CPU Distribution

Incoming packets can be sorted to 128 host buffers based on flow, L2-L4 filter, or a combination thereof, resulting in efficient utilization of all server CPU cores.

Traffic Generation

Generate traffic based on PCAP files, or open-source applications like TRex at all line speeds up to 100Gbps.

Join Us at the Next Napatech Breakout Session

Packet Capture as the Market Moves to 100G and Beyond

What are the issues encountered when moving from 10/40G to 100G?
How to capture and accurately replay the recorded market data?

Stay Connected With Napatech

napa:tech;

RECONFIGURABLE COMPUTING

On the Web:



napa:tech;
RECONFIGURABLE COMPUTING
www.napatech.com

Social Media:



Other News:

Google 
www.google.com/alerts


www.napatech.com/newsletter

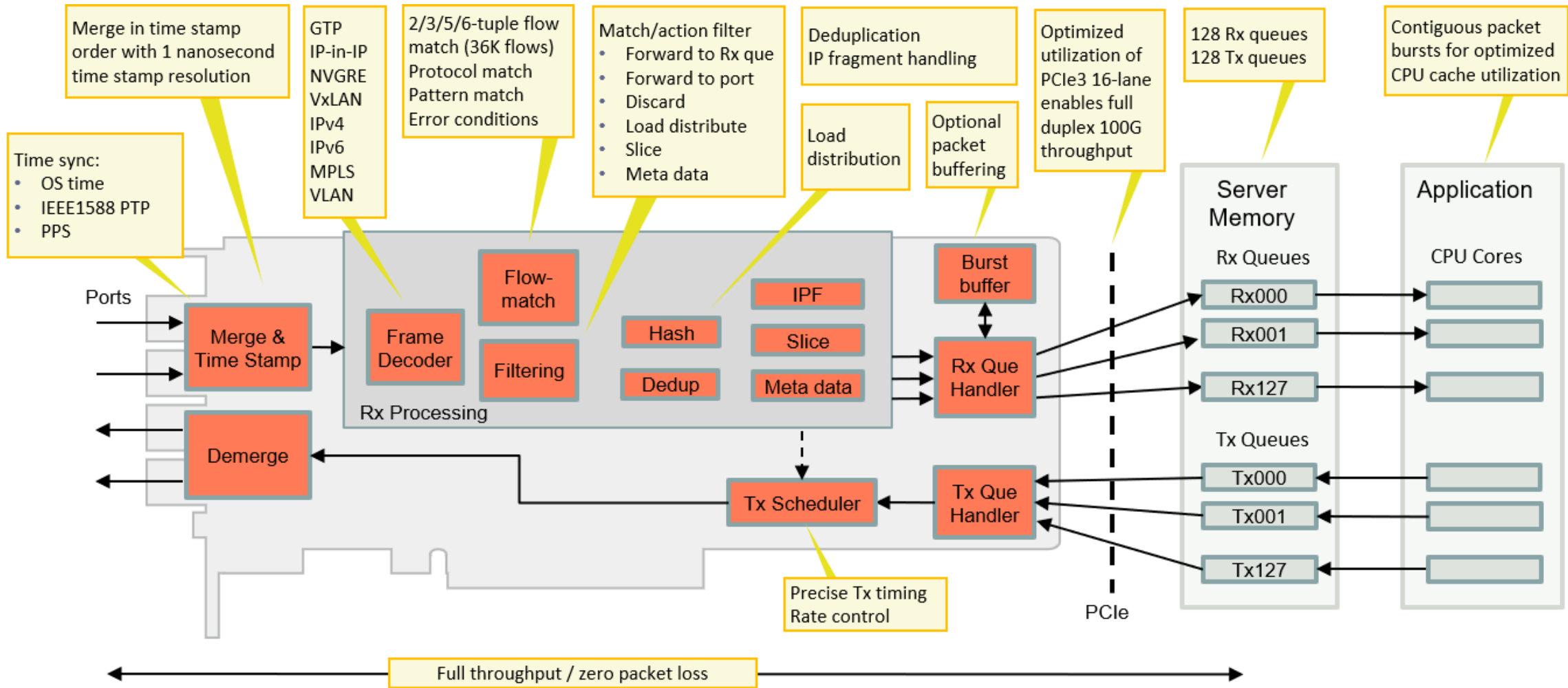
napatech

napa:tech;
RECONFIGURABLE COMPUTING

Thank You!

• APPENDIX

Link™ Capture Software Packet Processing Overview



NIC Usage Comparison Chart

| | | Standard NIC | Latency Optimized NIC | Capture NIC | |
|--------------------------|--|---|--|---|---|
| Primary use | | General data center deployment | Financial trading servers | Network performance and security analytics | |
| Key requirements | Line rate throughput (any packet size) | Less important | Somewhat important | Very important | |
| | Zero packet loss (100% packet capture) | Less important | Somewhat important <ul style="list-style-type: none"> Packet loss can cause retransmission | Very important <ul style="list-style-type: none"> Packet loss critical for security applications | |
| | Low latency | Somewhat important | Very important (100s nsec) | Less important (u secs) | |
| Implementation | Hardware type | Standard Ethernet ASIC | FPGA or specialized ASIC | FPGA | |
| | Hardware latency* | Low | 300 + nsec | N/A | |
| | FPGA configuration** | (N/A) | <ul style="list-style-type: none"> Optimized for low latency Per packet processing | <ul style="list-style-type: none"> Optimized for high throughput: Store and Forward Optimized for zero packet loss: Large buffers Optimized for CPU utilization: All frame sizes | |
| | | Standard Linux network driver | Proprietary driver <ul style="list-style-type: none"> Optimized for low latency | Proprietary driver <ul style="list-style-type: none"> Optimized for high throughput Optimized for low CPU utilization | |
| Integration & Deployment | Considerations | Requires seamless I&D Deployment | Customization is acceptable <ul style="list-style-type: none"> Low latency is top priority | Some customization is acceptable <ul style="list-style-type: none"> Packet capture is top priority | *FPGA SmartNIC Hardware has same low latency as a standard NIC (ASIC) |
| | Driver | Standard <ul style="list-style-type: none"> Linux network driver | Proprietary <ul style="list-style-type: none"> Optimized for low latency | Proprietary <ul style="list-style-type: none"> Optimized for high throughput Optimized for CPU utilization | **FPGA code (loaded on the FPGA SmartNIC) can configure the SmartNIC to be either a Standard NIC, an HFT NIC or a Capture NIC |
| | API | Standard OS API <ul style="list-style-type: none"> Linux/Windows | Proprietary <ul style="list-style-type: none"> Optimized for low latency | De facto and proprietary <ul style="list-style-type: none"> PCAP, DPDK, vendor proprietary | |
| Cost (relative) | | \$ | \$\$ | \$\$\$ | |