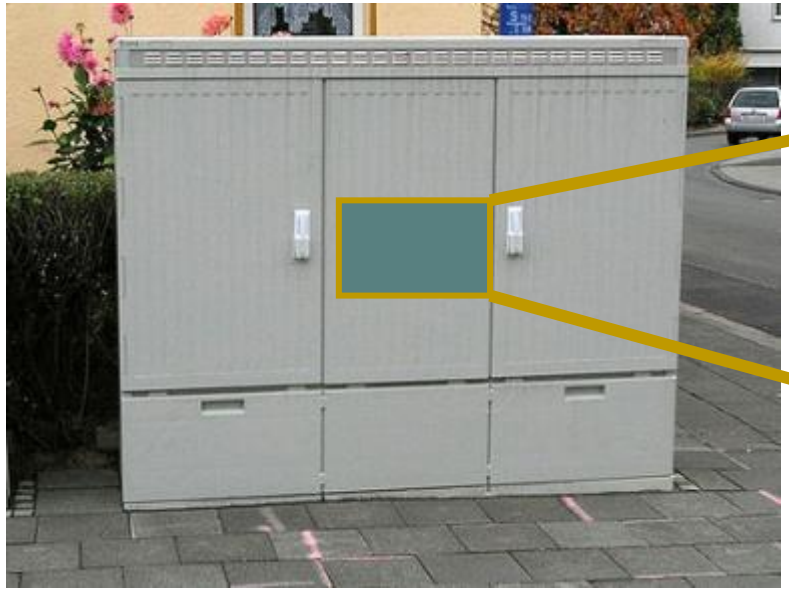
Architecture of a Synchronized Low-Latency Network Node Targeted to Research and Education

Christian Liss[1], Marian Ulbricht[1], Umar Farooq Zia[1], Hartmut Müller[2]
[1]InnoRoute GmbH
[2]Verilogic

19 June 2017

OUR HISTORY: IP-CORE FOR DSL ACCESS MULTIPLEXERS (DSLAM)



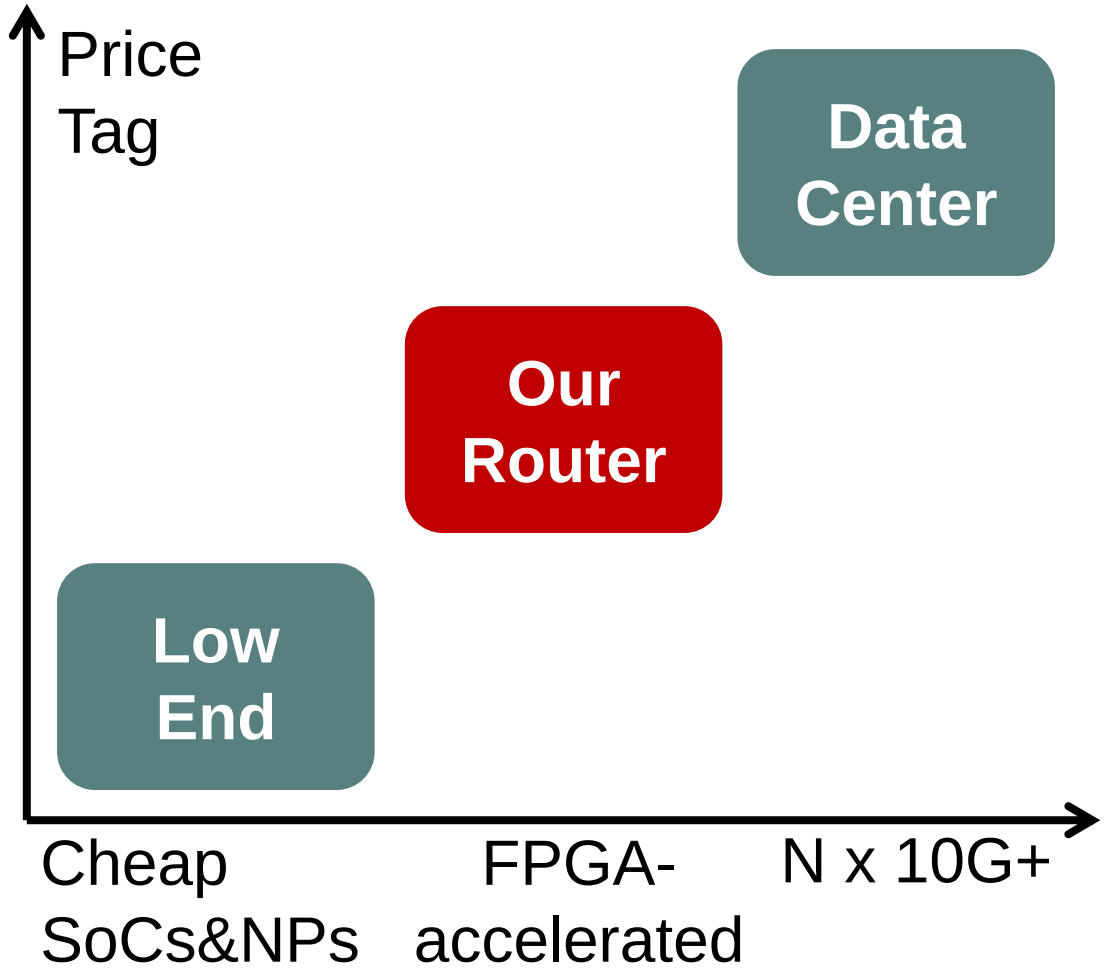
First version of the **FlowEngine**:

- QoS packet processor on **FPGA**
- Very **compact**: Spartan-6 75T
- Full **line speed**, up to 10Gbps
- **Store-and-Forward**
- VDSL2 and G.fast

Source of upper left picture: Wikipedia

DECISION: CREATING AN UP-TO-DATE ROUTER, BASED ON FLOWENGINE

- Target Applications:
  - Research & Education: 5G, IoT …
  - Small and Medium Businesses
  - Industrial Networks, e.g., Automation
- Product Outline:
  - Versatile & Affordable Ethernet Interfaces
  - Lowest Latency & Jitter
  - FPGA & Software
  - Time Synchronization



REQUIREMENTS: 5G MOBILE NETWORKS & INDUSTRY 4.0

- **Latency** (100µs e2e) & **Jitter** (64ns e2e):
  - **many nodes** in lines/trees/rings
  - challenging **control loops**
  - avoid **bufferbloat**
- Hierarchical **time synchronization**:
  - **Phase**: timestamps
  - **Frequency** (4.7ppm to 2ppb): recovered clock + quality signaling
- **5G** access networks require at least 1 Gbps/antenna
- **Industry 4.0** networks (IT, OT) require 1 Gbps links (& slower for legacy devices)
- **Engineers** need **free ports** for monitoring, administration, …
- **Experimenters** want a **precise and robust device**

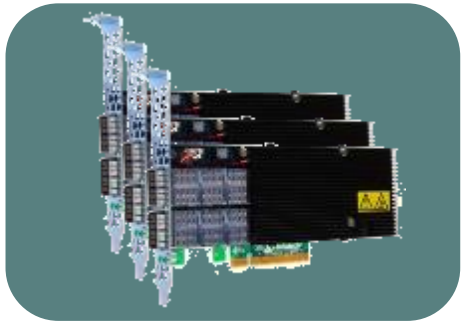
- TSN for fronthaul
- 32 or 64 hops
- Very low error rate

STATE OF THE ART

Dedicated **Switch/Router**

3µs

Missing control & visibility for custom functionality or high latency



PC with **FPGA cards optimized for low latency**

Few ports, only 10G+ interfaces.

Inefficient for 1G

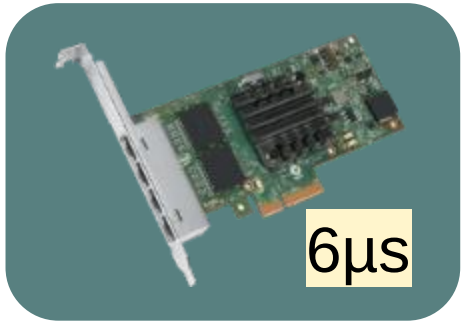
Latency depends on software stack and input/output port combination.

PC with one or two **NetFPGA** cards

High latency (store & forward) and few ports.

No frequency synchronization option.

Huge latency for large frames



XEON **servers with standard NICs**

Easily available and scalable, but higher latency

SOLUTION: TrustNode – 19“ RACK-MOUNTABLE/DESKTOP NETWORK NODE

Versatile Local User Interface

Device Status

GigE ports with SyncE Master/Slave support

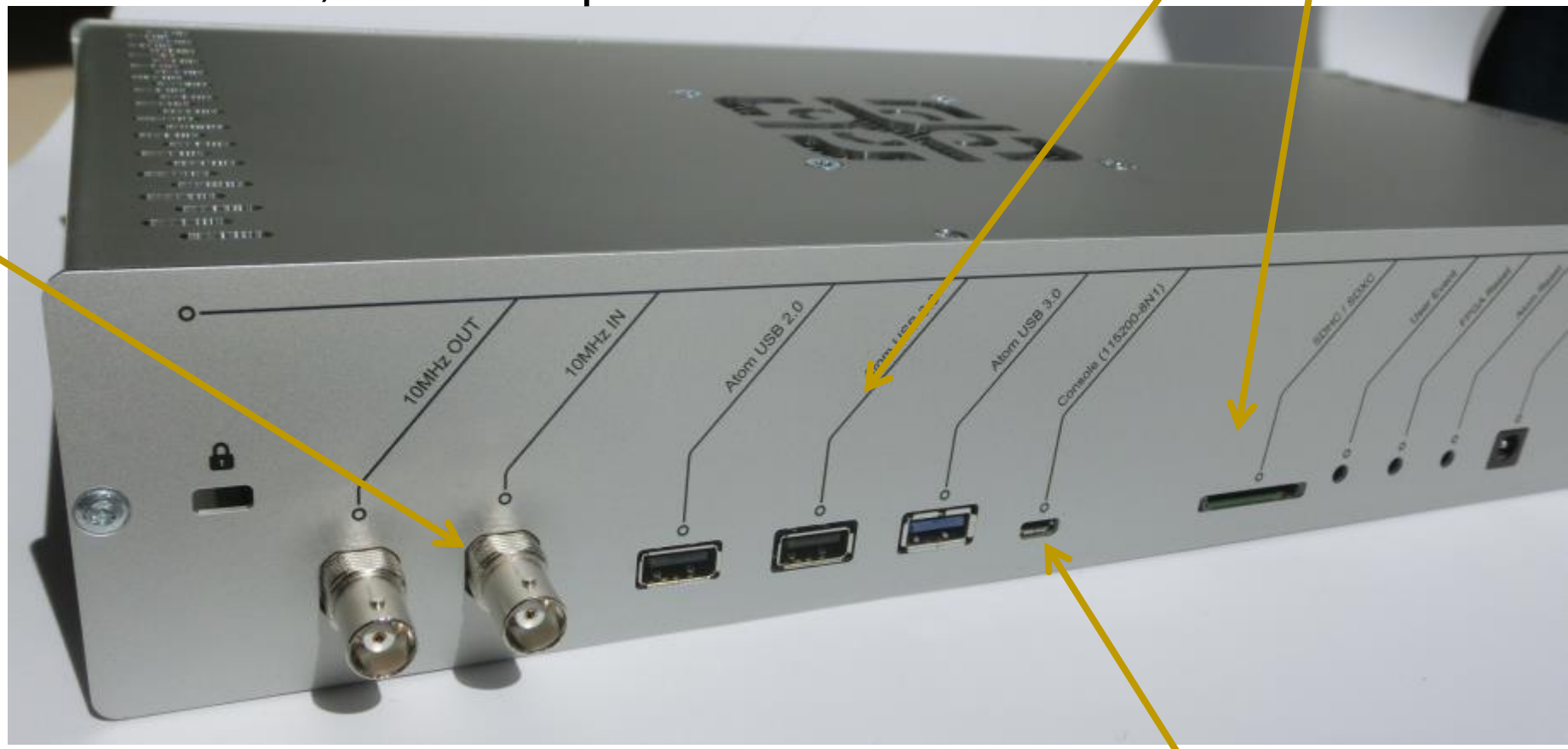
GigE ports as SyncE Master

TRUSTNODE INTERFACES

Mechanically robust, EMI checked, and ESD protected

Boot options

GPS Clock Reference
(programmable,
stackable)



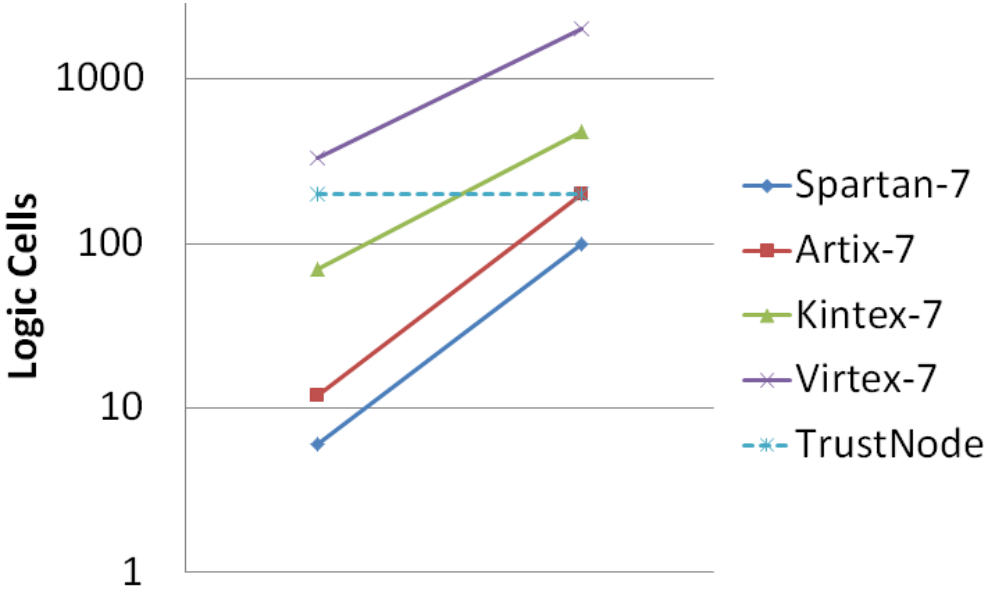
As little persistent state as possible
→ Devices shared by researchers or students

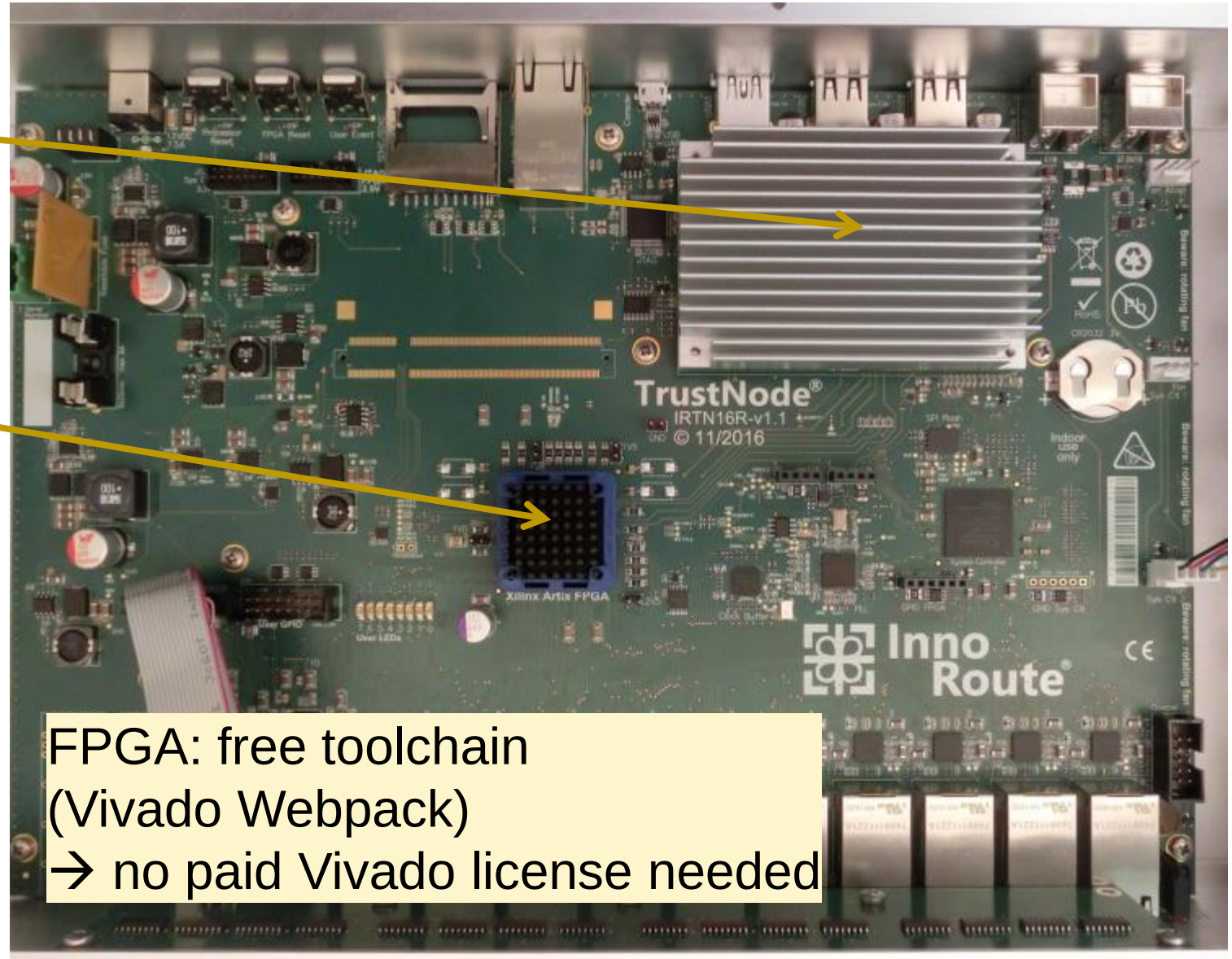
Local Console
to access the Atom

MAIN BOARD

Intel Atom E3845 SoC
Quad-core with 4GB RAM
- Control Plane
- Software-based Data Plane

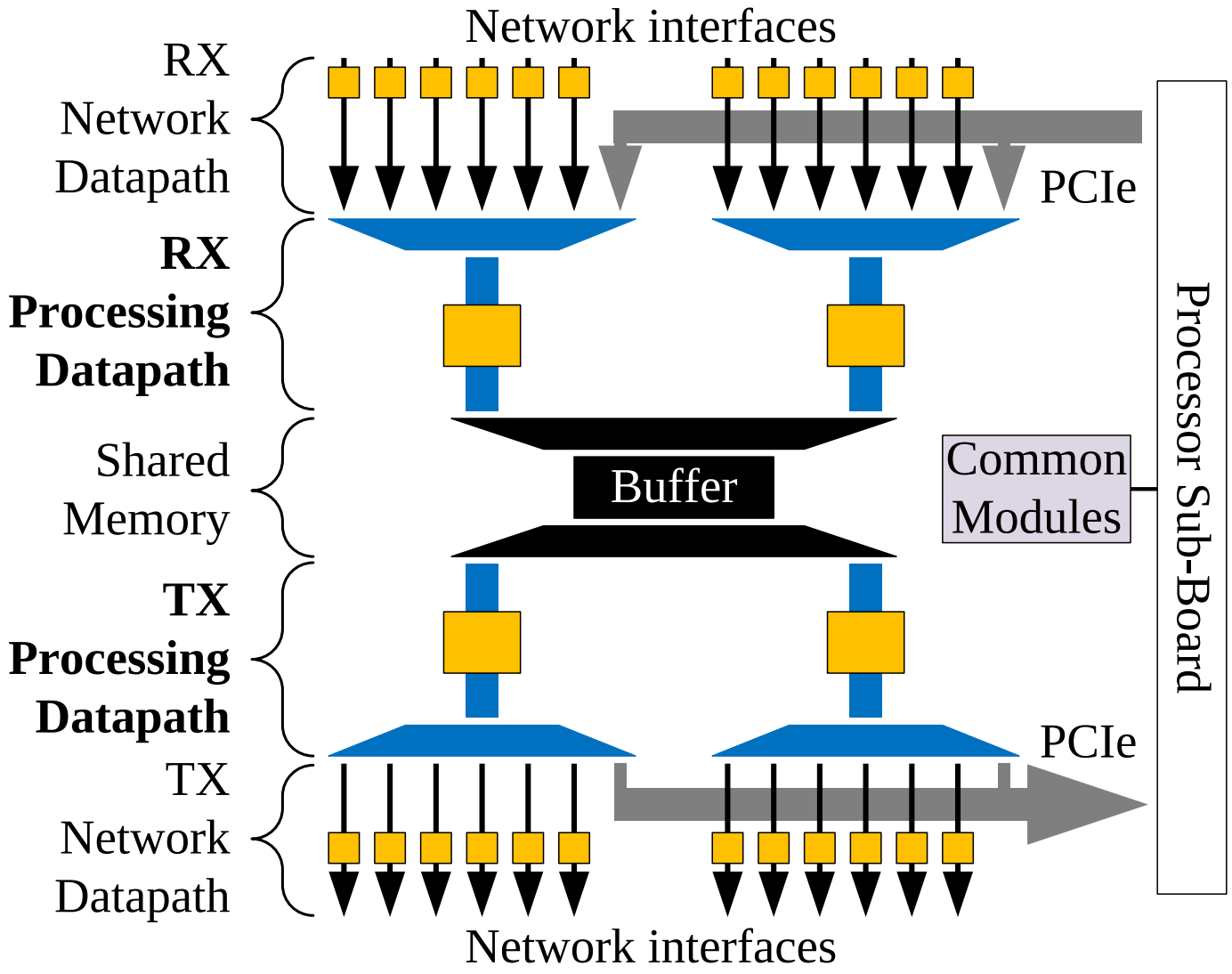
Xilinx Artix-200T FPGA
Low-latency Data Plane





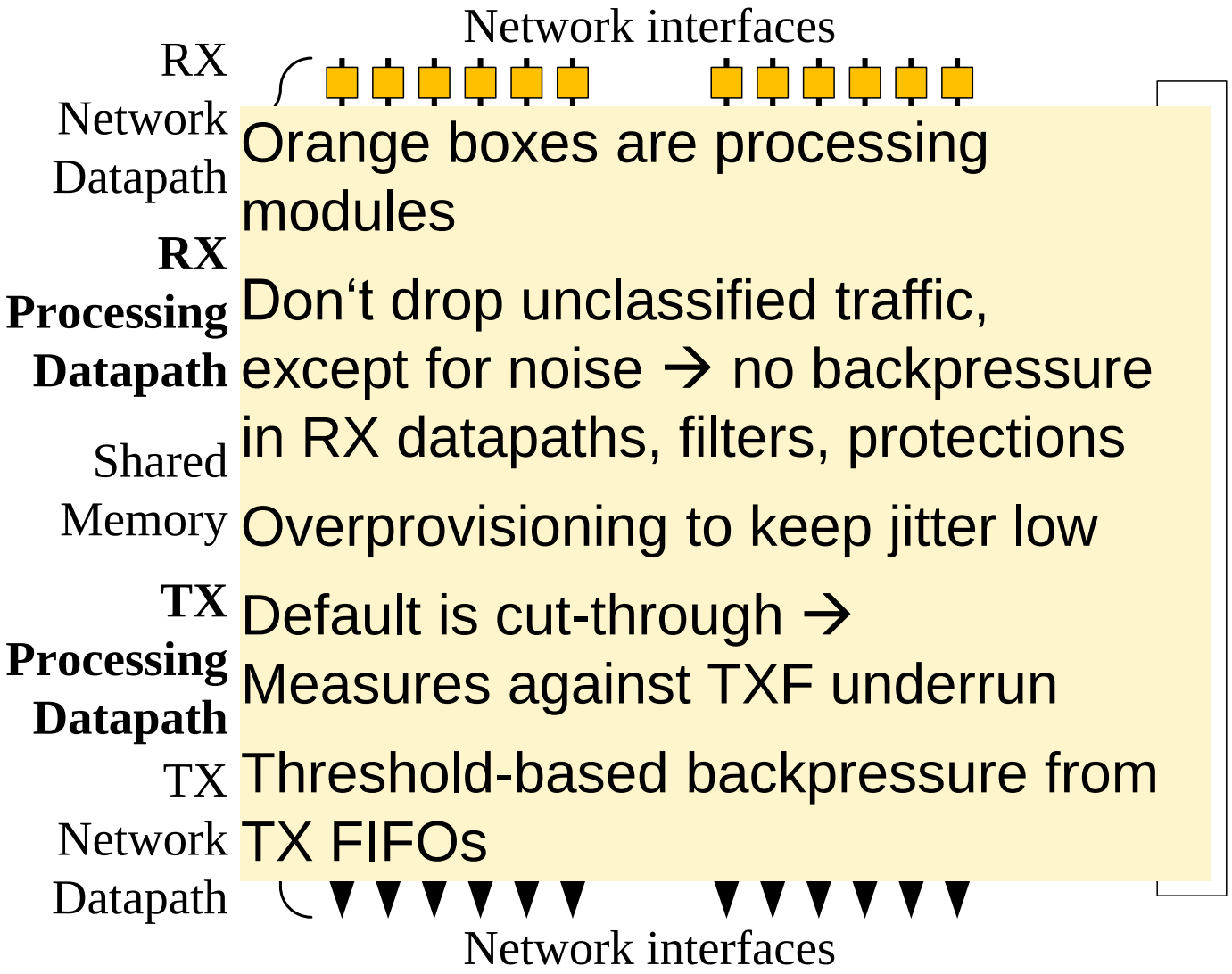
FPGA: DATAPLANE ARCHITECTURE

1. Ethernet PHYs
2. RGMIIs
3. Eth. RX MACs
4. RX FIFOs
5. RX Arbiters
6. RX Datapaths
7. Buffer Manager/ Queue Manager/ Scheduler
8. TX Datapaths
9. TX FIFOs
10. Eth. TX MACs
11. RGMIIs
12. Ethernet PHYs



FPGA: DATAPLANE ARCHITECTURE

1. Ethernet PHYs
2. RGMIIs
3. Eth. RX MACs
4. RX FIFOs
5. RX Arbiters
6. RX Datapaths
7. Buffer Manager/ Queue Manager/ Scheduler
8. TX Datapaths
9. TX FIFOs
10. Eth. TX MACs
11. RGMIIs
12. Ethernet PHYs



Orange boxes are processing modules

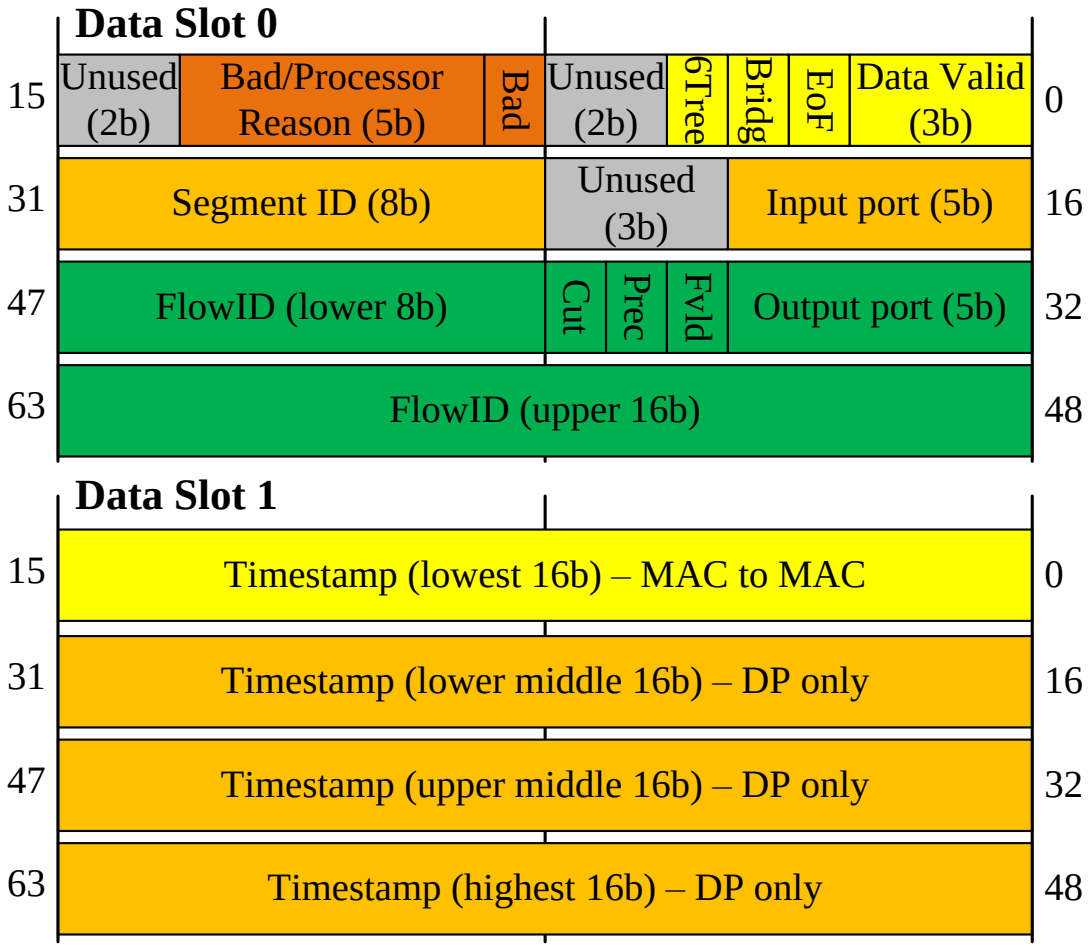
Don‘t drop unclassified traffic, except for noise → no backpressure in RX datapaths, filters, protections

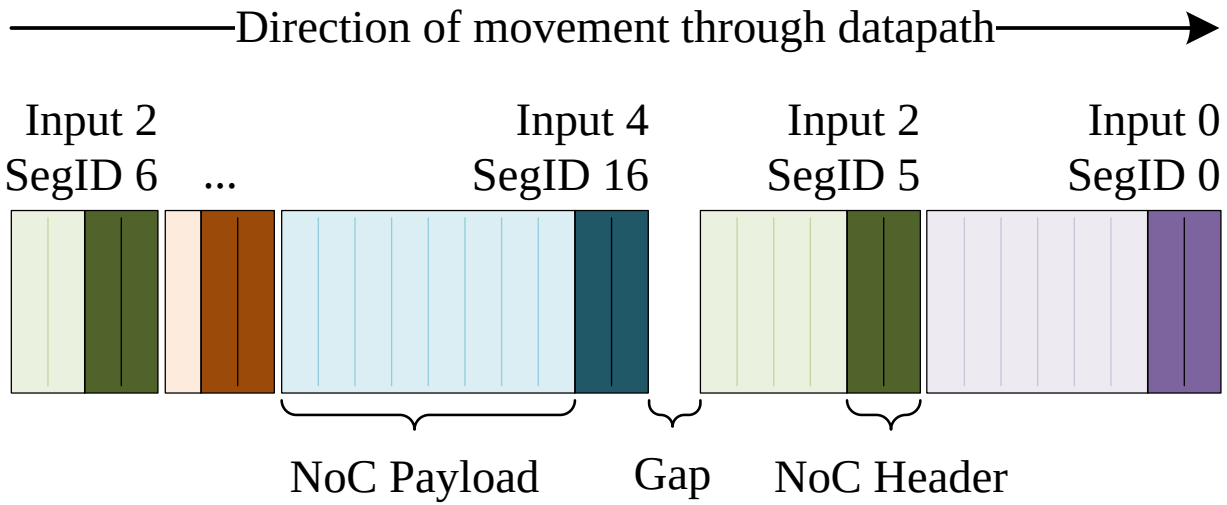
Overprovisioning to keep jitter low

Default is cut-through → Measures against TXF underrun

Threshold-based backpressure from TX FIFOs

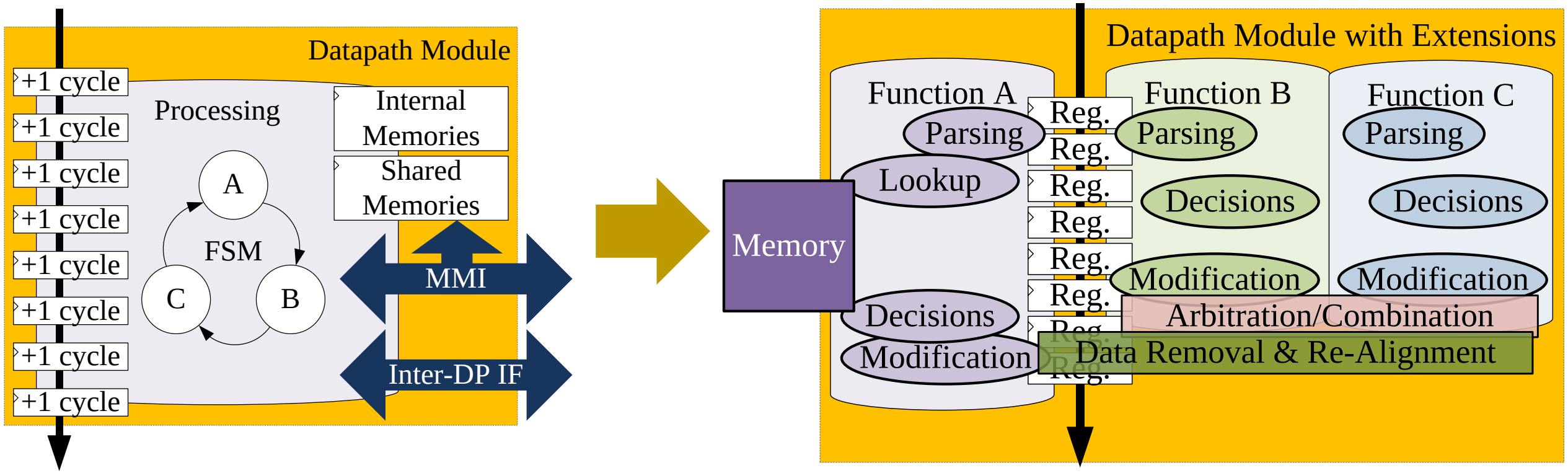
TRUSTNODE NETWORK-ON-CHIP: HEADER AND SEGMENTS





- 2 NoC header data slots are followed by 8 segment data slots (64B segments) → process new segments every 10 cycles – more challenging for smaller last segments
- Precision scheduling
- Drop reason analysis
- Up to 16 FPGA PHY ports and up to 16 software network ports

NETWORK-ON-CHIP MODULES: GENERAL STRUCTURE & EXTENSIONS



- **Segmented** frames **interleaved** over Network-on-Chip. Segment size: **64 Bytes** (Range: 1-128 Bytes)
- **Timestamping of all** received frames (125MHz clock, 20ns uncertainty)

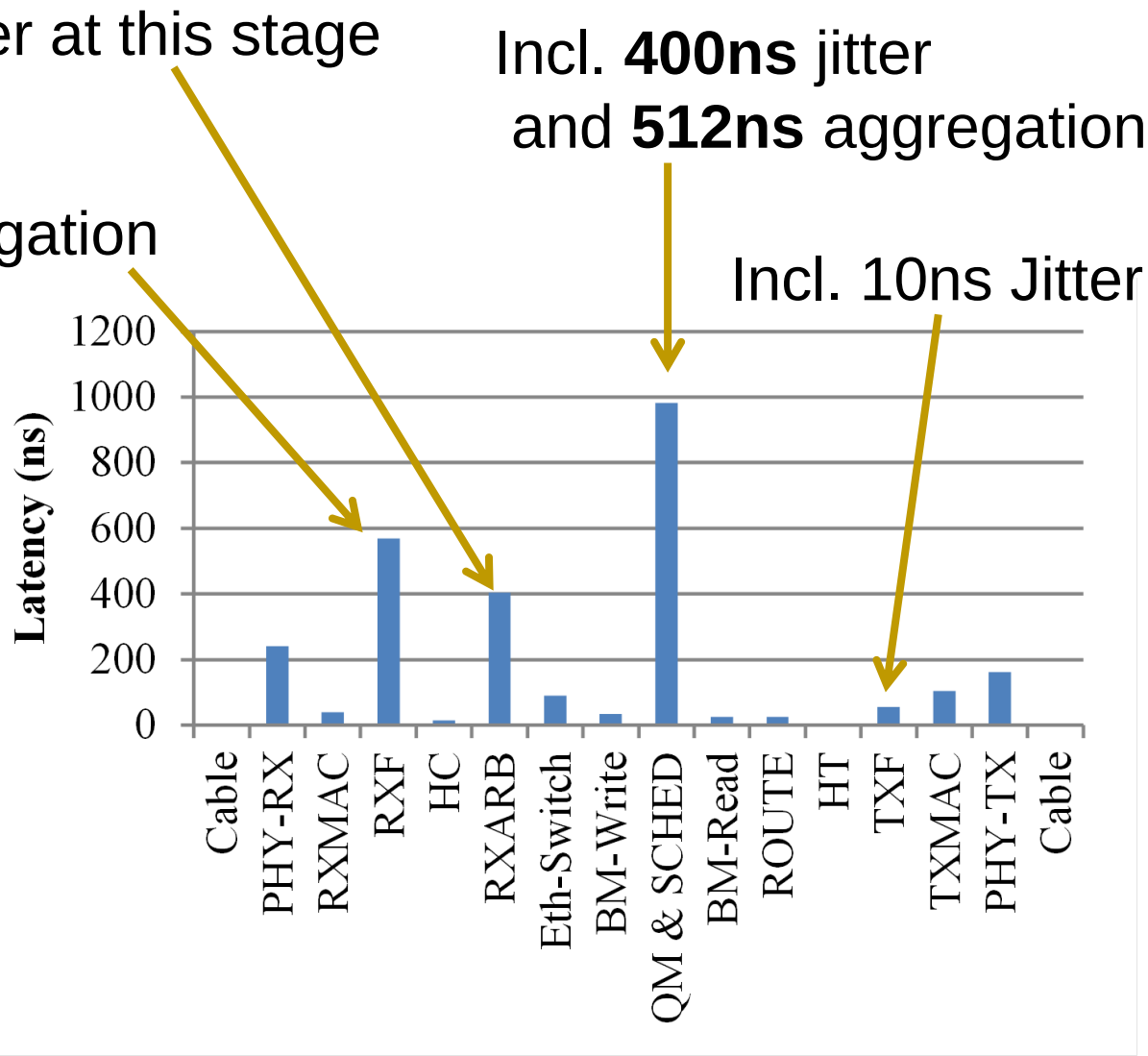
MMI = Memory Mapped Interface: Software Control Interface

Each and every module optimized for latency

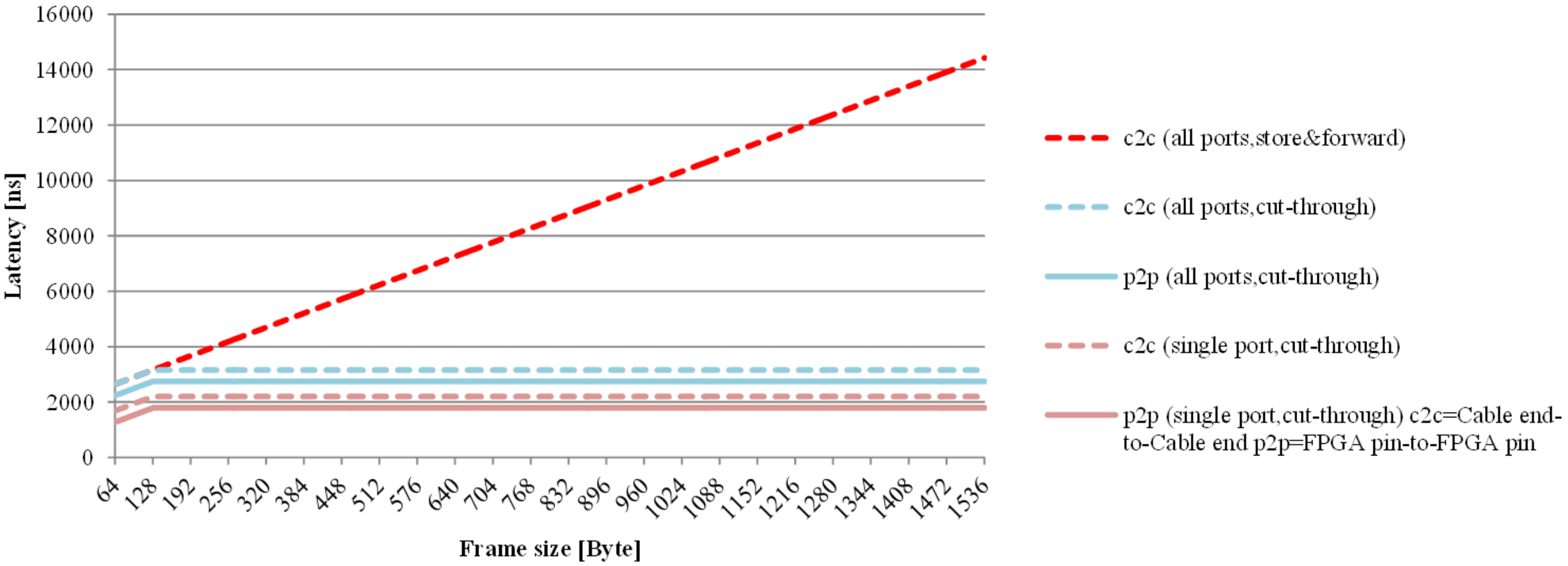
ANALYTIC WORST-CASE LATENCY MODEL

Model guided development
Dejittering in QM expensive: time and space
→ new arbitration schemes

- Most of the jitter is caused by **arbiters**:
  - RX Arbiter
  - QM Scheduler
- **Aggregation** to segments and number of **pipeline stages** are major latency factors

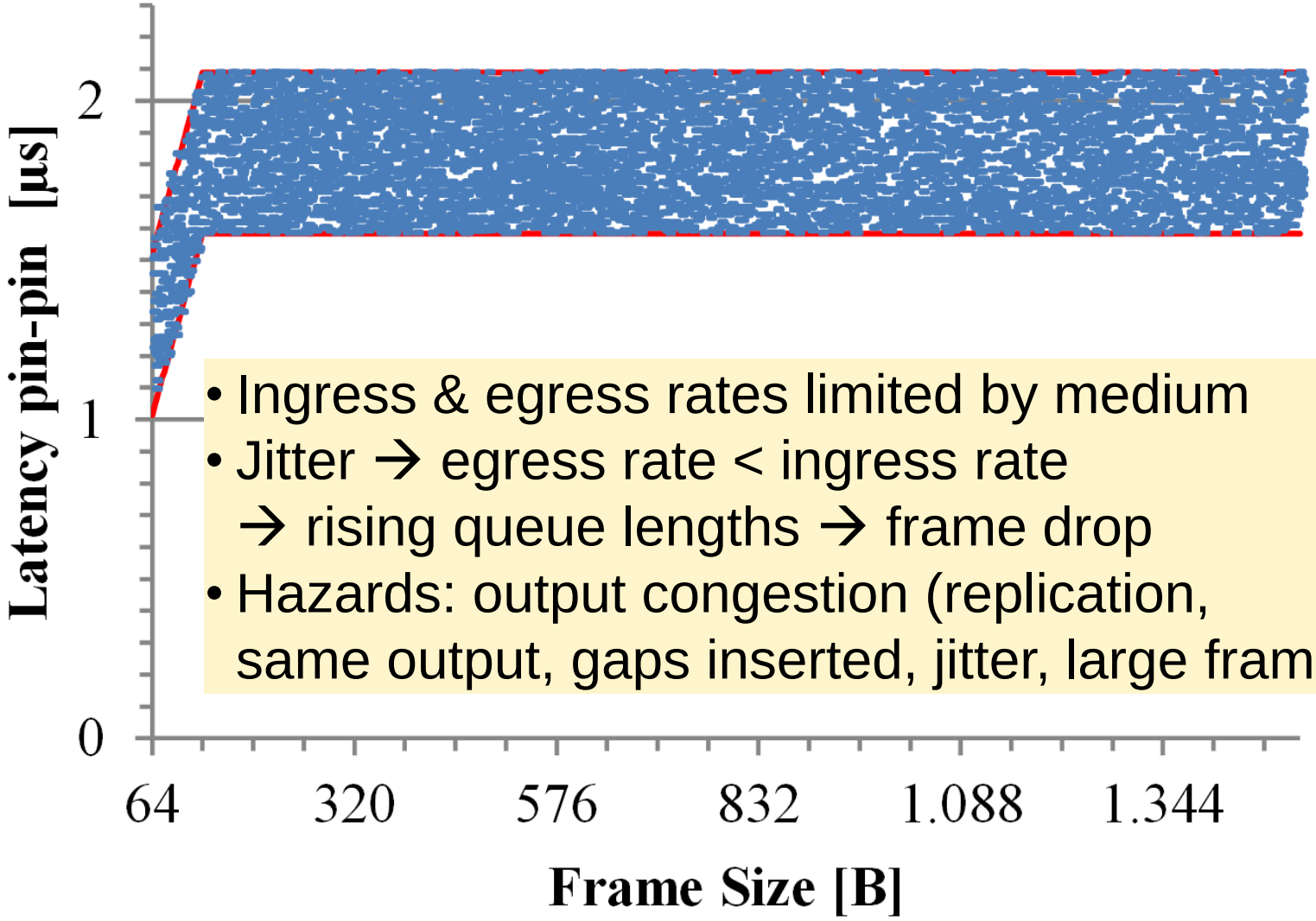


ANALYTIC LATENCY MODEL: CUT-THROUGH VS. STORE-AND-FORWARD



LATENCY & JITTER: FIRST MEASUREMENTS

- **FPGA-only** measurements (pin-to-pin)
- Low-speed, single-port, to have **empty queues**
- PHYs add additional **401-721ns**
- Constant latency & jitter **above 127 Bytes** (two segments)



CONCLUSION & SUMMARY

- **Lowest-latency Gigabit Ethernet** Switch/Router
- Processing at **line rate**
- **More than 50%** of the FPGA available for extensions

- Analytic model **validated** by measurements:
  port-to-port cut-through latency stable and below 3µs
  in case of empty output queues

- Low latency allows for **multi-hop** industrial/automotive/5G networks
- Precision allows for **repeatable experiments**
- Designed for **usability**

Special Versions:

- SODIMM up to 8GB
- SFP Ports
- GPIO Extender
- Admin Port

and many more …

Selfnet

CHARISMA

Contact Us

Christian Liss, Head of R&D

InnoRoute GmbH
Marsstrasse 14a
80335 Munich
Germany
+49 89 4524199 - 02
liss@innoroute.de
Visit us: www.Innoroute.de

CEO Andreas Foglar
Registrations: Amtsgericht München
VAT ID: DE 271566134
WEEE: DE 84823388