Zone-Based Automotive Backbones Tunneling PCIe®

Dr. Endric Schubert
Chief Technology Officer
Missing Link Electronics, Inc.

PCI-SIG Virtual Developers Conference 2021

Members Implementation Session

Disclaimer

- Presentation Disclaimer: All opinions, judgments, recommendations, etc. that are presented herein are the opinions of the presenter of the material and do not necessarily reflect the opinions of the PCI-SIG®.

Acknowledgements

- Ulrich Langenbach
  Dir. Eng. Missing Link Electronics GmbH (MLE)
- Marcus Pietzsch
  Group Mgr, Fraunhofer Institute for Photonic Microsystems (IPMS), Dresden, Germany

Presentation Outline

WHY?

- More safe and eco-friendly vehicles drive automotive connectivity towards so-called Zone-Based Architectures. Inside those Zone Gateways PCIe® connects multiple SoC for scalable performance. Zone Gateways connect with each other via the emerging IEEE standards “Time Sensitive Networking” (TSN).

WHAT?

- A solution that fulfills the need for “tunneling” PCIe over TSN (working title PCIe-over-TSN), supporting CPU-to-CPU communication (PCIe NTB) and NVMe storage.

HOW?

- A digital circuit & system stack to encapsulate and to decapsulate PCIe TLPs (along with other protocols) over real-time automotive TSN 10G/25G Ethernet

Automotive Network History (1967)

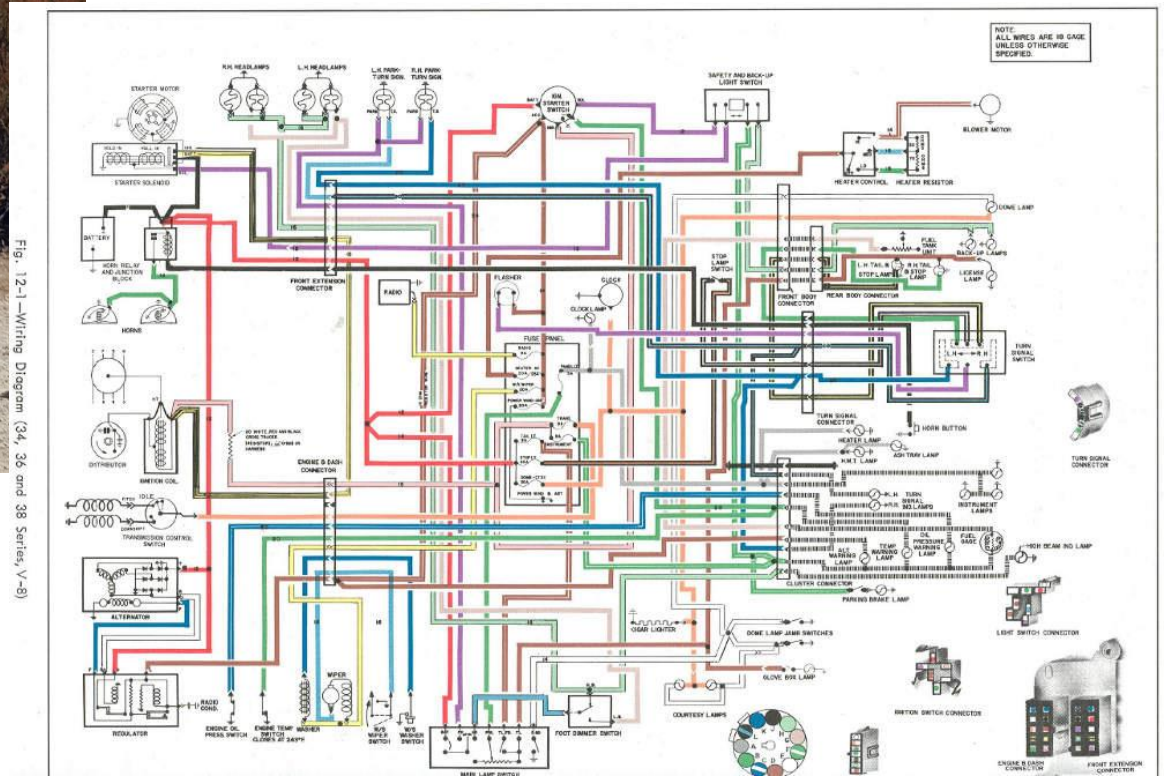
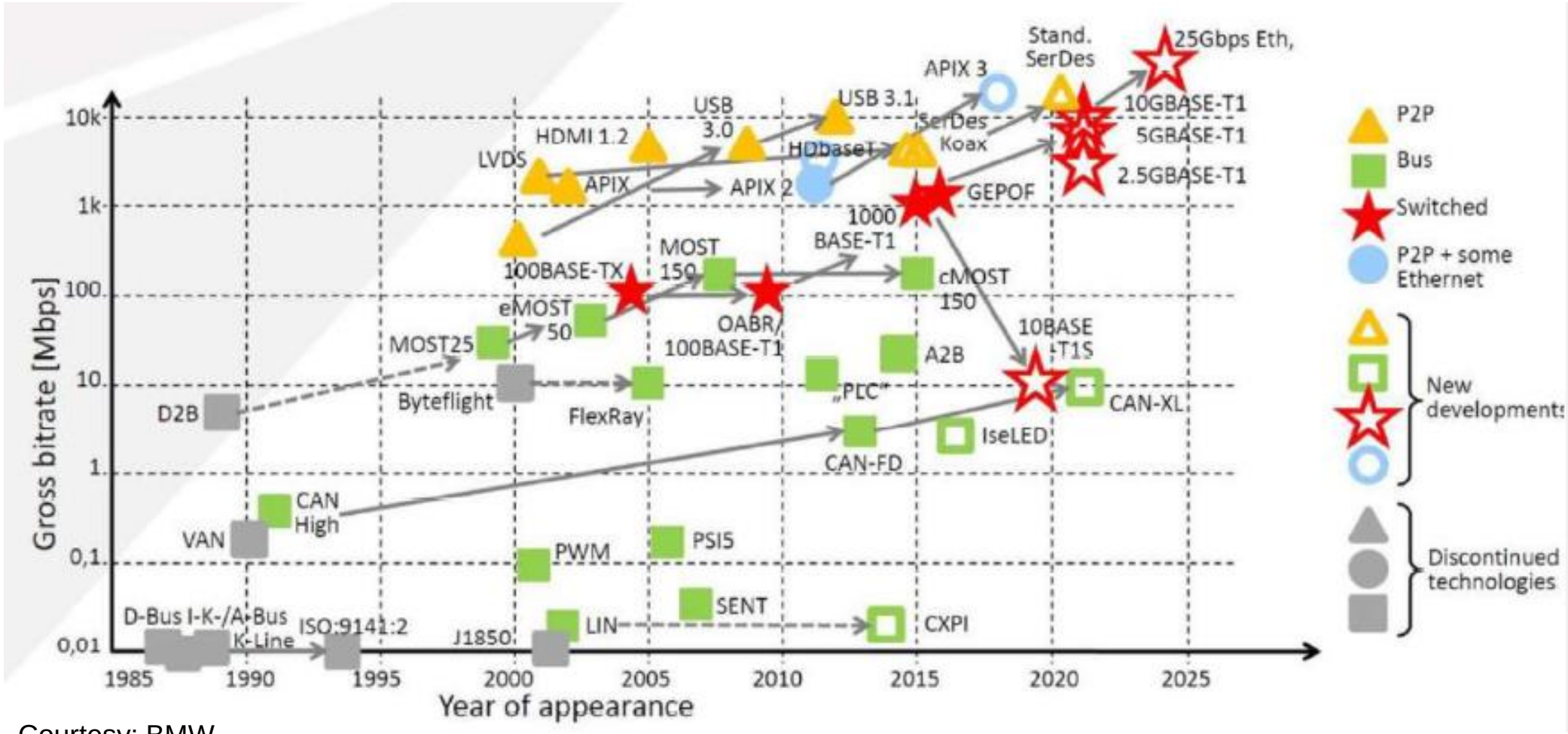


Fig. 12-1—Wiring Diagram (34, 36 and 38 Series, V-8)

Evolution of Automotive Networks



Courtesy: BMW

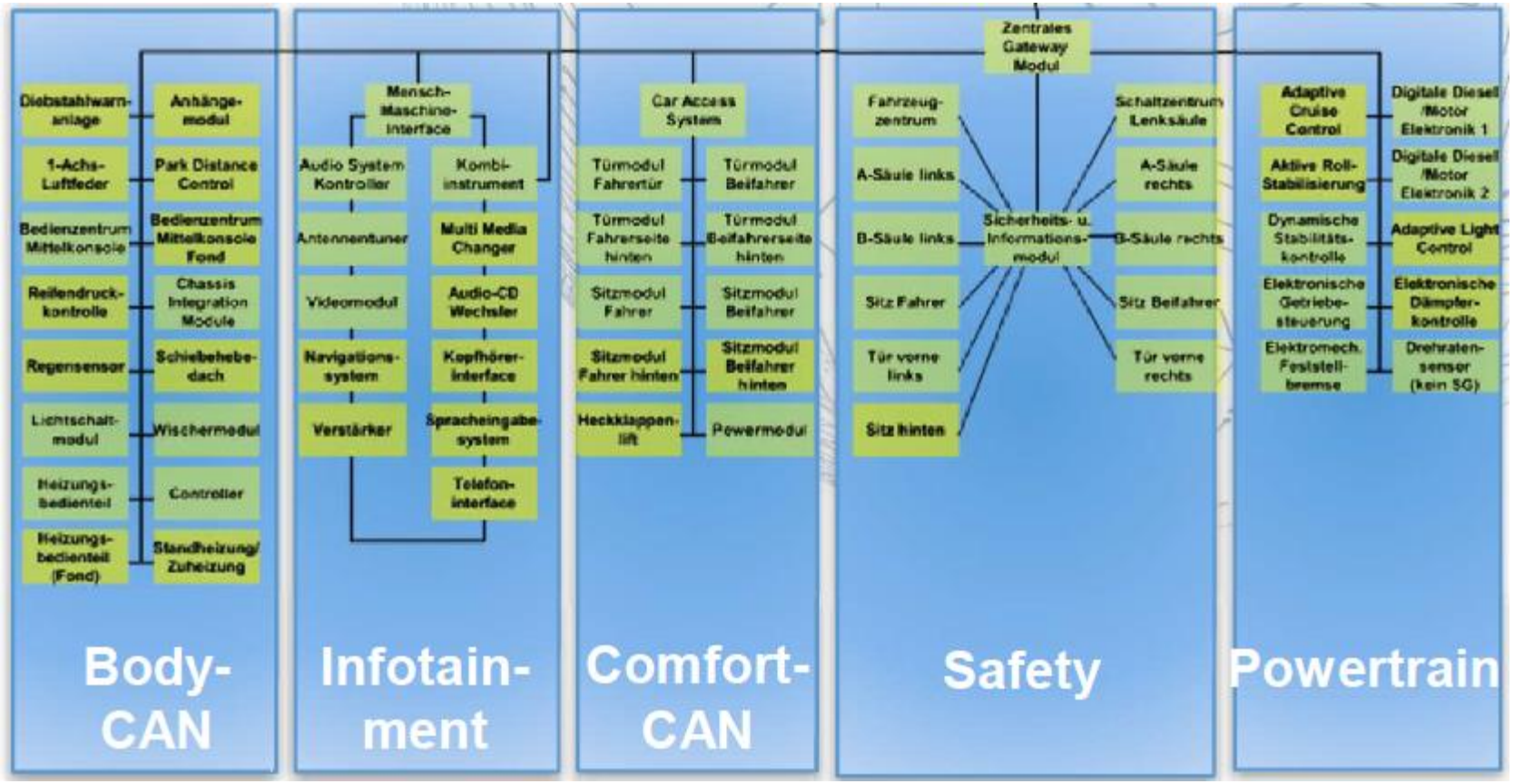
Today's Expensive Wiring Nightmare

Automotive Networks Today (Domain-Based)

Organized in Functional Domains

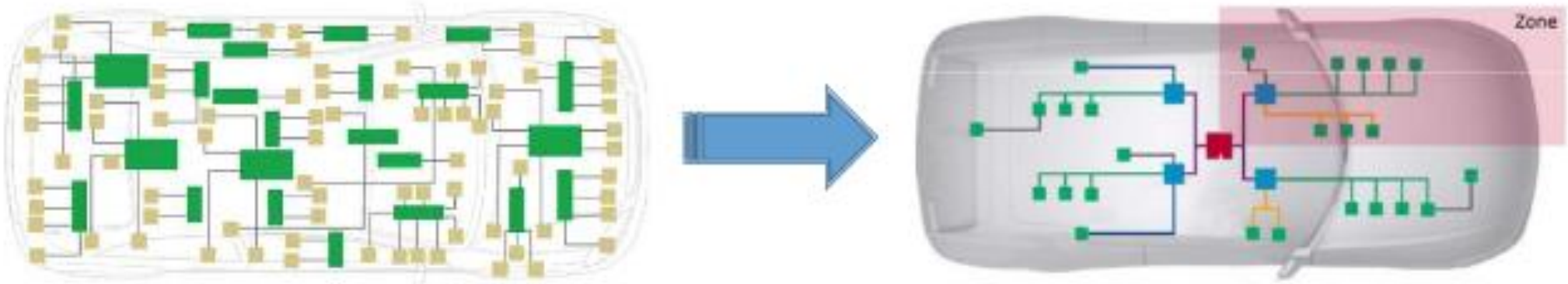
- Powertrain, Body Control, Comfort, Infotainment, etc
- Bus Gateways for Cross-Domain Communication

100 ECUs, or more via many different busses and protocols



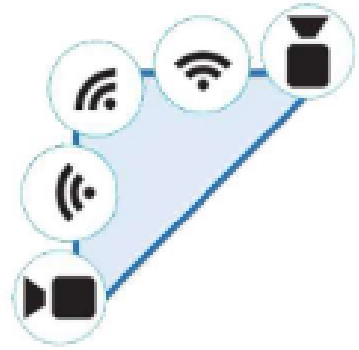
Next: Zone-Based Architectures

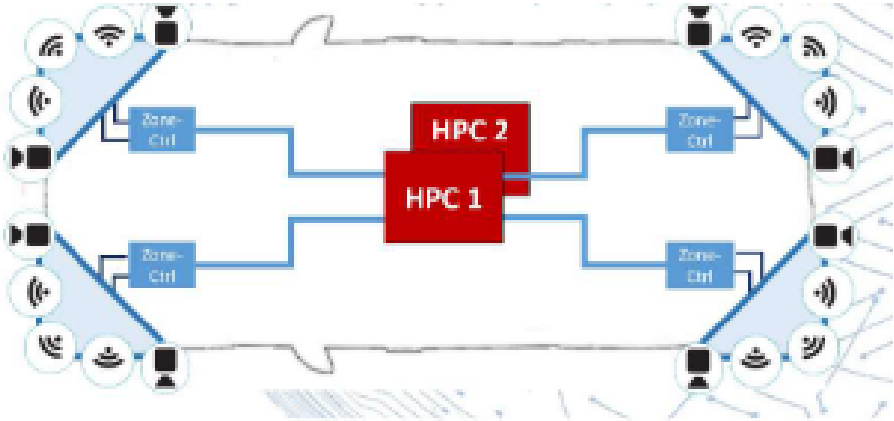
- Network connectivity is not based on functional domains, but on *physical location and proximity* inside the vehicle, i.e. “Zones”
- Data aggregation and preprocessing in Zone Controllers:
- High-bandwidth connectivity towards a central “High Performance Computer” HPC



Application Example: “Smart Corner”

- Smart Corner:
  integrates all data sources and sinks located at one corner of a vehicle
- Smart Corner Node contains e.g.:
  - 2 Lidars
  - 2 Radars
  - 2 Cameras
  - 2 Ultrasonics
  - 1 Lighting Unit





Traffic Type	Characteristics
Scheduled Traffic	High-priority real-time traffic transmitted according to a time schedule (time-driven), no interference from other traffic
Stream Reservation	Periodic, guaranteed
Event-driven Traffic	Aperiodic bursts, generated by sporadic events, with real-time constraints
Best-effort Traffic	No guarantees; statistical performance

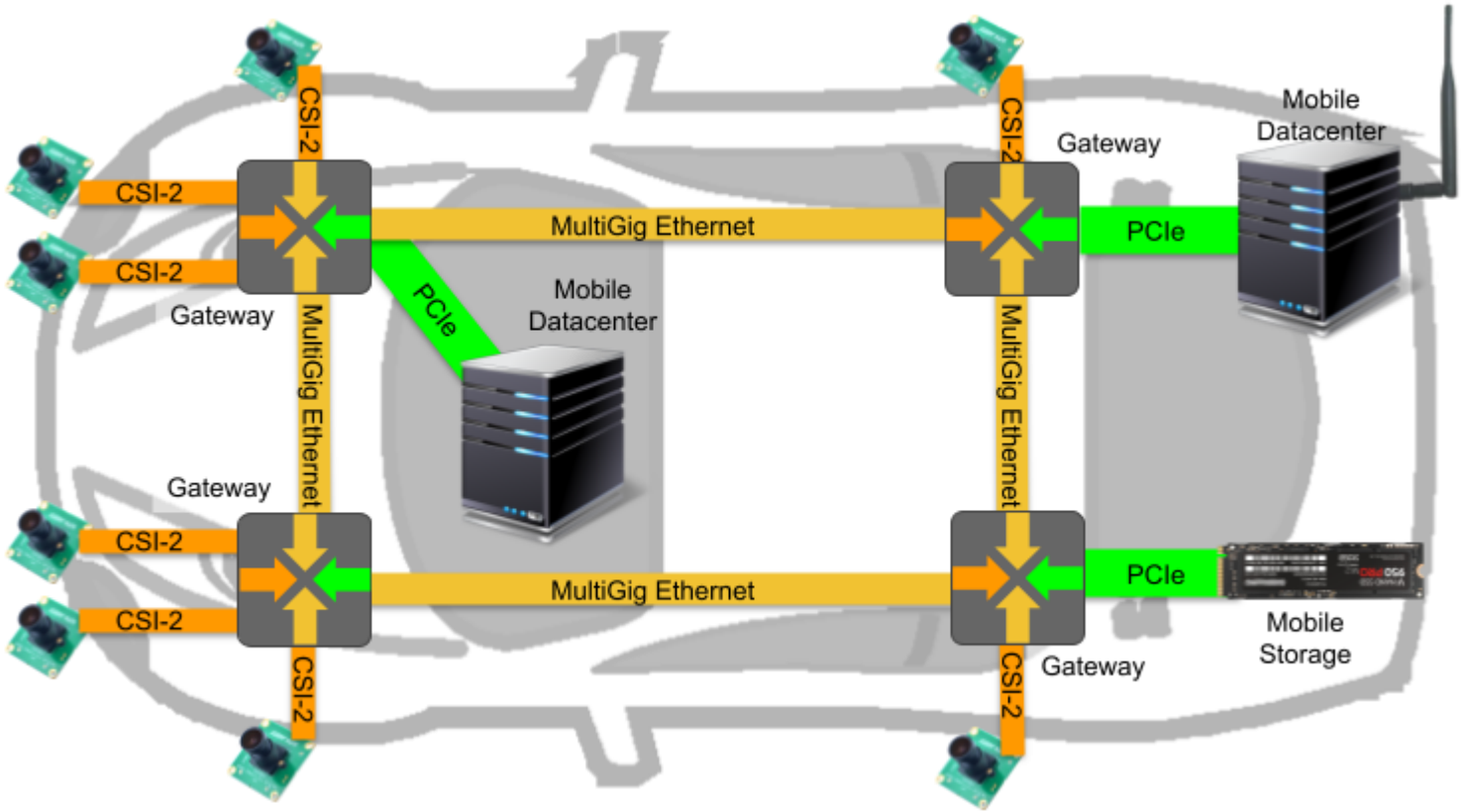
Zone-Based Automotive Network Needs to Transport PCIe

Driven by Cost/Performance, i.e. Centralized Compute & Storage:

- PCIe (for Embedded CPUs, GPUs, FPGAs and SoCs)
- NVMe (for SSDs)

Driven by Compliance

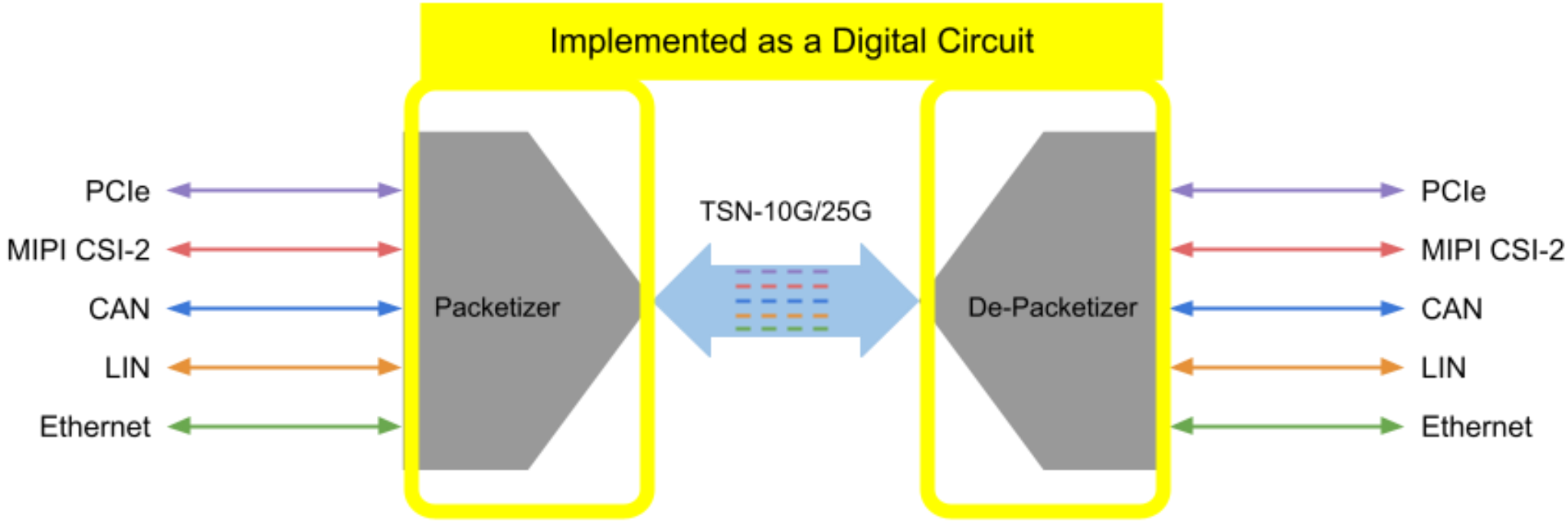
- FuSa ISO 26262
- Security ISO/SAE 21434
- SOTIF ISO 21448
- etc



PCIe-over-TSN is Based on Open Standards

PCIe from PCI-SIG, TSN from IEEE

Symmetric for CPU-to-CPU (e.g. PCIe NTB) or Asymmetric Sensor-to-CPU



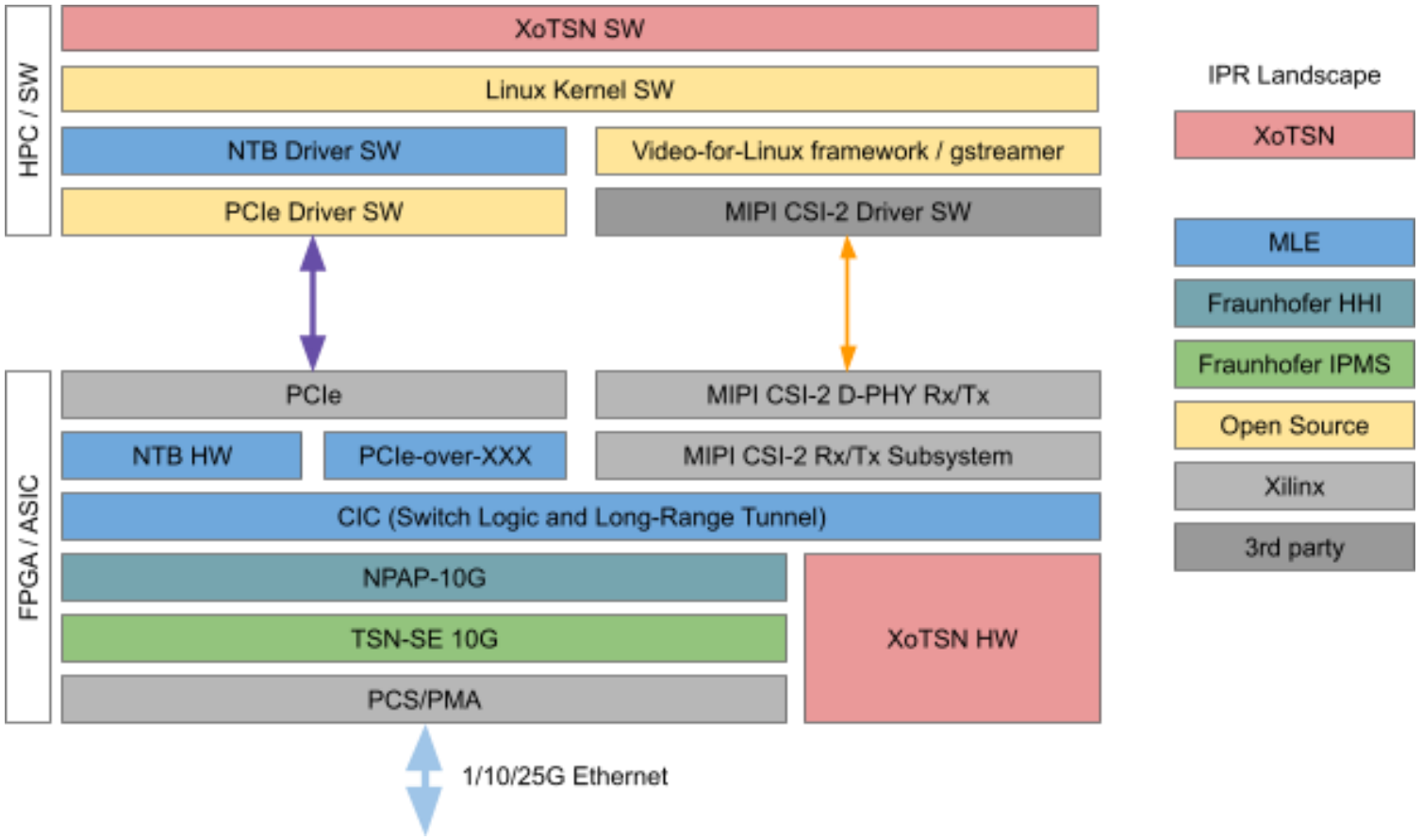
PCIe-over-TSN System Stack Example

System Stack is

- Hardware (Digital Circuit)
- Software (Drivers)

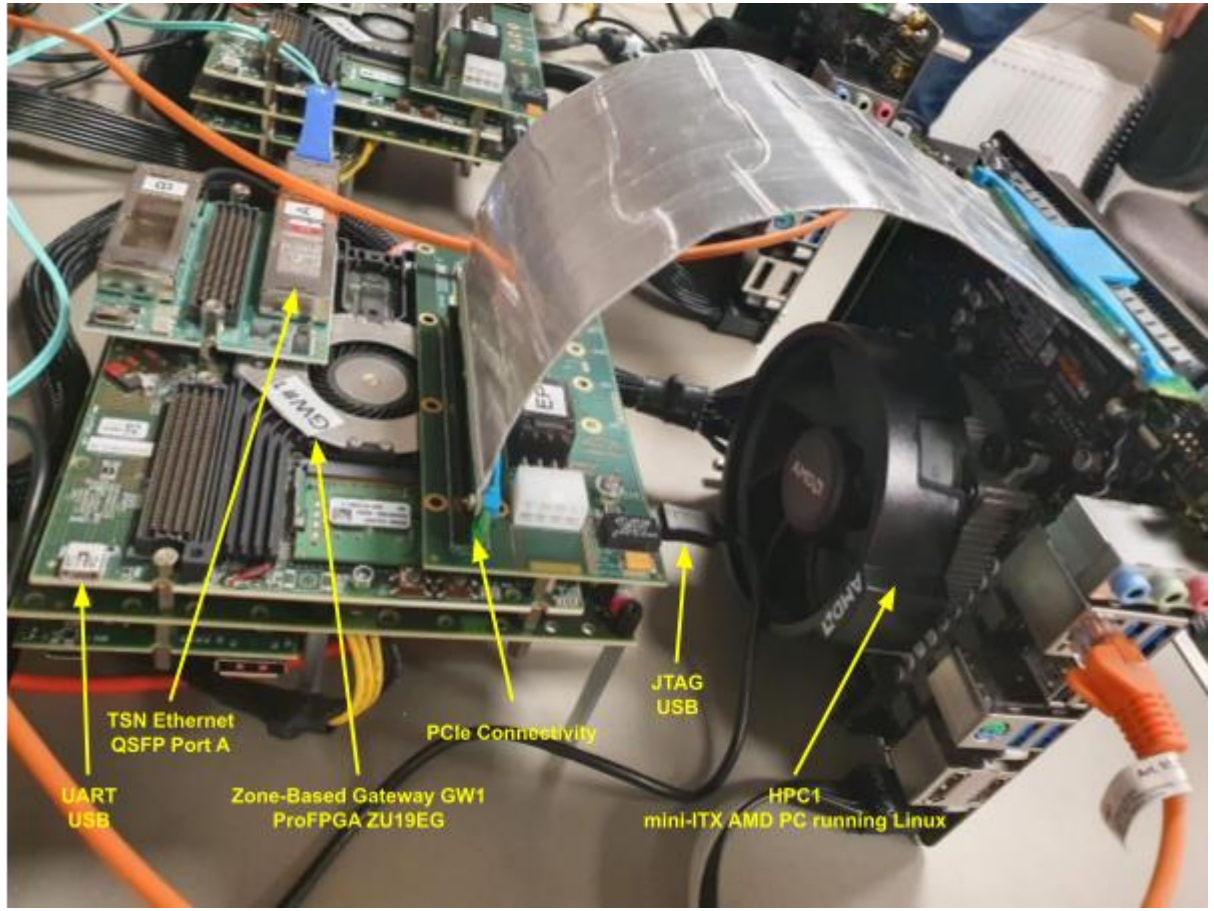
3rd Party (IP Cores)

- Fraunhofer IPMS
- Fraunhofer HHI
- MLE
- Open Source



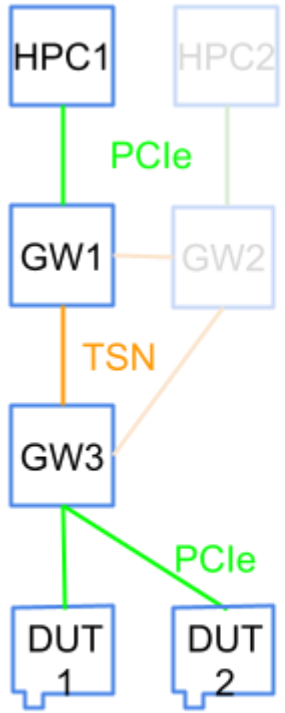
PCIe-over-TSN System Stack Implemented

Labcar Setup w/ PCIe Connect to HPC

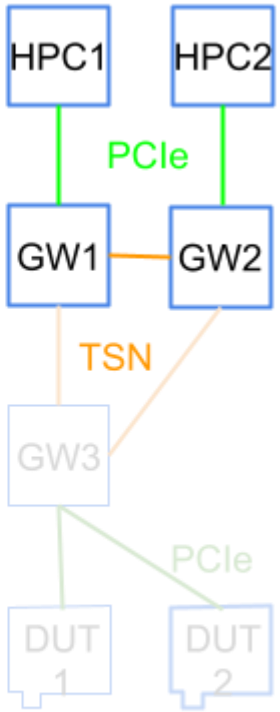


Labcar Setup for Experiments

PCIe/NVMe



PCIE NTB



PCIe = 8GT/s x4

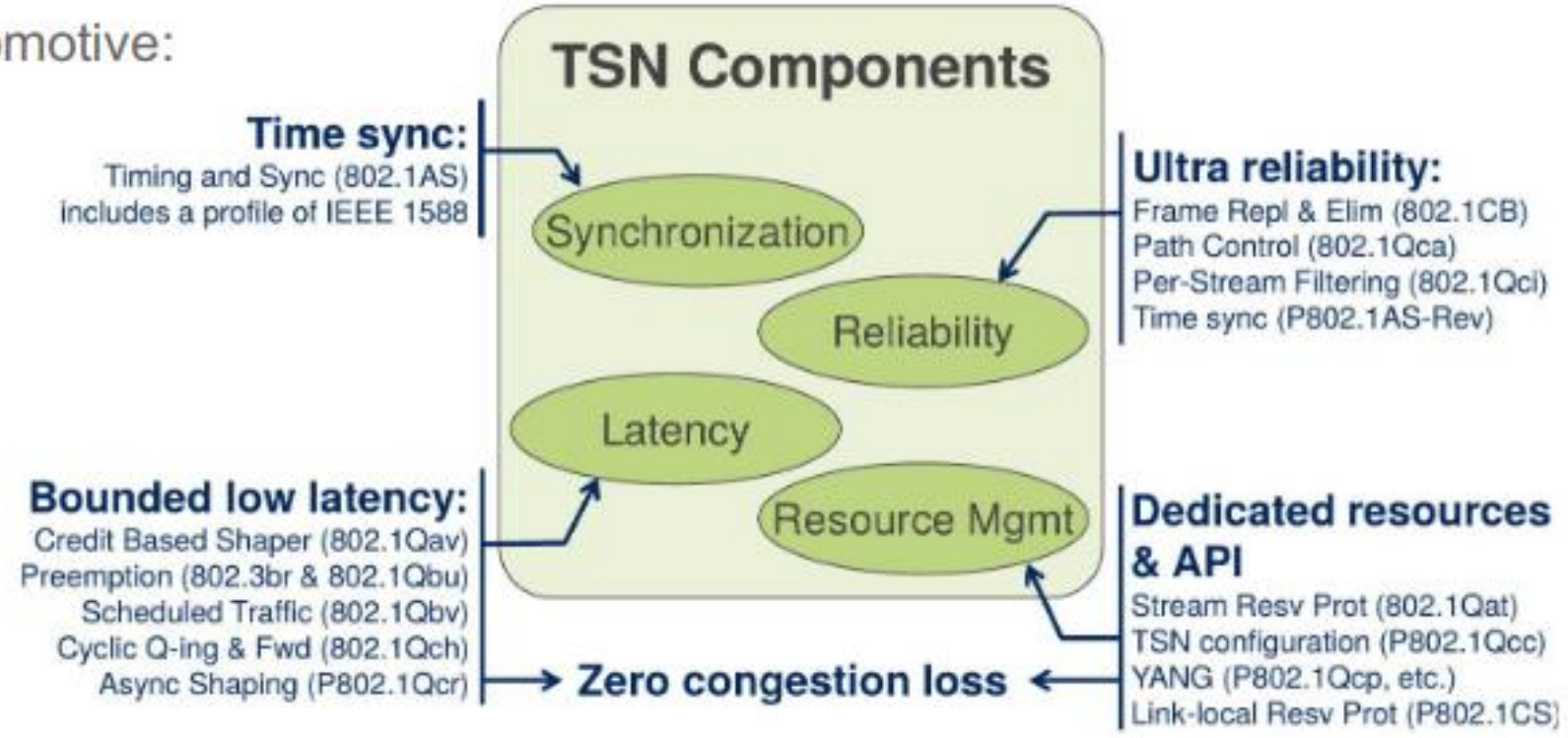
TSN = 1GE and 10GE

Backgrounder TSN

TSN is Time Sensitive Networking which is a collection of IEEE Standards for (Layer 2) Ethernet to deliver reliability and bounded latency

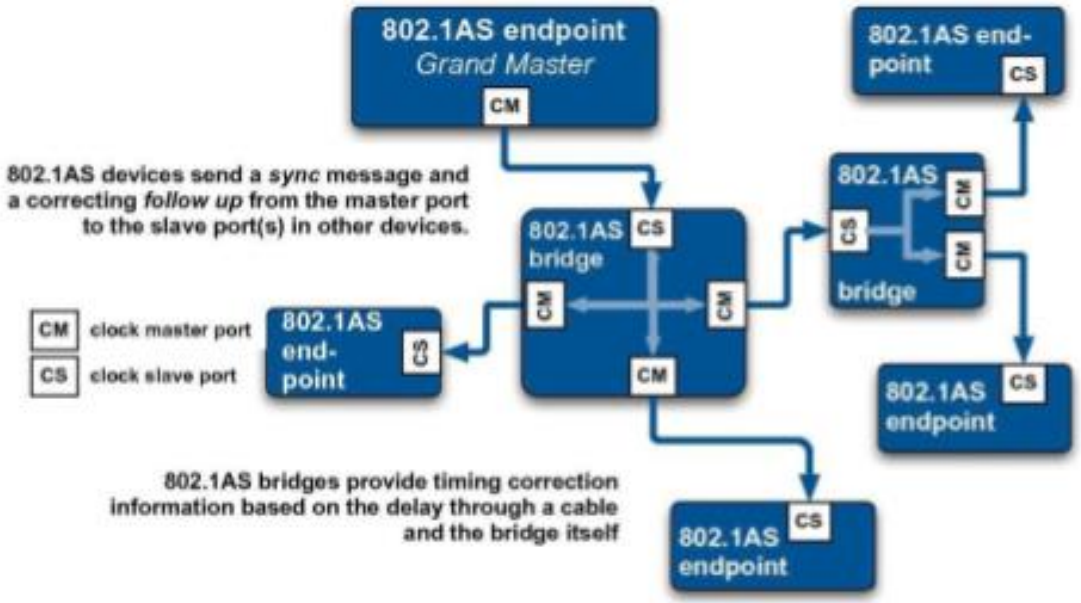
Important for automotive:

- 802.1AS
- 802.1Qbv
- 802.1CB

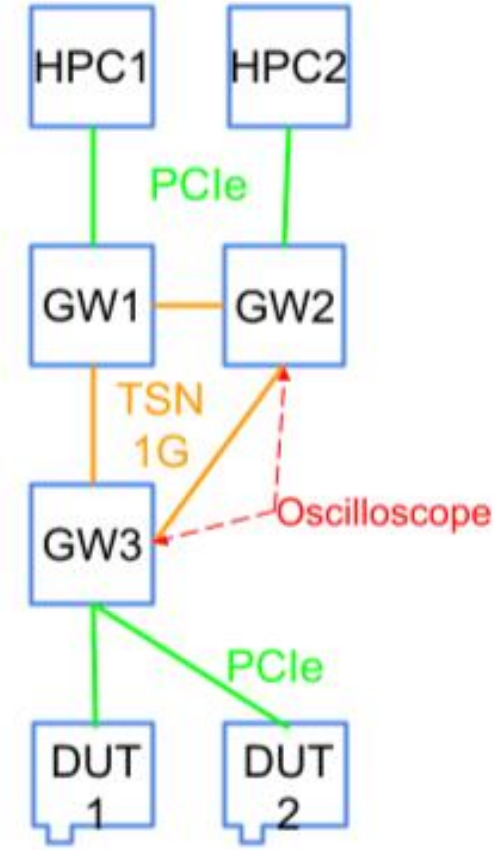


IEEE 802.1AS Precision Timestamping

Concept of Time Synchronization

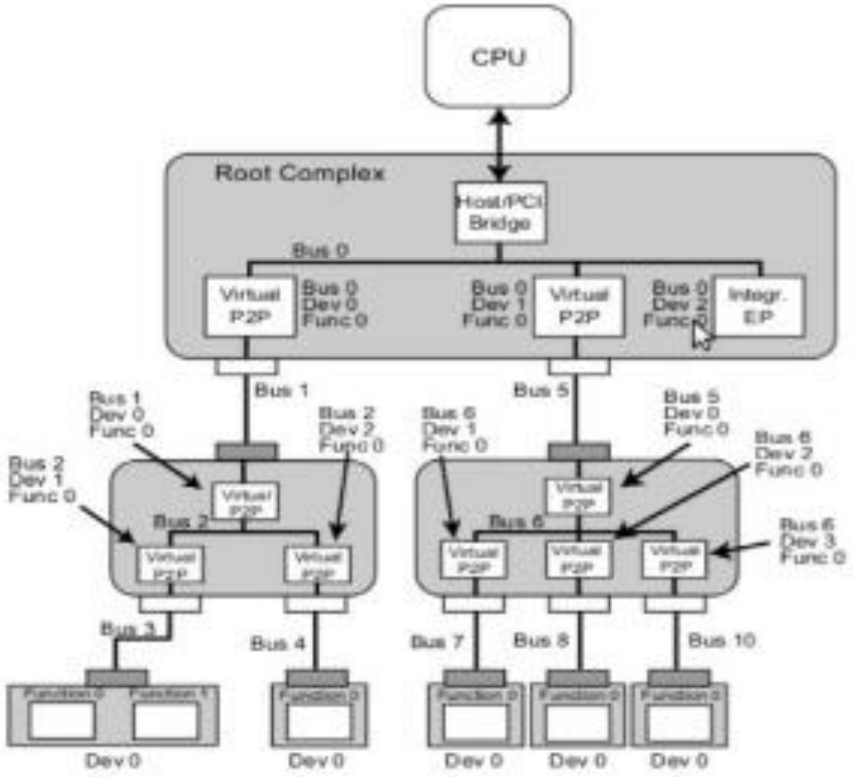


Outcome measured with Oscilloscope is ~22 nanoseconds in FPGA (better in ASIC)



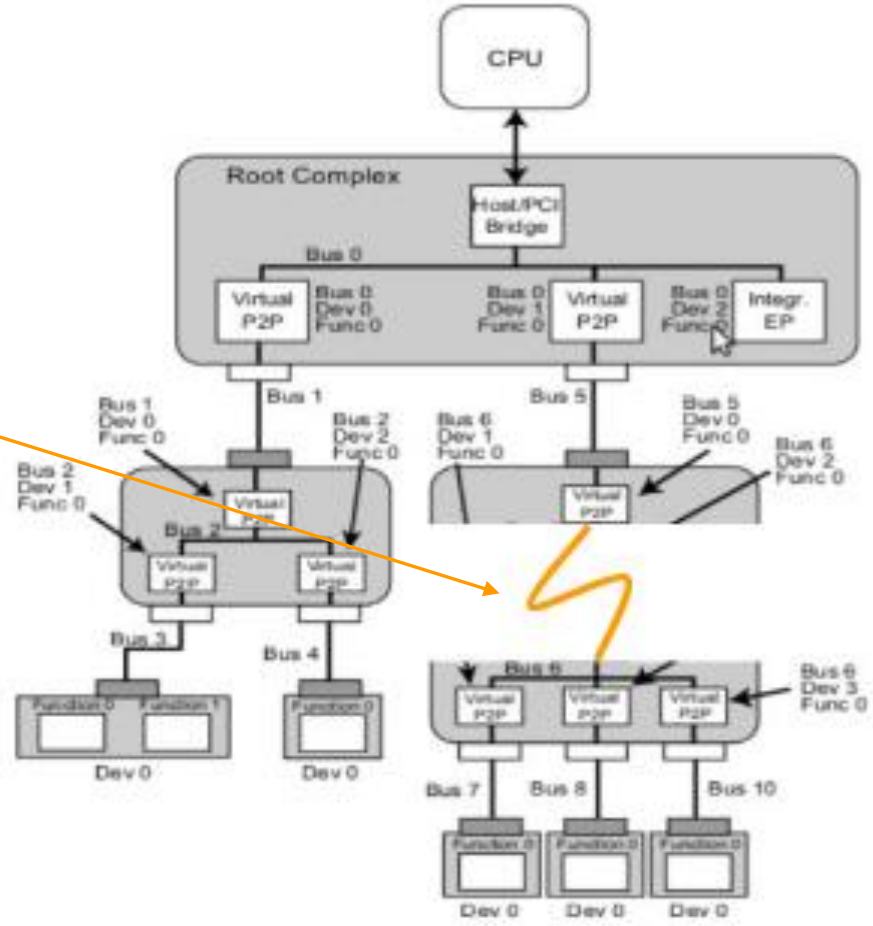
PCIe-over-TSN = “Distributed PCIe Switch”

PCIe Hierarchy with PCIe Switches



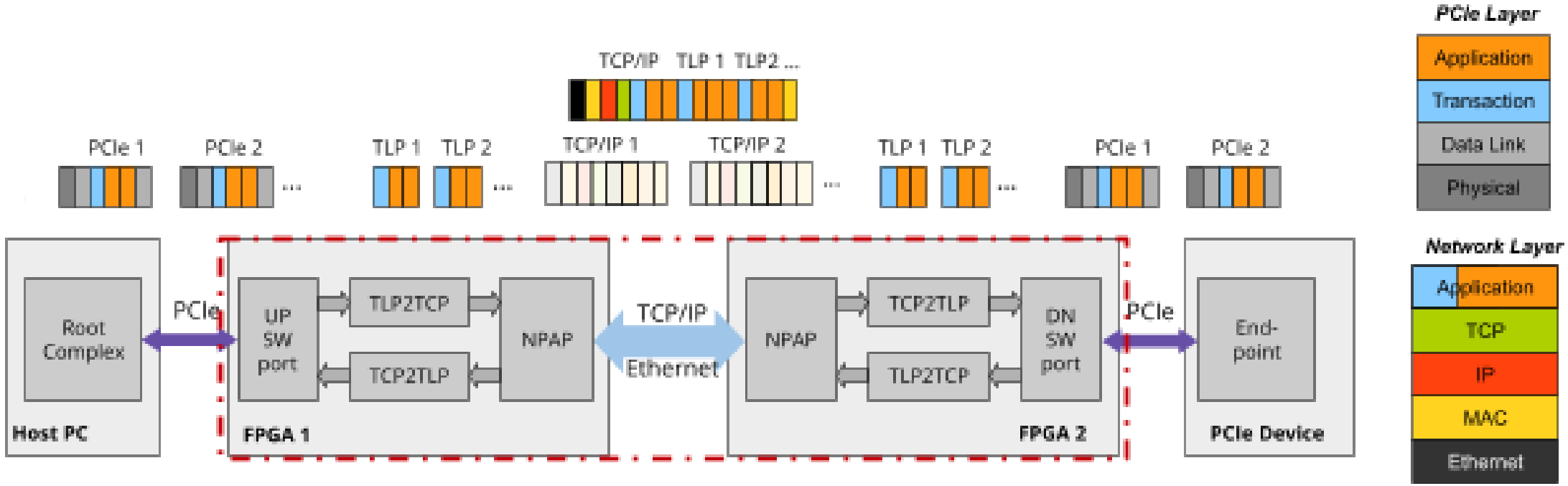
TSN

PCIe Long Range Tunnel “cuts open” PCIe Switch



PCIe-over-TSN Concept

Encapsulate and Decapsulate PCIe TLPs. PCIe demands reliability, therefore we transport TLPs over TCP/IP over TSN over Ethernet.

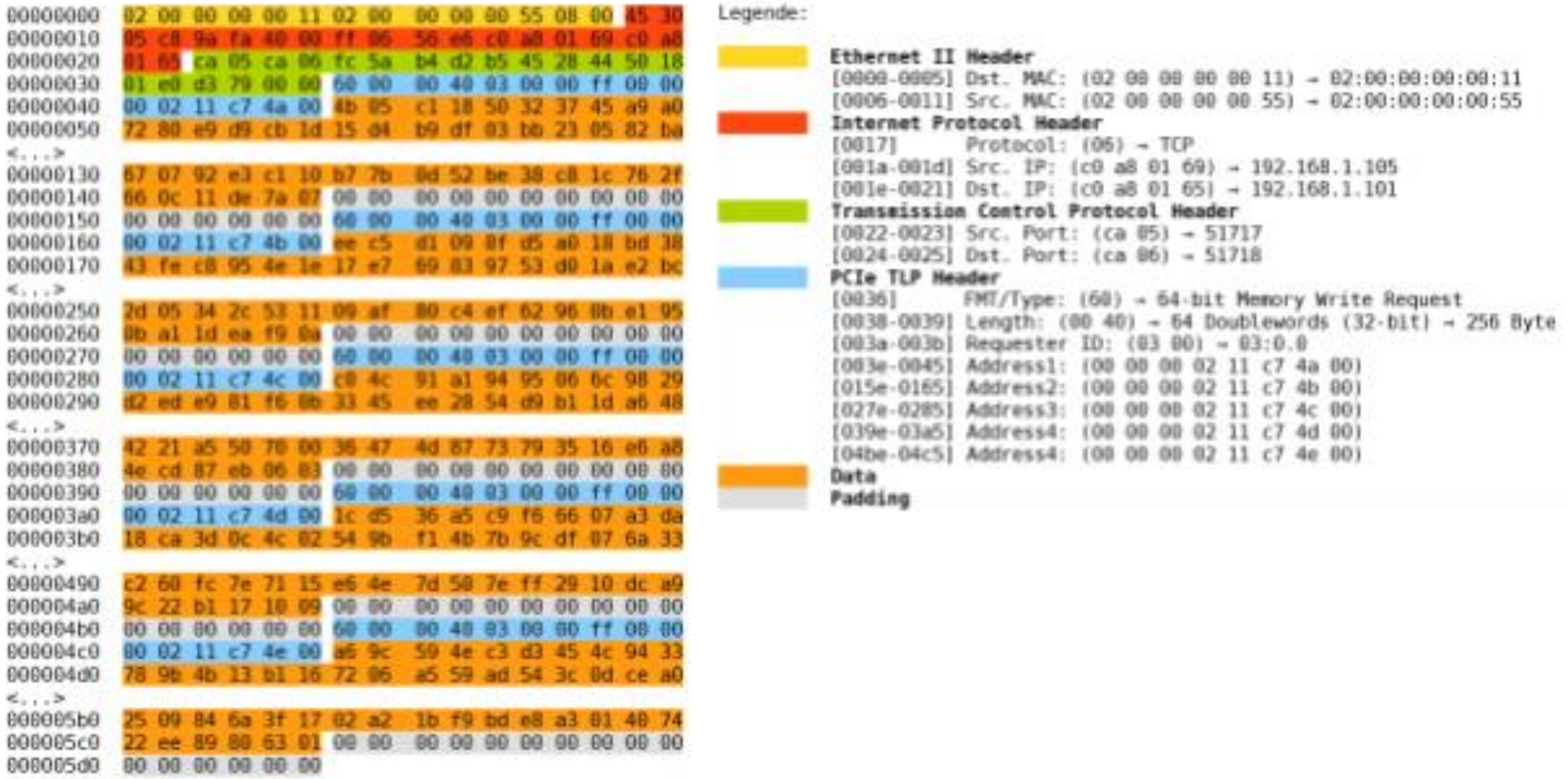


PCIe-over-TSN Smart TLP Aggregation

Aggregate multiple TLPs to optimize bandwidth dominated communication

Do not aggregate (i.e. send immediately) for latency oriented communication

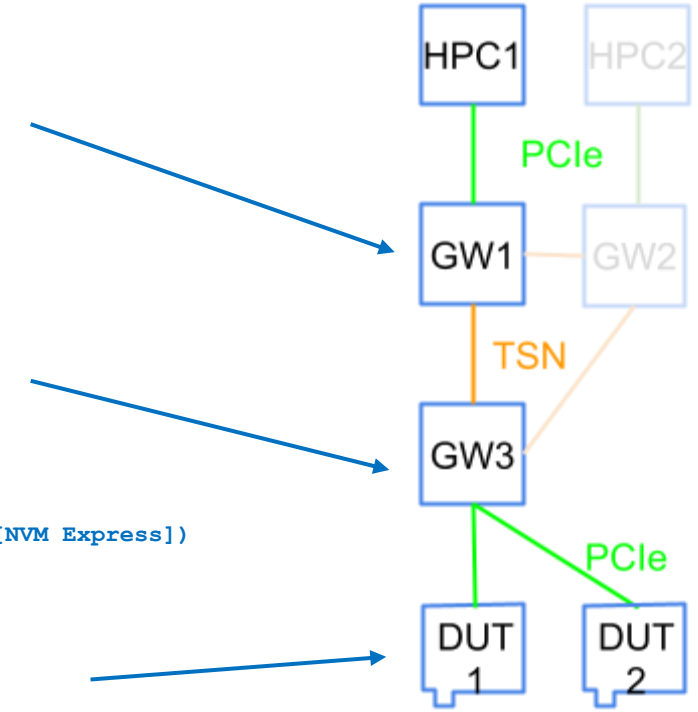
Obey PCIe ordering rules by handling TLPs “as a stream”



PCIe-over-TSN Results with NVMe SSD

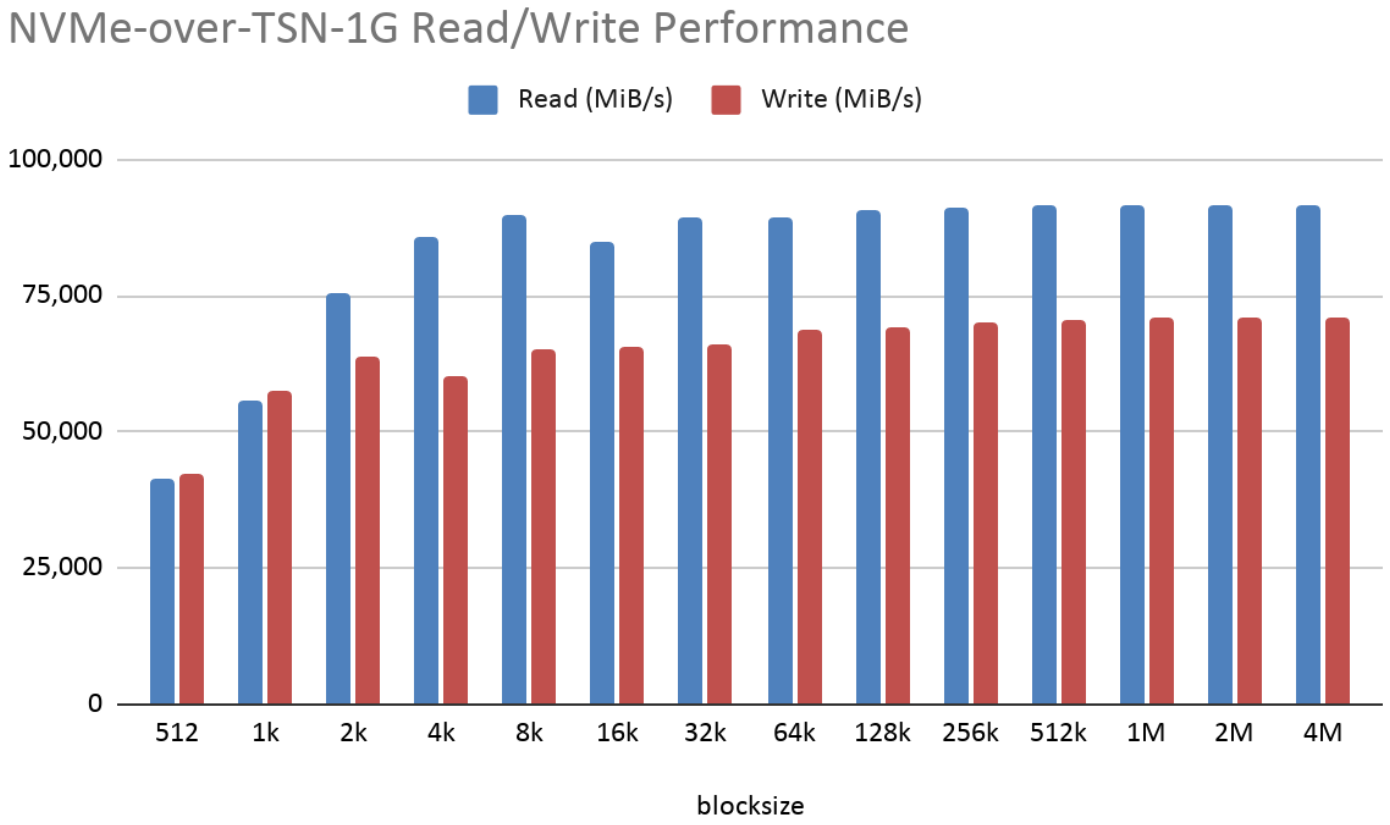
Linux lspci

```
00:00.0 Host bridge [0600]: Advanced Micro Devices, Inc. [AMD] Device [1022:15d0]
00:01.0 Host bridge [0600]: Advanced Micro Devices, Inc. [AMD] Device [1022:1452]

01:00.0 PCI bridge [0604]: Xilinx Corporation Device [10ee:9034] (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr- Stepping- SERR- FastB2B- DisINTx-
Status: Cap+ 66MHz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR- INTx-
Latency: 0, Cache Line Size: 64 bytes
Bus: primary=01, secondary=02, subordinate=03, sec-latency=0
Capabilities: [70] Express (v2) Upstream Port, MSI 00
Capabilities: [1c0 v1] #19
Kernel driver in use: pcieport
Kernel modules: shpchp

02:00.0 PCI bridge [0604]: Xilinx Corporation Device [10ee:9134] (prog-if 00 [Normal decode])
Control: I/O+ Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr- Stepping- SERR- FastB2B- DisINTx-
Status: Cap+ 66MHz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR- INTx-
Latency: 0, Cache Line Size: 64 bytes
Bus: primary=02, secondary=03, subordinate=03, sec-latency=0
Capabilities: [70] Express (v2) Downstream Port (Slot+), MSI 00
Capabilities: [1c0 v1] #19
Kernel driver in use: pcieport
Kernel modules: shpchp

03:00.0 Non-Volatile memory controller [0108]: Samsung Electronics Co Ltd Device [144d:a808] (prog-if 02 [NVM Express])
Subsystem: Samsung Electronics Co Ltd Device [144d:a801]
Control: I/O- Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr- Stepping- SERR- FastB2B- DisINTx+
Status: Cap+ 66MHz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR- <PERR- INTx-
Latency: 0, Cache Line Size: 64 bytes
Interrupt: pin A routed to IRQ 30
NUMA node: 0
Region 0: Memory at fcf00000 (64-bit, non-prefetchable) [size=16K]
Kernel driver in use: nvme
Kernel modules: nvme
```



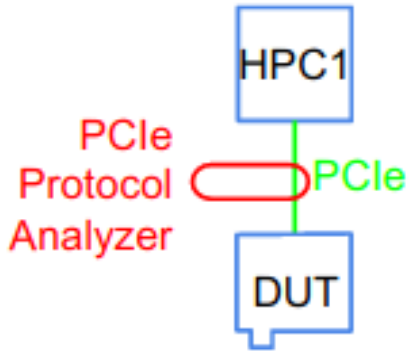
PCIe-over-TSN Robustness

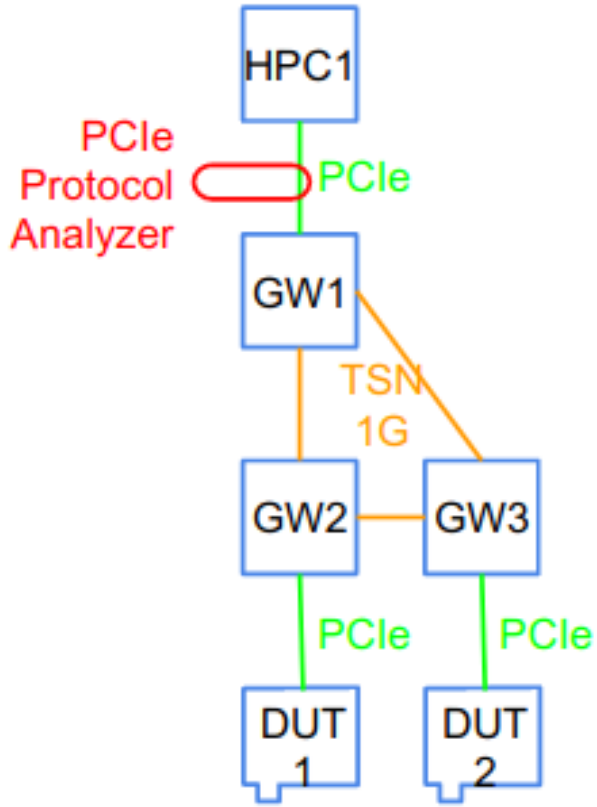
Robust transport, even if TSN “throttles” bandwidth – test results for 1GE

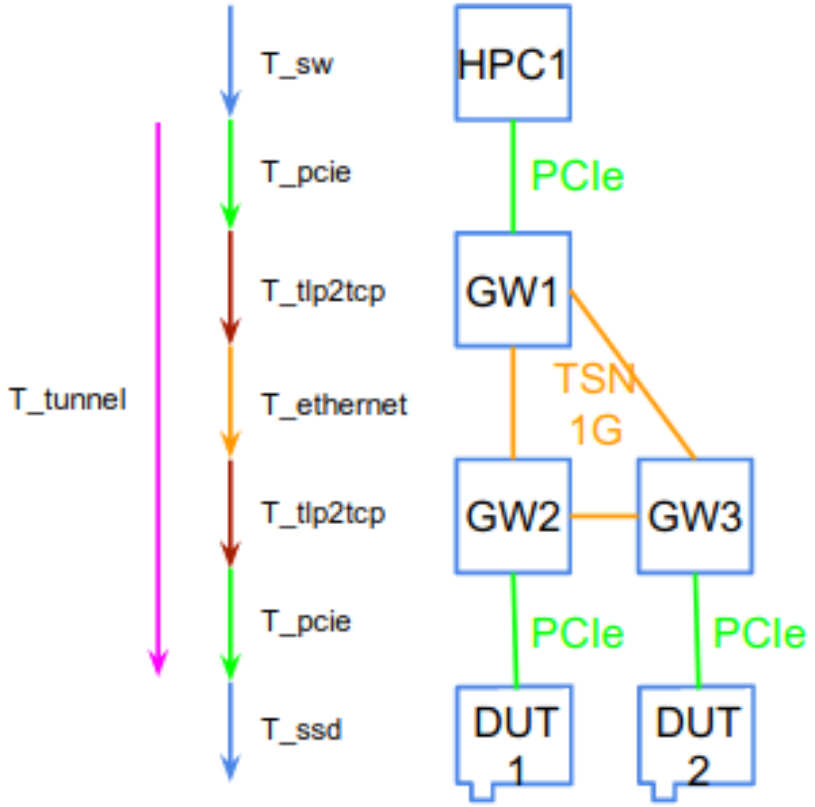


PCIe-over-TSN Latency Analysis

Compare PCIe Latency for a directly connected SSD against the “tunneled”
Use LeCroy PCIe Protocol Analyzer for PCIe Configuration TLP

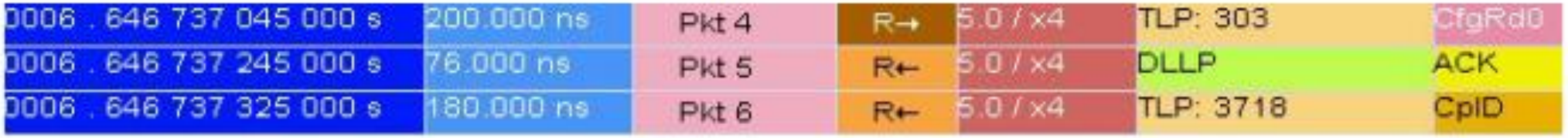




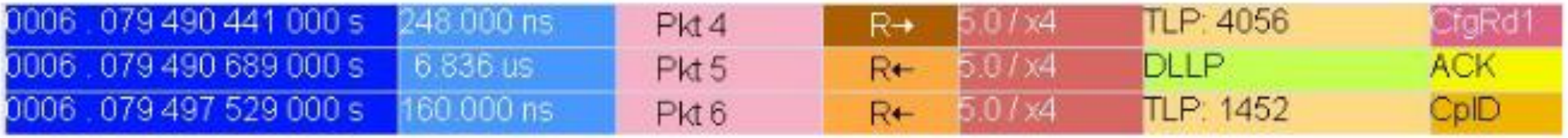


PCIe-over-TSN Latency – Results for 1 GigE

Direct-attached SSD: Baseline one PCIe Config TLP is 646,737,325 - 646,737,045 = 280 ns



PCIe-over-TSN: Round-trip time for one PCIe Config TLP is 079,497,529 - 079,490,441 = 7,088 ns

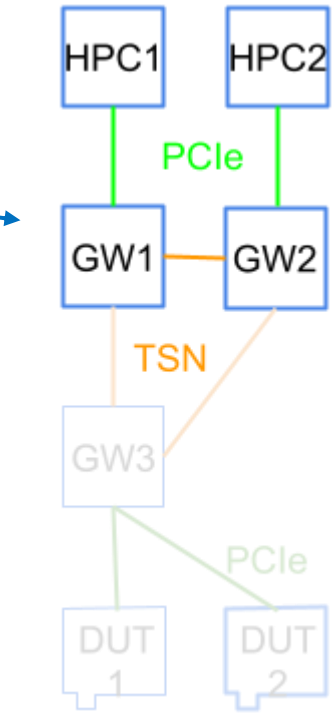


⇒ T_tunnel = (7,088 - 280) / 2 = 3,404 ns

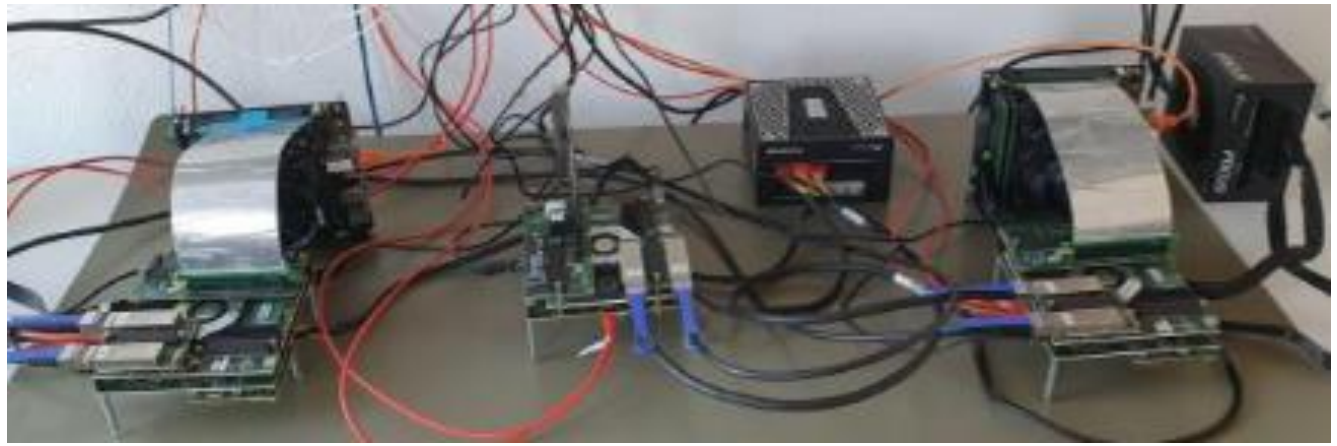
10 Gig Ethernet likely shortens this to < 3,000 ns

PCIe-over-TSN for Non-Transparent Bridging (NTB)

```
> lspci -vt

-[0000:00]-+-00.0  Advanced Micro Devices, Inc. [AMD] Raven/Raven2 Root Complex
           +-00.2  Advanced Micro Devices, Inc. [AMD] Raven/Raven2 IOMMU
           +-01.0  Advanced Micro Devices, Inc. [AMD] Family 17h (Models 00h-1fh) PCIe Dummy
                   Host Bridge
           +-01.1-[01]----00.0  Missing Link Electronics Device 22fb
...


> lspci -vv -s 01:00.0


01:00.0 Memory controller: Missing Link Electronics Device 22fb
        Subsystem: Xilinx Corporation Device 0007
        Control: I/O- Mem+ BusMaster+ SpecCycle- MemWINV- VGASnoop- ParErr- Stepping- SERR-
          FastB2B- DisINTx+
        Status: Cap+ 66MHz- UDF- FastB2B- ParErr- DEVSEL=fast >TAbort- <TAbort- <MAbort- >SERR-
          <PERR- INTx-
        Latency: 0, Cache Line Size: 64 bytes
        Region 0: Memory at d0000000 (64-bit, prefetchable) [size=256M]
        Region 2: Memory at c0000000 (64-bit, prefetchable) [size=256M]
        Region 4: Memory at fcf10000 (32-bit, non-prefetchable) [size=64K]
        Region 5: Memory at fcf00000 (32-bit, non-prefetchable) [size=64K]
        Capabilities: [1c0 v1] #19
        Kernel driver in use: ntb_hw_mle
        Kernel modules: ntb_hw_mle
```



PCIe-over-TSN “Labcar”



Conclusion

- In-vehicle connectivity is quickly evolving, driven by cost/performance and compliance
  - Migration towards centralized so-called Zone-Based Architectures
- PCIe key choice for short-distance local connect between CPUs, GPUs, FPGAs, SoCs
- Our working proof-of-concept underlines feasibility
  - A layered stack of open-standard protocols extends reach of PCIe over real-time Ethernet / Automotive Ethernet – working title PCIe-over-TSN
- Implementation as a digital circuit delivers robust behavior at deterministic, low latencies

Thank you for attending the PCI-SIG Virtual Developers Conference 2021!

For more information about PCIe technology, please go to www.pcisig.com