New IP for Future Vehicular Networking

Lin Han

Futurewei, USA

lin.han@Futurewei.com

Bio: Lin Han

- Lin Han worked in electronic research institute and communication industry for more than 25 years including Southeast University (China), University of Toronto (Canada), New Bridge Network (Canada), Cisco System and Huawei USA. His previous academia research focused on the modeling, simulation and analysis of Electro-Magnetic-Field for device and antenna. His industry career has been with the evolution path for data communication from Circuit switch, TDM/ATM network to IP based Internet.
- He worked in European ETSI ISG NGP (Next-Generation Protocols) project from January 2016 to December 2019 and served as rapporteurs for two working items: “New Transport Technologies” and “Network Layer Multi-Path Support”.
- He has over 20 USA and international patents and several research publications (IEEE/IETF). He is a Principal Engineer at Futurewei Inc. now and focusing on the research for Future Network Technologies. His current interests are in the new architecture, protocols, solutions to solve new problems emerged in 5G and beyond.

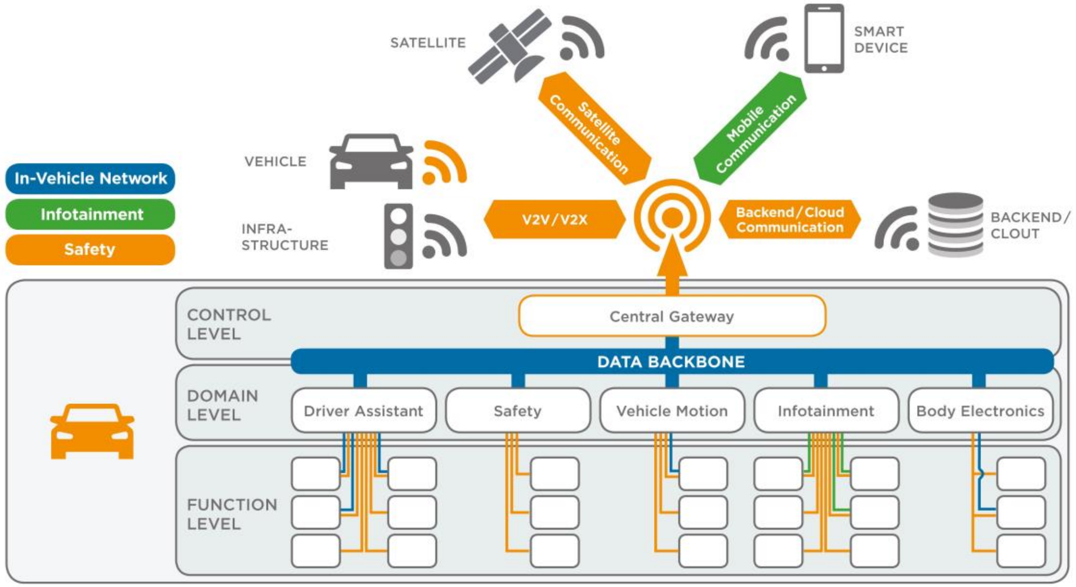
Agenda

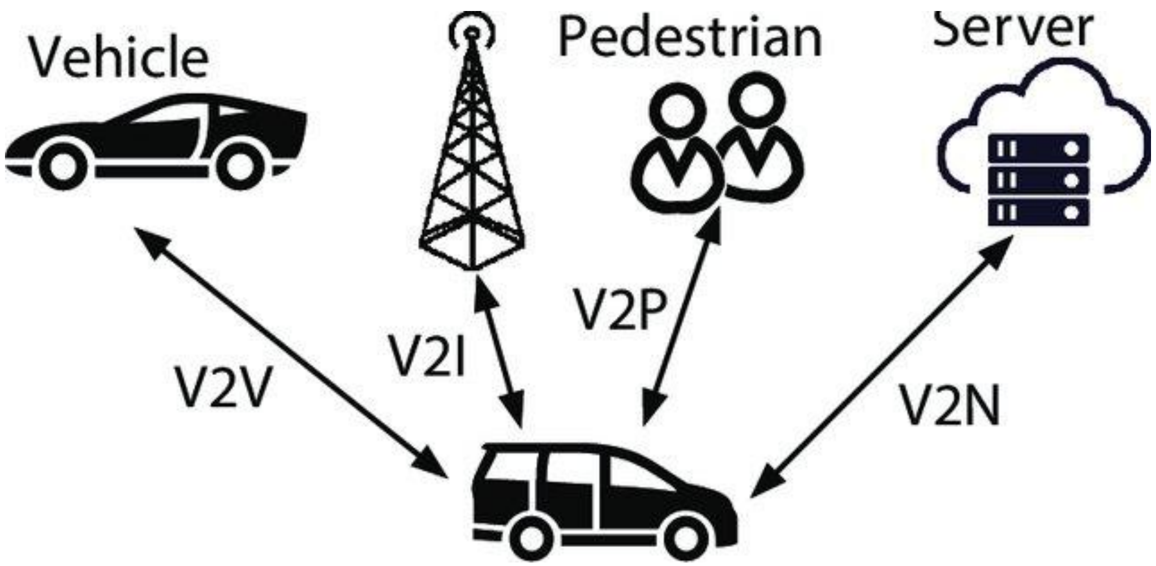
- **Trends of Vehicle Evolution**
- New IP Overview
  - New IP Introduction
  - New IP for New Service: Goal and Key Technologies
- Use case for In-Vehicle-Network (IVN)
  - Review of Current IVN Technologies
  - New IP Based IVN
- Use case for V2X
  - Latency Analysis
  - New IP for True E2E uRLLC for 5G and Beyond

Trends of Vehicle Evolution

FUTUREWEI Technologies

- Power (EV), Software (Self-Driving) and Networking (V2X)





Source: "6 Key Connectivity Requirements of Autonomous Driving“ https://spectrum.ieee.org/transportation/advanced-cars/6-key-connectivity-requirements-of-autonomous-driving

•**Source:** **"**An Overview of 3GPP Cellular Vehicle-to-Everything Standards**".** November 2017. GetMobile Mobile Computing and Communications 21(3):19-25

Agenda

- Trends of Vehicle Evolution
- **New IP Overview**
  - **New IP Introduction**
  - New IP for New Service: Goal and Key Technologies
- Use case for In-Vehicle-Network (IVN)
  - Review of Current IVN Technologies
  - New IP Based IVN
- Use case for V2X
  - Latency Analysis
  - New IP for True E2E uRLLC for 5G and Beyond

New IP Introduction

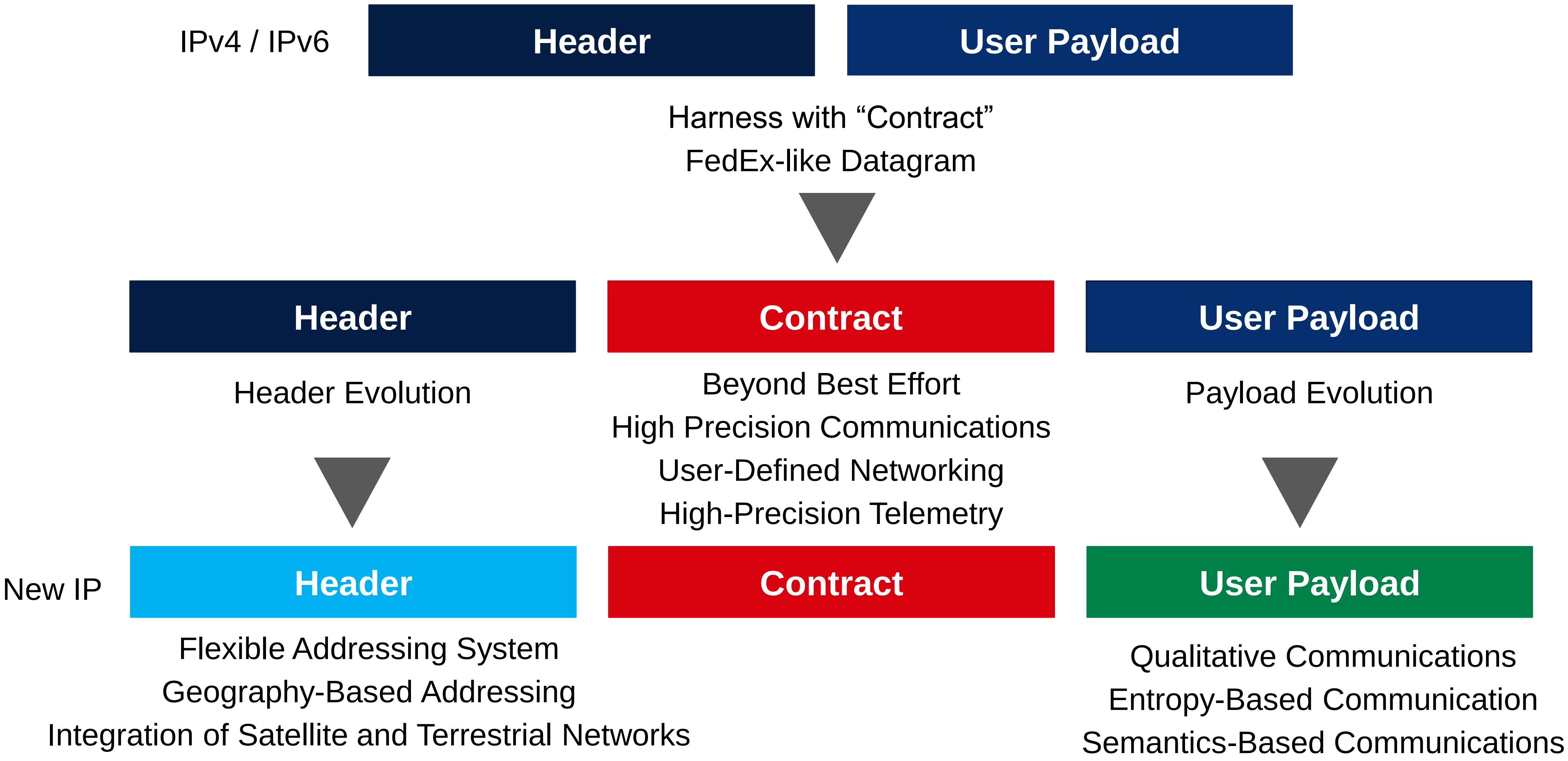
- Richard Li's presentation in Globecom 2019

https://globecom2019.ieee-globecom.org/program/keynotes

- Huawei's presentation in ITU

https://www.itu.int/en/ITU-T/Workshops-and-Seminars/2019101416/Documents/Sheng_Jiang_Presentation.pdf

New IP: Evolution Map



New Features – Open and Extensible

FUTUREWEI Technologies

NewIP Addr	NewIP Contract	User Payload

Variable length and heterogeneous address

New Features
- Service oriented routing
- Interconnect heterogeneous network
- Dynamic and auditable anonymous IP and ID
- Decentralized ID-based Key
- Authenticity based on minimum trust model
- Joint Inter-AS auditing and attack prevention

Advantages
- No more address type and size issue in the future
- Balanced privacy and security
- Balanced network sovereignty and openness
- Guaranteed integrity and confidentiality
- Guaranteed E2E security and privacy
- Routing is more than achievability

Contract and Meta Data

New Features
- Guaranteed E2E service for different QoS Metrics
- Improved user-network interface
- User-defined network behavior
- Ultra-high network throughput
- Concurrent multipath network support
- Network coding-based flow control
- New transport protocol stack

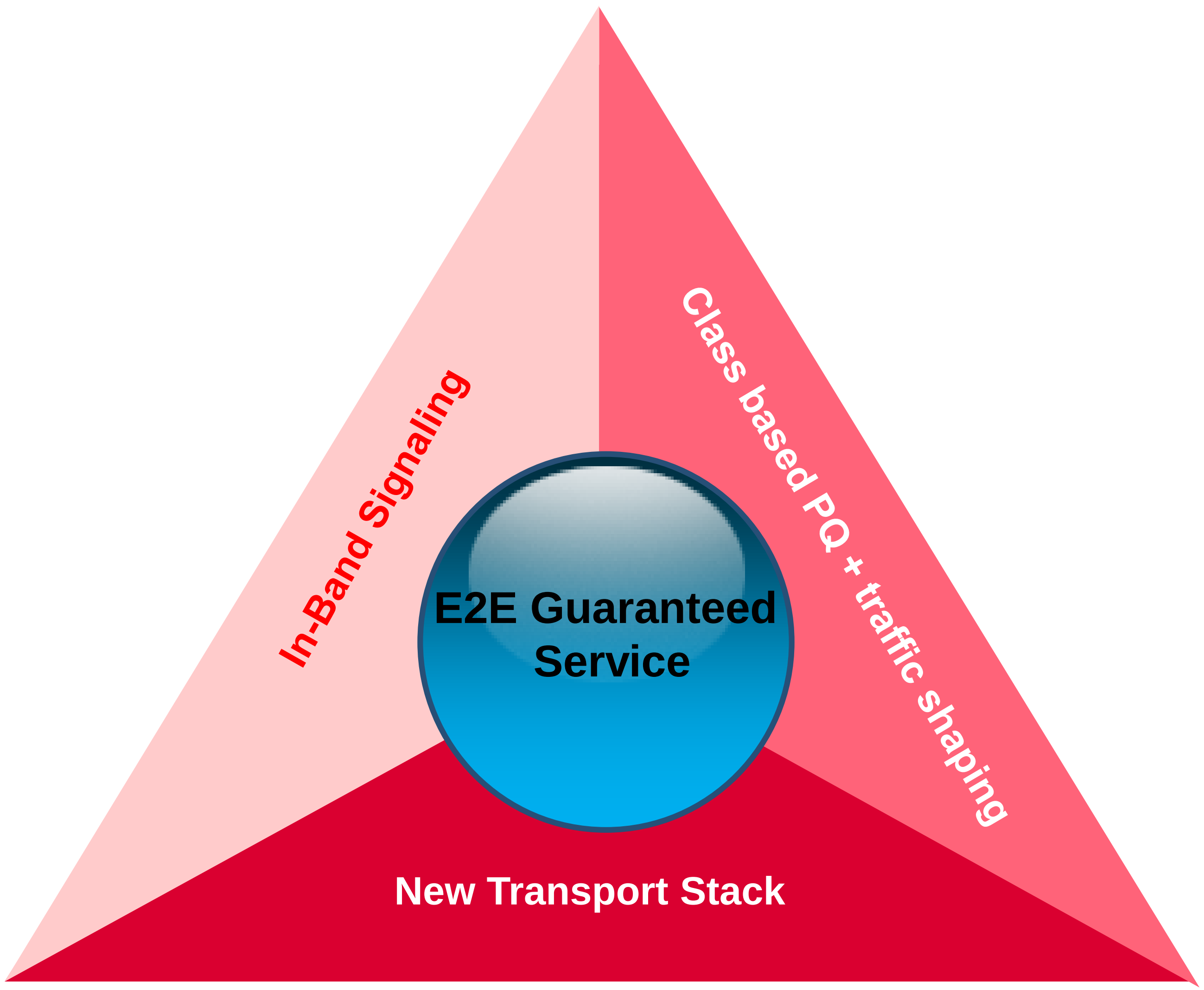
Advantages
- More service directly access-able to end user APP
- APP is aware of the network state and service and select on-demand
- True E2E uRLLC for 5G and beyond
- Industry control can be over Internet finally
- Congestion free and lossless to critical communication

- Not standardized yet
- IPv6 can use extension header
- IPv4 has to extend IPv4 first

Agenda

- Trends of Vehicle Evolution
- **New IP Overview**
  - – New IP Introduction
  - – **New IP for New Service: Goal and Key Technologies**
- Use case for In-Vehicle-Network (IVN)
  - – Review of Current IVN Technologies
  - – New IP Based IVN
- Use case for V2X
  - – Latency Analysis
  - – New IP for True E2E uRLLC for 5G and Beyond

New IP for New Service: Goal and Key Technologies



Goal: E2E Guaranteed Service in Internet

- E2E
  - From one end-device's APP to another end-device's APP
- Guaranteed Service
  - Granularity can be as fine as IP flow level, i.e. defined by 5 tuples
  - Guaranteed Service to satisfy SLA (Service Level Agreement)
  - BGS: Bandwidth Guaranteed on all device on the path
  - LGS: E2E Latency Guaranteed (also guarantee the bandwidth).
    - In-Time: The maximum E2E latency is below a given upper limit
    - On-Time: The maximum E2E latency and jitter is below a given upper limit
  - Lossless: Not packet drop (caused by queueing and congestion) for the guaranteed flow
  - Congestion-less: No congestion for the guaranteed flow
- Internet
  - 1st step is to support a closed network
  - Gradually adopted.

In-band Signaling - Overview

FUTUREWEI Technologies

In-band Signaling
Control msg carried in IP data

Purpose
- New service setup
- Admission Control
- Dynamic state management for the new service

What to Signal
- Between user device and network device and content provider
- Service expectation from user end-device's APP to Network device and Content provider
- Provisioning state and offered service from Network Device and Content Provider to user's APP
- Service state event or report from Network and Content Provider to user's APP

Simplicity and Effectiveness
- Simplified control mechanism based on data flow loop
- Good enough to support flow level QoS w/o using RSVP
- Advanced features (protection, non-shortest path, etc) can be enhanced independently

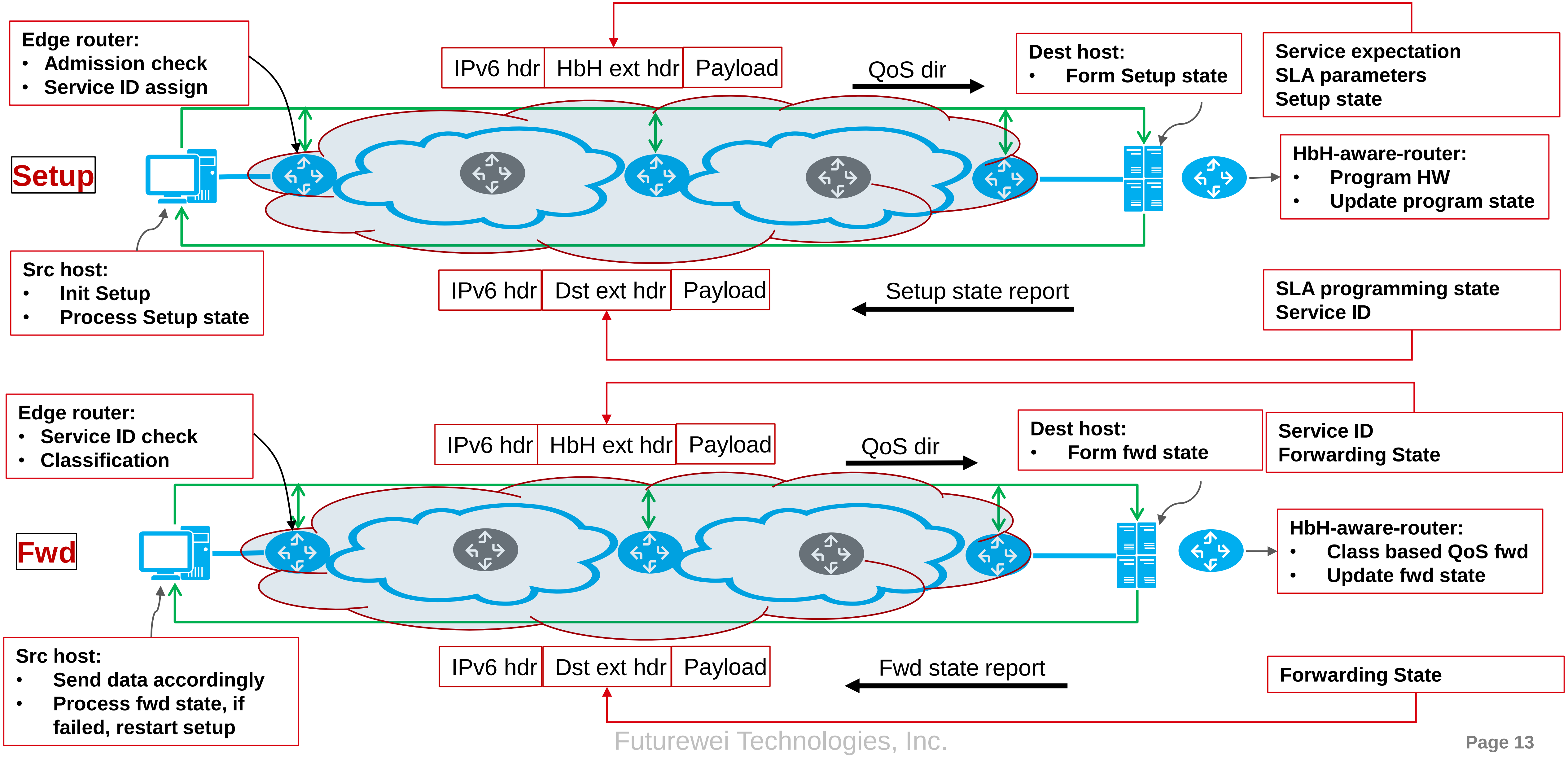
Scalability and Performance
- No control protocol is running, no signaling storm
- Signaling is with data, E2E and crossing domain
- Signaling can be once or periodical
- Light overhead in forwarding and TCAM consumption
- Data flow functions as path refreshment
- Entry-auto-expire for resource releasing

Agnostics and Compatibility
- Independent of upper layers
- Not impacted by user data encryption and Ipsec
- Backward compatible
- Coexist with all other services

Other Advantages
- Essential for E2E service, control, accounting and billing (may not be free service)
- Potential for other features (in-band OAM, new congestion control, traffic control, etc)

Critical to Queuing and Scheduling
- Make the admission control and resource dynamic mgmt. simple, economical and efficient
- Key parameters calculation for queuing and scheduling
  - Aggregation rate for LGS class for latency estimation when using SPQ
  - W(Weight) calculation for BGS, Afxy and BE classes when using DWRR

In-band Signaling – Example for IPv6



Class Base Priority Queueing + Traffic Shaping

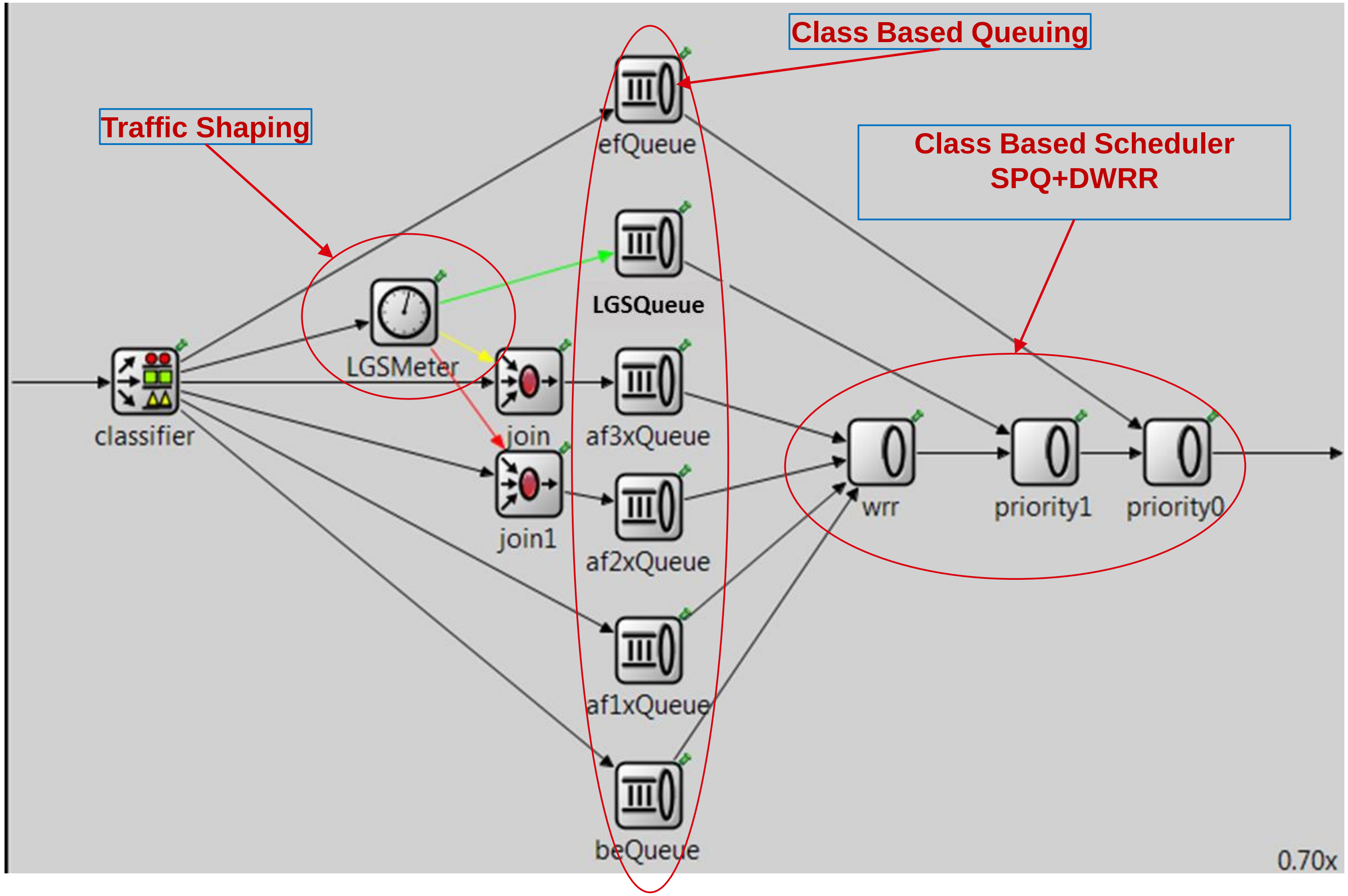
Class Base Priority Queueing

- Packets are grouped, processed and scheduled by its class
- DSCP values
  - Reuse existing DSCP
  - New values LGS, BGS for backward compatibility
  - Mapping to other protocols
    - MPLS: Exp (3bits)
    - 802.1Q: User Priority (3bits)
    - OTN: OSU (still in research)
- Strict Priority Queueing
  - Latency guaranteed flows are in 1st or 2nd (if EF is used) highest priority queue.
  - Bandwidth guaranteed flows
    - In a dedicated queue
    - in lowest priority with BE and use WFQ (Weighted Fair Queueing)

Traffic Shaping

- Purpose
  - Traffic conformity checking
  - Traffic policing for flow or class
- Different algorithm can be used
  - Two rate three color; single rate three color
  - Other algorithms
- Deployment
  - Ingress and/or Egress
- Granularity
  - Per flow and/or per class

Example

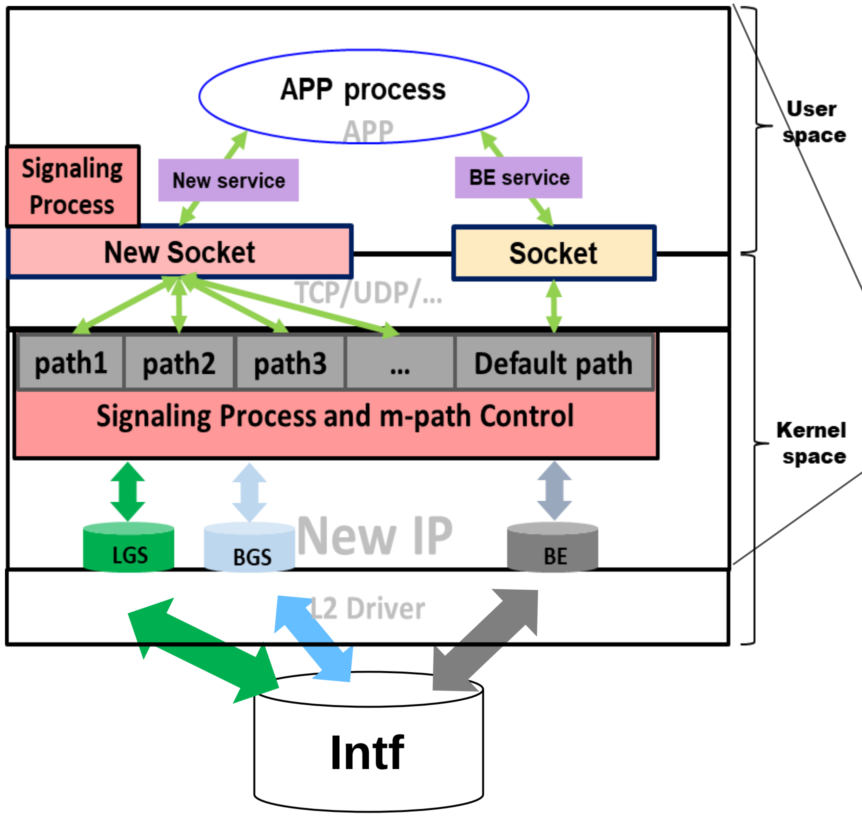

Class Based Queuing
Traffic Shaping
Class Based Scheduler
SPQ+DWRR
efQueue
LGSQueue
LGSMeter
classifier
join
af3xQueue
join1
af2xQueue
af1xQueue
beQueue
wrr
priority1
priority0
0.70x

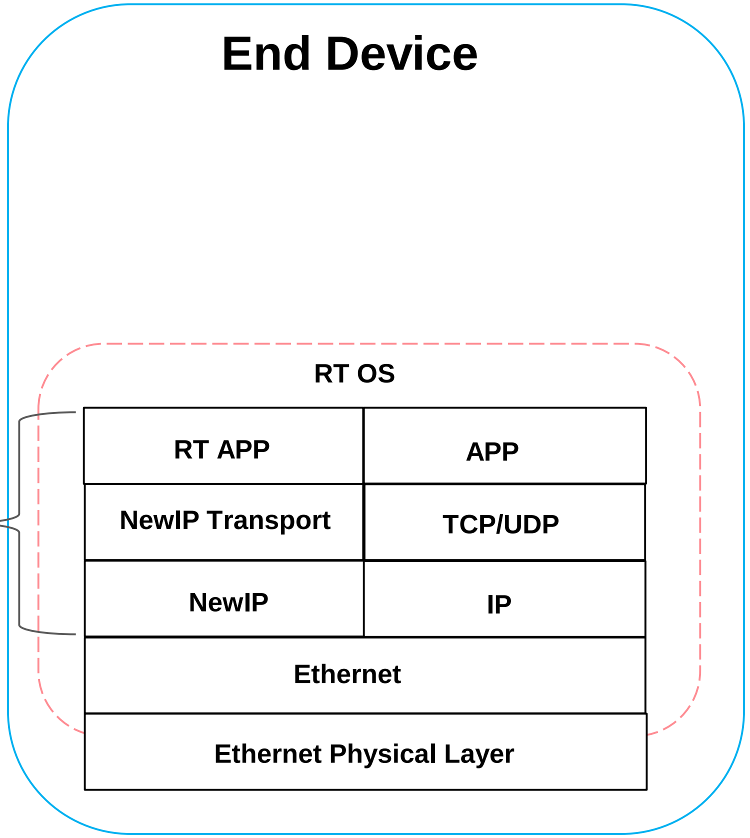
New Transport Protocol Stack

- On end-user device
- Make the traditional transport protocol TCP/UDP use the new service
- Control stack for in-band signaling
  - Kernel space, for protocol that control is in kernel space, like TCP/UDP
  - User space, for protocol that control is in user space, like QUIC
  - User space can override the kernel space
- New User Socket API
  - APP expectation
  - Service offered from network device and content provider
  - Path property
  - Network state (link/node/QoS forwarding, etc)

NewIP protocol support in host

- RT OS
- New service interface
- NewIP transport protocol
- NewIP socket
- Signaling Process in both kernel and user space
- M-path control
- Differentiated queuing and scheduler

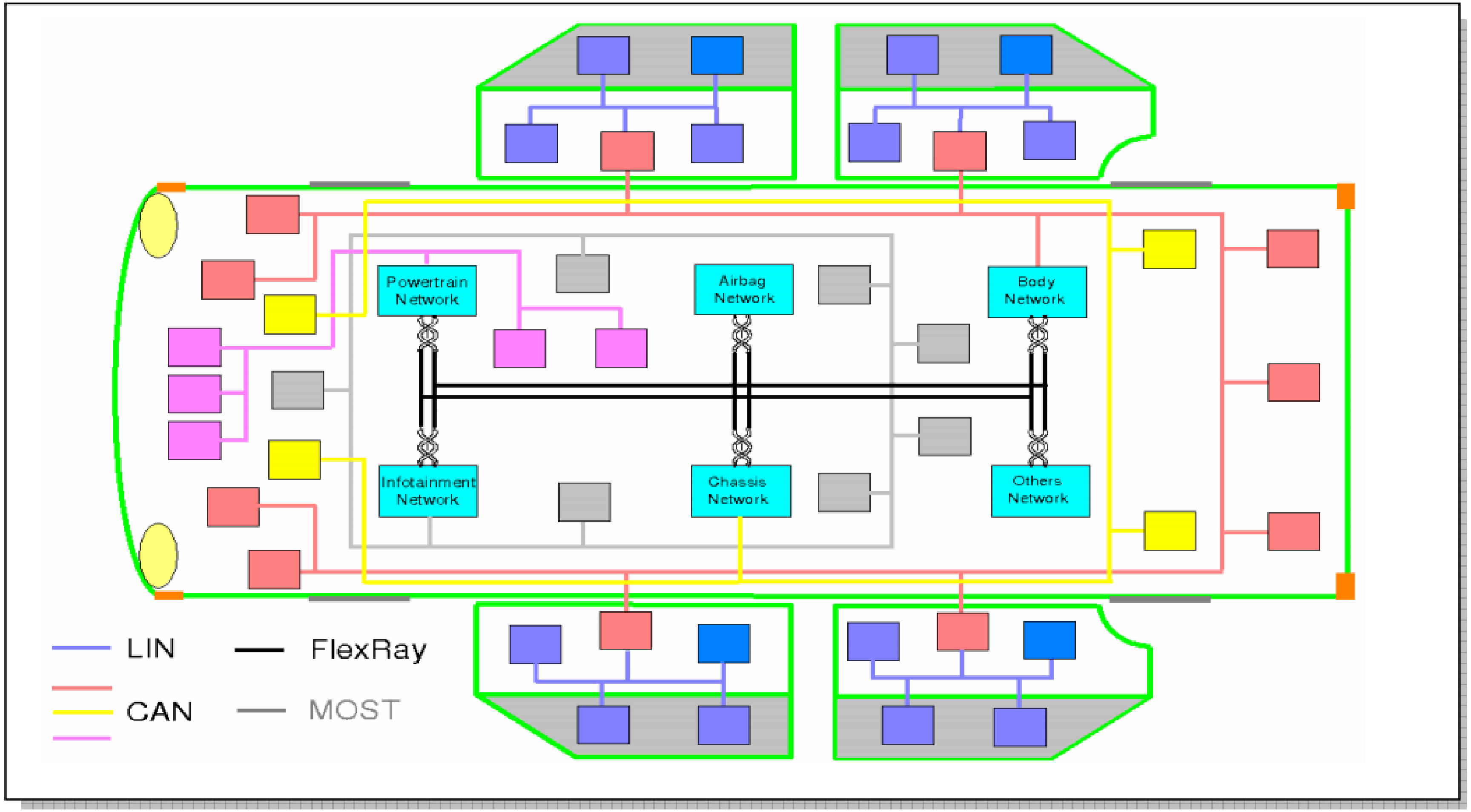




Agenda

- Trends of Vehicle Evolution
- New IP Overview
  - New IP Introduction
  - New IP for New Service: Goal and Key Technologies
- **Use case for In-Vehicle-Network (IVN)**
  - **Review of Current IVN Technologies**
  - New IP Based IVN
- Use case for V2X
  - Latency Analysis
  - New IP for True E2E uRLLC for 5G and Beyond

Current In-Vehicle Network



Protocols Comparison

BUS	LIN	CAN	FLEXRAY	MOST
Cost/Node [$]	1.50	3.00	6.00	4.00
Used in	Subnets	Soft real-time	Hard real-time	Multimedia
Applications	Body	Chassis, Powertrain	Chassis, Powertrain	Multimedia, Telematics
Message transmission	Synchronous	Asynchronous	Synchronous & Asynchronous	Asynchronous & Synchronous
Data rate	20 kbps	1 Mbps	10 Mbps	24 Mbps
Physical layer	Single Wire	Dual-Wire	Dual-Wire (Optical-Fiber)	Optical-Fiber (Dual-Wire)
Latency jitter	Constant	Load dependent	Constant	Data stream
Extensibility	High	High	Low	High

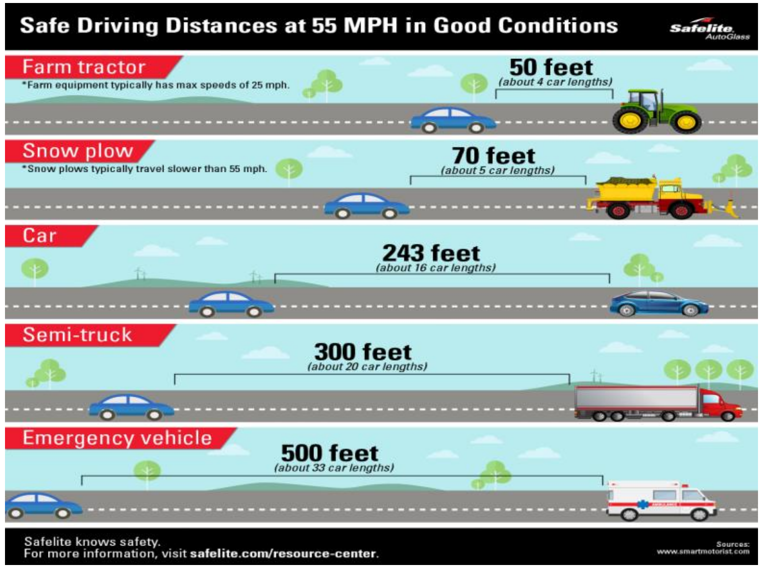
Source: https://vehicle-electronics.biz/content/vehicle-networking-opportunities

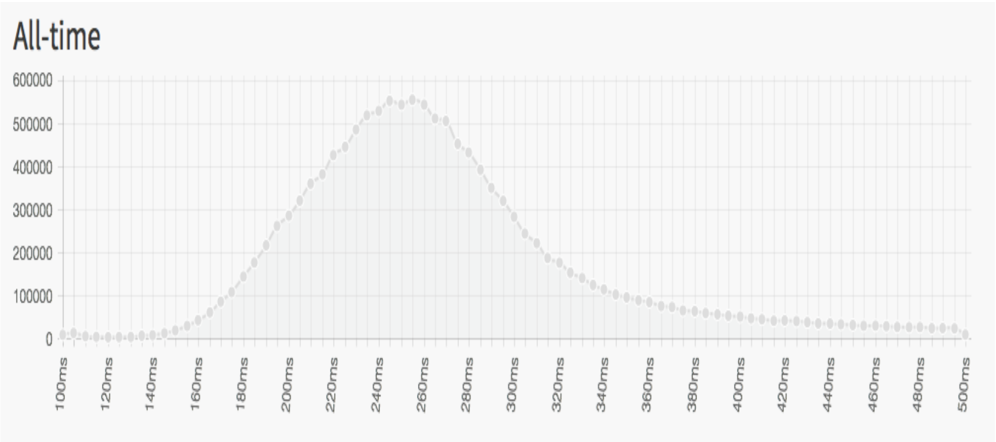
Latency requirement for IVN

- No standard yet
- From the driving safety distance:
  - 55MPH/50FT: 0.613s; 110MPH/50FT: 0.306s
- From Human reaction time: 250ms
- 10us was mentioned in paper: https://ieeexplore.ieee.org/document/8315204, but it does not define it. It should be E2E latency for network perspective;
- Use the fastest FlexRay as example,
  - For minimum FlexRay pack size (9bytes), the serialization latency is 9*8/10M = 7.2us;
  - For max Flexray pack size (262bytes), the serialization latency is 262*8/10M = 209us
- Also use FlexRay's cycle to predict, For real-time perspective, if there is any sporadic msg needs to be sent over FlexRay, the max delay by FlexRay is one cycle (about 1ms)
- On paper: https://ieeexplore.ieee.org/document/5981915

Table 2: Application constraints [9]

Priority Value	Traffic Type	Max. End-to-End Delay [ms]
3	Control	≤ 10 [16, 5]
2	Driver Assistance CAM	≤ 45 [5, 17]
1	Navigation	≤ 100
0	Multimedia	≤ 150 [18]

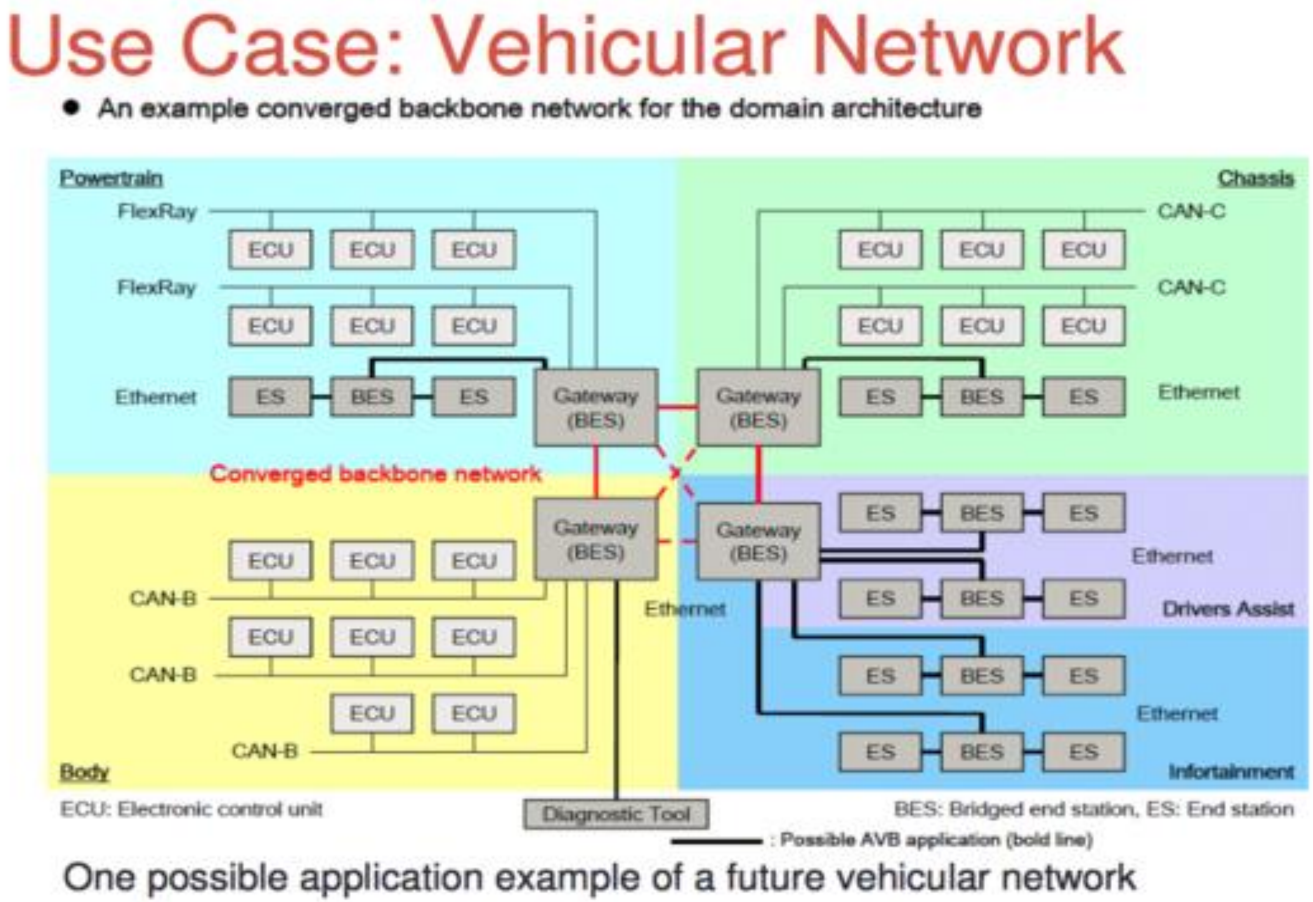




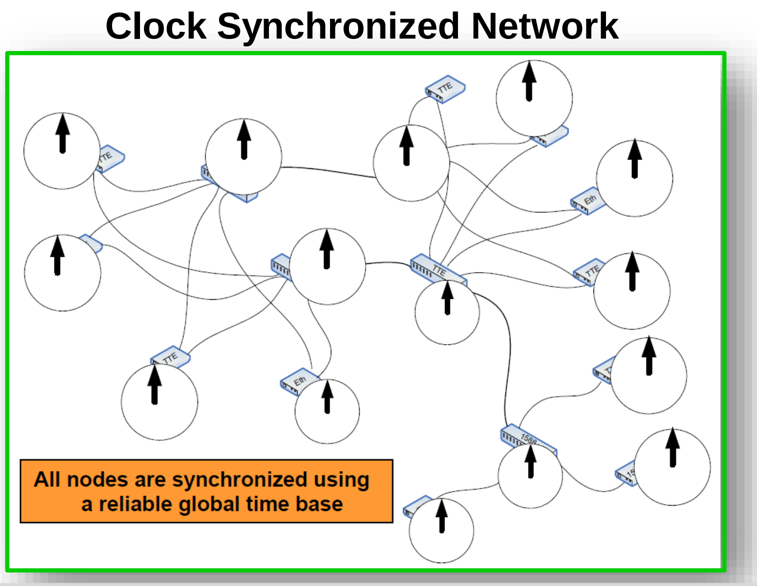
https://www.pubnub.com/blog/how-fast-is-realtime-human-perception-and-technology/

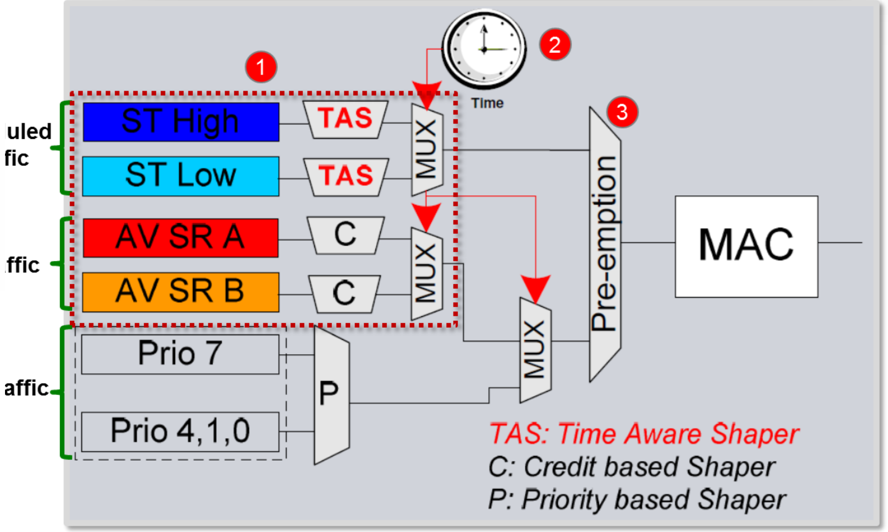
So, 1ms E2E delay for any p2p IVN communication is good enough!

Industry proposals - TSN



Source: IEEE 802.1 AVB TG presentation





Agenda

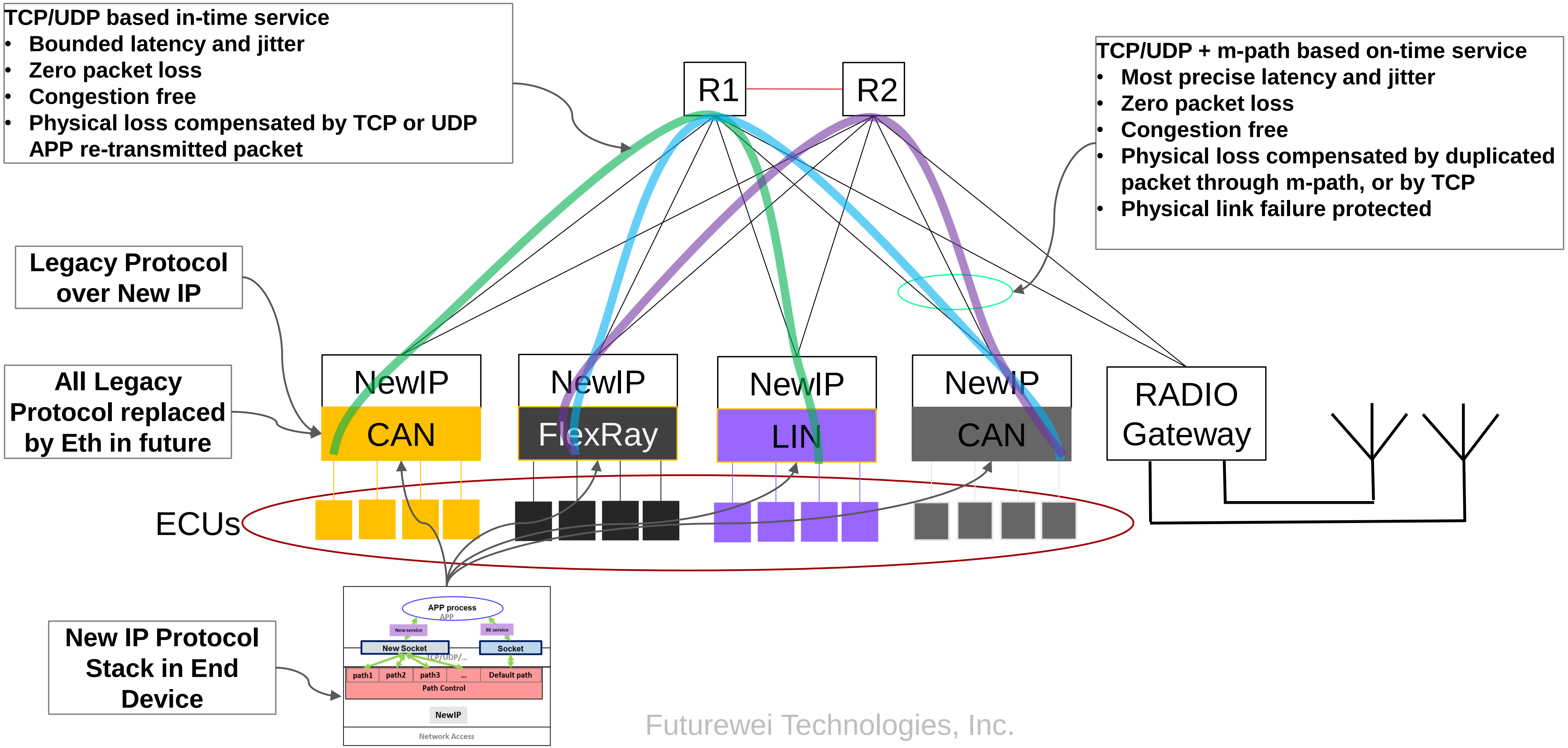
- Trends of Vehicle Evolution
- New IP Overview
  - New IP Introduction
  - New IP for New Service: Goal and Key Technologies
- **Use case for In-Vehicle-Network (IVN)**
  - Review of Current IVN Technologies
  - **New IP Based IVN**
- Use case for V2X
  - Latency Analysis
  - New IP for True E2E uRLLC for 5G and Beyond

Four types of services by New IP for IVN

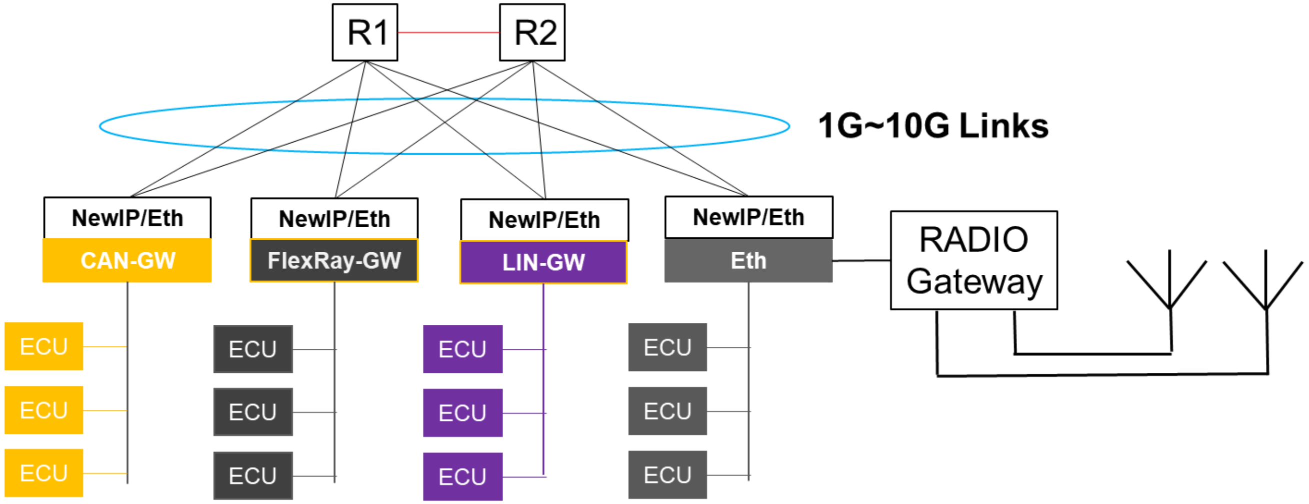
Service Type	Bandwidth	Latency	Jitter	Packet Loss	Use case
On-time (for Scheduled Traffic)	CIR Provided by APP is guaranteed by network	Most precise, Network guarantees E2E bounded latency	Approximately zero	Approximately Zero • m-path to prevent drop due to physical failure • Congestion-free • Lossless (due to queuing)	Synchronous communication: Critical sensor and control data
In-time (for Real-time Traffic)	CIR Provided by APP is guaranteed by network	Minimized, Network guarantees E2E bounded latency	½ of E2E bounded latency	Minimized • Congestion-free • Lossless (due to queuing) • Only drop due to physical failure	Asynchronous communication: Critical sensor and control data
Bandwidth Guaranteed (Bandwidth sensitive, but not time critical)	CIR Provided by APP is guaranteed by network Could reach PIR	Less important	Less important	Don't care	Un-critical data
Best Effort	Don't care	Don't care	Don't care	Don't care	Other data

Scheduled Traffic: Data size, data start time, data rate are fixed, used for critical sensor data and control data in polling mechanism
Real-Time Traffic: Data size and rate are fixed, but when the data starts is unknow, used for urgent sensor data change that polling mechanism did not catch.

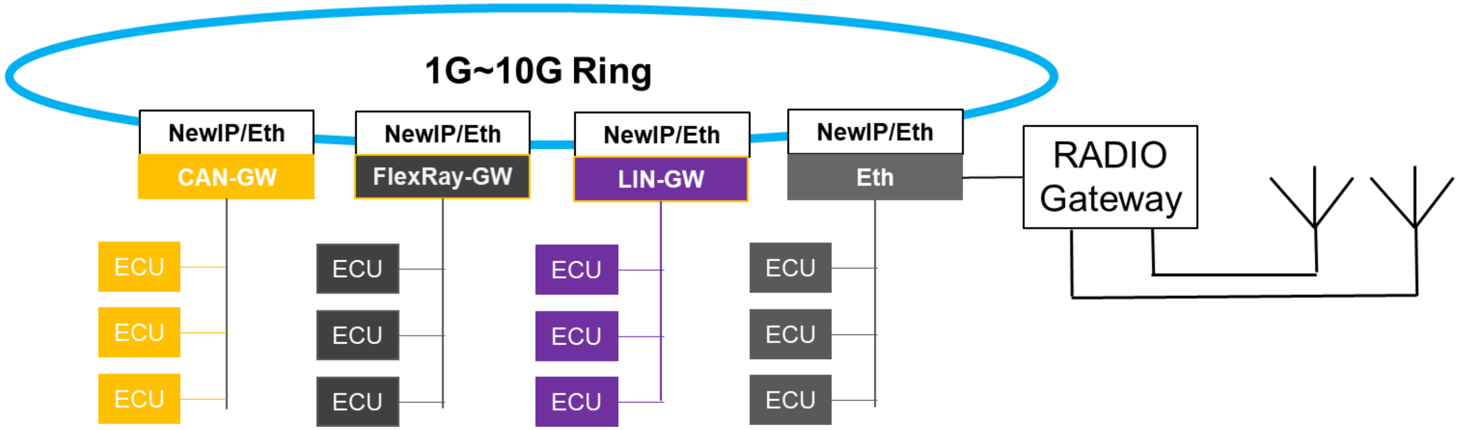
NewIP Based In-Vehicle Network



NewIP Based In-Vehicle Network topo

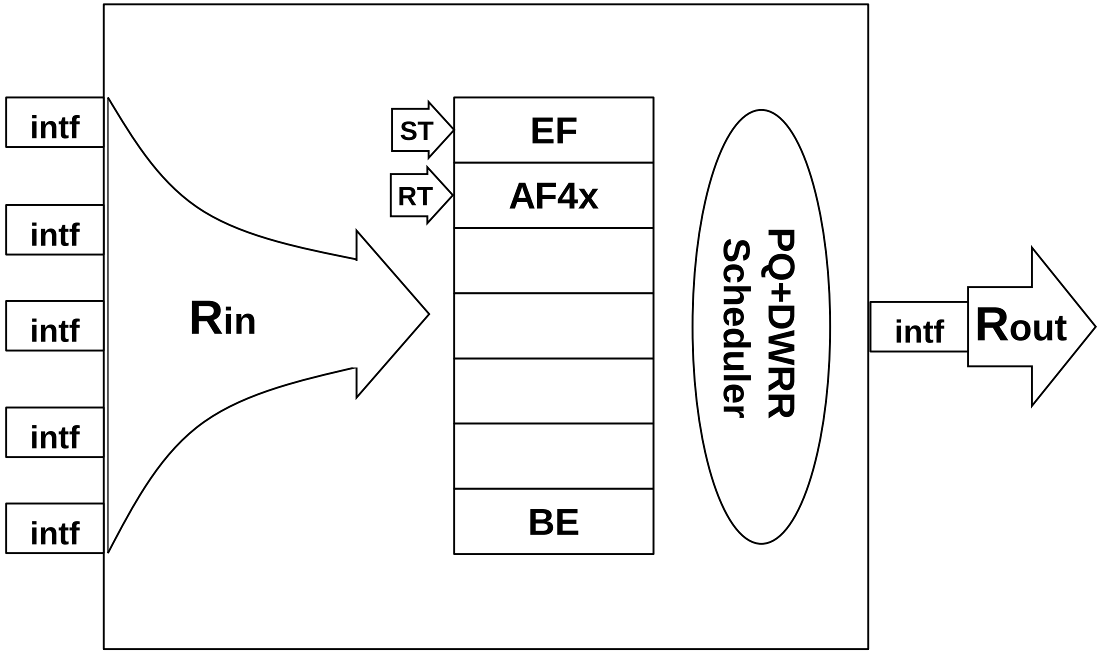


There will be at most two hops between IP and any other protocol communication

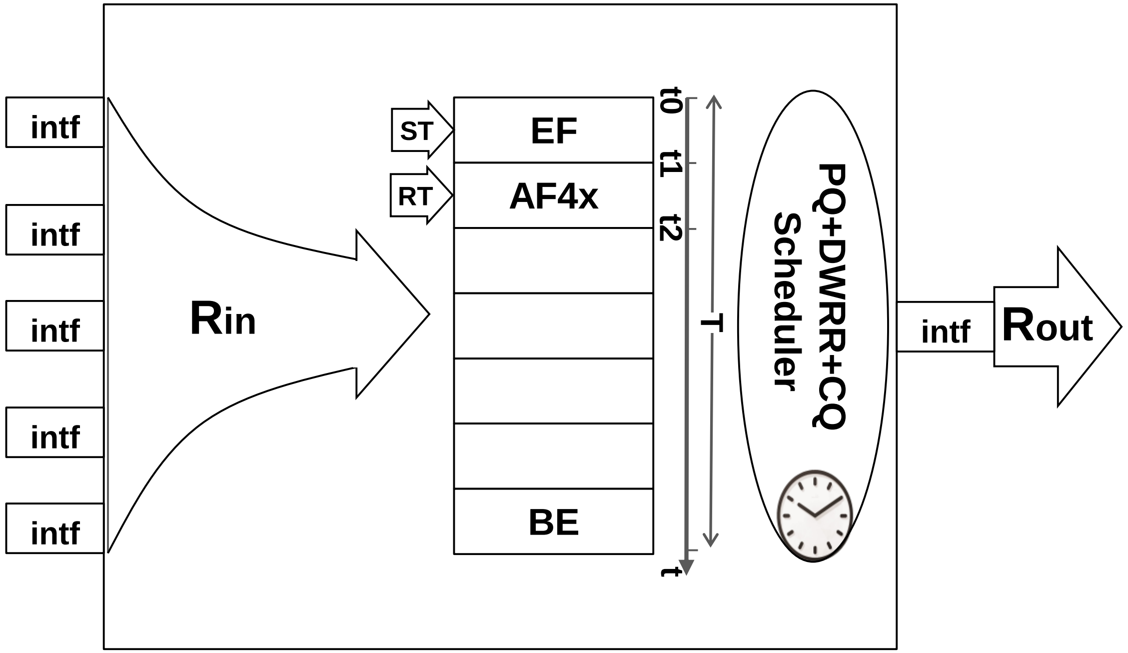


There will be at most two hops between IP and any other protocol communication

Class Based PQ and its Variation

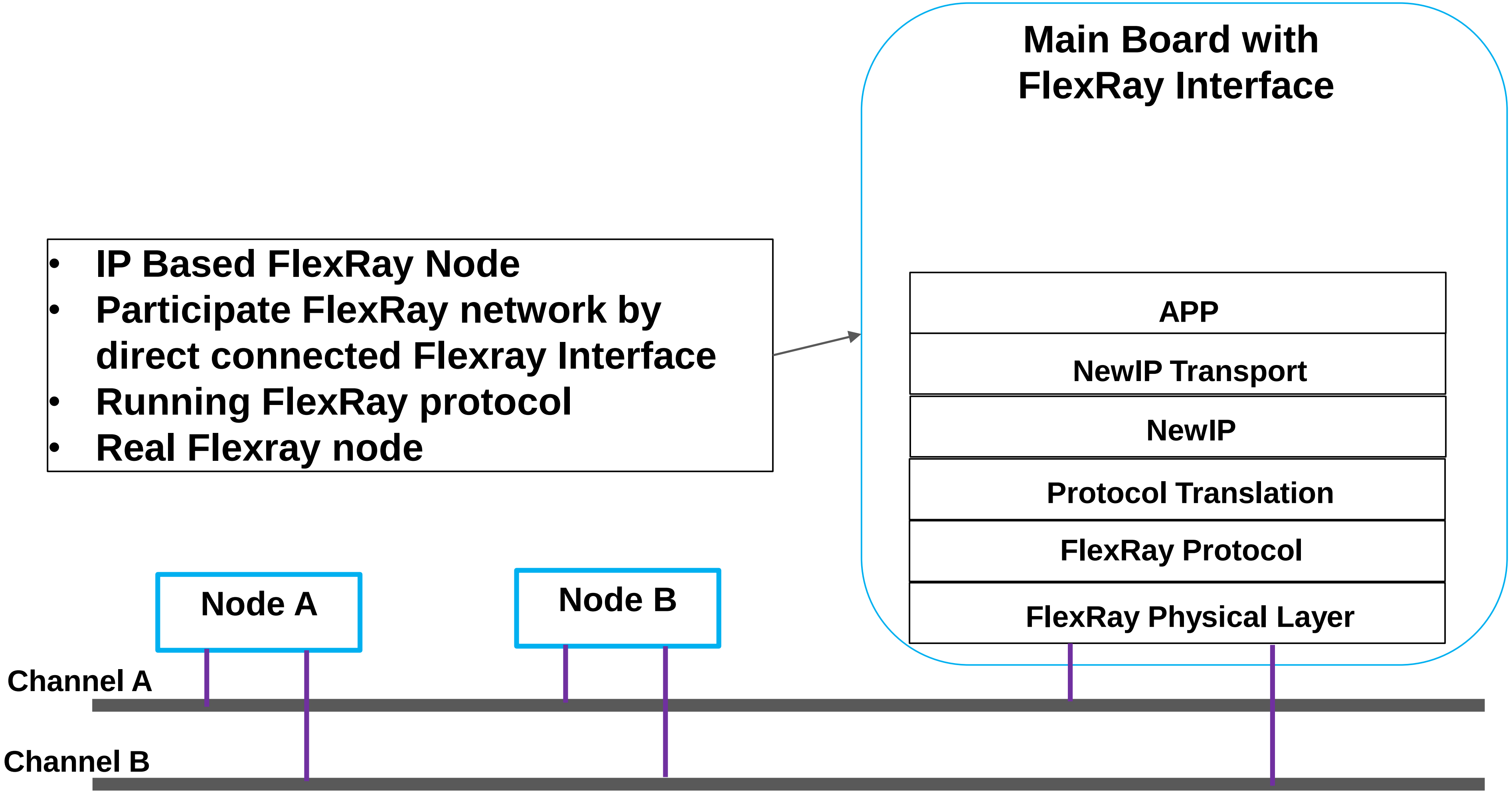


Asynchronous Solution

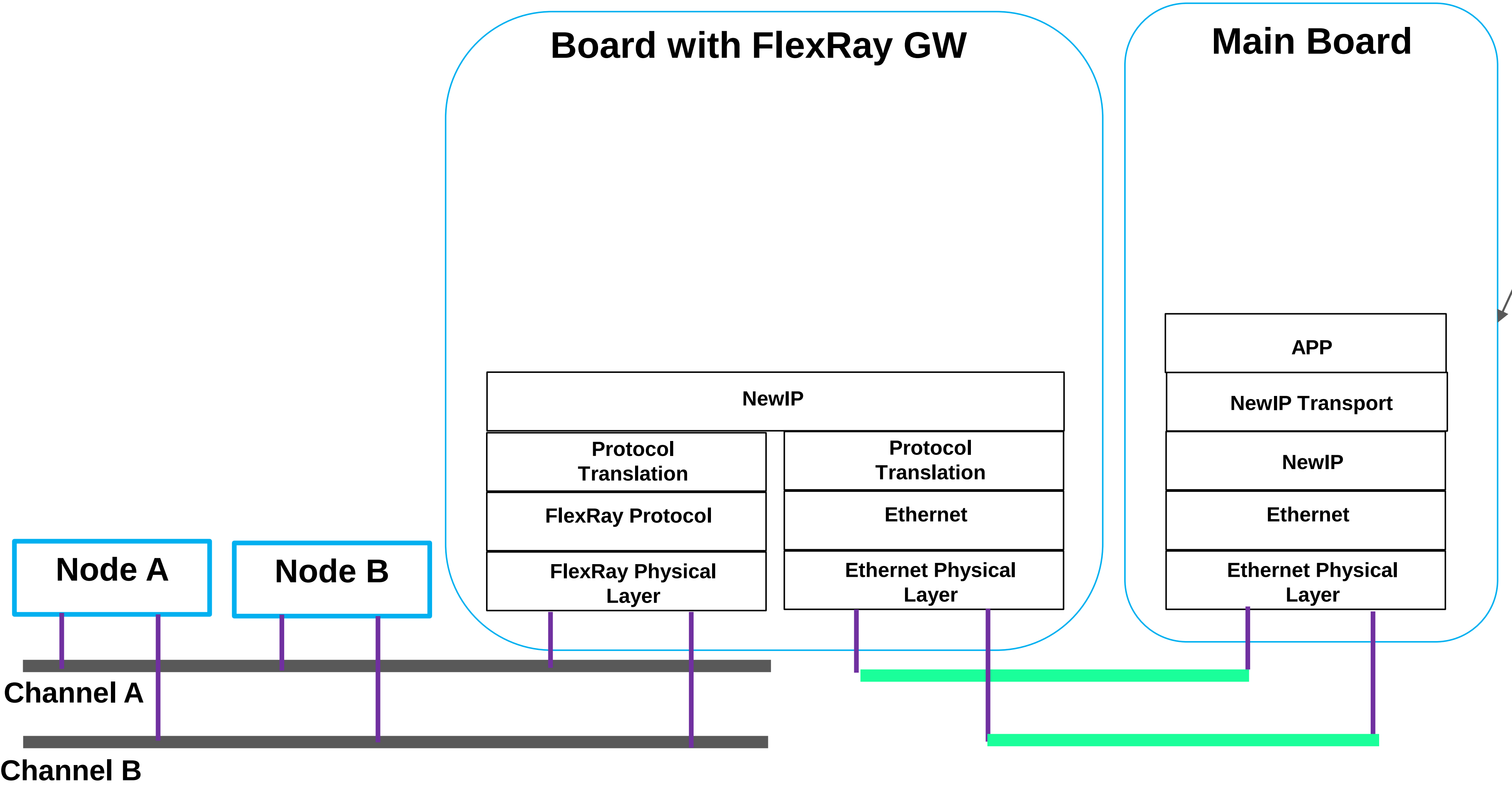


Synchronous Solution

IP based APP communicate with FlexRay Nodes directly

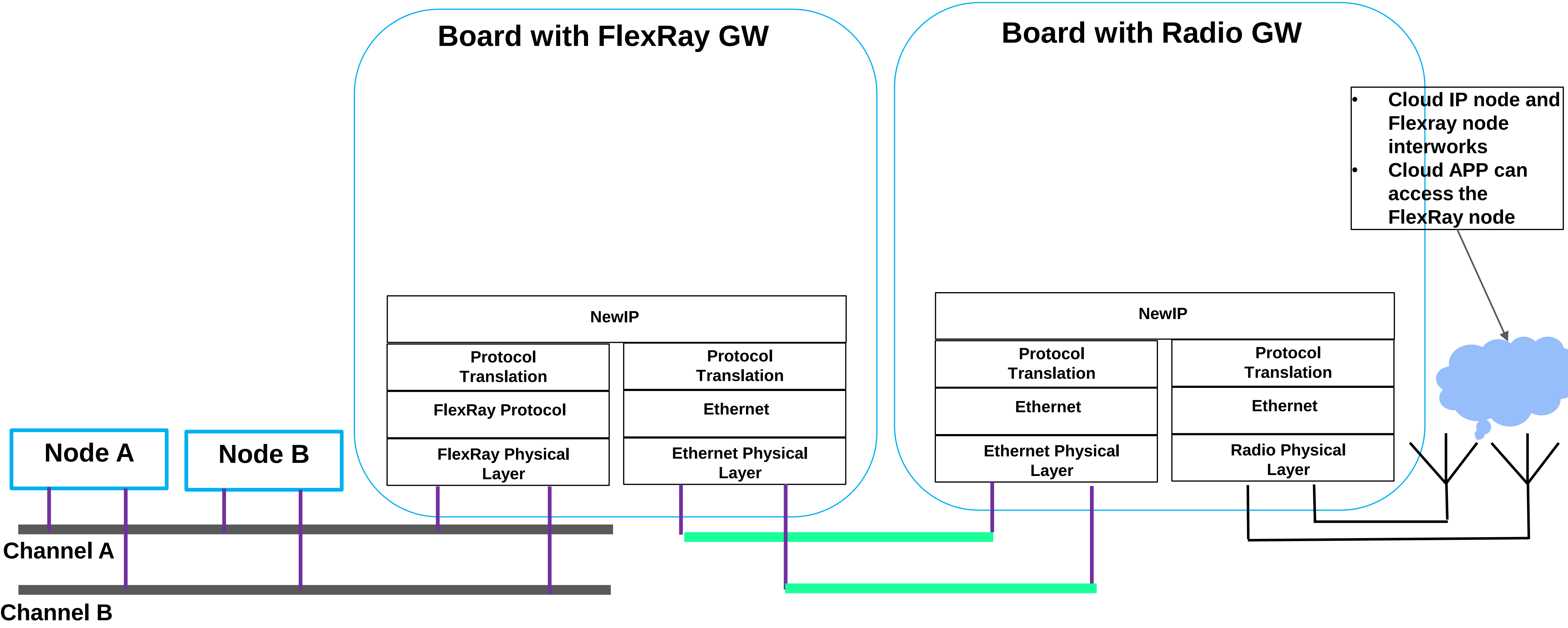


IP based APP communicate with FlexRay Nodes by interworking



- IP node and Flexray node interworks
- Remote IP as virtual Flexray node
- Remote APP can access the FlexRay node
- M-path for more reliability

IP based APP communicate with FlexRay Nodes by interworking between GWs



Agenda

FUTUREWEI Technologies

- Trends of Vehicle Evolution
- New IP Overview
  - New IP Introduction
  - New IP for New Service: Goal and Key Technologies
- Use case for In-Vehicle-Network (IVN)
  - Review of Current IVN Technologies
  - New IP Based IVN
- **Use case for V2X**
  - **Latency Analysis**
  - New IP for True E2E uRLLC for 5G and Beyond

Latency Improvement: 4G to 5G

FUTUREWEI Technologies

Metric	Requirement	Comments
Peak Data Rate	DL: **20 Gb/s** UL: 10 Gb/s	Single eMBB mobile in ideal scenarios assuming all resources utilized
Peak Spectral Efficiency	DL: **30 b/s/Hz** (assuming 8 streams) UL: 15 b/s/Hz (assuming 4 streams)	Single eMBB mobile in ideal scenarios assuming all resources utilized
User Experienced Data Rate	DL: 100 Mb/s UL: 50 Mb/s	5% CDF of the eMBB user throughput
Area Traffic Capacity	Indoor hotspot DL: 10 Mb/s/m²	eMBB
User Plane Latency	eMBB: 4 ms URLLC: **1 ms**	Single user for small IP packets, for both DL and UL (eMBB and URLLC)
Control Plane Latency	20 ms (encouraged to consider **10 ms**)	Transition from Idle to Active (eMBB and URLLC)
Connection Density	1M devices per km²	For mMTC
Reliability	99.9999% success prob.	32 L2 bytes within 1 ms at cell edge
Bandwidth	>100 MHz; up to **1 GHz** in **> 6 GHz**	Carrier aggregation allowed

Source: https://spectrum.ieee.org/telecom/wireless/3gpp-release-15-overview

Source: Samsung

- 5G latency is 1/10 of 4G; bandwidth is 20x of 4G
- Latency definition: RTT for user plane
- 5G: air latency < 1ms; E2E latency < 5ms

Progress and Reality

- Rel 15 technologies support air latency: 0.7(DL)+2.0(UL) (lots restrictions)
- E2E mobile latency: UE <-> PGW (UPF)
- E2E Internet Latency: UE <-> UE
- Wiki: In 5G, the "air latency" in equipment shipping in 2019 is 8–12 milliseconds. The latency to the server must be added to the "air latency" for most comparisons. Verizon reports the latency on its 5G early deployment is 30 ms: Edge Servers close to the towers can reduce latency to 10–20 ms; 1–4 ms will be extremely rare for years outside the lab.
- Long way to go for 5ms E2E latency
- No clear solution in Core and Internet

TABLE II: Latency results

		Rel. 14 SF	Rel. 15 SF & n+3	Rel. 15 slot	Rel. 15 subslot
DL	initial transmission	4	4	2	0.7
	1st retransmission	12	10	6	2.0
	2nd retransmission	20	16	10	3.3
	3rd retransmission	28	22	14	4.7
UL	initial transmission	12	10	6	2.0
	1st retransmission	20	16	10	3.3
	2nd retransmission	28	22	14	4.7
	3rd retransmission	36	28	18	6.0

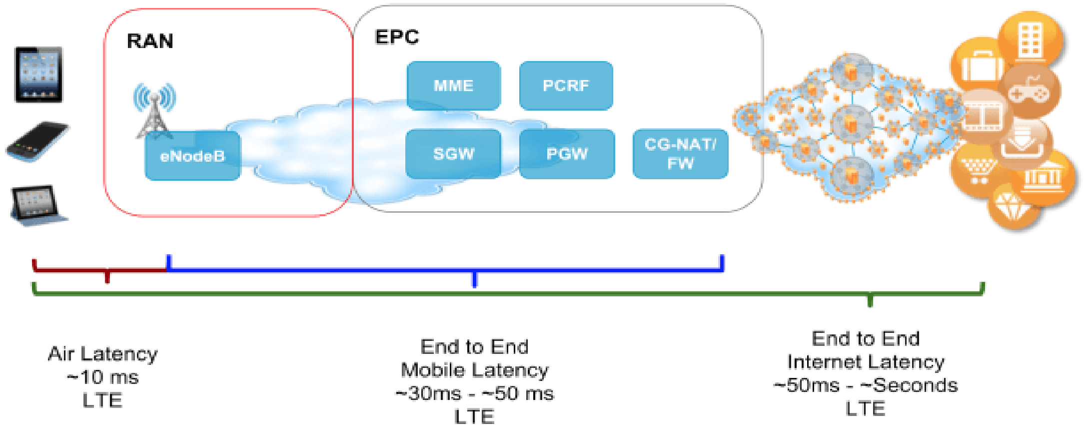
(a) with HARQ retransmissions

		Rel. 14 SF	Rel. 15 SF & n+3	Rel. 15 slot	Rel. 15 subslot
DL	initial transmission	4	4	2	0.7
	1st repetition	5	5	2.5	0.8
	2nd repetition	6	6	3.0	1.0
	3rd repetition	7	7	3.5	1.2
UL	initial transmission	12	10	6	2.0
	1st repetition	14	12	7	2.3
	2nd repetition	16	14	8	2.7
	3rd repetition	18	16	9	3.0

(b) HARQless repetition

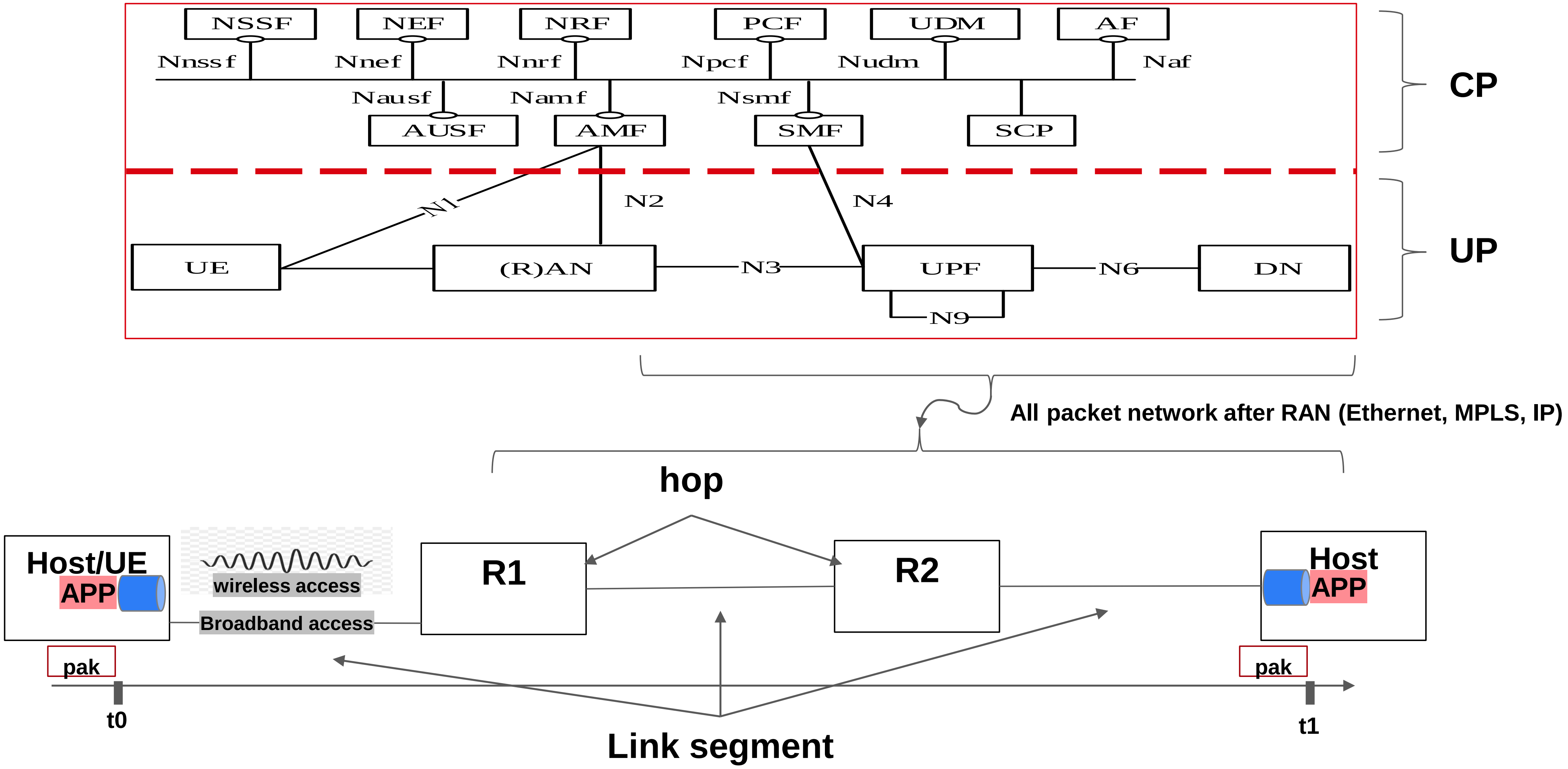
Calculated results for Downlink (DL) and Uplink (UL) for the LTE Rel. 14 SF (subframe) 1 ms TTI as well as LTE Rel. 15 short processing time, slot and subslot configurations. The circles indicate the fulfillment of the 10 ms HRLLC (H) requirement and 1 ms URLLC (U) requirement respectively.

Source: Thomas Fehrenbach, et.at. “URLLC Services in 5G Low Latency Enhancements for LTE”,



Source: "5G Latency – Reality Checks". SENKI. December 9, 2018.

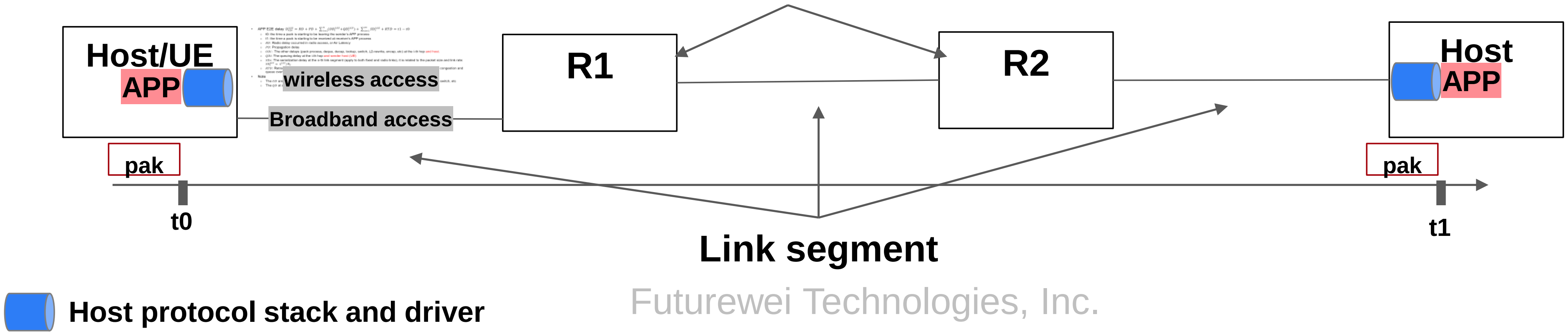
E2E Latency Analysis



Host protocol stack and driver

E2E Delay Estimation

- APP E2E delay $D_{e2e}^{LGS} = RD + PD + \sum_{i=1}^{n}(OD_i^{LGS} + QD_i^{LGS}) + \sum_{s=1}^{m} SD_s^{LGS} + RTD = t1 - t0$
  - t0: the time a pack is starting to be leaving the sender's APP process
  - t1: the time a pack is starting to be received at receiver's APP process
  - RD: Radio delay occurred in radio access, or Air Latency
  - PD: Propagation delay
  - ODi : The other delays (pack process, deque, decap, lookup, switch, L2-rewrite, encap, etc) at the i-th hop and host.
  - QDi: The queuing delay at the i-th hop and sender host (UE)
  - SDs: The serialization delay at the s-th link segment (apply to both fixed and radio links), it is related to the packet size and link rate: $SD_S^{LGS} = L^{LGS}/R_S$
  - RTD: Retransmission Delay when there is packet loss, Caused by packet loss 1) physical fault or failure in media 2) congestion and queue overflow
- Note
  - The OD are usually and relatively fixed, it is dependent on HW; OD at host are delays caused by pack process, task switch, etc
  - The QD at sender host is similar to hop



Strategy for E2E Latency Reduction

FUTUREWEI Technologies

Strategy

- RD: Air Latency, significant
  - The major contributor for E2E delay
  - Refer to all technologies in 5G New Radio (NR),
- PD: Propagation delay, insignificant
  - Deduce the distance of all links
  - Place the server to UE as close as possible. The closest point is directly connecting to UPF, or even gNB, (MEC)
- ODi : Other delays, insignificant
  - Enhance hardware and chip performance,
  - Optimize/speed up all process in packet switch/forwarding
- QDi: Queuing delays, significant
  - The major contribute for the E2E delay
  - The new service for packet switching and forwarding from Best-Effort to Guaranteed service
- SDs: The serialization delay, significant
  - Increase Link Speed, Reduce the packet size
- RTD: The retransmission delay, significant
  - Redundancy to mitigate the packet loss
  - New technology to eliminate the congestion and queue overflow.

What New IP can contribute

- **New IP can Provide New Service**
  - **E2E and Guarantee**
  - **Granular to IP Flow Level**
  - **Bandwidth (BGS) and Latency (LGS)**
    - **Guaranteed Bandwidth**
    - **Bounded Latency (in-time/on-time)**
  - **Congestion Free and lossless**
- RD: Air Latency
  - Signaling integration with NR to achieve the true E2E uRLLC
- ODi : Other delays
  - Optimize/speed up lookup in packet switch/forwarding by using the Service ID
- QDi: Queuing delays
  - The new service for packet switching and forwarding from Best-Effort to Guaranteed service
- RTD: The retransmission delay
  - Eliminate the congestion and queue overflow.
  - Multiple-path for packet loss protection
  - Other network coding technologies

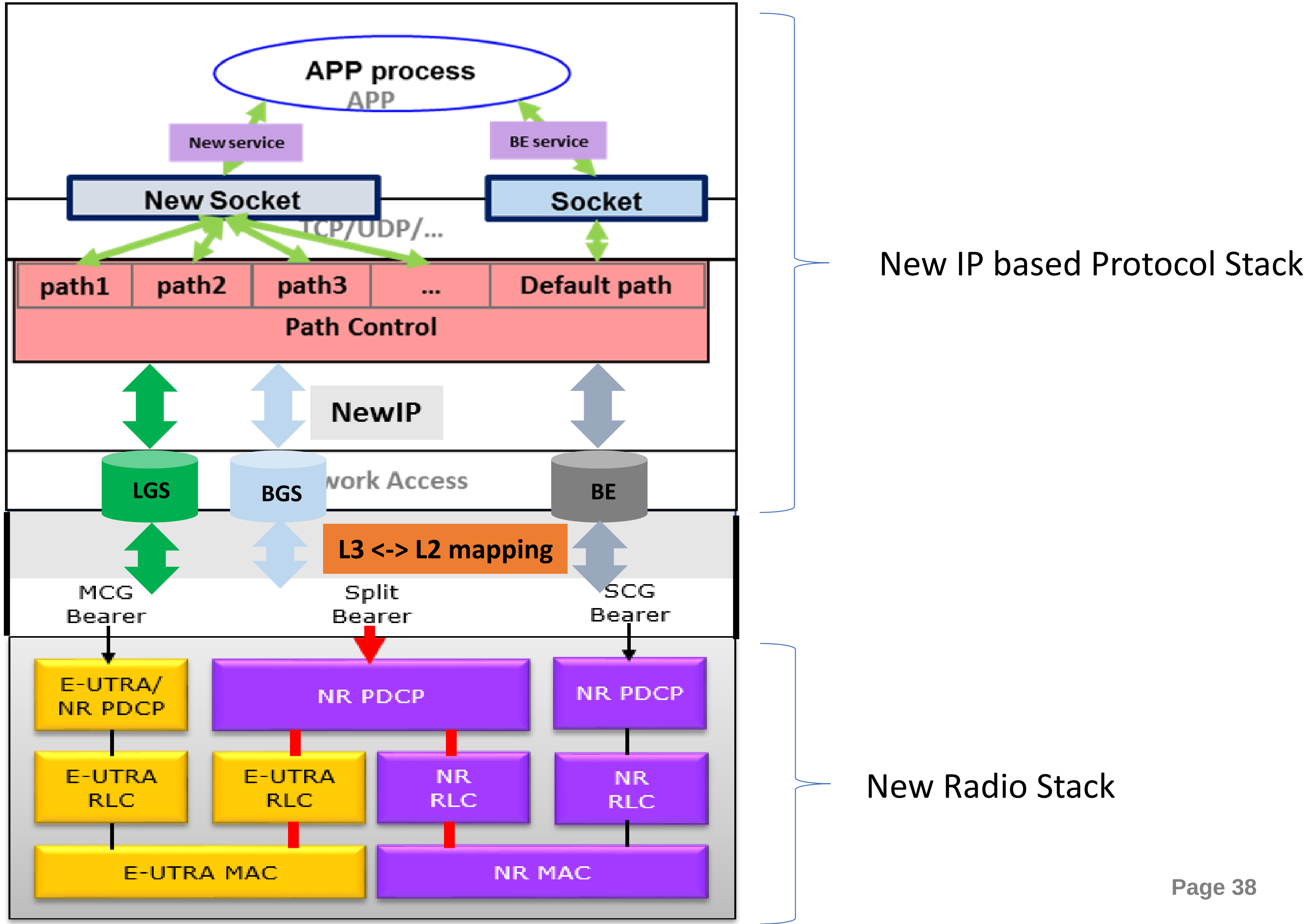
Agenda

FUTUREWEI Technologies

- Trends of Vehicle Evolution
- New IP Overview
  - New IP Introduction
  - New IP for New Service: Goal and Key Technologies
- Use case for In-Vehicle-Network (IVN)
  - Review of Current IVN Technologies
  - New IP Based IVN
- **Use case for V2X**
  - Latency Analysis
  - **New IP for True E2E uRLLC for 5G and Beyond**

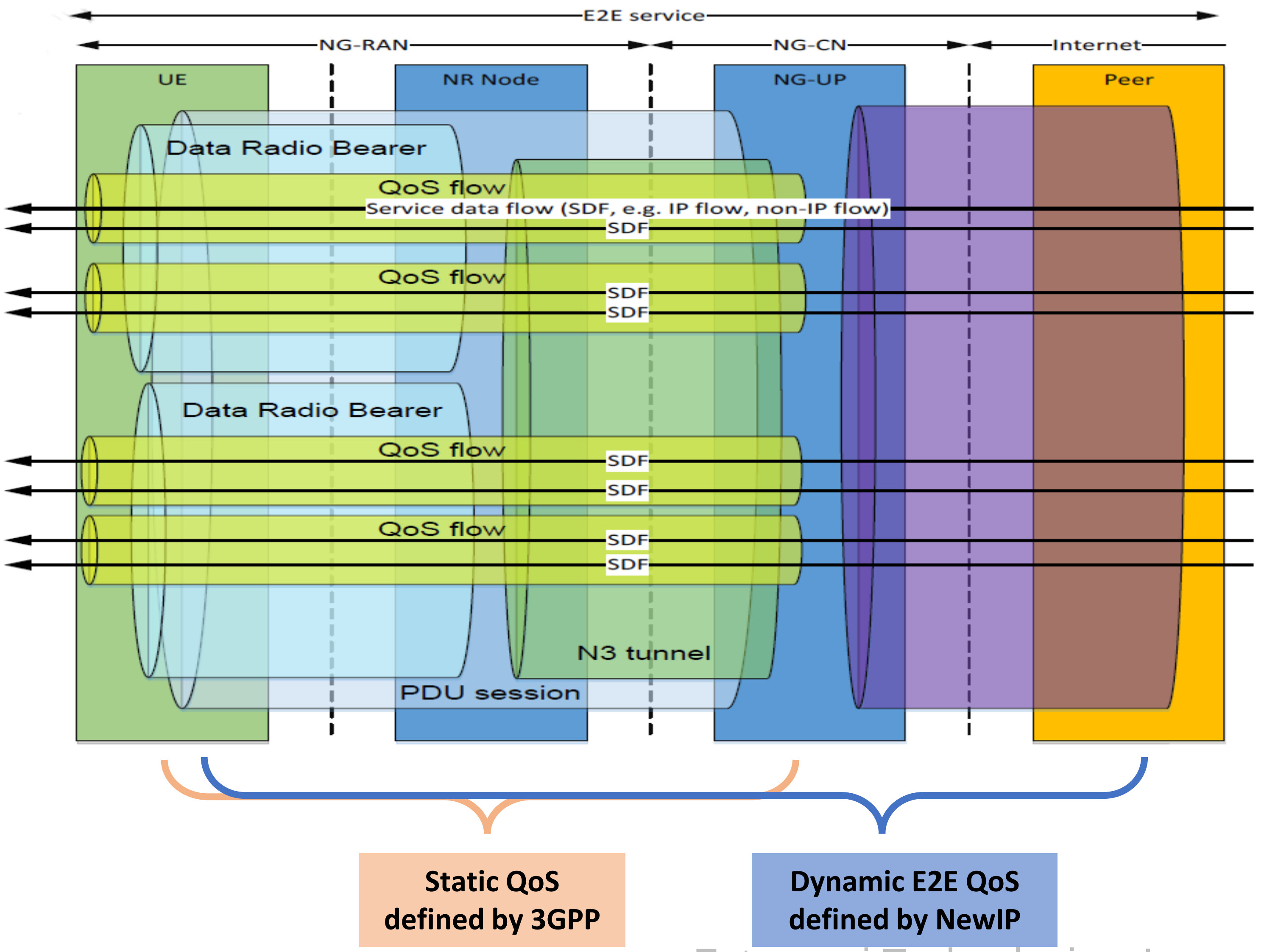
New IP stack integrated with NR in UE

FUTUREWEI Technologies



Dynamic QoS by New IP for True E2E uRLLC

FUTUREWEI Technologies

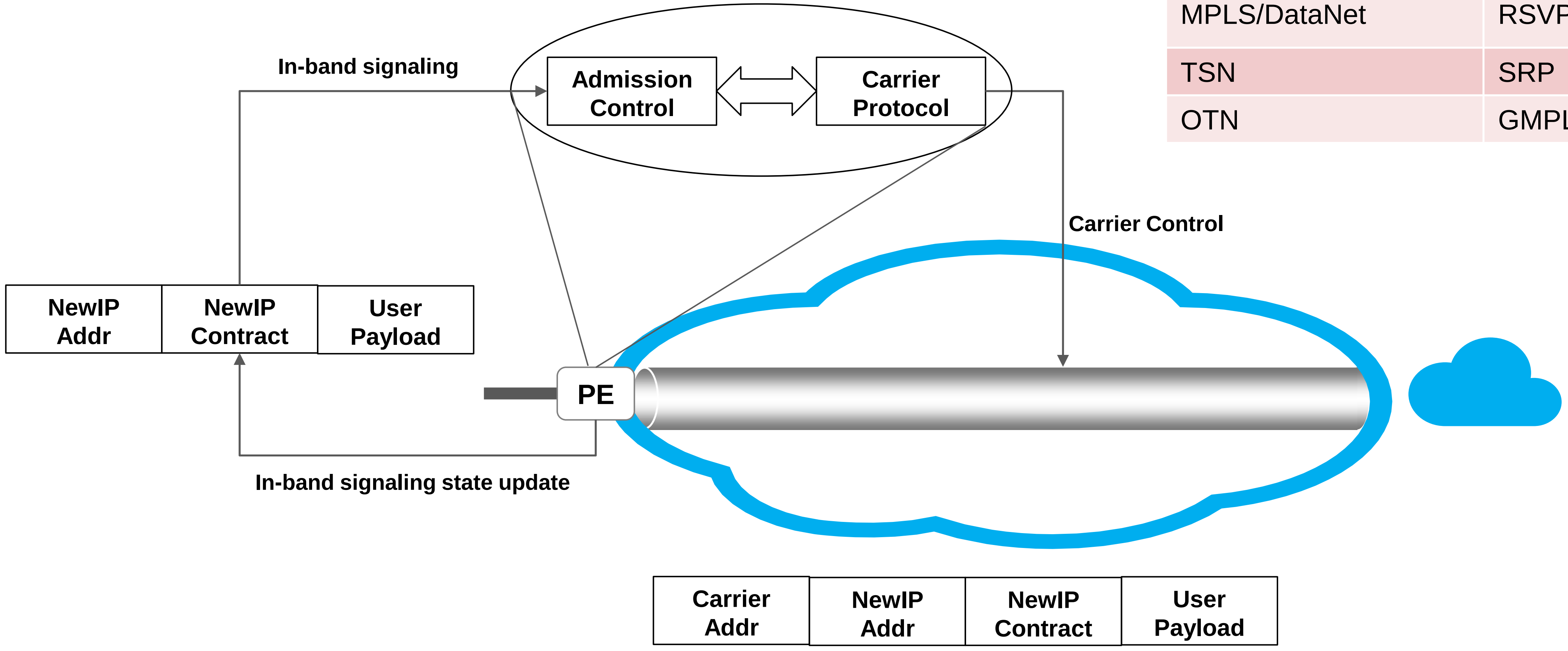


Dynamic QoS

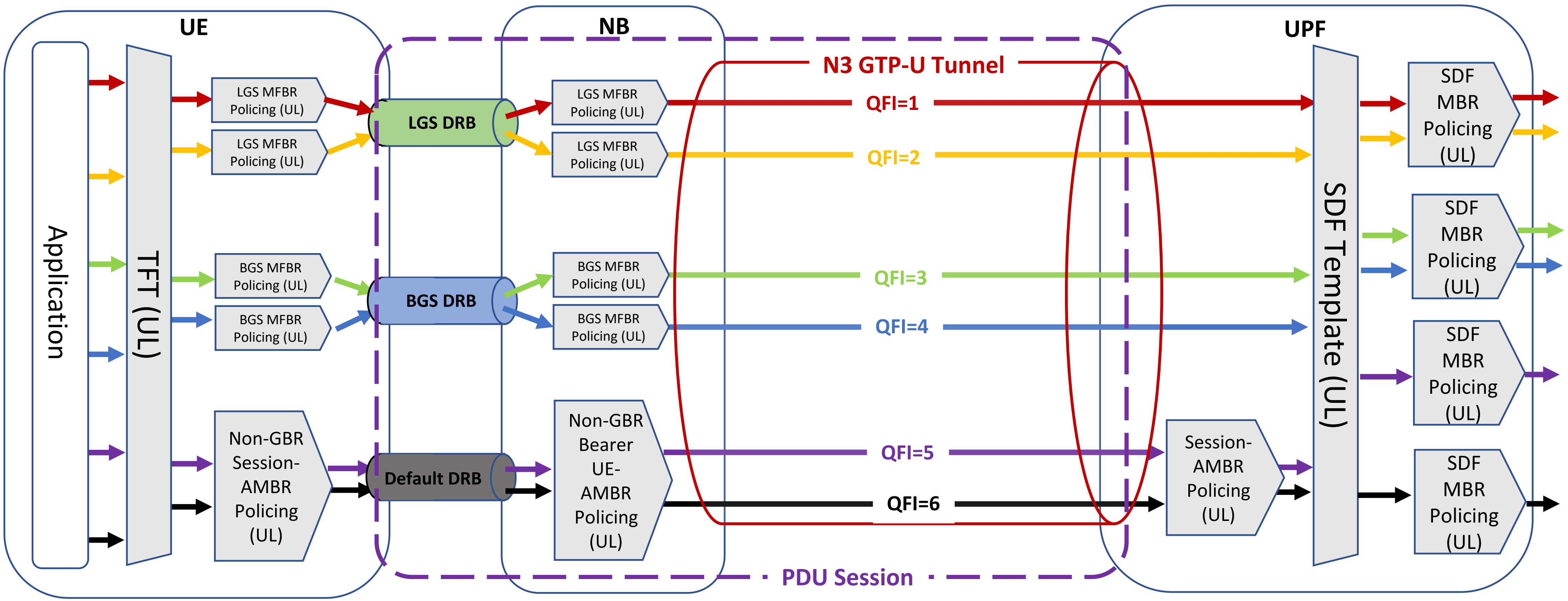
- True E2E uRLLC
- APP driven
  - Per APP session or Per user session
- Controlled by
  - User device and APP
  - Service Provider
  - Content Provider
- Dynamic and on-demand
  - Reservation does not consume network resource (Spectrum, bandwidth)
  - Only use the resource when there is traffic
- Flexible granularity
  - 5 IP tuples
  - Less than 5
- Flexible Service
  - BGS
  - LGS
  - Lossless

In-band Signaling Integration with Carrier

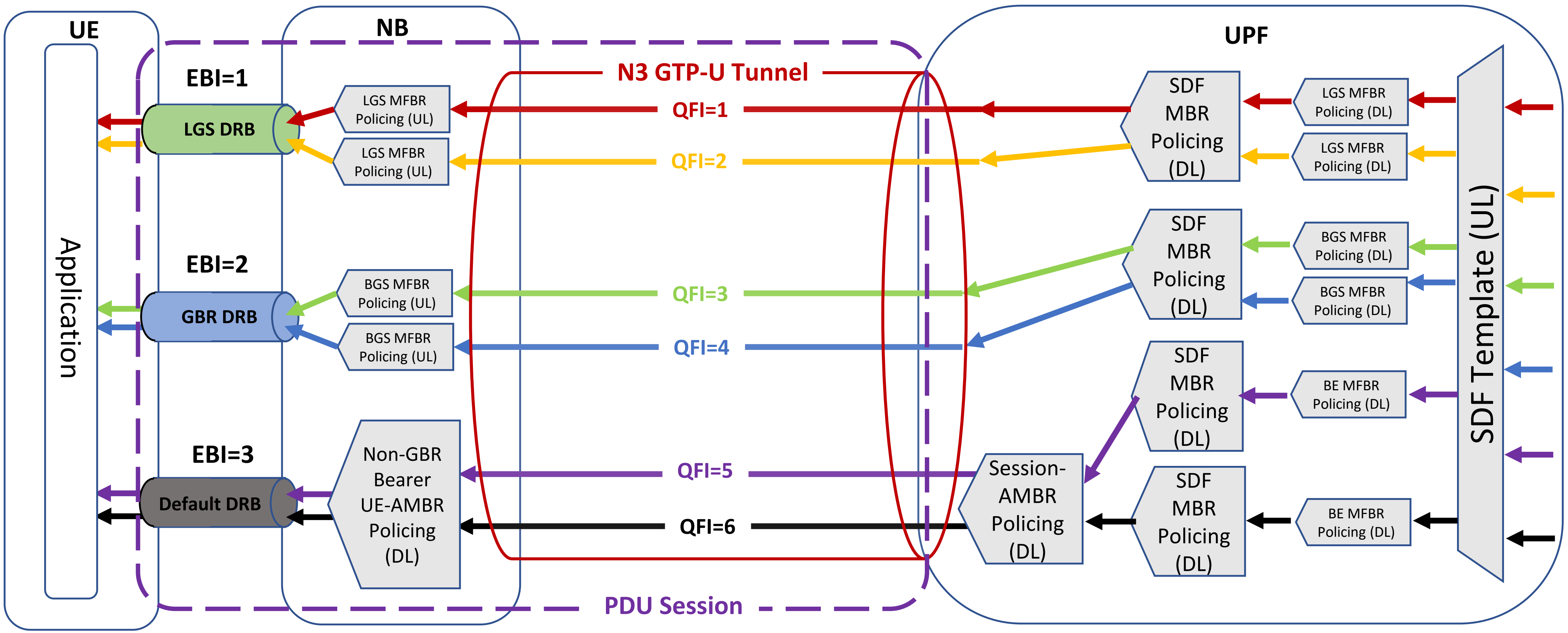
Carrier Type	Carrier Protocol
Radio Date Bearer	PDCP
MPLS/DataNet	RSVP-TE
TSN	SRP
OTN	GMPLS



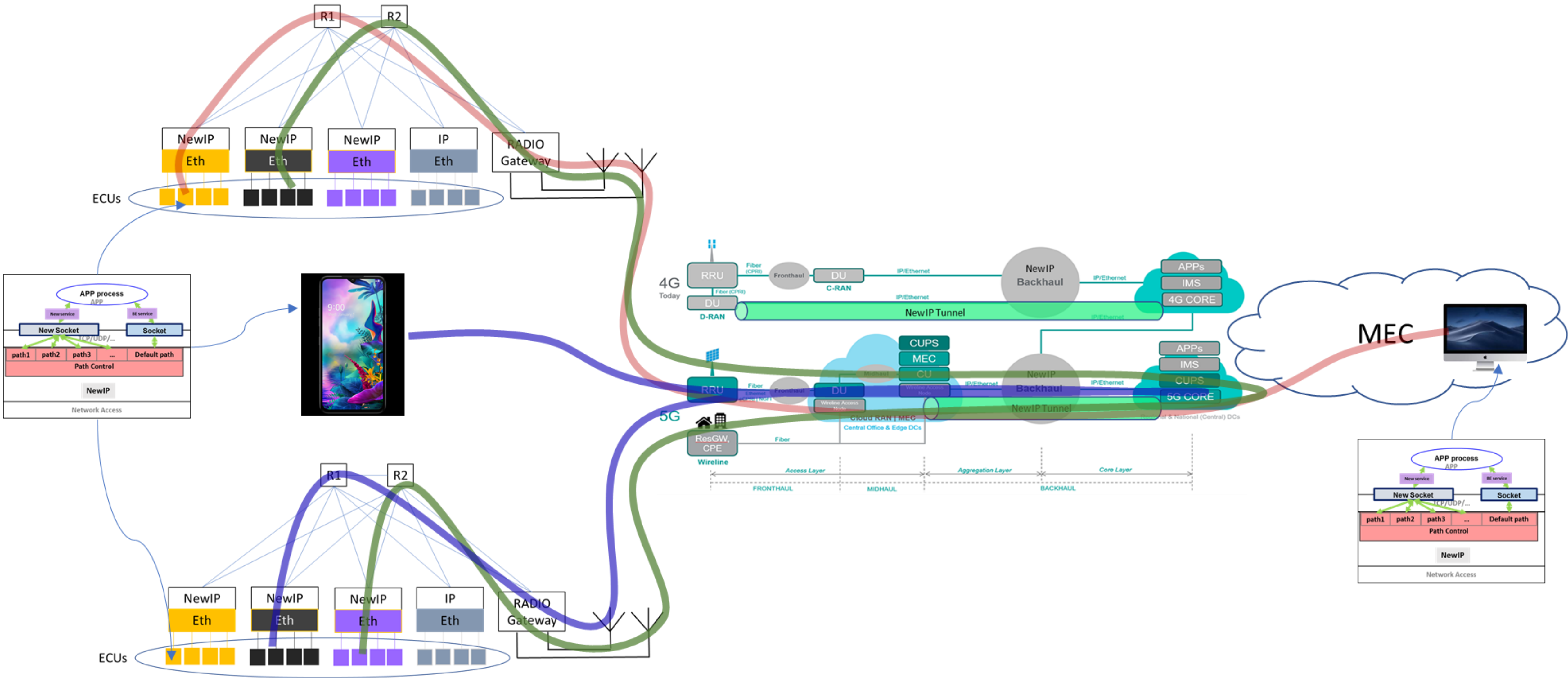
5G: Up Link QoS Policy Enforcement



5G: Down Link QoS Policy Enforcement



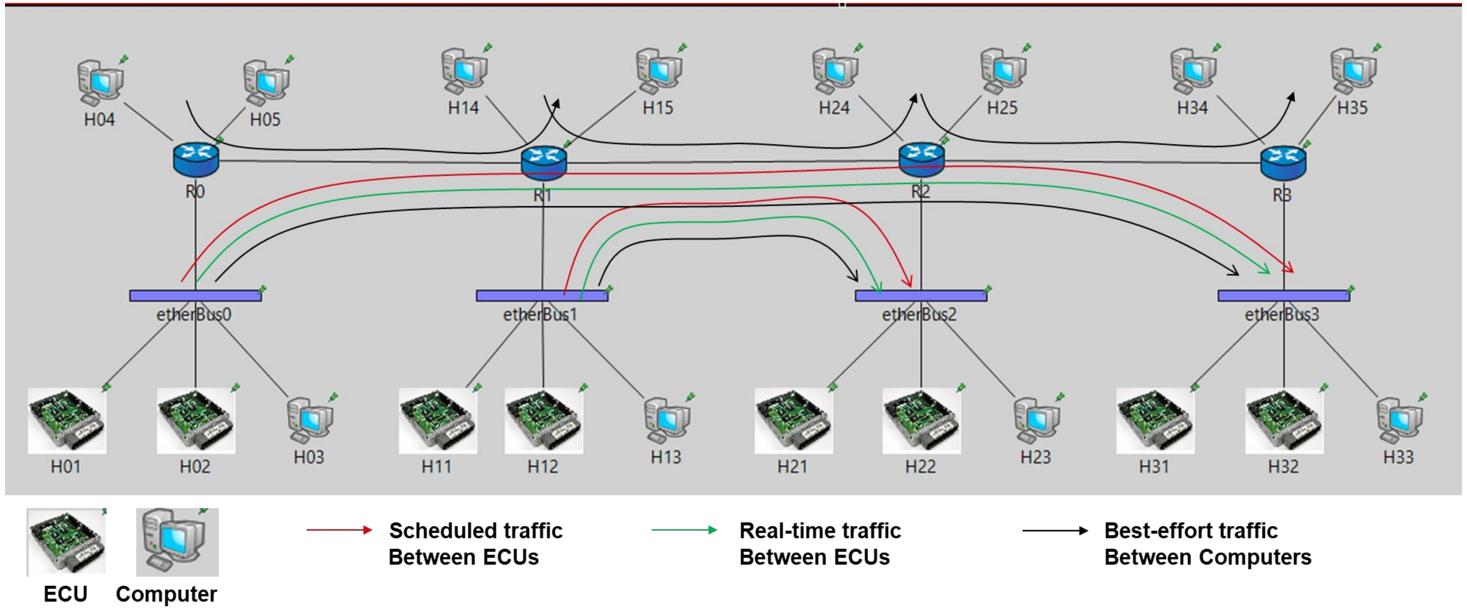
NewIP for E2E uRLLC



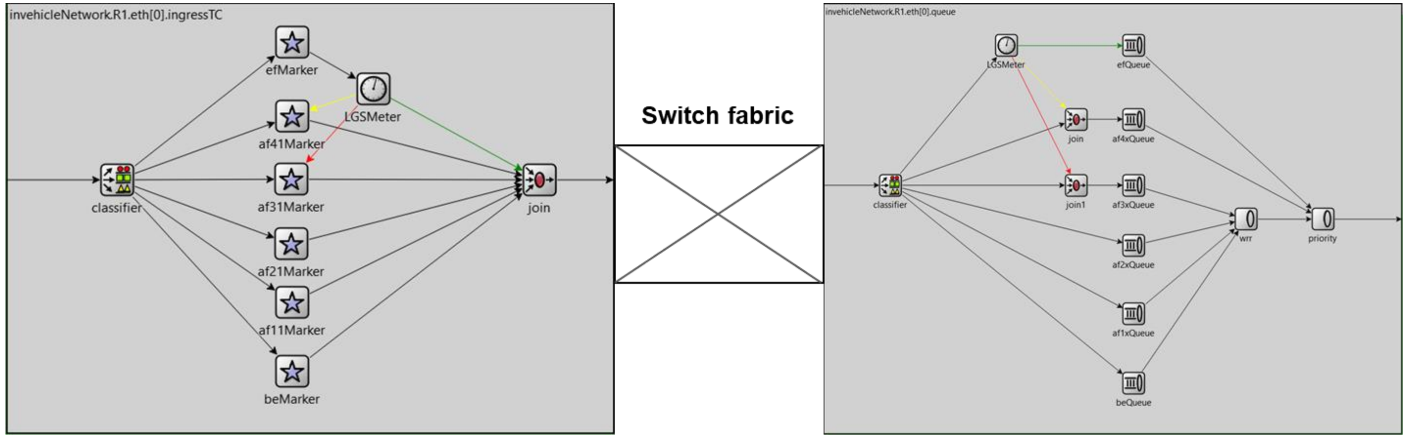
Q&A

Simulation Results – Topo and Modeling

FUTUREWEI Technologies



- 100 ST + 100 RT: UDP, 254 bytes, 40Mbps
  H01/H02-> H31/H32; H11/H12->H21/H22
- 250 BE: TCP/UDP, 254bytes+1500bytes, >>60Mbps
  H03->H33, H13->H23, H04->H14, H15->H23, H24->H34
- Measure the most severely congested R1's Q depth
- Measure the worst performed flow E2E latency



Estimation and Simulation Results

FUTUREWEI Technologies

Algo	Class and traffic	Estimated max number of pack in Egress Q					Estimated Total Queuing Latency (us)	Calculated Total Serialization Delay (each hop has 20 us)	Estimated Total E2E Delay (us)
		Host	R0	R1	R2	R3			
PQ+DWRR	EF for ST	0	3 (r_{EF}=2)	6 (r_{EF}=4)	3 (r_{EF}=1)	3 (r_{EF}=1)	305	100	405
	AF4x for RT	0	4 (r_{AF4x}=2)	6 (r_{AF4x}=4)	4 (r_{AF4x}=1)	4 (r_{AF4x}=1)	365	100	465
PQ+DWRR +CQ	EF for ST	0	2 (r_{EF}=1)	2 (r_{EF}=1)	2 (r_{EF}=1)	2 (r_{EF}=1)	162	100	262
	AF4x for RT	0	2 (r_{AF4x}=1)	2 (r_{AF4x}=1)	2 (r_{AF4x}=1)	2 (r_{AF4x}=1)	162	100	262

Algorithm	Min/Max E2E Delay (us) for the worst performed flow carrying ST betweenH01/H02 to H31/H32		Min/Max E2E Delay (us) for the worst performed flow carrying RT betweenH01/H02 to H31/H32	
	Experiment Value (us)	Estimation Value (us)	Experiment Value (us)	Estimation Value (us)
PQ+DWRR	108/391 (Figure 7)	100/405	278/542 (Figure. 8)	100/465
PQ+DWRR+CQ	109/152 (Figure. 9)	100/262	169/169 (Figure. 10)	100/262

Formulas refer to: “A Framework for Bandwidth and Latency Guaranteed Service in New IP Network”, New IP workshop INFOCOMM 2020.

E2E Latency, Lossless for Worst Performed flow with ST and RT