

## Study for In-Vehicle-Network and New V2X Architecture by New IP

Lin Han, Lijun Dong, Richard Li

*Futurewei Technologies, Inc.*

Santa Clara, California, U.S.A

email: {lin.han, lijun.dong, richard.li}@futurewei.com

**Abstract—** For many In-Vehicle-Network (IVN) and Vehicle-to-everything (V2X) applications in the latest vehicle, the higher Quality of Service (QoS) and more deterministic networking are mandatory requirements. The paper proposes an architecture to support latency sensitive communication that is based on New IP technology. The new architecture and technologies can provide the End-to-End (E2E) Latency Guaranteed Service (LGS) and Bandwidth Guaranteed Service (BGS) for any granularity of IP flow(s). It can be used for IVN and V2X communication combined with 5G for future Internet. This paper will use IVN as an example to prove that the New IP can replace other legacy protocols and is able to provide satisfactory service in terms of the critical QoS metrics (Bandwidth, Latency, Jitter and Packet loss). The paper will analyze the challenge of latency requirements for IVN, it focuses on the design of new IVN control plane and data plane especially queuing and scheduling. The theoretical latency analysis, estimation and experimental verification are provided.

**Keywords-** IVN; V2X; TCP; IP; UDP; QoS; New IP; Deterministic Networking; In-band signaling; Guaranteed Service; Class Based Queueing, Priority Scheduling; Cyclic Queueing, End-to-End; Traffic Shaping; Congestion; Packet loss; Bandwidth; Latency; Jitter; eMBB; mMTC; uRLLC

### I. INTRODUCTION

This paper is an extended version of [1], which investigates the latency requirements for IVN, proposes a New IP based IVN architecture, and presents a detailed study and emulations. The paper will provide more details about the New IP based V2X architecture, the algorithms, and experimental results.

Recently, a trend in vehicle industry is that electrical or hybrid motors are gradually replacing the combustion engine and power transmission. The major components of Electrical Vehicle (EV) are battery and electrical motors. They are simpler, more modular, and easier to be manufactured with standard and thus lower the manufacturing threshold and cost. This results in tougher competitions in other areas, such as Tele-driving, Self-driving, Infotainment System, etc. All those advanced futures are computing driven and require advanced networking technologies in following two areas:

- In-Vehicle-Network (IVN): this is the network inside vehicle to connect different electronic devices, such as Sensors, Actuators, Electrical controller unit (ECU), GPS, Camera, Radar, LiDAR, Embedded computer, etc.
- Vehicle-to-Everything (V2X): This is a technology that allows moving vehicle to communicate with other moving vehicles, the traffic control system along roads, and everything in Internet, such as Cloud, home, environment,

people, etc. The traditional V2X term only represents the wireless technologies DSRC defined in IEEE802.11p [2], and C-V2X defined in 3GPP [3]. DSRC is a modification of Wi-Fi and allows wireless devices communicate directly without intermediate device. C-V2X supports two modes: Direct C-V2X (Devices communicate directly) and Indirect C-V2X (Device communicate via wireless network). In this paper, V2X is defined as a general term that is End-to-End communication between any applications within a car and another application running outside of that car, that application could be running in another car, in a cell phone, in cloud or in Internet.

There are different types of applications using IVN or V2X. Based on the requirements for network, traffic can be categorized as three types:

- The time sensitive: For this type of communication, the latency requirement is stringent, but the data amount is limited. This includes the communication for sensor data, control data, such as the control for powertrain system, braking system, security system, etc. The data rate is up to Mbps per flow. This type of traffic normally could be within a car on top of In-Vehicle-Network (example a), it could also be between applications in a car and remote applications on device outside the car using V2X (example b):
  - a. For Self-driving car, some critical sensor data and control data are very time sensitive, the IVN must provide the guaranteed service for shortest E2E latency and zero packet loss.
  - b. Tele-driving system will control a car remotely by human being, or by an automatic AI system in cloud. The feedback data from a car and associated control signal from remote site must experience the shortest latency.
- The bandwidth sensitive: For this type of communication, the latency requirement is not stringent, but the data amount is higher. It includes GPS display, Radar, LiDAR data feeding. The data rate could be up to tens of Mbps per flow. Like the 1<sup>st</sup> type traffic, some of this type of traffic is within a car, but some is between a car and a remote application.
- Best-Effort: This is the traditional IP traffic that is not belonging to above two types. Network will deliver the traffic to destination without any guarantee.

For above three types of traffic, the 1<sup>st</sup> one is the most challenging to support by the current technologies for V2X and Internet. This is because the current V2X only addresses the wireless technologies by DSRC or C-V2X but does not

consider other wired network segments. From the perspective of E2E effects for V2X, the latency, jitter and packet loss happened in the segments of wired network are not negligible. Since the IP network can only provide the Best-Effort service, the queuing latency and packet loss due to congestion in IP network is very normal.

The paper proposes to use New IP technology for new architecture of IVN and V2X. New V2X architecture will integrated 5G and New IP to obtain the true E2E guaranteed service in terms of bandwidth, latency, jitter, and packet loss. The remained paper has three parts:

- 1<sup>st</sup> part discusses the basics of New IP. Section II introduces the New IP. Section III will talk about New IP based V2X architecture.
- 2<sup>nd</sup> part focus on the new IP based IVN details that includes Sections IV to IX. Section IV reviews the current technologies for IVN. Sections V, VI and VII will discuss the basics, architecture for control plane, and data plane respectively. Section VIII addresses the latency analysis and estimation. Section IX describes the network modeling and experiments.
- 3<sup>rd</sup> part is in Section X that will describe the conclusions.

## II. NEW IP INTRODUCTION

New IP is a broad technology set dedicated to solving requirements from future Internet, it is still in research stage and not mature. It was first proposed in ITU [4], and some research papers were published [5][6][7].

Compared with the existing IPv4 and IPv6, New IP has many forward-looking visions and will support some new features, such as

- Free Choice Addressing. Different size of IP address can be used for different use case. For the scenario that packet header overhead is a concern, such as in IOT network, a shorter than IPv4 or IPv6 address can be selected. For the extreme secured environment, invisible source address or longer than 128-bit randomized address can be used. This paper will not discuss this feature in detail. We still assume to use IPv4 for IVN. For IVN experiment in Sections VIII to IX, 32-bit IPv4 address is used for simulation.
- Deterministic E2E IP service. It can provide the guaranteed service to satisfy the pre-negotiated Service Level Agreement (SLA). New IP can be used for IVN and E2E V2X since both have very strict QoS requirements especially in bandwidth, latency, jitter, and packet loss that the current IP technology cannot meet.

New IP can coexist with other technologies in Internet, the traditional IP packets can still be processed and delivered in New IP networks. The interworking between New IP and IP networks can be easily provided by a proper gateway device between different networks. Migration to New IP network can take step by step gradually, we only need to upgrade the

network required to support new services that traditional IP network cannot support, so, the cost is limited.

As a summary, New IP is for Future Internet to provide services that the current Internet cannot provide. It is like the New Radio (NR) for 5G [8] in objectives, solutions, and technologies, see TABLE I for comparison.

TABLE I. 5G NR for 5G and New IP for Future Internet

	5G	Future Internet
Purpose and Requirements	<ul style="list-style-type: none"> <li>• eMBB [9]</li> <li>• mMTC [9]</li> <li>• uRLLC [9]</li> </ul>	<ul style="list-style-type: none"> <li>• Ultra-high through put</li> <li>• All things connected</li> <li>• High Precision Communication</li> </ul>
Solutions	<ul style="list-style-type: none"> <li>• New Radio (5G NR)</li> <li>• Service Based Architecture (SBA) [10]</li> </ul>	<ul style="list-style-type: none"> <li>• New IP</li> </ul>
Technologies	<ul style="list-style-type: none"> <li>• New spectrum</li> <li>• MIMO [8]</li> <li>• New protocol stack at UE</li> <li>• 5G NR QoS [8]</li> <li>• Grant Free Dynamic Scheduling</li> </ul>	<ul style="list-style-type: none"> <li>• Flexible addressing</li> <li>• Network Layer Multiple path</li> <li>• New protocol stack at host and UE</li> <li>• In-band signaling</li> <li>• New queuing and scheduling</li> </ul>

There could be different technologies developed for New IP for different use cases. The paper [7] has proposed key New IP technologies to realize the E2E guaranteed service for Internet, details are as following:

- In-band signaling. This is a control mechanism to provide a scalable control protocol for flow level guaranteed service. The key part of In-band signaling is that the control messages are embedded into the user data packets. With such binding, when the user data packets travel through a network, the control messages can be fetched by each network device on the path and control the behaviors of expected devices accordingly. Since all QoS metrics (bandwidth, latency, jitter, packet loss) are majorly determined by each network device on how user data packets are processed, accurately control network devices on path is the best way to achieve the best service a network can provide to applications. In traditional way, such controls are provided by separate protocols (sometimes called out-of-band signaling), the complexity is high and the scalability are limited. Through in-band signaling, the QoS path setup, SLA negotiation, Resource Reservation, QoS forwarding state report and control are accomplished without running extra control protocol like RSVP [11] for IP, or Stream Reservation Protocol (SRP) [12] for TSN [13]. The details of In-band signaling is described in [7].
- Class based queuing and scheduling. It uses the concept of Class as defined in Differentiated Service (DiffServ) [14] to identify different types of traffic. Different class of traffic is queued into different queues for differentiated service. Priority Queuing (PQ) combined with Deficit Weighted Round Robin (DWRR) or any other Weighted

Fair Queuing (WFQ) are used. Compared with other algorithms, this is the simplest to be implemented in high-speed hardware, and can achieve very satisfactory QoS in bandwidth, latency, jitter, and packet loss ratio. It also solves the scalability issue in Integrated Service (IntServ) [15] where the per-flow queueing was used.

- New TCP/UDP transport stack for end devices. The current TCP/UDP transport protocol stack was designed based on the best-effort service from IP. Enhanced protocol stacks are expected to obtain the benefits if the network can provide guaranteed service while keep the backward compatibility.

Above technologies set can be used by different combinations for IVN and V2X. For V2X, all technologies could be used. But for IVN, control methods (such as SDN controller) other than In-band signaling can also be used.

### III. NEW IP BASED V2X ARCHITECTURE

5G has defined that the End-to-End latency (uRLLC) is the Round-Trip Time (RTT) of IP packets transmitted from User End Device (UE) to the N6 interface in the 5G network [16]. The N6 interface is the reference point between UPF and Data Network (DN). It is obvious that uRLLC does not include the latency occurred in UE and DN.

The latency in UE is that when IP packet left application, it takes some time before the scheduler will send the packet to outgoing physical interface, this delay is significant when the UE has multiple applications running since different IP flow will compete the resource to get service from Operating System.

The latency for Data Network is the time spent for IP packets traveling from N6 interface to the IP (IPv4 or IPv6) destination. The destination can be any IP address in Internet, for example, a server inside a cloud. Normally, this latency is significant and is much bigger than inside a 5G network.

Same behaviors will apply to other QoS characters. The insufficient bandwidth, waiting for resource, and resulted jitter and packet loss happened in DN is normal and significant.

The root cause of above QoS degradation in data network for IP is because all IP packets are treated equally on the path the packet is traveling. Every IP packet is competing for the network resource, this will result in unexpected congestion, queue built up and even packet loss when queue is full. Even there are many technologies to mitigate or fix the problem, such as different congestion avoidance algorithms studied for long time [17], TSN [13], L4S [18], MPLS traffic engineering [19], etc. All these solutions are only working in a specific network but cannot be applied to Internet from real end to end (IP source to IP destination). It is insufficient to solve E2E latency issue in Internet if only considering specific network segments, such as wireless access network by 5G uRLLC [16] or Ethernet network by TSN [13].

The paper proposes to combine New IP technologies with 5G wireless technologies for the new architecture of future V2X communication.

To minimize the latency in UE, a new IP protocol stack is needed for UE. Figures 1 and 2 illustrate these stacks in wired and wireless device. The major changes for the new protocol stack are new socket or API. It is introduced for applications that require new service which is different with the traditional best-effort service using traditional socket. The new socket will pass application's service expectation to the network. The different flow with different service expectation will be queued to different queues, Latency Guaranteed Service (LGS) queue, Bandwidth Guaranteed Service (BGS) queue, or Best-Effort (BE) queue. System scheduler will serve different queue based on the priority and resource. Signaling Process module is to process the setup and forwarding state for in-band signaling. M-path control is for the multi-path support, it could split one flow into different network path, or replicate one flow couple of times to send to multiple network path. Multi-path feature can either increase the total bandwidth for application or compensate the packet loss due to the physical failure on one path. For a wireless device, an extra module will provide the interworking between New IP and New Radio (Figure 2). This module will coordinate the mapping between L3 multi-path and multiple Bearer introduced in 5G NR.

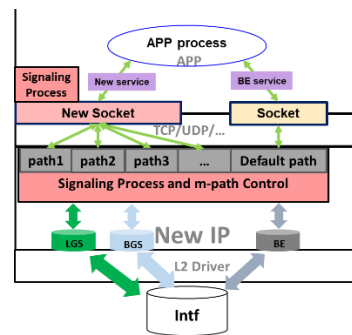


Figure 1. The New IP protocol stack for a computer or ECU.

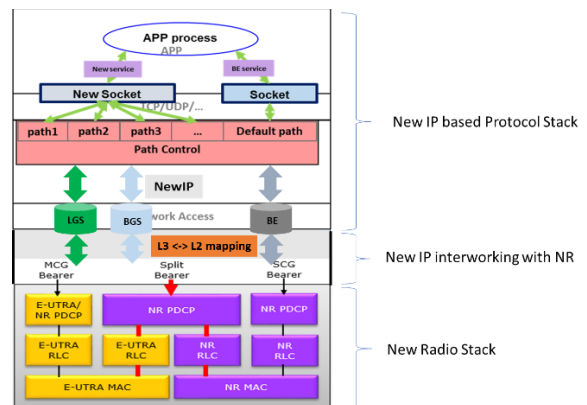


Figure 2. The New IP protocol stack integrated with 5G New Radio (NR) protocol stack for wireless device.

To minimize the latency in Data Network, the in-band signaling initiated from UE can pass through all network and reach the IP destination. This mechanism provides a simpler and more scalable control mechanism to provision a true end-to-end guaranteed service for any IP based application. When encountering a heterogeneous network (Ethernet, MPLS or other types), the in-band signaling carried in IP packet can be retrieved and used to interwork with other protocols, such as SRP for TSN, RSVP-TE for MPLS, etc.

Figure 3 illustrates New IP enabled V2X architecture in future Internet where IVN, 5G and wired data network in Internet all need New IP enabled, with such architecture, the true E2E deterministic service can be realized. It should be noted, for the case of directly communication (DSRC or Direct C-V2X), the architecture will only have IVN and V2X.

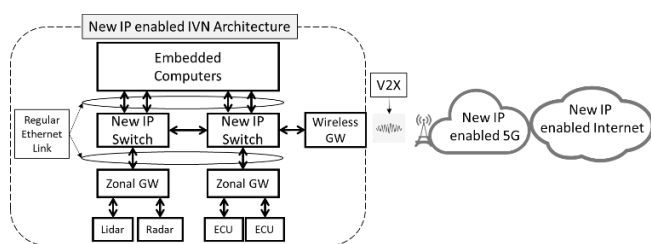


Figure 3. New IP enabled IVN architecture in future Internet.

Compared to the traditional V2X architecture which only address the wireless technologies, the new architecture shown in Figure 3 has New IP enabled networks including IVN, 5G and Internet. Only after the integration of those new IP enabled network, the true E2E service can be guaranteed for new applications.

In above picture, how to use New IP for each segment of network has many technical details. Due to the space limit, the paper cannot go to details for each, but will only focus on the case that New IP is used for IVN. We select IVN as an example is because the traditional IVN did not use IP, it normally uses some legacy protocols because of the stringent latency and packet loss requirement. The paper will demonstrate and prove that the New IP can provide the satisfactory deterministic service that the traditional IP cannot provide, and this service will satisfy the latency requirement of IVN.

#### IV. REVIEW OF CURRENT IVN TECHNOLOGIES

The section will brief the networking protocols used in current IVN and analyze the latency requirement for IVN.

##### A. Network technologies in current IVN

Most of the current IVN uses the legacy protocols, such as Local Interconnect Network (LIN) [20], Controller Area Network (CAN) [21], FlexRay [22]. These are specifically L2 technologies, they use the special designed physical media, signaling to manage strictly and timely for data to satisfy the requirements for communications inside car.

When more and more IP based applications come to IVN, the disadvantage of above legacy protocols is obvious. Its cost is normally higher than the TCP/IP plus Ethernet based network, IP based application must re-write the interface with new underlayer network if it is not Ethernet. AutoSAR [23] has proposed all IP based interface for IVN, and IP based IVN was proposed in [24][25].

However, without special technology, traditional TCP/IP and Ethernet cannot satisfy the requirement of IVN in terms of QoS. That is why IEEE TSN [13] was also proposed for IVN [26].

##### B. Requirement for IVN

The most important requirement in terms of QoS for IVN is the communication latency, jitter, and packet loss ratio.

The latency is crucial to the safety of vehicle and will determine if a new technology can be used in IVN. So far, there is no industry standard or requirement for the latency for IVN. Below are some existing publications about the topic:

- From the perspective of fastest human reaction time, the IVN latency must not be slower than that. It is said the fastest human reaction time is 250ms [27]. Some papers gave lower values but not shorter than 100ms if human brain is needed to process the input signal.
- The paper [26] mentioned the latency for control data must be less than 10ms. The papers [24] and [28] said the latency for control data must be less than 2.5ms.

Based on all available analysis, it is safe to assume that the qualified IVN must support the E2E latency not bigger than 2.5ms. During this short time, a car with a speed of 200 km/s will only move 0.138m.

There is no requirement for the jitter from current research. Theoretically, jitter can be removed by buffering technology when the maximum latency is within the target.

The zero-packet-loss is expected for control data. In a packet network (Ethernet or IP), the packet loss is normally caused by two factors: (1) the congestion in network (2) physical failure, such as link, node, hardware. The 1<sup>st</sup> factor has much higher occurrence probability and higher packet loss ratio than the 2<sup>nd</sup> factor. Thus, it must be eliminated for control data in New IP based IVN. The loss by 2<sup>nd</sup> factor can be mitigated or eliminated by sending the same data to two or multiple disjointed paths to reach the same destination, and/or, sending the same data more than one time as long as the time period is chosen below the upper bound of the latency.

#### V. THE IVN ARCHITECTURE - INTRODUCTION

The new architecture of IVN is based on New IP technologies and consists of Control plane and Data Plane. This section will discuss some basics for architecture.

##### A. Topologies

The topologies of new IVN can be any type, but to reduce the complexity and to provide a redundant protection, the paper proposes to use two topologies, one is the Spine-Leaf

topology, and another is Ring topology. They are shown in Figure 4 and Figure 5, respectively.

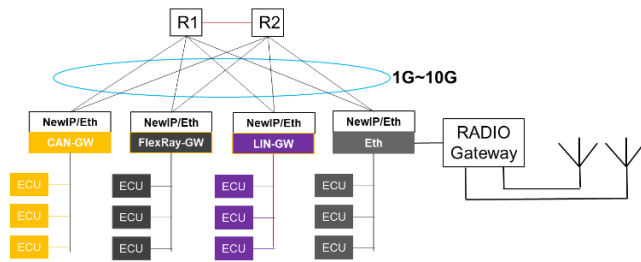


Figure 4. The Spine-Leaf IVN topology.

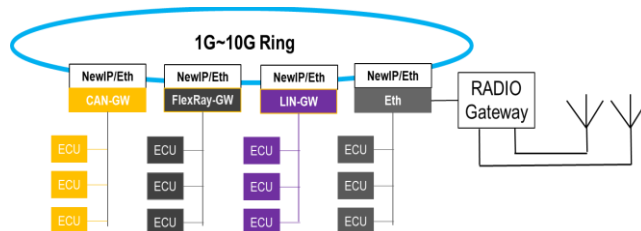


Figure 5. The Ring IVN Topology.

In the topologies illustrated in Figures 4 and 5, there are always two disjointed physical paths between any network devices. Also, the Ethernet Bus is supported. The advantages of such design are:

- The protection of physical link. Any failure of any link does not completely stop the communication.
- The higher reliability for zero packet loss. Multiple paths of New IP can be used to transport critical packet to two paths to compensate possible packet loss due to temporary failure or fault in one physical transmission media.
- Ethernet Bus can make the plug-and-play possible for most of sensors, ECU, computers, etc.

#### B. Network Device and Link

The network device can be either IP Router or Ethernet Switch. IP router is more powerful to provide more features in networking, such as more flexibility in routing and network state changes, higher link utilization, secured communication, etc.

When Ethernet Switch is selected, DPI (Deep Packet Inspection) should be configured to check the IP level information (address, port, protocol, DSCP values) for admission control for IP flows.

The Physical Link and protocol can be any type of Layer 2 link, Normal Ethernet or IEEE802.1 with the speed higher than 100 Mbps is minimum, and 1G ~10G is better to achieve a shorter latency. There is no need to select any special IEEE802.1Q serials, such as TSN. This is one of the advantages of the new architecture compared with TSN and other legacy protocols (LIN, CAN, FlexRay, etc). It not only provides more flexibility in device development and technology selection, but also save the cost for V2X

applications, since IP is more general technology that fits most of existing application's interfaces. In addition to that, IP device is normally cheaper than legacy device especially in higher speed.

#### C. Backward Compatibility

The legacy protocol LIN, CAN and FlexRay are still supported in the new IVN architecture. As shown in Figures 4 and 5, legacy ECUs used for legacy protocols can still be attached to the legacy bus. The New IP based network node will have an interworking function to support the legacy protocols. Figure 6 illustrates a Gateway board with two interfaces: Ethernet and FlexRay, and another board only has Ethernet interface. Two board can be connected by Ethernet interface. The ECU attached to the FlexRay bus can communicate with any application running in both boards on top of New IP.

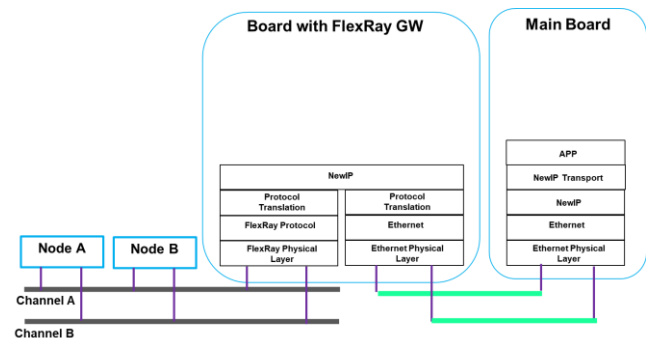


Figure 6. Interworking between Ethernet and FlexRay.

#### D. New Service

The new service provided by New IP based IVN is "E2E flow level guaranteed service for bandwidth, latency, jitter and packet loss". Following is detail about the new service:

- The E2E is defined as "From Application(s) of one end-user device to other Application(s) of another end-user device. For IVN, the end-user device is any device connected to IVN that supports TCP/IP protocols, and application is running on top of TCP/IP, such as TCP/IP capable ECU, Embedded computer, Infotainment system, Mobile device, etc.
- The Flow can be any granularity, for example, it can be an IP flow defined by 5 tuples (source/destination address, source/destination port number, protocol), or a group of flows defined by less tuples, such as source/destination address.
- The Guaranteed service means that the service provided by system will go through some crucial steps like Service Level Agreement (SLA) negotiation or provisioning, admission control and user traffic conformity enforcement, etc. After all procedures are accomplished, the promised service will meet the negotiated bandwidth, latency, jitter, and packet loss defined in SLA.

- Different application may need different guaranteed service. For example, critical sensor and control data may need the guaranteed service for both bandwidth and latency. The new service is like the service for Scheduled traffic and Real-time traffic defined in FlexRay [22]. For these types of traffic, the strictest service is needed to achieve the minimum latency, jitter, and packet loss ratio. almost all other type of data does not need any guaranteed service, the best-effort service is good enough. For any application, weather it needs the new service is case by case and up to the application's requirement from the networking.

## VI. ARCHITECTURE- CONTROL PLANE

This section discusses the aspects of control plane for new IVN architecture including the Control Plane Candidates, and Control Plane Functions.

### A. Control Plane Candidates

The control plane could select the following candidates:

- Central controller: such as SDN controller or network management controller. For IVN, it is normally a controller's responsibility to provision some basic function for IVN, such as address assignment, routing protocol configuration (for dynamic routing) and static routing table installation (for fast and simple system boot up). Central controller can also be used for the static provisioning for the guaranteed service, such as scheduled and real-time traffic configuration on ECUs,
- In-band signaling protocol [7] is an alternative control method distributed to all network nodes. It can be used for connections between IVN and cloud for critical data in V2X scenario, it can also be used in IVN for dynamic service state report, network state OAM and network problem diagnosis. In-band signaling is not mandatory for communication within IVN.

### B. Control Plane Functions

In addition to the static provisioning from a central controller described in A, another key function for the control plane to achieve the guaranteed service support is the Admission Control. All flows requesting new service, except the Best Effort, must obtain the approve for the admission from central controller or from in-band signaling process. This includes three steps:

- An application requesting new service specifies the expectation of service type (BGS, LGS), the traffic pattern (rate specification) and expected End-to-End latency.
- System (Central controller or the network device) will process the request and try to reserve the resource for the flow, and notify the application about the CIR (Committed Information Rate), PIR (Peak Information Rate), bounded end-to-end latency and jitter values, packet loss ratio, etc.
- The application agreed the offered service will send traffic according to the system notification, i.e., send traffic no

more than CIR, and monitor the notification from network to adjust the traffic pattern accordingly.

## VII. ARCHITECTURE - DATA PLANE

This section discusses the aspects of data plane for new IVN architecture including the Protocol Selection, Queuing and Scheduling Algorithm, Traffic shaping, Latency estimation.

### A. Protocol Selection

As new IVN is IP based, IPv4 is proposed to be the basic protocol for New IP, a protocol extension is needed if in-band signaling is used [29]. All data process, such as forwarding, traffic classification, traffic shaping, queuing, and scheduling, are for IPv4 data. It is noted that New IP's "Free address choice" feature can provide address shorter than IPv4 that can benefit the latency, but it is not discussed here.

### B. Traffic Classification and Services

This paper will propose to classify all IVN traffic as four types:

- Scheduled traffic (ST). This type of traffic has fixed data size, exact time of when the data is starting and what is the interval of the data. Normally, all sensor data report and control data belong to this type. Typically, IVN can configure the polling mechanism for all sensors to make use of this type of traffic. The service associated with this type of traffic will get LGS. This type of traffic is classified as EF class in DSCP value defined in DiffServ.
- Real-Time Traffic (RT). This type of traffic has fixed data size, but the time of the data starting, and the data rate is unknow. Normally, all urgent sensor data report and control data belong to this type. IVN can configure the critical sensors to send data to controller in the situation of emergency and the polling mechanism did not catch the latest data changes. The service associated with this type of traffic is also LGS. But the latency and jitter might be a little bigger than for the ST depending on the algorithm and burst of RT. This type of traffic is classified as AF4x class in DSCP value.
- Bandwidth Guaranteed Traffic. This type of traffic has special requirement from the network bandwidth, but not the latency, jitter, and packer loss ratio. Normally, the IVN software update from cloud, diagnosis data uploading to cloud, on-line gaming and streaming for infotainment system, etc., belong to this type. It can be classified as any AFxy class (other than EF and AF4x) in DSCP value.
- Best-Effort Traffic. This is a default class of traffic, all applications that do not require any special treatment from network perspective can be classified as this type of traffic, Best Effort Class is used.

There are four types of services in IVN corresponding to the above four type of traffic. TABLE II shows QoS Characters and Use Case for different type of services. Both

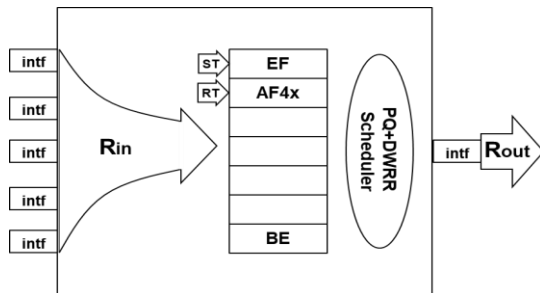
Scheduled Traffic (ST), Real-Time Traffic (RT) are treated by Latency Guaranteed Service (LGS) as described in [7]. The traffic that only needs the bandwidth guarantee is treated by Bandwidth Guaranteed Service (BGS). Other types of traffic are treated by Best-Effort Service (BES)

TABLE II. FOUR TYPE OF SERVICE AND QOS CHARACTERS

Service Type	QoS Characters	Use Case
LGS for Scheduled Traffic	<p><b>Bandwidth:</b> Network guarantees the bandwidth is within (CIR, PIR)</p> <p><b>Latency:</b> Most precise. Network guarantees E2E bounded latency</p> <p><b>Jitter:</b> Approximately zero</p> <p><b>Packet Loss:</b> Almost Zero</p> <ul style="list-style-type: none"> <li>• Congestion-free</li> <li>• Lossless queuing</li> <li>• Multi-path to prevent drop from physical failure</li> </ul>	Asynchronous or Synchronous communication: Critical sensor and control data
LGS for Real Time Traffic	<p><b>Bandwidth:</b> Network guarantees the bandwidth is within (CIR, PIR)</p> <p><b>Latency:</b> Minimized. Network guarantees E2E bounded latency</p> <p><b>Jitter:</b> 1/2 of E2E bounded latency</p> <p><b>Packet Loss:</b> Minimized</p> <ul style="list-style-type: none"> <li>• Congestion-free</li> <li>• Lossless queuing</li> <li>• Only drop when physical failure</li> </ul>	Asynchronous communication: Critical sensor and control data
BGS for bandwidth sensitive traffic	<p><b>Bandwidth:</b> Network guarantees the bandwidth is within (CIR, PIR)</p> <p><b>Latency:</b> Less important</p> <p><b>Jitter:</b> Less important</p> <p><b>Packet Loss:</b> Don't care</p>	Un-critical data
BES for other type of traffic	Don't care	Other data

C. Queuing and Scheduling Algorithm

The paper proposes two types of algorithms illustrated in Figures 7 and 8. One is for asynchronous environment that there is no clock sync for network. Another is synchronous environment that clock is synced with certain accuracy for IVN including all devices. Below are details, also, the experiment section is based on the two algorithms discussed here.

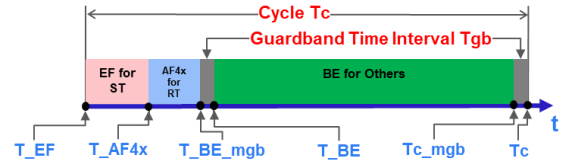
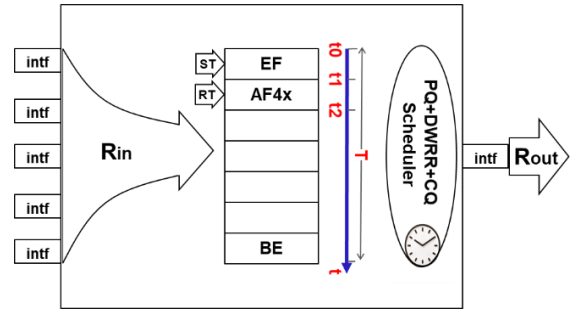


Algorithm 1: Asynchronous Class Based Scheduler

```

Packet *pkt scheduler () {
//Scheduler function. The EFQ has the highest priority, AF4xQ and other Q
have lower priority and are served by DWRR
1. while EFQ.length() > 0 do //serve the EF queue
2.   EFQ.dequeue(pkt)
3.   return(pkt)
4. while AF4xQ.length() > 0 do //serve the AF4x queue
5.   W_AF4x -> W_AF4x' //update weight W for AF4x
6.   if W_AF4x' < W_AF4x then //updated W < assigned W for AF4x
7.     AF4xQ.dequeue(pkt)
8.     return(pkt)
9.   else
10.    continue
11. while BEQ.length() > 0 do //serve the BE queue
12.   BEQ.dequeue(pkt)
13.   return(pkt)
14.}
    
```

Figure 7. 1<sup>st</sup> Algorithm and psudo code: Asynchronous Solution.



Timer:

- T<sub>EF</sub>**: The time when EF class is started to be served
- T<sub>AF4x</sub>**: The time when AF4x class is started to be served
- T<sub>BE</sub>**: The time when BE class is started to be served
- T<sub>c</sub>**: Cycle time interval
- T<sub>gb</sub>**: Time interval for Guard-band
- T<sub>BE\_mgb</sub> = T<sub>BE</sub> - T<sub>gb</sub>**
- T<sub>c\_mgb</sub> = T<sub>c</sub> - T<sub>gb</sub>**

Algorithm 2: Synchronous Class Based Scheduler

```

Void timerProcess (TIMER ExpiredTimer) {
//Timer process function, process events when a timer expired. When a timer
expired, the associated gate is open, then the scheduler can schedule the traffic
for the class. Example only shows three classes.
1. if ExpiredTimer == T_EF then //Timer for T_EF is expired
2.   openGate = EF //Open the gate for EF
3.   if isTimerRunning() != true then
4.     startTimer(T_AF4x) //Start next timer for T_AF4x
5. else if ExpiredTimer == T_AF4x then //Timer for T_AF4x is expired
6.   openGate = AF4x //Open the gate for AF4x
7.   if isTimerRunning() != true then
8.     startTimer(T_BE_mgb) //Start next timer for T_BE_mgb
    
```

```

9.  else if ExpiredTimer == T_BE_mgb then
10.                                     //Timer for T_BE_mgb is expired
11.  openGate == NONE                    //Close the gate for all
12.  isTimerRunning() != true then
13.  startTimer(T_BE) /                  //Start next timer for T_BE
14.  else if ExpiredTimer == T_BE then  //Timer for T_BE is expired
15.  openGate = BE                       //Open the gate for BE
16.  isTimerRunning() != true then
17.  startTimer(Tc_mgb)                 //Start next timer for Tc_mgb
18.  else if ExpiredTimer == Tc_mgb then
19.                                     //Timer for Tc_mgb is expired
20.  openGate = NONE                    //Close the gate for all
21.  isTimerRunning() != true then
22.  Increment all timer by Tc          //Increase all timer by one Tc
23.  startTimer(T_EF)                  //Start next timer for T_EF
24.}
25.
26.Packet *pkt scheduler () {          //Scheduler function
27. while EFQ.length() > 0 and
28. openGate == EF do                  //serve the EF queue
29. EFQ.dequeue(pkt)
30. return(pkt)
31. while AF4xQ.length() > 0 and
32. openGate == AF4x do                //serve the AF4x queue
33. W_AF4x -> W_AF4x'                  //update weight W for AF4x
34. if W_AF4x' < W_AF4x then          //updated W < assigned W for AF4x
35. AF4xQ.dequeue(pkt)
36. return(pkt)
37. else
38. continue
39. ....                               //serve other queues
40. while BEQ.length() > 0 and
41. openGate == BE do
42. BEQ.dequeue(pkt)                  //serve the BE queue
43.}

```

Figure 8. 2<sup>nd</sup> Algorithm and psudo code: Synchronous Solution.

- For asynchronous environment, Priority Queuing (PQ) combined with Deficit Weighted Round Robin (DWRR) or any type of Weighted Fair Queuing (WFQ) are used. It is called the 1st Algorithm in the document thereafter. Normally, the time sensitive flows, i.e., scheduled traffic (EF class) and real-time traffic (AF4x class) are put into the 1st and 2nd priority of the queue, and other classes of traffic, BGS and Best Effort class of traffic, are put into the lower priority queues. For admission control and scheduler configuration, the total CIR for LGS class, and the WEIGHT values of BGS class can be calculated from the sum of CIR of all flows in the same class. This algorithm has already deeply analyzed in [7].
- For synchronous environment, above asynchronous PQ+DWRR algorithm is combined with Cyclic Queuing (CQ). It is called the 2nd Algorithm in the document thereafter. Each class of traffic has a dedicated time window to be served by the scheduler. The service time is associated with the sum of CIR of all flows in the same service. The Scheduler will calculate and adjust the serving time window for each class when a flow's state is changed, such as new flow is added, or old flow is removed. The guard-band is added for lower priority classes to guarantee the EF class traffic, when served, is not blocked by lower priority traffic on wire. In another word, when EF class is served, the wire is always available

for transmission. The guard-band timer interval can be calculated as the required time to transmit one maximum size of packet on wire speed.

#### D. Traffic Shaping

Traffic shaping is used to absorb the overflow and burst of the traffic in the class and its objectives are: (1) the packet in the class is never built up, thus reducing the latency (2) traffic in lower priority class is never starved by higher priority traffic. Existing Single Rate Three Color Marker [30] or Two Rate Three Color Marker [31] could be used for traffic shaping. Other type shaping like leaky bucket shaping can also be used. Traffic shaping deployment is very flexible. It can be configured in both ingress and egress interface. It can be per flow based, or per class based.

Flow-level traffic shaping in ingress interface can also be used as the policy enforcement module, it will check the user's traffic to see if it is allowed to pass or trigger some policy, such as discard or put into lower priority to process.

### VIII. LATENCY ANALYSIS AND ESTIMATION

To provide the Latency Guaranteed Service (LGS) for ST and RT, the network must be able to estimate the latency for a network path and offer to user in the provisioning stage. This is the requirement for SLA negotiation. This section will analyze all factors that can result in network latency and discuss some basic formulas.

#### A. The Latency Analysis for IP Network

In this paper, the latency estimation is for E2E from the perspective of user's application. The latency must include all delay occurred in network and hosts. This is illustrated in Figure 9. The formula for the latency is as in (1) and (2). The superscript "LGS" denotes LGS packet.

$$D_{e2e}^{LGS} = PD + \sum_{i=1}^n (OD_i^{LGS} + QD_i^{LGS}) + \sum_{s=1}^m SD_s^{LGS} = t1 - t0 \quad (1)$$

$$SD_s^{LGS} = L^{LGS} * 8/R_{out} \quad (2)$$

- $t0$ : the time the 1<sup>st</sup> bit of a pack is leaving the application process on the source host.
- $t1$ : the time the 1<sup>st</sup> bit of the pack is received by the application process on the destination host.
- $PD$ : Propagation delay, this delay is limited by the speed of signaling in a physical media. For example, it is approximately 200k KM/s in optical fiber.
- $OD_i$  : The other delays (pack process, deque, de-capsulation, lookup, switch, L2-rewrite, encapsulation, etc.) at the  $i$ -th hop and source host. This delay is related to the Forwarding Chip and hardware, it is normally and relatively steady for a specified router or switch and can be easily measured. This delay is insignificant compared with  $QD$  and  $SD$  described below.



- $QDi$ : The queuing delay at the  $i$ -th hop and source host.
- $SDs$ : The serialization delay at the  $s$ -th link segment, it can be calculated by the formula (2).  $L^{LGS}$  is the packet length (byte) for the LGS flow.  $R_{out}$  is the link speed.

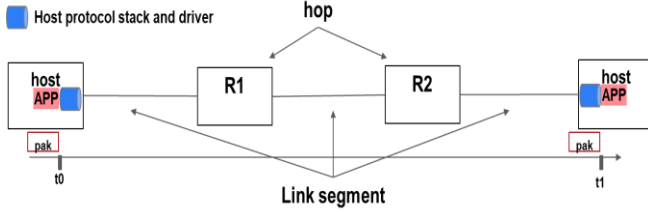


Figure 9. The End-to-End Latency for IP Applications.

### B. Estimation for the Queuing Latency (QD)

The formulas for the queuing latency estimation (for the same packet size) have been derived in [7] for the 1<sup>st</sup> Algorithm. In this paper, different packet size for two class is used, thus formulas are different as in [7]. The maximum number of packet and queuing time for a queue (EF or AF4x) under the worst scenario for a hop are shown in equations from (3) to (8).

$$N_{max}^{EF} = \lceil R_{in}^{EF} / R_{out} * (L_{max}^{LOW} / L_{max}^{EF} + 1) + 1 \rceil \quad (3)$$

$$D_{max}^{EF} = N_{max}^{EF} * L^{EF} * 8 / R_{out} \quad (4)$$

$$N_{max}^{AF4x} = \lceil R_{in}^{EF} / R_{out} * (L_{max}^{LOW} / L_{max}^{EF} + 1) + 1 \rceil + \lceil (R_{in}^{AF4x} / R_{out} * (L_{max}^{LOW} / L_{max}^{AF4x} + 1) + 1) * (R_{in}^{AF4x} / R_{out}) \rceil \quad (5)$$

$$D_{max}^{AF4x} = N_{max}^{AF4x} * L^{AF4x} * 8 / R_{out} \quad (6)$$

$$R_{in}^{EF} = r_{EF} \sum_{i=1}^m cir_i^{EF} \quad (7)$$

$$R_{in}^{AF4x} = r_{AF4x} \sum_{i=1}^n cir_i^{AF4x} \quad (8)$$

For the 2<sup>nd</sup> Algorithm, the packet in any queue is served on a pre-allocated time window, and this will guarantee that flows will not be interfered by any packets in other queues. So, it is easy to estimate that the maximum number of packets in a queue is as in (9), (10). The associated queuing time is the same as in (4) and (6). However, for the worst scenario when a packet is out of the allocated window for some reason, the maximum latency will be as the (11).

$$N_{max}^{EF} = \lceil R_{in}^{EF} / R_{out} + 1 \rceil \quad (9)$$

$$N_{max}^{AF4x} = \lceil R_{in}^{AF4x} / R_{out} + 1 \rceil \quad (10)$$

$$D_{max}^{EF} = D_{max}^{AF4x} = T \quad (11)$$

The symbols and parameters in the formulas above are described as below,

- The symbol “ $\lceil \ ]$ ” is the rounding up operator.
- $N_{max}^{EF}$ : the maximum queue depth for EF queue.

- $N_{max}^{AF4x}$ : the maximum queue depth for AF4x queue.
- $D_{max}^{EF}$ : the maximum queuing time for EF queue.
- $D_{max}^{AF4x}$ : the maximum queuing time for AF4x queue.
- $R_{in}^{EF}$ : the ingress rate for EF queue.
- $R_{in}^{AF4x}$ : the ingress rate for AF4x queue.
- $cir_i^{EF}$ : the Committed Information Rate (cir) for the  $i$ -th flow for EF queue.
- $cir_i^{AF4x}$ : the Committed Information Rate (cir) for the  $i$ -th flow for AF4x queue.
- $r_{EF}$ : the burst coefficient for the traffic of EF queue.
- $r_{AF4x}$ : the burst coefficient for the traffic of AF4x queue.
- $T$ : the cycle time for the scheduler when CQ is used.

## IX. NETWORK MODELING AND EXPERIMENTS

To verify and analyze the New IP based IVN architecture can meet the requirements of IVN, OMNeT++ [32] is used to simulate the network, the detailed bandwidth, E2E latency, pack loss, etc., can be retrieved from tests. OMNeT++ is very popular to simulate time driven events and activities involved in networking technologies, it can accurately calculate and simulate the life of each individual packet traveling from source to destination via different intermediate devices. So, its results in QoS metrics are very close to the theoretical estimations.

### A. Network Topology

The network is illustrated in Figure 10. It is a ring topology but with the cut of another ring link to focus on the latency simulation under the worst scenario (longer latency). All links speed is 100 Mbps. The network consists of ECU, computers, and routers. ECU is to simulate the sensors with control connected on Ethernet Bus. It has a full TCP/IP stack and is responsible for the ST and RT generation and process. The ST and RT are simulated by UDP packets. Computers are simulating the generation and process of Best-Effort traffics (TCP and UDP) that are used to interfere ST and RT between ECUs.

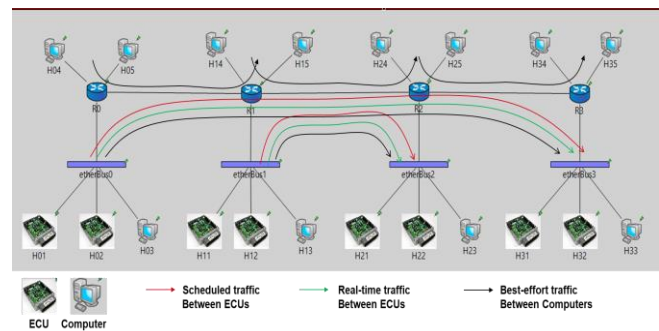


Figure 10. Network Topology and traffic.

The purpose of simulation is to illustrate the new architecture can provide the E2E guaranteed service for ST and RT flows when the network is severely congested and interfered by the Best-Effort traffic. The E2E guaranteed service includes three criteria: (1) bounded latency (2) bounded jitter (3) congestion free and lossless. Moreover, the tested latency and jitter for ST and RT should be close to the estimated latency described in section VIII.

### B. Network Devices

Each router consists of Ingress Modules, Switch Fabric and Egress Modules that are illustrated in Figure 11. The Ingress Modules simulate the traffic classification and ingress traffic shaping functions; The Egress Modules simulate the egress traffic shaping, queuing, and scheduling functions. The Switch Fabric Modules simulate the IP lookup, switching and L2 re-writing functions. Two types of schedulers are used. Only class level traffic shaping is used for ST for ingress and egress.

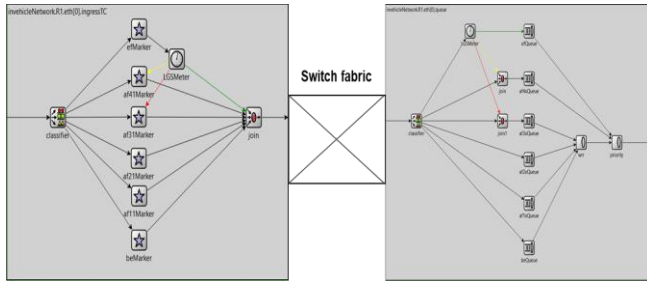


Figure 11. Router structure.

### C. Traffic Configuration

To simulate the worst scenario, very heavy traffic for the IVN simulation is configured as below:

- There is total 100 ST flows and 100 RT flows using UDP, each flow has the packet size 254 bytes (200 bytes data, 54 bytes of UDP and Ethernet header), the send interval is 10ms. So, each flow has a rate of 203.2 Kbps. Both rate for ST flows and RT flows are 20.32Mbps, it means the remained bandwidth for BGS, and BE is about 60Mbps.
- 50 ST flows and 50 RT flows are from ECU H01 and H02 to H31 and H32, these flows' results are checked and compared with the estimation. 50 ST flows and 50 RT flows are from ECU H11 and H12 to H21, H22.
- There is total 250 interference flows configured between other computers. The interference flows will cause all links between routers congested, R1 link Eth[0] is the most severely congested router and link. All flows packet size are 200 bytes or 1500 bytes. Both TCP and UDP are configured for interference flows.

### D. Cyclic Queueing and Scheduler Configuration

For the 2<sup>nd</sup> algorithm, the detail of the cyclic queuing is configured as in Figure 12.



Figure 12. The Cyclic Queueing Configuration.

- o The cycle T for all router and hosts are 10ms.
- o A guard-band of 1500 bytes or 120 us are configured for both AF4x and BE classes. 120 us is the time to transmit 1500 bytes packet on 100M bps link.
- o The time window size for EF and AF41 are 22% and 32% of the cycle T respectively.

### E. Experiment Results and Analysis for E2E Latency/Jitter

This sub-section will analyze the E2E latency/jitter for different type of traffic, compare the experiment results with the theoretical estimation made in Section VIII.

TABLE III shows the detailed calculation for the E2E latency estimation. First, estimate the maximum number of packets queued in each egress link of all routers on the path, then calculate the maximum queuing delay. The minimum E2E latency means there is no queuing latency in each hop, so it is determined by the sum of all link segment's serialization latency on the path. Each 100M link will have 20.3 us serialization latency for 254 bytes ST or RT traffic. The burst coefficient for each case is also shown in Table III. Higher coefficients for router R0 and R1 are selected since there are aggregation of the traffic for the routers. For other routers, the coefficient is selected as 1, or no burst effect.

TABLE III. THE E2E DELAY ESTIMATION OF ST AND RT FLOWS

Algorithm	Class and traffic	Estimated max number of packet 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+CC	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

TABLE IV shows the Min/Max E2E Delay for the worst performed flow, and estimation values also compared. The worst performed flow is defined as that the flow's Max E2E delay is the biggest in all flows in the same class.

Jitter is not shown in the table, but it can be easily calculated by the variation of mean and Min/Max value, the mean value can be simply calculated by the average of Min/Max values.

TABLE IV. THE COMPARISON OF EXPERIMENT RESULT AND ESTIMATION

Algorithm	Min/Max E2E Delay (us) for the worst performed flow carrying ST between H01/H02 to H31/H32		Min/Max E2E Delay (us) for the worst performed flow carrying RT between H01/H02 to H31/H32	
	Experiment	Estimation	Experiment	Estimation
PQ+DWR	108/391	100/405	278/542	100/465
PQ+DWR+CQ	109/152	100/262	169/169	100/262

Figures 13-16 illustrate the E2E delay changes with time for the worst performed flows shown in TABLE IV.

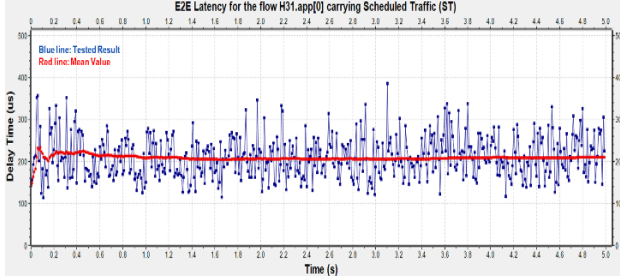


Figure 13. The 1<sup>st</sup> Algo: The E2E Latency (min=108us, max=391us) for the worst performed ST flow.

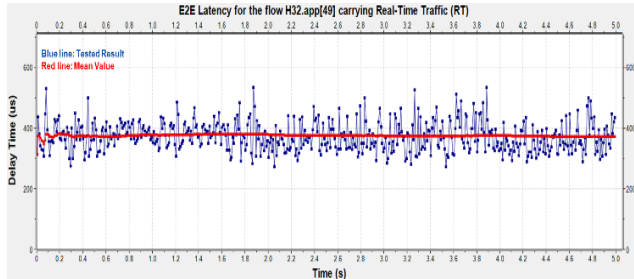


Figure 14. The 1<sup>st</sup> Algo: The E2E Latency (min=278us, max=542us) for the worst performed RT flow.

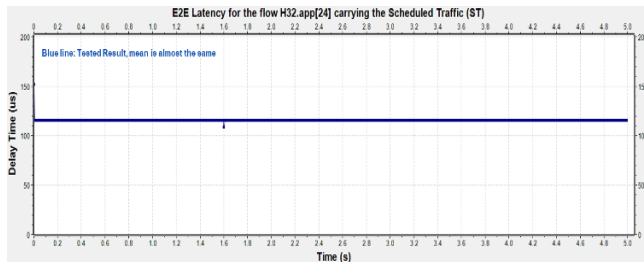


Figure 15. The 2<sup>nd</sup> Algo: The E2E Latency (min=109us, max=152us) for the worst performed ST flow.

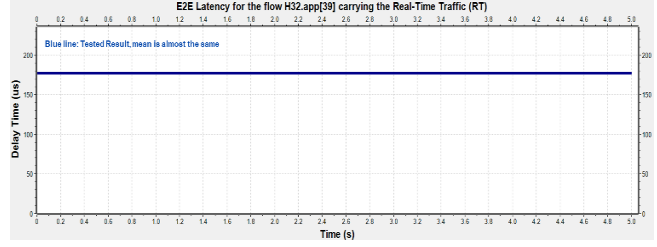


Figure 16. The 2<sup>nd</sup> Algo: The E2E Latency (min=169us, max=169us) for the worst performed RT flow.

### F. The Receiver's Instantaneous Bandwidth and Packet Loss Verification

This sub-section will verify there is no bandwidth loss for every flow. "No bandwidth loss" is verified by checking if receiver's instantaneous rate or bandwidth is similar to the sender's rate for every flow.

The receiver's Instantaneous Bandwidth (B) is calculated for each received packet at receiver side by the formulars (12) to (13), there are three scenarios :

- When there is only one packet received:

$$B = 0 \quad (12)$$

- When there are two packets received with different size in byte. At  $t_0$ , received a packet and its size is  $L_{t_0}$ . At  $t_1$ , received a packet and its size is  $L_{t_1}$ :

$$B = 0.5 * (L_{t_0} + L_{t_1}) * 8 / (t_1 - t_0) \quad (13)$$

- When there are more than two packets received with different size in byte. Three packets are sampled for calculation: At  $t_0$ , received a packet and its size is  $L_{t_0}$ . At  $t_1$ , received a packet and its size is  $L_{t_1}$ . At  $t_2$ , received a packet and its size is  $L_{t_2}$ :

$$B = (0.5 * L_{t_0} + L_{t_1} + 0.5 * L_{t_2}) * 8 / (t_2 - t_0) \quad (14)$$

For the test for Algorithm 1, five ST flow's sending rate are set differently at source, two have constant rate and three have variable rate.

For the test for Algorithm 2, five ST flow's sending rate are constant. It is hard to set the rate to be variable for algorithm 2 since if a packet is not sending at its allocated time window, there will be extra delay of time cycle. This will impact the analysis for the instantaneous bandwidth.

The paper only demonstrates the bandwidth for ST flows. The results for RT flows are similar.

Figures 17 to 20 illustrate the instantaneous rate or bandwidth for the five ST flows for two algorithms respectively. It is obvious that each flow for two algorithms has almost same wave shape. It indicates that the receiver's instantaneous rate is almost the same as the sender's rate, so there is no bandwidth loss for the network.

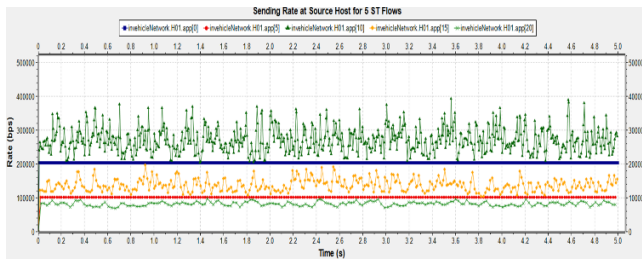


Figure 17. The 1<sup>st</sup> Algo: The Sender’s Instantaneous Bandwidth for 5 ST flow.

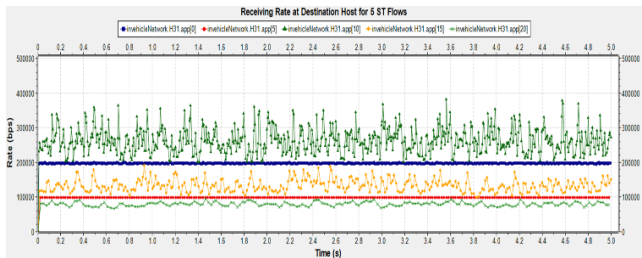


Figure 18. The 1<sup>st</sup> Algo: The Receiver’s Instantaneous Bandwidth for 5 ST flows.

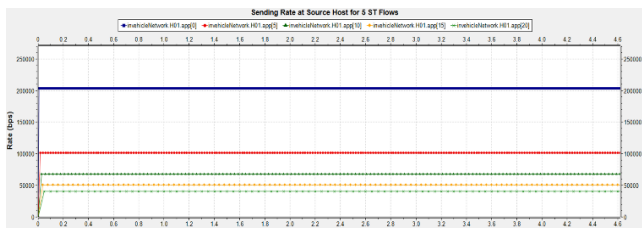


Figure 19. The 2<sup>nd</sup> Algo: The Receiver’s Instantaneous Bandwidth for the worst performed ST flow.

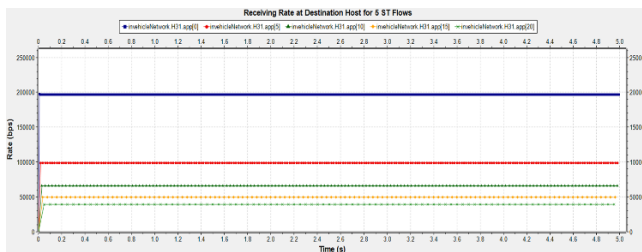


Figure 20. The 2<sup>nd</sup> Algo: The Receiver’s Instantaneous Bandwidth for the worst performed RT flow.

This sub-section will also verify there is no packet drop from queuing and congestion. To demonstrate the lossless and congestion-free for ST and RT flows, Figure 21 shows the statistics of all queues in R1 for two algorithms. No packet dropped in EF and AF4x queues while there are packets dropped in BE queue. This is as expected, congestion should only happen for BE traffic, ST and RT flows are not impacted and are lossless and congestion-free. R1 is the most severely congested, other Router’s queues also have similar pattern. No packet drops for EF and AF4x.

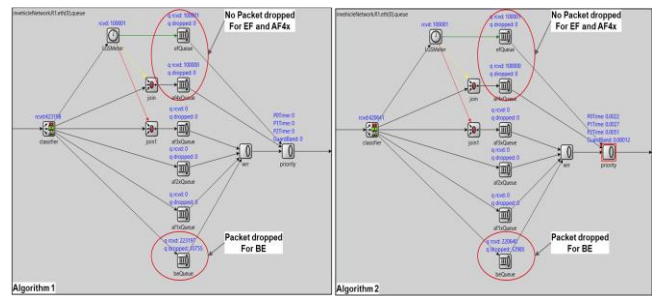


Figure 21. The statistics for all Queues for two algorithms

Here is a summary from the test results:

- The queuing latency of higher priority queues by PQ is very short and is not impacted by the congestion of lower priority class of traffic. E2E Maximum latency estimation in Section VIII can be used as the rough prediction for almost all traffic’s real maximum E2E latency.
- Lossless and congestion free can be achieved for ST and RT flows if the admission control is done for the flows. When the total rate for ST and RT flows are below the CIR of service expectation has claimed, there will be no packet drop caused by queue overflow.
- The E2E latency shown in the experiment does not include “Other Delay” and “Propagation Delay” described in Section VIII. “Propagation Delay” is very trivial in IVN, but “Other Delay” should be considered and added up if they are significant compared with the final queuing latency. For most of forwarding chip, “Other Delay” is very small and below hundred microseconds, but for x86 based virtual router, it might not be true depending on the forwarding software design.
- The latency per hop is inversely proportional to the link speed. For example, the experiment using 100M link with 4 hops network can achieve hundreds microsecond for E2E latency. It is expected that the corresponding latency for the same network is about tens of microsecond and couple microseconds for 1G and 10G link, respectively. Higher link rate will not only reduce latency, but also provide more bandwidth for non-time-sensitive applications. So, the paper proposes to use at least 1G link for the IVN in the future.

## X. CONCLUSIONS

The paper has proposed a new architecture for future V2X communication, that is based on the integration of New IP and 5G Technologies. Unlike the 5G uRLLC that is only limited in wireless network for its end-to-end definition, The new V2X architecture can provide a real end-to-end guaranteed service for bandwidth, latency, jitter and packet loss. The “real end-to-end” will cover all segments of network including user end device (UE) associated with IP source, wireless access, wireless core network, data network and to another user end device or computer in Internet associated with IP destination.

The paper also analyzed the detailed requirements for the In-Vehicle-Network in terms of QoS characters. The paper proposed to use New IP for future IVN. Class based queuing

and scheduling plus traffic shaping can provide per-hop LGS and BGS. Combined with Central Controller or In-band Signaling, the E2E guaranteed service for new IVN can be achieved by enforcing the per-hop guaranteed service on all network devices on the IP forwarding path. The solution is backward compatible as the existing IP traffic and traditional best effort service can coexist with the new classes of traffic and new services.

To prove the concept, the paper also discussed in detail about the experiments of network modelling on New IP based IVN. The simulation has demonstrated that the New IP can satisfy very stringent QoS requirements for IVN. The results indicate the future IVN can obsolete diversified legacy protocols and unify to one protocol: New IP. This will dramatically reduce the cost of IVN.

The paper investigated two algorithms for scheduling, asynchronous and synchronous solutions. If the accurate clock can be provided, the synchronous solution by using CQ could improve the latency and jitter significantly. But it must be noted that costs of synchronous solution are not trivial, following tasks are mandatory:

- The crucial requirement of using CQ is the clock sync in the IVN, this is a different topic, and the paper does not address it. Basically, a central controller or distributed protocol, such as IEEE1588 can be used to sync all device clock with a certain accuracy.
- Cycle value selection. The cycle value and the clock accuracy requirement depend on each other, both will determine the granularity of the served packet size, the link utilization, the maximum latency, and the cost of the scheduler design.
- Time window allocation for different flows with different constraints in bandwidth and latency. The optimized solution needs complicated math and cause an overhead for the solution.

As a conclusion, the New IP based IVN can satisfy very well the requirements for the communications of different applications. It opens the door for future IVN and V2X.

Further research is still needed in the following areas:

- Burst effect analysis: The burst coefficient value (Section VIII) will directly impact the accuracy of queuing latency estimation at each hop and will finally determine the accuracy of E2E latency estimation. More study is needed for the burst analysis. A better and more accurate quantitative estimation to the queuing behavior by burst traffic is expected.
- TCP congestion control: The congestion control for different service is expected to be different. New algorithms are critical for application to effectively utilize the new guaranteed service provided by network.
- Algorithm for network resource planning and allocation for synchronous solution, such as optimized cycle number, fast and efficient time slot allocation, scheduler management, etc.

- Simpler method than preemption is needed to eliminate the extra latency and jitter for higher priority traffic caused by a lower priority packet on hardware that is in transmission. This unfinished packet is the root cause of jitter for high priority traffic. Preemption is hard to realize in hardware. Without preemption, the only way to eliminate such effect is to use CQ, but CQ has to sacrifice the link utilization.

## REFERENCES

- [1] L. Han, L. Dong, R. Li, "A Study of In-Vehicle-Network by New IP", INTERNET 2021, The Thirteenth International Conference on Evolving Internet, ISBN: 978-1-61208-880-8
- [2] "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments". IEEE 802.11p published standard. IEEE. July 15, 2010.
- [3] 3GPP, "Cellular Vehicle-to-Everything (V2X)," [https://www.3gpp.org/ftp/tsg\\_ran/TSG\\_RAN/TSGR\\_73/Info\\_for\\_workplan/revised\\_WID\\_22/RAN1\\_6/RP-161894.zip](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_73/Info_for_workplan/revised_WID_22/RAN1_6/RP-161894.zip)
- [4] S. Jiang, S. Yan, L. Geng, C. Cao, and H. Xu, "New IP, Shaping Future Network: Propose to initiate the discussion of strategy transformation for ITU-T", TSAG C-83
- [5] R. Li, A. Clemm, U. Chunduri, L. Dong, and K. Makhijani, "A New Framework and Protocol for Future Networking Applications," ACM Sigcomm NEAT workshop, 2018, pp 21–26.
- [6] L. Han, Y. Qu, L. Dong and R. Li, "Flow-level QoS assurance via IPv6 in-band signalling," 2018 27th Wireless and Optical Communication Conference (WOCC), 2018, pp. 1-5, doi: 10.1109/WOCC.2018.8372726..
- [7] L. Han, Y. Qu, L. Dong and R. Li, "A Framework for Bandwidth and Latency Guaranteed Service in New IP Network," IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2020, pp. 85-90, doi: 10.1109/INFOCOMWKSHPS50562.2020.9162747.
- [8] "3GPP specification series: 38series". 3GPP. Retrieved 2018-10-31.
- [9] "Minimum requirements related to technical performance for IMT-2020 radio interface(s)", Report ITU-R M.2410-0.
- [10] "System Architecture for the 5G System", 3GPP TS 23.501 version 15.2.0 Release 15
- [11] R. Braden, L. Zhang., S. Berson, S. Herzog, and S. Jamin, "RFC 2205: Resource ReSerVation Protocol (RSVP)-Version 1 Functional Specification", IETF, Sept. 1997.
- [12] "Stream Reservation Protocol (SRP)", IEEE 802.1Qat
- [13] "IEEE 802.1 Time-Sensitive Networking Task Group".
- [14] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "RFC 2475: An Architecture for Differentiated Services," IETF, Dec. 1998.
- [15] R. Braden, D. Clark, and S. Shenker, "RFC 1663: Integrated Services in the Internet Architecture: an Overview," IETF, Jun. 1994.
- [16] "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Management and orchestration; 5G end to end Key Performance Indicators (KPI)", 3GPP TS 28.554
- [17] P. Yang, J. Shao, W. Luo, L. Xu, J. Deogun and Y. Lu, "TCP Congestion Avoidance Algorithm Identification," in

- IEEE/ACM Transactions on Networking, vol. 22, no. 4, pp. 1311-1324, Aug. 2014, doi: 10.1109/TNET.2013.2278271.
- [18] B. Briscoe, K. Schepper, M. Bagnulo, and G. White, "Low Latency, Low Loss, Scalable Throughput (L4S) Internet Service: Architecture", Work in Progress, Internet-Draft, draft-ietf-tsvwg-l4s-arch-08, 15 November 2020, <<http://www.ietf.org/internet-drafts/draft-ietf-tsvwg-l4s-arch-08.txt>>.
- [19] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G. Swallow, "RFC3209: RSVP-TE: Extensions to RSVP for LSP Tunnels", IETF, Sept. 2001
- [20] LIN, "ISO/AWI 17987-8"
- [21] CAN: "Road vehicles - Controller area network (CAN) - Part 1: Data link layer and physical signalling", ISO 11898-1:2003
- [22] FlexRay: ISO 17458-1 to 17458-5
- [23] AUTOSAR: AUTomotive Open System ARchitecture
- [24] H. Lim, L. Völker, and D. Herrscher, "Challenges in a future IP/Ethernet-based in-car network for real-time applications", 48th ACM/EDAC/IEEE Design Automation Conference (DAC), 2011
- [25] R. Steffen, R. Bogenberger, J. Hillebrand, W. Hintermaier, A. Winckler, and M. Rahmani, "Design and Realization of an IP-based In-car Network Architecture", Proceedings of "1st International ICST Symposium on Vehicular Computing Systems", 2008
- [26] "P802.1DG – TSN Profile for Automotive In-Vehicle Ethernet Communications". [1.ieee802.org](http://1.ieee802.org).
- [27] E. Ackerman, "Upgrade to Superhuman Reflexes Without Feeling Like a Robot" <<https://spectrum.ieee.org/enabling-superhuman-reflexes-without-feeling-like-a-robot#toggle-gdpr>>
- [28] S. Tuohy, M. Glavin, C. Hughes, E. Jones, M. Trivedi and L. Kilmartin, "Intra-Vehicle Networks: A Review," in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 2, pp. 534-545, April 2015, doi: 10.1109/TITS.2014.2320605.
- [29] "Supporting internet protocol version 4 (IPv4) extension headers", United States Patent, 10,742,775.
- [30] J. Heinanen and R. Guerin, "RFC 2697: A Single Rate Three Color Marker", IETF, Sept. 1999.
- [31] O. Aboul-Magd and S. Rabie, "RFC 4115: A Differentiated Service Two-Rate, Three-Color Marker with Efficient Handling of in-profile Traffic", IETF, Jul. 2005.
- [32] "OMNeT++ Discrete Event Simulator"