Dual-Plane Switch Architecture for Time-Triggered Ethernet

Meng Dong, Zhiliang Qiu, Weitao Pan*,
Hongbin Zhang, Chenglei Kong
ISN, Xidian University, Xi'an, Shaanxi, China
mdong@stu.xidian.edu.cn,{zlqiu,wtpan}@mail.xidian.edu.cn

Hui Jin, Jianlei Yang
SCSE, BDBC, Beihang University, Beijing, China
jianlei@buaa.edu.cn

ABSTRACT

Time-triggered Ethernet (TTE) technology introduces the concept of time-triggered on the basis of traditional Ethernet, so that it can achieve conflict-free and deterministic service forwarding without sacrificing compatibility. However, storage resources in industrial, aviation, aerospace and other equipment are limited. Therefore, it is important for TTEthernet to develop switching technologies with high storage efficiency and scalability. This paper proposes a dual plane switching (DPS) architecture for TTEthernet, which divides time-triggered services and event-triggered services into two planes for data forwarding. Experimental results show that using the TTE switch of this architecture has the advantages of high clock synchronization accuracy, high throughout, low transmission delay and small jitter of TTE service.

CCS CONCEPTS

• **Hardware** → *Very large scale integration design*; *Network on chip.*

KEYWORDS

Time-triggered Ethernet, dual plane switching architecture, high throughout, low transmission delay

ACM Reference Format:
Meng Dong, Zhiliang Qiu, Weitao Pan, Hongbin Zhang, Chenglei Kong and Hui Jin, Jianlei Yang. 2020. Dual-Plane Switch Architecture for Time- Triggered Ethernet. In *ACM Great Lakes Symposium on VLSI 2020 (GLSVLSI'20), September 7–9, 2020, Virtual Event, China.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3386263.3407597

1 INTRODUCTION

Terminals of distributed real-time systems, such as aerospace electronic equipment and industrial control equipment, are usually located in spatially dispersed sites. Therefore, reliable networks are required to ensure the certainty and low latency of business exchanges between terminals. To meet the real-time and stability requirements of the network in different fields, field buses such as CAN, 1553B, SpaceWire, etc.[8] are used in distributed real-time systems. However, due to the slow transmission rate and poor compatibility of these field buses, they are no longer suitable for modern applications with higher bandwidth requirements.

*Corresponding author is Weitao Pan. Email: wtpan@mail.xidian.edu.cn.
This work is supported by National Natural Science Foundation(61306047) and Equipment Advance Research Aerospace Science and Technology Joint Fund (6141B060911).

ACM GLSVLSI '20, September 7–9, 2020, Virtual Event, China

ACM ISBN 978-1-4503-7944-1/20/09…$15.00
https://doi.org/10.1145/3386263.3407597

In the last ten years, standard Ethernet has been widely used in aerospace, aviation, railway and other fields due to its high bandwidth, relatively low component prices, and strong scalability [5, 12]. However, in the early days of standard Ethernet design, it was not considered to support real-time communication, which allows the connected terminals to exchange data at any time, resulting in competition among the terminals [11]. At present, there are many Ethernet-based protocols that enable standard Ethernet to meet the deterministic network time requirements of distributed real-time systems, collectively referred to as real-time Ethernet, such as: Ethernet POWERLINK, PROFINET, Ethernet/IP, EtherCAT, SERCOSIII, etc [4]. However, this real-time Ethernet cannot coexist in the same network as traditional Ethernet. Therefore, the Institute of Electronics and Electrical Engineering (IEEE) has extended the 802.3 standard and proposed TimeTriggered Ethernet.

IEEE 802.3 standard: describes the architecture of best-effort (BE) services in switched Ethernet networks [7].

IEEE 1588 standard: describes the precise clock synchronization method of network measurement and control systems. The hardware and software is used to synchronize the internal clock of each terminal with the master clock of the master machine, providing applications for synchronization establishment time less than 10us [6].

AS 6802 standard: defines a high-precision and fault-tolerant time synchronization method, which can provide TT messages with low-latency and low-jitter high-precision global synchronization clock, thereby improving the quality of TTEthernet service [1].

Aeronautical Radio, Incorporated (ARINC) 664 Part7 Standard: describes a service architecture that supports rate constrained (RC) services in avionics switched Ethernet [3].

The contributions of this paper are as follows: i) We propose a DPS architecture that supports real-time Ethernet, which divides data forwarding into two independent planes, thereby ensuring the delay and jitter of TT services. ii) In the DPS architecture, clock synchronization is isolated from data forwarding. This paper implements IEEE 1588 and AS6802 synchronization modes, which can be flexibly switched according to different usage scenarios.

The remaining chapters of this article are arranged as follows. The second section introduces the related work, the third section introduces the design and implementation of the dual-plane real-time triggered Ethernet switch architecture, and the fourth section introduces the evaluation structure. Section V summarizes the work of this paper.

2 RELATED WORK

In traditional switched Ethernet, best-effort strategies are used for data transmission. At present, some switching structures have been proposed to increase network throughput, reduce frame latency in the network, and support controllable services.

Figure 1a) shows the switching structure of the output queue. In this structure, the data frame enters the corresponding output port directly through the Crossbar after entering from the input port, and then enters the output port buffer and waits for transmission [9]. It has better latency and throughput than the input queued switching structure. If there are situations where data packets are sent from N different input ports to the same output port, to ensure that no frames are dropped, the processing capability of the output ports should be N times that of the switching capacity, but such a requirement is impractical.

To solve this problem, the output queue (VOQ) switching architecture shown in Figure 1b) is proposed. It sets a corresponding buffer at each input port, and divides the buffer into several logical queues, while each logical queue corresponds to a real output port [10]. The data frame first goes through the steps of table lookup and flow classification, and then enters the corresponding logical queue according to its destination port number.

This switching structure is the first choice of wire-speed switching, but it will reduce the throughput in switching.

Figure 1c) proposes a crossbar switching structure with a buffer [2, 13]. The buffer of the data frame is placed in the cross node of each crossbar. This switching structure guarantees high throughput in switching, but for an $N * N$ switching, the required storage cost is N^2. Therefore, when the number of ports is large, this architecture will cause a large amount of cache waste.

Figure 1d) proposes a shared buffer switching architecture that aggregates data from all ports together and writes it into a cache area shared by all ports. In the buffer area, different logical queues uses storage according to the destination port number of different packets. Because the cache area is shared by all ports, the switching structure has a high cache utilization rate, but it is difficult to meet large-capacity exchanges due to its complex control and high read/write speed requirements for the cache area.

Due to the high performance of many existing scheduling algorithms such as iSLP, wire-speed switching can be achieved. At the same time, in the distributed real-time control system, there are many restrictions on on-chip memory, and the consumption of high storage space is not acceptable. Therefore, the VOQ + input buffer architecture is most widely used in Ethernet switches. To ensure the throughput of the switch and reduce the consumption of storage at the same time, in this paper, the architecture of input and cross-node joint queuing is adopted in the ET plane.

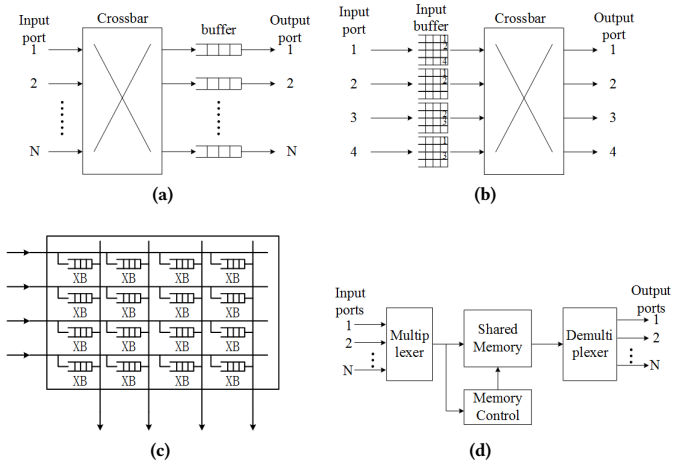


Figure 1: Architecture of a switch with N input and output ports

3 PROPOSED SWITCH ARCHITECTURE

Based on the division of services by the TTE network, the DPS architecture proposed in this paper is shown in Figure 2. It is mainly composed of the TT switching plane, the ET switching plane, and the clock synchronization processing unit. Among them, the TT switching plane and the ET switching plane respectively serves the TT service and the ET service. The input processing, switching network, buffer management, and output processing of the two planes work completely independently. The data received by the switch first enters the filter module which distinguishes the data according to the frame header of the received data frame, and the TT, ET, or synchronization frame will be sent to its respective processing unit. In addition, an output interface module is added at the output port to solve the output conflict between the TT plane and the ET plane.

3.1 TT Switch Plane

The structure of TT switching plane is shown in Figure 3. It is mainly responsible for forwarding TT services entering the switching unit, including input processing, fully interconnected switching architecture and output processing.

1) Input Processing

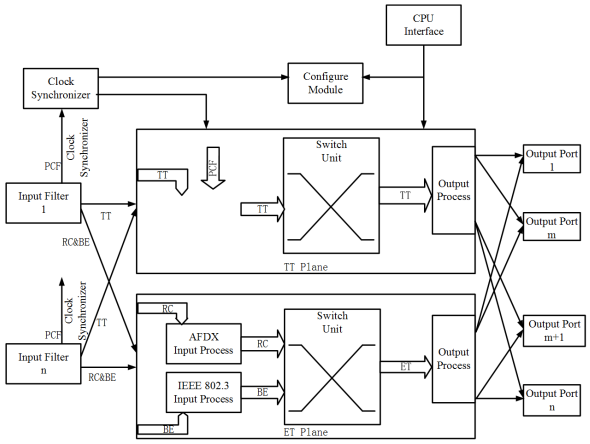


Figure 2: Proposed DPS architecture

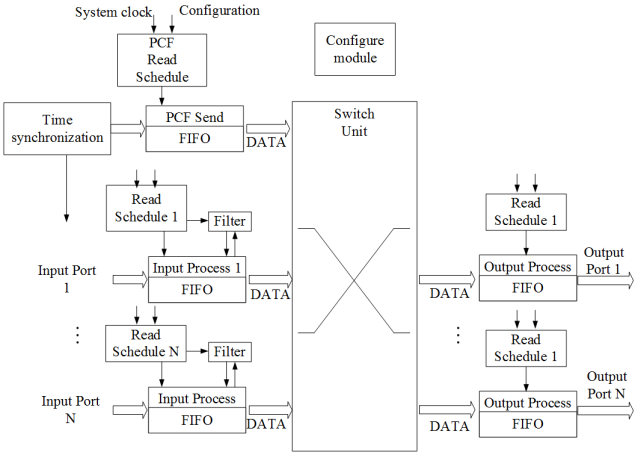


Figure 3: Architecture of TT plane

Follow the time-triggered Ethernet protocol specification, and according to the time schedule plan, first open the data receiving window within the specified time and discard the data frames arriving outside the window. Second, check the service ID number of the received data frame and discard the data frame that does not match the ID number. In the above manner, filtering of TT frames is implemented, thereby ensuring that error frames do not spread within the exchange.

2) Fully Interconnected Switching Architecture

To ensure the low-latency and low-jitter characteristics of the TT service, the switching unit of the TT plane adopts a fully interconnected switching architecture as shown in Figure 4. It establishes an independent data channel between each pair of input and output ports, so as to ensure that there is no blocking problem during the exchange of TT service data between different ports. In addition, the switching unit immediately takes out the data frame after input processing, and forwards the legitimate service through its corresponding data channel. Therefore, there is no queuing problem between the data frames of the same input port. By avoiding data blocking and queuing, the TT plane provides switching services for TT services with low transmission delay and delay jitter.

3) Output Processing

The output processing mainly completes the functions of rate adaptation, secondary filtering, and transmission control for TT data frames. First, there may be a rate mismatch between the input and output ports of the switch. To ensure complete packet reception and continuous transmission, data frames need to be completely stored in the TT plane before transmission.

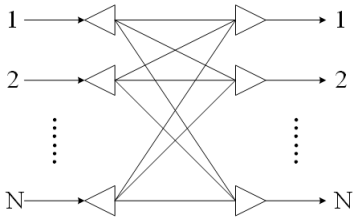


Figure 4: Switching architecture of TT plane

Because the Ethernet frame length is 64-1518 bytes, the delay caused by buffering complete data frames is within an acceptable range. Since the input processing needs to filter the TT service frames, if the TT frames are buffered during the input processing, more delay and waste of buffering will be introduced. Therefore, the received data frame is completely stored once in the output processing section. Second, to enhance the fault tolerance and error isolation capabilities of the switch,the length of the frame and the CRC check value will be calculated when a data frame is registered. And if the frame length is not in the range of 64-1518 bytes, or a CRC error occurs, the frame will be discarded. Finally, at the sending time point specified in the time schedule, the registered complete data frame is sent out of the TT switching plane to complete the data exchange.

3.2 ET Switch Plane

The ET switching plane is mainly responsible for processing and forwarding the RC services specified in the ARINC 664 part7 protocol and the ordinary Ethernet BE services. Similar to the TT plane structure, the ET switching plane is also divided into three parts: input processing, switching unit and output processing.

1) Input Processing

The ET switching plane input processing mainly completes the following functions: reception of RC and BE service data frames, identification and classification of RC service frames and BE service frames, filtering and policing services of RC service frames, and processing of VLAN frames. The structure of input processing is shown in Figure 5.

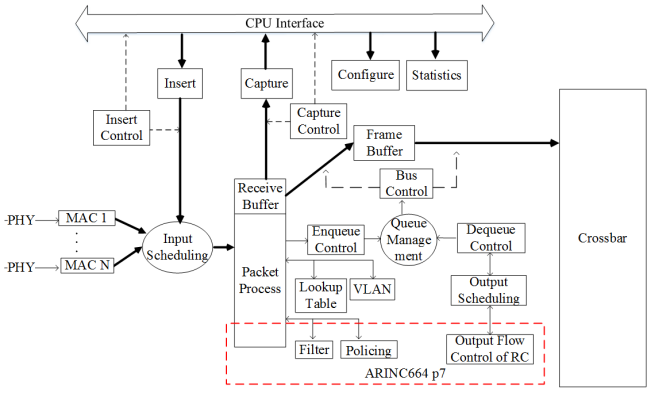


Figure 5: Input process architecture of TT plane

For data frame reception, first, convert the received data frame bit width from 8bit to 32bit to improve the bus rate and throughput of the ET plane. Second, a CRC check is performed on the input data frames, and data frames with CRC check errors are discarded, thereby preventing error diffusion. At the same time, the data frame type is distinguished according to the frame header information of the correctly received data frame. In addition, the input processing also supports the flow control function specified by the IEEE 802.3 protocol. When the receiving buffer space is insufficient, the control instruction is sent to the output processing, and the ET plane sends a "pause" frame to stop the other party from sending data.

For the filtering and policing of RC frames, in order to reduce the time of looking up the table, this design uses two-level table lookup strategy. The first level is an index table corresponding to the virtual link space, and the content in the index table is the address in the second level configuration table. The second level configuration table stores all configuration information of the virtual link. By setting up a two-level lookup table, not only the time complexity of the lookup table is reduced to $O(1)$. At the same time, excessive invalid storage is avoided, and the demand for storage is reduced.

2) Switch Unit

To further improve the utilization of the cache and avoid the head-of-line blocking (HOL) problem, switching unit of the ET switching plane adopts the input and cross-node joint queuing structure shown in Figure 6. Corresponding buffers are set at the data input ports and the cross nodes, and at the input port, we adopt VOQ buffer technology.

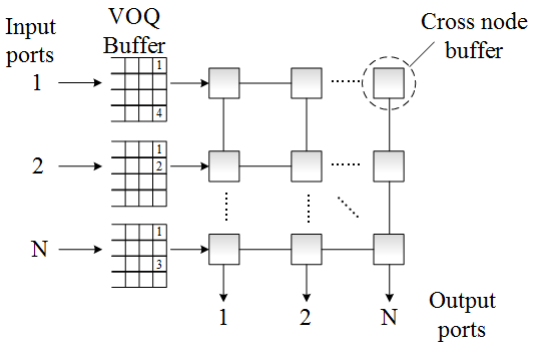


Figure 6: switching architecture of ET plane

First, set the queue according to the priority, output port number and frame type of the service frames to be exchanged. The storage space is divided into a shared area and a guarantee area for each queue. The guarantee area is only used by the queue itself, and the shared area can be borrowed by any queue. Second, the data frame is stored in a fixed-length buffer. This method divides storage space into several fixed-capacity buffers. According to the IEEE 802.3 standard, the frame length is between 64Byte and 1518Byte, so the capacity of the buffer in this design is 64Byte.

3) Output Processing

The output processing of the ET switching plane mainly completes the work of bus arbitration. The arbitration standard is that the TT switching plane has absolute priority. The working mechanism of the arbiter is further explained below from the situation that there are data frames to be sent on the ET plane and the TT plane.

First, when a data frame is ready to be sent on the TT plane, the arbiter reserves the data channel for the TT switching plane by blocking the output of the ET switching plane. According to the IEEE 802.3 standard, the shortest frame interval is required between two Ethernet frames. Therefore, ET services should be blocked 96ns earlier than the sending time of TT frame arrives.

Second, assume that the time required for the ET service to be transmitted is $Tlast$, the current system time is $Tcurrent$, the start time of the TT frame closest to the current time is $Tstart$, the end time of the current TT frame is $Tend$, and the minimum frame interval between two frames specified by Ethernet is IFG. When the transmission time of the ET service overlaps with the transmission window of the TT service, as shown in formulas (1) and (2), the transmission of the ET frame will not be allowed.

$$Tstart < Tcurrent < Tend + IFG \quad (1)$$

$$Tcurrent + Tlast + IFG > Tstart \quad (2)$$

3.3 Clock Synchronization Processing Unit

To better support multiple synchronization protocols, the synchronization processing unit is separated from the data forwarding unit in DPS architecture. This paper designs and implements two synchronization methods based on IEEE 1588 and AS 6802 protocols. The registers can be configured through the upper-layer software or the on-board CPU to select the synchronization method used for the switch, thereby satisfying more application scenarios. To obtain higher synchronization accuracy, the following two measures are taken.

1. According to the clock synchronization period, synchronous data frames such as PTP frames in IEEE 1588 protocol and PCF frames in AS 6802 protocol are treated as TT services in the system. Plan the time synchronization data frame together with the TT service to ensure the priority of the synchronization frame. At the same time, the processing delay of the synchronization frame in the exchange is reduced, and the synchronization accuracy of the system is improved.

2. In the design of the synchronization protocol, the pure FPGA implementation method to avoid using the CPU or a dedicated PHY chip is adopted. This synchronization processing method can effectively reduce synchronization delay and improve synchronization accuracy.

In addition, for systems that use the AS6802 protocol for synchronization, the message permanence time point generated by the message permanence function is usually passed to the compression function for screening. However, the screening delay increases linearly with the increase of the number of terminals in the system. To reduce the delay of the compression function in screening the message permanence time, this paper adds a pre-integration module between the message permanence function and the compression function. First, the module sorts the message permanence information according to the port number, divides the three ports into a group and pre-integrates each port in the group. Second, continue to group and integrate the message permanence information after finishing pre-integration until only one set of message permanence information is left. Through the process of pre-integration, the delay of the message permanence time point that screened by the compression function is reduced from n to $\lceil \log_3 n \rceil$ when there are n ports in the system that need to be cured.

4 EXPERIMENT AND RESULT

4.1 Test System

We implemented the proposed DPS-based real-time Ethernet switch in the Xilinx FPGA Virtex-7 690T based on Figure 7. The switch is an 8-port 1000 / 100M adaptive switch, which is implemented by the vivado 2017.2 integrated design environment. Figure 8 shows the resource occupation of the TT plane when the DPS architecture is used to implement different number of ports Ethernet switches.

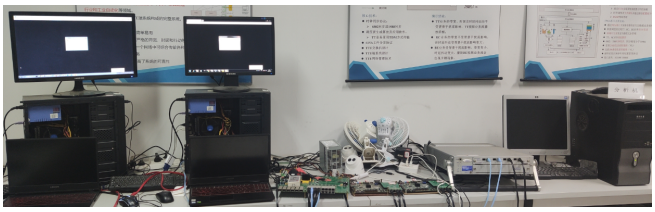
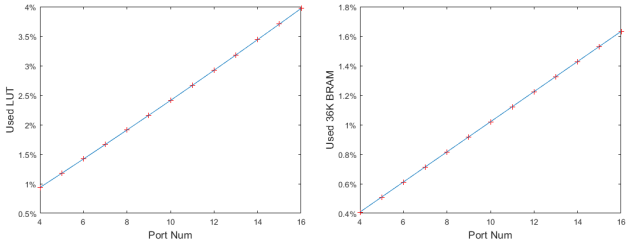


Figure 7: switching architecture of ET plane



(a) (b)

Figure 8: TT plane resource occupancy

Set up a test system as shown in Figure 9, where the ports 1 and 2 of the switch are connected to the TTE end systems #1 and #2. The #1 end system encapsulates the video data collected by the camera into a TT service and sends it to the #2 end system. In addition, the BE data stream is generated internally in the #1 end system and sent to the #2 end system. At the same time, the TT service rate, BE service rate, synchronization accuracy in the #2 end system are remotely measured, and the results are displayed in the network monitoring system.

Table 1: Test Results of Mixed Transmission of TT and BE Services

	Test Results
TT Service	$3340Kbps$
BE Service	$727Mbps$
Synchronization Accuracy	$20ns$

The port 3 and port 4 of the switch are connected to the TTE end systems #3 and #4 respectively. The #3 and #4 end systems complete the sending and receiving of TT and RC services through the upper-layer application software. At the same time, end system #4 dials the 4k video resources in end system #3, and the video resources are transmitted through the BE service.

The port 5 and port 6 of the switch are connected to the Spirent Testcenter network tester. The network tester generates RC service data from 5 to 6 ports.

The port 7 and port 8 of the switch are connected to the Spirent Testcenter network tester, and the network tester injects background interference flow through the switch port.

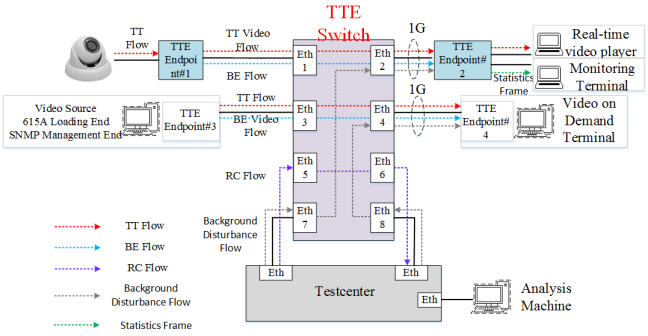


Figure 9: switching architecture of ET plane

4.2 Test Scheme and Result

In the test system shown in Figure 9, the main test contents are performance tests such as mixed transmission of TT, RC and BE services, bandwidth of TT services, delay and delay jitter, and bandwidth guarantee tests of RC services. Among them, the network monitoring terminal corresponding to end system #1 monitors the performance index of TT service. The end-system #4 on-demand video stream is the BE service, and the end-system #1 plays the video stream in real-time as the TT service.

1) Mixed transmission of TT and BE services

Test scheme: The #1 end system continuously sends fixed-rate TT video services and BE services to the #2 end system, and the sum of the rates of the two services is not higher than the upper limit of the rate of the #2 end system by 1 Gbps.

Test results: The TT video service stream captured by the camera is played smoothly. The network test terminal monitors the TT service, BE service, and synchronization accuracy information received by the #2 end system as shown in Table 1.

2) Mixed transmission of TT, RC and BE services

Test scheme: The #3 end system continues to send BE video service flow, TT service flow and RC service flow to #4 end system. The sum of the rates of the three service flows does not exceed the upper limit of the rate of the #4 end system by 1Gbps.

Test results: The on-demand BE video service stream in the #4 end system plays smoothly, and the real-time monitoring results of the three service transmission and reception rates are shown in Table 2.

Table 2: Test Results of Mixed Transmission of TT, RC and BE services

	Send Speed	Receive Speed	Delay	Jitter
TT Service	4.5$Mbps$	4.5$Mbps$	18us	0-8ns
RC Service	4.5$Mbps$	4.5$Mbps$	15us	0-22000ns
BE Service	4.0-6.8$Mbps$	3.0-7.5$Mbps$	—	—

Table 3: TT Service Assurance Test

	Test Results
TT Service	3340$Kbps$
BE Service	577$Mbps$
Synchronization Accuracy	20ns

Table 4: Bandwidth Guarantee Test of RC Service

	Send Speed	Receive Speed	Delay	Jitter
TT Service	4.5$Mbps$	4.5$Mbps$	18us	0-15ns
RC Service	4.5$Mbps$	4.5$Mbps$	15-60us	5000-32000ns
BE Service	4.0-6.8$Mbps$	2.5-5.3$Mbps$	—	—

3) TT service assurance test

Test scheme: On the basis of the scheme described in 1), the network tester sends a 900M background interference stream to the port 2 through the port 7 of the switch.

Test results: The TT video service stream captured by the camera is played smoothly. The network test terminal monitors the TT service, BE service, and synchronization accuracy information received by the #2 end system as shown in Table 3. The comparison between Table 1 and Table 3 shows that in these two cases, the reception rate, transmission delay, delay jitter, and synchronization accuracy of the TT service flow remain unchanged, while the BE service rate decreases significantly. It is proved that the DPS architecture designed in this paper can well ensure the certainty of TT services.

4) Bandwidth guarantee test of RC service

Test scheme: Based on the test scheme described in 2), the network tester sends a 900M background interference stream form the port 8 of switch to the port 4.

Test results: As the service rate sent to the #4 end system in the switch is higher than its reception rate, the on-demand BE video service in end system #4 shows a significant freeze. It can be seen from the comparison between Table 2 and Table 4 that the reception rate of the RC service stream is unchanged, but the transmission delay is increased, and the reception rate of the BE video service is significantly reduced.

5 CONCLUSION AND FUTURE WORK

In this paper, we propose a DPS architecture for real-time Ethernet. Based on the traditional Ethernet, this architecture introduces the TT plane and time synchronization processing, and adds support for time-triggered services. At the same time, in the ET plane, a joint queuing structure of input and cross nodes is used to make full use of the shared buffer to reduce the storage requirements of the switch.

REFERENCES

[1] SAE AS6802. 2011. Time-triggered ethernet. *SAE International* (2011).

[2] S-T Chuang, Sundar Iyer, and Nick McKeown. 2005. Practical algorithms for performance guarantees in buffered crossbars. In *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, Vol. 2. IEEE, 981–991.

[3] Airlines Electronic Engineering Committee et al. 2002. Aircraft Data Network Part 7, Avionics Full Duplex Switched Ethernet (AFDX) Network, ARINC Specification 664. *Aeronautical Radio* (2002).

[4] Jean-Dominique Decotignie. 2009. The many faces of industrial ethernet [past and present]. *IEEE Industrial Electronics Magazine* 3, 1 (2009), 8–19.

[5] Sinisa Derasevic, Manuel Barranco, and Julián Proenza. 2016. Designing fault-diagnosis and reintegration to prevent node redundancy attrition in highly reliable control systems based on FTT-Ethernet. In *2016 IEEE World Conference on Factory Communication Systems (WFCS)*. IEEE, 1–4.

[6] John C Eidson, M Fisher, and J White. 2002. IEEE standard for a precision clock synchronization protocol for networked measurement and control systems. *IEEE Std* 1588 (2002).

[7] IEEE802.3-2005. 2005. IEEE standard for ethernet. https://standards.ieee.org/about/get/802/802.3.html.

[8] ISO ISO. 2003. 11898-2, Road vehicles Controller area network (CAN) Part 2: High-speed medium access unit. *International Organization for Standardization* (2003).

[9] Srinivasan Keshav and S Kesahv. 1997. *An engineering approach to computer networking: ATM networks, the Internet, and the telephone network*. Vol. 116. Addison-Wesley Reading.

[10] Nick McKeown. 1999. The iSLIP scheduling algorithm for input-queued switches. *IEEE/ACM transactions on networking* 7, 2 (1999), 188–201.

[11] Ahmed Nasrallah, Akhilesh S Thyagaturu, Ziyad Alharbi, Cuixiang Wang, Xing Shao, Martin Reisslein, and Hesham ElBakoury. 2018. Ultra-low latency (ULL) networks: The IEEE TSN and IETF DetNet standards and related 5G ULL research. *IEEE Communications Surveys & Tutorials* 21, 1 (2018), 88–145.

[12] Paulo Pedreiras, Paolo Gai, Luís Almeida, and Giorgio C Buttazzo. 2005. FTT-Ethernet: A flexible real-time communication protocol that supports dynamic QoS management on Ethernet-based systems. *IEEE Transactions on industrial informatics* 1, 3 (2005), 162–172.

[13] Jonathan Turner. 2008. Strong performance guarantees for asynchronous buffered crossbar schedulers. *IEEE/ACM Transactions on Networking* 17, 4 (2008), 1017–1028.