Nuclear Engineering and Technology 54 (2022) 3587–3594

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original Article

Development of a distributed high-speed data acquisition and monitoring system based on a special data packet format for HUST RF negative ion source

Dong Li, Ling Yin*, Sai Wang, Chen Zuo, Dezhi Chen

State Key Laboratory of Advanced Electromagnetic Engineering and Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan, 430074, Hubei, China

ARTICLE INFO

Article history:
Received 20 January 2022
Received in revised form
29 April 2022
Accepted 6 May 2022
Available online 11 May 2022

Keywords:
Distributed data acquisition
ZeroMQ
Data storage
RF negative ion source

ABSTRACT

A distributed high-speed data acquisition and monitoring system for the RF negative ion source at Huazhong University of Science and Technology (HUST) is developed, which consists of data acquisition, data forwarding and data processing. Firstly, the data acquisition modules sample physical signals at high speed and upload the sampling data with corresponding absolute-time labels over UDP, which builds the time correlation among different signals. And a special data packet format is proposed for the data upload, which is convenient for packing or parsing a fixed-length packet, especially when the span of the time labels in a packet crosses an absolute second. The data forwarding modules then receive the UDP messages and distribute their data packets to the real-time display module and the data storage modules by PUB/SUB-pattern message queue of ZeroMQ. As for the data storage, a scheme combining the file server and MySQL database is adopted to increase the storage rate and facilitate the data query. The test results show that the loss rate of the data packets is within the range of 0–5% and the storage rate is higher than 20 Mbps, both acceptable for the HUST RF negative ion source.

1. Introduction

Neutral beam injection (NBI) is one of the most important mechanisms for plasma heating and current driving of the controlled fusion experimental reactor, and the ion source is a key module of the NBI system to provide charged particle beam [1]. Supported by the project “CFETR N-NBI beam optics optimization of research and experimental evaluation” of the Ministry of Science and Technology in China, Huazhong University of Science and Technology (HUST) has built a small RF negative ion source platform (HUST-RNIS) [2,3]. HUST-RNIS is designed for pulse operation and adopts the three-grid extraction system – the plasma grid (PG), extraction grid (EG) and grounded grid (GG), with a maximum extraction voltage of 10 kV and the acceleration voltage up to 20 kV.

During the operation of HUST-RNIS, monitoring the potentials of PG and EG could obtain their rates of rise and synchronization performance, and monitoring the currents of PG, EG, GG as well as the calorimeter could analyze the extracted beams and their transient characteristics. These physical signals change fast in the transient process, for example, it only takes about 6 ms for PG or EG potential to rise. In addition, the RF power and gas pressure also change fast during the plasma excitation while it is necessary to sample the outputs of the directional coupler and gas gauge to analyze the relationship between plasma parameters and the RF power or gas pressure. Therefore, a high-speed data acquisition and monitoring system needs to be developed specifically.

Many large scientific devices realize centralized high-speed data acquisition based on the NI acquisition board, which is easy to build the time correlation among different acquisition channels, and use LabVIEW software to build the data processing platform [4,5]. In Ref. [6], a three-channel data acquisition and transmission system is designed based on the FPGA, DDR3 and PCIe bus, with a sampling rate of 5 Msps per channel, and the Qt framework is adopted to develop the interface application for the waveform display and data storage. In Ref. [7], a high-speed data acquisition card based on the Zynq processor and PCIe bus is designed, and the data transmission rate is more than 20 Gbps. Ref. [8,9] adopt the Gigabit Ethernet for long-distance data transmission.

The high-speed data acquisition and monitoring system for HUST-RNIS should not only sample the fast-changing signals, but

* Corresponding author.
E-mail address: m202071707@hust.edu.cn (L. Yin).

https://doi.org/10.1016/j.net.2022.05.008

also send the sampling data to the upper computer for real-time display and storage, so that users can observe the change of signals in real time and analyze the historical data later. Because the object signals are of different reference potentials, it is necessary to set up multiple data-acquisition modules for distributed acquisition. Besides, the subsystems work in accordance with a set time sequence, so it is necessary to build the time correlation of different signals for the subsequent analysis. In addition, to decentralize the load of data processing, the upper computer should also adopt the distributed structure. Finally, the speed of data storage may not match with the speed of data upload, hence the data buffer needs to be considered.

Thus, a distributed high-speed data acquisition and monitoring system is used to sample the multiple physical signals, and to upload and store the sampling information for HUST-RNIS. Its data acquisition modules provide an absolute-time label of the microsecond resolution for each sampling data to build the time correlation of different signals, and upload the data and time information together in the UDP message. A special data packet format is proposed for the data upload, which is convenient for packing or parsing a fixed-length packet, especially when the span of sampling time in a message crosses an absolute second. The upper-layer programs of data processing are developed based on the ZeroMQ messaging library, with which the data packets extracted from UDP messages are distributed to the real-time display module and data storage modules [10]. A storage scheme combining the file server and MySQL database is adopted to increase the storage rate and facilitate the data query.

2. System design and implementation

2.1. Overall structure

The overall structure of the high-speed data acquisition and monitoring system is shown in Fig. 1 [11].

The data acquisition module is responsible for high-speed acquisition of physical signals, the requirements of which are shown in Table 1. Based on its internal clocks, an absolute-time label of the microsecond resolution is provided for each sampling data, which will be packed with the sampling data in a prescribed packet format and uploaded to the corresponding data forwarding module through the Gigabit Ethernet over UDP. In order to synchronize the time information of multiple acquisition modules, a time synchronization system is used to correct the time of acquisition modules once per second.

All of the data forwarding modules, data storage modules and the real-time display module are running on virtual machines (VM). After receiving a UDP message, the data forwarding module directly forwards its data packet through ZeroMQ message queue. Then the corresponding real-time display module and data storage module will receive and parse the forwarded data packet, and the former displays the data in real time, while the latter stores the data in the storage server for subsequent review and analysis. The core idea of the software is to use the input/output pipe of ZeroMQ as the data buffer to resolve the mismatch between the speed of data upload and storage during the pulse operation of HUST-RNIS. Besides, the multi-connection feature of ZeroMQ is used to distribute the data packets to the real-time display module and corresponding data storage module in parallel. In addition, the management module is set up to monitor the running status of the data forwarding modules and the data storage modules, and to obtain the configuration information of HUST-RNIS set by the EPICS master control system, so as to analyze the running data in combination with the experimental configuration later.

2.2. Data acquisition module

The functional structure of a data acquisition module is shown in Fig. 2, which mainly includes the FPGA controller, photoelectric conversion module, A/D module, dual RAM cache module and Ethernet transmission module. The FPGA controller drives the A/D module to sample dual-channel signals at high speed. Then each sampling data is marked with a microsecond-resolution time label and stored in RAM temporarily. Finally, the data in RAM will be read out and uploaded by the Ethernet transmission module. Compared with the TCP protocol, in the UDP protocol, there is no mechanism of the confirmation packet which will reduce the utilization efficiency of the network. Although a small amount of UDP messages would be lost, the interpolation method can be used to estimate those lost data. Therefore, UDP protocol is adopted to upload the data.

The time synchronization system is a distribution center of the absolute time, which generates the serial IRIG-B code stream based on an FPGA, a clock module T302-3 and constant-temperature crystal oscillator.

As HUST-RNIS is designed for the pulse operation, the timing control system is adopted to enable every subsystem's operation with a logic link, including the data acquisition modules. The data acquisition modules work only when the corresponding subsystems operate, and this avoids the waste of the data buffer and storage space in the off-duty period. Therefore, every data acquisition module needs to provide a fiber optic interface with HFBR-2512 to receive the enable signal from the timing control system.

Because the speed of data acquisition is much higher than the speed of data processing, the scheme of uploading a UDP message per sampling data may not only keep the upper computers from reading UDP messages in time, but also result in a waste of network resources due to the low proportion of effective data in a UDP message. Therefore, a certain amount of sampling data and absolute-time labels will be temporarily stored in RAM and uploaded together later, based on the prescribed packet format shown in Fig. 3. The channel ID serves as the identity of the physical signal. The data number represents the total number of sampling data in the packet, which is used as the reference length for the data parsing in the upper computer. The break number indicates how many data (starting from the header) belong to the current second, the rest then being the next-second data. The current second, which represents how many seconds have passed from the beginning of the day, is the time when the first data of the packet is sampled. The data is encoded by IEEE 745 standard. The microsecond indicates the microsecond-resolution time of the corresponding data within an absolute second.

According to the UDP protocol, an Ethernet frame consists of the Ethernet header ($L_1 = 14$ bytes), IP header ($L_2 = 20$ bytes), UDP header ($L_3 = 8$ bytes), application data (referring to the prescribed data packet, L_d is the length of the packet), and the Ethernet footer ($L_4 = 4$ bytes). L_d is calculated as (1):

$$L_{\mathrm{d}} = \begin{cases} 18 & , N_{\mathrm{d}} = 1 \\ 8 + 8N_{\mathrm{d}} & , N_{\mathrm{d}} \geq 2 \end{cases} \tag{1}$$

where N_d represents the data number of the prescribed packet. When $N_d = 1$, the packet will be automatically supplemented to 18 bytes with zero according to the UDP protocol. The ratio R_e of

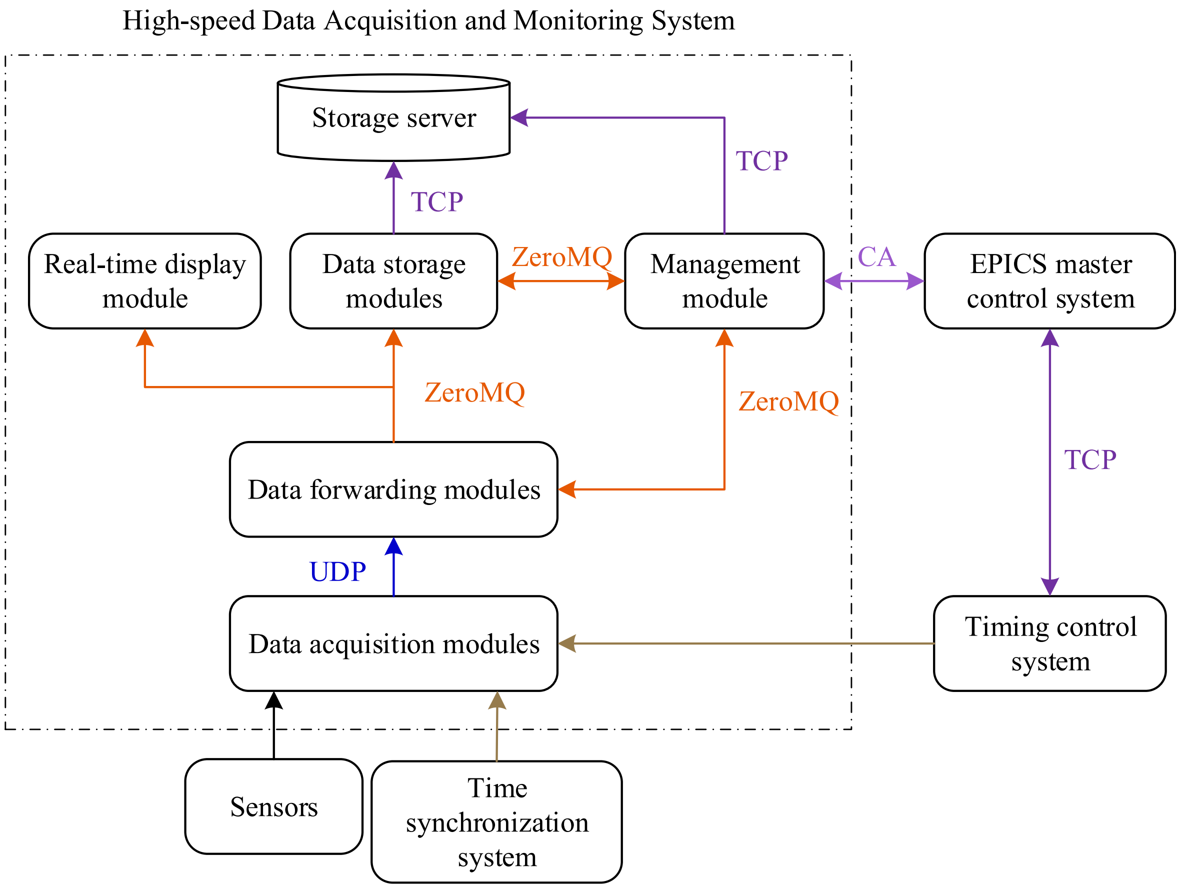

Fig. 1. Overall structure of the high-speed data acquisition and monitoring system.

Table 1
Requirements of the high-speed data acquisition.

System Name	Signal Name	Desired Sampling Rate	Data Type
High-voltage DC power supply	Potential of PG	1 Msps	Float
	Potential of EG	1 Msps	Float
	Current of PG	1 Msps	Float
	Current of EG	1 Msps	Float
	Current of GG	1 Msps	Float
RF power system	Incident power	1 ksps	Float
	Reflected power	1 ksps	
Gas pressure gauge	Gas pressure	1 ksps	Float

effective bytes in a UDP message is then calculated as (2):

$$R_e = \frac{L_d}{L_1 + L_2 + L_3 + L_d + L_4} \times 100\% \tag{2}$$

And the occupied bandwidth BW is calculated as (3):

$$BW = 8(L_1 + L_2 + L_3 + L_d + L_4) \cdot \frac{f_s}{N_d} \cdot N_c \tag{3}$$

where the unit of BW is bps, f_s represents the sampling rate of an A/D channel, N_c represents the number of channels in the A/D module.

The effects of N_d on R_e and BW are shown in Fig. 4. Given that L_d is within 1472 bytes to prevent the fragmentation at the IP layer, the maximum of N_d is 183. As N_d increases, the efficiency of data transmission is improved and tends to saturate after $N_d > 40$. However, larger N_d is conducive to the data forwarding, $N_d = 180$ is thus chosen in our packet. In this case, R_e is calculated to be 97% and BW is 127 Mbps. Noting that $R_e = 28\%$ and $BW = 977$ Mbps when each packet only contains one data, the efficiency of data transmission is therefore greatly improved with our packet format.

2.3. Data forwarding and processing modules

The functional structure of data forwarding and processing modules is shown in Fig. 5 [11]. Each data acquisition module has a corresponding data forwarding module to receive its UDP messages and forward the data packets in the messages one by one through ZeroMQ, and a data storage module to parse the data packets into sampling data and time labels, which will be stored in the storage server. In addition, the real-time display module can select the signals to be monitored and shown.

ZeroMQ is a message-based network communication library with the characteristics of high reliability, high efficiency and

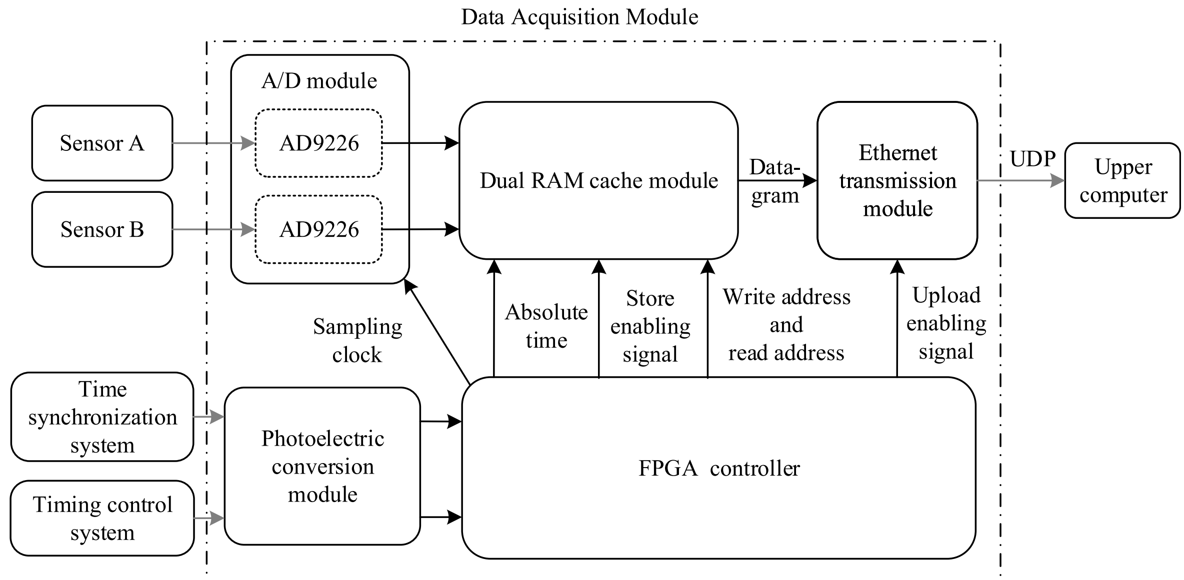

Fig. 2. Functional structure of the data acquisition module.

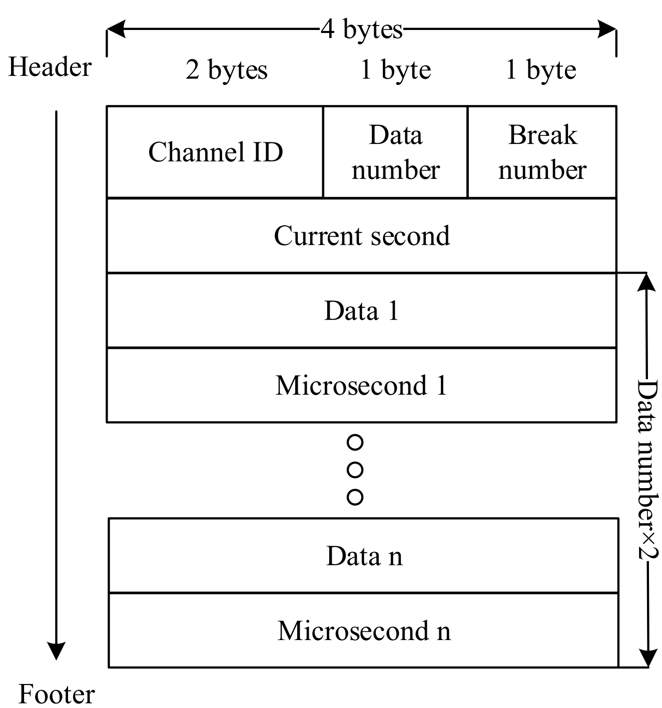

Fig. 3. Format of the prescribed data packet.

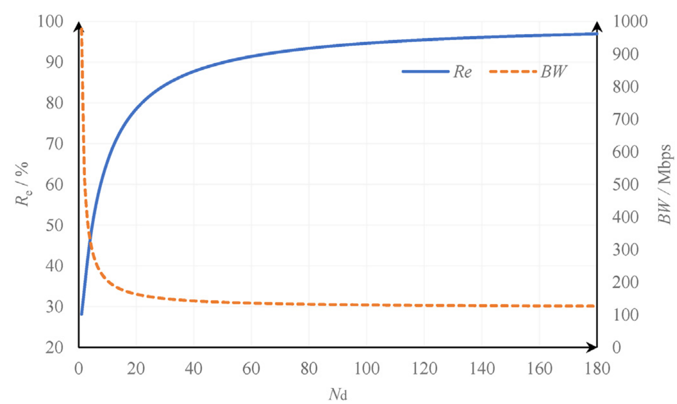

Fig. 4. Efficiency of data transmission with different L_d when f_s = 1 Msps.

asynchronous communication. ZeroMQ node establishes the input/output pipe as a buffer to realize asynchronous message transfer between the I/O thread and the main thread, which can resolve the mismatch between the speed of data upload and speed of data storage during the pulse operation of HUST-RNIS.

The forwarding thread of the data forwarding module receives the UDP message and directly forwards its data packet through ZeroMQ of PUB pattern, without time-consuming operations such as the data parsing, which ensures the data forwarding as soon as possible and reduces the packet loss.

Both the real-time display module and the data storage module establish the SUB-pattern ZeroMQ to subscribe to the corresponding data. They firstly parse packets according to the prescribed packet format into one-to-one data and microsecond-resolution absolute-time labels, and then the former extracts a certain proportion of the data and time labels with equal interval for dynamic curve display, and the latter stores all the data and time labels into the storage server.

In order to monitor the status of the data forwarding threads and the data storage threads, the daemon thread is set up for each data forwarding and storage thread to check whether they are alive by the attribute 'is alive'. At the same time, a management module is created to obtain the status of the data forwarding threads and the data storage threads from corresponding daemon threads. In addition, the management module also obtains the operation configuration set by EPICS master control system for HUST-RNIS through the CA bus, so as to analyze the operation data in combination with the configuration later.

The real-time display is implemented based on the graphical user interface library PyQt5 and the high-performance data graphical display library PyQtGraph of Python [12]. The workflow of the real-time display is shown in Fig. 6. Firstly, according to the data transmission rate of each signal, a corresponding extraction ratio of the data in a packet is set to match the capability of real-time

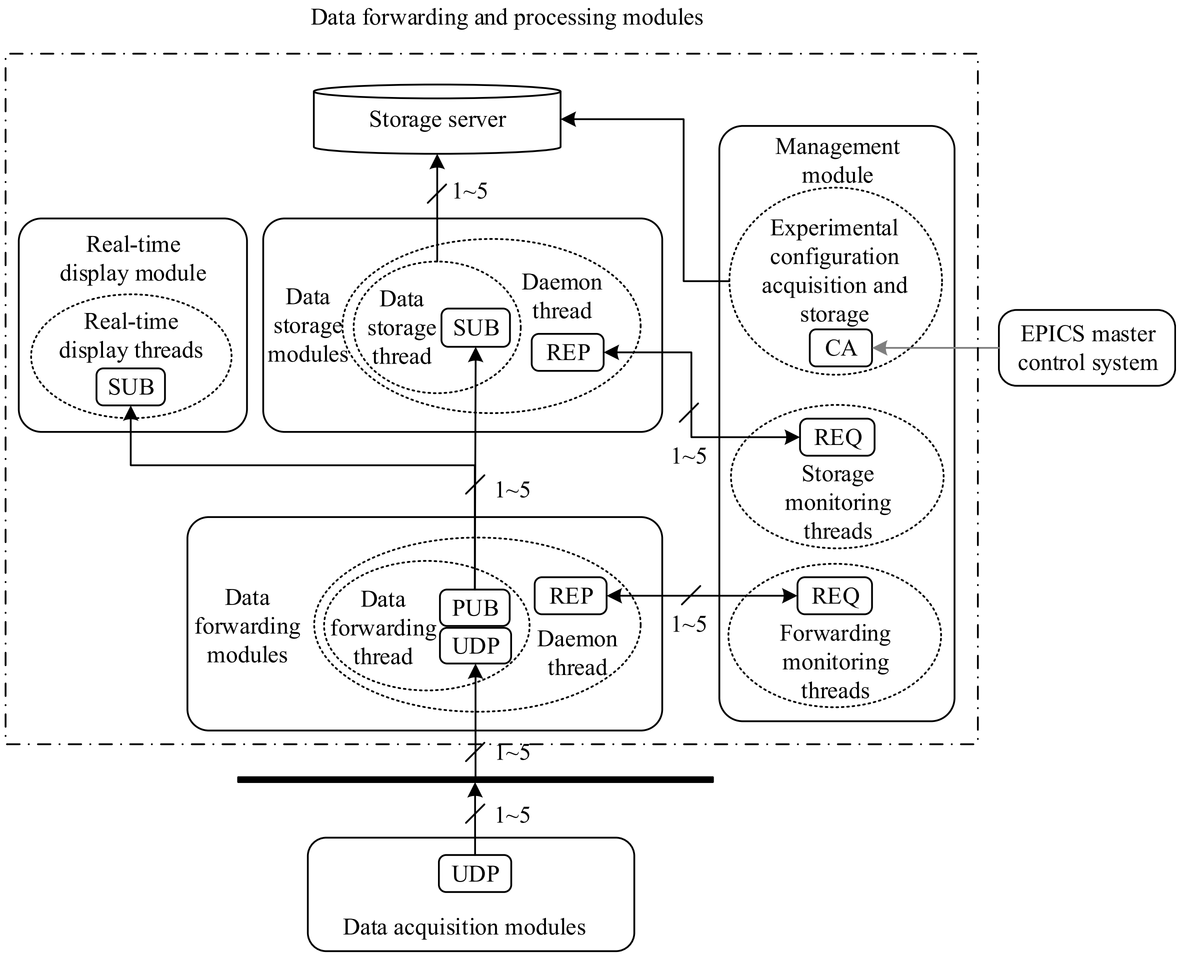

Fig. 5. Functional structure of the data forwarding and processing modules.

display. Once the 'Start receiving' button is pressed, the threads of data packet receiving and parsing start, and the set ratio of the data and time labels in each packet will be appended to the array and flushed to the curve. The function of exporting graphs is provided, as well as stopping the reception of data packets and clearing the display interface.

As for the data storage and query, Fig. 7 shows one simple scheme. Firstly, the packets are parsed and the data and time labels will be stored in the MySQL database. Then through the Web platform, the desired experimental data or curves can be retrieved from the MySQL database according to the experiment date, the sequence number and the signal name. However, the writing speed of MySQL is very slow compared with the speed of the data upload. The running period of HUST-RNIS is about tens of seconds each time, and it is impossible for MySQL to complete the multi-channel data storage within the off-duty period of several minutes. Therefore, it is necessary to improve the storage scheme with the goal of high-speed storage.

Given that many enterprises set up file servers for central storage and the file management [13,14], a data storage and query scheme for the high-speed data acquisition and monitoring system is proposed based on the file server (shown in Fig. 8). The data storage module will parse the data packets and save the data and time to the file server as CSV files. And the MySQL database only stores the basic file information such as the file ID, file path and description for Web retrieval for its advantage of convenient query. Then the data storage rate can be significantly increased.

3. Experimental verification of the system performance

3.1. Test platform introduction

The structure of the test platform for the distributed high-speed data acquisition and monitoring system is shown in Fig. 9. The data acquisition modules are separate entities based on the FPGA, and all of the upper-layer data forwarding and processing modules run in the different virtual machines on the same VLAN, which communicate with each other through a Gigabit switch.

3.2. Transmission performance

The data transmission has two phases: uploading data from the data acquisition modules to the data forwarding modules based on the UDP protocol, and forwarding data from the data forwarding modules to the data processing modules based on the ZeroMQ message queue.

Firstly, the loss rate of the data packets between one data acquisition module and the corresponding data forwarding module is tested. To test the transmission performance under maximum network load, the sampling rate of each A/D channel is set as 1 Msps according to the requirements in Table 1, and every 180 sampling data are packed in a prescribed packet for uploading. The number of data packets N_{acq} that uploaded by each acquisition module is set to 200,000. The number of data packets received by one data forwarding module is then counted as N_{fwd} and the packet

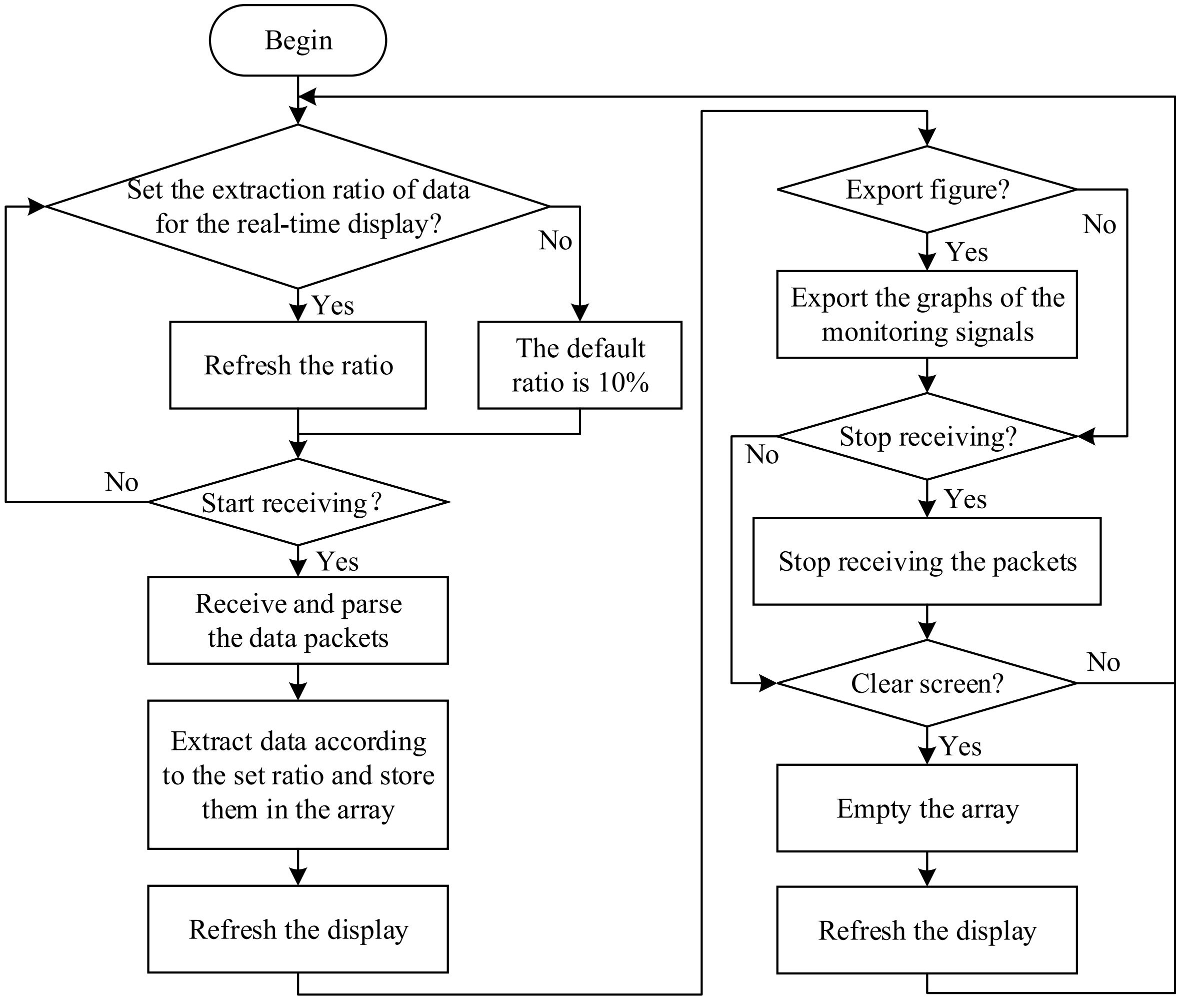

Fig. 6. Workflow of the real-time display.

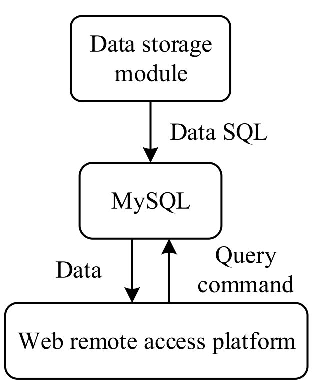

Fig. 7. Data storage and query scheme with MySQL database.

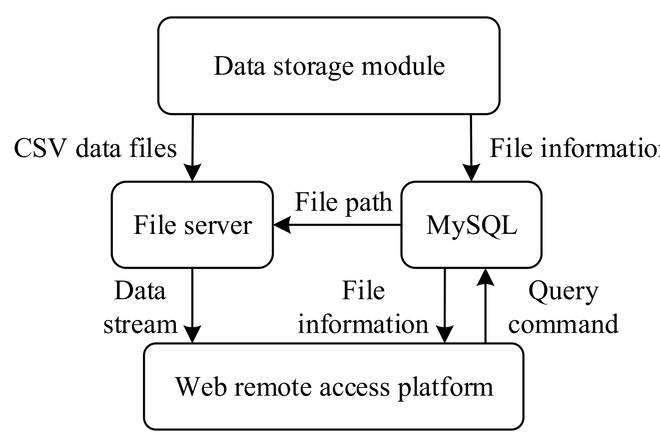

Fig. 8. Data storage and query scheme with the file server and MySQL database.

loss rate r_{loss} is calculated as (4):

$$r_{loss} = \frac{N_{acq} - N_{fwd}}{N_{acq}} \times 100\% \quad (4)$$

As is shown in Fig. 10, the horizontal axis represents the number of tests, and the loss rate is within the range of 0–5%, which is acceptable for HUST-RNIS.

Next, the loss rates of the data packets between the data forwarding modules and the corresponding data processing modules are tested. By uploading data packets with the data acquisition modules for a short duration, and comparing the number of the data packets received by each data forwarding module with that of the corresponding data storage module, it is found that the two numbers are always consistent, indicating there is no packet loss in this phase. The reason may be that ZeroMQ adopts the TCP transmission protocol which has high transmission reliability, and SUB-

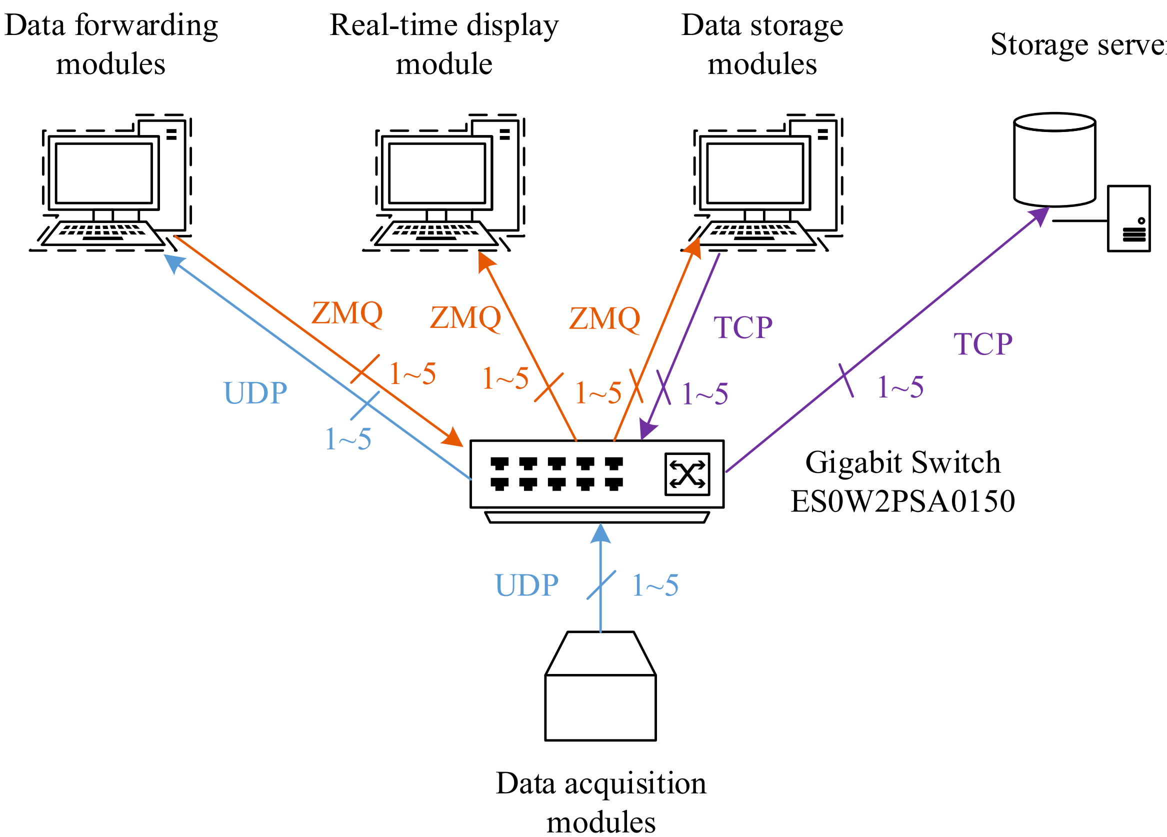

Fig. 9. Test structure of the high-speed data acquisition and monitoring system.

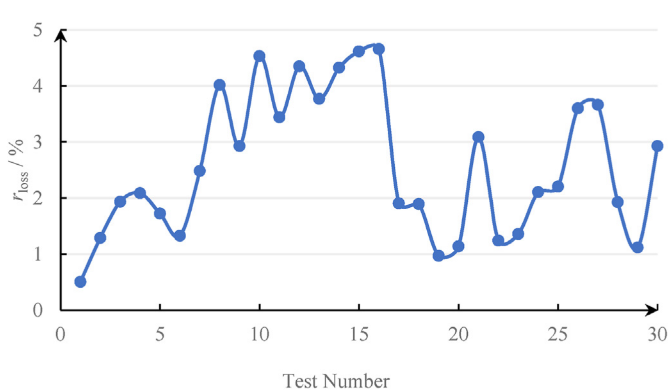

Fig. 10. Test results of average packet loss rate in UDP transmission.

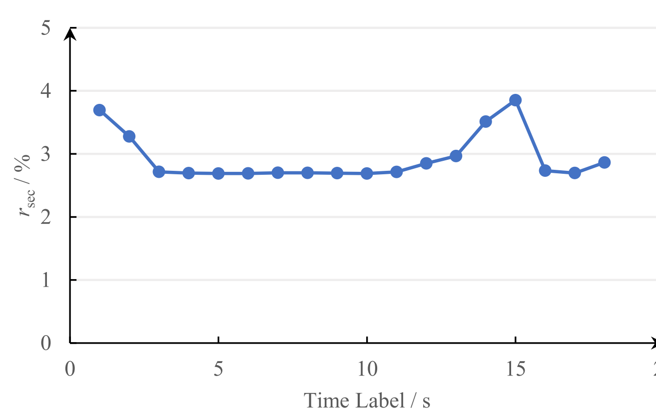

Fig. 11. Test results of packet loss rate per second in the transmission.

pattern ZeroMQ has the receiving buffer which can temporarily store the unprocessed packets.

In addition, the data loss rate per second during the data transmission is tested. One data acquisition module uploads data in the same way as above, so the number N_{acq_sec} of the uploaded data per second is 1 million. After receiving the data packets, the corresponding data storage module parses them and counts the number N_{sto_sec} of the received data belonging to each second (according to the absolute-time labels), and then the loss rate per second r_{sec} of the data is calculated as (5):

$$r_{\ sec} = \frac{N_{acq\_sec} - N_{sto\_sec}}{N_{acq\_sec}} \times 100\% \tag{5}$$

Fig. 11 shows the data loss rate per second in a certain test. As is shown, the loss rate fluctuates within the range of 2–4%, and there is no abnormally large value.

3.3. Storage rate

In order to test the storage rate of the whole system, all the data acquisition modules, data forwarding modules, and data storage modules are used. The sampling rate of every data acquisition module is set to 1 Msps according to the requirements in Table 1 and every 180 sampling data are packed in a prescribed packet for uploading. The two storage schemes in section 2.3 are respectively tested with different spans of the acquisition time, and the corresponding average durations for the data storage are shown in Table 2. In comparison, the storage rate of the latter scheme is 40

Table 2
Test results of data storage rate.

Span of Acquisition Time / s	Data Save Duration / min	
	MySQL	File Server
4.5	107.9	2.6
9.0	210.6	4.5
13.5	299.7	7.5
18.0	443.0	10.8

times higher than the former. The storage rate of the latter scheme is more than 20 Mbps and acceptable for HUST-RNIS.

4. Conclusion

A distributed high-speed data acquisition and monitoring system for HUST-RNIS is developed, which consists of data acquisition, data forwarding and data processing. Based on the FPGA and high-speed A/D, the data acquisition modules sample physical signals at high speed and upload the sampling data with the absolute-time labels, which builds the time correlation among different signals. In order to improve the ratio of effective bytes in UDP message, a special packet format is proposed for the data upload. After receiving the UDP messages, PUB/SUB-pattern message queue of ZeroMQ is adopted by the data forwarding modules to distribute the data packets to the real-time display module and the data storage modules, whose receiving buffer could resolve the mismatch between the speed of data upload and storage during the pulse operation of HUST-RNIS. As for the data storage, compared with storing the data stream directly in MySQL database, storing the data in CSV files can increase the storage rate significantly and is more suitable for HUST-RNIS.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Key Research and Development Program of China [grant numbers 2017YFE0300105]; the National Natural Science Foundation of China [grant numbers 11775088, 11705063]. The authors would also like to thank Fei Song and Ruoyu Wu for their constructive suggestions on my article.

References

[1] B. Heinemann, H. Falter, U. Fantz, P. Franzen, M. Fröschle, R. Gutser, W. Kraus, R. Nocentini, R. Riedl, E. Speth, A. Stäbler, D. Wünderlich, P. Agostinetti, T. Jiang, Design of the “half-size” ITER neutral beam source for the test facility ELISE, Fusion Engineering and Design 84 (2008) 915–922, https://doi.org/10.1016/j.fusengdes.2008.11.076.
[2] P. Zhao, Research on Plasma Production and Detection Techniques for a High-Power RF Negative Ion Source, Huazhong University of Science and Technology, 2018. https://d.wanfangdata.com.cn/thesis/D01546351.
[3] C. Zuo, System Design and Analysis of High Power RF-Driven Negative Ion Source, Huazhong University of Science and Technology, 2018. https://d.wanfangdata.com.cn/thesis/D01546353.
[4] Jun Wu, The Research on the Data Acquisition and Processing System of the Large Scientific Facilities, University of Science and Technology of China, 2017. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CDFD&dbname=CDFDLAST2017&filename=1017047814.nh&uniplatform=NZKPT&v=LXxZtYFQVKDEPhD3CRUlW_XoALs-fSCo0N4HKHYUhLbXLocrQlSTMj7dNTWo9Jk3.
[5] Yang-yang Gao, Chun-dong Hu, Peng Sheng, Yuan-zhe Zhao, Qing-long Cui, Ya-hong Xie, Shan Yu, Yu-qian Chen, Rui Zhang, Signal acquisition system for neutral beam injector RF ion source, Computer Systems & Applications 25 (2016) 63–67. https://d.wanfangdata.com.cn/periodical/jsjxtyy201604011.
[6] Y. Sun, Development of PCIe-based High-Speed Multi-Channel Data Acquisition System, Donghua University, 2021. https://d.wanfangdata.com.cn/thesis/D02373238.
[7] Z.G. Han, Design and Implementation of High Speed PCIE Data Acquisition Card Based on Zynq, Xidian University, 2020. https://d.wanfangdata.com.cn/thesis/Y3790895.
[8] X. Li, Design and Implementation of FPGA High-Speed Data Acquisition and Transmission System Based on Gigabit Ethernet, Beijing University of Posts and Telecommunications, 2019. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201902&filename=1019113042.nh&uniplatform=NZKPT&v=qUfzncG32mAGGQxFyMMhAGch2FEh7puLyudHj22VS5j5u0Rts_HV_2IBnzJGD0y_.
[9] Kai Chen, Qingshan Tang, Yajie Li, Zhen Luo, Huan Li, Weiwu Shan, Design method of gigabit ethernet transmission based on FPGA, Electronic Science and Technology 30 (2017) 102–104+108. https://d.wanfangdata.com.cn/periodical/dzkj201706028.
[10] P. Hintjens, ZeroMQ: Messaging for Many Applications, O'Reilly Media, 2013.
[11] S. Wang, D. Li, Design of high-speed data acquisition and management system for negative ion source, in: 2020 IEEE 1st China International Youth Conference on Electrical Engineering (CIYCEE), Wuhan, China, Nov 2–4, 2020, https://doi.org/10.1109/CIYCEE49808.2020.9332705.
[12] S. Sepúlveda, P. Reyes, A. Weinstein, Visualizing physiological signals in real-time, in: Python in Science Conference, Austin, the United States, July 6–12, 2015, https://doi.org/10.25080/Majora-7b98e3ed-01c.
[13] Baihui Ma, Based on B/S Mode of Enterprise File Sharing Office System Design and Implementation, Jilin University, 2018. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201901&filename=1018288685.nh&uniplatform=NZKPT&v=u0VXB-Usi5msV1Mn2bSsgot35XEdx14zqVrNiAQjXhrniO6GSwW7d0o_JsmtKdjO.
[14] Li Wu, Hybrid Cloud Based File Storage and Sharing System for Enterprise, Shanghai Jiaotong University, 2015. https://d.wanfangdata.com.cn/thesis/D785090.