25th IEEE Real Time Conference
2026/05/25 – 2026/05/29

FPGA based RDMA for BEE Readout

Chang Xu [1,2], Sheng Dong[2], Hongyu Zhang[1,2] , Yunpeng Lu[2], Kejun Zhu[2]

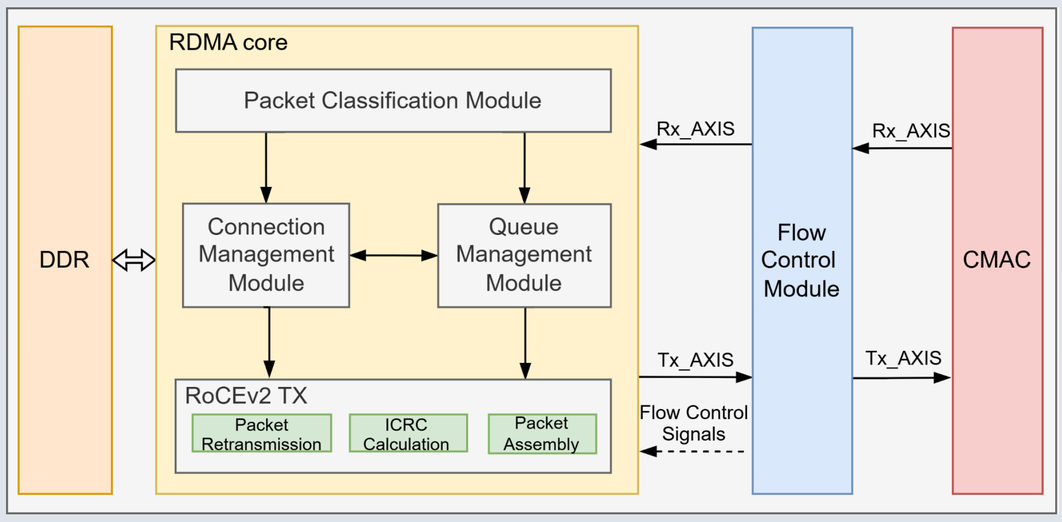
(1) University of Chinese Academy of Sciences, Beijing 100049, China
(2) Institute of High Energy Physics, CAS, Beijing 100049, China

Introduction

Traditional TCP/IP protocols rely heavily on CPU processing for data packet encapsulation, parsing, and transfer, resulting in significant latency and resource consumption. In contrast, RDMA (Remote Direct Memory Access) enables direct data transfer between the network adapter and memory, bypassing the operating system kernel. This reduces CPU overhead while delivering high bandwidth and ultra-low latency, offering an efficient solution for high-performance and data-intensive applications. For high-energy physics experiments confronting exponentially increasing data acquisition rates, RDMA-based Back End Electronics (BEE) readout can alleviate the computational burden on data acquisition (DAQ) systems, optimize resource allocation, and enhance bandwidth utilization to meet the stringent real-time transmission requirements.

Network Stack Design

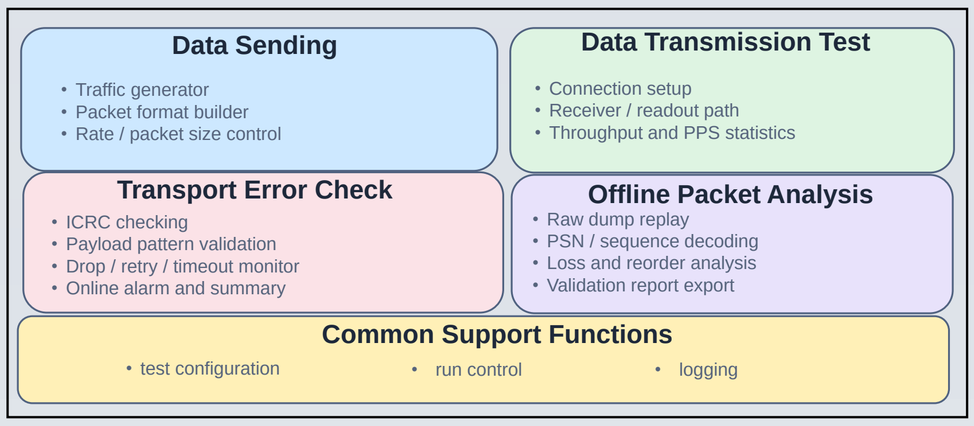
▸ Firmware Design



- **CMAC:** Link layer data transmission
- **Packet Classification Module:** Packet filtering, RoCEv2 protocol verification and ARP response
- **RoCEv2 TX :** Encapsulates data and constructs packets complying with RoCEv2 specifications
- **Connection & Queue Management Module:** RDMA connection establishment and maintenance

▸ Software Design

- Development Goals:
  - Support max throughput test, in-transit & offline packet error check
  - Equipped with sender/receiver for PC-to-PC data transmission test



Research Progress

▸ Software Implementation

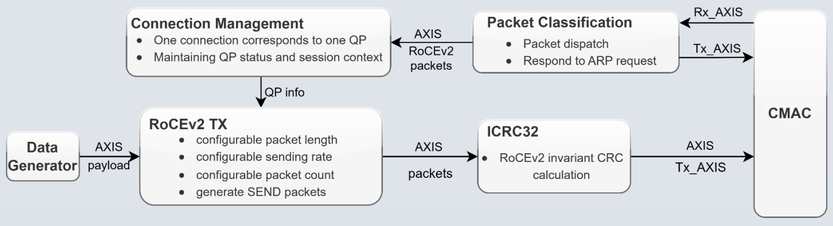
- Developed based on Verbs and RDMA Connection Manager, with complete error detection functions

▸ Customized Firmware Implementation

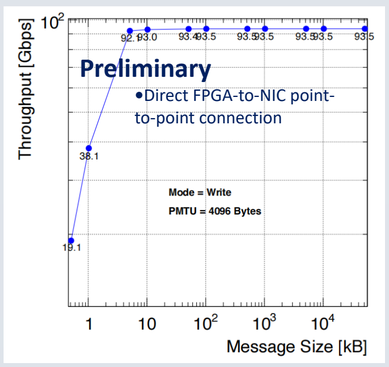
- Implement connection establishment and RDMA send function with no packet fragmentation
- Support adjustable packet length, transmission frequency and sending quantity
- Achieve a transmission rate of **75 Gbps**

▸ 100G ROCEv2 Transmission

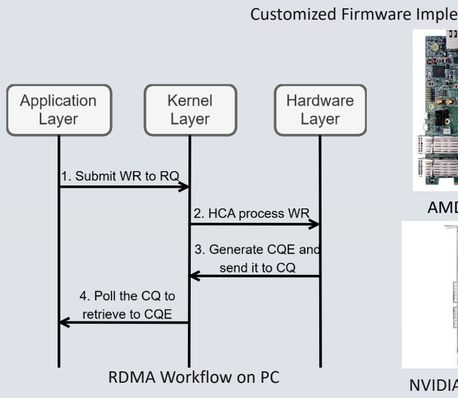
- Developed Based on Open-source Project: *https://github.com/Gabriele-bot/100G-verilog-RoCEv2-lite*
- Support RDMA Send/Write operations
- Throughput reaches **~92 Gbps** for packets larger than 5 KB
- Single-packet latency is ~730 cycles, about **2.27 µs**



Customized Firmware Implementation

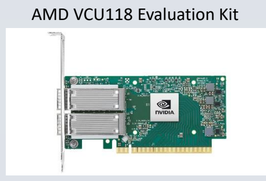


Throughput Test for 100G ROCEv2 Transmission

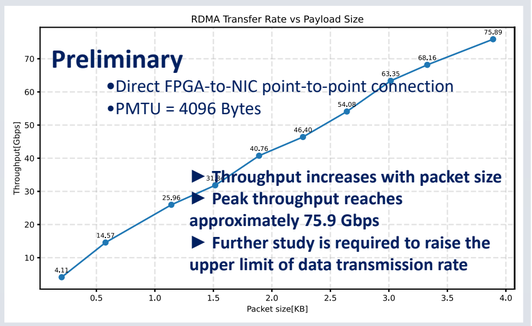


RDMA Workflow on PC

AMD VCU118 Evaluation Kit



NVIDIA Mellanox ConnectX-5 EN



Throughput Test Result for Customized Network Stack

Future Application of RDMA in BESIII

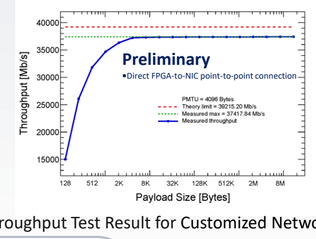
▸ 40G ROCEv2 Transmission

- Hardware: BESIII Trigger Core Board, Nvidia Mellanox Connectx-4
  - ✓ For details regarding the BESIII trigger system, please refer to Poster No.192 by Haoxin Wang: *Design of a Full Trigger Data Readout Scheme for the BESIII MDC Sub-trigger System*
- Support: RDMA Send/Write, Retransmission module with on-chip RAM
- Throughput: 37.4 Gbps maximum, 93.5% bandwidth utilization

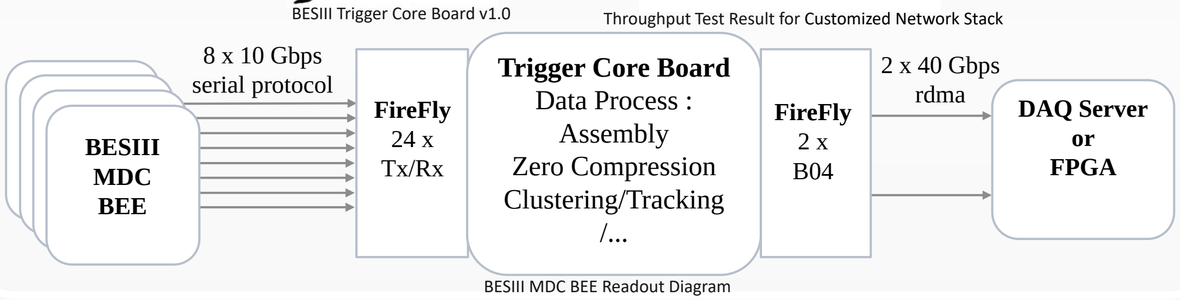
▸ Beijing Spectrometer III (BESIII) BEE Readout

- Applying 40G ROCEv2 transmission to the MDC (Main Drift Chamber) detector for full-data readout links

BESIII Trigger Core Board v1.0



Throughput Test Result for Customized Network Stack



BESIII MDC BEE Readout Diagram

Future Application of RDMA in CEPC

RDMA-based data forwarding from BEE to DAQ farm

▸ Design Goals of RDMA-Based BEE Readout System for Circular Electron-Positron Collider (CEPC)

- Use RDMA instead of traditional protocols for BEE-to-DAQ data transmission, improving efficiency and reducing computing resource usage
- L1 assembly implemented in BEE readout module, enables the dispatch of event data segment to individual HLT nodes

Conclusion and Outlook

RDMA enables high-throughput, low-latency data transfer for BEE readout. The custom protocol stack achieves 75 Gbps peak throughput, while the open-source stack reaches 92 Gbps with only 2.27 µs single-packet latency, demonstrating excellent transmission efficiency. This design validates the feasibility of both firmware and software. Future work will focus on optimizing the custom stack for CEPC DAQ applications and implementing 40 GbE RoCEv2 for BESIII BEE readout.