v1.0

JLab LDRD2410 Quarterly Report FY24Q1

Proposal Name:
Streaming Readout Real-Time Development and Testing Platform

Report Date:
Jan. 8, 2024

Principal Investigator:
David Lawrence PhD

1 Work-scope Highlights

Work is well underway with 5 different people contributing directly to the project in different capacities.

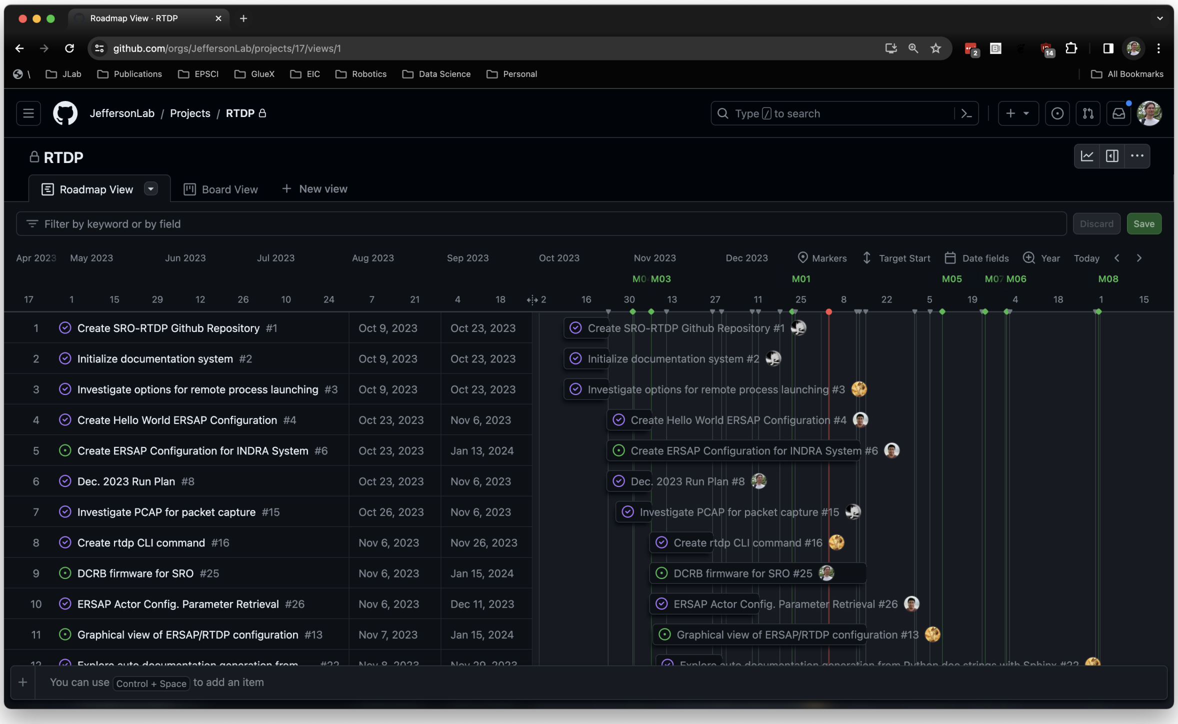
The major highlights are:

- CLAS12 Data capture at multiple beam currents ranging from 25nA to 150nA
- Github repository established with connected Project for task tracking
- Review of field for remote project launch options
- Visualization R&D for YAML configurations
- Preliminary ERSAP configurations

Significant progress has been made on the milestone schedule as detailed in section 2. A significant event was the data capture exercise with CLAS12 using 3 hours of beam time during the Run Group K commissioning period on Dec. 17, 2023. The time was used to capture TCP packets from two CLAS12 VTP modules reading out the ECAL and PCAL for two sectors of CLAS12. The data was streamed directly from the VTP modules to CEBAF Center to an EJFAT computer where the data was captured. The EJFAT NIC (Networking Interface Card) placed hardware timestamps on the packets which were recorded as part of the metadata attached to the packets. This is significant since the metadata and payload data were captured together in a form that can be replayed offline with the same time and packet structure that was present online. A diagram of the setup can be seen in figure 2. Figure 3 shows a graph of the time difference between subsequent packages for each of the TCP ports in the exercise.

We have also established a GitHub repository for code related to the project and have begun to populate it with various programs and scripts relevant to it. The relatively new GitHub “Projects” feature is also being used to list and track tasks and milestones. See figure 1 for a screencap showing part of the roadmap view of the project in that platform.

A review of some options for launching multiple processes distributed across multiple distributed compute facilities was done. This is a capability required for this project. It was determined that a combination of technologies will be needed with JIRIAF and ssh supplying most (if not all) of them. It was also determined that ssh launching would be a useful feature in general for ERSAP so will be added to that under that project. A brief report reviewing some options can be found on the RTDP wiki here (see “Draft review of MQTT, Ansible, Globus, SuperFacility and FireWorks” by Xinxin Mei and Jeng-Yuan Tsai):



1	Create SRO-RTDP Github Repository #1	Oct 9, 2023	Oct 23, 2023
2	Initialize documentation system #2	Oct 9, 2023	Oct 23, 2023
3	Investigate options for remote process launching #3	Oct 9, 2023	Oct 23, 2023
4	Create Hello World ERSAP Configuration #4	Oct 23, 2023	Nov 6, 2023
5	Create ERSAP Configuration for INDRA System #6	Oct 23, 2023	Jan 13, 2024
6	Dec. 2023 Run Plan #8	Oct 23, 2023	Nov 6, 2023
7	Investigate PCAP for packet capture #15	Oct 26, 2023	Nov 6, 2023
8	Create rtdp CLI command #16	Nov 6, 2023	Nov 26, 2023
9	DCRB firmware for SRO #25	Nov 6, 2023	Jan 15, 2024
10	ERSAP Actor Config. Parameter Retrieval #26	Nov 6, 2023	Dec 11, 2023
11	Graphical view of ERSAP/RTDP configuration #13	Nov 7, 2023	Jan 15, 2024
12	Explore auto documentation generation from #22	Nov 8, 2023	Nov 29, 2023

Figure 1: Screencap of the Github project. Tasks are entered and associated with individual project milestones. They are assigned to individuals so progress may be tracked. This itself applies to milestone M04.

https://wiki.jlab.org/epsciwiki/index.php/SRO_RTDP#Documentation

Preliminary work has started on visualization of the YAML configurations that can be dynamically updated with telemetry from a working system. Similarly, the new CS postdoc, Ayan Roy, has started to familiarize himself with ERSAP configurations and has written an example "Hello World" along with some example actors. This is important as ERSAP will be a key part of the software stack on which RTDP will be built.

2 Milestone Progress

Table 1 shows the status of the FY24Q1 milestones. Two have been partially completed and two have been fully completed. The scope of milestone M01 may be increased as it was determined that RTDP configurations will encompass a combination of CODA and ERSAP configurations. The M02 milestone regarding CLAS12 data was perhaps the most significant task for FY24Q1 and is complete. Doing the same for the setup in the INDRA lab will be simple and can be done at anytime so was deemed lower priority.

3 Budget

Figure 4 shows the project spending for FY24Q1. This is considered on track with the project expectations.

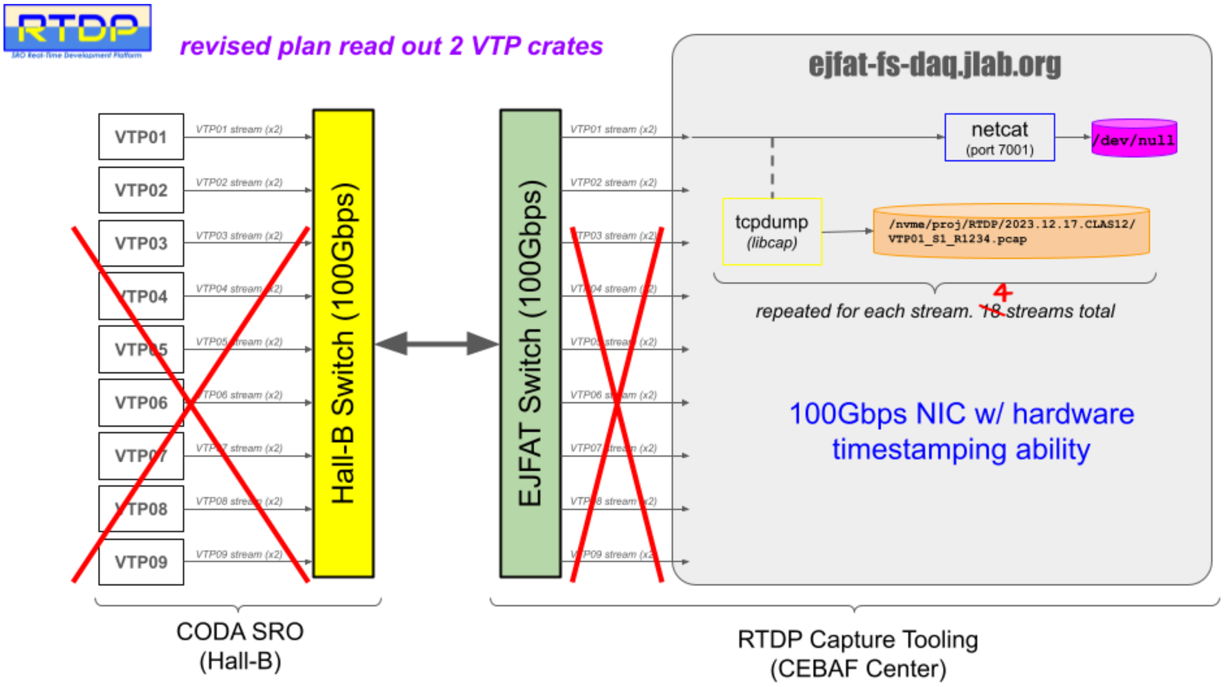


Figure 2: Diagram illustrating the Dec. 17th Data Capture Exercise with CLAS12. Technical issues with the firmware restricted the number of VTPs that could be used to 2 as opposed to the 9 originally envisioned. (Hence, the red “X”s.)

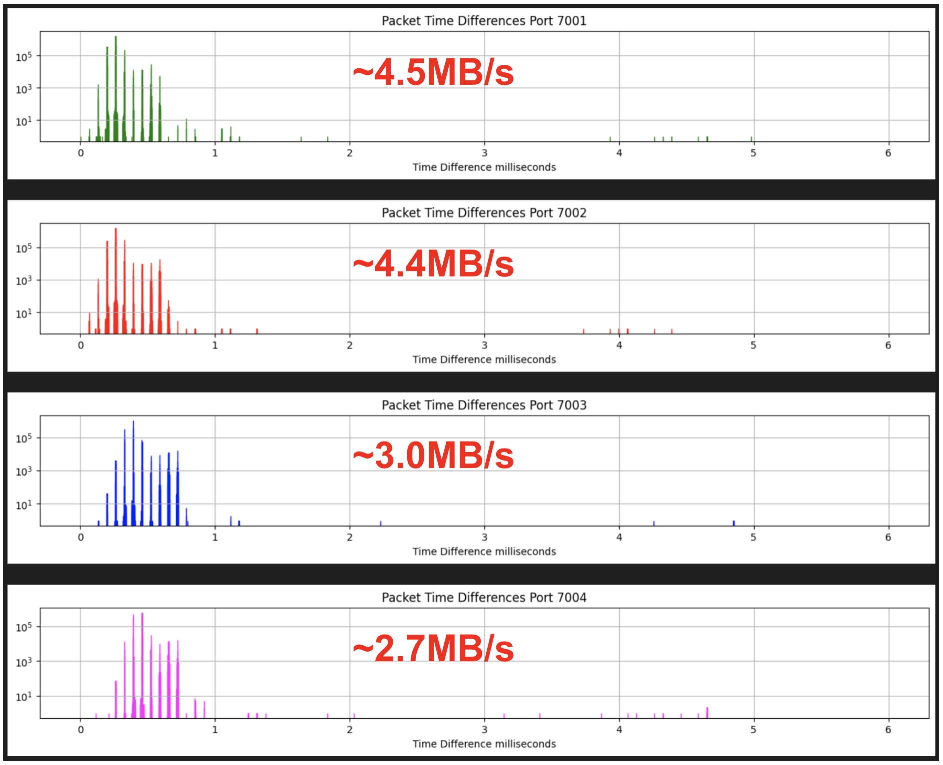


Figure 3: Time difference in milliseconds between packets arriving at ejfat-fs node in CEBAF center streamed from Hall-D DAQ system. The different plots/colors represent different destination ports.

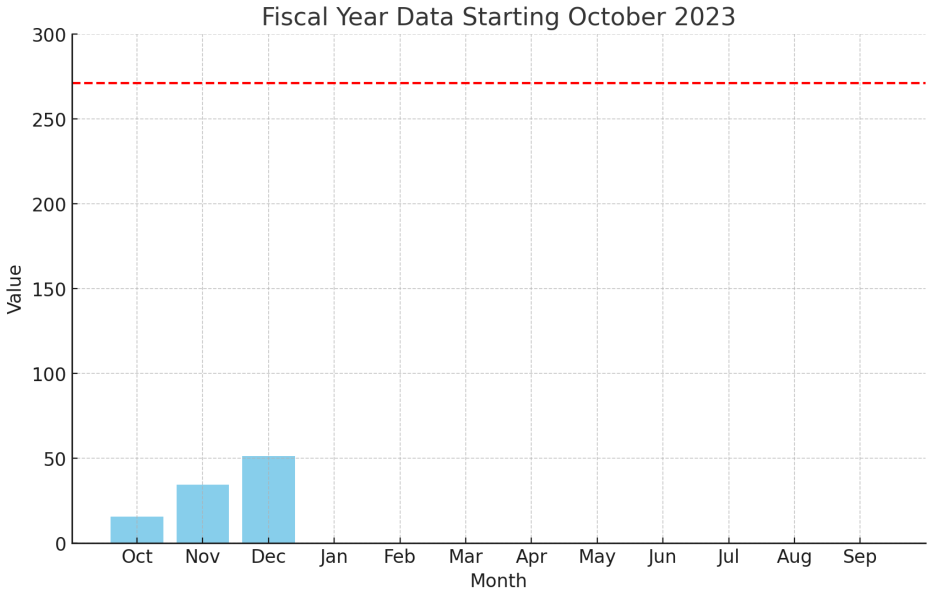


Figure 4: LD2410 Project Spending through FY24Q1. Values are in $K.

ID	**Task**	**status**	**Comments**
M01	Create prototype ERSAP configurations for INDRA and CLAS12 test systems	○	A CLAS12 example and ”Hello World” example have been placed in Github. INDRA has not been done yet.
M02	Identify or capture SRO formatted data from CLAS12 and INDRA test systems with data tag/filtering capability (output data ready for further offline processing)	○	Data was captured at vrious beam current from CLAS12 on Dec. 17. No INDRA data captured yet.
M03	Evaluate existing solutions for configuring and launching remote distributed processes	✓	see evaluations in document on EP-SCI wiki.
M04	Establish code repository(s), project site, and method of documentation	✓	This has been done here: https://github.com/JeffersonLab/SRO-RTDP

Table 1: FY24Q1 Milestones

4 Concerns

No concerns at this time.

Acknowledgements

The research described in this report was conducted under the Laboratory Directed Research and Development Program at Thomas Jefferson National Accelerator Facility for the U.S. Department of Energy.

Appendix: Full Project Milestones

- **Y1Q1**
  - M01: Create prototype ERSAP configurations for INDRA and CLAS12 test systems
  - M02: Identify or capture SRO formatted data from CLAS12 and INDRA test systems with data tag/filtering capability (output data ready for further offline processing)
  - M03: Evaluate existing solutions for configuring and launching remote distributed processes
  - M04: Establish code repository(s), project site, and method of documentation
- **Y1Q2**
  - M05: Create stream splitter program for EVIO or HIPO data formatted files
  - M06: Create stream splitter program for simulated data in PODIO for ePIC
  - M07: Create VTP emulator using files produced by stream splitter
  - M08: Create controller program to synchronize multiple VTP emulators
- **Y1Q3**
  - M09: Determine appropriate schema for all aspects of monitoring system.
  - M10: Establish databases for monitoring system using existing JLab servers.
  - M11: Integrate Hydra as monitoring component.
- **Y1Q4**
  - M12: Integrate off-line data analysis framework into platform for CLAS12 data
  - M13: Integrate off-line data analysis framework into platform for ePIC or GlueX simulated data
  - M14: Integrate example JANA2 analysis into platform
- **Y2Q1**
  - M15: Create configurable CPU proxy component
  - M16: Create configurable GPU proxy component (hardware and software)
  - M17: Create configurable FPGA proxy component (hardware and software)
  - M18: Create functioning hardware GPU component (e.g., CLAS12 L3)
  - M19: Create functioning hardware FPGA component (e.g., ML4FPGA)
- **Y2Q2**
  - M20: Impose artificial time structure on stream sources to mimic beam-like conditions
  - M21: Configure simulation of full SRO system using existing JLab hardware resources
- **Y2Q3**
  - M22: Establish working test of system that transfers ¿=100Gbps from CH to compute center
  - M23: Establish working test of system that includes GPU component for portion of the stream
  - M24: Establish working test of system that includes FPGA component for portion of the stream
  - M25: Test system with remote compute facility (e.g., BNL or NERSC) at limits of available resources
- **Y2Q4**
  - M26: Configure system that results in stream(s) being received by JLab from external source
  - M27: Collaborate with HPDF group to evaluate processing SRO data at JLab for external experiments
  - M28: Complete documentation for platform to be used by non-experts