100GbE2FE – Evaluation of Ethernet as a detector front-end readout link

Valentin Stümpert[1,2], Adam Klekotko[1,3], Alberto Perro[4], Daniel Hernandez Montesinos[1], Francesco Martina[1], Paschalis Vichoudis[1], Sophie Baron[1], Stefan Biereigel[1]

[1] CERN [2] Institute for Data Processing and Electronics (IPE), Karlsruhe Institute of Technology (KIT) [3] KU Leuven
[4] Universite d'Aix-Marseille III

Abstract

New-generation detectors, with a potentially less demanding radiation environment, are creating a need for flexible, high-speed data links. As a part of EP RnD WP6 and in collaboration with WP9, an examination of commercial-grade Ethernet as a front-end readout link is underway, potentially enabling off-the-shelf hardware to be used in data readout systems and cutting down complexity using modern data center technologies.
We present the encouraging first results of this RnD effort, evaluating 100Gb/s Ethernet for data readout in the context of typical High-Energy Physics detector requirements. Due to asymmetries in data rate requirements in up- and down-links, unidirectional Ethernet is examined. Results from a recent radiation study, allowing a preliminary assessment of radiation hardness via statistical analysis, are provided. The application is verified with realistic traffic using a demonstrator to translate from lpGBT to Ethernet and a road map for future demonstrators is presented.

Motivation

Current DAQ systems are based on custom links.

- Require specially designed hardware back-ends
- In some systems: a second custom link going to a dedicated interface stage before the data center

Solution may lie with Ethernet:

- fully compatible with commercial hardware
- Physics data can directly offload to the data center
- **But:** Not designed for HEP requirements

Challenges:

- Timing accuracy and latency
- Radiation hardness
- Unidirectionality

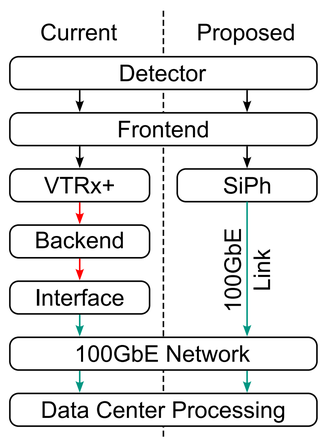

Figure 1: Comparison between current and proposed architecture. Custom links are marked in red, industry standard links in green.

Timing Accuracy and Latency

Ethernet may not fulfill HEP requirements for high timing accuracy and low latency:

- Inherent long buffer- and switch traversal times
- Variable latency and packet order changes due to packets traversing the network on arbitrary paths
- This can be mitigated via architectural measures eliminating cycles in the network (see Figure 2)
- High jitter allowance (SR4: 0.39 Unit Intervals)
- Ethernet is not bunch clock synchronous, potentially necessitating a CDC in the front-end

Because of these factors, Ethernet is best suited for deployment on self-triggered or trigger-less systems with dedicated timing distribution.

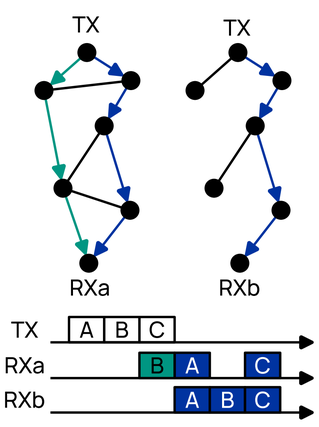

Figure 2: Effect of network architecture on latency and packet order. In case A, cycles in the network create different paths with varying latencies, causing order changes. In case B only one path exists between any two nodes.

Radiation Hardness

CERN links (e.g. lpGBT) are rad-hard by design:

- Small frames, small "attack surface"
- Strong FEC schemes: $\mathrm{RER}^* = 10\%$
- but large overhead: $\mathrm{OH} = 20\%$

Ethernet does not have these measures

- Large frames, large "attack surface"
- Weaker FEC scheme: $\mathrm{RER}^* = 1.3\%$
- More efficient overhead: $\mathrm{OH} = 3\%$

Results from radiation studies are promising.

* RER: Recoverable Error Ratio - Maximum correctable ratio of errored bits

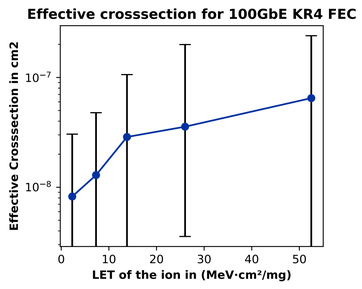

Figure 3: Expected vulnerable crosssection of a 100GbE transmitter extracted from the DART28 demonstrator ASIC. Note that these figures are not yet adjusted to the more favorable LET distribution in the LHC.

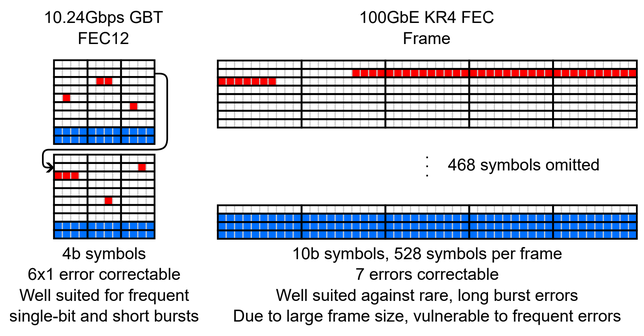

Figure 4: Comparison between lpGBT FEC12 and 100GbE KR4

Unidirectionality

- Full-speed duplex link is too complex for small down-link bandwidth
- No asymmetric 100GbE standard exists, only for lower speeds
- **But:** Optical link can operate with only one fibre
- This has been demonstrated with a wide range of commercial hardware

Proof-of-Concept

- FPGA-based demonstrator for 100GbE traffic
- Proof-of-concept and technology familiarization
- Evaluation of multiple commercial transceiver modules
- Unidirectional traffic demonstration
- Full-speed saturation testing with multiple network hops

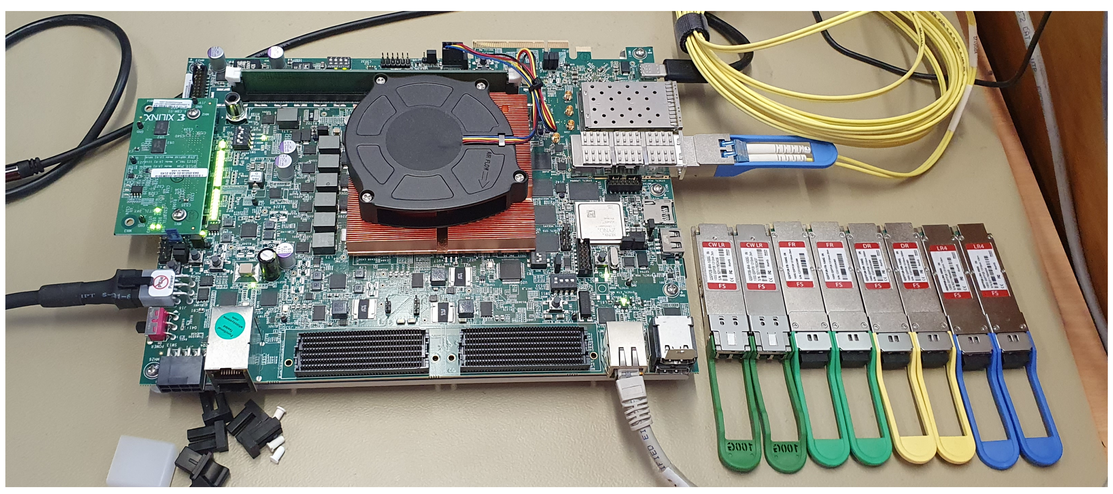

Figure 5: Versal FPGA with LR QSFP+ module in 100GbE testing setup. Additional module types shown (sets of two, from left to right): IR4, FR, DR, LR4

Protocol Translator

- FPGA-based translator between lpGBT and Ethernet
- Resolves lpGBT FEC and extracts payload
- Generates UDP packets from accumulated lpGBT traffic
- Effective data rate of 9.3Gbps from a 10.24Gbps lpGBT link due to lower overhead
- Implemented in two FPGA families: AMD Artix Ultrascale+ and Microchip Polarfire

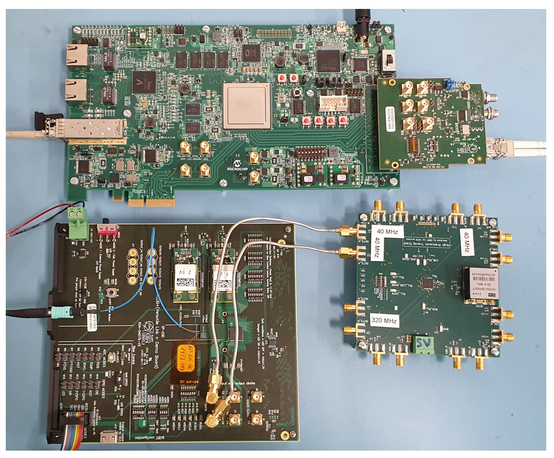

Figure 6: Polarfire protocol translation test bench with lpGBT and clock generator.

Conclusion and Outlook

- Standard evaluation and first results show no major roadblocks
- Caveats for timing: separate timing link and CDC may be needed
- Early results from radiation tests indicate sufficient radiation hardening due to built-in FEC
- Further efforts focusing on demonstrating practical viability:
  - Plans to package protocol translator in industry-standard SFP+ housing
  - Potential for multi-link translators in QSFP form factor
- **End Goal:** Prototype system integrating native 100GbE front-end demonstrator based on silicon photonics with legacy links using protocol translators

Acknowledgements

This work has been sponsored by the Wolfgang Gentner Programme of the German Federal Ministry of Education and Research (grant no. 13E18CHA)

EP R&D
WP6 High Speed Links

Valentin Stümpert
EP-ESE-BE, CERN
valentin.stumpert@cern.ch

SPONSORED BY THE
Federal Ministry of Education and Research

Karlsruhe Institute of Technology

EP R&D Day, May 22nd 2024, CERN, Poster #32