High performance, low latency data acquisition for AO RTC

RTC4AO^{6th}

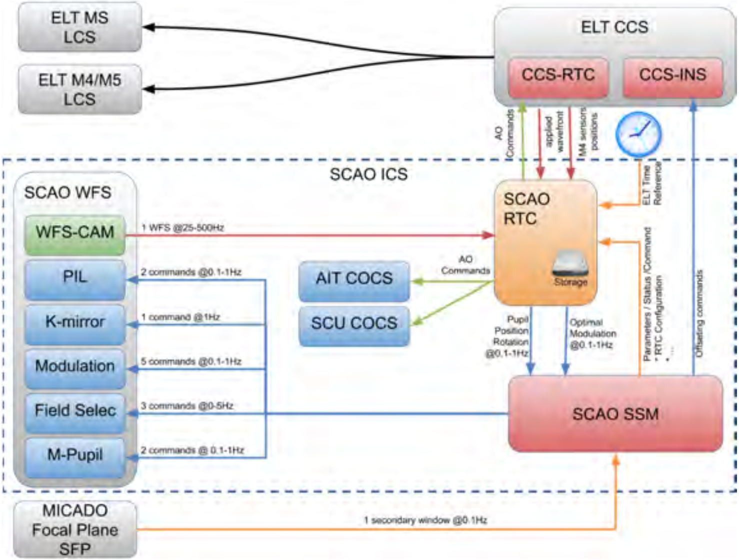
Julien Plante
julien.plante@protonmail.com

07-10-2023

Outline

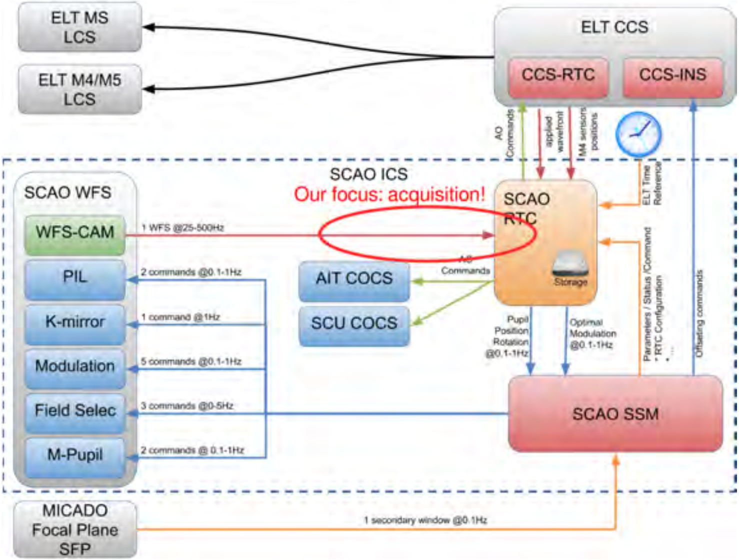
1. Context
   - MICADO RTC
   - State-of-the-art Ethernet data acquisition methods
2. Feeding a GPU from a NIC
   - DPDK
   - GPUDirect RDMA
   - gpudev
   - Final design
3. Results
   - Adaptive Optics
   - Others
4. Conclusion

MICADO SCAO RTC



(Clénet et al. 2019)

MICADO SCAO RTC



(Clénet et al. 2019)

State-of-the-art data acquisition methods

- Considering latency from packet reception at the NIC/FPGA and availability in GPU memory
- Interesting metric in the case of pipelining

Method	Latency (µs)	ptp (µs)
FPGA GPUDirect (Perret et al. 2016)	5.24	2
DPDK + `cudaMemcpy`	22.44	16

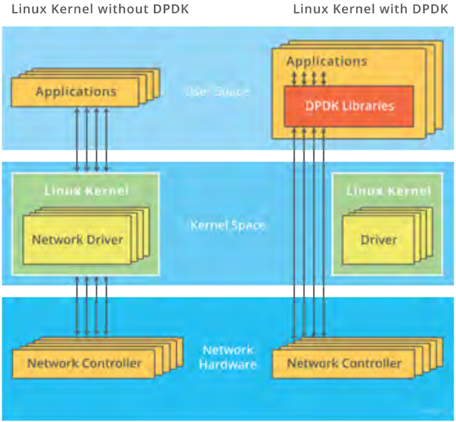
Goal

Very good solution based on FPGA, but can we achieve the same on COTS hardware ?

- Use of Smart NIC
- Simpler system
- Different maintenance plan

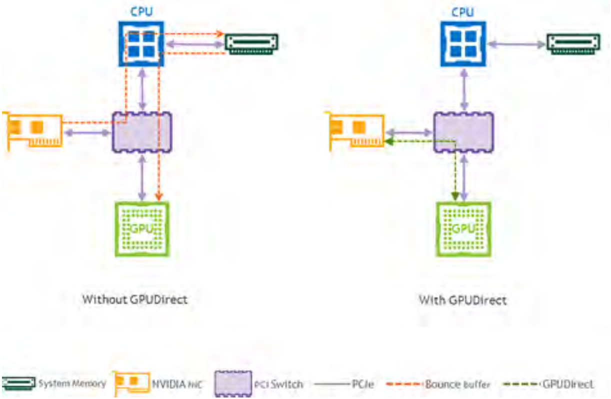
DPDK

- Open-source userspace networking library
- Less overhead than the Linux networking stack
- Access point for advanced features of Smart NICs



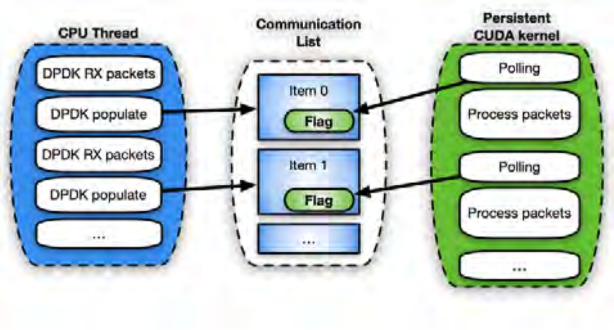
GPUDirect RDMA

- DMA between a GPU and a third party device over PCIe
  - → Less transfers, less overhead
- Implemented in DPDK for Nvidia Ethernet NICs



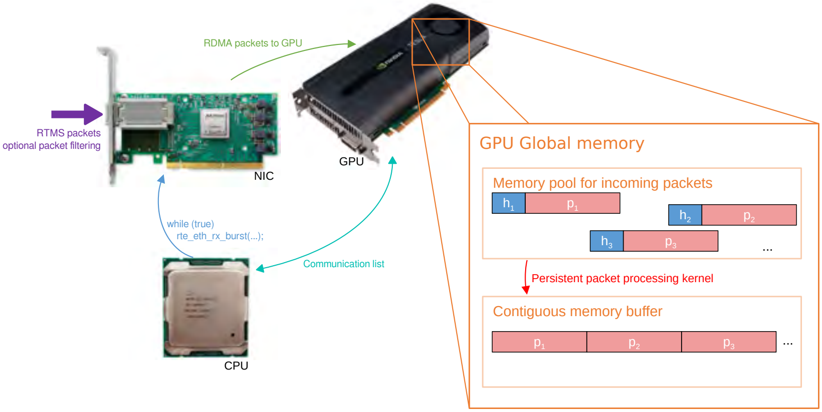
gpudev

- DPDK module dedicated to GPU control
  - Abstractions (malloc, memset, ...)
  - Utilities (mapped flag, communication list)
- Framework guiding to best performance
  - Persistent kernels
  - Data locality



Final design overview

- Semi-persistent kernel: launched for each image frame



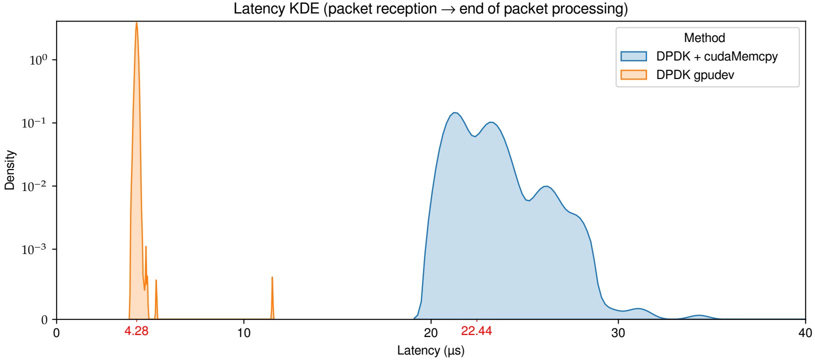
Final design overview

Feature set:

- Support for multiple WFS → scalability for other instruments, such as MAVIS
- One CUDA block per WFS → minimal GPU usage
  - Possibility to overlap acquisition with computations
- Partial support (as of today) of the RTMS protocol
  - ✓ Packet Type, Endianness, Payload Size, Payload Tag, Padding, Incomplete frame management, WFSCs ALICE and LISA
  - ✗ Stream Error, Version, IsSimSource, Extended Info, WFSC FREDA
- Ongoing validation with the ESO WFS Simulator

Adaptive Optics

- Benchmark over ALICE WFSC frame transmission
- Point-to-point communication over a 40 GbE link
- Nvidia Connectx-5 NIC + A100 GPU
- Custom sender, written in pure DPDK (no GPU)



Adaptive Optics

Method	Latency (μs)	ptp (μs)
FPGA GPUDirect (Perret et al. 2016)	5.24	2
DPDK + `cudaMemcpy`	22.44	16
DPDK gpudev	4.28	7

- Comparable performance to FPGA GPUDirect
- Higher observed jitter
- Much better latency than DPDK + `cudaMemcpy`

Radioastronomy

- Real-time Fast Radio Burst (FRB) detection on NenuFAR
  - Strong reliability constraints
  - Medium data rate (10 Gbit$/$s)
  - Strong scientific goal
- Deployed on site, first iteration functional
- Validated on pulsar data

Radar

- Real-time acquisition for Primary or Secondary radars
  - Strong reliability constraints
  - Huge data rate
  - Medium latency constraints
- Validated up to 99.7 Gbit/s on 100 GbE hardware (no packet loss)
- Desire to scale up

Future work

- More benchmarks
  - Different latency measures (round-trip, per-packet vs per image frame, ...)
  - Latency at high bandwidth
  - Closer latency estimation, considering neglected effects
    - 10 GbE link
    - Readout pipelining
    - Computations
  - Better jitter measurement
- Complete support for MUDPI/RTMS
- Higher level protocol description
- Experiment using non-Nvidia hardware
- Scale up
- Jitter reduction
- Offload packet processing (DPU ? DOCA ?)

Conclusion

- High performance, general purpose data acquisition system for COTS hardware
  - Low latency (~ 4 µs)
  - High bandwidth(~ 100 Gbit/s)
- NIC $\leftrightarrow$ GPU DMA using DPDK gpudev
- Partial support for MUDPI/RTMS, under completion
- Currently under testing and integration in the COSMIC framework (module `streams`)
- Already deployed for radioastronomy on NenuFAR

General considerations

- Is it really simpler than FPGA ?
  - Complexity of DPDK
  - Dependency to Nvidia
- Tradeoff between reusability and maintainability

References I

Clénet, Yann et al. (Nov. 2019). "MICADO-MAORY SCAO Preliminary design, development plan & calibration strategies". In: *Adaptive Optics for Extremely Large Telescopes conference, 6th edition*. Québec, Canada. URL: https://hal.science/hal-03078430.

Perret, Denis et al. (2016). "Bridging FPGA and GPU technologies for AO real-time control". In: *Adaptive Optics Systems V*. Ed. by Enrico Marchetti, Laird M. Close, and Jean-Pierre Véran. Vol. 9909. International Society for Optics and Photonics. SPIE, p. 99094M. DOI: 10.1117/12.2232858. URL: https://doi.org/10.1117/12.2232858.

References II

Plante, Julien et al. (2022). “A high-performance data acquisition on COTS hardware for astronomical instrumentation”. In: *Software and Cyberinfrastructure for Astronomy VII*. Ed. by Jorge Ibsen and Gianluca Chiozzi. Vol. 12189. International Society for Optics and Photonics. SPIE, 121890U. DOI: 10.1117/12.2627827. URL: https://doi.org/10.1117/12.2627827.

Thank you for your attention!

Rising STARS

RISE International Network
for Solutions Technologies
and Applications of
Real-time Systems

l'Observatoire de Paris | PSL

Région îledeFrance

THALES