Green Flash

High performance computing for real-time science

Workpackage 7: Ecosystem

Mid-term Review, 1st February 2016

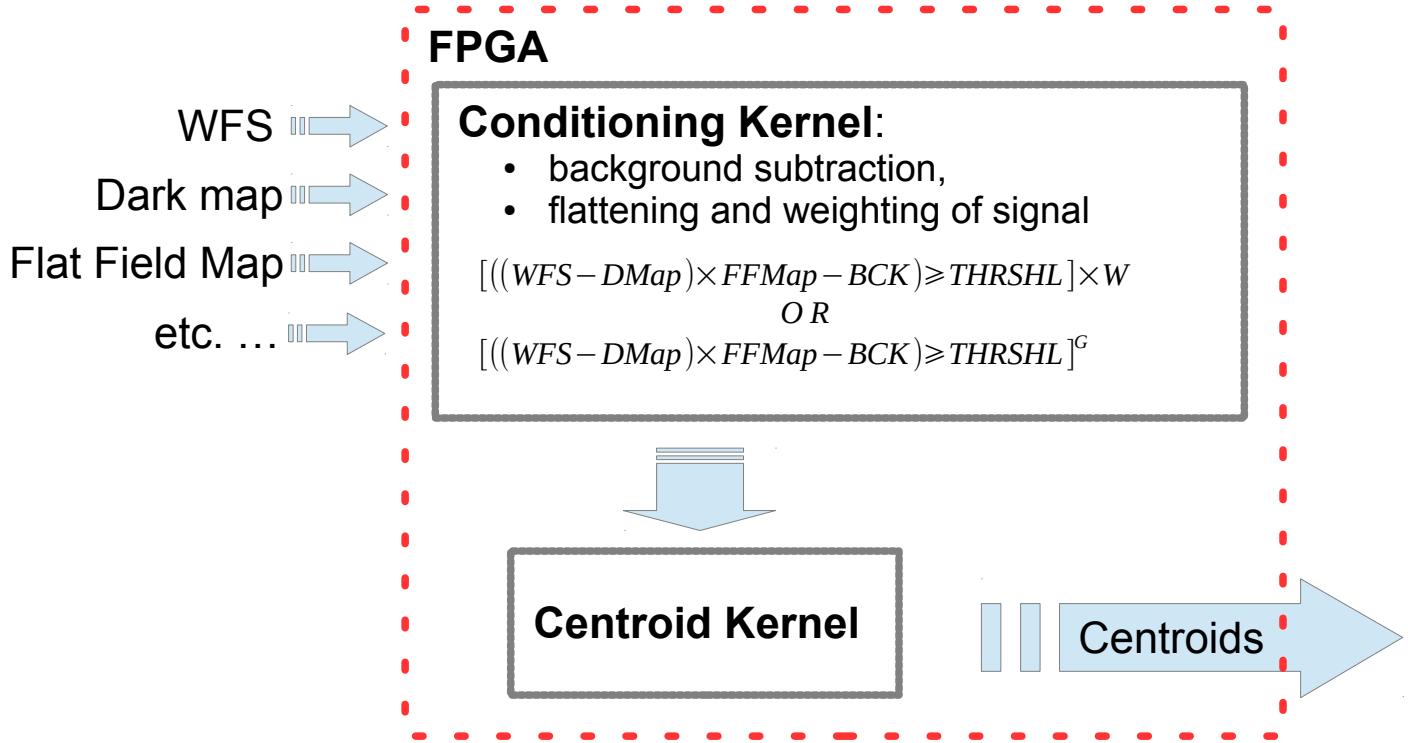
Project #671662 funded by European Commission under program H2020-EU.1.2.2 coordinated in H2020-FETHPC-2014

Tasks

- FPGA Development environment - PLDA
- Middleware – DU
- Algorithms and Libraries – OdP

FPGA development

- Data pipeline based on SPARTA specs
- FPGA programmed using QuickPlay entirely
  - Engineer had no experience in FPGA programming



Middleware

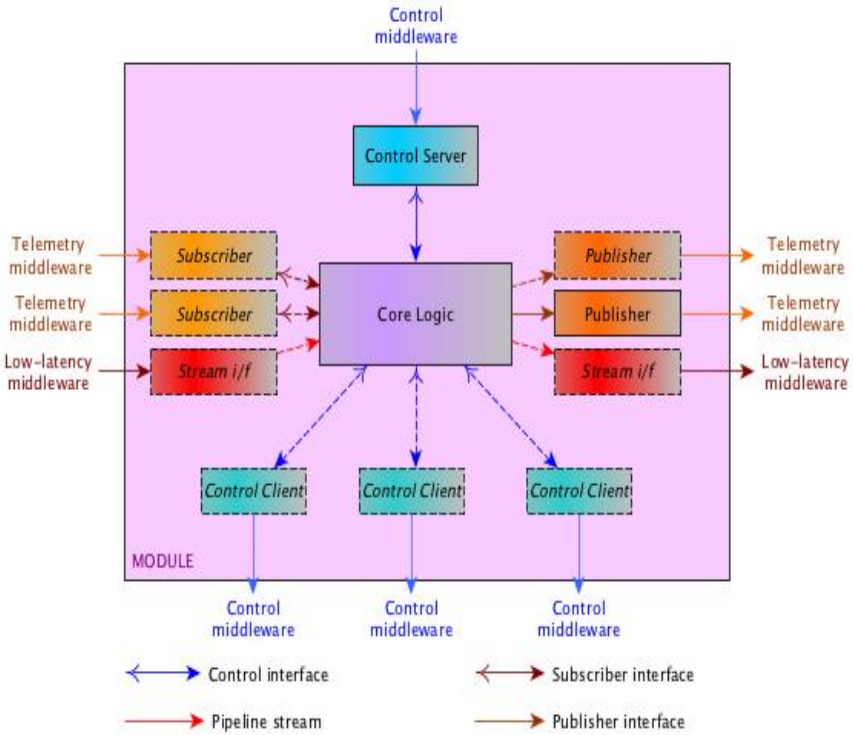
- 3 Middleware domains:
  - Control
  - Telemetry
  - Low-latency pipeline
- Closely related to design abstraction

Design Concept(1)

- E-ELT instrument control systems in 2024 are likely to use technologies which are not yet available
- Abstraction – must be possible to realise high-level design on different hardware/middleware
- Design interfaces which are independent of both core logic and middleware

Design Concept(2)

- Interfaces abstracted from core logic
- Module internal interfaces independent of middleware
- Middleware dependencies in separate components



Middleware Requirements

- Control:
  - Request/reply pattern
  - Service discovery/location transparency
- Telemetry:
  - Publish/subscribe pattern
  - Hard throughput, weak latency/determinism reqts
- Low-latency pipeline:
  - Hard latency, determinism, throughput reqts
  - Fan-out/fan-in pattern – distribute workloads

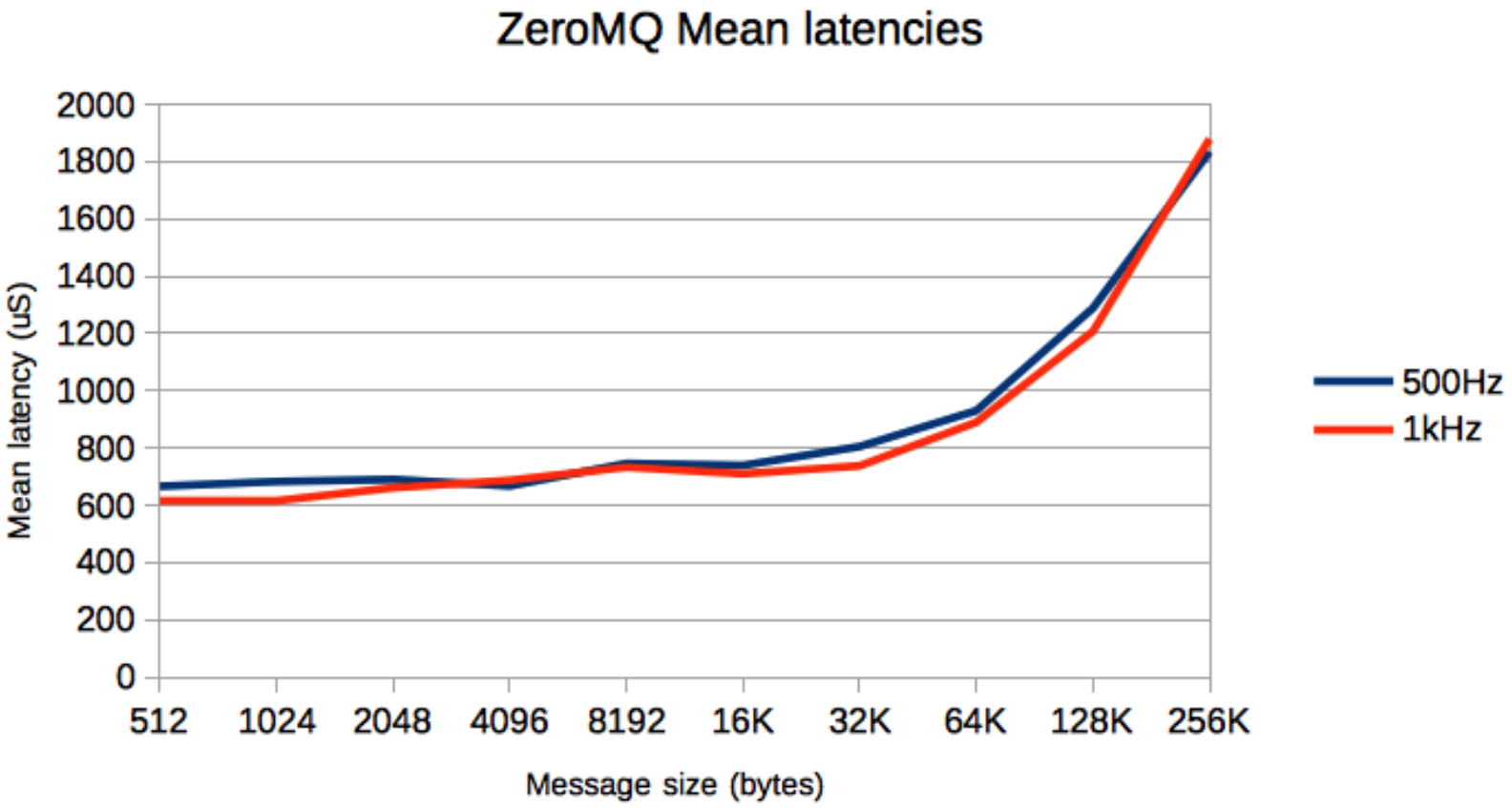
Technology evaluations

- Evaluations in progress
- Control:
  - DDS, ICE
- Telemetry
  - DDS, ZeroMQ/Google protocol buffers
- Real-time pipeline
  - ZeroMQ, MPI

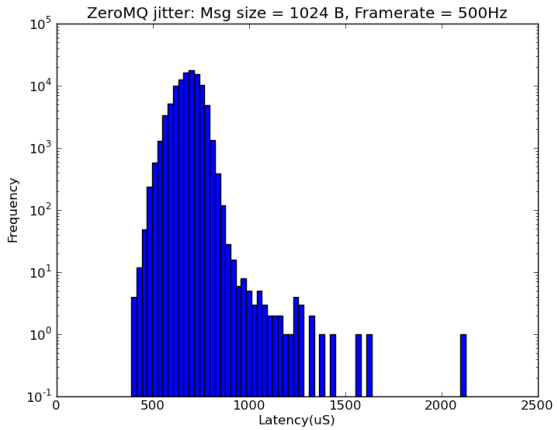
Requirements: real-time pipeline

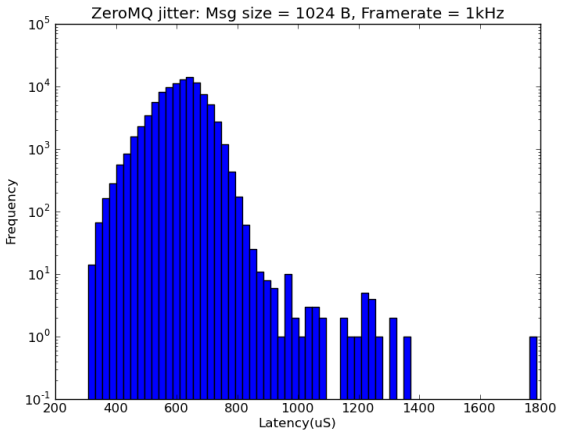
- Latency:
  - Goal is 1 frame latency from arrival of last pixel to beginning of DM output
  - At goal frame-rate of 1000Hz :- 1000uS
- Jitter:
  - Goal is 100uS peak-to-peak in any 1 second period

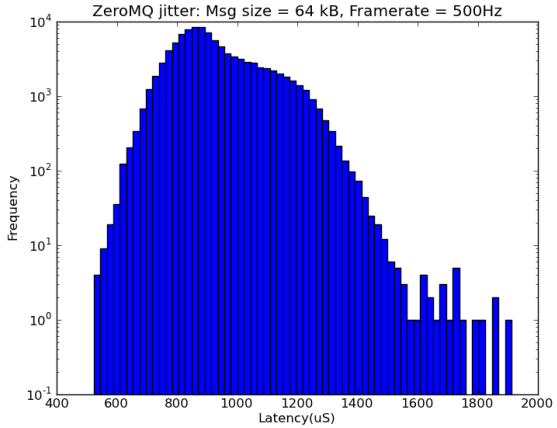
ZeroMQ

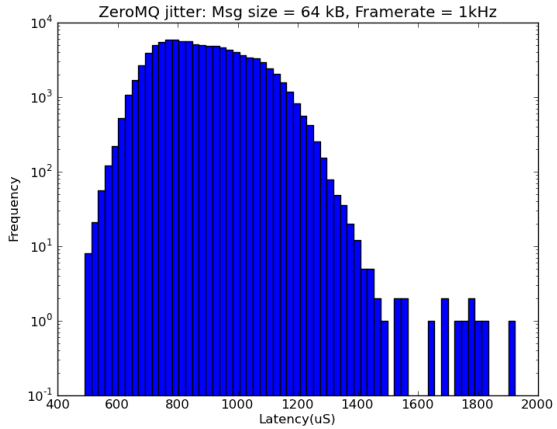

ZeroMQ Mean latencies
Mean latency (uS)
2000
1800
1600
1400
1200
1000
800
600
400
200
0
512
1024
2048
4096
8192
16K
32K
64K
128K
256K
Message size (bytes)
500Hz
1kHz

ZeroMQ


ZeroMQ jitter: Msg size = 1024 B, Framerate = 500Hz
Frequency
Latency(uS)

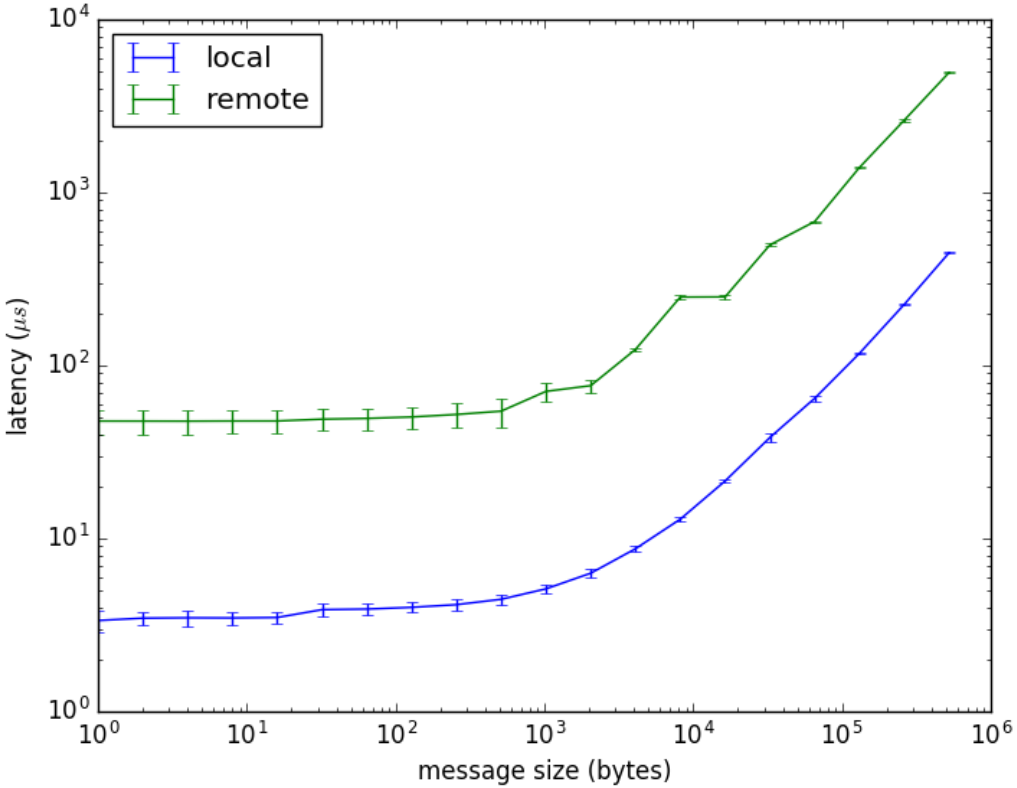

ZeroMQ jitter: Msg size = 1024 B, Framerate = 1kHz
Frequency
Latency(uS)


ZeroMQ jitter: Msg size = 64 kB, Framerate = 500Hz
Frequency
Latency(uS)


ZeroMQ jitter: Msg size = 64 kB, Framerate = 1kHz
Frequency
Latency(uS)

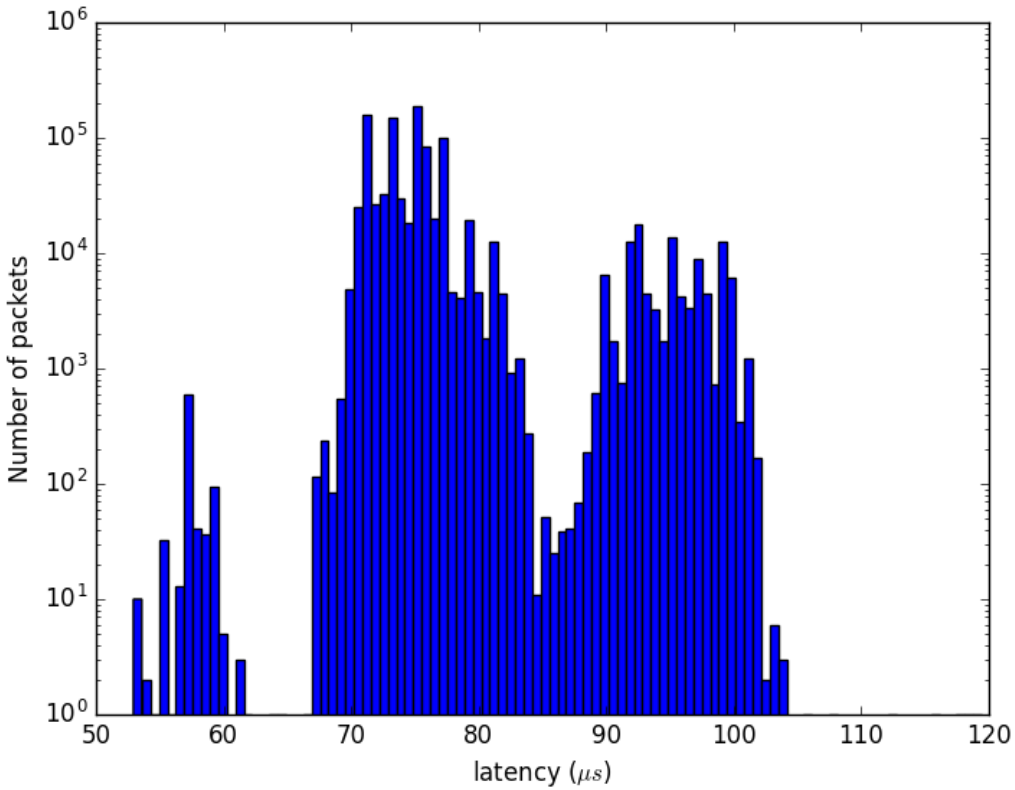
MPI

Mean latency vs. message size



MPI

Jitter, 1kB messages



Conclusions

- ZeroMQ: unsuitable for real-time pipeline
  - Excessive latency: > 3 x budget, probably owing to internal buffering and message aggregation
- MPI: latency and jitter adequate
  - ~ 5% of latency budget, for small messages
  - Hence, limited number of network hops allowed
  - Hence, some constraints on implementations using MPI

Plans

- Complete evaluation, downselect technologies - M24
- Middleware development for prototype – M24-
- Prototype performance report - M36

WP 7 : Ecosystem

Task 7.3 : algorithms and libraries

Definition of various algorithms and libs used in different sub-systems

- » RT data pipeline : compare custom code and standard libs
- » Supervisor : only based on standard libs

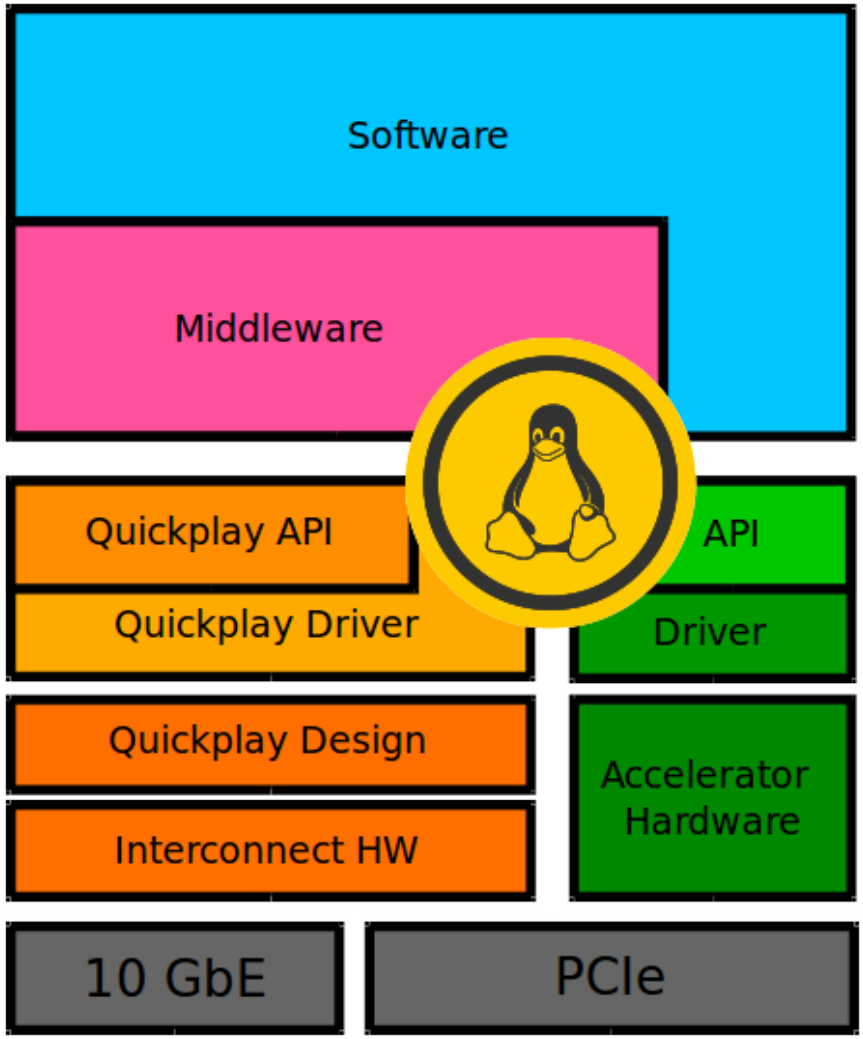
Use of accelerators / distributed memory systems

portability is key to mitigate obsolescence risk

use vendor-based (possibly proprietary) core libs for performance

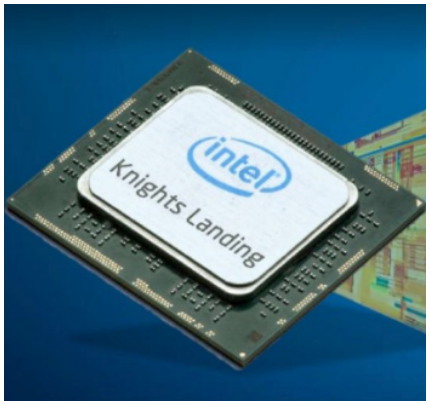
SW / MW stack

- Under development in Paris & Durham
- Run a standard HPC ecosystem on the prototype boards and clusters
- Performance assessed w/r AO application (mainly linear algebra)



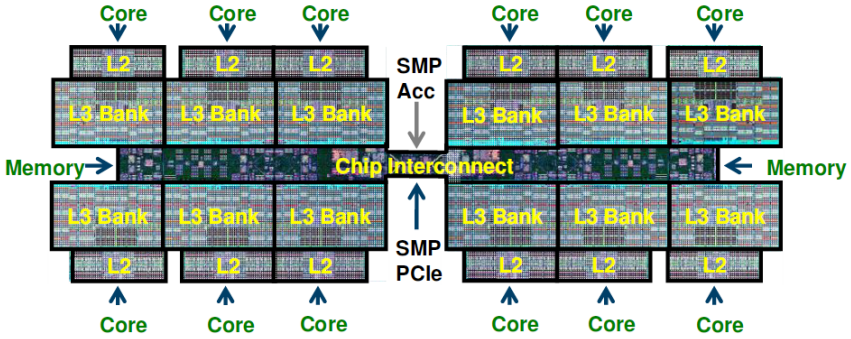
HW landscape

- Intel KNL



- Nvidia GPUs

- IBM Power 8



Unified framework

- Supervisor module : compute bound, most demanding stack
- Based on the Chameleon library

Tile algorithms

Sequential task-based programming model

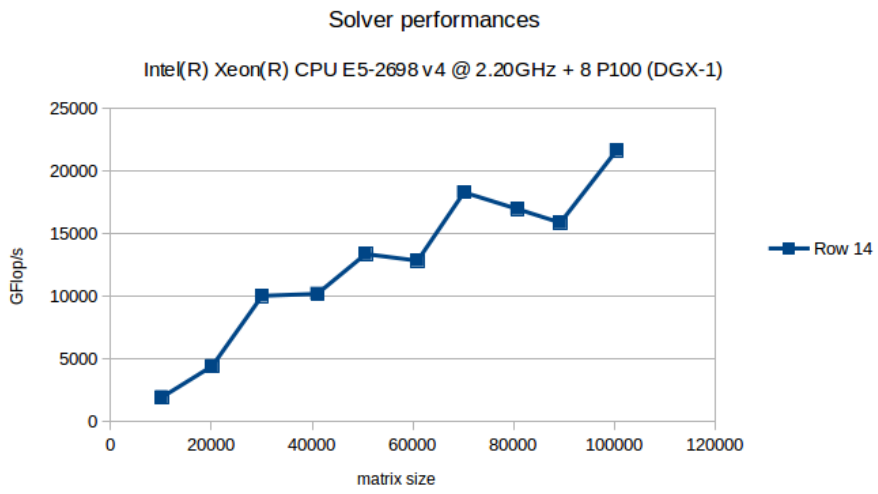
Dynamic runtime systems: StarPU, Quark, OpenMP, PaRSEC

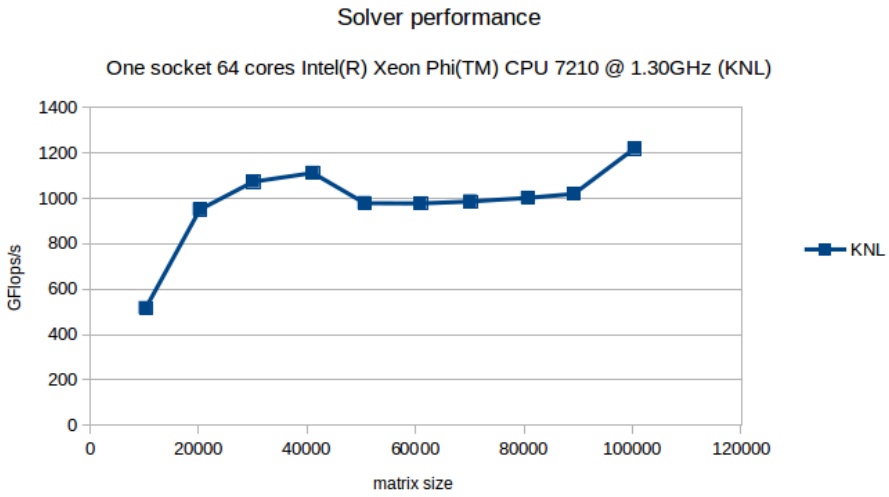
Oblivious to the underlying hardware

Unified framework

Performance for reconstructor matrix computation (“apply” process)

- Direct comparing between last generation of GPU (NVIDIA P100) and last generation of Intel Xeon Phi (KNL). **Same SW stack**





Green Flash

l'Observatoire de Paris — LESIA
Laboratoire d'Études Spatiales et d'Instrumentation en Astrophysique

Durham University

MICROGATE

PLDA