



# Green Flash

High performance computing for real-time science

WP5 Status – PLDA / Accelize

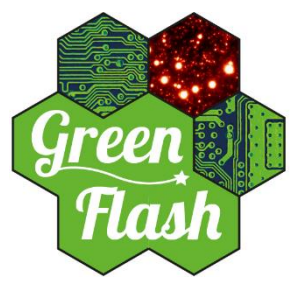
## Final Design Review – Meudon April 2018



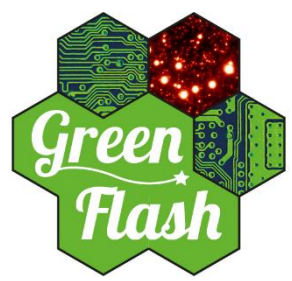


# Agenda

- WP5 – Past Achievements
- WP5 – Present Outcomes
- WP5 – Applications
- WP5 – Future
- Discussion



# WP5 – Past achievements



# WP5 – Context Reminder

- WP5 = Smart Interconnects

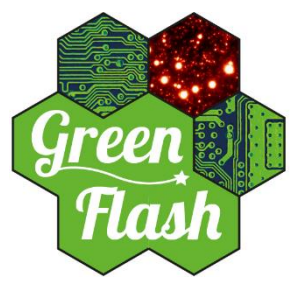
***“The goal of this WP is to provide a comprehensive study of a Smart Interconnect concept, in the context of the AO application, including hardware, firmware, middleware and development environment considerations”***

- WP5 is mainly structured around :
  - Development of new features for QuickPlay tool
  - Design of Smart Interconnect using QuickPlay tool



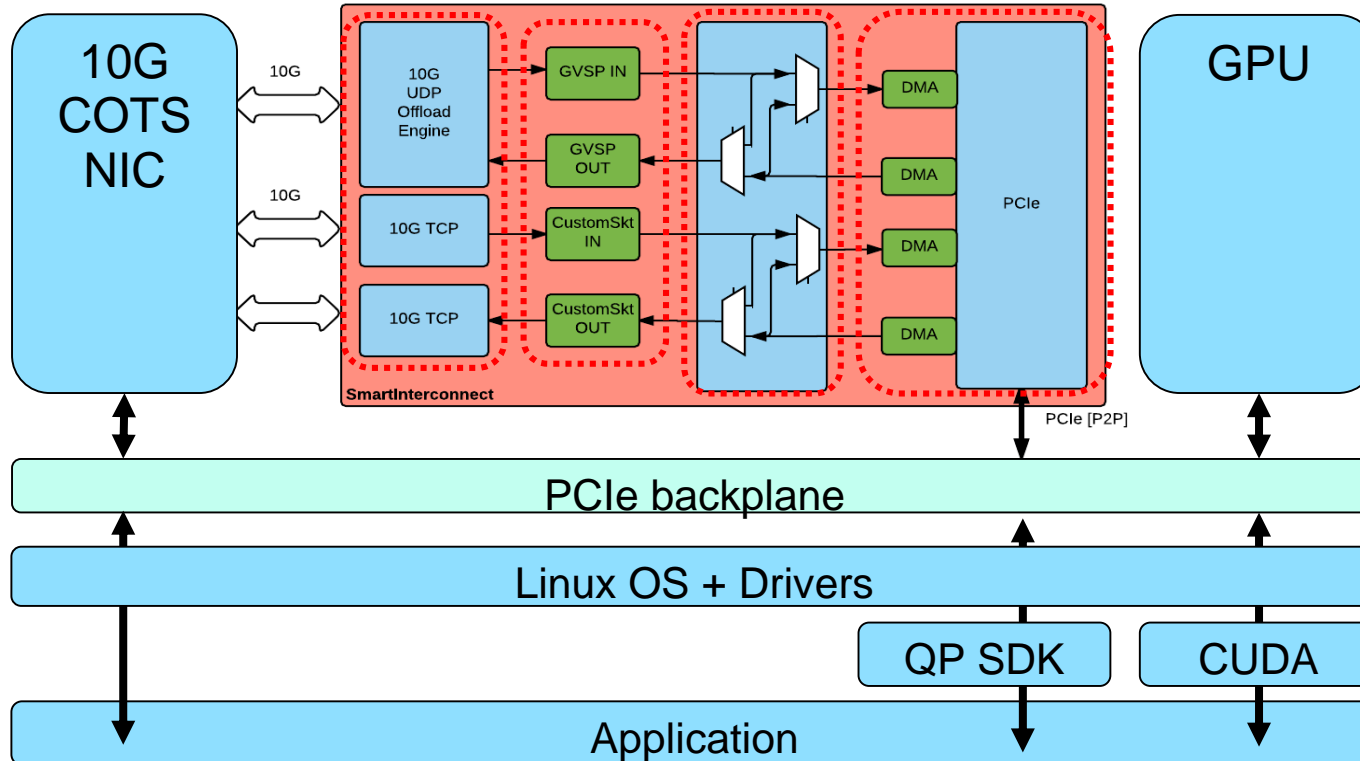
# MTR - WP5 Status

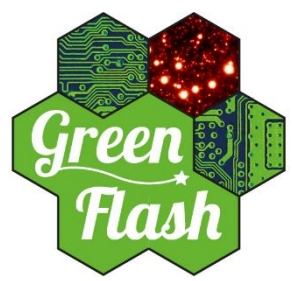
- 1<sup>st</sup> half of the project allowed to
  - Explore feasibility of different solutions
  - Extract solutions that appeared to be the most realistic and pertinent ... and discarding others
  - Start development and integration of these solutions
- Main outcomes
  - QuickPlay
    - improved tool maturity, improved HLS performances
    - Additional target support (Arria 10)
    - Integration of new IPs : UDP, UDP multicast, DDR4, PCIe Peer-to-Peer (basic)



# MTR - WP5 Status – cont'd

- Main outcomes
  - Smart NIC : Materialized through Smart Interconnect concept





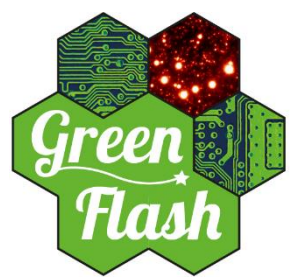
# MTR – WP5 Remaining objectives

- Remaining objectives after D5.1
  - Implementation of a new Smart Interconnect with :
    - Mitigation of limitations (PCIe and TCP/UDP performances)
    - Complete P2P feature integration
    - Microgate uXComp support
  - $\mu$ -server concept support (Arria10 SOC support for execution of QP designs)
  - Adaptative Optics implementation using QuickPlay



# WP5 – Present Outcomes





# WP5 – QuickPlay Features & Perf.

- Multiple implementations support (Eased migration of a same project to various FPGA families / boards)

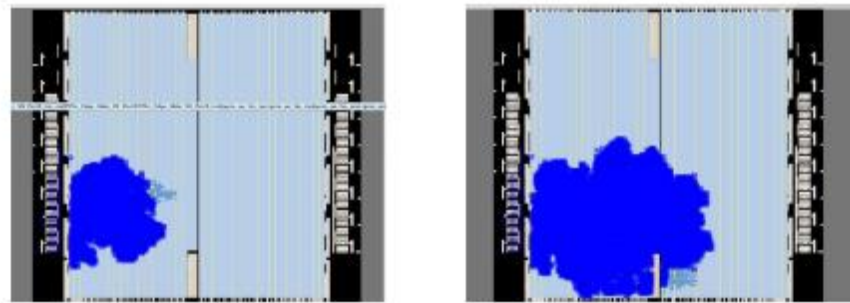
The screenshot shows the QuickPlay IDE interface. The top bar displays the project name: "20180319 - QuickPlay Development - MatrixVectorProduct\_DDR\_qpp/source/TopLevel.cpp". The Project Explorer on the left shows a tree view with "MatrixVectorProduct\_DDR\_qpp" selected, containing sub-items like "TopLevel", "Kernels", "Applications", "Reports", and "Implementations". The "Implementations" folder is expanded, showing "ARRIA10 (built)", "uXComp (built)", and "XpressKUS (not executed)". A red box highlights the "uXComp (built)" implementation. The Navigator on the right shows the current board: "Board: Microgate - uXComp" with a small image of the board. A red box highlights this information, with a red text annotation: "Targeted board for current implementation". The main workspace displays a hardware diagram with a central block "VectorMatrixProdu... VectorScalarProduct\_1" connected to "PCie data\_in" and "PCie data\_out" blocks, and a "Memory" block below it.

1 QuickPlay design, multiple implementations



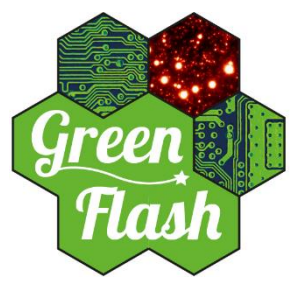
# WP5 – QuickPlay Features & Perf.

- Configurable PCIe :
  - Dynamic PCIe (built for each project), allowing reduced resources usage



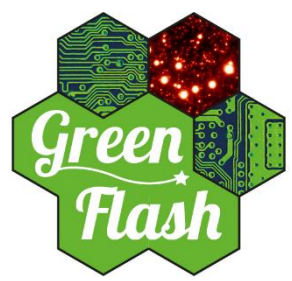
*Figure 1 : XpressGX5 layout for PCIe Loopback on 1 or 4 streams*

- Enumeration options (BAR opening for P2P, 32/64 bit enumeration)
- PCIe registers exposed to internal logic (eg for DMA configuration by embedded processor)
- Support of dynamic parameters in emulation (registers), allowing more complex/accurate C models
- But also... support of CSP providers platforms : Amazon AWS (Xilinx) / OVH (Intel DCP)



# WP5 – QuickPlay Features & Perf.

- SDK improvements :
  - Multi-board support
  - Improved TCP performances (around 9 Gbps)
  - Improved PCIe performances (around 5.5 Gbps for Gen3x8)
  - Addition of UDP streams support
  - Addition of TCP and UDP configuration and management methods
  - Peer-2-Peer support for Read/Write streams operations
- D5.10 (Scalability of QuickPlay designs) demonstrated tool versatility

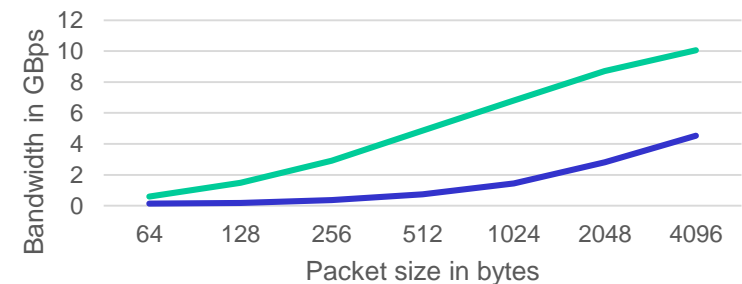


# WP5 – QuickPlay Boards & IPs

- Additional boards support :
  - Intel Arria 10 boards : Bittware A10PL4, Microgate uXComp, Intel RushCreek (DCP)
  - Xilinx board : Virtex Ultrascale (AWS F1 instances)
- Microgate  $\mu$ XComp board, including :
  - PCIe Gen3x8
  - 1 QSFP (4x10G) with TCP and UDP offloading
  - HMC memory with limitations :
    - Access must be multiple of 64 bytes
    - Only 1 lane (out of 4)
    - Bandwidth below expectation (18 GBps)



HMC bandwidth (1st tests)



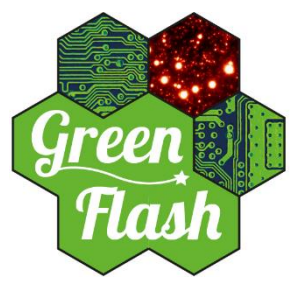
— Write — Read

➔ Already working on limitation resolution



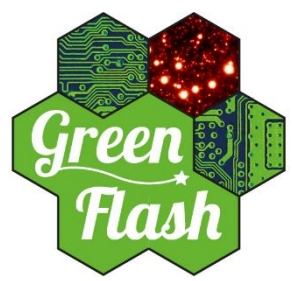
# WP5 – Smart Interconnect

- Microgate uXComp board support
- Peer-to-Peer solution for reduced CPU interaction and improved BW
- Improved performances (figures obtained for GF application)
  - TCP (8 Gbps)
  - UDP (9,5 Gbps)
  - PCIe (35 Gbps)
  - C Kernels for GVSP images encoding/decoding (up to 10 Gbps)
- Internal definition :
  - Logic addition for enhanced modes of operation and testability (more loopback modes)
  - Improved C kernels for dynamic images/matrix handling



# WP5 – Protocols support

- D5.11 : Support of UDP, Infiniband and RTPS
  - UDP fully supported under QuickPlay
    - UDP offload engine IP Core
    - UDP stream support by SDK)
  - Infiniband :
    - Not a requirement for AO.
    - Not supported under QuickPlay. Long and expensive developments
    - Could be supported by existing COTS product such as Mellanox Smart NIC solutions (Innova 2 or Bluefield)
  - RTPS :
    - Initial requirement of the project.
    - Not reported in ESO requirements
    - Could also be supported by existing COTS



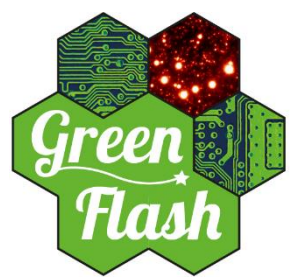
# WP5 – Middleware

- D5.6 and D5.9 outcomes:
  - Standard NIC adaptation to Smart Interconnect : out of scope considering manpower and skills in the project
  - OpenMPI : deeply relies on standard NIC architecture. The abstraction level of a custom transport layer is not adequate with manpower & skills
  - OpenDDS : Custom UDP transport layer concept has high level of abstraction, allowing support of QuickNET
  - QuickPlay SDK enhancements (already mentioned in slide 11)



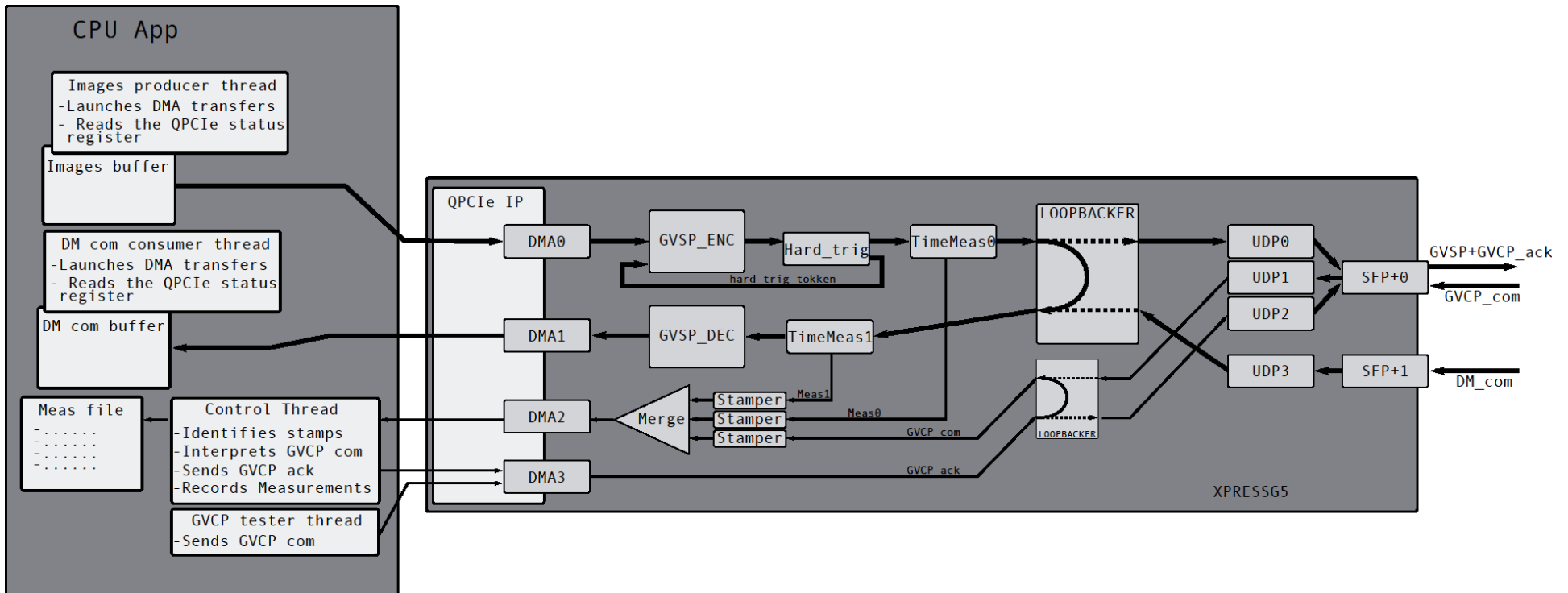
# WP5 – Applications





# WP5 – OBSPM usage of SI

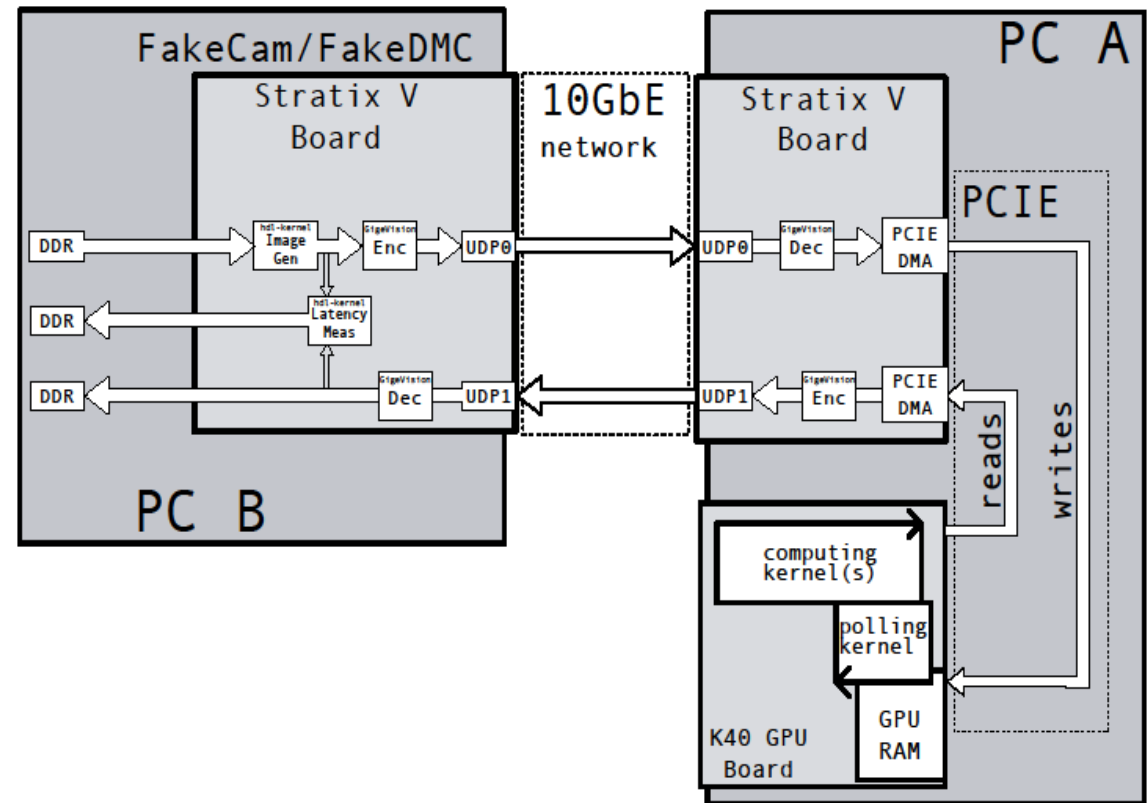
- Fake CAM / DMC prototype

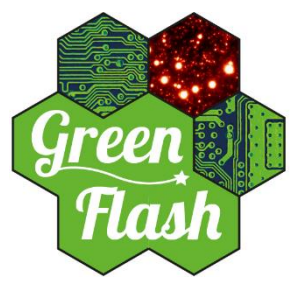




# WP5 – OBSPM usage of SI

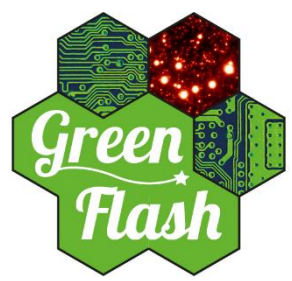
- RTC prototype
  - Fake CAM / DMC
  - Image acquisition
  - GPU interfacing
  - DM matrix computation





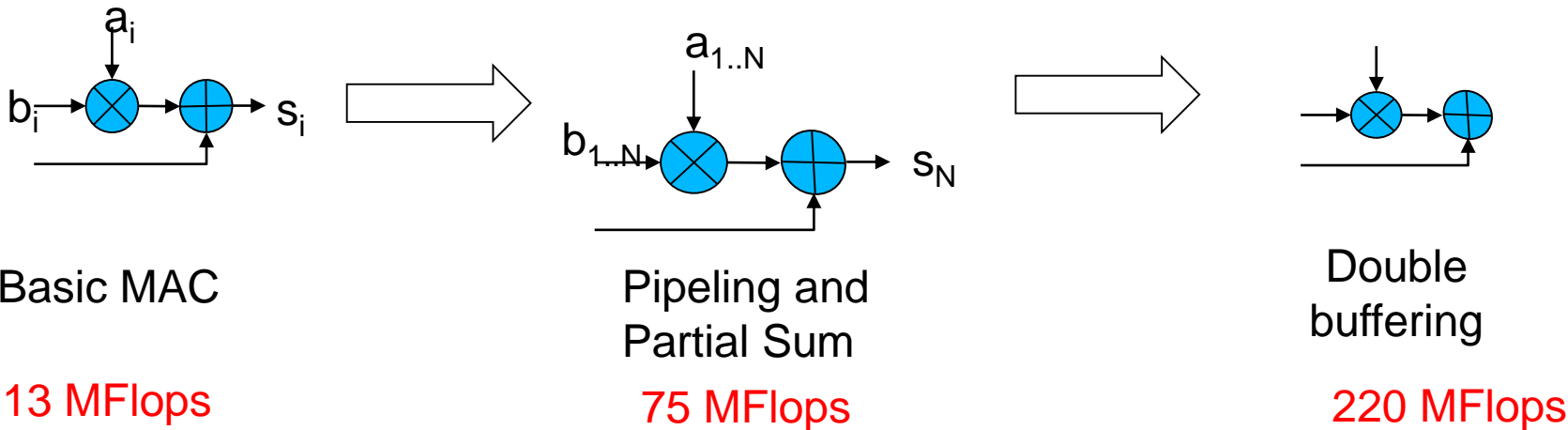
# WP5 – Adaptative Optics (1/4)

- Objectives
  - Investigate FPGA boards computation capability
  - Validate QuickPlay capability to explore various solutions
  - Target : MVM on  $\mu$ XComp board using HMC memory (Intermediate results presented were obtained on Arria10 boards with DDR4 memory)
- Approach :
  - Step by step architecture, from scalar product to MVM distributed over 8 kernels



# WP5 – Adaptative Optics (2/4)

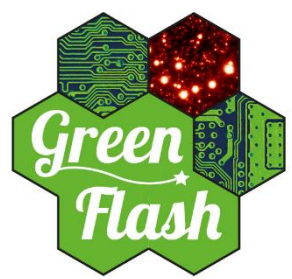
Scalar Product :  $S = \sum_{i=1, N} \text{step 1} (a_i b_i)$



## Matrix Vector Product:

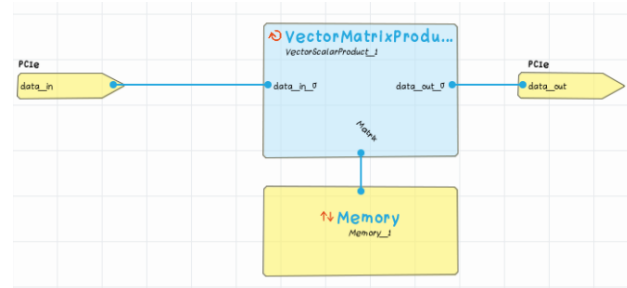
- \* Vector send through stream
- \* Matrix stored in external memory

@ Each clock cycle, memory can exchange 2x64 bits, ie 4 floats => 4 MAC at the same time



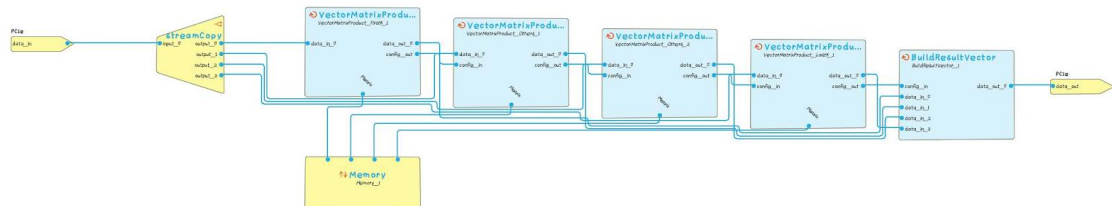
# WP5 – Adaptative Optics (3/4)

MVM  
 1 kernel with  
 4 MAC  
 822 MFlops



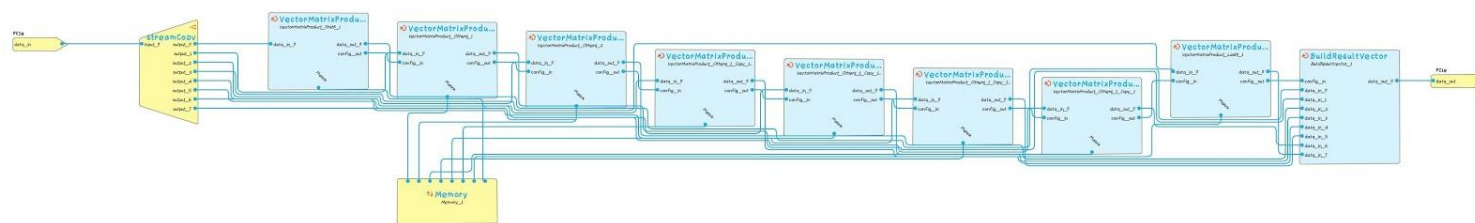
MVM  
 4 kernels with  
 4 MAC each  
 3,2 GFlops

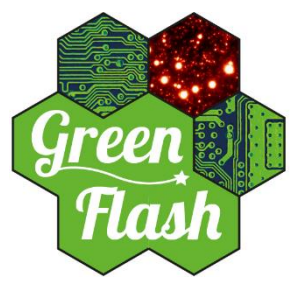
Project: MatrixVectorProduct\_2024\_01\_14\_app  
 Board: Reflex - socrates



MVM  
 8 kernels with  
 4 MAC each  
 3,9 GFlops

Project: MatrixVectorProduct\_2024\_01\_14\_app  
 Board: Strix - ADP4





# WP5 – Adaptative Optics (4/4)

- Figures interpretation :
  - Kernel multiplication should allow proportional computation BW, in the limit of memory BW
  - Could not reach DDR4 memory BW limit (6 GFlops) with 8 kernels. C Kernel stall policy could be the root cause :
    - DDR access by 1 kernel is delayed
    - It causes its streams RD/WR and computations operations to be stalled,
    - then in cascade slowing all the computation
  - Memory BW / Latency is another limiting factor :
    - Arria10 device used should have enough resources for 32 kernels (8 kernels occupy around 20 % of resources)
    - HMC memory is a good candidate to support this computation bandwidth
      - 144 GBps announced over 4 lanes
      - First MVM tests targeting HMC are inline with DDR4 results



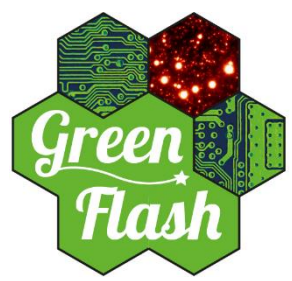
# WP5 – Future



# WP5 – Next developments

- D 5.4 : Support of Arria10 SOC devices
  - SDK compiled for ARM
  - Root Port design to host QuickPlay designs
  - Boards support : Altera Arria10 SOC DevKit + Microgate  $\mu$ XLink
- D 5.5 : Smart Interconnect performances report
  - Final report on Microgate  $\mu$ XComp board
  - Solving last limitations (HMC, P2P)
- D 5.8 : IPs implementing Adaptative Optics algorithms
  - Targeting  $\mu$ XComp boards with HMC memory (4 lanes)
  - FPGA Cluster tentative ?





# WP5 – Deliverables Status

ID	Content	Delivery
D5.1	Smart interconnect prototype 1	Delivered
D5.2	Smart interconnect prototype 2	<b>Delivered</b>
D5.3	Smart interconnect prototype 3	<b>Delivered</b>
D5.4	Smart interconnect prototype 4	July 2018
D5.5	Smart interconnect performance report	Sept. 2018 (1)
D5.6	Smart features to middleware test report	<b>Delivered</b>
D5.7	Prototype Boards Support Package	Delivered (2)
D5.8	IPs implementing AO control algorithms	June 2018
D5.9	System level API primitives	<b>Delivered</b>
D5.10	Scalability of QuickPlay designs	<b>Delivered</b>
D5.11	Support for UDP, Infiniband and RTPS	UDP : Done

(1) : Final report. Intermediate report comes with each delivery

(2) : New boards support when available



# WP5 – Discussion

