Green Flash

High performance computing for real-time science

WP5 (Smart Interconnect) Status
PLDA / Accelize

Mid-Term Review – Meudon Feb. 2017

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

Agenda

- Work Package WP5 Highlights
- State of the art
  - Status
  - Findings
  - Issues & Concerns
- Overall planning for the remaining activities
- Discussion

WP5 - Definition

- 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"

- Tasks identification
  - Task 6.1 : High Bandwidth FPGA NIC
  - Task 6.2 : Smart Features to Middleware
  - Task 6.3 : Development environment and IPs

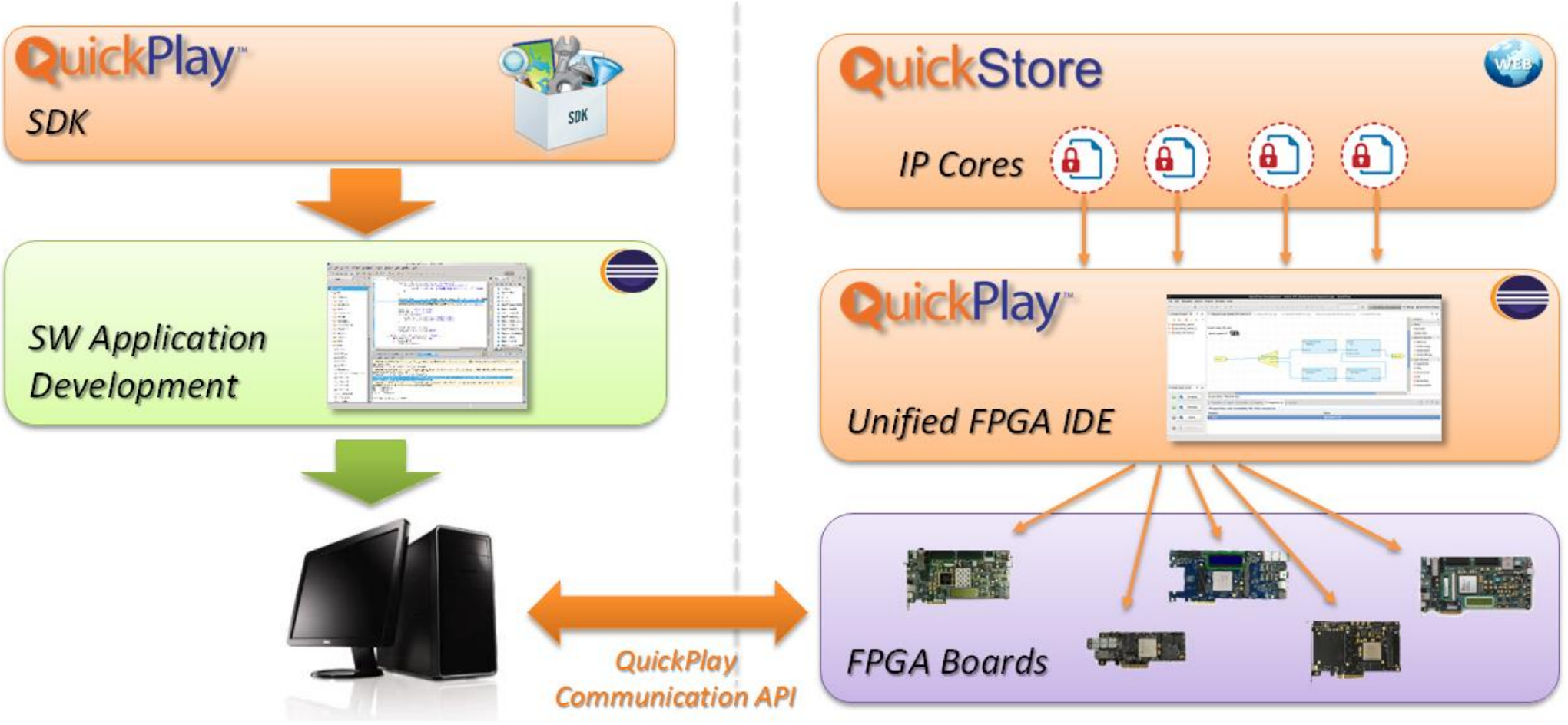
➔ **Task 6.3 is actually driving the WP development**

WP5 – Tasks vs Deliverables

ID	Content	6.1	6.2	6.3
D5.1	Smart interconnect prototype 1	X		X
D5.2	Smart interconnect prototype 2	X		X
D5.3	Smart interconnect prototype 3	X		X
D5.4	Smart interconnect prototype 4	X		X
D5.5	Smart interconnect performance report	X	X	X
D5.6	Smart features to middleware test report		X	
D5.7	Prototype Boards Support Package	X		X
D5.8	IPs implementing AO control algorithms		X	X
D5.9	System level API primitives	X	X	X
D5.10	Scalability of QuickPlay designs	X	X	X
D5.11	Support for UDP, Infiniband and RTPS	X		X

Tasks achievement : materialized by Smart Interconnect deliveries

WP5 – Task 6.3 – Environment (1/3)



WP5 – Task 6.3 – Environment (2/3)

- Improved Tool maturity (independently from GreenFlash)
  - Improved IP and Boards development flow
    - Eases & fastens COTS IPs and boards integration
    - Now open to 3rd Party providers (Bittware, CAST, …)
  - High Standard development methodology (continuous delivery, automated validation and analysis capability)
- Improved Performances & Features
  - Improved proprietary HLS through directives usage, allowing higher computing performances. (Eg : x10 Gbps for GeV image and DM commands processing)
  - Xilinx Vivado HLS integration

WP5 – Task 6.3 – Environment (3/3)

- Tool accessibility :
  - Ease of use (see Durham University testimonial)

- Coming soon :
  - Multi-board support by SDK (then addressing D5.10) **– Q2 2017**
  - Increased emulation capability for C and HDL kernels (static and dynamic properties made available at C level) **– Q2 2017**

WP5 – Task 6.3 – Boards

- BSP delivery for several boards (D5.7)
  - Reflex XpressGX5 (Altera Stratix V) + XpressKUS (Xilinx Kintex Ultra Scale) : Delivered
  - MicroGate µXComp (Altera Arria10) + µXLink (Altera Arria10 SOC) support delayed, pending upon delivery of the boards
  - Backup solution for Arria10 devices: Reflex XpressGXA10
    - All useful interface IPs integrated (PCIe, Ethernet, TCP/UDP, DDR4)
    - Currently fixing FPGA timing issues

➔ **Increased Scalability / Genericity through several COTS or custom boards enabled by usage of QuickPlay**

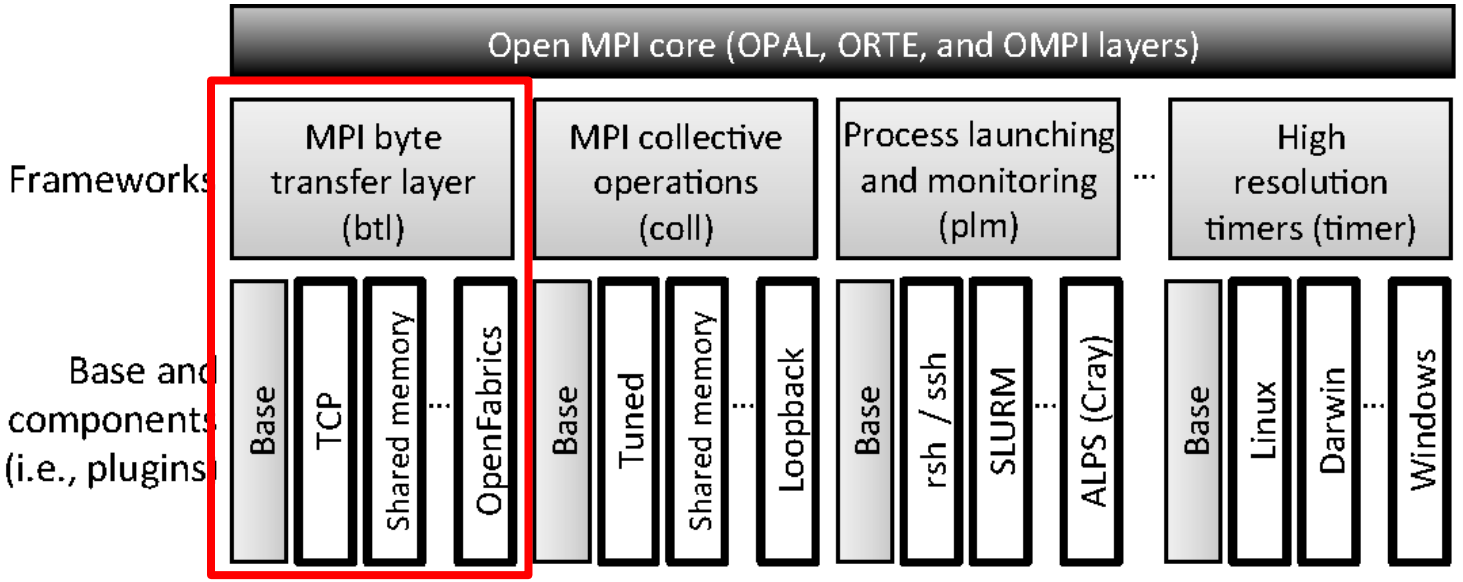
WP5 – Task 6.3 – IPs

- QuickPlay IPs integration (D5.11)
  - UDP, UDP Multicast, DDR4, PCIe Peer-to-Peer
  - Enriched with C-Kernels collection :
    - GVSP (GeV Image) and CSKT (matrix exchange) CODECs
  - Coming next :
    - HMC (Microgate µXComp board), PTP1588 **– Q2 2017**
    - Candidates : Infiniband, Camera-Link, 40G

➔ **Increased IP Portfolio, with high perf. interfaces**

WP5 – Task 6.2 – Middleware (1/2)

- OBSPM studied support of Open MPI on top of FPGA NIC

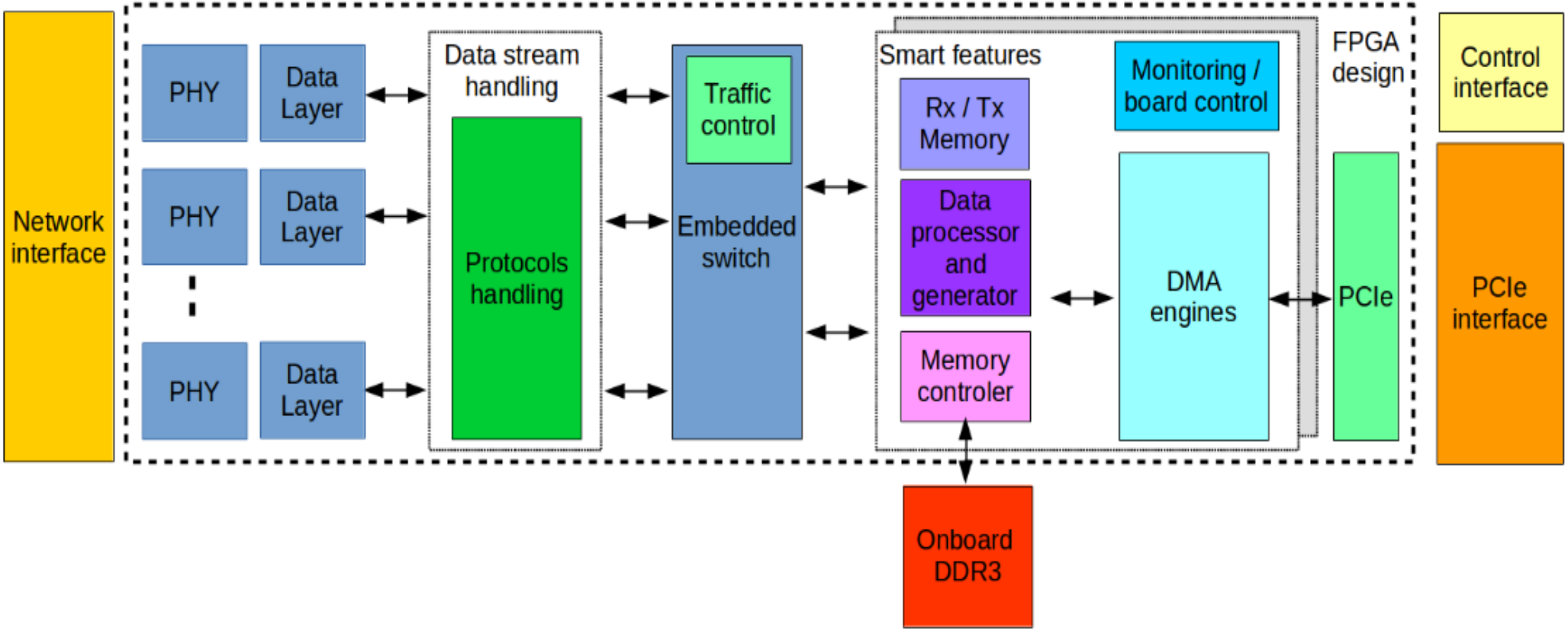


- btl as preferred framework
- Requires development of MPI components

WP5 – Task 6.2 – Middleware (2/2)

- Like earlier presented in Periodic Report, study did conclude on difficulty to interface openMPI with FPGA network card.
- Difficulties reside in high level of expertise and significant manpower required in order to :
  - Develop a standard NIC driver for FPGA-based network card (TCP and UDP being off-loaded). Neither OBSPM nor PLDA have sufficient expertise to develop and even correctly specify
  - Map open MPI on top of such proprietary driver (poor documentation)
- Considering DDS support currently. More confident on being able to get a custom support layer before the end of the project

WP5 – Task 6.1: FPGA NIC (1/3)



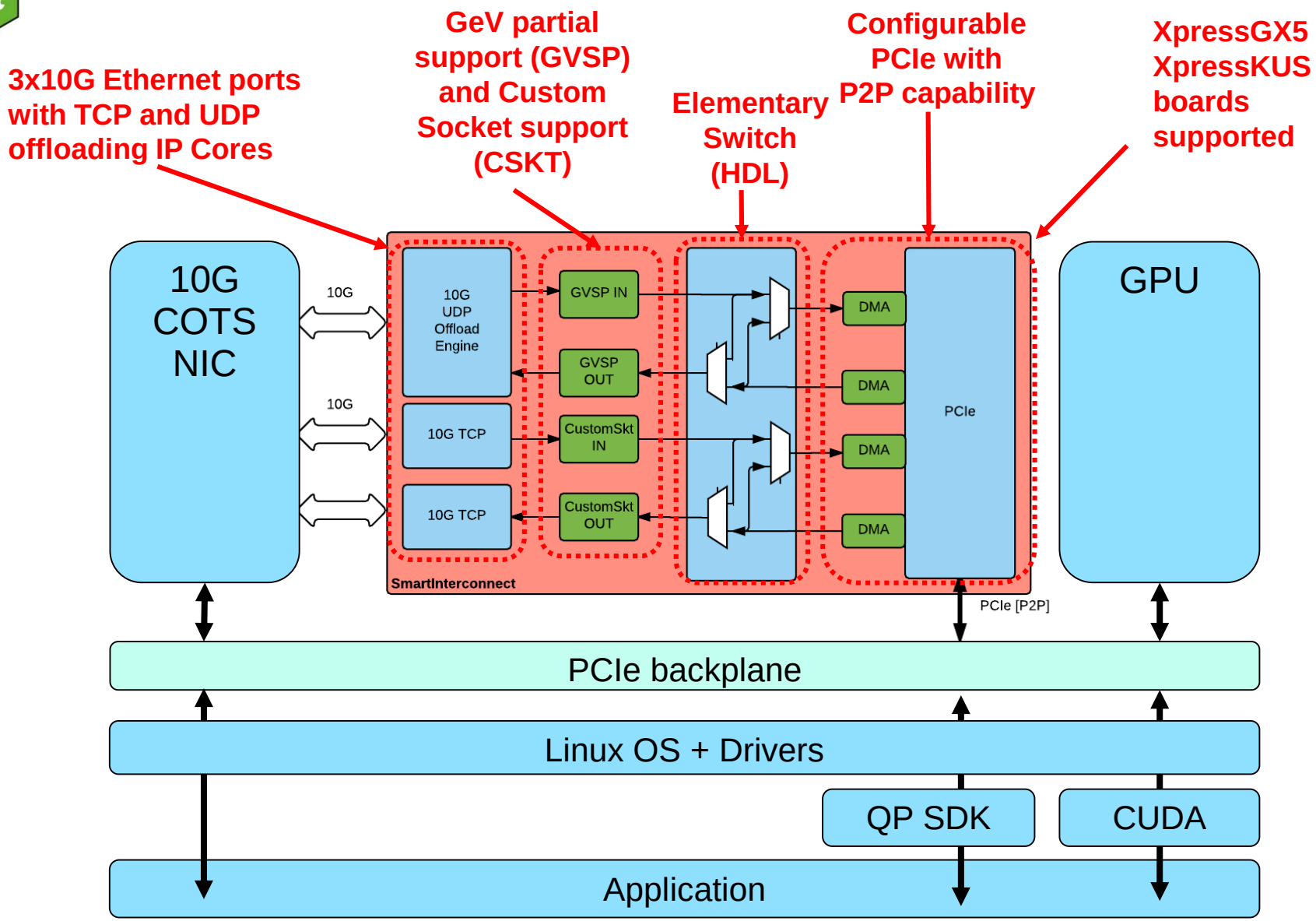
WP5 – Task 6.1: FPGA NIC (2/3)

- Did benefit from the use of QuickPlay development environment (see task 6.3) :
  - BSP delivery for multiple boards (D5.7), demonstrating flexibility allowed by QuickPlay
  - IPs integration (D5.11) : High bandwidth support through TCP/UDP, DDR4, PCIe and HMC
  - NIC made “Smart” through
    - Modular, parallel and scalable architecture
    - Ease of effective implementation and validation of various algorithms (in QuickPlay Kernels)

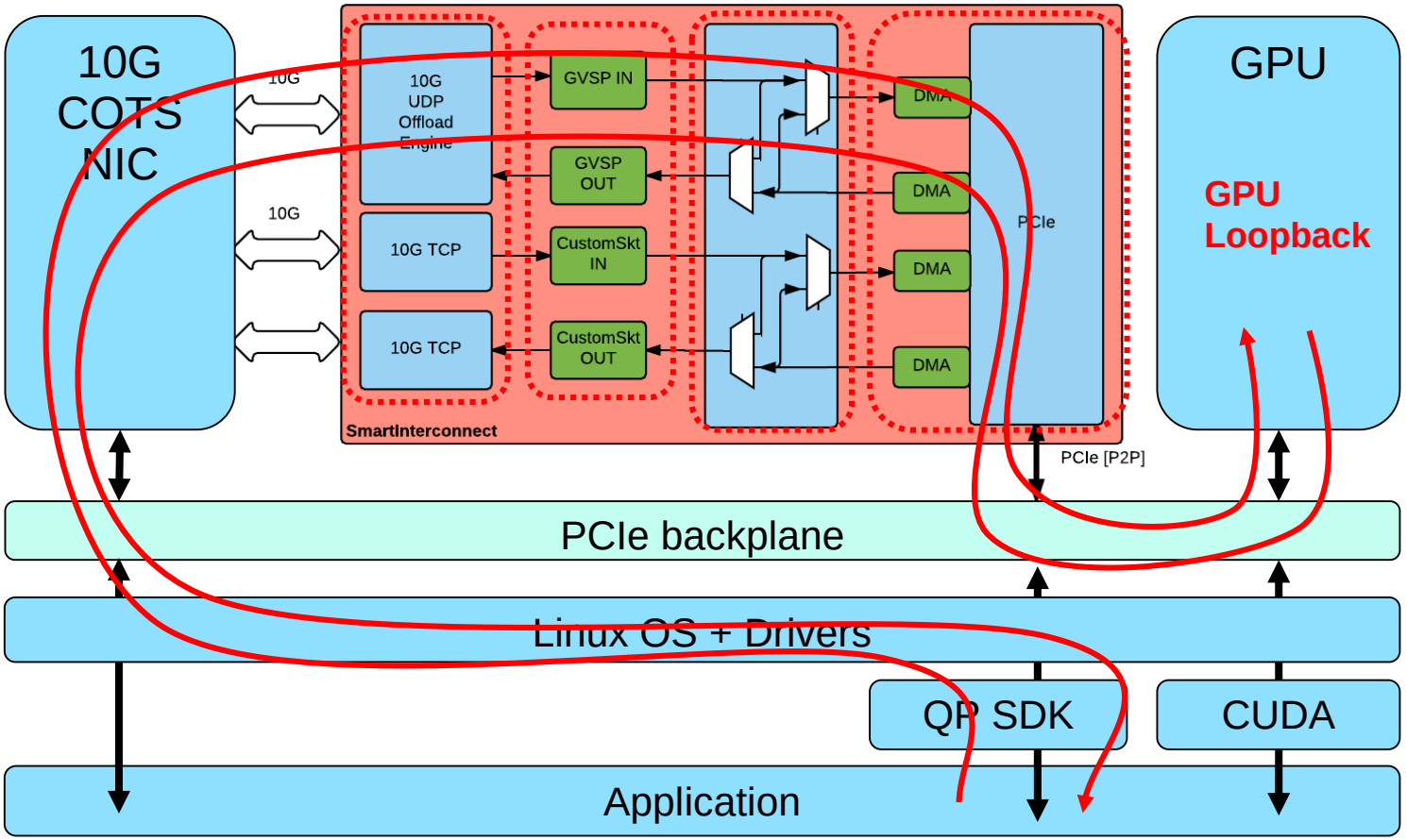
WP5 – Task 6.1: FPGA NIC (3/3)

- Qualified through Smart Interconnect prototypes deliveries (D5.1)
  - Board “agnostic” : No architecture/performance difference depending on targeted board
  - Network Performances :
    - IP Core : 9 Gbps TCP and UDP (see D5.1 report)
    - Limited by SW layers (around 1 Gbps)
    - Latency / Jitter consistent with GF requirements (see Denis Perret work)

WP5 – Smart Interconnect Proto #1



WP5 – Smart Interconnect Proto #1



GeV or CSKT Loopback

16

WP5 – Smart Interconnect Findings

- Reduced dependency to BSP support : FPGA technology might impact Kernel performances (frequency)

- Performances (see D5.1 test report)
  - GeV Image Encoding/Decoding capability **> 10 Gbps**
  - Matrix Emission/Reception (CSKT) at **up to 20 Gbps**
  - TCP/UDP : Suffer from SW limitations (API or test application ?). We did limit to **1 Gbps**
  - PCIe : Suffers from SW limitations (API) ➔ **5.5 Gbps**. New SDK solves this issue (to be tested with Smart Interconnect)
  - P2P : Not fully characterized. Limited set of features, with required host application to configure & start FPGAs DMA engines

➔ QuickPlay environment flexibility and performances have been demonstrated

WP5 – Conclusions

- QuickPlay FPGA development environment conforms with most of the requirements :
  - Scalability / Genericity
  - Standard protocols compliance
  - Performance (latency / data processing)
- …with some improvements to be done on :
  - API with interface IPs (PCIe & TCP/UDP)
  - P2P capability
- But did fail regarding Middleware compliance (NIC drivers / Open MPI)

WP5 – Actions

- Actions :
  - Solve performance issues affecting PCIe and TCP/UDP (SW / API)
  - Solve FPGA timing issues
  - Integrate µXComp board (including HMC)
  - Complete P2P support (FPGA to FPGA, generic P2P)
  - Integrate µXLink board (including SOC)
  - Complete IP integration plan

WP5 – Plans

- Smart Interconnect Proto deliverables redefinition
  - Proto 2 : µXComp support + Arria10 + DDR4 + Perfs improvement
  - Proto 3 : HMC + Generic Peer-to-Peer
  - Proto 4 : µXLink support + SOC support + PCIe Root Port
  - Proto 5 : PTP 1588 + 1G support
- See next slide

WP5 – Planning

ID	Content	Delivery
D5.1	Smart interconnect prototype 1	Done
D5.2	Smart interconnect prototype 2	31/03/2017
D5.3	Smart interconnect prototype 3	15/05/2017
D5.4	Smart interconnect prototype 4	29/09/2017
D5.12	**Smart interconnect prototype 5**	**?**
D5.5	Smart interconnect performance report	Q4 2017 (1)
D5.6	Smart features to middleware test report	Q2 2017
D5.7	Prototype Boards Support Package	Done (2)
D5.8	IPs implementing AO control algorithms	Q4 2017
D5.9	System level API primitives	No need identified
D5.10	Scalability of QuickPlay designs	Q2 2017
D5.11	Support for UDP, Infiniband and RTPS	UDP : Done. Others : ?

(1) : Final report. Intermediate report comes with each delivery
(2) : New boards support when available

WP5 – Discussion