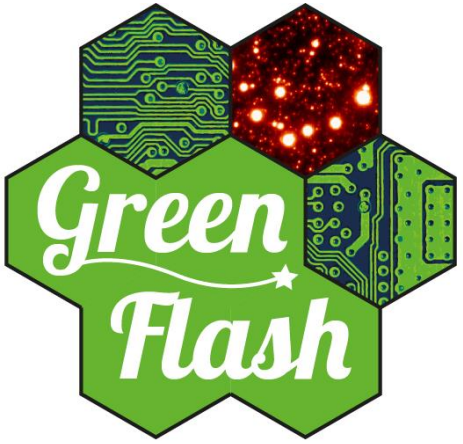


Green Flash

High performance computing for real-time science

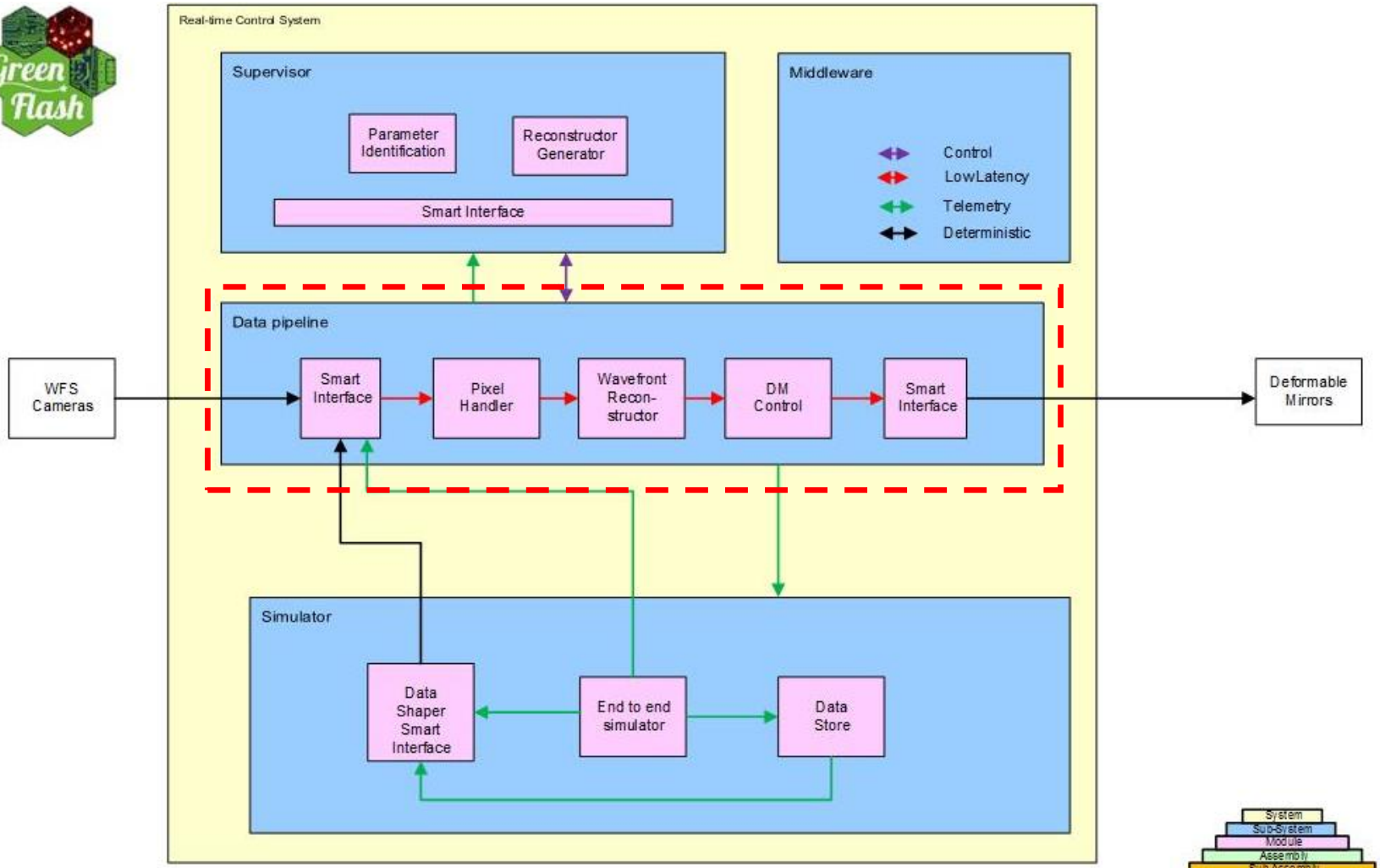
MTR 01/02/2017 Paris

WP3 – FPGA solutions for Hard Real Time

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

Work package 3: FPGA solutions for hard real-time

Microgate is responsible for WP3 (FPGA solutions for Hard Real Time)
In this frame we will develop a solution for the **hard real-time data pipeline** with **FPGA** based boards for realizing a **stackable, energy efficient microserver** for data-intensive applications.



Work package 3: FPGA solutions for hard real-time

This **WP 3** under the responsibility of **Microgate** provides a concept study based on FPGA boards for the stackable, energy efficient stand-alone microserver for data-intensive applications. It involves the prototyping of one main board with an SoC FPGA containing a hard-wired ARM processor and several interface and the production of several FPGA based computational boards to be clustered in a microserver. The performance in terms of communication bandwidth and computation throughput will be assessed for the AO application on a single board and on the small scale cluster.

Tasks description

Task 3.1: Prototype microserver based on FPGA

Develop a prototype microserver based on FPGAs with ARM HPS. This activity aims to develop a prototype of a high throughput computational board and a communication board dedicated to hard real-time control, data pre/post processing and communication with acquisition and control units.

Task 3.2: Extension of microserver with PCIe bus

Extend the prototype of microserver board to include PCIe root complex

Task 3.3: Firmware implementation

Develop IPs + drivers for the microserver boards

WP 3: Project deliverables

Deliverable	Due date	Internal/public
FPGA microserver design report	T_0+12Mo	int
FPGA microserver HW test report	T_0+12Mo	int
FPGA microserver prototypes	T_0+12Mo	Int
FPGA microserver including PCIe interface design report	T_0+24Mo	int
FPGA microserver including PCIe interface HW test report	T_0+24Mo	int
FPGA microserver including PCIe interface prototypes	T_0+24Mo	int
FPGA microserver firmware design	T_0+36Mo	int
Cluster HW test report	T_0+36Mo	pub

WP 3: Summary of requirements

- Stand-alone system using SoC FPGA-CPU still preserving compatibility with servers
- Capability to interface to different accelerator cards (FPGA based, GPU based, CPU based)
- Modularity to adapt to the Real Time Reconstructor requirements of different AO instruments
- Providing different interfaces to wavefront cameras and deformable mirrors
- High Computational throughput
- Low latency and jitter
- Energy efficient
- PLDA QuickPlay compatible

Instrument	*Frame Rate (Hz)*	*Latency (ms)*	*Jitter (us)*	*Pixel Rate (Gb/s)*	*TMAC/s for MVM*
HARMONI	800	2.5	125	139.3	0.2
MAORY/ MICADO	500	4	200	138.2	0.46
METIS SCAO	1000	2	100	10.2	2.2 GMAC/s
METIS LTAO	1000	2	100	194.6	0.37
MOSIAC	250	8	400	74.2	1.4
HIRES	500	4	200	138.2	0.28
EPICS	3000	0.667	33	122.9	11.05

Required Goal	**Required**	**Goal**
R9.10 Frame Rate (Hz)	800	1000
R9.20 Pixel Rate (Gb/s)	200	250
R9.30 TMAC/s	1.5	3
R9.40 Max. Latency (ms)	2	1
R9.50 Max. Jitter (us)	100	100

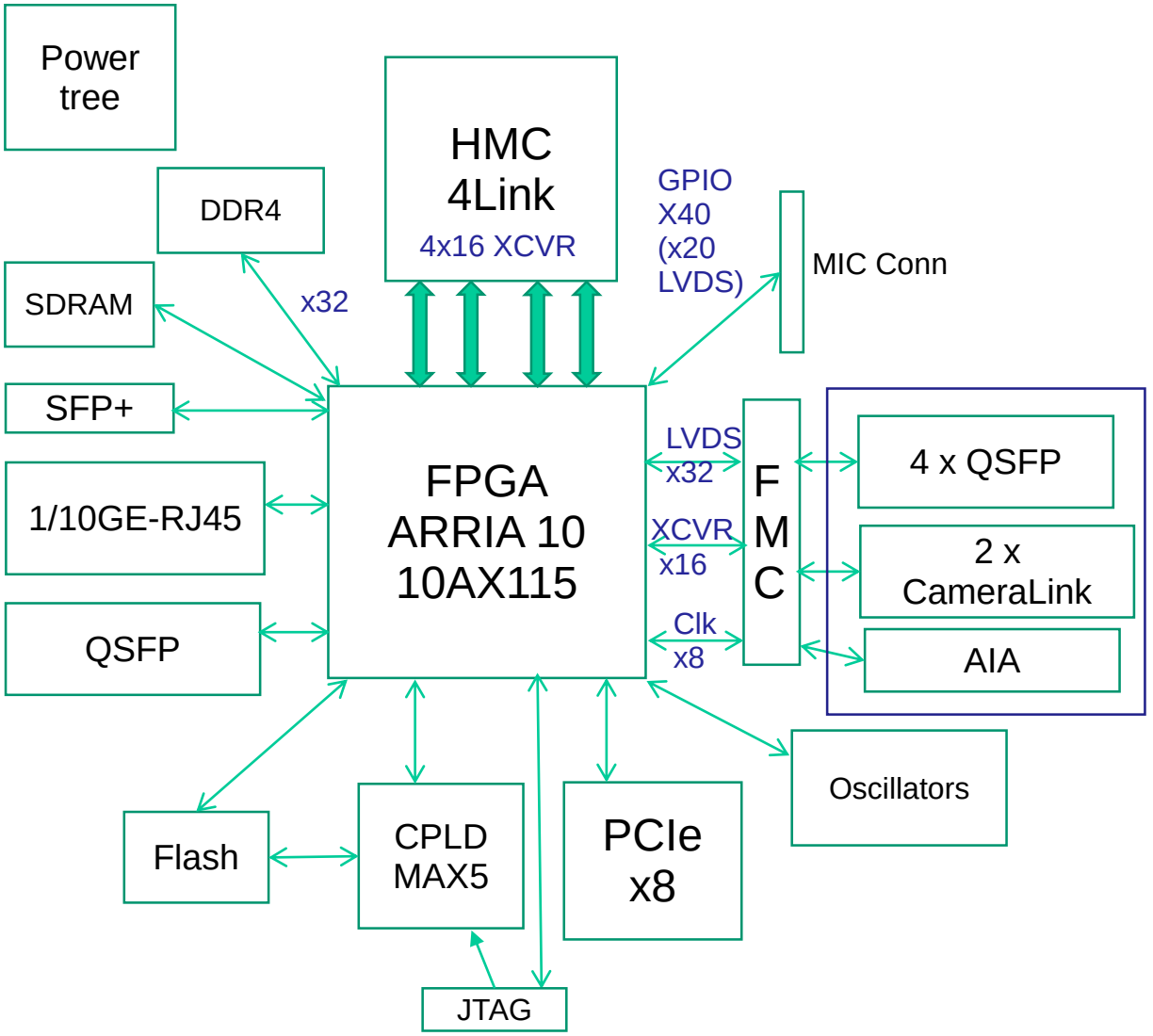
WP 3: Solution

To realize the stand-alone microserver and fulfill the requirements we will develop 2 different boards based on FPGAs.
One board called **µXComp** that act as a computational board, which can perform the **real-time computation** in a **deterministic** way with low latency, low jitter and high energy efficiency.
The second board called **µXLink** that contains an FPGA with an hard-wired ARM processor in the same chip (SoC) and is used as an **interface and control board** to connect to serval computational boards and to WFSs and DMs.
The hardware and the firmware will be designed to be compatible with PLDA QuickPlay

The highlight of these boards:

- Both based on Altera ARRIA 10 FPGAs – Newest Altera FPGA Chip on the market
- High number of hard-wired DSPs and Transceivers
- Each DSP can perform a full single precision (32-bit) floating-point MAC
- The large number of transceivers allows to realize a large number of different interfaces e.g. 10G Ethernet, Infiniband …
- Backplane communication interface based on PCIe x8 up to Gen3
- Include a novel external memory chip HMC (Hybrid Memory Cube) – fast DRAM memories stacked vertically using true-silicon-via combined with up to 64 high-speed transceiver serial links (up to 120GB/s each direction)
- Low power consumption

WP3: Computational board µXComp



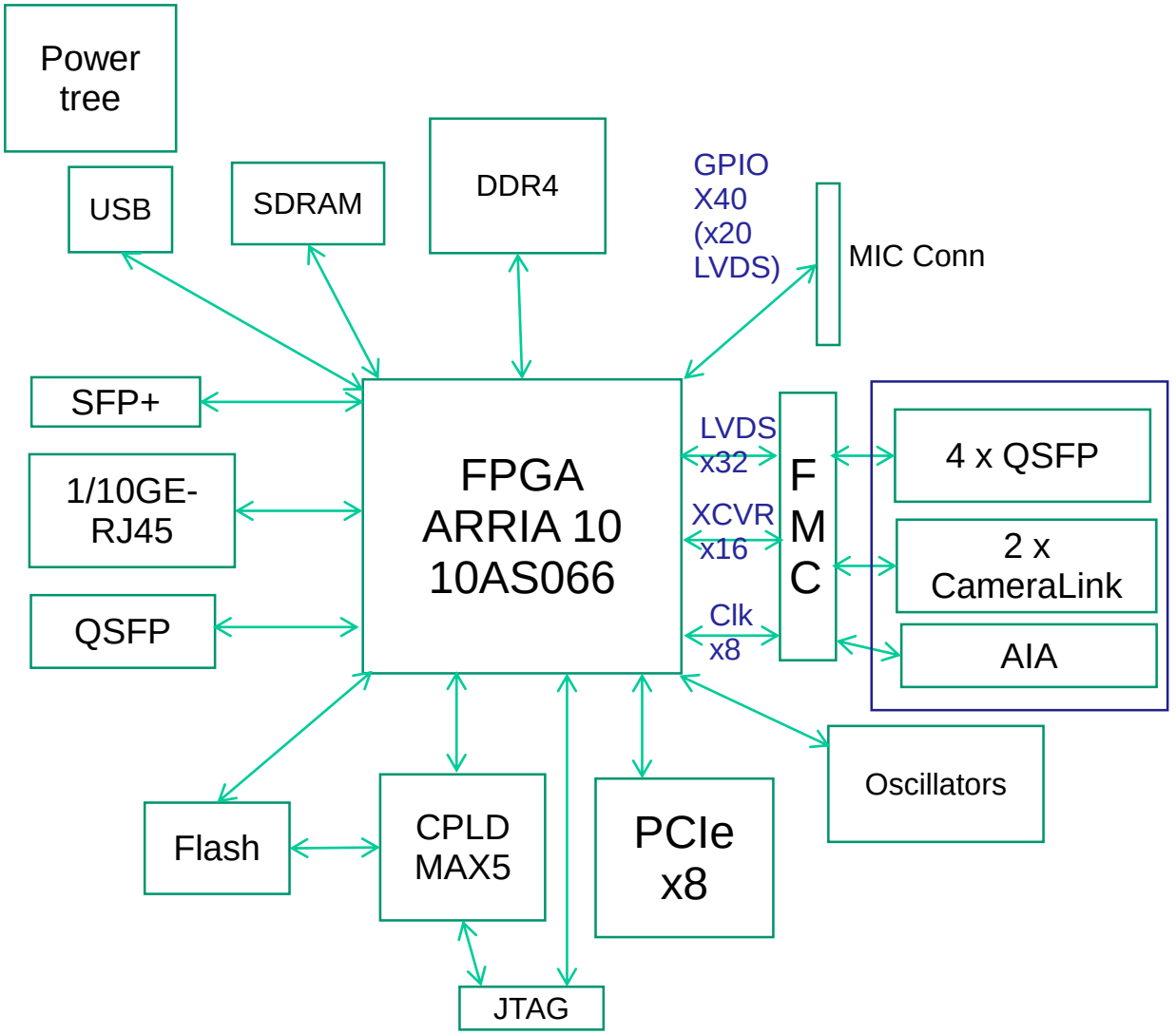
Based on ARRIA 10AX115:

- 1518 DSP blocks
- 6.6MB int. RAM
- 96 XCVR

Board features:

- Optimized for **heavy deterministic computation** in floating-point
- **Large Bandwidth between HMC and FPGA** - 4 links 16 lanes/link up to 15Gbps/lane = 120GB/s bidirectional
- Extremely **low jitter**
- More **power efficient** compared to GPUs
- Offers a lot of different interfaces on board or via the FMC connector and extension cards

WP3: Microserver/interface board µXLink



Based on
ARRIA 10AS066 SoC:

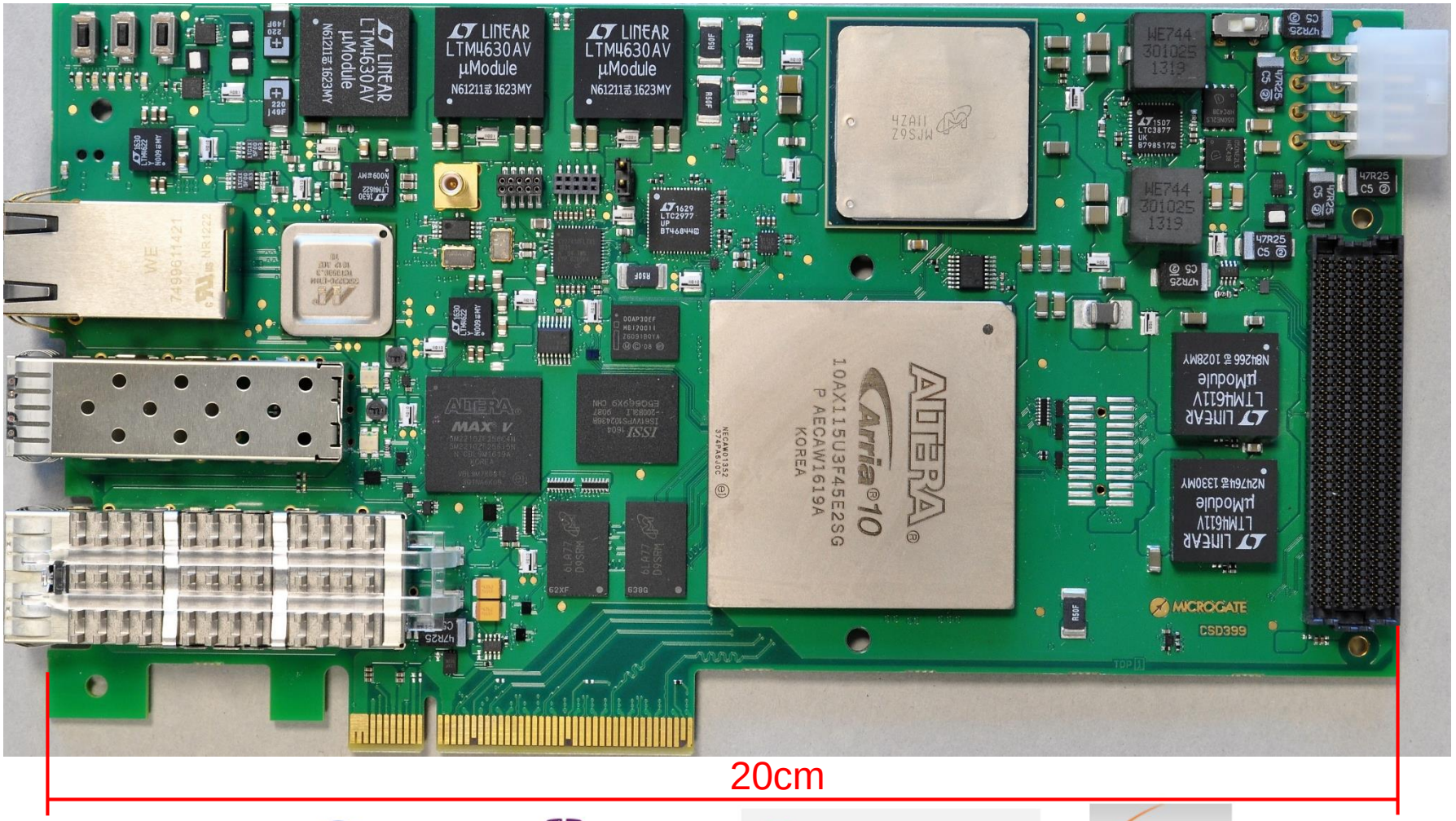
- 1.5GHz ARM dual-core Cortex-A9 on-chip processor
- 1855 DSP blocks
- 5.2MB int. RAM
- max. 48 XCVR

Board features:

- ARM **embedded processor** for **stand-alone** real-time box
- Powerful **PCIe root port** because of ARM and OS
- Management of accelerator cards on the PCIe interface
- Running control software using a full OS (e.g. Linux)
- Easy implementation of different communication protocols
- Offers a lot of different interfaces on board or via the FMC connector and extension cards

WP3: Current status

The first prototype of the two FPGA boards, the µXComp board is manufactured and is currently under test. After the validation of the interfaces and the communication between FPGA and HMC some more boards of this type will be produced and made available to the team.



WP3: Current status

Features:

- PCIe up to x8 Gen 3 (64Gb/s each direction)
- Front-panel interfaces: 1/10Gb Ethernet via fiber (SFP+) and via copper (RJ45), 40Gb Ethernet or Infiniband (QSFP)
- FMC standard extension boards attachable on the back (up to length 130mm to fit in a PCIe full length slot (320mm)
- Sustained performance 30 GMAC/s (single precision floating-point)
- Memory transfer rate 120GByte/s each direction
- Compatible with QuickPlay tool requirement

Board facts:

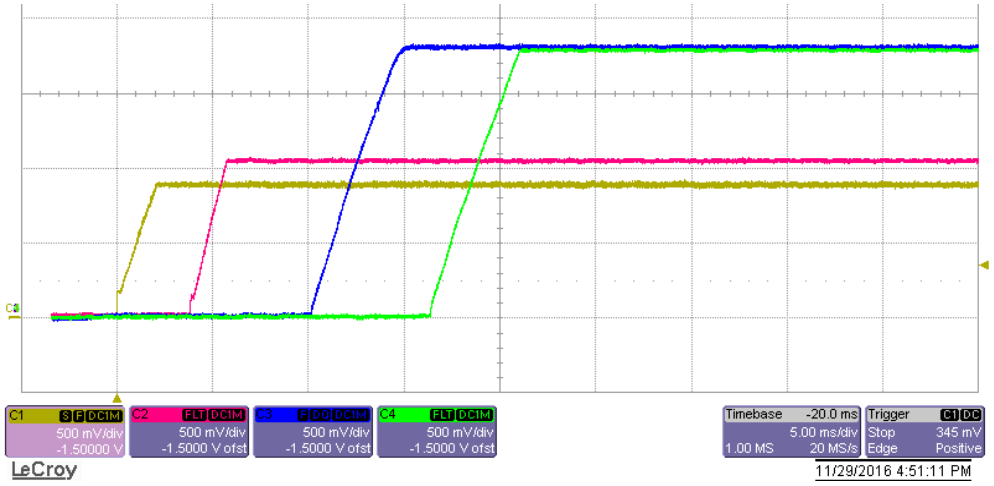
- Board size (111x200 mm) compliant with PCIe standard single slots full height and >= ¾ length
- # Layers: 18 (9 signal, 9 power-ground)
- # Components: 1442
- # Tracks: 71388 (300 LVDS pairs)
- # Vias: 12305

Tests completed:

- Voltages measured of all Power rails
- Current driving capability for each rail measured (up to 100W total power consumption tested)
- Power-up and Power-down sequence programmed and measured
- External PLL chip programmed and clk outputs measured
- Max5 CPLD logic development started for housekeeping
  - SPI interface implemented for ADC to read out voltages and currents of the different power rails.
  - I2C interface implemented to read out the temperature sensors and set the thresholds
  - Flash interface implemented to access the flash memory via JTAG.
  - JTAG switches tested adding the HMC JTAG.
- PCIe interface x8 Gen 2 implemented in new ARRIA 10 and tested -> works
- MVM calculation 222X5316 floating-points implemented for DP control -> works good with execution time of <30µs (only internal memory)

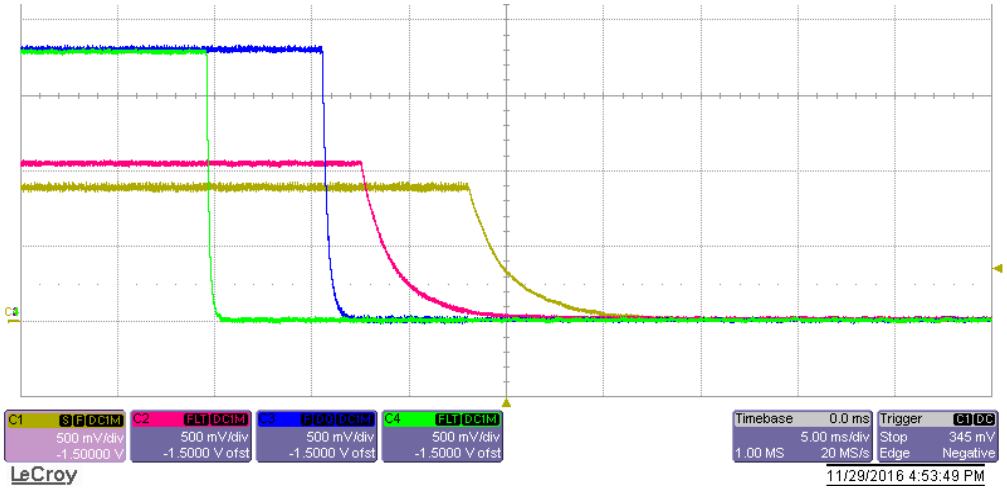
WP3: Current status

Power-up Sequence



Ch1: EN_A10_GROUP1 (U3) A10_VCC_0.9V
Ch3: EN_A10_1.8V (U50) A10_VCCPT_1.8V

Power-down Sequence



Ch2: EN_A10_GROUP2 (U34) 10_VCCRT_GBX_1.03V
Ch4: EN_A10_VCCIO (U39) A10_VCCIO_1.8V

Tests still to be performed:

- HMC interface testing up to 4 links with 16 Transceiver lanes per link
- DDR4 memory testing
- 10G Ethernet
- Real-Time Interface to AO mirrors

After these tests are done and show the main features of the board working we will launch the production of other 2 µXComp boards – one for PLDA and one for Observatoire de Paris

WP3: Criticalities (issues)

Issues:

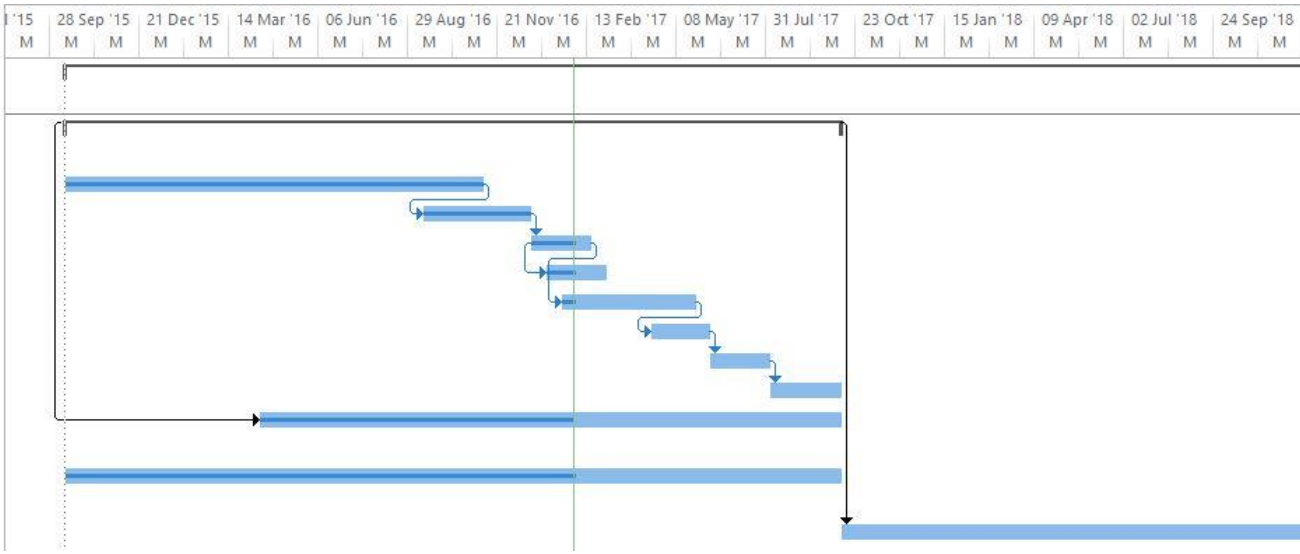
1. (FPGA and Power supply chips marked introduction delayed significantly from original timeline)
2. Delay of the µXComp board accumulated because of a problem at the PCB manufactured that created bubbles during assembly
3. Due to 2. delay also in the delivery of the hardware test reports

Solutions:

1. (Enhanced simulation activity, pressure to the suppliers and development on evaluation boards)
2. PCBs re-manufactured: µXComp board will undergo a targeted test campaign to kick out the manufacturing of two more boards as soon as possible and further tests will be carried out later. In addition expected reduction of development time of the second kind of board the µXLink because of lessons learned with the current board.

WP3: Schedule summary

Task Name	Duration	Start	Finish	Predecessor
WP3 - FPGA for Hard-RT	**910 days**	**Tue 13/10/15**	**Mon 08/04/19**	
FPGA-based microserver & PCIe extension to microserver	**520 days**	**Tue 13/10/15**	**Mon 09/10/17**	
µXComp board development	280 days	Tue 13/10/15	Mon 07/11/16	
µXComp board prototyping	73 days	Tue 13/09/16	Thu 22/12/16	3FS-40 days
µXComp board testing	40 days	Fri 23/12/16	Thu 16/02/17	4
Test reports	40 days	Fri 06/01/17	Thu 02/03/17	5SS+10 days
µXLink board development	90 days	Fri 20/01/17	Thu 25/05/17	5FS-20 days
µXLink board prototyping	40 days	Fri 14/04/17	Thu 08/06/17	7FS-30 days
µXLink board testing	40 days	Fri 09/06/17	Thu 03/08/17	8
Microserver integration	47 days	Fri 04/08/17	Mon 09/10/17	9
PCIe extension to microserver	390 days	Tue 12/04/16	Mon 09/10/17	2SS+130 days
Firmware implementation	520 days	Tue 13/10/15	Mon 09/10/17	
Clustering	390 days	Tue 10/10/17	Mon 08/04/19	2



WP3: Conclusions

The first prototype of the µXComp is produced by Microgate and is one of two types of FPGA boards to realize the hard real-time data pipeline in a microserver. The board is currently under test expecting the completion of the tests in the next weeks.
It is a complex board with 18 layers and a high number of components.
For the design of the second µXLink board already a preliminary design is started and the final design will start in March. A lot of knowledge gained with the µXComp can be reused for the design of the µXLink that will reduce the development time of the second board significantly. The first prototype of µXLink is expected to be available in summer.
The assembly of a Microserver containing one µXLink and one µXComp is predicted by the end of the year.