
# Green Flash

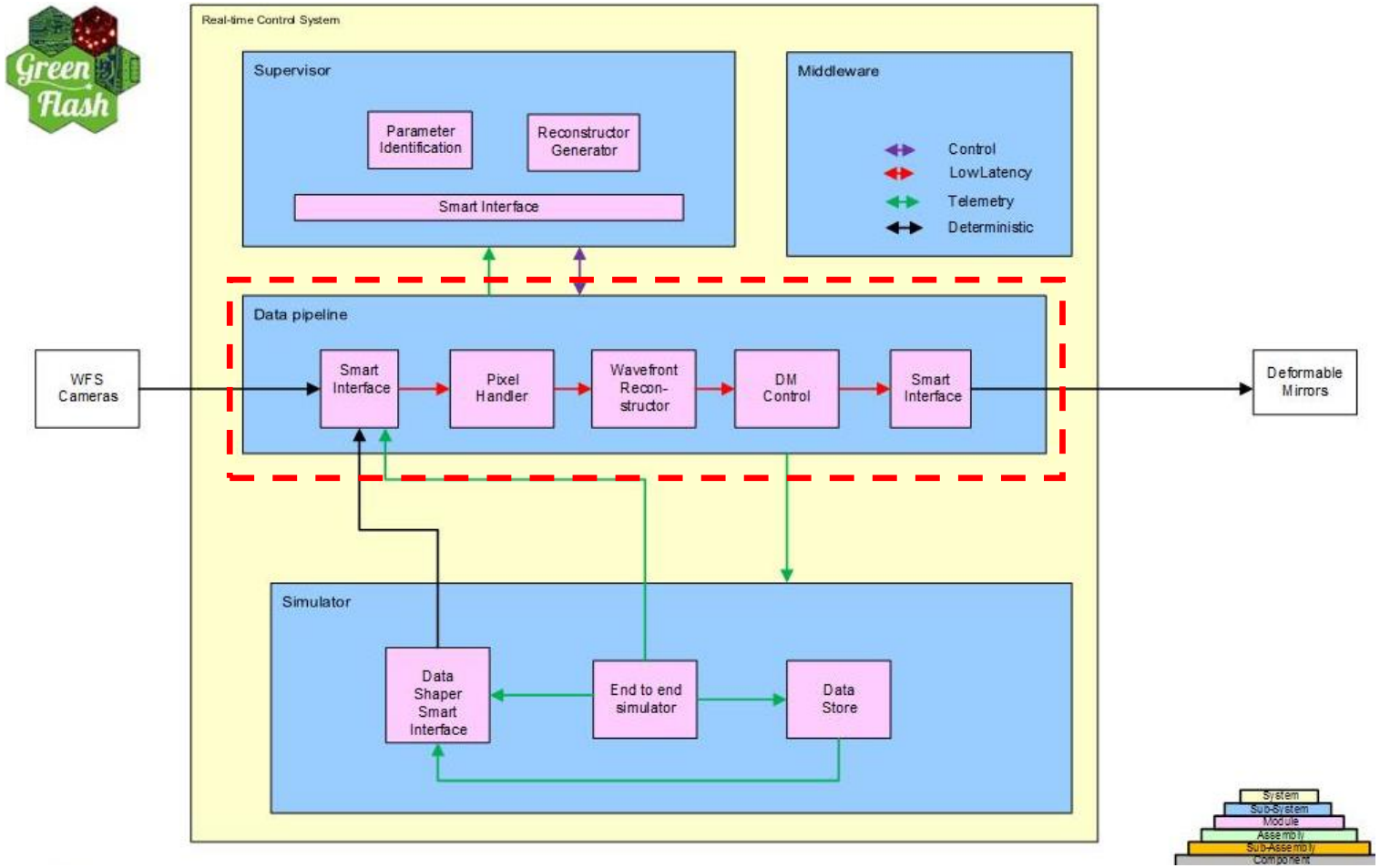
High performance computing for real-time science

Prototype Review 06/04/2018 Paris

## WP3 – FPGA solutions for Hard Real Time

# 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.



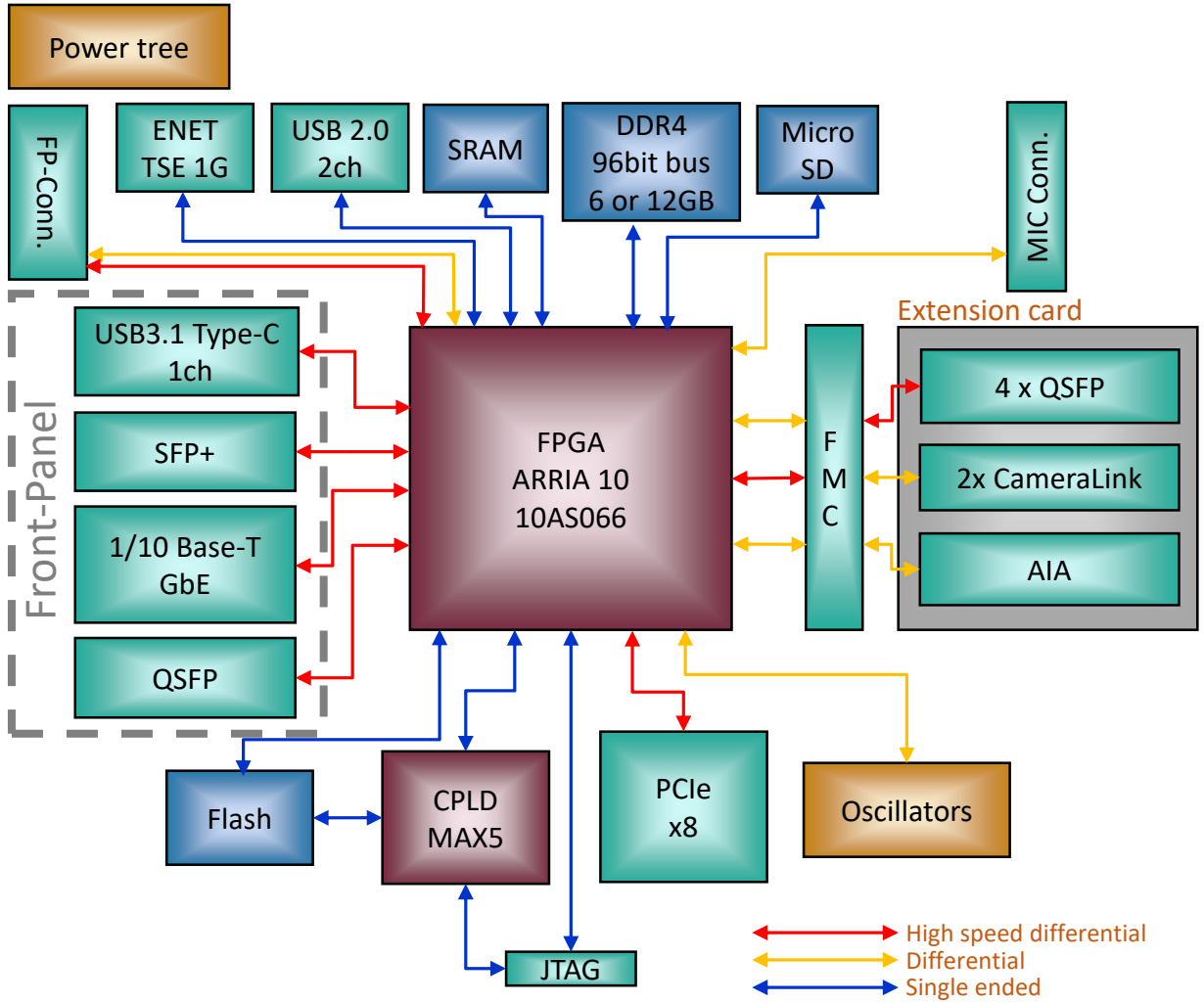
# WP 3: Solution

To realize the stand-alone microserver and fulfill the requirements of GreenFlash we will develop 2 different boards based on FPGAs.
One board called **µXComp** that acts 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 a hard-wired ARM processor in the same chip (SoC) and is used as an **interface and control board** to connect to serval types of computational boards, WFSs and DMs.
The hardware and the firmware will be designed to be compatible with PLDA QuickPlay
The µXComp board prototype is already manufactured and tested and was already presented at MTR.
The µXLink board is under development and first fully assembled prototypes are expected this month.

**The highlight of these boards:**

- Both are based on Intel ARRIA 10 FPGAs and the µXLink contains a ARM dual-core co-processor in the same chip with the FPGA
- 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, 40G Ethernet …
- 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

# µXLink board outlook




**Board features:**

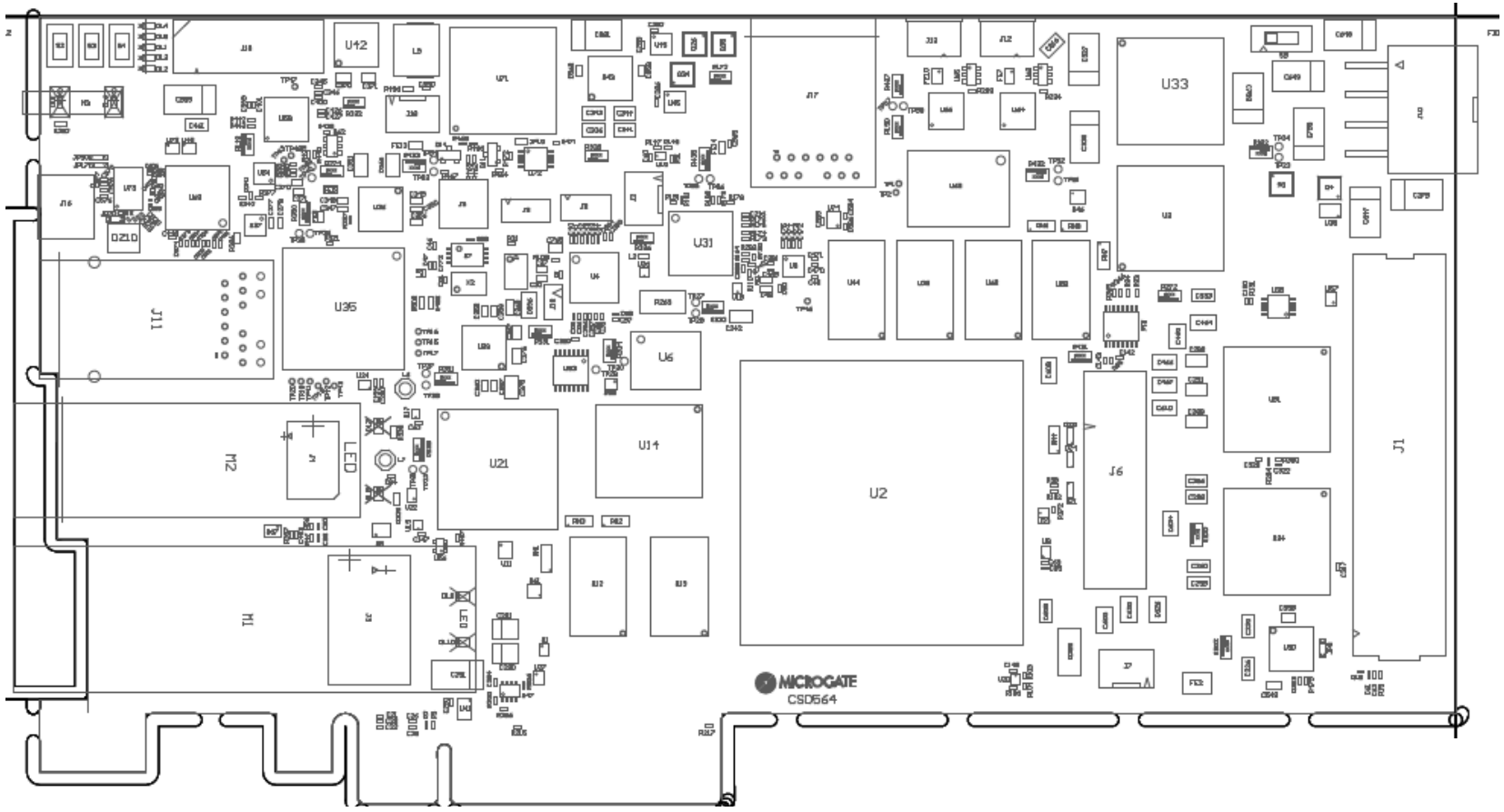
- Based on the Intel ARRIA 10AS066 SoC
- Dual core **ARM Cortex-A9 embedded processor** (with OS)
- Max. nr. Transceivers -> 48 (used 43)
- 1855 DSP blocks
- 5.2MB FPGA internal RAM
- 6 GB DDR4 (no HMC because only 48 transceivers)
- Can be configured as well as **powerful PCIe root complex** because of ARM and OS (e.g. Linux) for a stand-alone box

**On-board interfaces:**

- 8 lanes PCIe up to Gen3 as endpoint or root
- SFP+ for 1G/10GbE (or different interfaces)
- RJ-45 for 1G/10G Ethernet over copper
- QSFP for 40 GbE or Infiniband (or different interfaces)
- USB3.1 Gen 2 Type-C to connect external hard drives, monitors, docking stations
- 2x USB2.0 directly to ARM co-processor
- 1G Ethernet directly to ARM co-processor
- MicroSD card slot directly to ARM co-processor
- Configurable digital I/O connector , e.g. to handle synchronization signals (MIC Connector)
- FP-Connector to attach a second front-panel
- FMC connector equal to µXComp for expansion boards on which several additional interfaces can be implemented, e.g. other 10GbE, CameraLink, S-FPDP, …

# µXLink facts

The µXLink board is designed as an **interconnect** board that provides a high number of different interfaces that can be easily handled by the user from an Operating System running on the ARM co-processor. The flexibility of the FPGA together with the processor power of the ARM allows to implement different communication protocols on different physical mediums (e.g. copper, fiber) and to realize a PCIe root port. With a Linux running on the ARM and the PCIe as root the µXLink is capable to interface to the µXComp using the PCIe backplane (also other PCIe endpoint cards as GPUs possible). In this configuration the µXLink allows to realize a stand-alone machine that dose not require a Host PC or server. The PCIe edge connector can be configured as well as endpoint if the µXLink is used as interface board in a standard host machine or server.




**The main components and interfaces on the µXLink board are:**

- the ARRIA 10 SX FPGA U2 that contains the ARM Cortax-A9 dual-core co-processor in it.
- The Board controller MAX V U21 that is the same as on the µXComp and programs the FPGA at power-up and performs the house-keeping
- The USB3.1 Gen 2 controller U69 from ASMedia that is connected via PCIe to the FPGA and provides up to two USB3.1 ports.
- Marvell Phy U35 that provides 1G/10G Ethernet over copper
- The 6 DDR4 chips
- FP-Connector J18
- FMC connector J1
- MIC connector J6

# µXLink facts

**Board facts:**

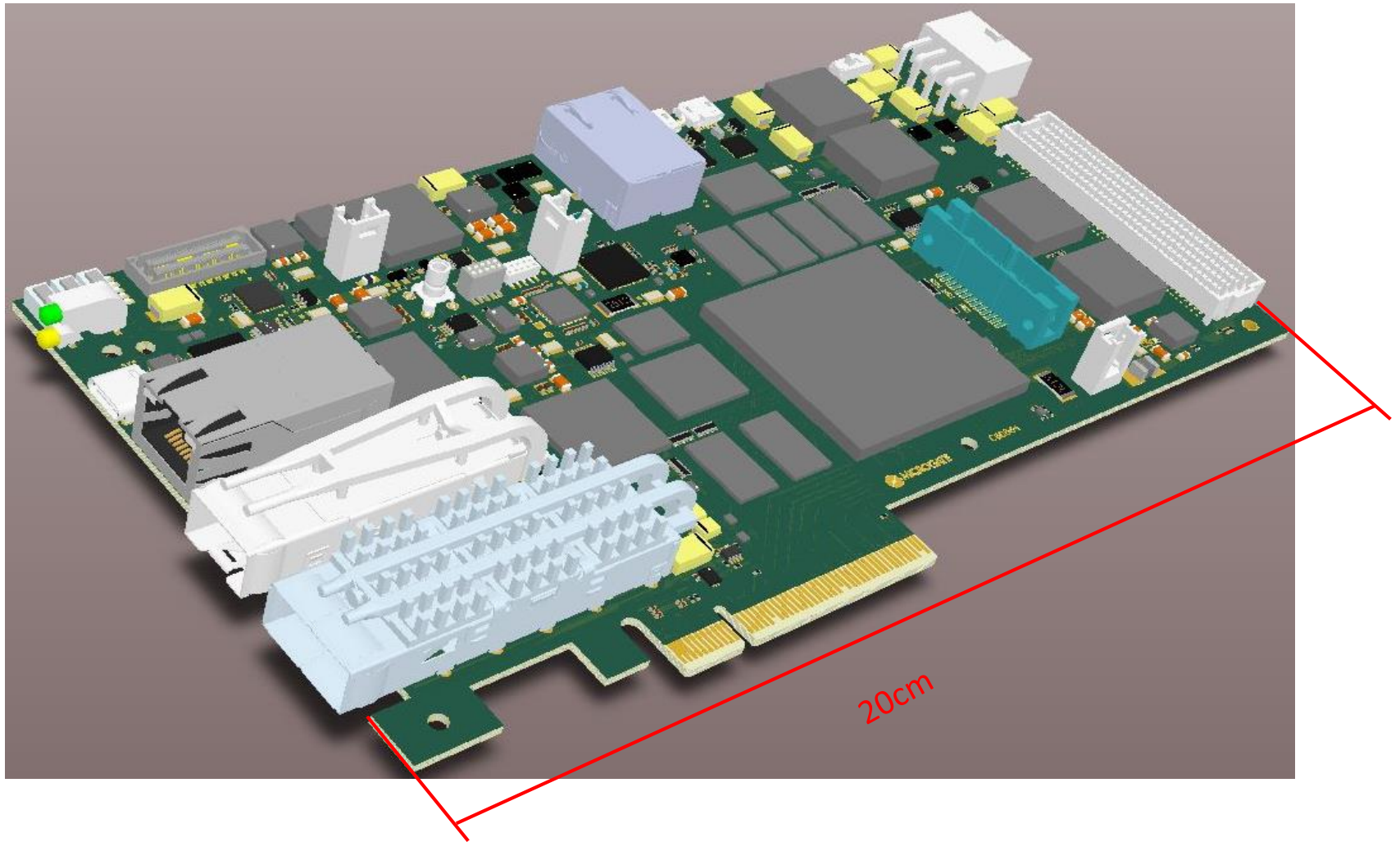
- Board size (111x200 mm) equal to µXComp compliant with PCIe standard full height and < ¾ length. With heat-sink requires double width PCIe slots
- Expansion Board of length 130mm in full-length PCIe standard
- # Layers: 18 (9 signal, 9 power-ground) = µXComp
- # Components: 1752 > µXComp
- # Connections: 7155 > µXComp
- # Vias: 10207 < µXComp

**Main performance goals:**

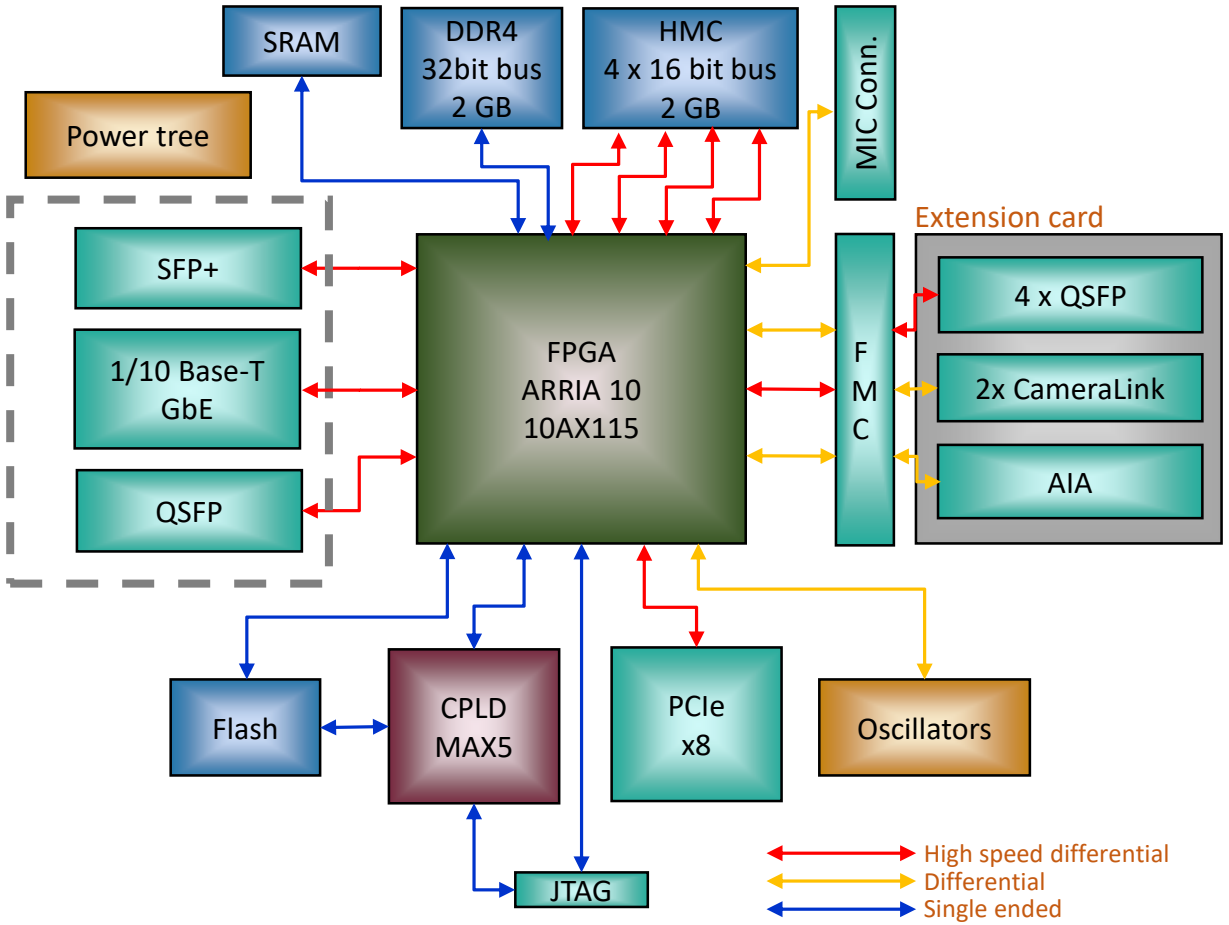
- ARM co-processor with 32bit dual-core Cortex-A9 running at a clock frequency of up to 1.2GHz
- Operating System on ARM e.g. Linux build on Yocto project
- 32-bit DDR4 RAMs directly connected to ARM and 64-bit DDR4 RAMs connected to the FPGA logic -> total memory bandwidth of 27GByte/s
- 10/40G Ethernet, Infiniband over fiber and copper (SFP+, QSFP, RJ45)
- USB 3.1 Gen 2 up to 10Gbps on modern Type-C connector to attach fast external hard drives or docking stations to the OS
- PCIe Gen3 x8 64Gb/s as root or endpoint performance by using a root adapter and some jumpers on board
- Power-Up and Power-Down sequence requirements met
- Total power consumption <60W
- FMC fully compatible with µXComp with equal performance

# µXLink: Current status

The first prototype of the µXLink board is designed and the PCBs are in production. The first fully assembled prototypes are expect soon and the start of the test campaign is expected by the end of the month. After the validation tests of the interfaces and the communication between FPGA and ARM we will start with the connection of the µXLink with one µXComp via the PCIe bus.



# µXComp board outlook




**Board features:**

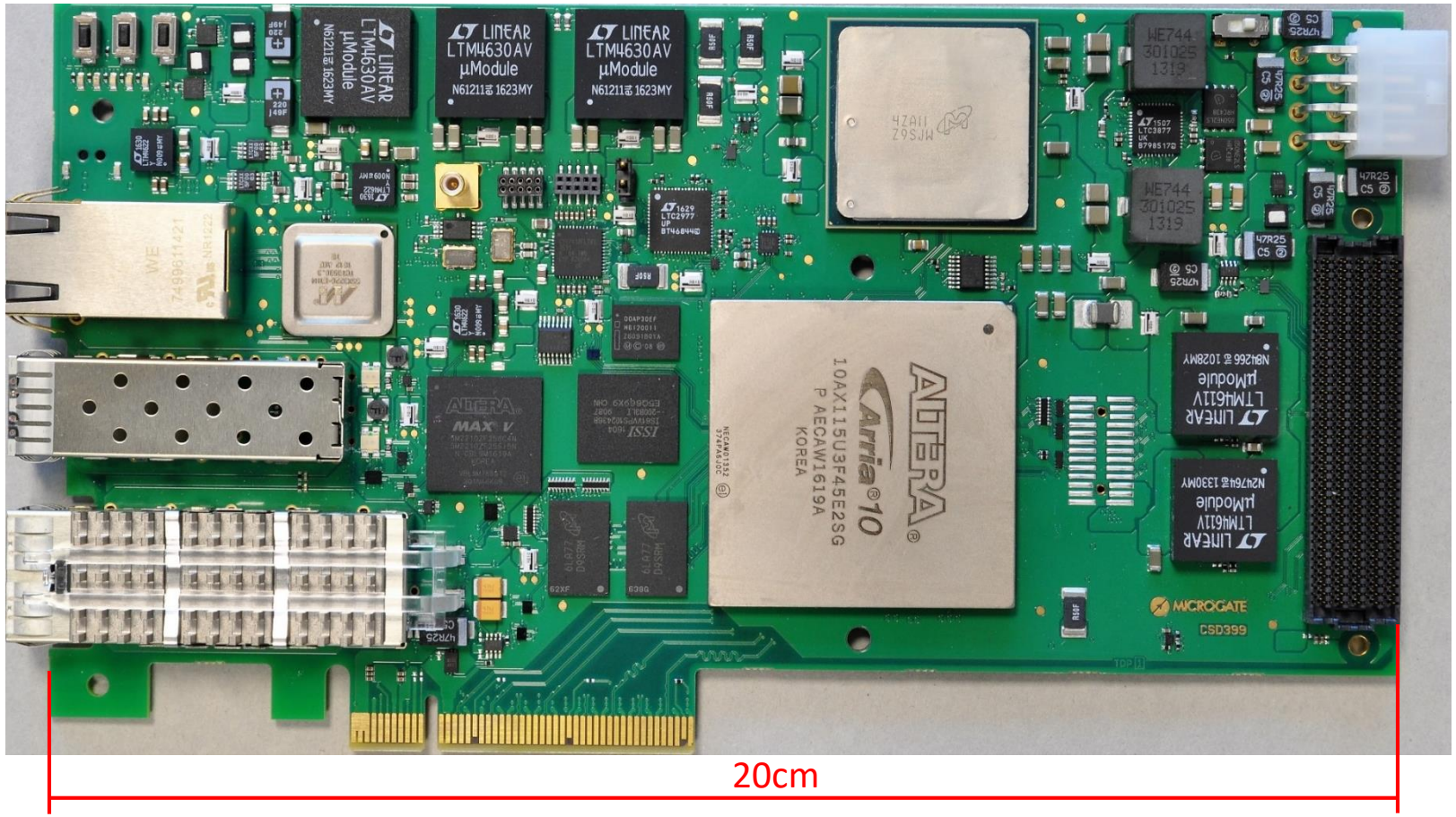
- Based on the Intel ARRIA 10AX115
- Max. nr. Transceivers -> 96 (used 95)
- 1518 DSP blocks
- 5.3MB FPGA internal RAM
- MAX V board controller for programming and housekeeping
- PLL chip with 8 differential and 4 single-ended outputs to program variable refclks for different interface standards
- 1Gb NOR Flash
- 2 GB DDR4
- 2 GB HMC memory with 4 links (64 transceivers)
- PCIe endpoint with Gen3 x8 edge connector

**On-board interfaces:**

- 8 lanes PCIe up to Gen3 for endpoint
- SFP+ for 10GbE (or different interfaces)
- RJ-45 for 1G/10G Ethernet over copper
- QSFP for 40 GbE or Infiniband (or different interfaces)
- Configurable digital I/O connector , e.g. to handle synchronization signals (MIC Connector)
- FMC connector for expansion boards on which several additional interfaces can be implemented, e.g. other 10GbE, CameraLink, S-FPDP, …

# µXComp: Current status

The first pre-production series of 6 µXComp board is manufactured and is currently under validation.
3 of the 6 produced µXComps have the faster FPGA mounted and can be used to boost the performance of the HMC and DDR4. The board has a PCIe standard height and with 20cm length it is less than ¾ length standard. With the heat-sink and fans mounted a PCIe double-width slot is required. The dimension is equal to the µXLink. In a full-length PCIe slots a extension card with up to 13cm length can be attached to the FMC.




**Board facts:**

- # Layers: 18 (9 signal, 9 power-ground)
- # Components: 1442
- # Connections: 6860
- # Vias: 12305

# µXComp: Test results

**Test results:**

- 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 developed for housekeeping (temperature and power measurements) and FPGA programming at power-up
  - 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.
  - JTAG switches tested adding the HMC JTAG.
- **PCIe** interface **x8** Gen2 and Gen3 tested and Gen2 works at full throughput of 3.2GB/s in write and read and **Gen3** at **6.4GB/s in Write** and **5.48GB/s in Read**
- **SFP+** tested with **10G Ethernet**
- **QSFP** tested with **10G Ethernet** on each of the **4** channels
- **RJ45** with Marvell Phy tested with 1G and **10G Ethernet** over cat. 6A copper cable
- FMC transceivers successfully tested with Transceiver toolkit and loopback board at 10Gbps for each of the 16 channels
- MIC connector tested with debug signals for clocks and timing measurement signals
- DDR4 32-bit bus evaluated up to 1067MHz clock with the Intel EMIF toolkit resulting in a memory bandwidth of 8.5GB/s
- **HMC** one Link and all 4 Links tested with transceiver speeds of 10Gbps and 12.5 Gbps: all 4 Links working in parallel gives a memory bandwidth of **71GB/s** for 10Gbps and **88.6GB/s** for 12.5Gbps in each direction. With the faster FPGA device further tests with 15Gbps transceiver speed can be performed.
- MVM calculation 222X5316 floating-points implemented for DP control -> works good with execution time of <30µs (only internal memory) corresponds to 40GMAC

More details on the obtained results and the performed tests can be found in the document GF-D3.2

# WP3: Conclusions

The first pre-production of the µXComp is produced by Microgate and the second FPGA based board, the µXLink is designed. With these two boards the hard real-time data pipeline execution can be realized in a microserver structure. The µXComp board is fully tested and the test campaign of the µXlink board is expected to stared end of the month.

Both boards are complex cards with 18 layers and a high number of components.
A high number of transceivers is used on both boards to realize a high number of flexible interfaces and a large computational throughput.

The ARM co-processor on the µXLink allows to realize a stand-alone microserver that can be used in various adaptive optics mirror systems and instruments.

The assembly of a Microserver containing one µXLink and one or more µXComps is predicted for summer this year.


# Green Flash

High performance computing for real-time science

Prototype Review 06/04/2018 Paris

## Microserver prototype concept

# Microserver: FPGA solutions for hard real-time

The **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*

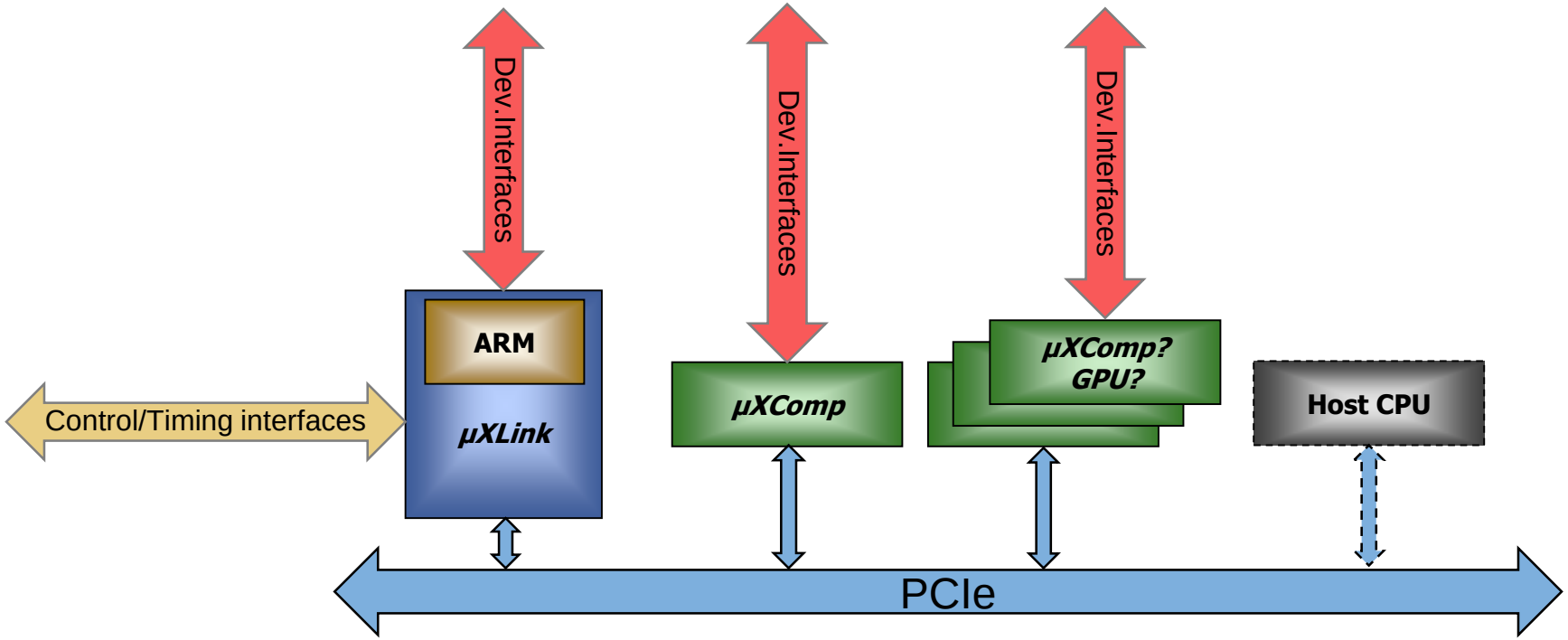
Extend the prototype of microserver board to include PCIe root complex

*Task 3.3: Firmware implementation*

Develop IPs + drivers for the microserver boards

# Microserver: FPGA solutions for hard real-time

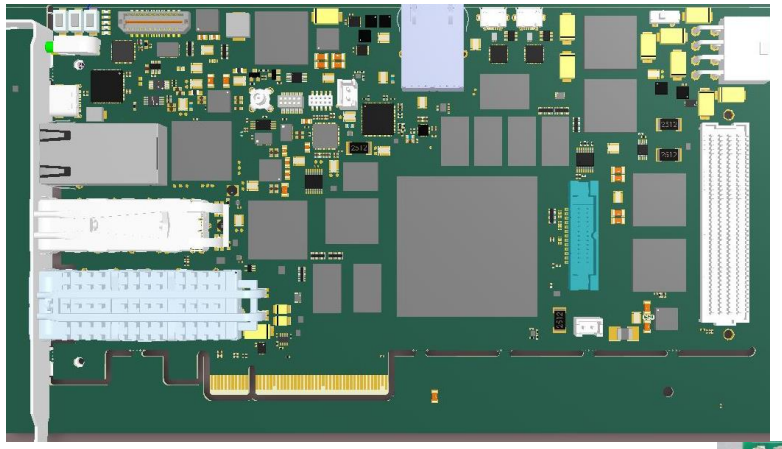
Our implementation of the Microserver concept comprises a PCIe backplane, on which the FPGA-based board µXLink that provides most of the interfaces the ARM co-processor with a OS and features PCIe root complex capability. In this way, the Microserver can operate in **standalone mode**, without requiring an additional host CPU. Several µXComp or GPUs can be inserted in the PCIe backplane and communicate with the µXLink. This allows achieving **optimal energy efficiency, low data transfer latency, low jitter and efficient management of telemetry data**.



# Microserver: FPGA solutions for hard real-time

The stand-alone Microserver can look like the example below: A box with the PCIe cards µXLink and µXComp or GPU inserted in the PCIe backplane and with the power supply and cooling fans. The interfaces are exiting on the front-panels on the back.

**Only an Example**



**µXLink**

**µXComp**


# Microserver: Goals

**Main goals:**

- Connecting µXLink with 1 to several µXComps using the PCIe backplane (requires root adapter for µXLink and PCIe backplane)
- RDMA and or GPUDirect for direct data transfer without requiring the processor
- Connecting µXLink with GPUs as accelerator cards
- Distributing the sensor data to the accelerator cards starting the computation on the accelerator cards µXComp or GPU and collecting the results
- Testing several interfaces to sensors and DMs
- Power consumption analysis
- Testing the configuration of the Microserver boards with QuickPlay
- Running QuickPlay on the ARM OS of the µXLink

# Microserver: Conclusions

- The current status of the prototype development allows us to say that we are still aiming to achieve the full goal of the project (extension of project until December would help)
- Major activities still to be performed:
  - Validation tests of µXLink board
  - Integration into Microserver (first 1x µXLink + 1x µXComp)
  - Implementation of a test case into the self-standing Microserver
  - Initial interfacing of other boards, in particular GPUs

- Deployment of the concept and related hardware already planned on several real systems (see presentation by Damien)