Scalable and Reliable Platform for AI-based Image Acquisition and Processing

DSc Dariusz Makowski, Associate Professor
Lodz University of Technology

Imaging Plasma Diagnostics

Diagnostic Systems provide data for:

- **Plasma Diagnostics, physics study (non-real-time)**
  - Physical studies, observation, measurements
  - Archiving measurements, raw data – **could be challenging** (transferring and archiving ~TBs - PBs of data)
    - **Stored data:** raw data, pre-processed, calibrated, measurements, meta-data (labels, tags)
    - **Important for ML and AI**
- **Machine and Plasma Control (real-time, soft real-time)**
  - Diagnostic systems supply information for Plasma Control System
  - Primary and supplementary systems
  - **Latency from 10 - 100 ms**
- **Machine Protection (hard real-time)**
  - Trigger interlocks to protect machine against damage
  - **Latency from 100 - 1000 ms**

Dariusz Makowski (dmakow@dmcs.pl)

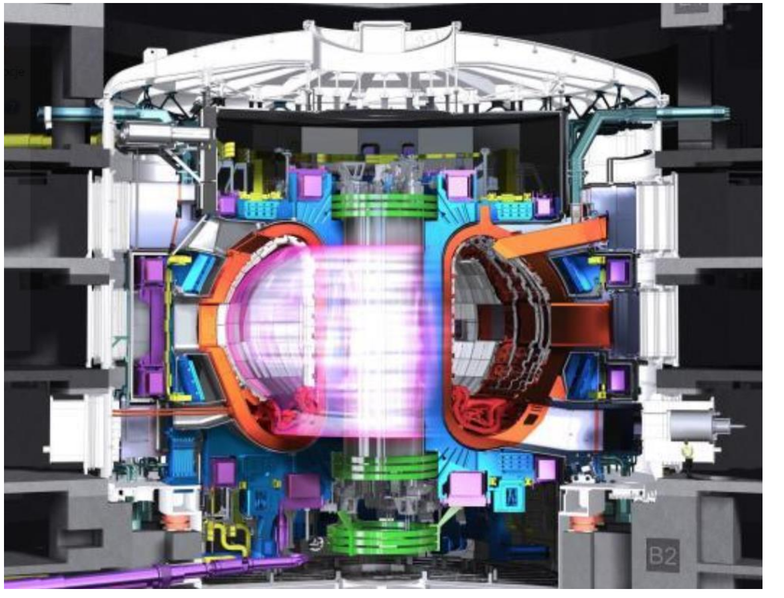
Fusion Projects - Plasma Diagnostics

Imaging diagnostics:

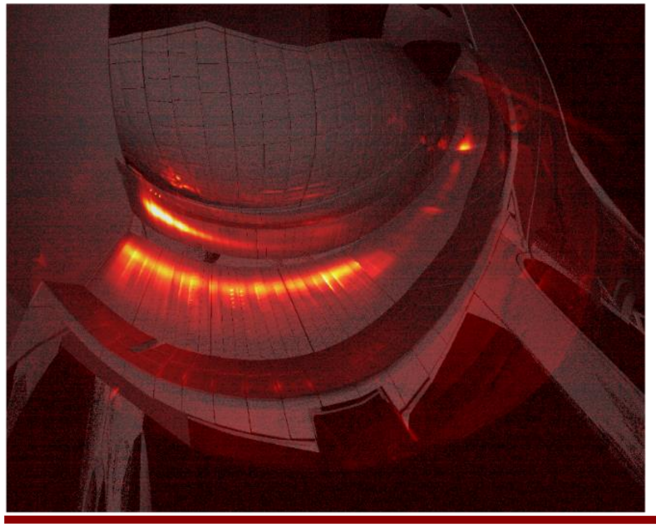
- **W7-X:**
  17 VIS and 13 IR cameras (OP2.1/OP2.2)

- **ITER:**
  more than 200 cameras

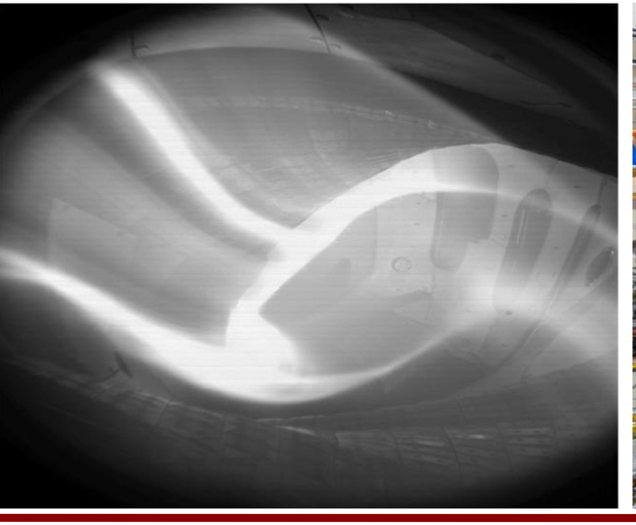
ITER Tokamak



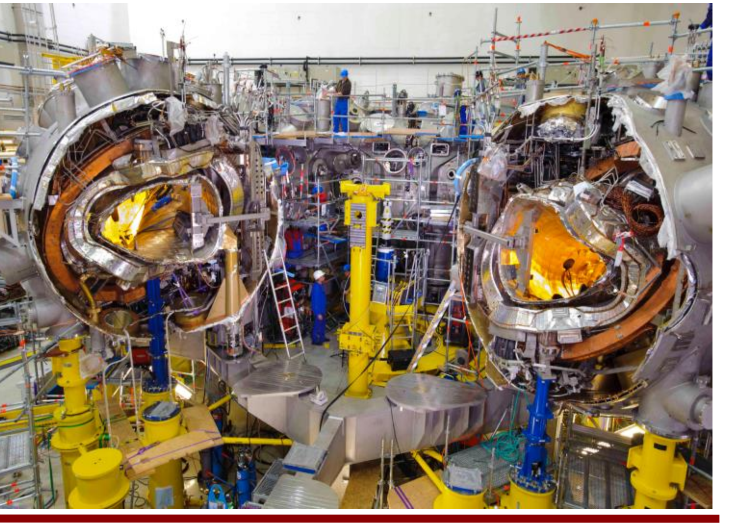
IR Diagnostics



VIS Diagnostics



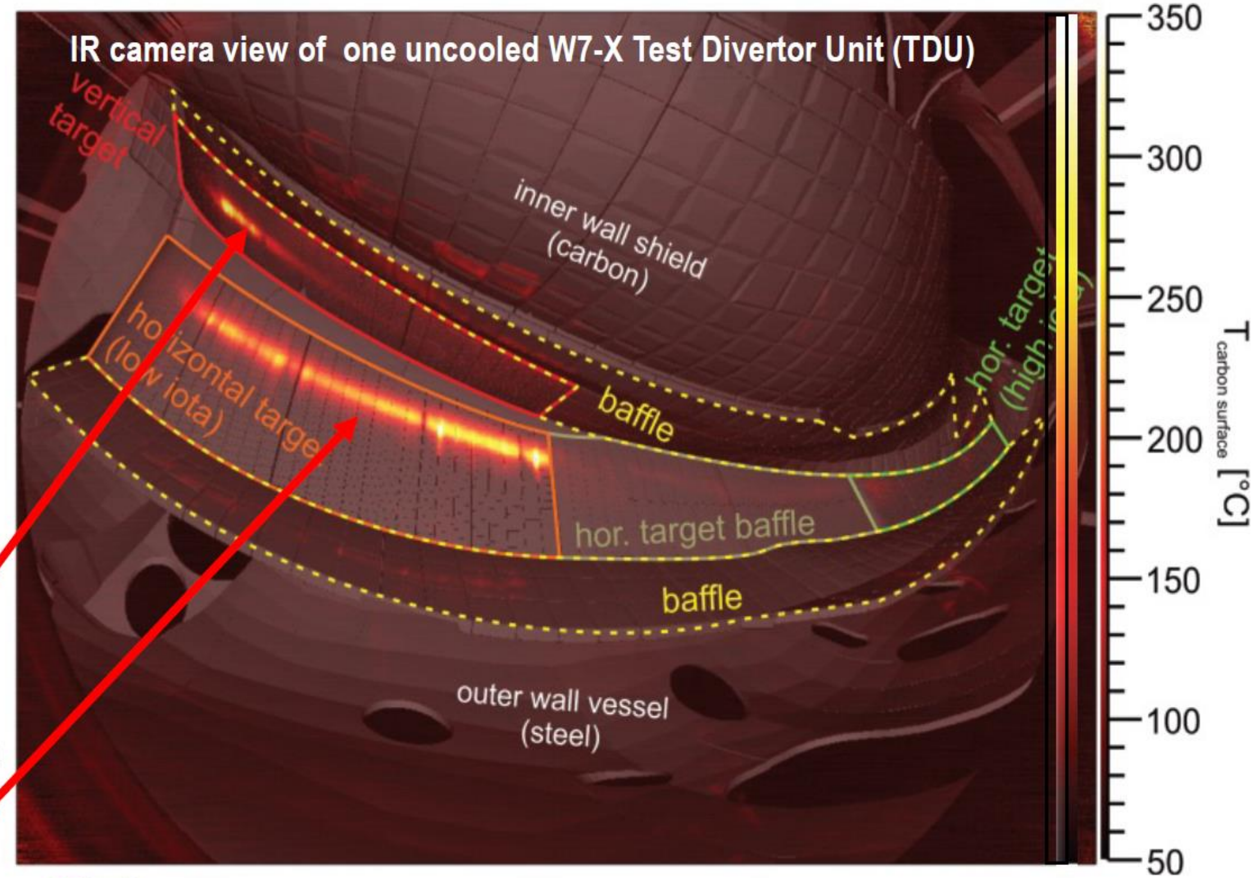
Wendelstein 7-X Stellarator

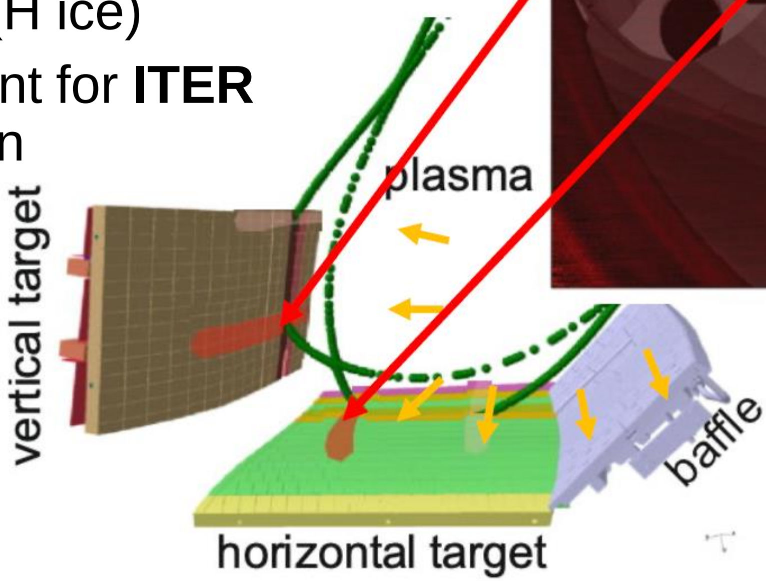


Dariusz Makowski (dmakow@dmcs.pl)

Wendelstein 7-X Divertor Protection

- Uses images from **IR cameras**
- One of the first machines with **cooled divertor**
- Divertor rated to heat flux: **10 MW/m²**
- Steady state operation at 10 MW
- **Continues pellet fuelling** system (H ice)
- Important for **ITER** operation





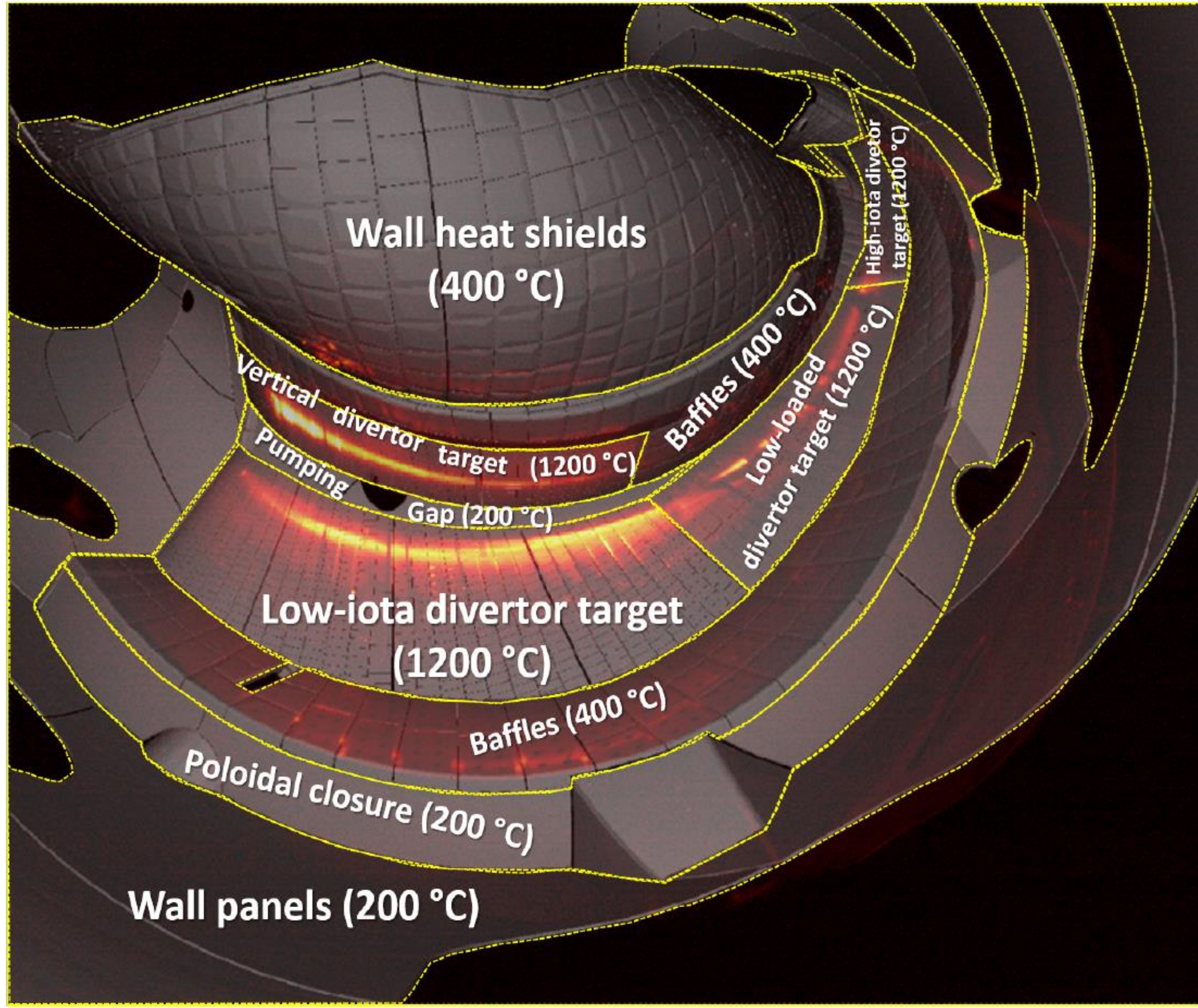
IR camera image of the divertor

T. Pedersen, IAEA FEC 2021

Protection of Plasma Facing Components

- **Divertor tiles**: Carbon Fibre Composite (CFC) joined to CuCrZr cooling structure
- Max. Operational temperature is limited by a **Cu to 475 °C**
- **Max. surface temperature** is **1200 ° C** for 10 MW/m2
- PFCs (graphite tiles) up to 400 °C
- Wall and pumping gap panels up to 200 °C

A. Puig, IAEA 2021



Dariusz Makowski (dmakow@dmcs.pl)

ITER Imaging Diagnostics

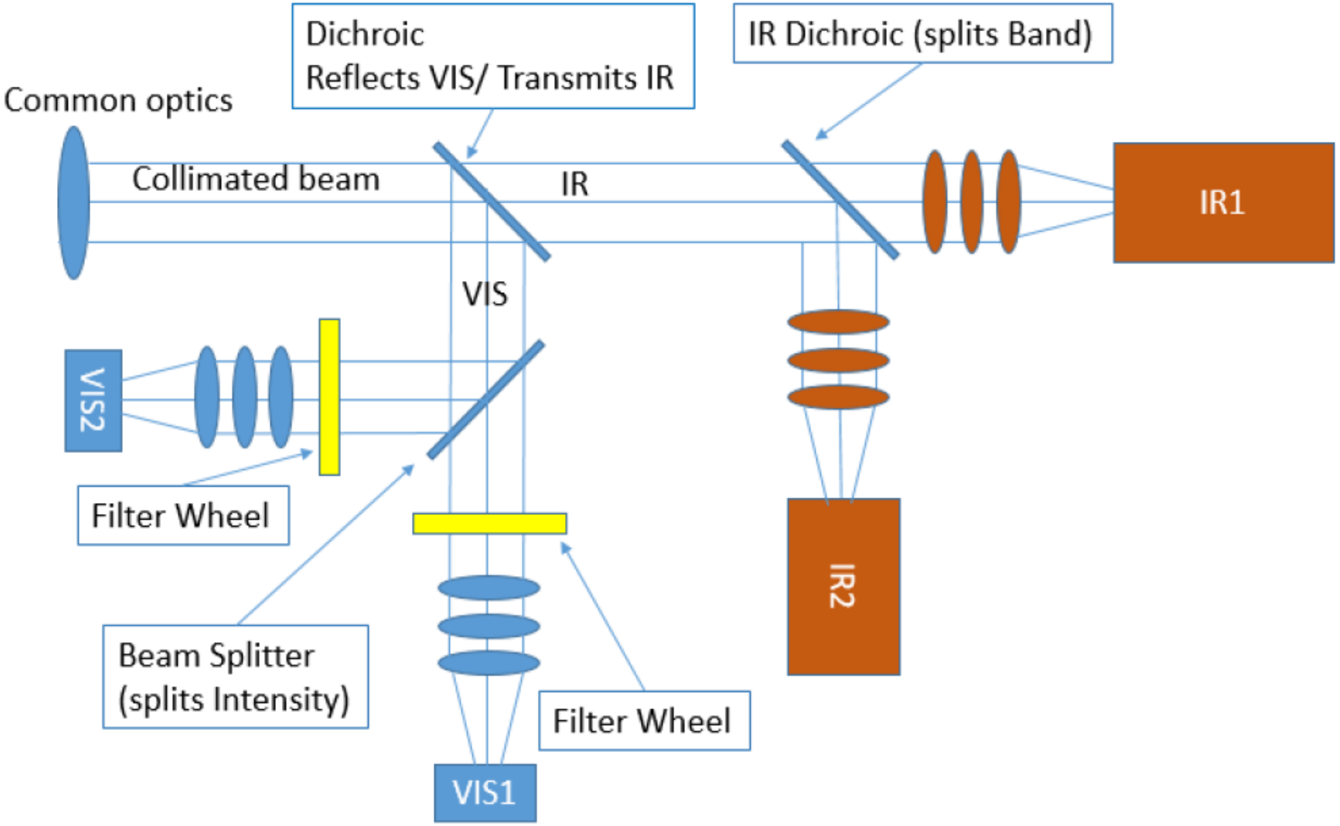
- ITER will use more than **16 diagnostic** systems based on cameras
- More than **200 digital VIS and IR cameras**
- **Needed a dedicated methodology to design and integrate systems provided by various Das**
- Diagnostic systems **provided by 7 DA** from different countries
- The most important systems:
  - **55.G1** - VIS/IR Equatorial Port Wide-Angle Viewing System
    - **55.G1.C0** - Equatorial Visible & Infrared Wide Angle Viewing System (WAVS)
  - **55.GA** - Upper Port Visible and Infrared Viewing System (UWAVS)
  - **55.GE** - Divertor Flow Monitor
  - **55.E2** - H-alpha and Visible Spectroscopy Diagnostic
  - **55.G6 -** Divertor Thermography
  - **55.B9** - Lost Alpha Monitor

55.G1 - VIS/IR Equatorial Port Wide-Angle Viewing System (1)

- Role of the system: **Plasma Control and Machine Protection**
- Visible and IR **viewing** and **temperature** data of the **first wall** and **divertor** to protect it from damage

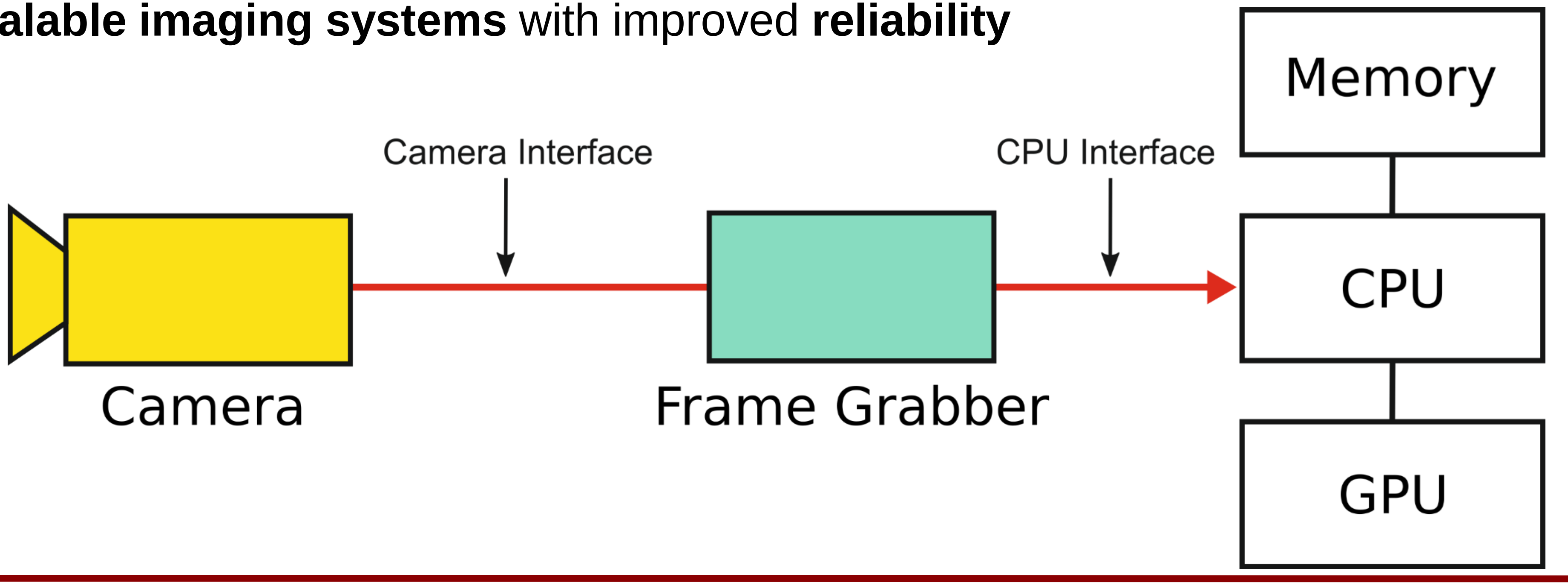
Uses:

- **30x IR** (3-5 µm)
- **30x VIS** (400-700 nm) cameras
- Optional fast IR / fast VIS operation (high framerate)
- **Redundancy** in both paths IR and VIS
- **Provides 25 various measurements**



Imaging Diagnostics – Image Acquisition and Processing (1)

- A single camera provides **one or more streams** of images
- **Frame grabber** configures camera, start and stop DAQ
- **All operations must work in real-time (hard real-time system)**
- Looking for **reliable, scalable and standardized solution (hardware/software)** suitable for **AI and ML real-time** applications
- Looking for a **methodology** to build **complex** (more than 50 cameras) and **scalable imaging systems** with improved **reliability**



Dariusz Makowski (dmakow@dmcs.pl)

Imaging Diagnostics – Image Acquisition and Processing (2)

- A single camera provides **one or more streams** of images
- **Frame grabber** configures camera, start and stop DAQ
- **All operations must work in real-time (hard real-time system)**
- Looking for **reliable, scalable and standardized solution (hardware/software)** suitable for **AI and ML real-time** applications

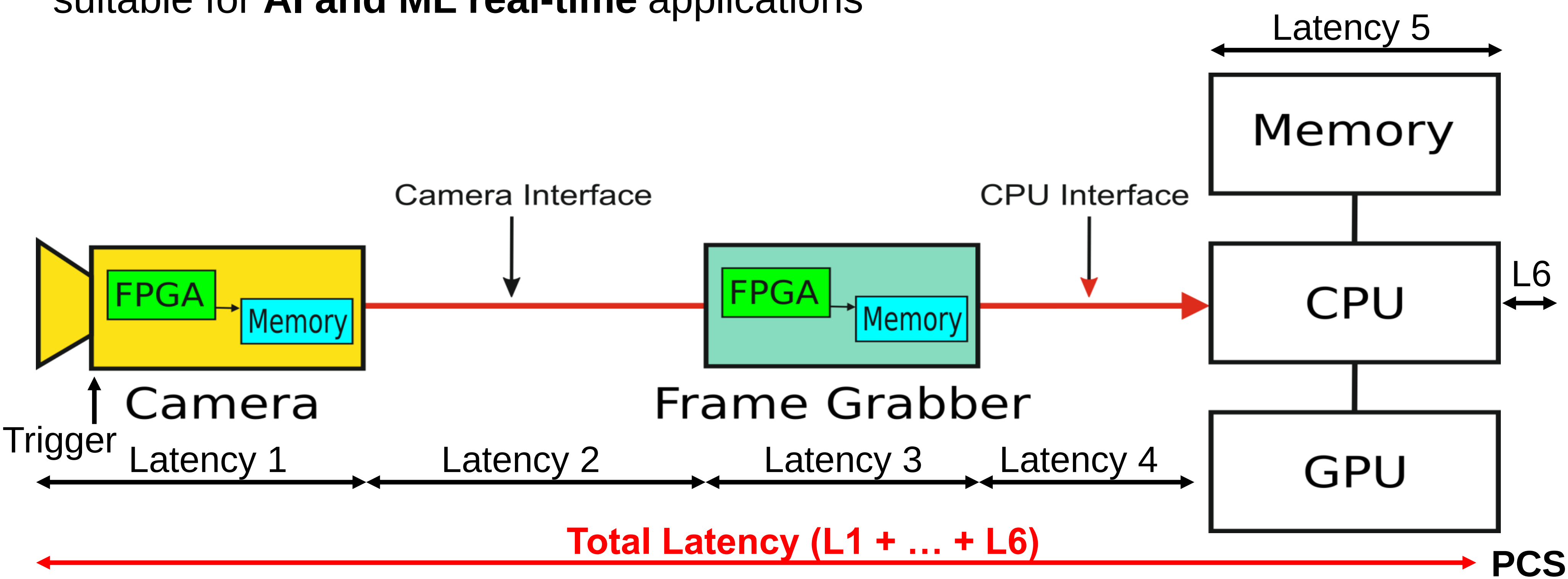
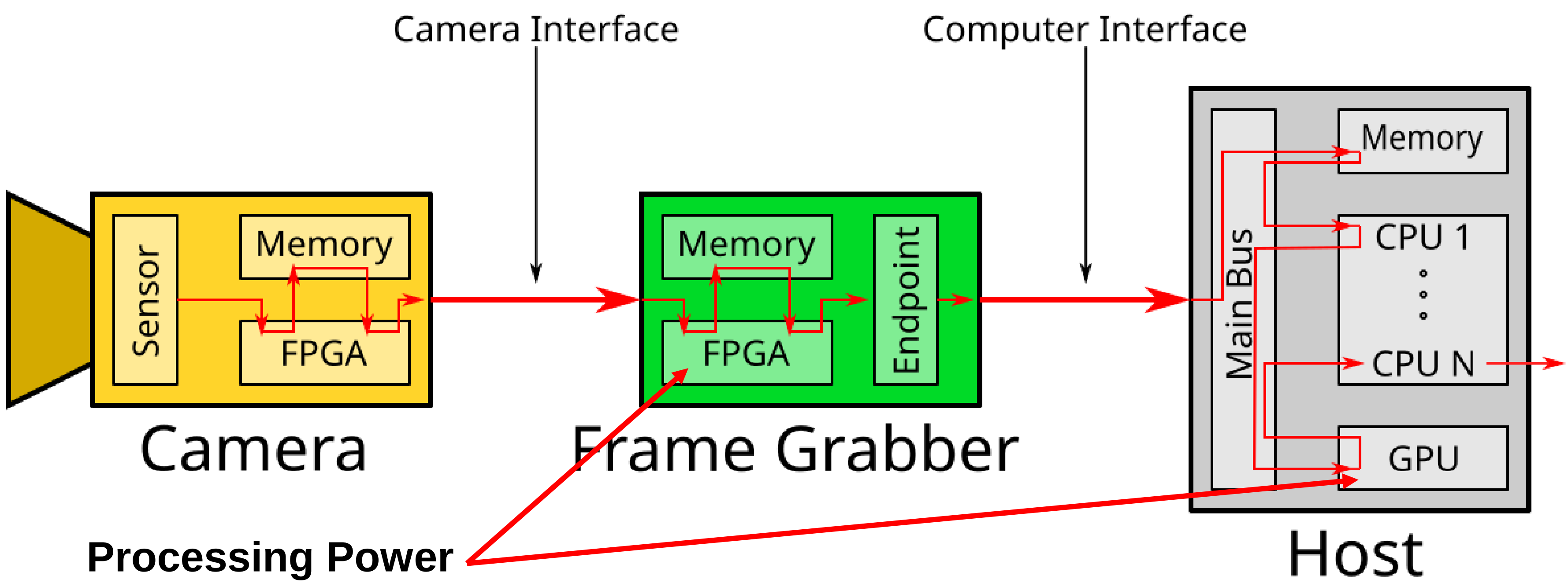


Image Processing Platform for AI

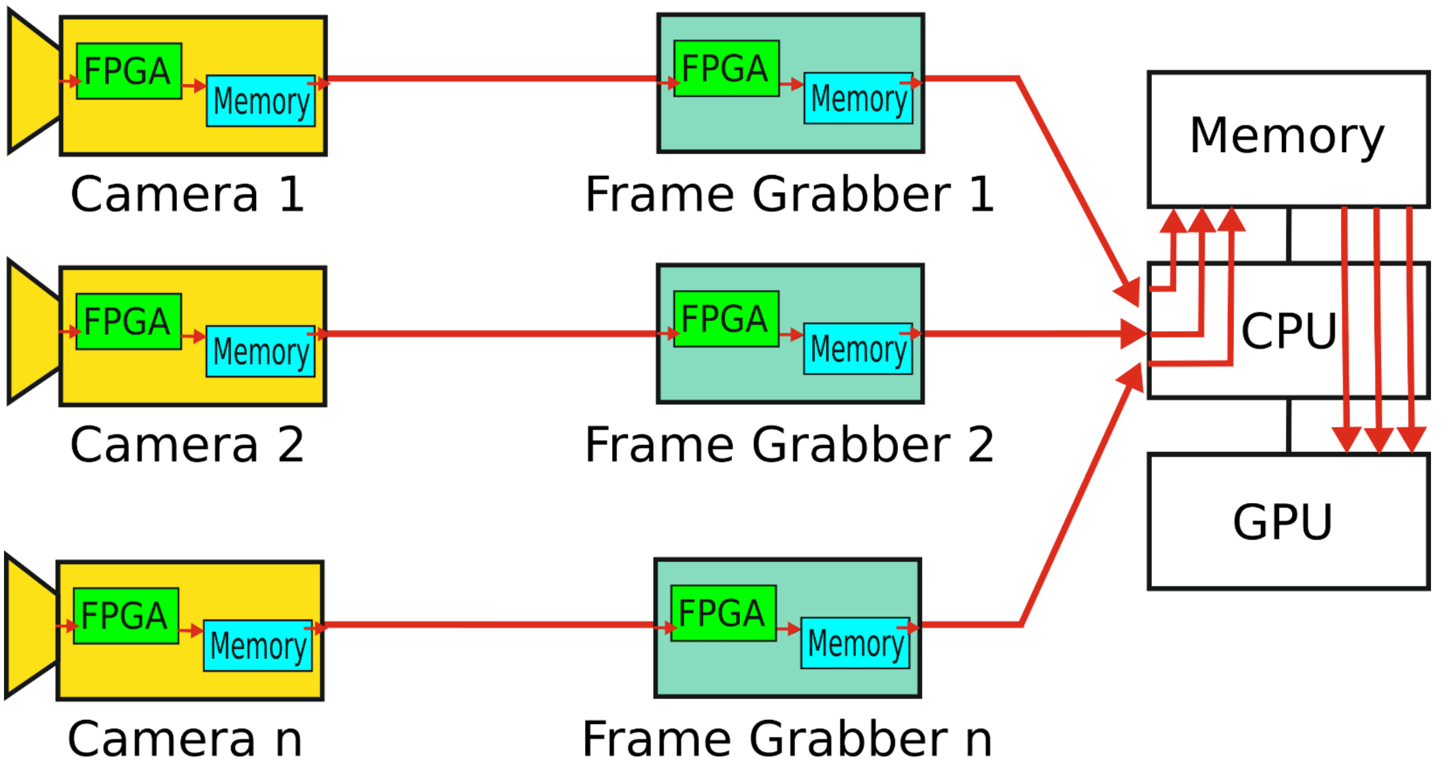
- Application of **AI algorithms** (Deep Neural Networks) requires **a significant processing power** for image processing
- **Frame grabber FPGA** (Field Processing Gate Array),

  GPU/GPGPU (Graphics Processing Unit) are used for **AI algorithms execution**
- **Standard solutions use multiple memory copying**
- **Still looking for low-latency real-time solution**



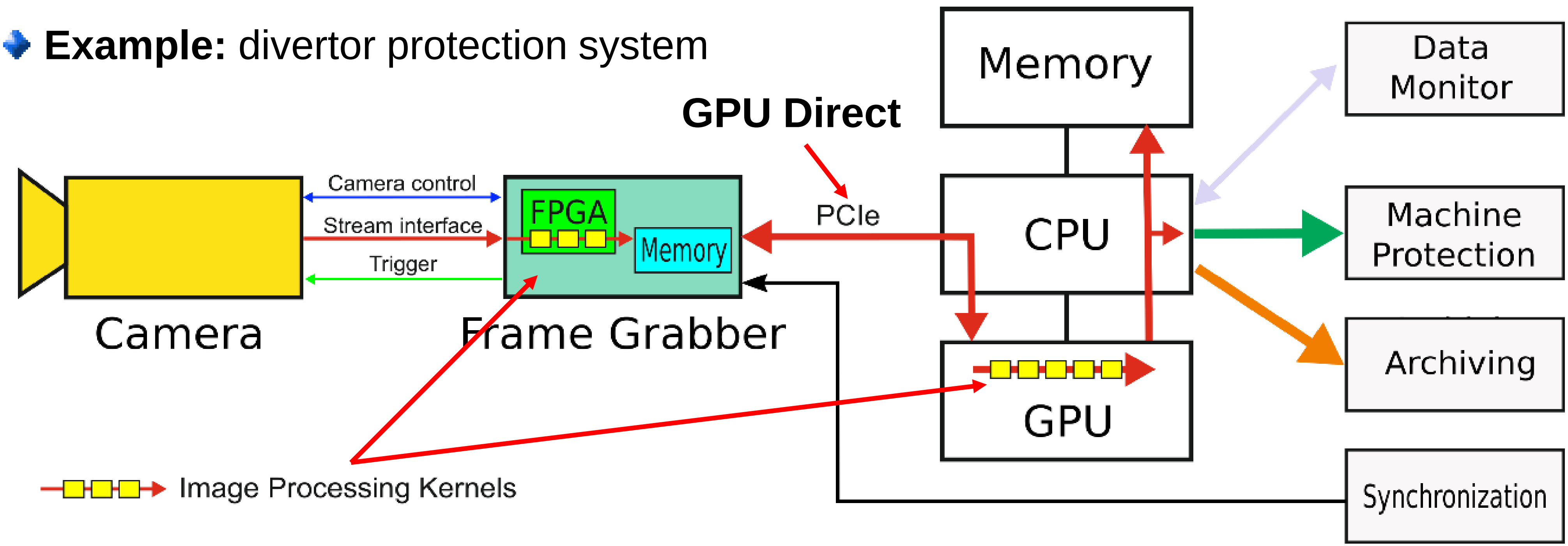
Imaging Diagnostics – Image Acquisition and Processing (3)

- In real system the situation is more complex, we have more cameras
- All devices must be **synchronised** with machine (each frame includes timestamp)
- **All operations must work in real-time (hard real-time system)**



Imaging Diagnostics – Image Acquisition and Processing

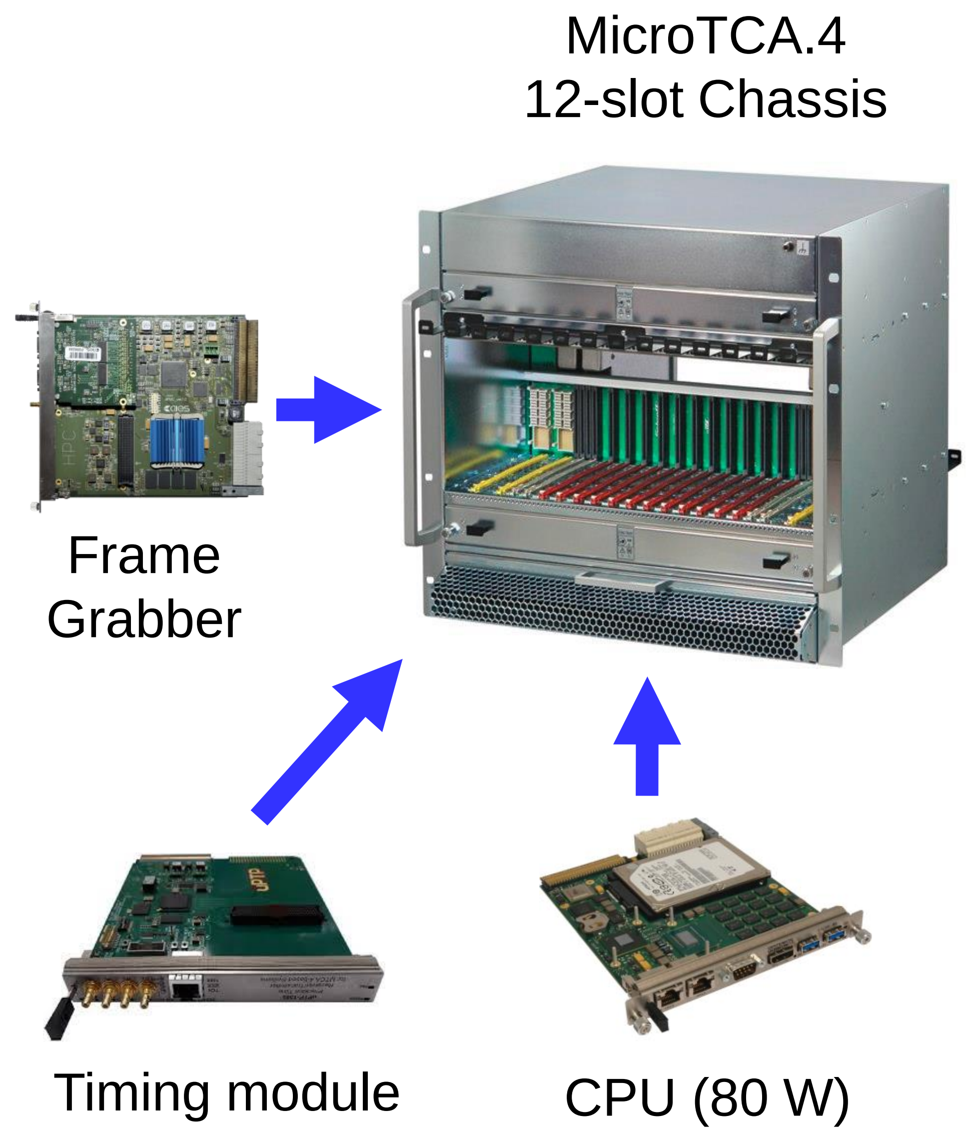
- Looking for a **standardized and scalable solution (hardware/software)** suitable for AI and ML real-time applications
- Requires **improved Reliability and Serviceability**
- Need **low-latency**, high-performance solution
- Could **AI-algorithms fulfil real-time requirements?** How to measure and guarantee this?
- **Example:** divertor protection system



MicroTCA – PICMG Industrial Standard

Micro Telecommunication Computing Architecture:

- High RAMI (**Reliability**, **Availability**, **Maintainability** and **Inspectability**)
- **Limited power** per slot to 80 W
- **Limited data throughput** to 32 Gbps
- **MTCA.0 @ 2006 (not-applicable)**
- **MTCA.4 @ 2011**
- **MTCA.4.1 @ 2016**
- PICMG is still working on new standard (release date: 2023/2024)

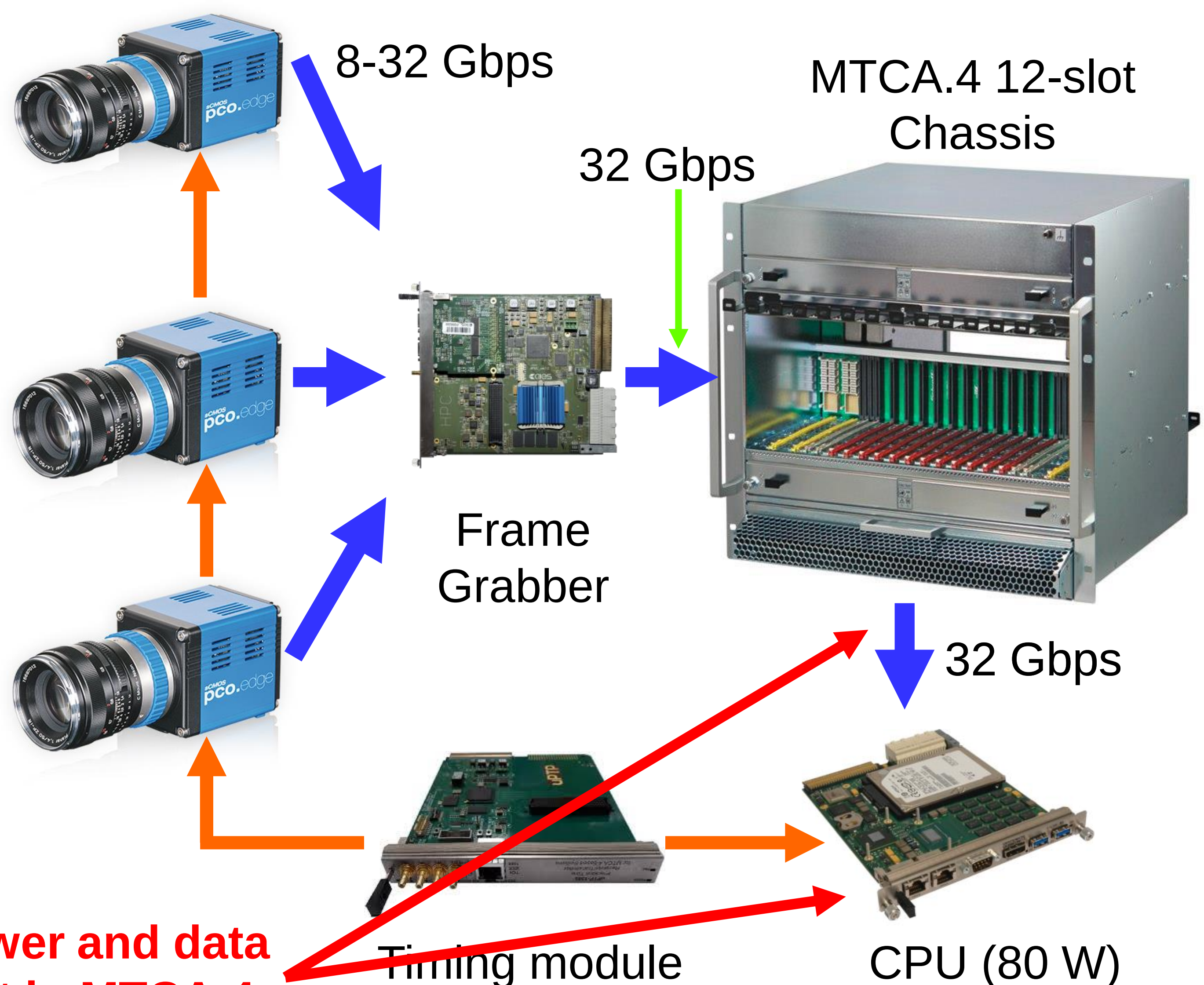


Dariusz Makowski (*dmakow@dmcs.pl)*

Image Acquisition and Processing with MicroTCA.4 (1)

Micro Telecommunication Computing Architecture:

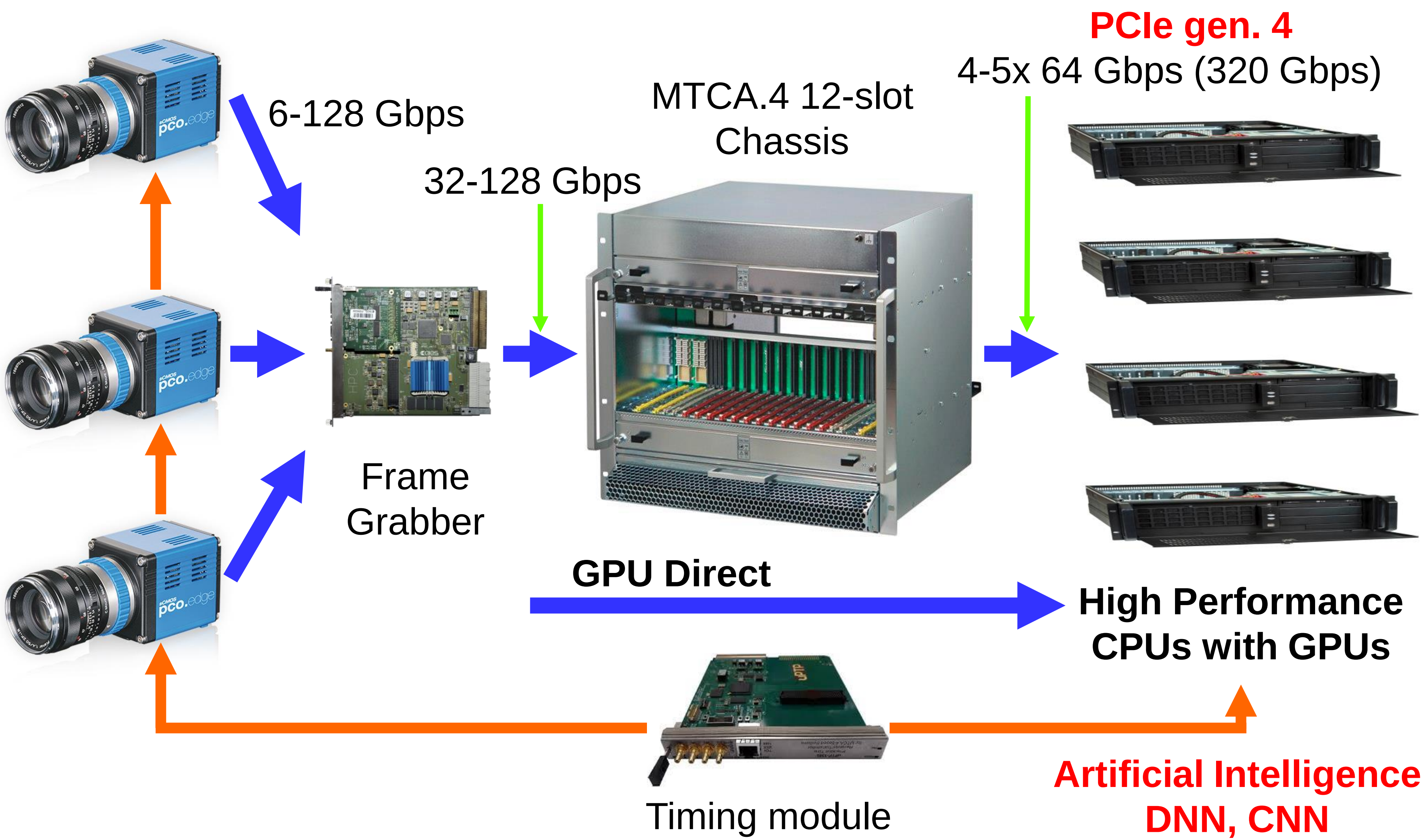
- **High RAMI** (Reliability, Availability, Maintainability and Inspectability)
- **Limited power** per slot to 80 W
- **Limited data throughput** to 32 Gbps
- **PICMG is working now on the NG-MTCA**
  - **PCIe gen. 5 (128 Gbps)**
  - **ca. 240 W/slot**



Limited power and data throughput in MTCA.4

Dariusz Makowski (dmakow@dmcs.pl)

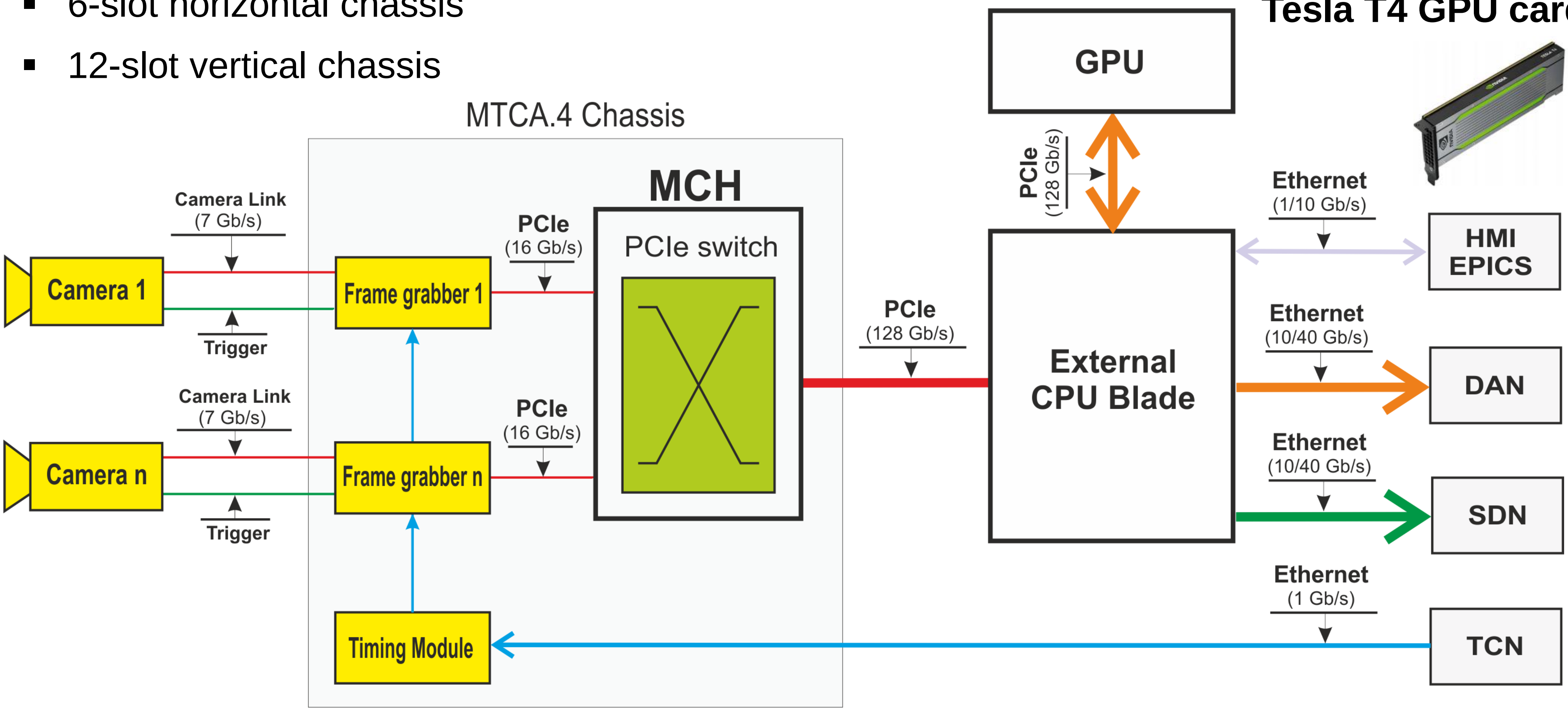
Image Acquisition and Processing with MicroTCA.4 (2)



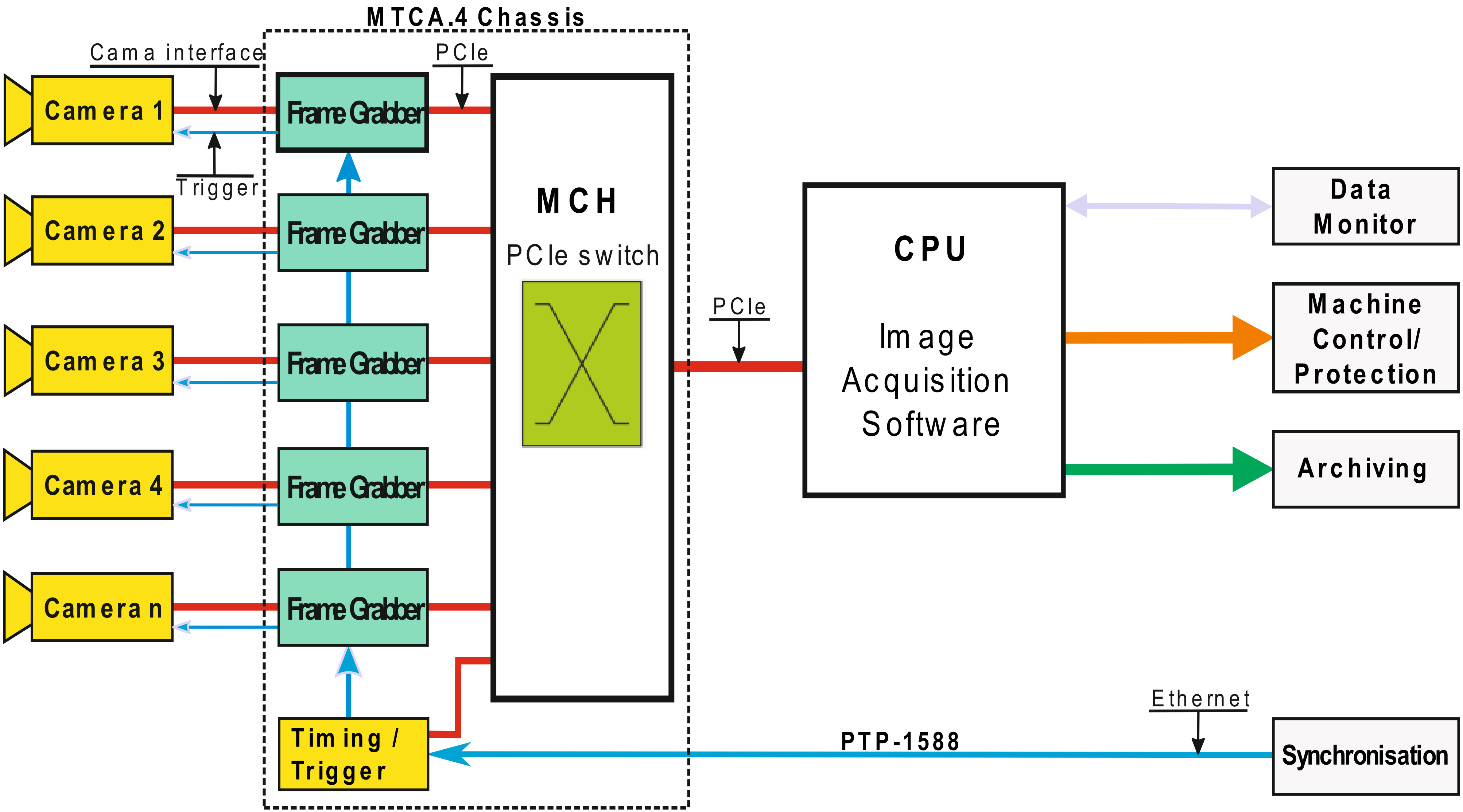
Dariusz Makowski (dmakow@dmcs.pl)

Image Acquisition and Processing with MicroTCA.4 (3)

- **Scalable, reliable and serviceable hardware architecture**
- Evaluation of **hardware components** for ITER fast controller catalogue
- Close collaboration with hardware manufacturers to integrate all components
  - 6-slot horizontal chassis
  - 12-slot vertical chassis



Block Diagram of Multi-camera Image Processing System



Dariusz Makowski (dmakow@dmcs.pl)

55.G1 - VIS/IR Equatorial Port Wide-Angle Viewing System

Proposed solution:

- Develop methodology
- **IO catalogue components**
- Example system with CL
- Example system with 1GigE Vision
- More…

Test stand at TUL-DMCS

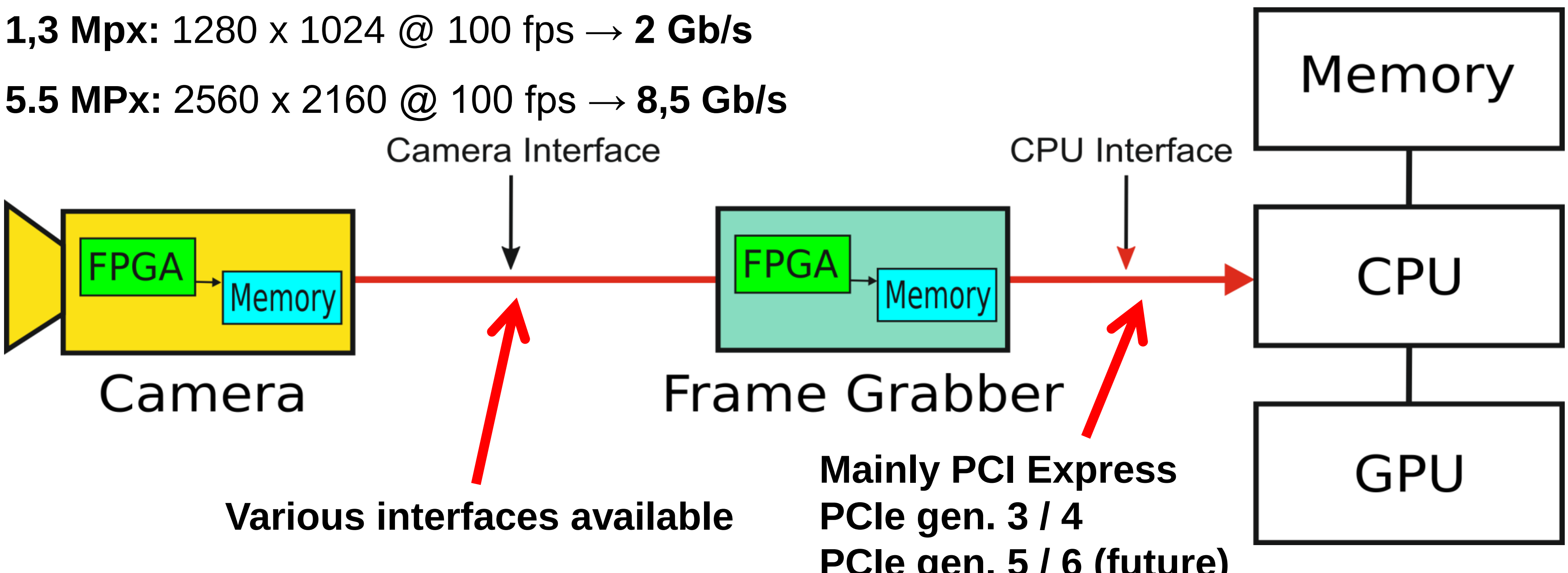
Dariusz Makowski (dmakow@dmcs.pl)

Interfaces

- **Various camera interfaces available**
  - Real-time and non-real-time (from fusion perspective)
  - Different parameters: reliability, latency, performance, etc.
  - Different electrical layer: cooper or **fiber**
- **PCI Express** is applied as main CPU interface – mainbone of computer

Data throughput:

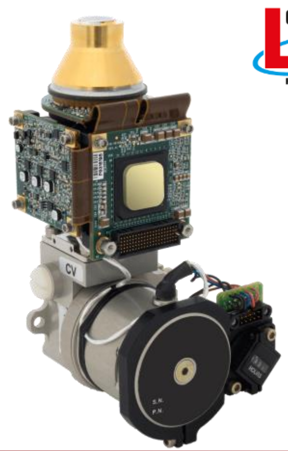
- **1,3 Mpx:** 1280 x 1024 @ 100 fps → **2 Gb/s**
- **5.5 MPx:** 2560 x 2160 @ 100 fps → **8,5 Gb/s**



Camera Interfaces Useful for Plasma Diagnostics

• <u>Camera Link</u>	2.04 Gb/s, 5.44 Gb/s, 6.8 Gb/s
• Camera Link-HS	2.4 Gb/s - 128 Gb/s
• CoaXPress 2.0	n x 6.25/12.5 Gb/s (n=4 → 25/50 Gb/s)
• 1 GigE Vision	800 Mb/s
• 10/25 GigE Vision	10/25/100 Gbps
• IEEE1394/Fire Wire	0.4 Gb/s (1394a) or 0.8 Gb/s (1394b)
• HD-SDI	1.45 Gb/s (max. 2.9 Gb/s)

SCD Hercules (CL)



Emergent HR-12000M camera with 10 GigE Vision interface

Dariusz Makowski (dmakow@dmcs.pl)

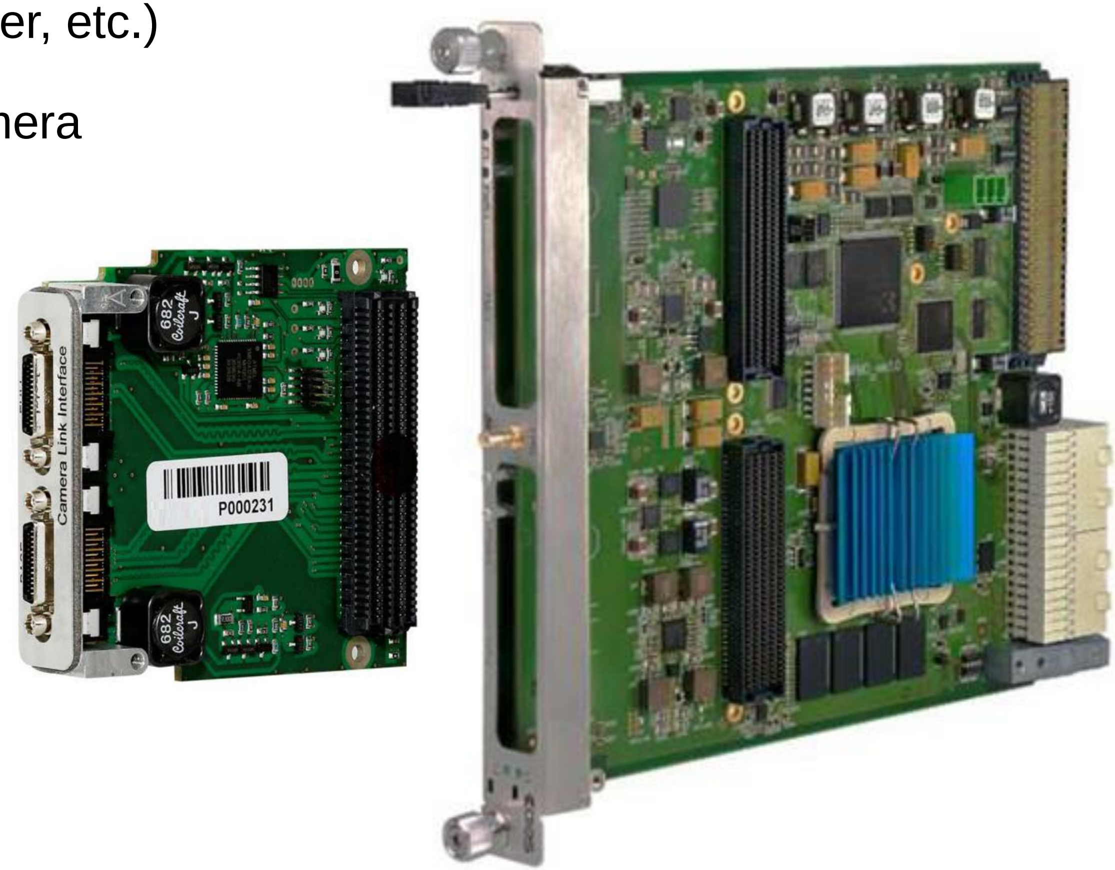
Universal Frame Grabber Module for MicroTCA.4

Frame grabber is composed of:

- **FMC carrier** (FPGA, DDR, PCIe, trigger, etc.)
- **FMC modules** supporting various camera interfaces (8 standards)

In addition, we need (**software**):

- **FPGA firmware**
  - **IP core** for selected camera interface
  - **Xilinx (IP Cores, Vivado)**
- **Linux driver**
  - Dedicated camera library
- **Image processing framework**
- **Additional tools**
  - Visualisation application



Mezzanine module

Carrier module

Dariusz Makowski (*dmakow@dmcs.pl)*

FMC Carrier Modules

Frame grabber is composed of:

- FMC carrier
  - Artix 7 FPGA (<6.5 Gb/s)
  - **Kintex US+ (>6.5 Gbps, 32 Gb/s per lane)**
- Mezzanine modules (FMCs) supporting various camera interfaces (8 standards)

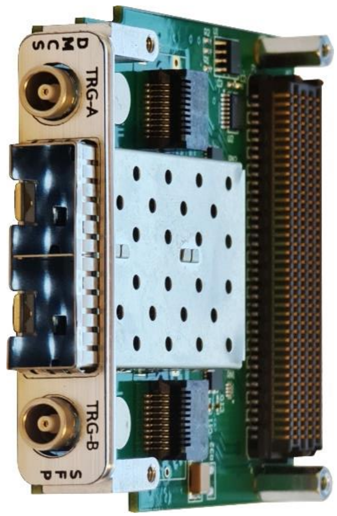
Software support:

- **IP cores** for selected **camera interfaces**
- Common Linux driver
- **Dedicated camera library (GenICam)**
- Real-time processing software
- **Algorithms (FPGA, CPU, GPU)**

Dariusz Makowski (*dmakow@dmcs.pl)*

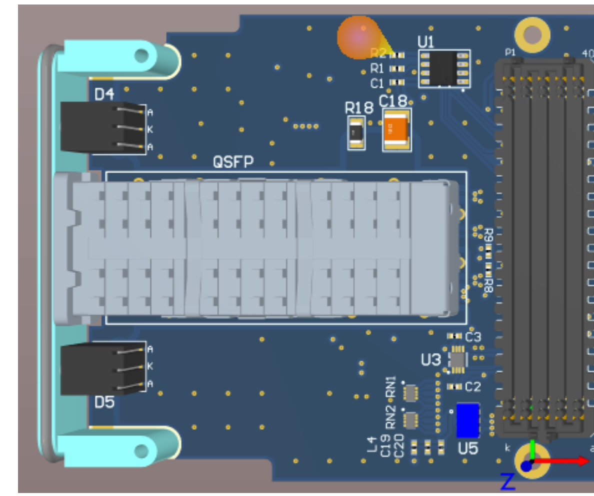
FMC Modules – Camera Interface (Selected Modules)

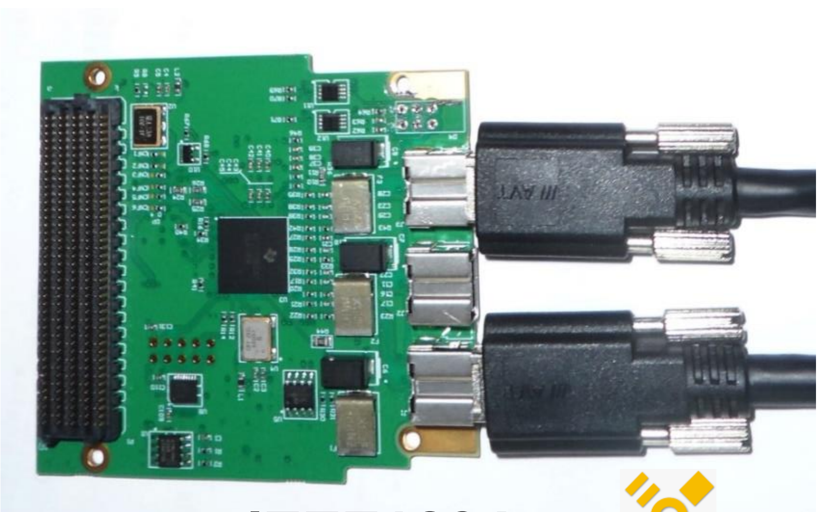
CAMERA Link



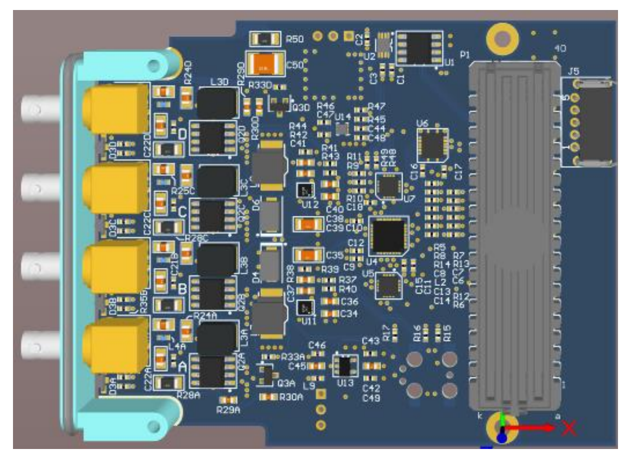
CAMERA Link HS

GiGE VISION





IEEE1394

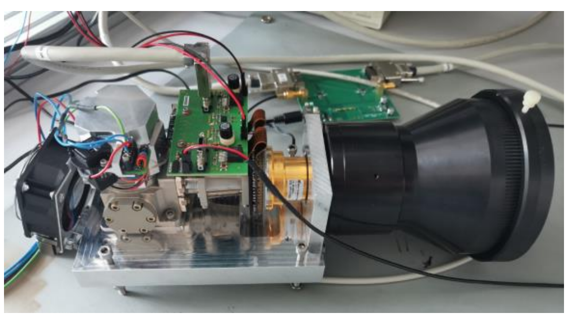


CoaXPress

Camera Interfaces

✓ Camera Link	2.04 Gb/s, 5.44 Gb/s, 6.8 Gb/s
✓ Camera Link-HS	2.4 Gb/s - 128 Gb/s
✓ CoaXPress 2.0	n x 6.25/12.5 Gb/s (n=4 → 25/50 Gb/s)
✓ 1 GigE Vision	800 Mb/s
✓ 10/25 GigE Vision	10/25/100 Gbps
✓ IEEE1394/Fire Wire	0.4 Gb/s (1394a) or 0.8 Gb/s (1394b)
✓ HD-SDI	1.45 Gb/s (max. 2.9 Gb/s)

SCD Hercules (CL)



Active Silicon (CXP-12)

Imperx Cheetah (10GigE Vision)

Supported Cameras with Developed Frame Graber

Mikrotron MC3010

Teledyne DALSA

Allied Vision PIKE F-145

Dalsa Genie T2505

Raptor Cygnet 2.1MP

PCO Edge 5.5

Andor Neo 5.5

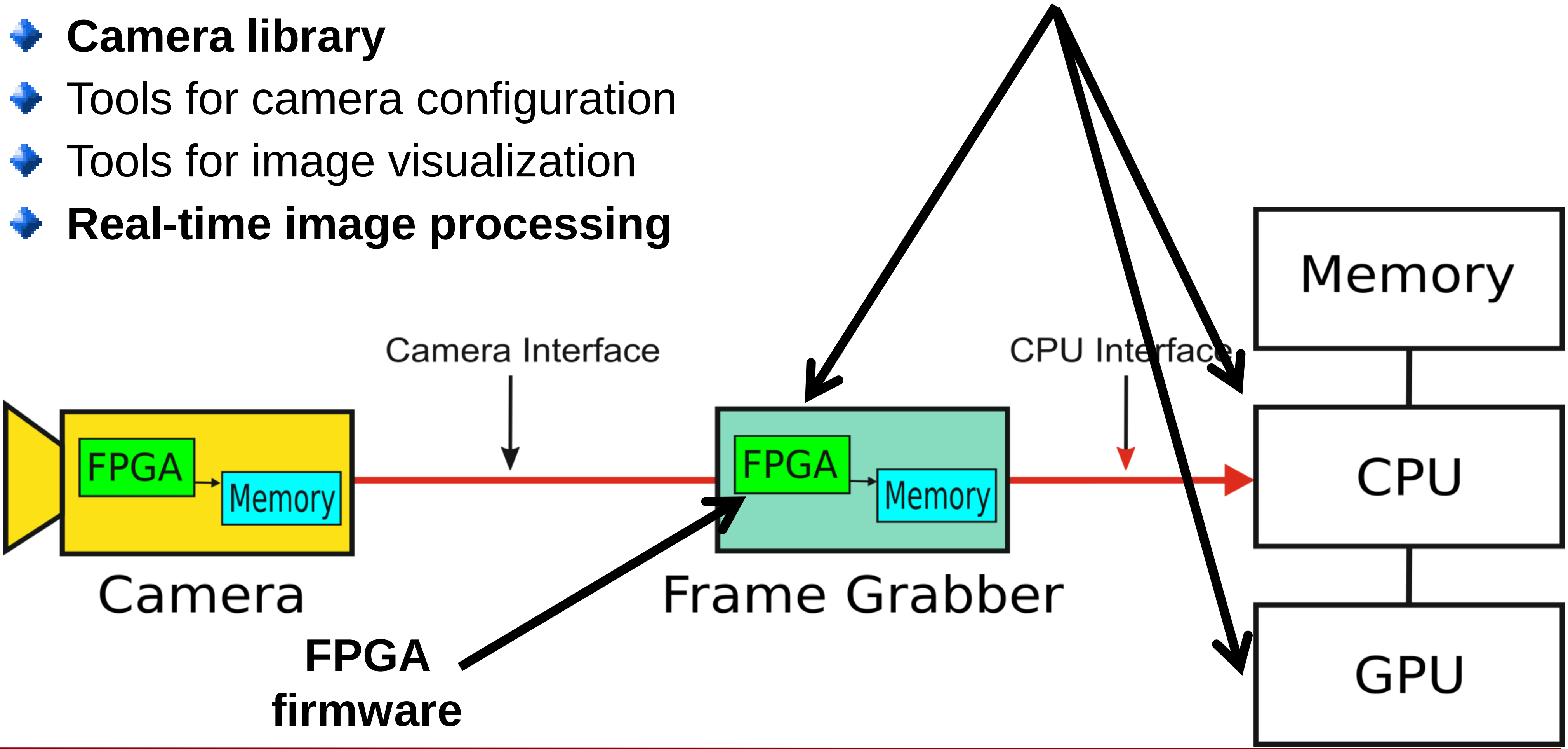
Emergent HR-12000-S-C

Dariusz Makowski (dmakow@dmcs.pl)

Universal Frame Grabber Module - Software

- **FPGA firmware**
- Linux driver
- **Frame grabber library**
- **Camera library**
- Tools for camera configuration
- Tools for image visualization
- **Real-time image processing**

Hardware we have now
Software is actually the main part of work
Working on an **universal IA and IP framework**



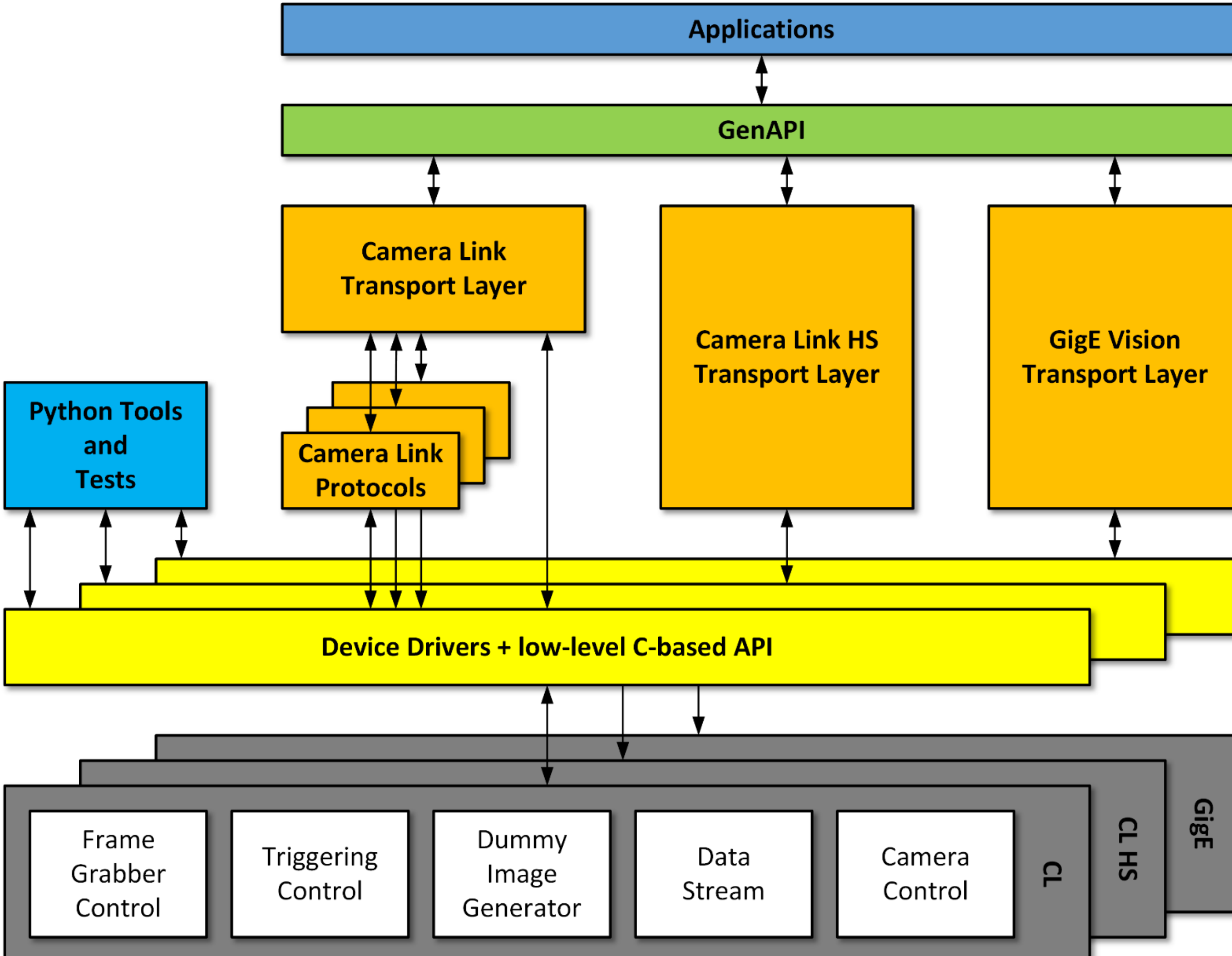
Dariusz Makowski (dmakow@dmcs.pl)

High Data Rates – Large Processing Power Required

- High-resolution/frame-rate cameras requires enormous processing power
  - **Programmable devices** (**FPGA**) are used for signal conditioning and low latency real-time processing
  - **CPU or Graphics Processing Unit** (**GPU**) are suitable for more complex algorithm, especially image processing
  - **Artificial Intelligence and Neural Network** for image analysis and recognition
- **Data copying is always expensive** (both processing power and memory)
  - Avoid data copying
  - Use Direct-Memory-Access when possible
  - **Ideal situation is a direct DMA transfer to data processing unit (GPU)**
- Use front-end FPGA for data pre-processing, filtering, decimation
- **Evaluate and reduce/optimise latency**

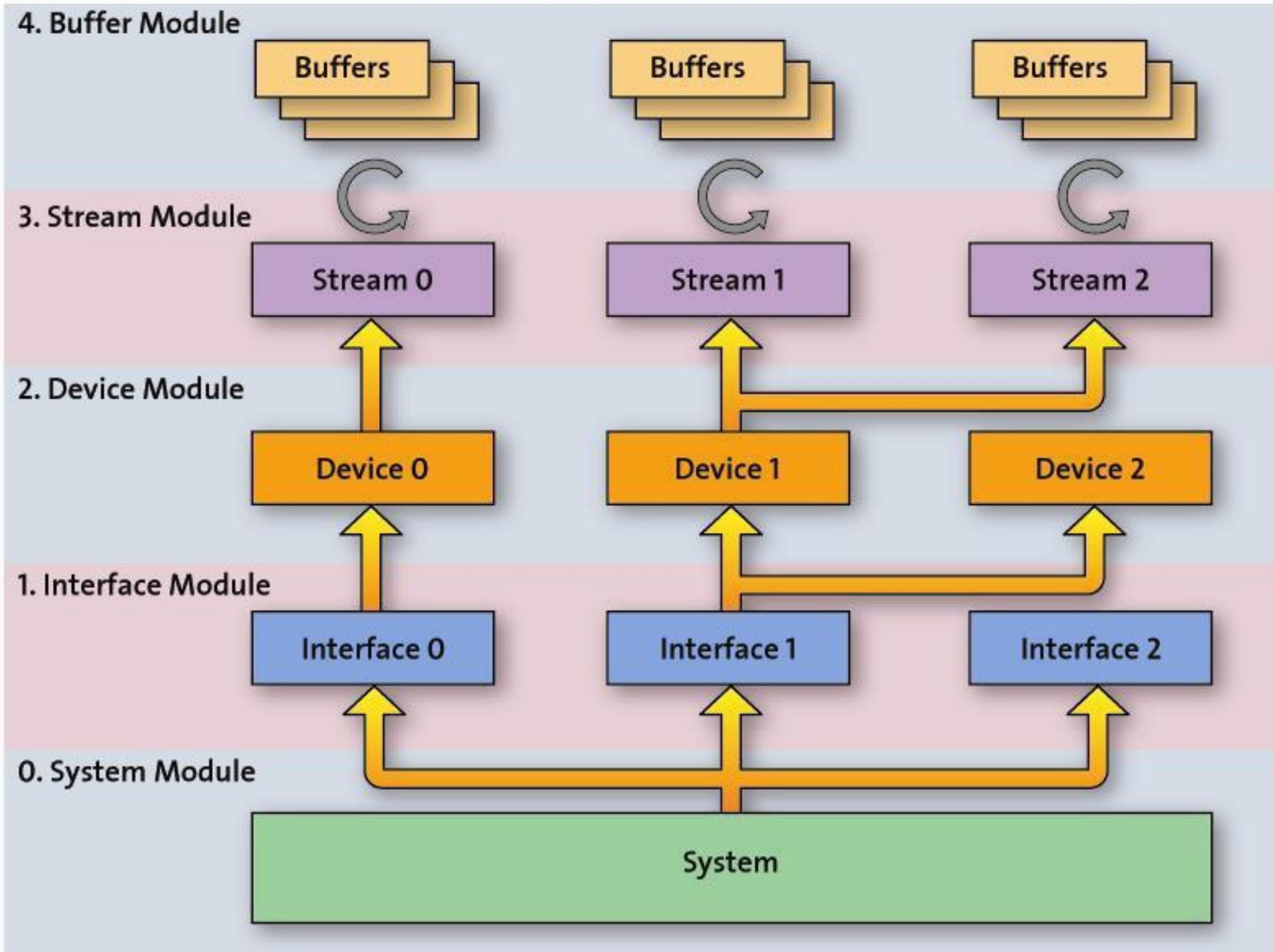
Dariusz Makowski (*dmakow@dmcs.pl*)

Image Acquisition Software Framework Based on GenICam



Dariusz Makowski (dmakow@dmcs.pl)

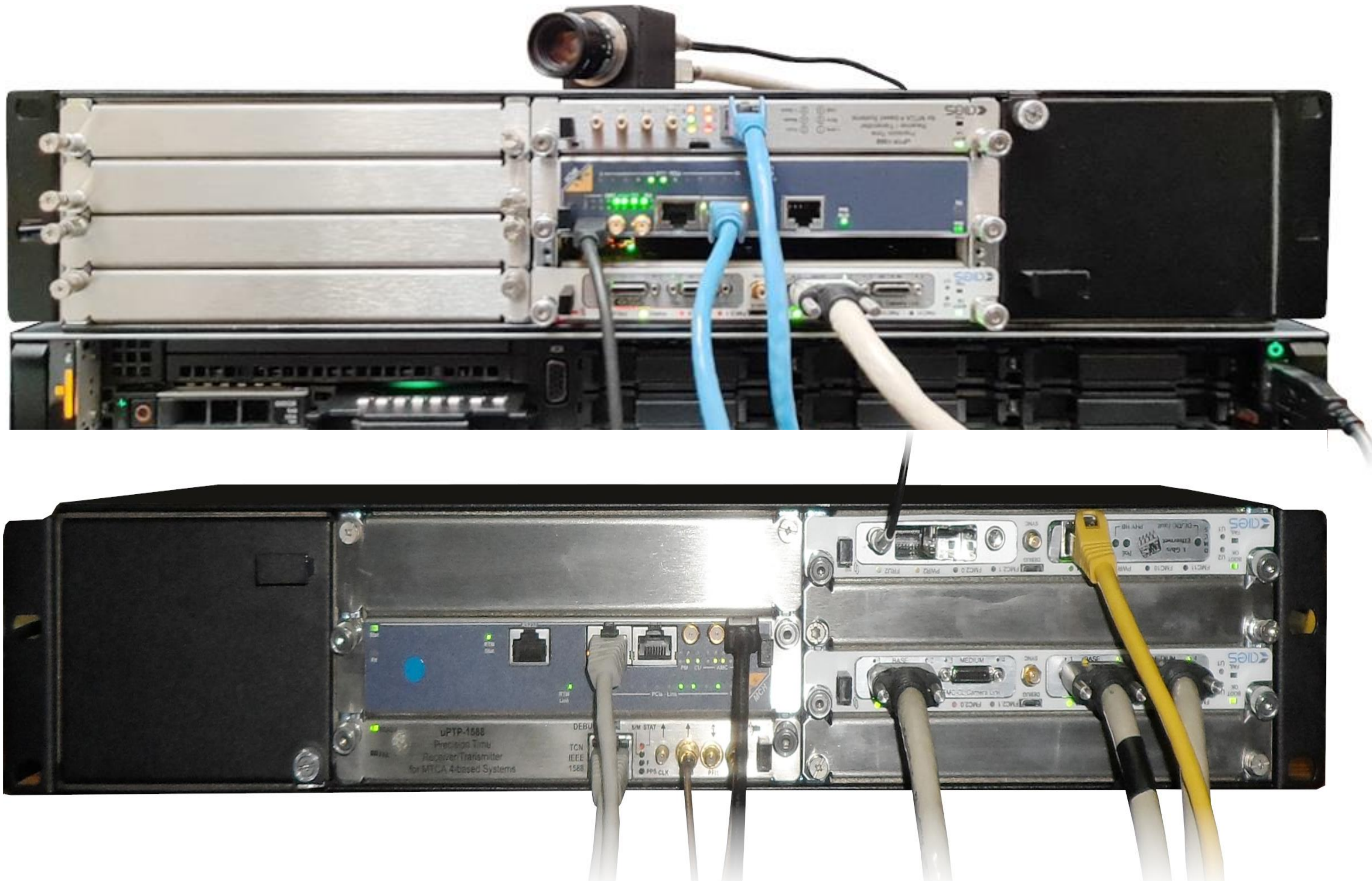
GenICam – Transport Layer



Source: GenICam GenTL Standard version 1.6

***Dariusz Makowski** (**dmakow@dmcs.pl**)*

Image Acquisition and Processing System



Dariusz Makowski (dmakow@dmcs.pl)

System at W7-X @ OP2.1 Campaign

CPU 1 (GPU)
CPU 2 (GPU)

2x Camera Link to fiber Converters

MTCA.4 Chassis
4x Frame grabber
MCH

2x Camera Link to fiber Converters

CPU 3 (GPU)
CPU 4 (GPU)

Dariusz Makowski (dmakow@dmcs.pl)

EUROfusion Engineering Grant EEG21-17

- **EEG21-17 Development of Infra-Red monitoring system using artificial intelligence techniques in view of ITER application**
- Started: June 1, 2022
- Evaluate classical (deterministic) and AI algorithms for plasma control and machine protection
- Knowledge transfer between various machines (W7-X → West)
- Research concerns also on real-time aspects

Candidate:
Bartłomiej Jabłoński

Software Engineer
Department of Microelectronics and Computer Science
Lodz University of Technology
Wólczańska 221
93-005 Łódź, Poland
Phone: +48 (42) 722 098 711
E-mail: jablonskiba@dmcs.pl

First Mentor:
Dariusz Makowski

Professor
Department of Microelectronics and Computer Science
Lodz University of Technology
Wólczańska 221
93-005 Łódź, Poland
Phone: +48 (42) 631 27 20
E-mail: dmakow@dmcs.pl

Scope of EEG21-17

1. GPU-Accelerated Real-Time Algorithms for Machine Safety

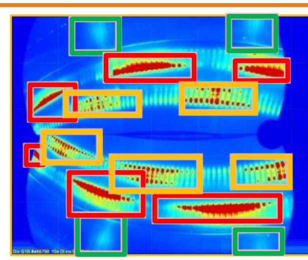
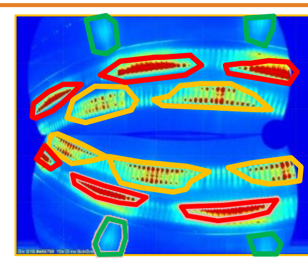
- Pre-Processing
- Real-Time Filtering
- Real-Time Analysis
- Deterministic alg.


on a GPU

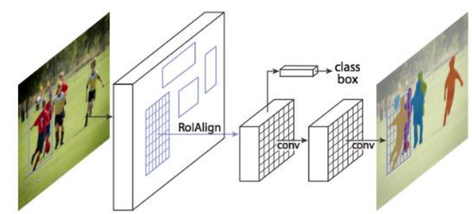
NVIDIA CUDA

2. Artificial Intelligence for Machine Protection

- Detection
- Instance Segmentation
- Classification
- NN selection for real-time processing

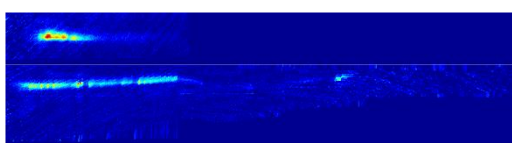



with Mask R-CNN

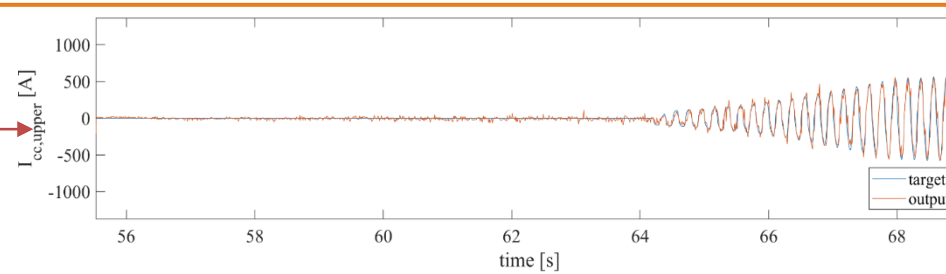


3. Artificial Intelligence for Machine Control

- Regression
- Reinforcement Learning?



CNN



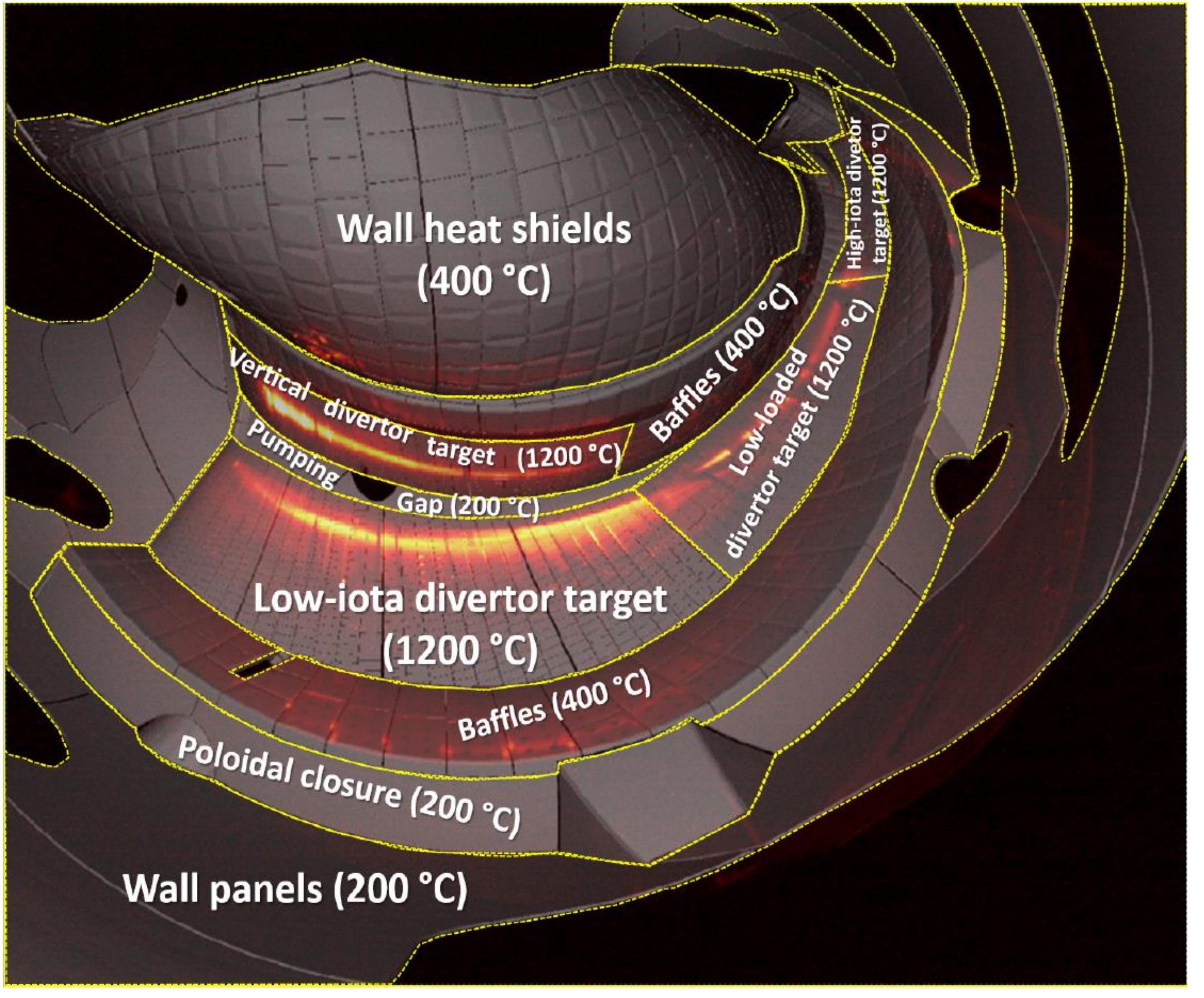
Dariusz Makowski (dmakow@dmcs.pl)

IAEA DMCS

Protection of Plasma Facing Components

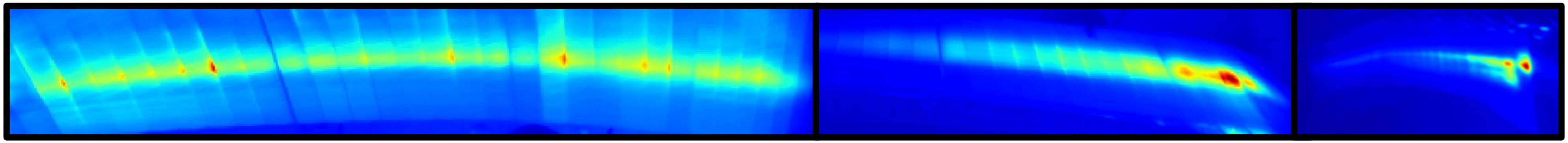
- **Divertor tiles**: Carbon Fibre Composite (CFC) joined to CuCrZr cooling structure
- Max. Operational temperature is limited by a **Cu to 475 °C**
- **Max. surface temperature** is **1200 ° C** for 10 MW/m2
- PFCs (graphite tiles) up to 400 °C
- Wall and pumping gap panels up to 200 °C

A. Puig, IAEA 2021

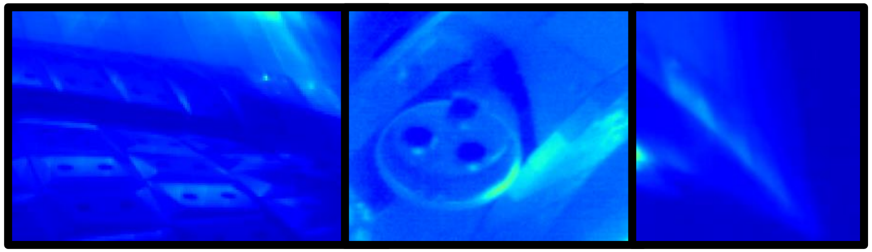


Thermal Events in Infrared Images

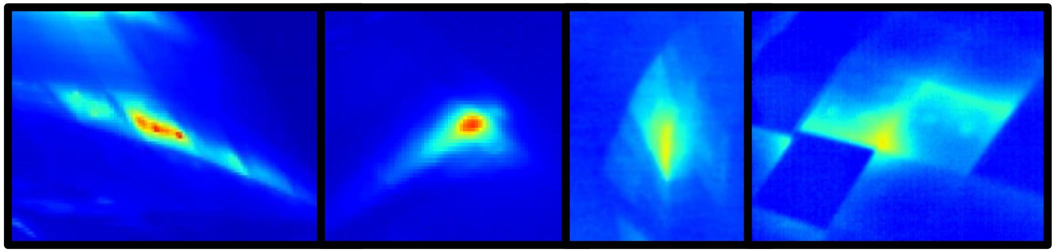
Strike-line



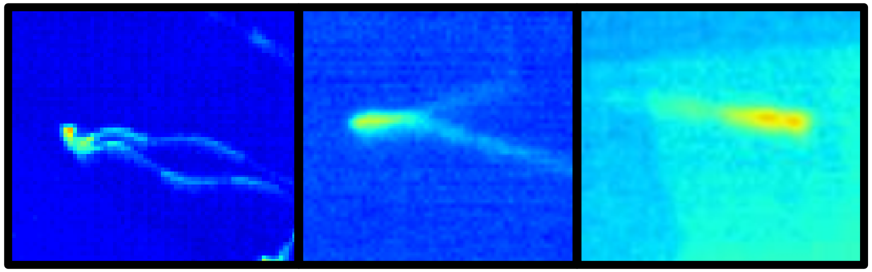
Reflection



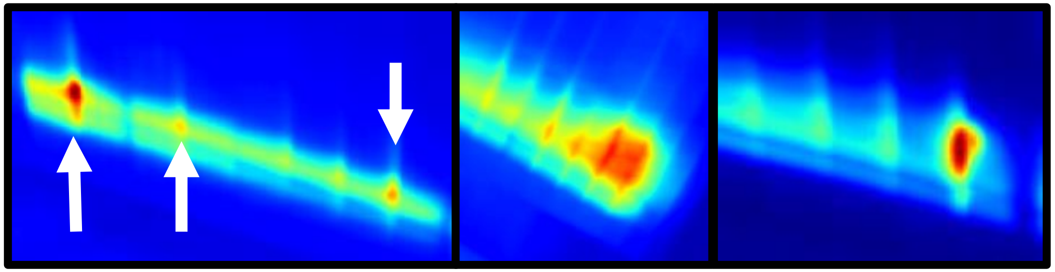
Hot-spot



UFO

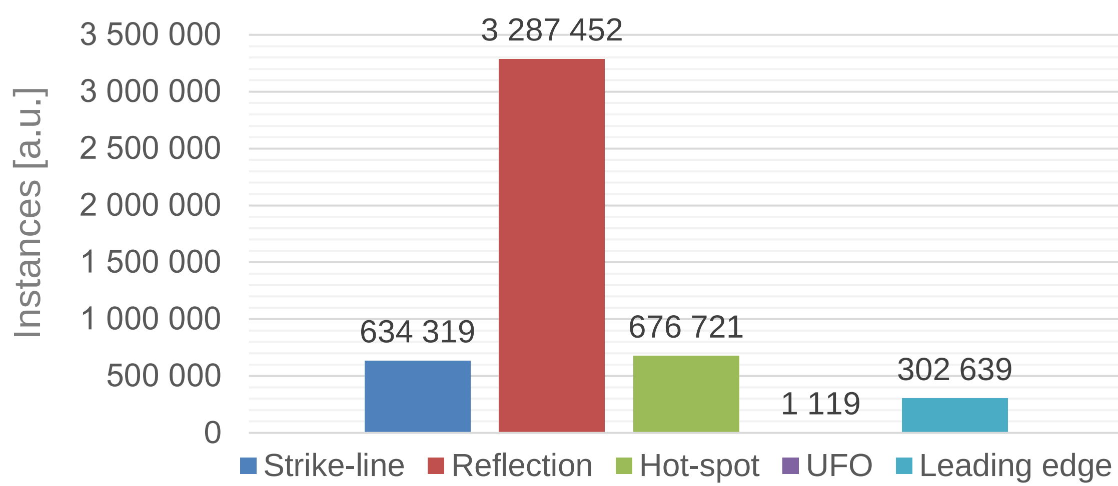


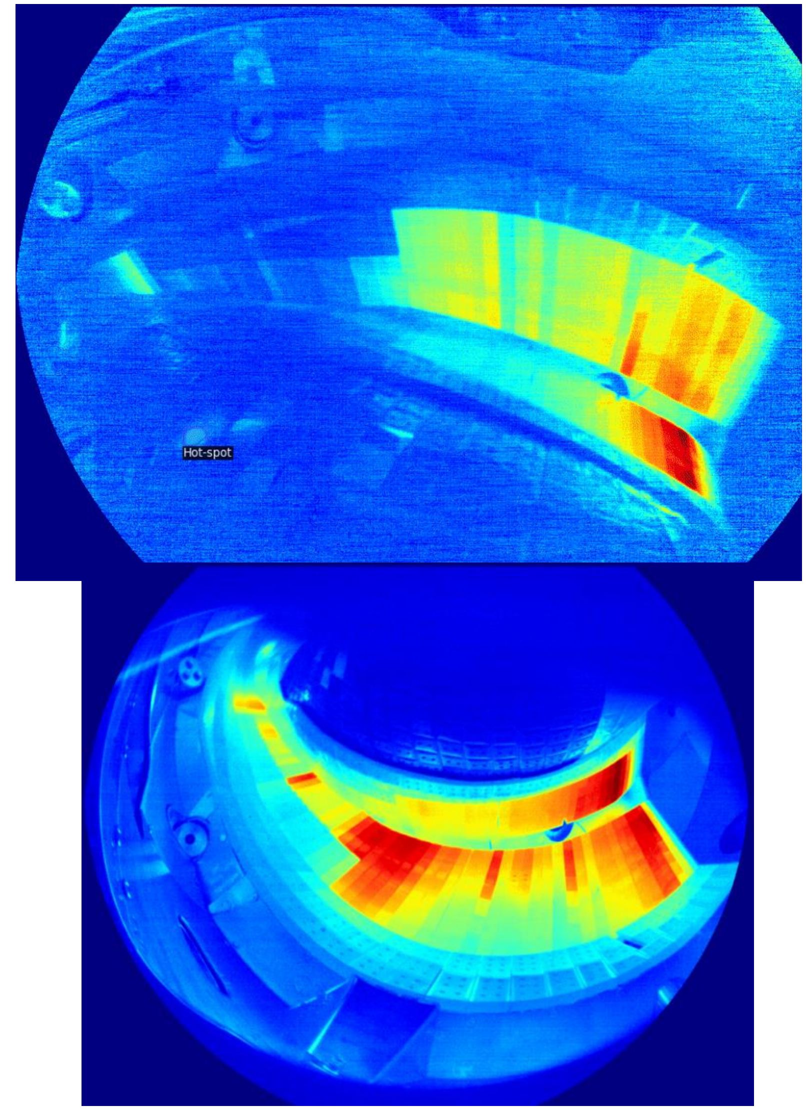
Leading edge



Annotated Dataset

- 109 training/21 test discharge sequences (**178 402** images from OP1.2)





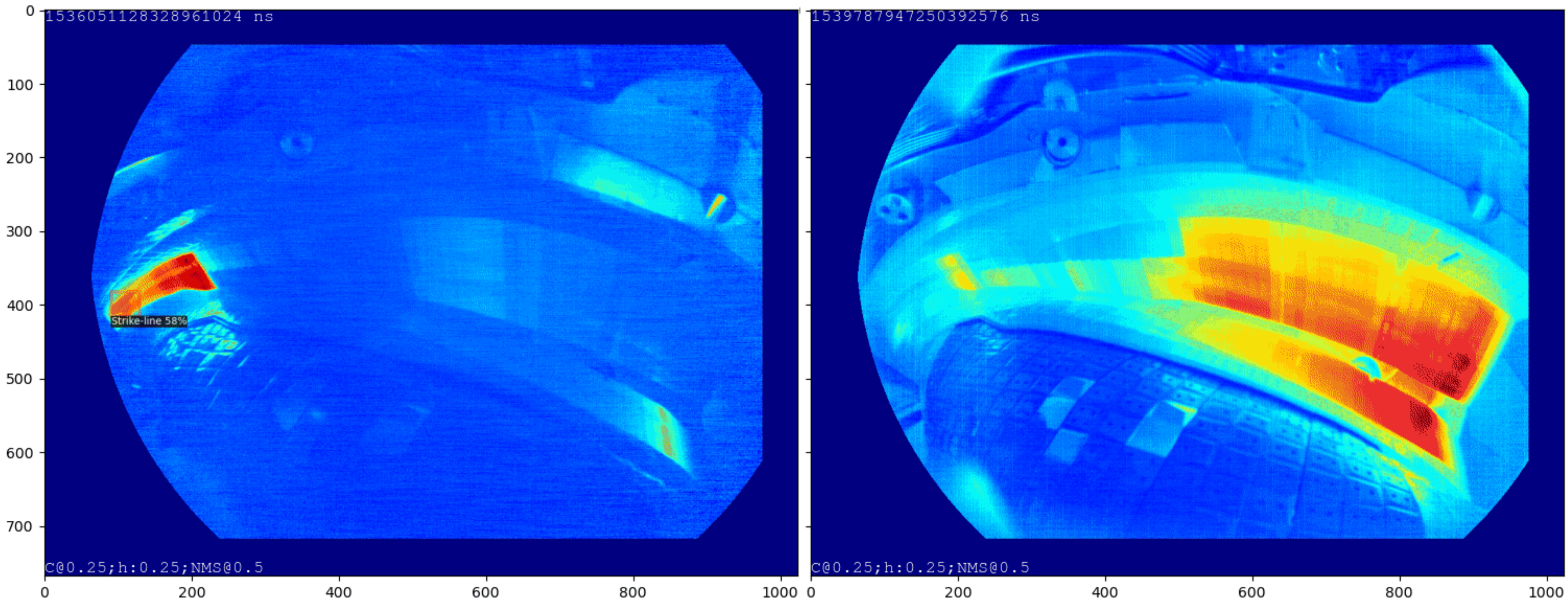
- COCO and YOLO annotation formats
- High similarity to ground truth (manual) annotations
- Annotation method described in B. Jabłoński, D. Makowski, A. Puig Sitjes, M. Jakubowski, "***Enabling Instance Segmentation: A Semi-Automatic Method for Thermal Event Annotation***", IEEE Transactions on Plasma Science (under review)

Dariusz Makowski (***dmakow@dmcs.pl***)

Instance Segmentation - Qualitative Results

20180904.007 (AEF10)
High-iota (FTM) configuration

20181017.038 (AEF10)
Standard (EJM) configuration



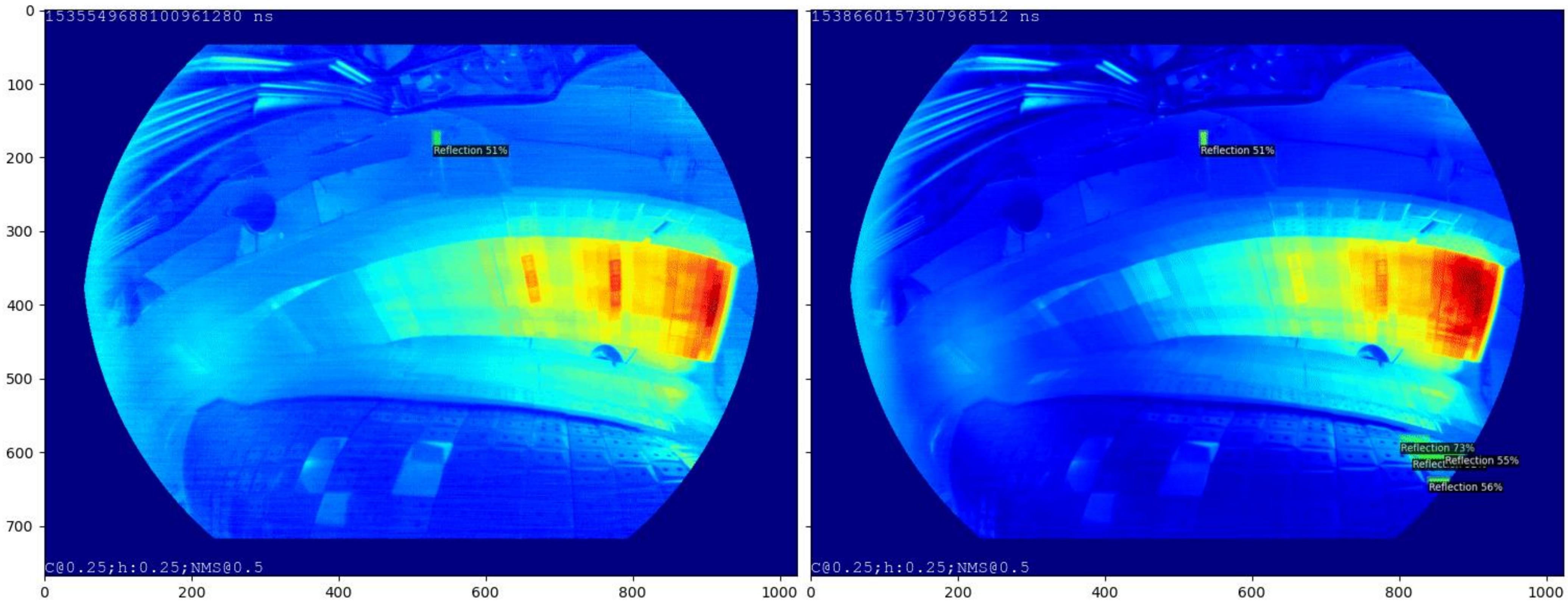
Mask R-CNN: T1 (heating start) → T4 (heating termination)
Visualize every 5th image

Dariusz Makowski (dmakow@dmcs.pl)

Instance Segmentation - Qualitative Results

20180829.040 (AEF51)
Low-iota (DBM) configuration

20181004.032 (AEF51)
Low-iota (DBM) configuration



YOLOv8: T1 (heating start) → T4 (heating termination)
Visualize every 5th image

Dariusz Makowski (dmakow@dmcs.pl)

Instance Segmentation - Quantitative Results

Model	# Params	Bounding-Box		Mask		TensorRT inference [ms] w/o pre- & post-processing
		mAP	AP@50	mAP↓	AP@50	
Mask R-CNN	45.3 M	29.89	62.92	**34.23**	66.58	-
YOLOv8 (medium)	27.2 M	**43.90**	71.10	33.20	63.50	20.76
Cascade Mask R-CNN	71.8 M	30.54	61.52	33.19	64.21	-
YOLOv8 (small)	11.8 M	41.60	68.90	31.50	62.20	**9.39**
MaskDINO (DETR)	43.8 M	22.66	54.62	25.43	62.05	-

- Smaller models with not significantly reduced performance might achieve **real-time processing**, i.e. faster than the acquisition rate: 100 Hz (10 ms)
- More annotated discharge sequences will be used for training to advance the performance
- Leading edges (few pixels) are significantly harder to detect/segment than other events; their annotations will be improved
- Transfer to different devices and experimental campaigns
- Additional data sources might be included

Dariusz Makowski (dmakow@dmcs.pl)

More About Machine Learning – see Our Poster

Multi-device dataset of infrared images for the control of thermal loads with machine learning

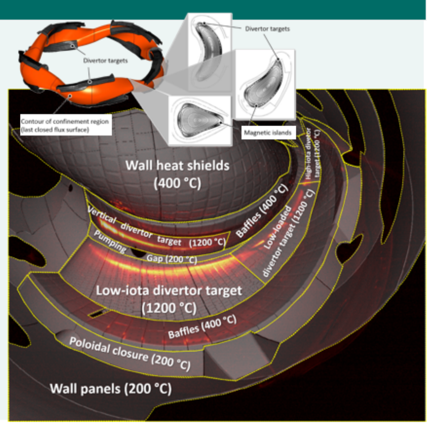
A. Puig Sitjes[1*], B. Jabłoński[2], M. Jakubowski[1], D. Makowski[2], E. Grelier[3], R. Mitteau[3], V. Moncada[3], and the W7-X Team.

[1] Max-Planck-Institut für Plasmaphysik, Germany.
[2] Lodz University of Technology, Department of Microelectronics and Computer Science, Poland.
[3] CEA, IRFM, F-13108 Saint-Paul-lez-Durance, France.
* Corresponding author: aleix.puig.sitjes@ipp.mpg.de

MAX-PLANCK-INSTITUT FÜR PLASMAPHYSIK

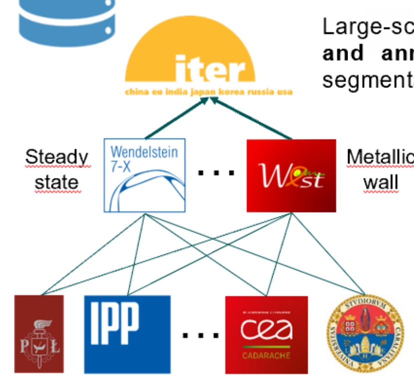
1. MOTIVATION

- **Control of the thermal loads** is required to guarantee a safe operation of high-performing fusion devices (W7-X or ITER)
- Thermal load control demands in-depth knowledge about thermal events to inform the feedback control. This is possible with **machine learning techniques**.
- ITER / DEMO necessitate thermal load **protection from day one**, leaving scarce time for gathering **sufficient data**.
- It is required to train models on current devices and **transfer large-scale models** with zero-shot learning to ITER.



2. INFRARED MULTI-DEVICE DATASET

Large-scale deep-learning models require **large diverse and annotated datasets** (very cumbersome for video segmentation) → **We propose:**



Multi-device collaboration

- A multi-device IR dataset (tokamaks and stellarators)
- Diverse first-wall materials (C and W)
- Initially, from W7-X and WEST (100s TB in HDF5 format)
- Metadata for physics, geometrical and material context information
- Semi-automated annotation in COCO format
- Compliant with FAIR principles

Summary

- Processing of images from VIS or IR cameras requires a **flexible and scalable hardware platform**
- **FPGA and GPU** could be used for executing AI-based algorithms
- Looking for **low latency solutions** working with MPx cameras
- Developed frame grabbers supporting various camera interfaces
- Developed universal software framework based on GenICam
- Developed a dedicated solution based on **NVIDIA GPUDirect RDMA** solution that significantly reduces the total image processing latency and releases the CPU
- Working on real-time AI algorithms and design methodology

Thank you for your attention

Dariusz Makowski (*dmakow@dmcs.pl)*