MIRCOTCA在中微子实验和同步辐射光源探测器中的应用

张杰 (中国科学院高能物理研究所)

第三届地下和空间粒子物理与宇宙物理前沿问题研讨会

2024-05-08

WHAT ABOUT FPGAS FOR PHYSICS EXPERIMENTS?

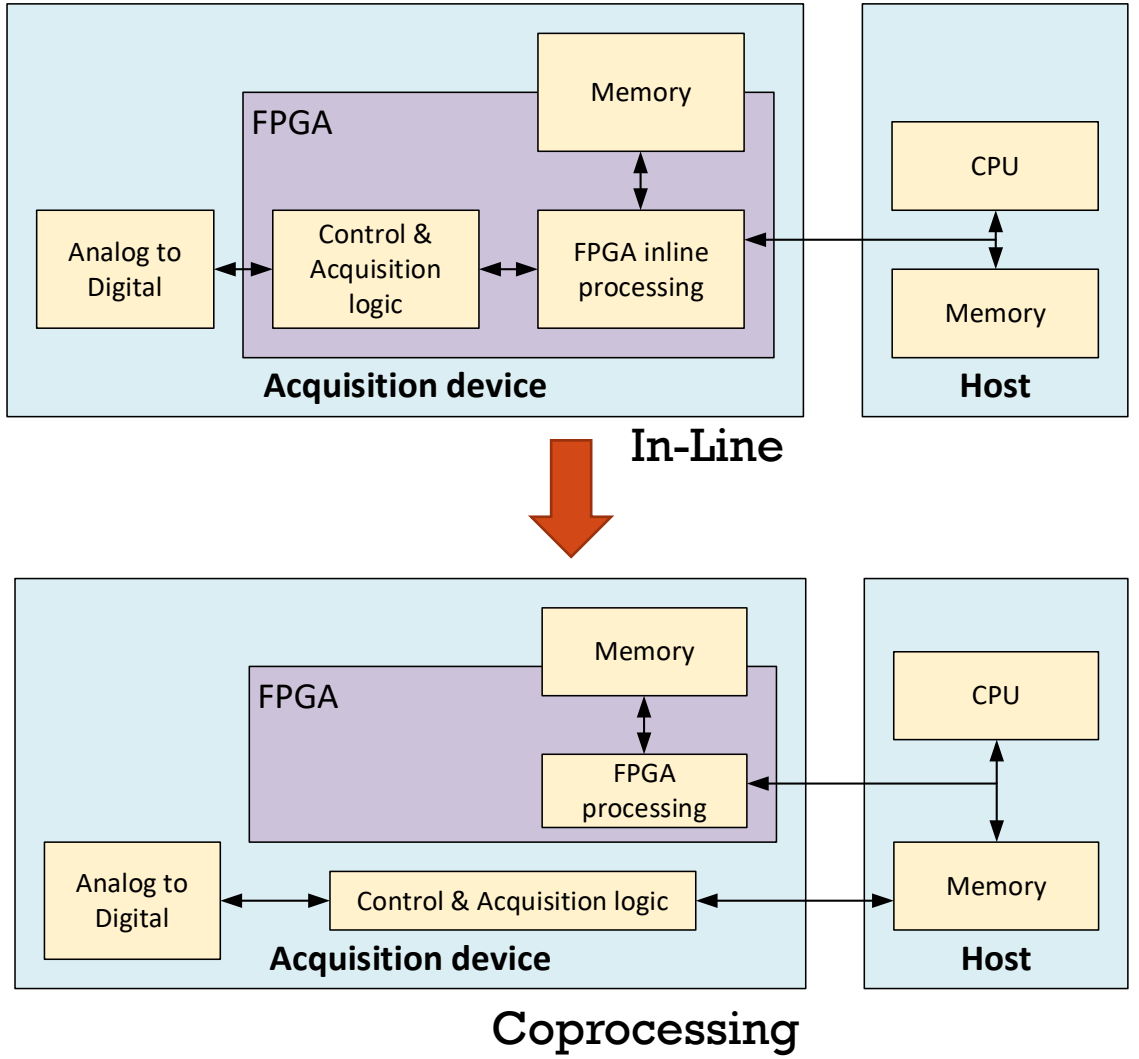
Workloads for FPGA:

- Front-end Electronics Control
  - Fast Control
  - Slow Control
- Clock Synchronization
- Monitoring
- Data Acquisition
  - Signal processing, filtering

Trends:

- Machine Learning (Deep Learning)
- Data Analytics

Source: Deeper, Faster Learning with FPGA Co-Processors

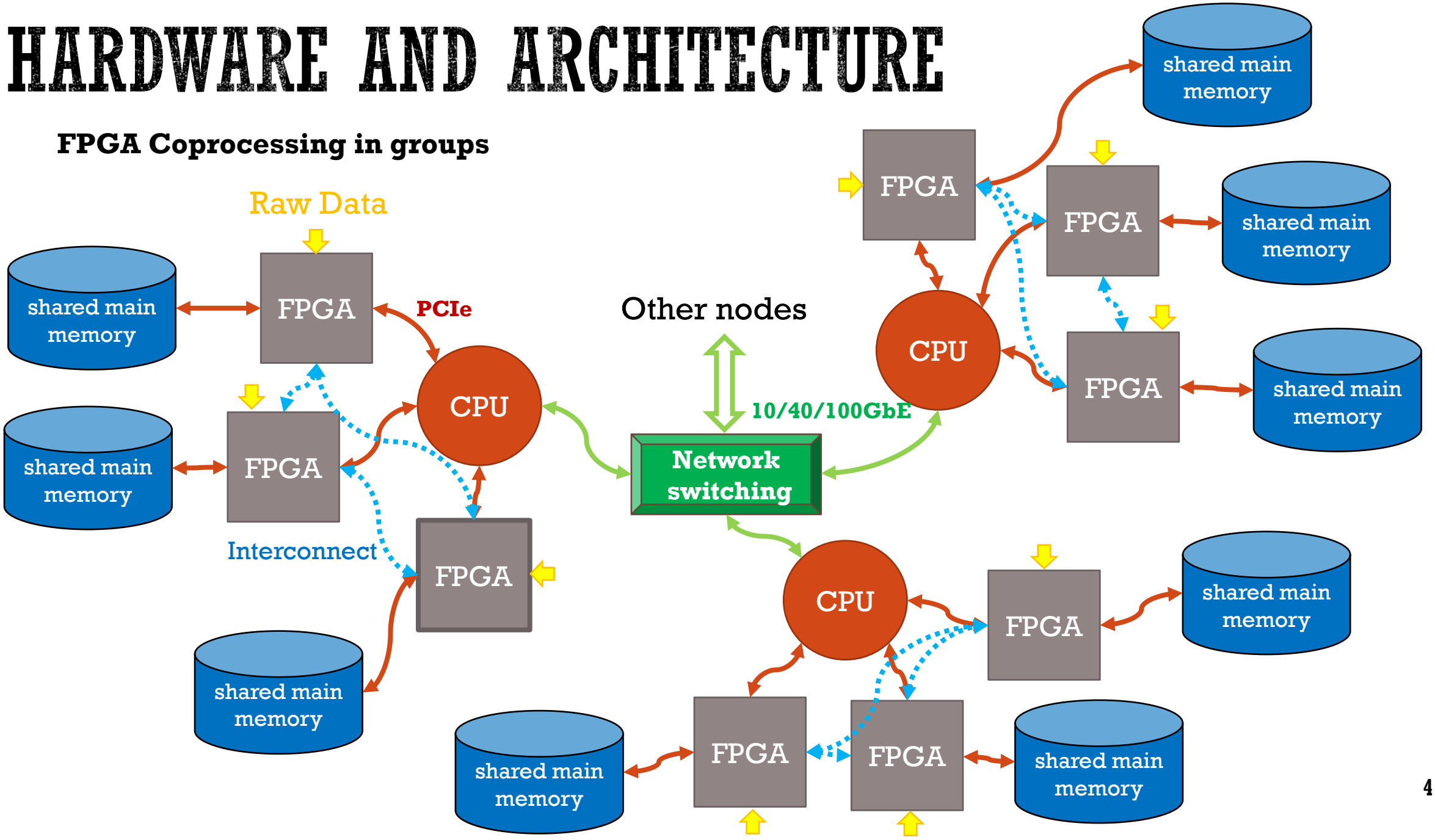


MOTIVATION

- The data volume of physics experiments need to reduce, PB/year to ?
- Finding new physics requires massive increase of processing power, much more flexible algorithms in software and much faster interconnects

HARDWARE AND ARCHITECTURE

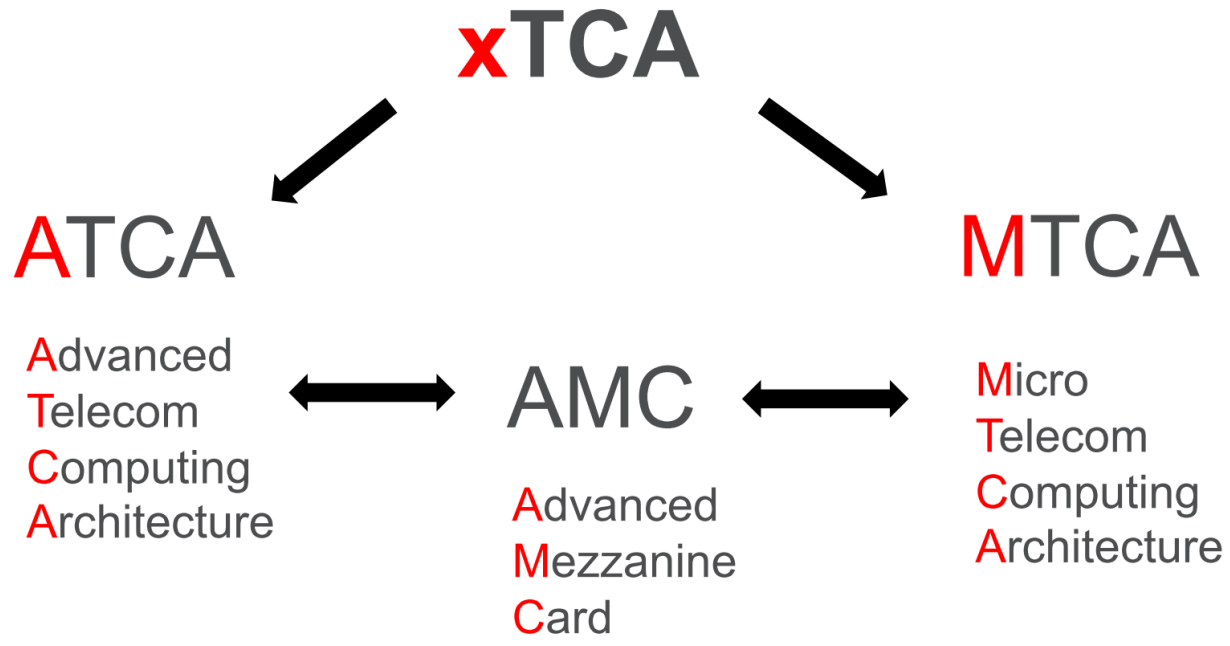
FPGA Coprocessing in groups



WHAT IS XTCA?

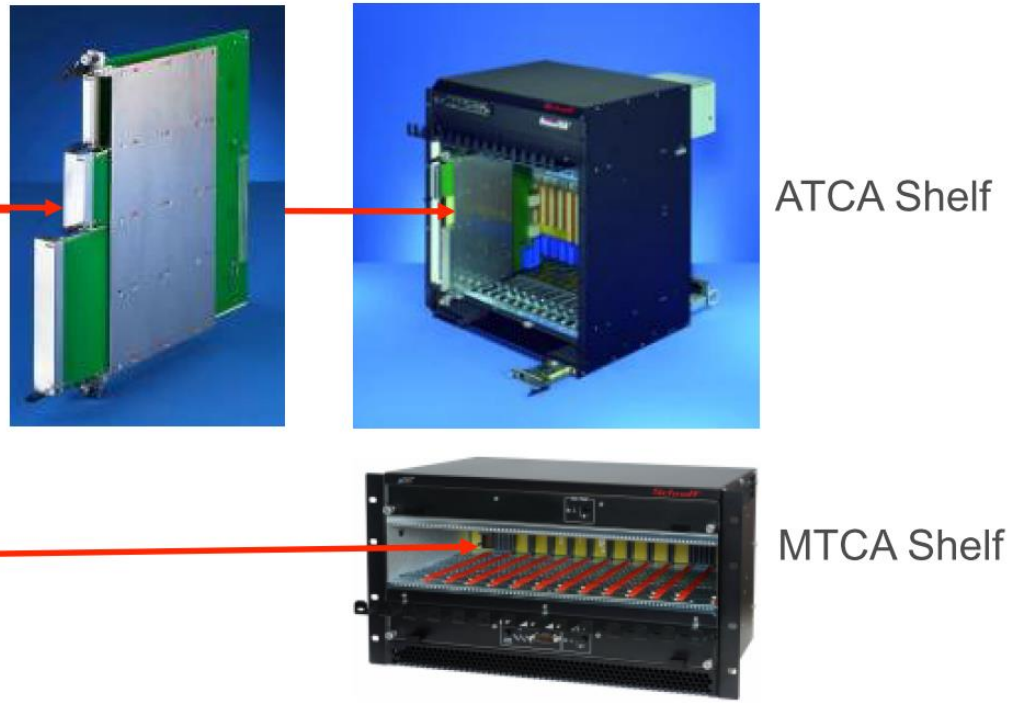
The dimension of a xTCA crate is depending on:

- Numbers and sizes of slots
- Cooling concept
- Heat dissipation
- Request for redundancy



AMC modules

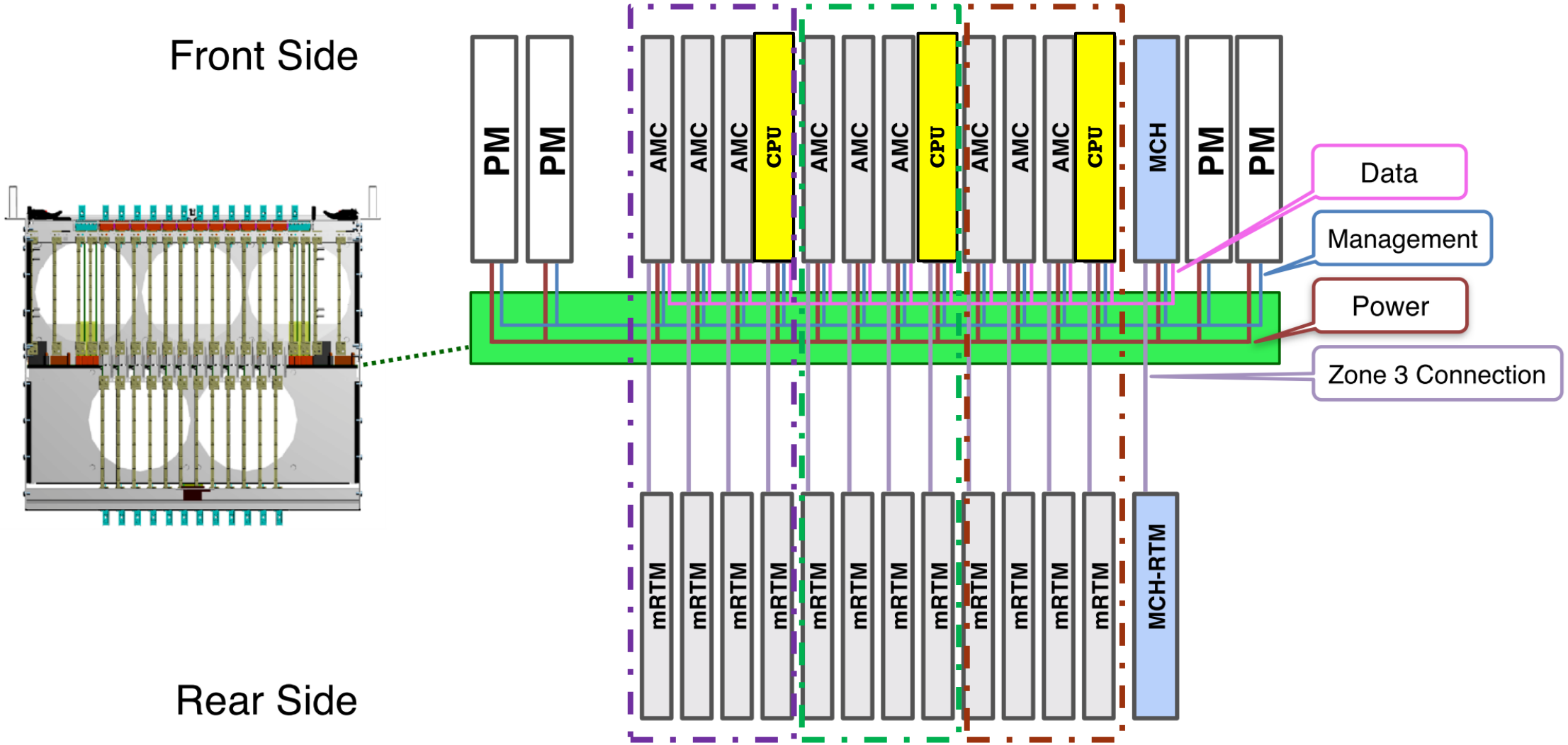
- Fully integrated into the ATCA IPMI management structure
- Hot Swap capability



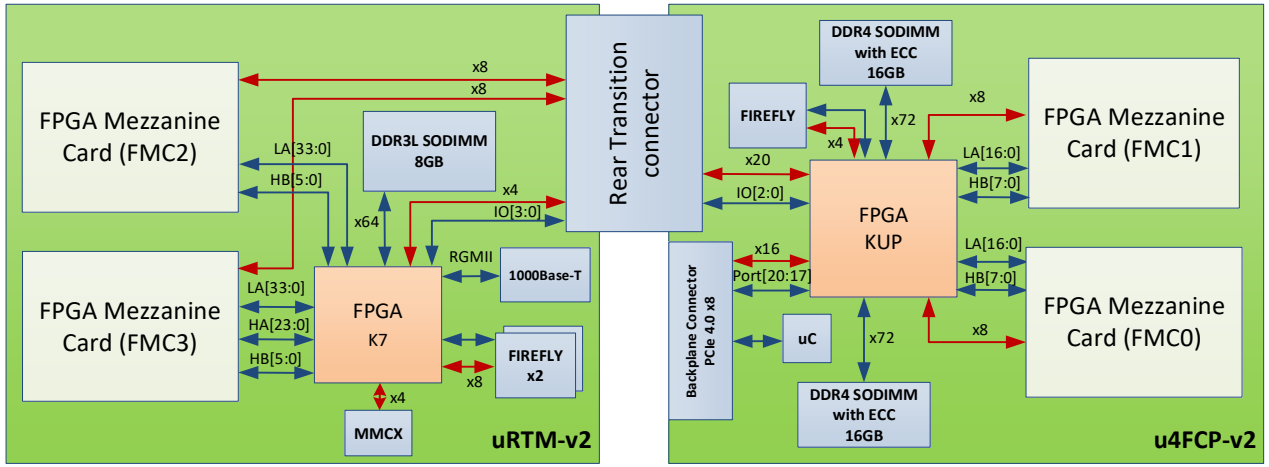
Source: N.A.T.

MICROTCA.4 GROUPING

- Divided into three groups, in each group:
  - 3 AMCs with RTM
    - FPGA boards
  - 1 CPU board
    - 4-core Xeon Processor E3-1505M (3GHz)
      - ConCurrent Technologies AM G64/472-51
      - Vadatech AMC725



FPGA BOARD (AMC+RTM)



- uRTM-v2
- u4FCP-v1

u4FCP & uRTM:

FPGA-based MicroTCA compatible AMC board

- For generic system control and data acquisition in HEP/HEPS experiments
- Conceived to serve a mid-sized system residing either
  - inside a **MicroTCA** crate or
  - **stand-alone** on desktop with high-speed optical links or Ethernet to PC
- HPC FMC sockets
  - Provide additional clock signals, user-specific I/O and high-speed transceivers that can be used to extend the connectivity as well as the I/O bandwidth
- The red lines are high-speed serial links connected to the gigabyte transceivers (GTY/GTH/GTX) of the FPGA. The blue lines are the general input/ outputs connected to the High Performance (HP), High Range (HR) or High Density (HD) banks of the FPGA.
- More details:
  - https://github.com/palzhj/u4FCPv2

FEASIBILITY

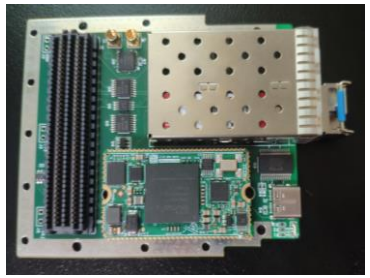
- AMC+RTM boards
  - With various FMC cards

FPGA node

FPGA node

FPGA node

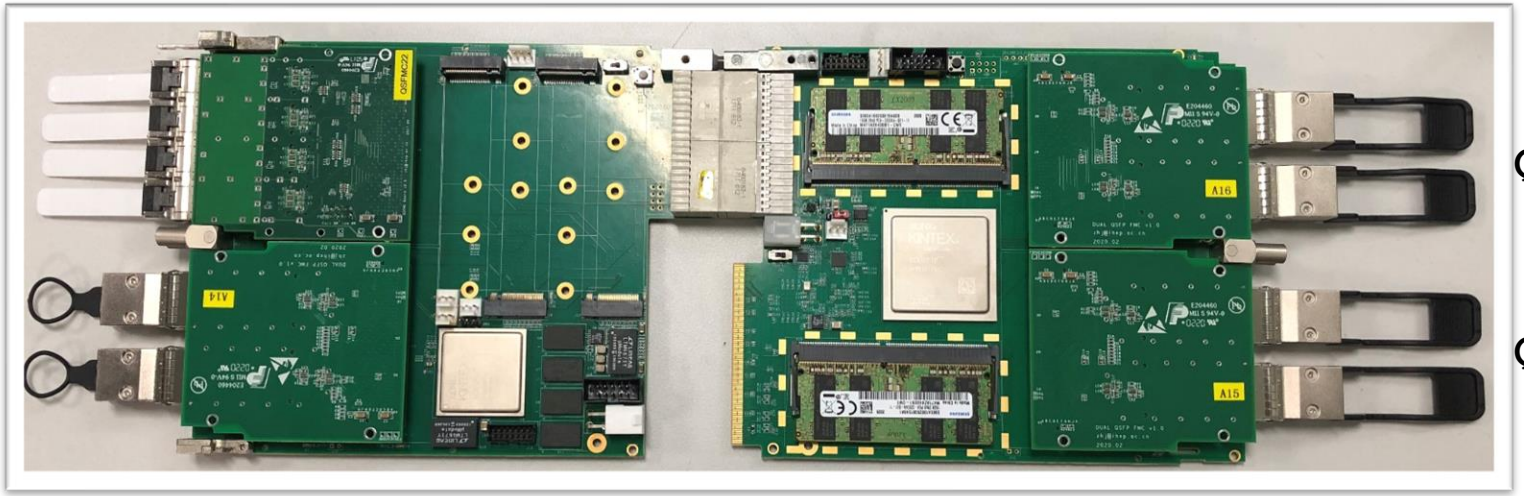
FPGA node



Mini-WR FMC

SFP+ x4

QSFP28 x2

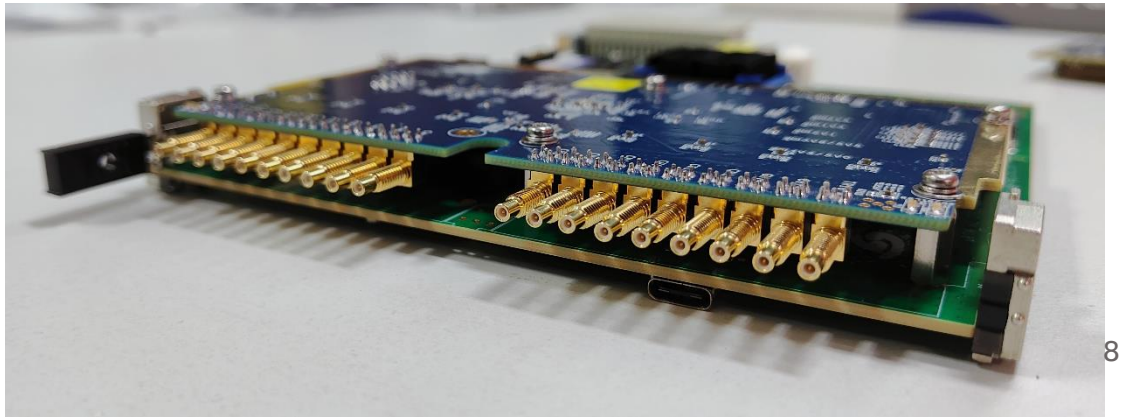


QSFP28 x2

QSFP28 x2

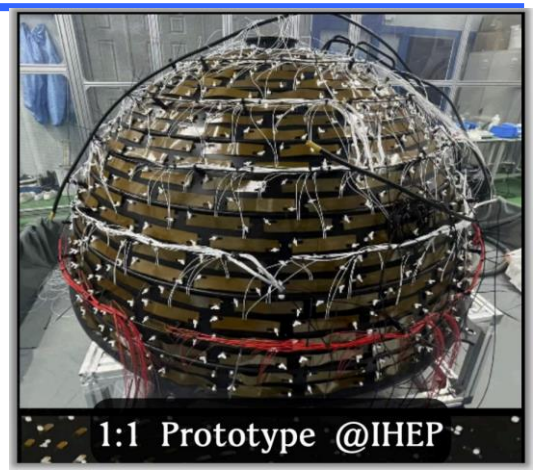
ADC FMC board

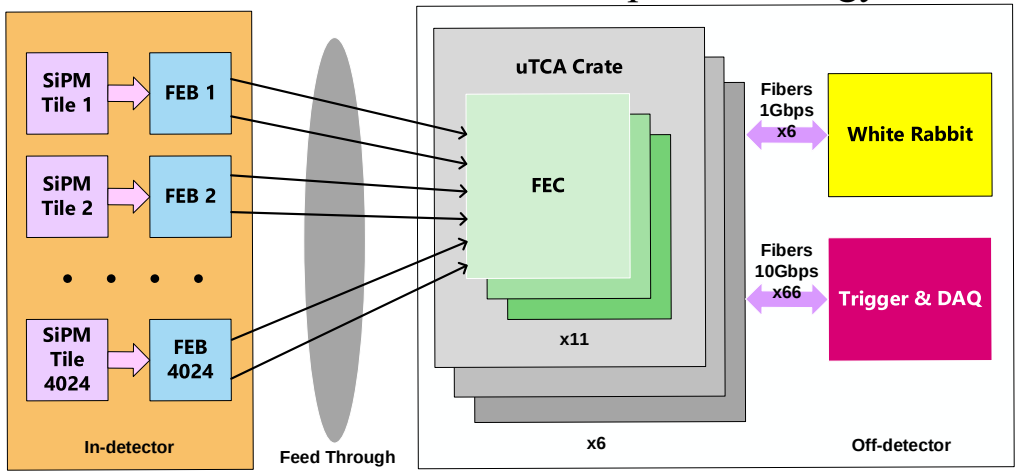
- 16 single-end channels
- 125 MHz analog bandwidth
- DC coupled analog input
- 12/16-bit Σ-Δ ADC
- Raw sample rate up to 2 GSps

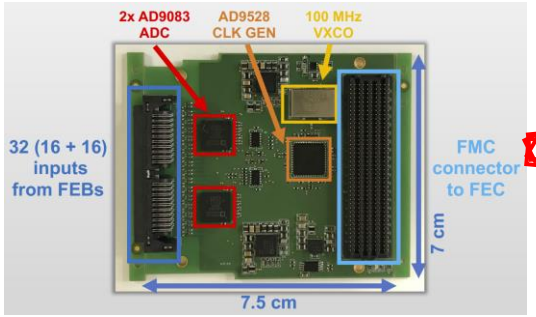


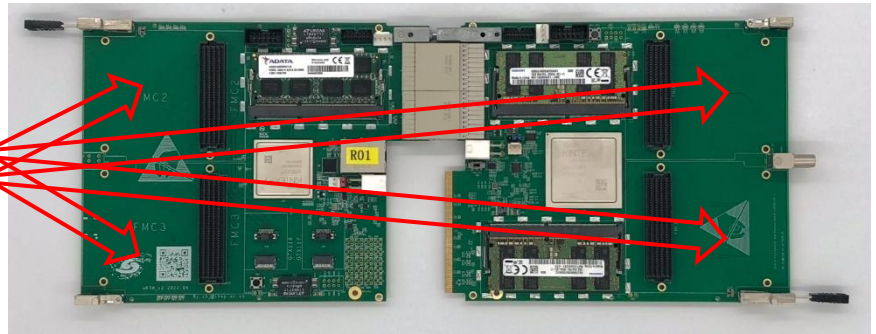
Applications – JUNO TAO Experiment

- The **Taishan Antineutrino Observatory (TAO)** is a satellite experiment of the **Jiangmen Underground Neutrino Observatory (JUNO)**, located in the southern China, expected to start collecting data in 2024.
- TAO consists of a spherical ton-level Gadolinium-doped **Liquid Scintillator** (Gd-LS) detector (1.8 m diameter) at ~30 m from a reactor core of the Taishan Nuclear Power Plant (4.6 GW) in Guangdong.
- By means of 10 m^2 SiPM covering the spherical LS, the **reactor antineutrino spectrum** will be measured with a sub-percent energy resolution ($\leq 2\% / \sqrt{E}$ MeV).









- **Electronics**
  - **In-detector**
    - **Discrete readout (ROMA TRE): 2 channels/SiPM tile**
  - **Off-detector**
    - **6 uTCA.4 crates**
      - **Each crate has 12 slots and will be mounted with 11 Front-end Electronics Controller (FEC) boards**
        - **9 FEC boards with 4 ADC FMC cards**
          - **Each ADC FMC card supports 32 channels (ROMA TRE)**
        - **2 FEC board with 3 ADC FMC cards and 1 WR FMC card**
        - **1 spare slot is reserved for veto/ redundancy/backup/debugging**

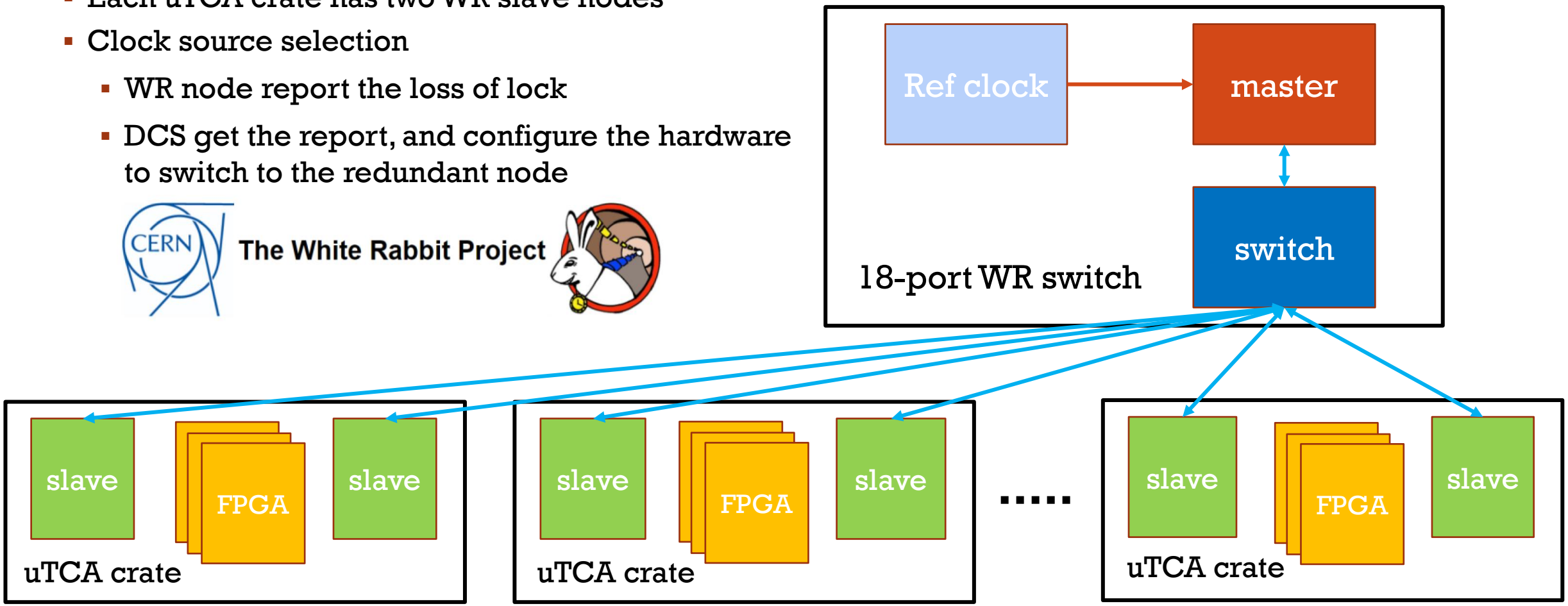
Total: up to 8064 ADC channels

- **Each crate supports up to 1344 ADC channels**
  - **6 x (9 x 4 x 32 + 2 x 3 x 32) = 6 x 1344 = 8064**
- **Note: project requirement is 8048 ADC channels**

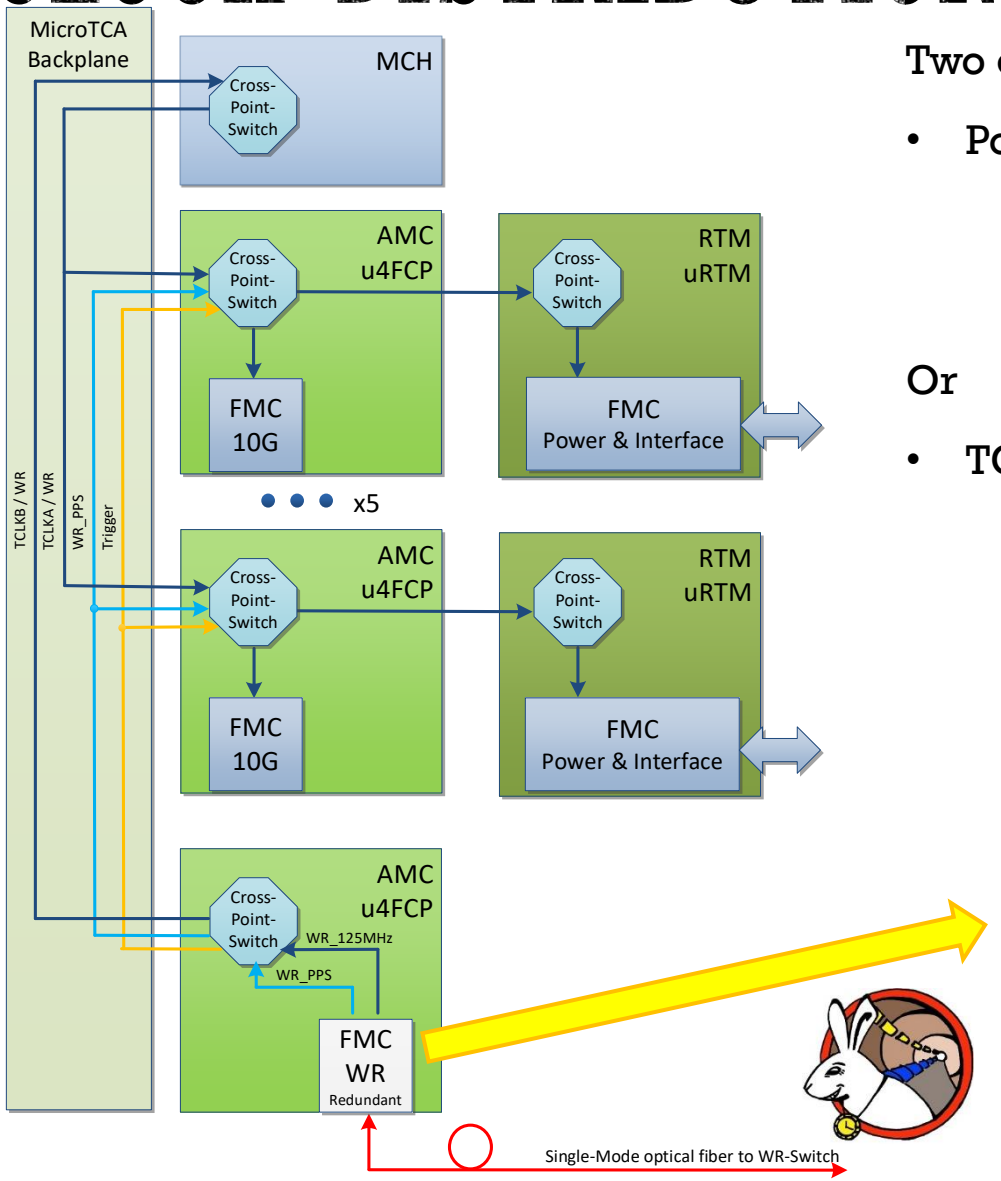
CLOCK REDUNDANCY CONSIDERATION

- Due to the life limit of fiber optic transceivers, we plan to use a redundant WR clock
- Each uTCA crate has two WR slave nodes
- Clock source selection
  - WR node report the loss of lock
  - DCS get the report, and configure the hardware to switch to the redundant node

CERN The White Rabbit Project

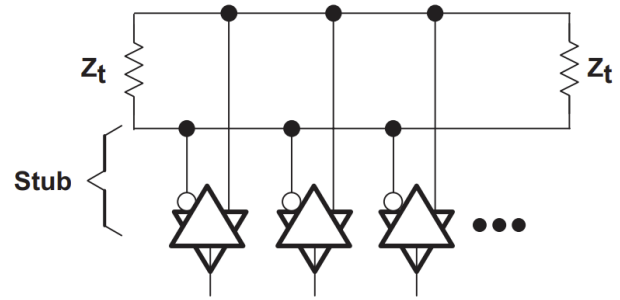


CLOCK DISTRIBUTION IN MICROTCA



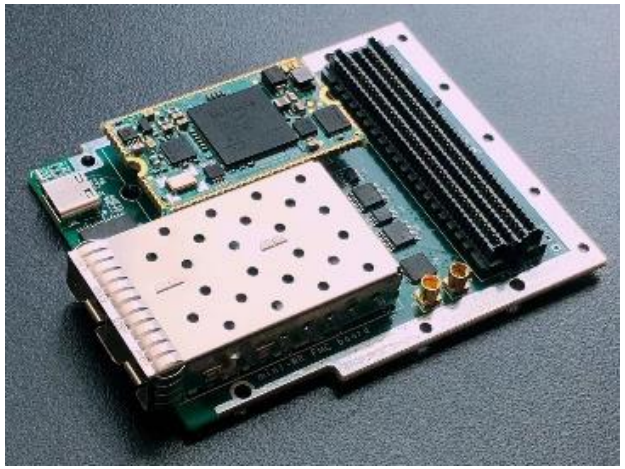
Two clock allocation schemes inside the crate

- Port 17 ?
  - Native R9 crate with WR Support
  - MLVDS multi-drop connection



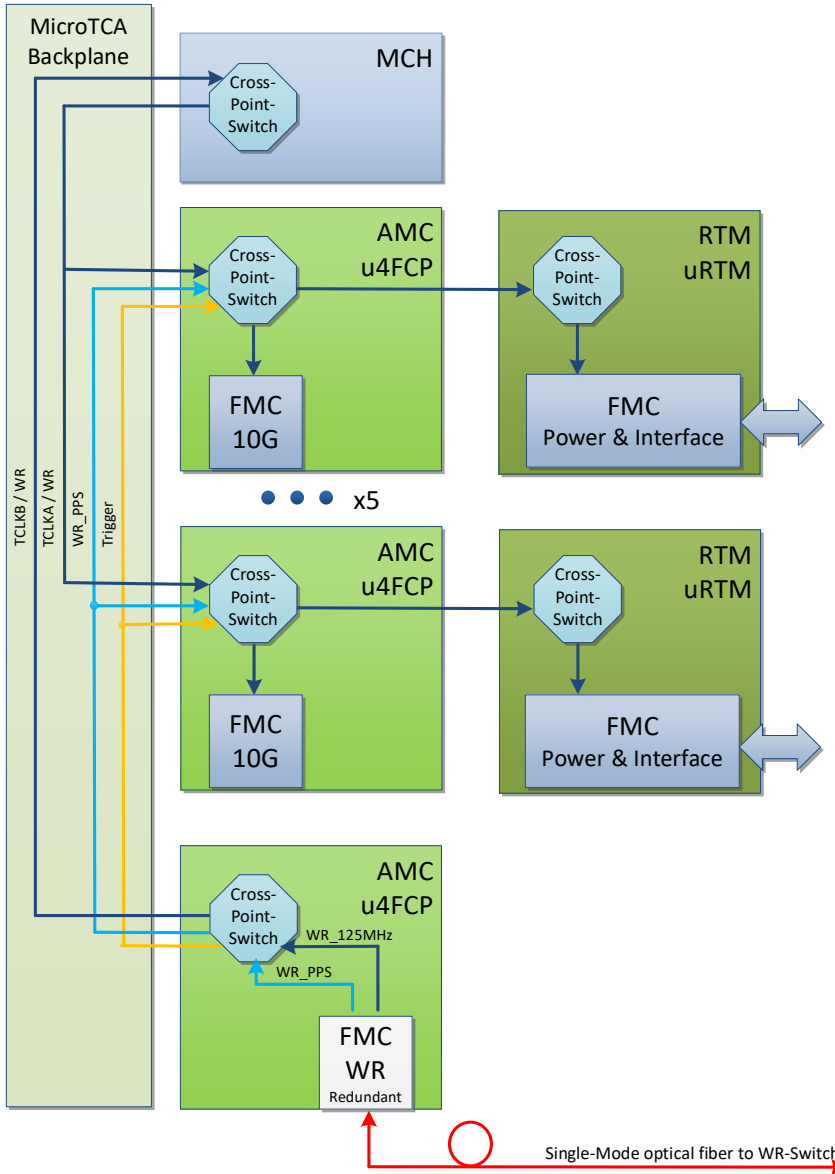
Or

- TCLK ?
  - TCLKB (slot12)-> MCH (NAT-MCH-PHYS80) -> TCLKA (slot1-11)

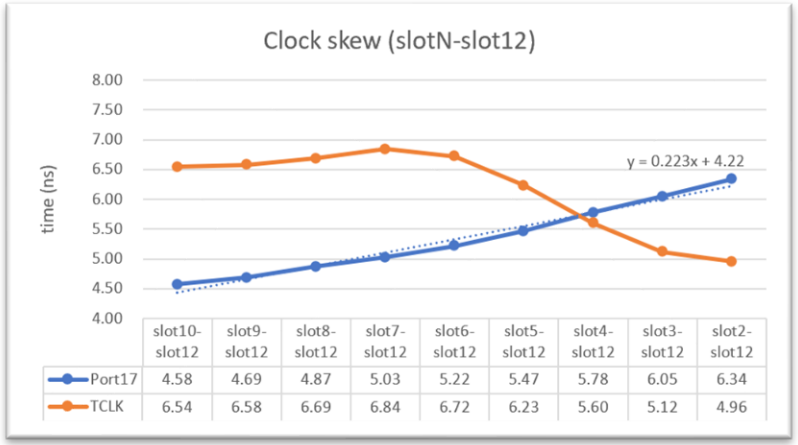


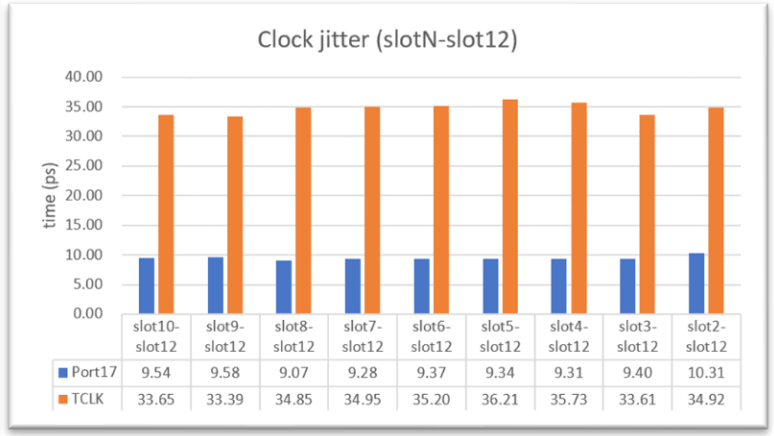
Mini-WR FMC for System-level clock synchronization

CLOCK DISTRIBUTION IN MICROTCA

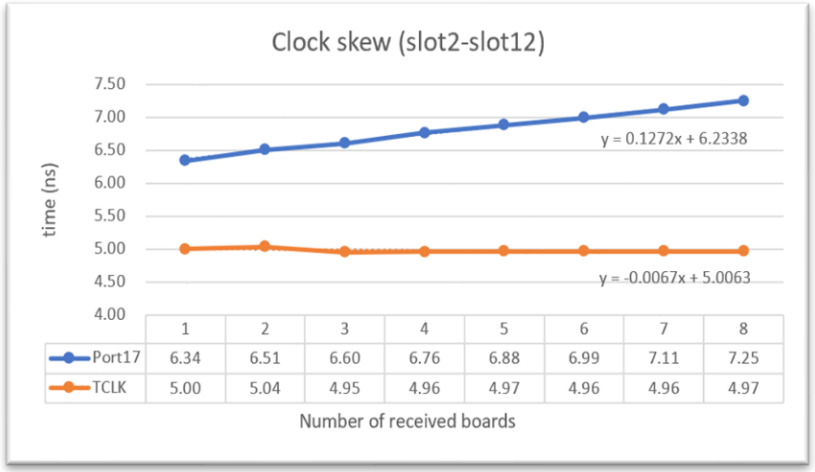


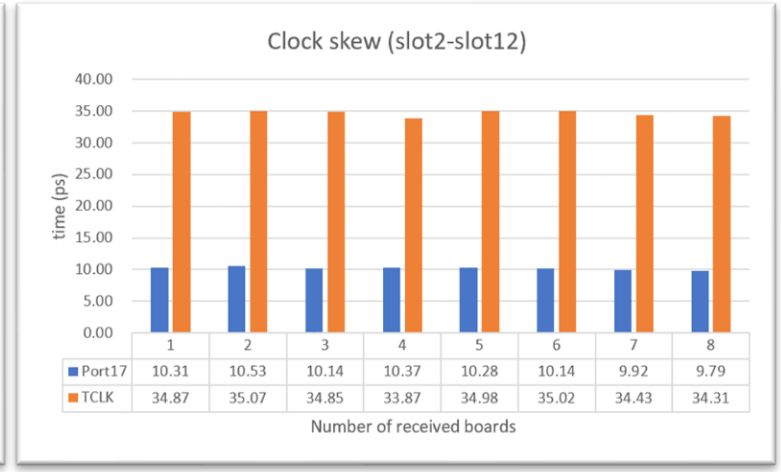
AMC with mini-WR fixed in slot 12, move AMC receiver





Receiver is fixed at slot 2, add more adjacent AMC boards



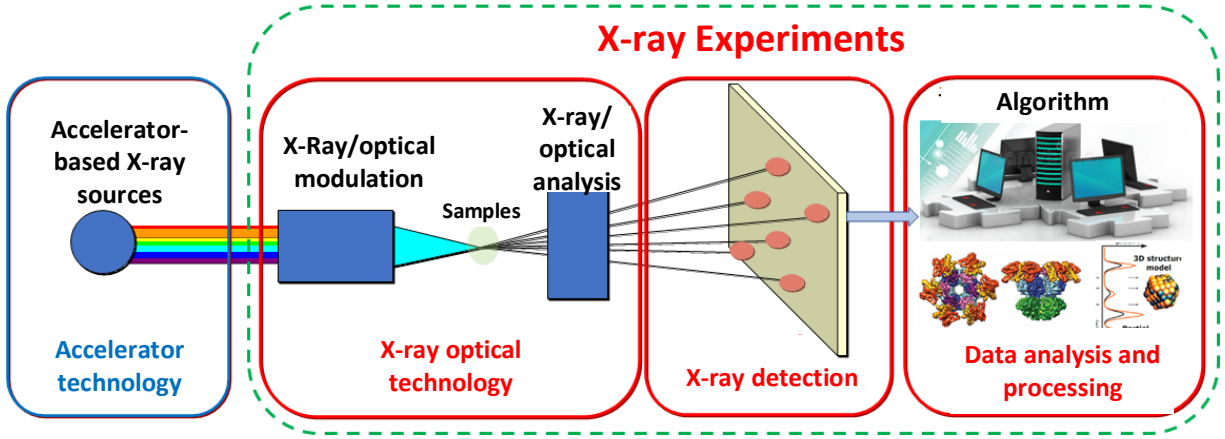


- TCLK scheme has deterministic skew, but more jitter
- Port17 scheme has smaller jitter, but the skew is related to the location and receiver quantity

High Energy Photon Source (HEPS)

- Fourth-generation synchrotron light source
- Under construction in Huairou District, Beijing
  - Start the user operation in 2026
- Key-parameters for beam

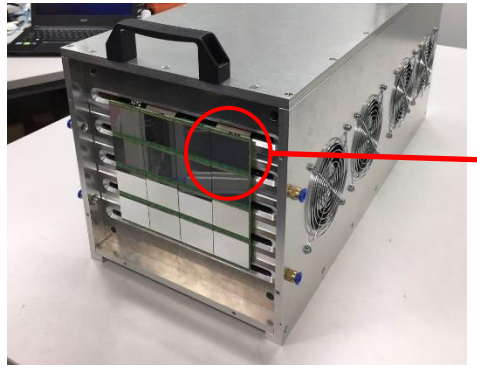
Parameters	Nominal
Beam energy	6.0 GeV
Emittance	better than 0.06nm×rad
Beam	Higher than 1×1022 phs/s/mm2/mrad2/0.1%BW
Spatial resolution	10 nm
Energy resolution	1 meV
Photon energy	Up to 300keV

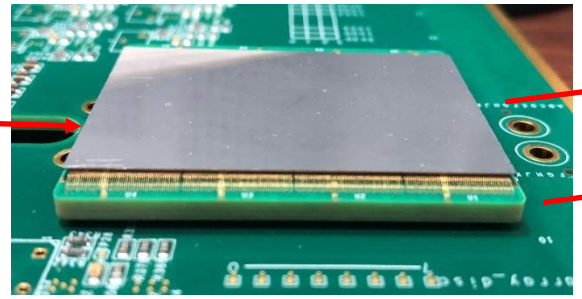


- HEPS Phase I plans 14 beamlines with the area array pixel detector in 2025
  - Detectable energy range: 8-20 keV
  - Spatial resolution: 140 μm
  - Total pixels: 6M, 2M, 1M, & 150K
  - Frame rate: Up to 2 kHz

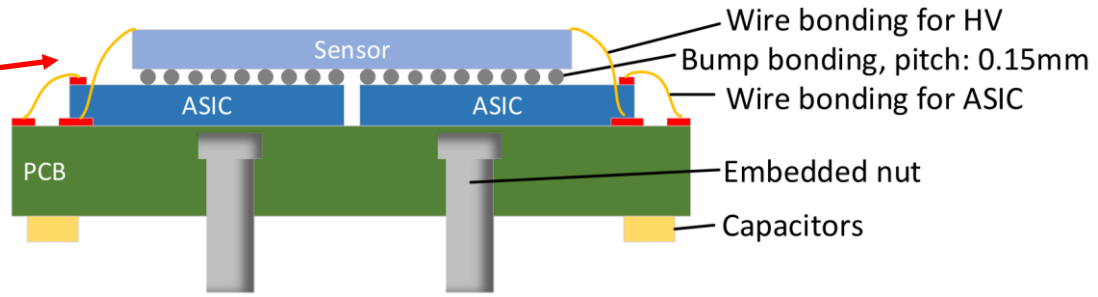
- Ref: http://english.ihep.cas.cn/heps/index.html

Hybrid Pixel Detector

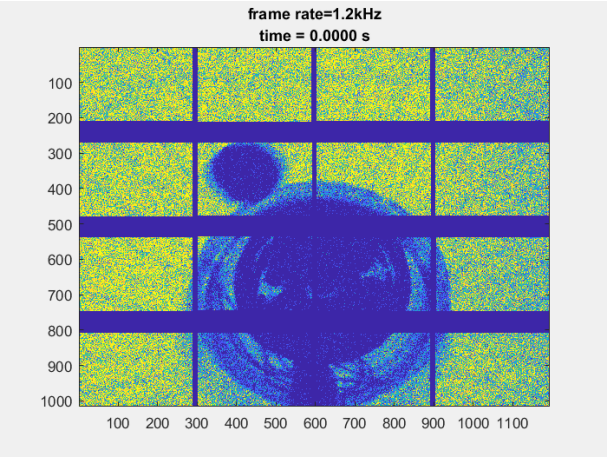




Front-end modules



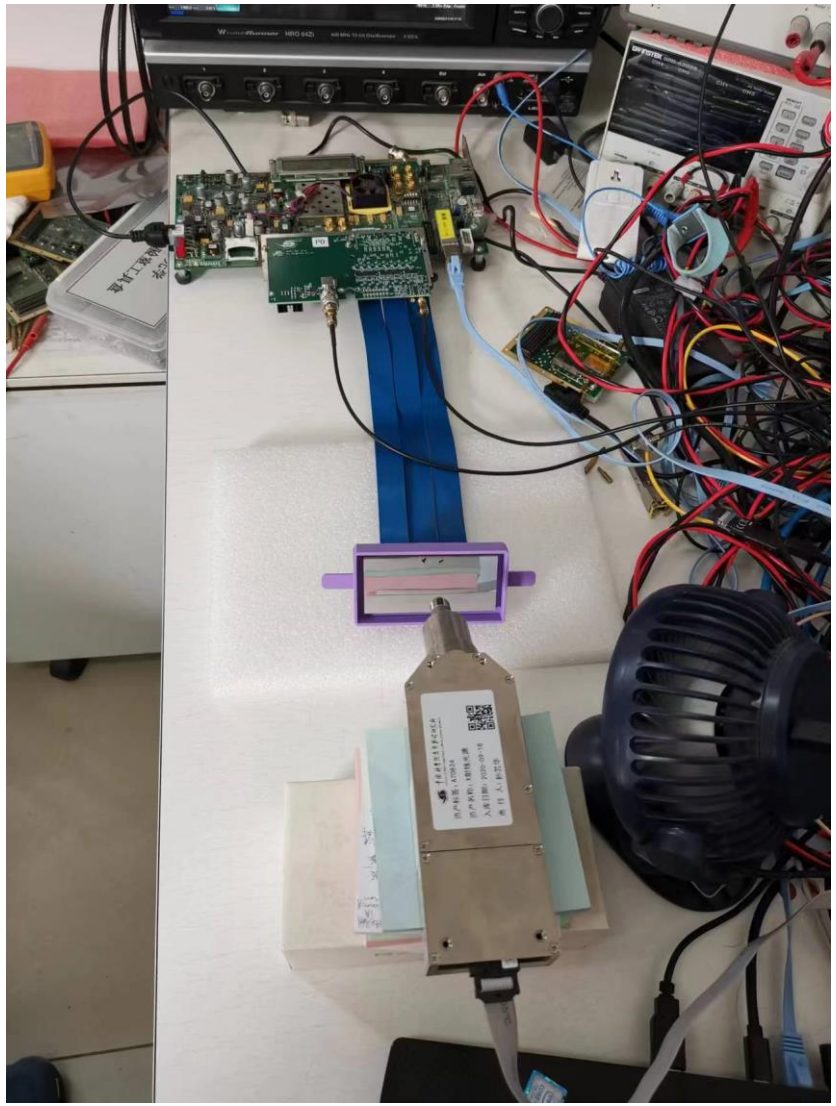
Cross-section of the front-end modules



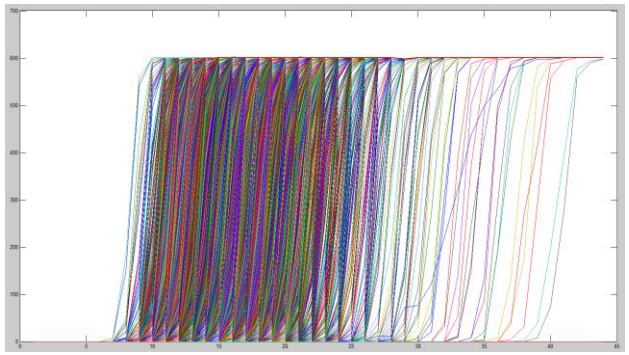
Dynamic image

- Image recorder
  - Assembled from multiple front-end modules
- Front-end modules
  - Sensor
    - An array of individual pixels arranged in a grid pattern
    - Convert the incident radiation into electrical signals
  - Readout ASIC
    - An array of individual pixels match with sensor
    - Amplify electrical signals and process to generate a digital image or data
  - Module hybridization
    - One sensor bump-bonded with multiple readout AISCs

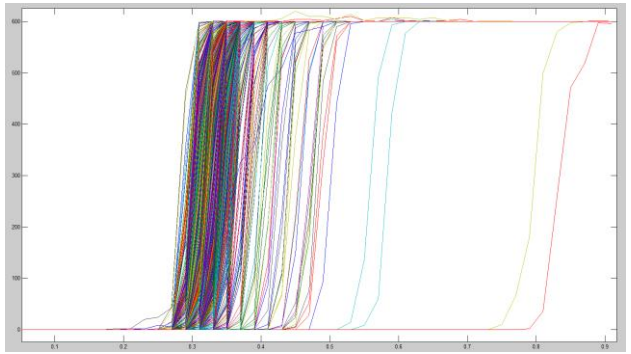
Single Module Test



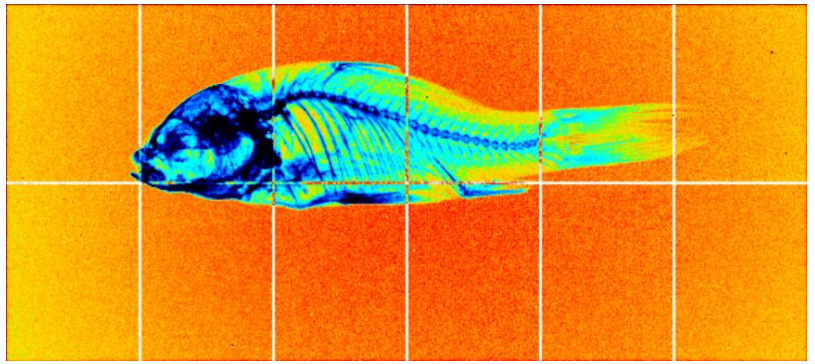
- **Single module passed the preliminary tests**
  - **All functions are as expected**
  - **Reliability tests in progress**

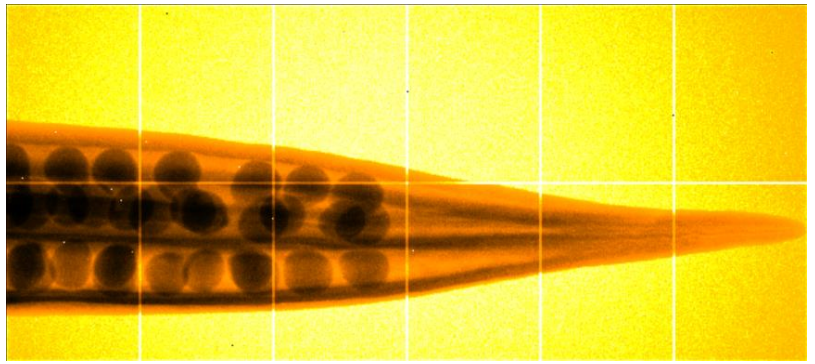


Threshold scan before calibration



Threshold scan after calibration

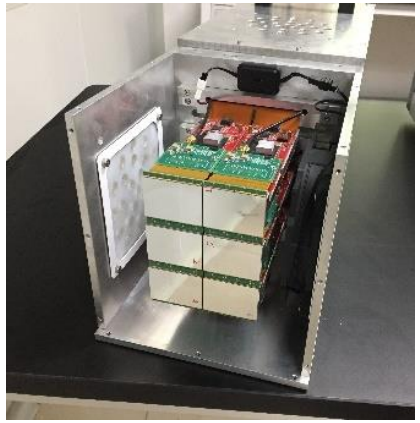


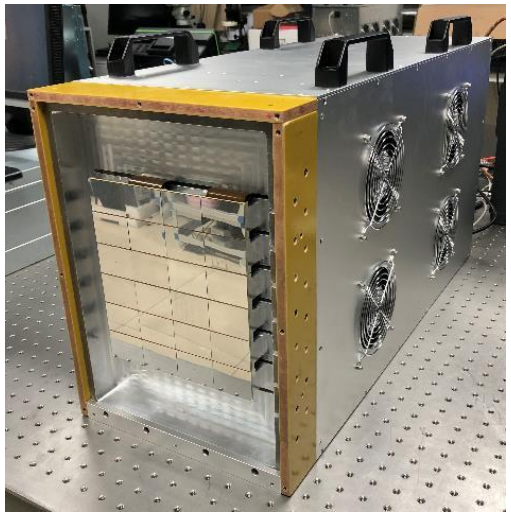


Preliminary X-ray image (uncalibrated)
A leaf exposed by an Au X-ray tube with 15 kV and 200 μA current for 0.3s

Prototype Systems Evolution

	1st BPIX（2015-2016）	2nd BPIX（2017-2018）	3rd BPIX（2019-2021）	4th BPIX（2022-now）
Modules	6	16	24	40
Pixels	360K	~1M	1.4M	6M
Assembly Scheme	Wire bonding & Rigid-flex PCB	Wire bonding & Rigid-flex PCB	Through Silicon Via (TSV) & Rigid-flex PCB with low CTE (coefficient of thermal expansion)	Advanced wire bonding & HTCC (High Temperature Co-Fired Ceramic)
Dead Area Occupancy	26.3%	26.3%	11.8%	~9.3%
FPGA Board	Spartan6 + SFP	Kintex7 + DDR3 + Molex Nano-Pitch I/O™ Cable	Kintex7 + DDR3 + Molex Nano-Pitch I/O™ Cable	UltraScale Kintex Plus + DDR4 + MicroTCA.4
DAQ Interface	1G Ethernet x12	1/10 G Ethernet x4	10G Ethernet x4	100G Ethernet x2
Power	100W	370W	500W	<2500W

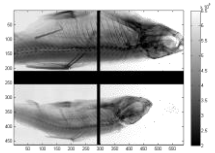


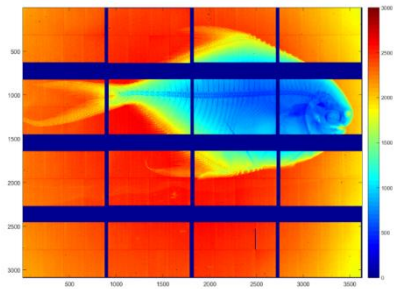


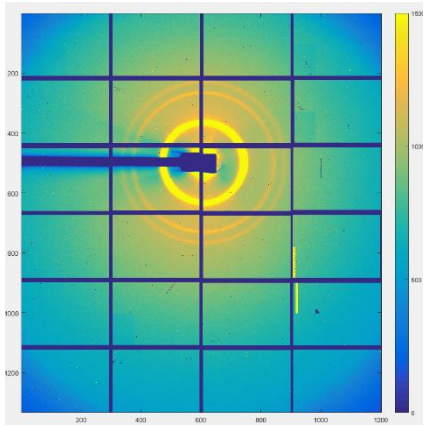
Under development

Prototype Systems Evolution

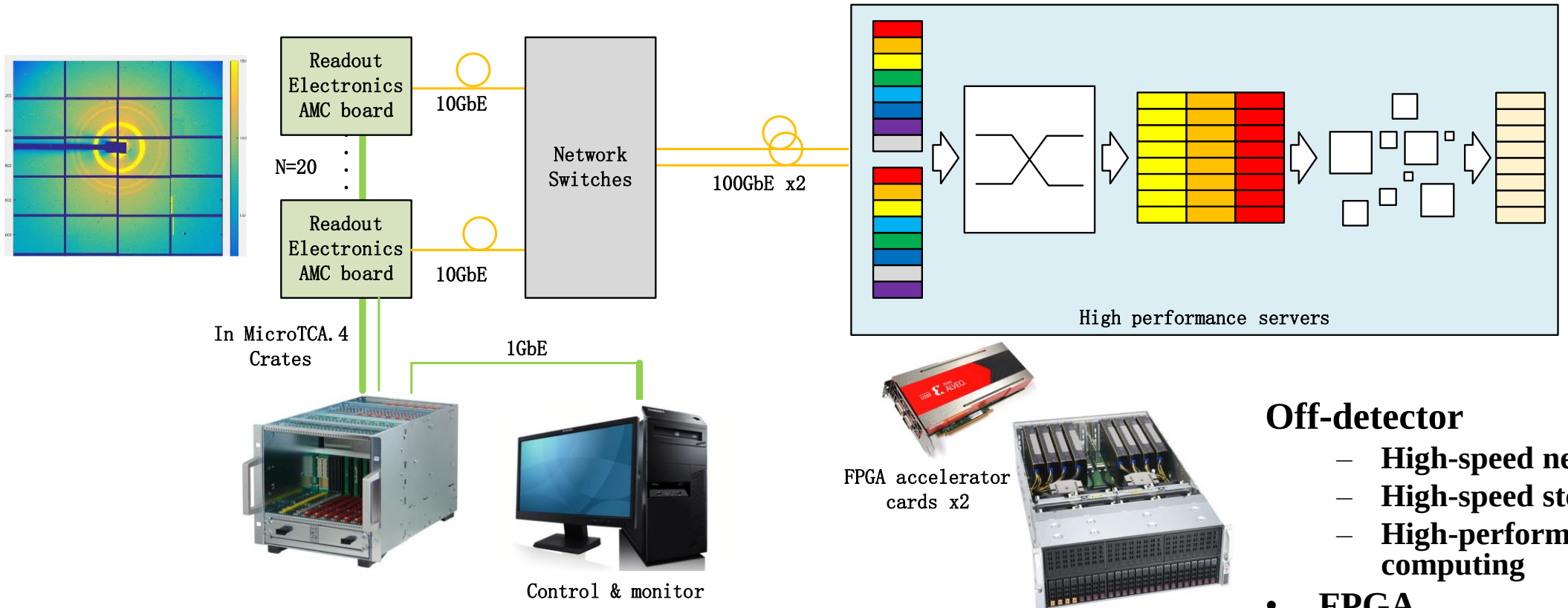
	1st BPIX（2015-2016）	2nd BPIX（2017-2018）	3rd BPIX（2019-2021）	4th BPIX（2022-now）
Modules	6	16	24	40
Pixels	360K	~1M	1.4M	6M
Assembly Scheme	Wire bonding & Rigid-flex PCB	Wire bonding & Rigid-flex PCB	Through Silicon Via (TSV) & Rigid-flex PCB with low CTE (coefficient of thermal expansion)	Advanced wire bonding & HTCC (High Temperature Co-Fired Ceramic)
Dead Area Occupancy	26.3%	26.3%	11.8%	~9.3%
FPGA Board	Spartan6 + SFP	Kintex7 + DDR3 + Molex Nano-Pitch I/O™ Cable	Kintex7 + DDR3 + Molex Nano-Pitch I/O™ Cable	UltraScale Kintex Plus + DDR4 + MicroTCA.4
DAQ Interface	1G Ethernet x12	1/10 G Ethernet x4	10G Ethernet x4	100G Ethernet x2
Power	100W	370W	500W	<2500W







Electronics System Architecture (Baseline)



On-detector

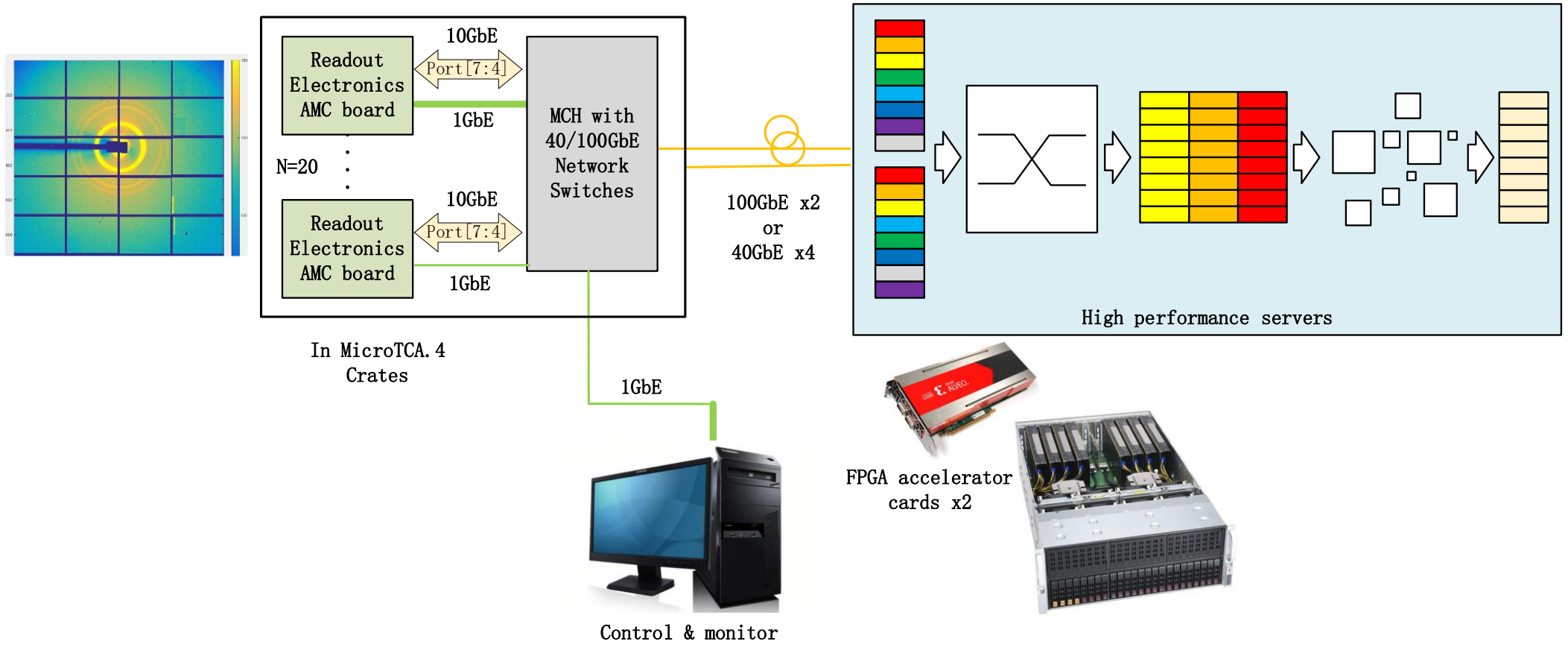
- **FPGA AMC board**
  - **Front-end Electronics Control**
  - **Clock Synchronization**
  - **Monitoring**
  - **Data Acquisition**

Off-detector

- **High-speed network**
- **High-speed storage**
- **High-performance computing**

- **FPGA**
  - **Data aggregation**
  - **Data sorting**
  - **Data compression**
  - **Real-time algorithm**

Electronics System Architecture (Alternative)



Use the MCH with 40/100GbE switch to replace the commercial network switch

- **More compact**

Options:

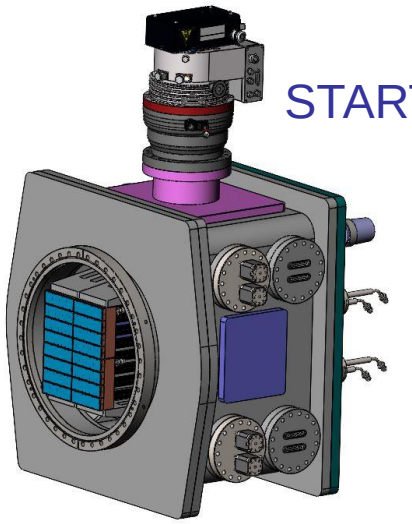
- **NAT, NAT-HUB-E**
- **Vadatech, UTC056-500-212-110**

More Applications – SHINE STARTLIGHT Detector

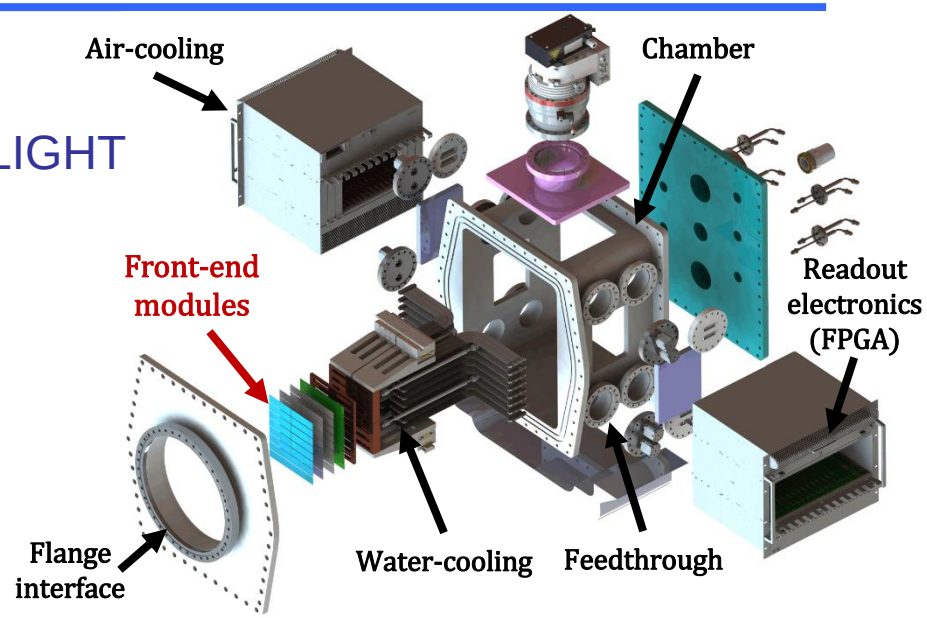
Shanghai HIgh repetitioN rate xfel and Extreme light facility (SHINE)

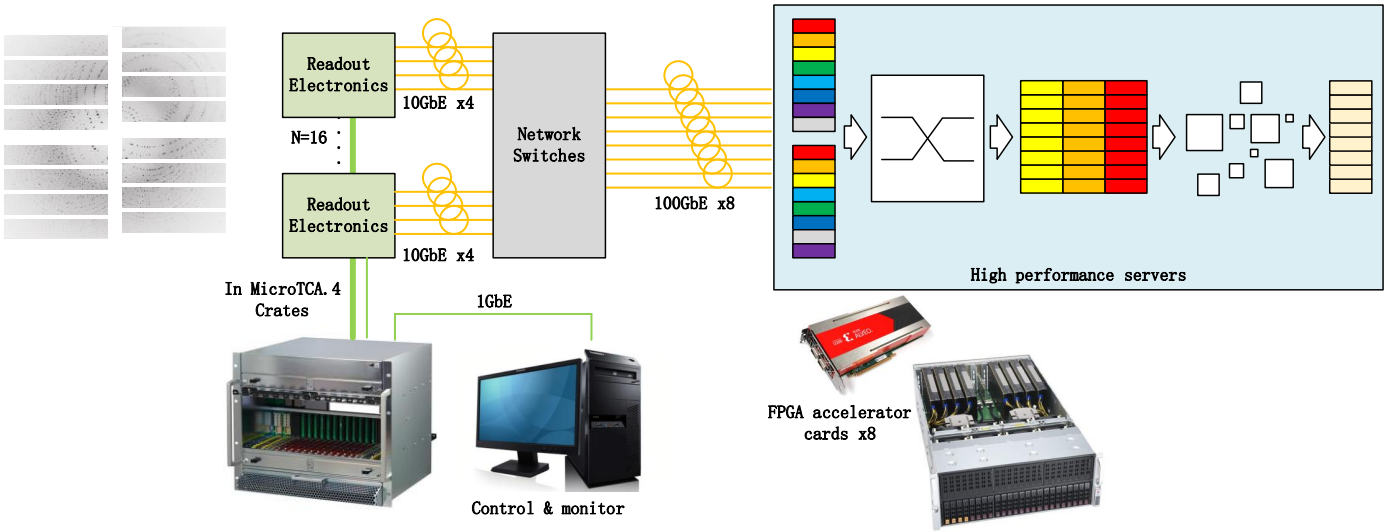
- Under construction in Zhangjiang, Shanghai
  - Start the user operation in 2026
- 3 beamlines and 10 end-stations
- One end-station for area array detector with silicon pixels
  - Detector: STARTLIGHT (SemiconducTor Array detectoR with Large dynamlc ranGe and cHarge inTegrating readout
  - ASIC: HYLITE (High dYnamic range free electron Laser Imaging deTEctor)

Specs	STARLIGHT
Pixel size	100 μm X 100 μm
ASIC Pixel Array	128 X 128
Gain	Self-adaptive 3 gains
Dynamic range	1 ~ 10000 photons/pulse @12 keV
Frame rate	12 kHz (continuous readout)
Detector	A 4.2M pixel detector in vacuum, quadrant movable



STARTLIGHT





Challenge: Continuously data stream up to 654.3 Gbps

SUMMARY

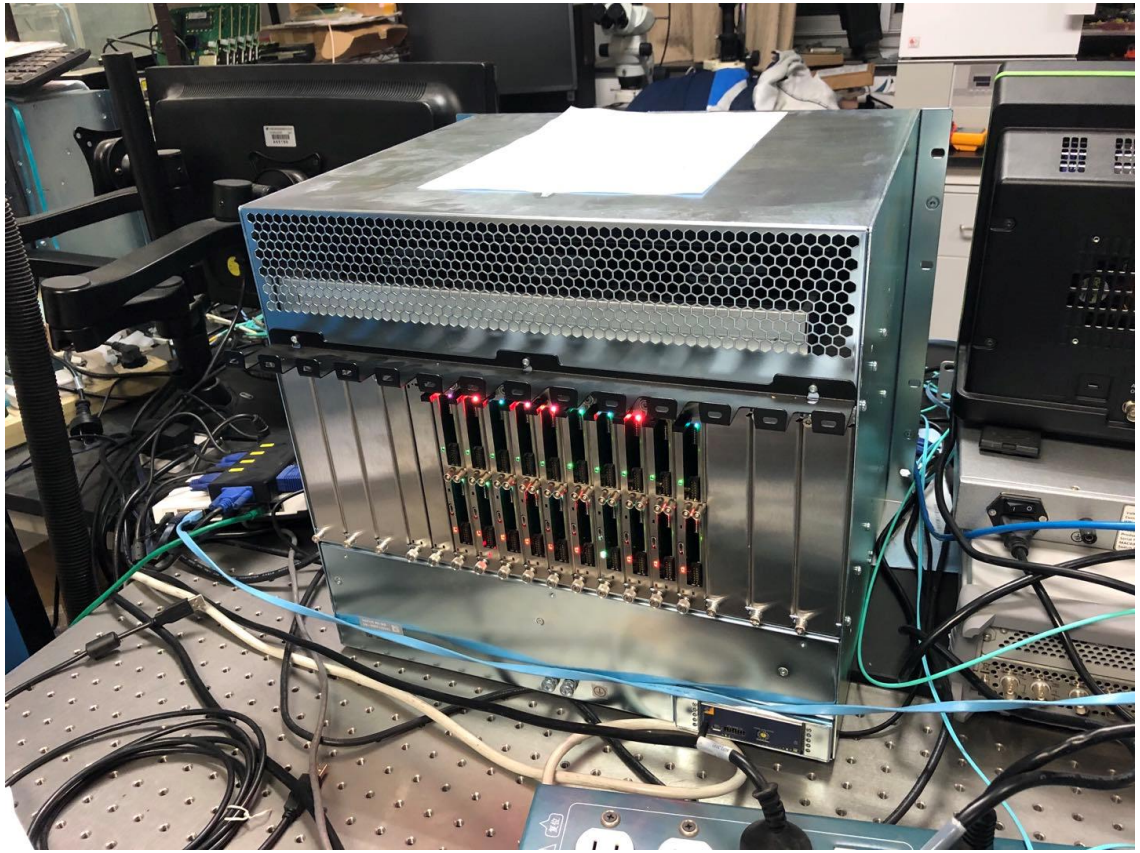
- MicroTCA architecture
  - Suitable for small and medium-sized experiments
- uFC series boards
  - FPGA-based MicroTCA compatible AMC board
    - For generic system control and data acquisition in physics experiments
  - HPC FMC sockets
    - Provide additional clock signals, user-specific I/O and high-speed transceivers that can be used to extend the connectivity as well as the I/O bandwidth
  - Successfully demonstrated the feasibility of the uFC

- Outlook
  - High-level tools for software development productivity
    - Vivado HLS, OpenCL, etc.
    - Applications in physics experiments
      - Need the cooperation with PHY/SIM/DAQ/Online-tracking groups
  - Long-term experience with respect to reliability and availability

THANKS TO YOUR ATTENTION

MICROTCA.4 CRATE WITH MCH, CPU AND FPGA BOARDS

Front View



Back View

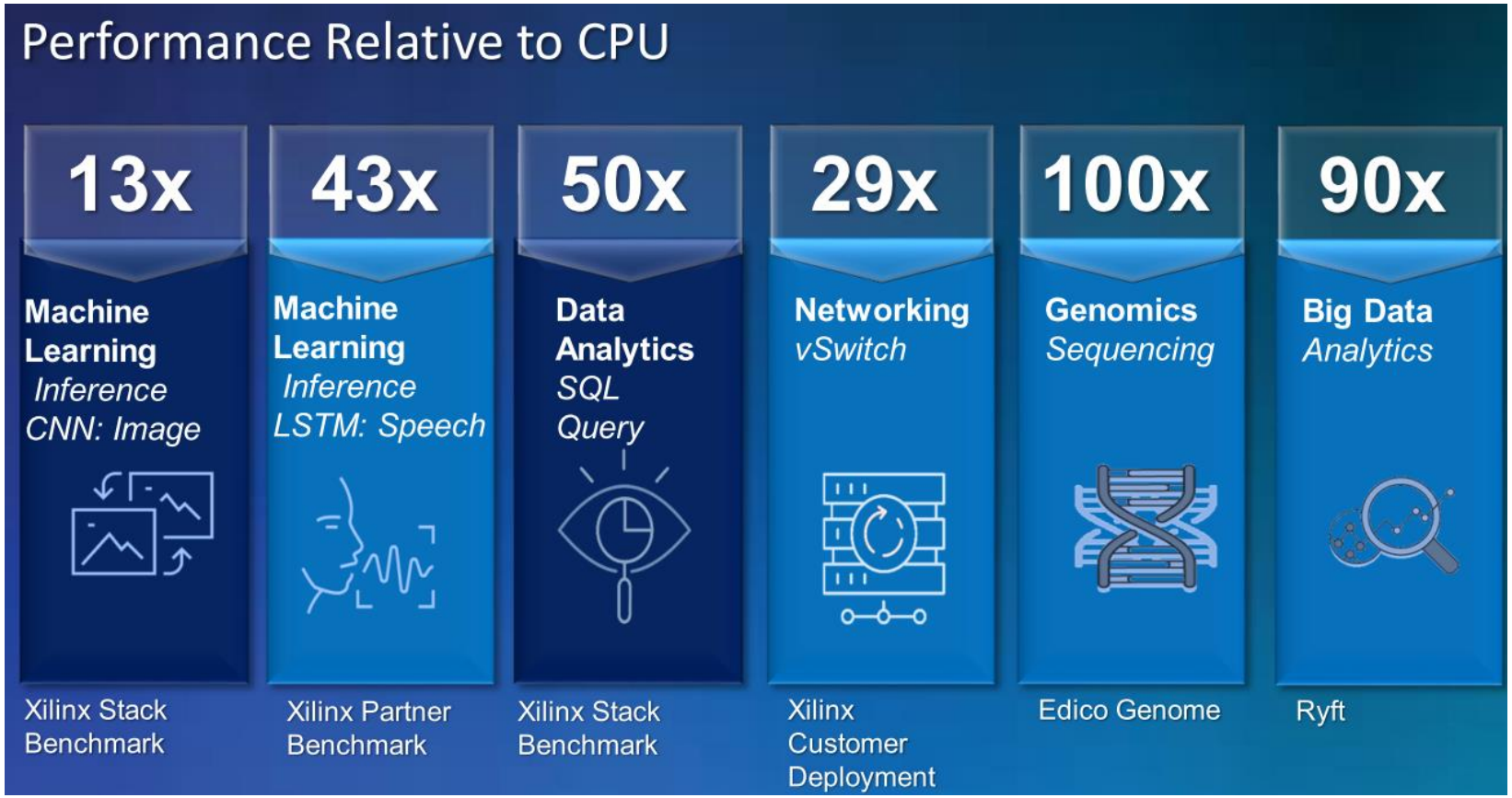
NOUNS

- **IP**, Intellectual Property
  In electronic design a semiconductor intellectual property core, IP core, or IP block is a reusable unit of logic, cell, or integrated circuit (commonly called a "chip") layout design that is the intellectual property of one party.
- **DMA**, Direct memory access
  DMA is a feature of computer systems that allows certain hardware subsystems to access main system memory (random-access memory), independent of the central processing unit (CPU).
- **XDMA**, DMA from Xilinx
- **RTL**, register-transfer level
  In digital circuit design, RTL is a design abstraction which models a synchronous digital circuit in terms of the flow of digital signals (data) between hardware registers, and the logical operations performed on those signals.
- **NVMe SSD**, Non-Volatile Memory Express Solid State Drives
- **Iperf**, is a network testing utility helpful for determining network performance.

Source: .wikipedia.org

KEY BENEFITS TO FPGA COMPUTING

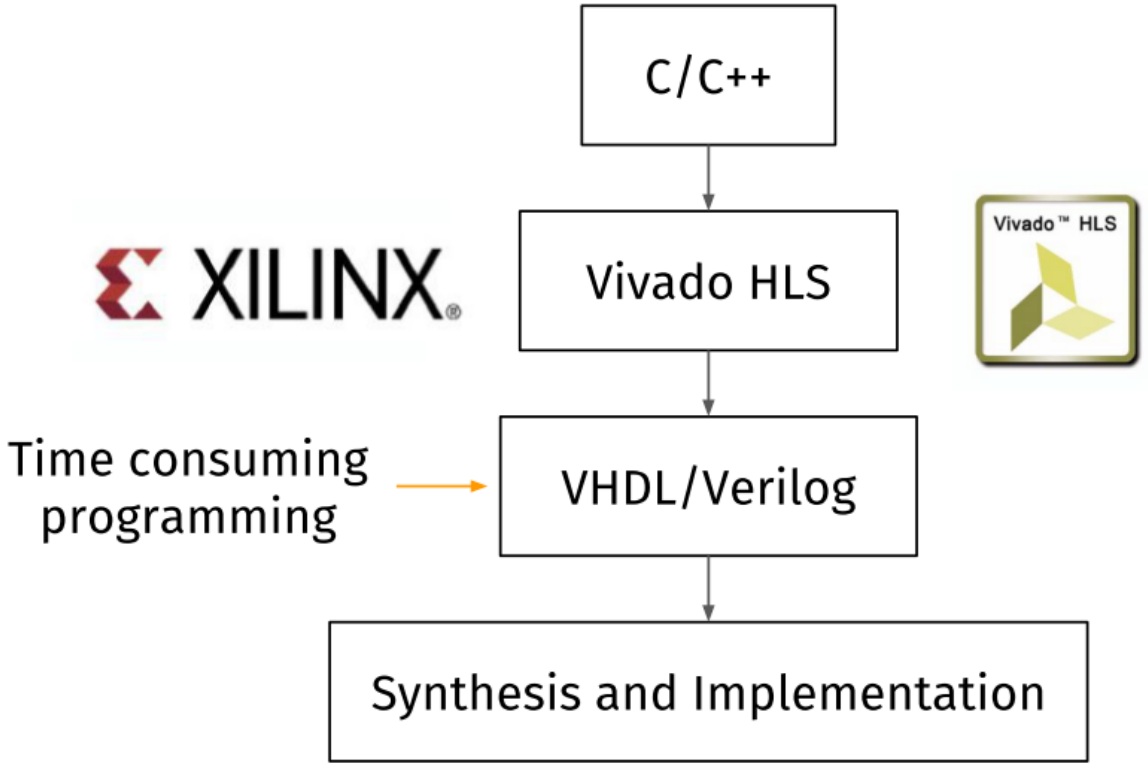
- Balances programmability and high performance for key workloads
- Utilizing FPGA technology as a utility, resulting in faster access to the newest technology

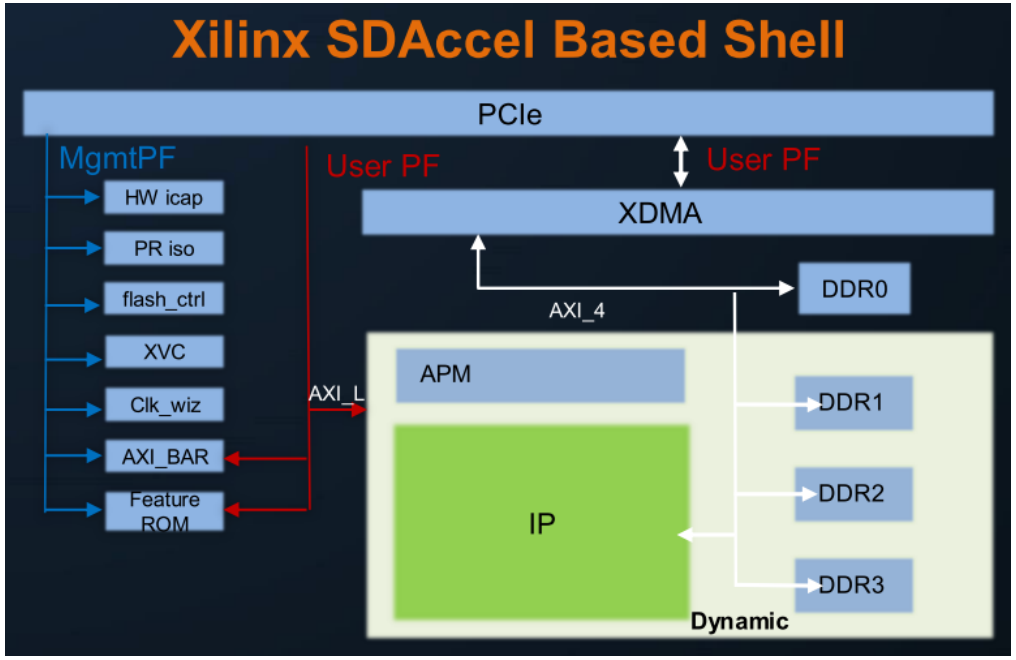


XILINX

Source: Xilinx Technical Marketing

FPGA SHELL OPTIONS



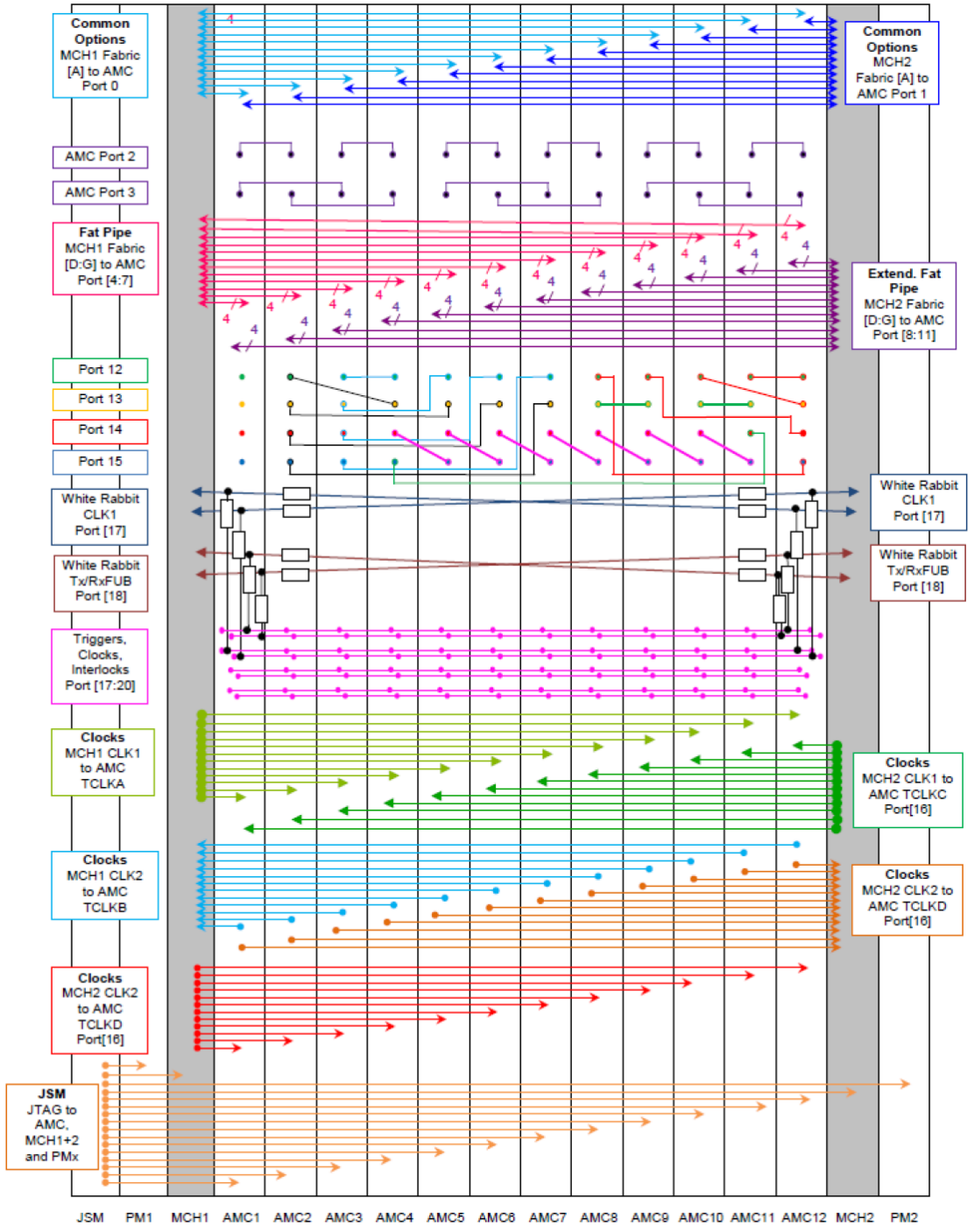


- Scenario： Rapid development, block computing
  - User logic: OpenCL C, HLS C and RTL supported
  - Suited for quick evaluation/porting of existing customer code
- Shell feature：
  - Xilinx scatter-gather XDMA optimized for big block data transfer
  - Serial message notification
  - Offload acceleration

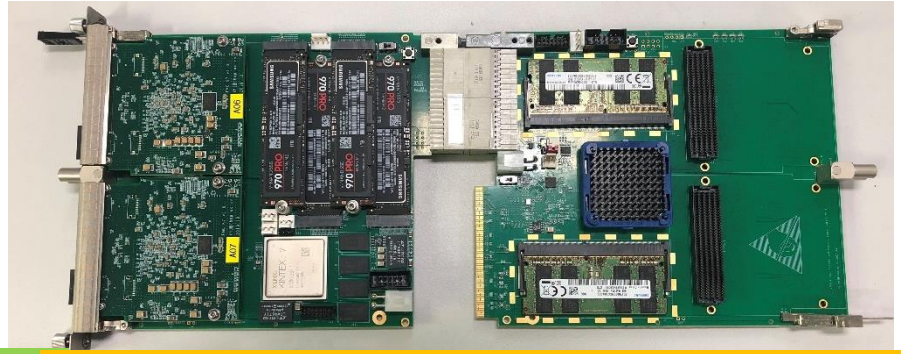
BACKPLANE TOPOLOGY

N.A.T NATIVE-R9-WR Crate

- **Port 0: 1GbE**
- Port 1: Redundant 1GbE
- Port 2~3: Internal links
- **Port 4~7: PCIe x4**
- Port 8~11: Redundant PCIe x4
- Port 12~15: Internal links
- **Port 17~20: Triggers, Clocks or Interlocks**
- **TCLKA, TCLKB: System clocks**
- TCLKC, TCLKD: Redundant system clocks
- TCLKD: PCIe reference clock



uFC Series Board



uFC v2
- Xilinx Kintex-7 28nm 7K325T
  - 0.32 Million System Logic
  - 840 DSP
- PCIe2.0 x4
- 8GB DDR3 800MHz SDRAM ECC
- (8+2)*10G High-Speed Serial Links

u4FC&P v1
- Kintex Ultrascale+ 16nm KU11P
  - 0.65 Million System Logic
  - 2928 DSP
- 4*PCIe4.0 x4 + PCIe4.0 x8
- 2*16GB DDR4 1200MHz SDRAM ECC
- 4*40G/100G High-Speed Serial Links

u4FC&P v2
- Kintex Ultrascale+ 16nm **KU15P**
  - 1.14 Million System Logic
  - 1968 DSP
- SAMTEC Firefly x3 + PCIe4.0 x8
- 2* 16GB DDR4 1200MHz SDRAM ECC
- 8*40G/100G High-Speed Serial Links

Name	Instance Specs					
	Status	FPGA	Memory	NVMe	PCIe BW	Network
uFC v2	Ready	7K325T	8GB	-	2 GB/s	10GbE
uFC&P v1	Ready	KU11P	2*16GB	4*(up to 4TB)	8 GB/s	40/100GbE
uFC&P v2	R&D	KU15P	2*16GB	-	8 GB/s	40/100GbE

Ref1: Zhang, J., et al. (2023). TNS **70(6): 935-940.**
Ref2: Zhang, J., et al. (2019). TNS **66(7): 1169-1173.**

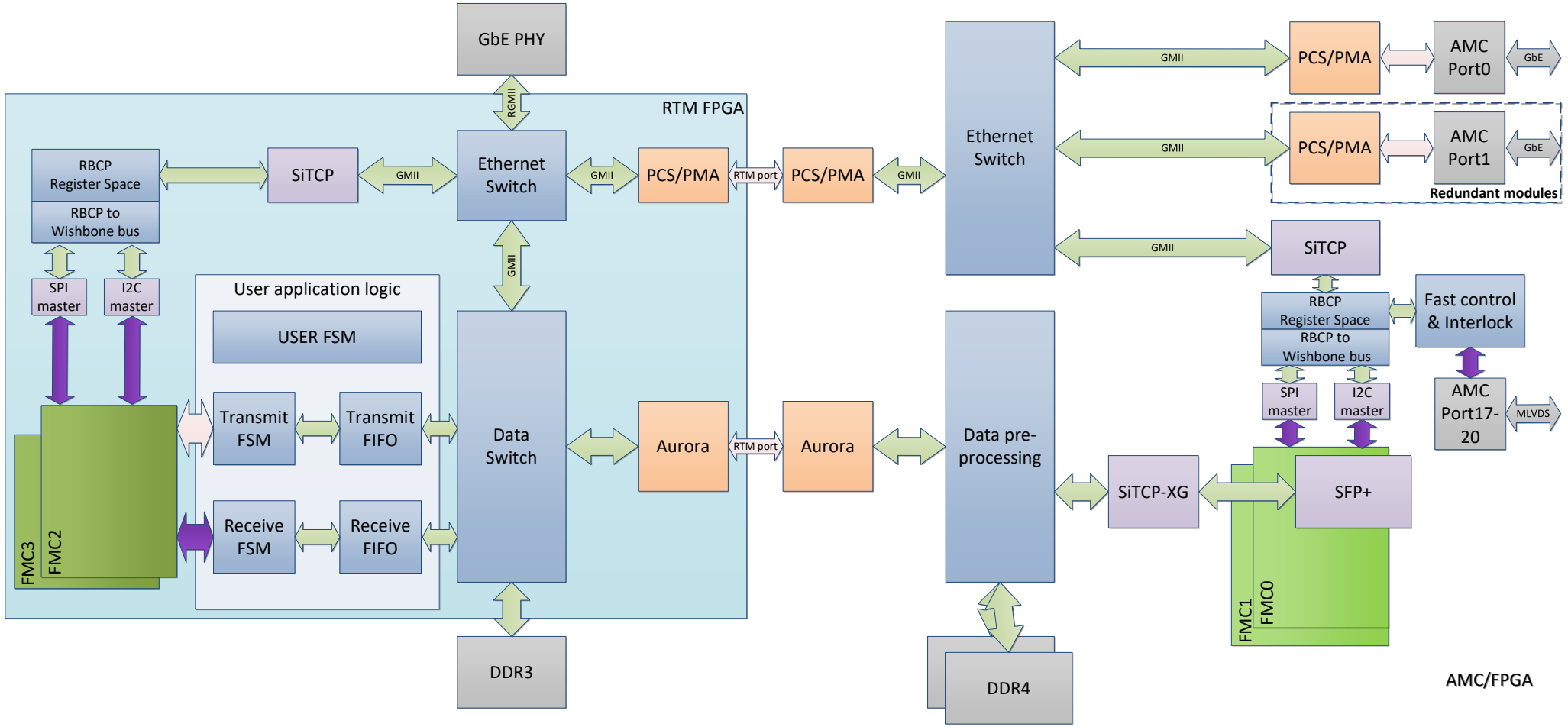
COMMUNICATION EVALUATION

- FPGA connects NVMe SSD directly with file system
  - Without CPU or external memory.
  - It is the best solution for applications which require huge capacity and ultra high-speed.
  - NVMe-IP: FAT32 or exFAT
    - From Design Gateway Co., Ltd
    - Tested via Xilinx KCU105 evaluation board

NVMe SSD M.2	HP EX900	Samsung 970 EVO	Samsung 970 PRO
Writing Speed from datasheet	1300 MB/s	Up to 2300 MB/s	Up to 2300 MB/s
NVMe-IP Tested Average Writing Speed @ Block size: 128 KB	80~100 MB/s *	~800 Mb/s *	~2200 MB/s

* Very fast at the beginning

FPGA Firmware



- **Config & monitor via UDP/IP**
- **Readout via TCP/IP**

Key Parameters of the Pixel Detectors for HEPS and SHINE

	HEPS BPIX-6M	SHINE STARLIGHT
Mode	Photon counting readout	Charge-integration
Pixel size	140 μm x 140 μm	100 μm x 100 μm
Total pixel	5.9M	4.2M
Modules	40	16
ASICs per module	6 x 2	8 x 2
ASIC Pixel Array	128 x 96	128 x 128
Data length per pixel	28 bit	13 bit
Total detectable area	403.2 mm x 286.72 mm	204.8 mm x 204.8 mm
Frame rate (Max.)	1 kHz	12 kHz
Peak data rate from electronics to DAQ	165.2 Gbps	654.3 Gbps