FPGA-based packet switch

UNIVERSIDAD DE GRANADA

Carlos Megías Núñez, PhD student

13th WR Workshop

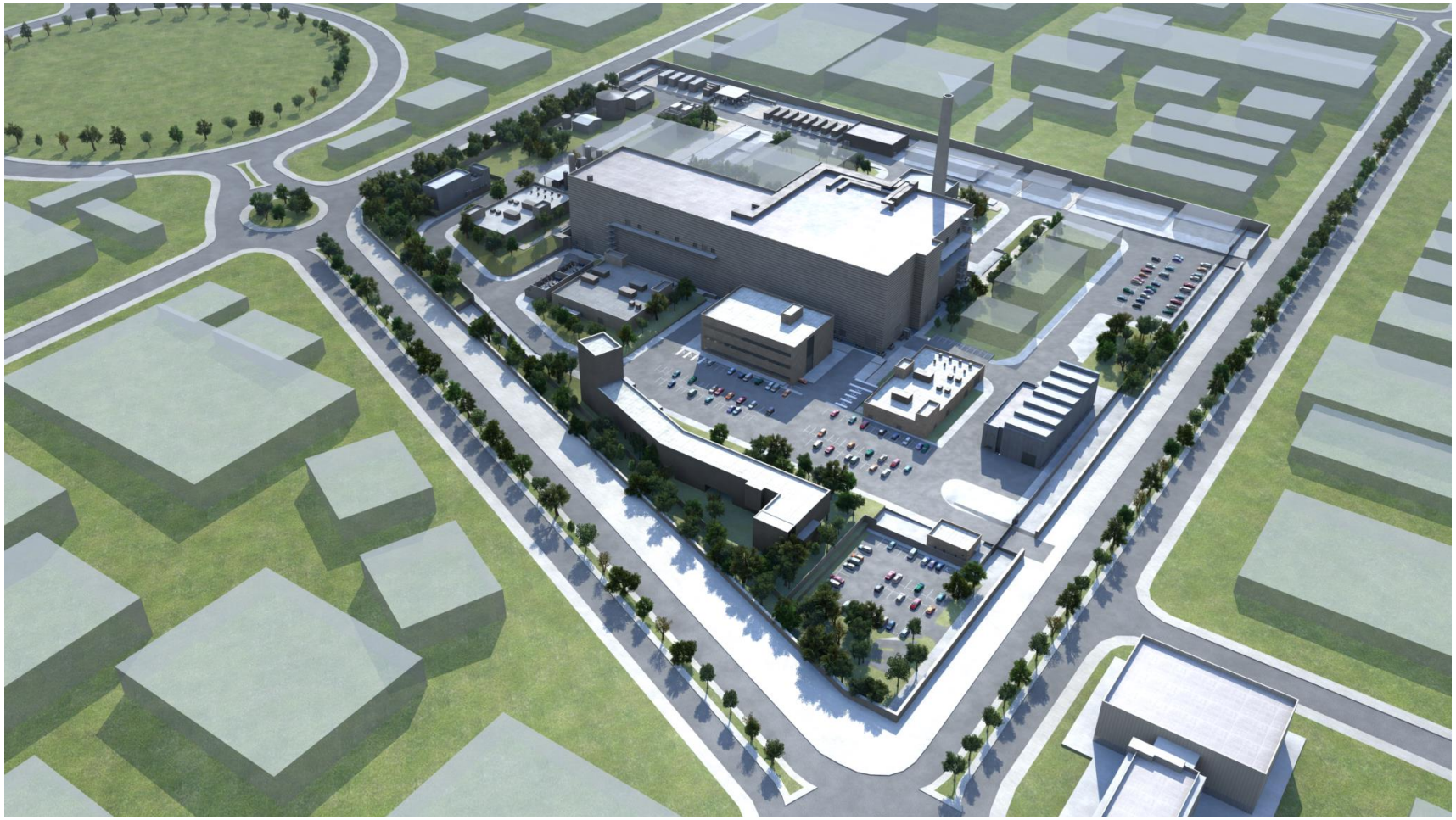
WRC Launch Event

Who are we?

- Time-based Technologies and Networks lab at the University of Granada, Spain
  - Working with timing related technologies since ~2009
- Most of the research activity for this project has been carried out at Research Centre for Information and Communications Technologies (CITIC-UGR)

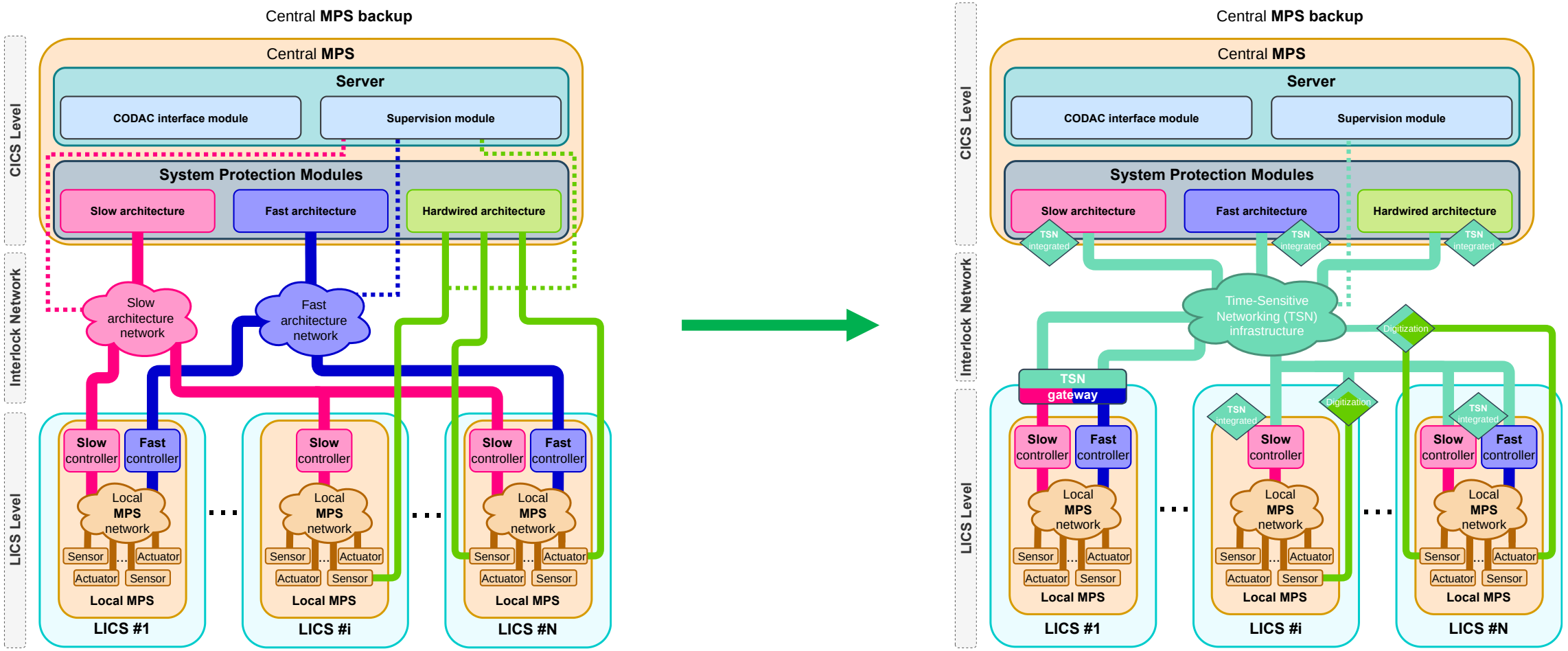
IFMIF-DONES coming to Granada…

- International Fusion Materials Irradiation Facility – Demo Oriented NEutron Source
- Part of the ITER and DEMO fusion program
- Our research group is responsible of some tasks related with:
  - Timing
  - Control
  - Networks



Interlock propagation in IFMIF-DONES

- We are proposing TSN technology for interlock propagation between the central control units and the local control units



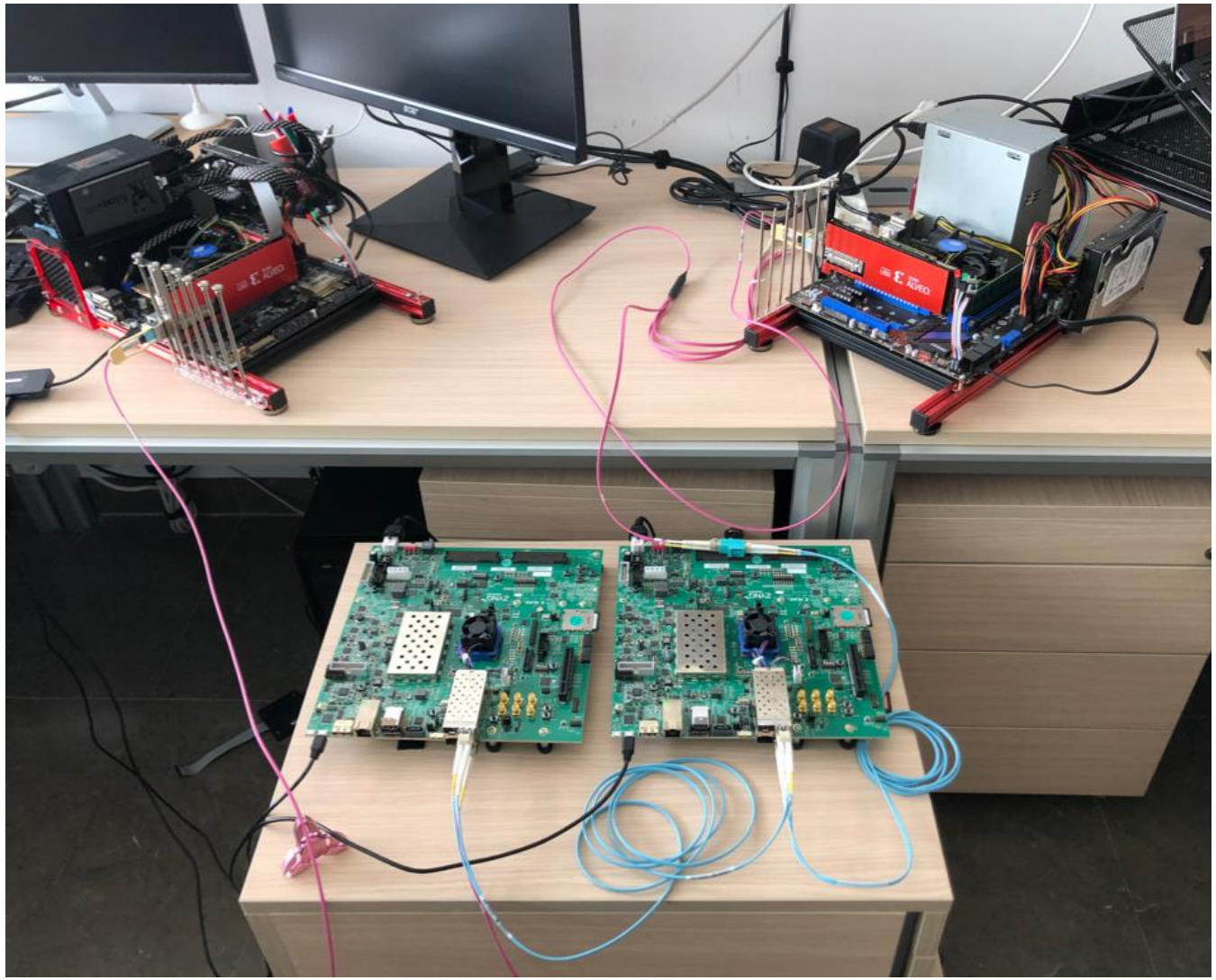
Time-Sensitive Networking (TSN)

- Collection of standards over conventional **Ethernet networks**
- Guarantee the determinism by keeping a **bounded end-to-end latency** and ensures an objective bandwidth
- **Aggregation of traffic flows** of different priorities and criticality from different systems on the same physical network infrastructure

- Additional features: redundancy, frame preemption, filtering and policing, etc
- For a complete TSN setup we need two fundamental types of nodes: **end-nodes and network switches**

Getting started with Corundum…

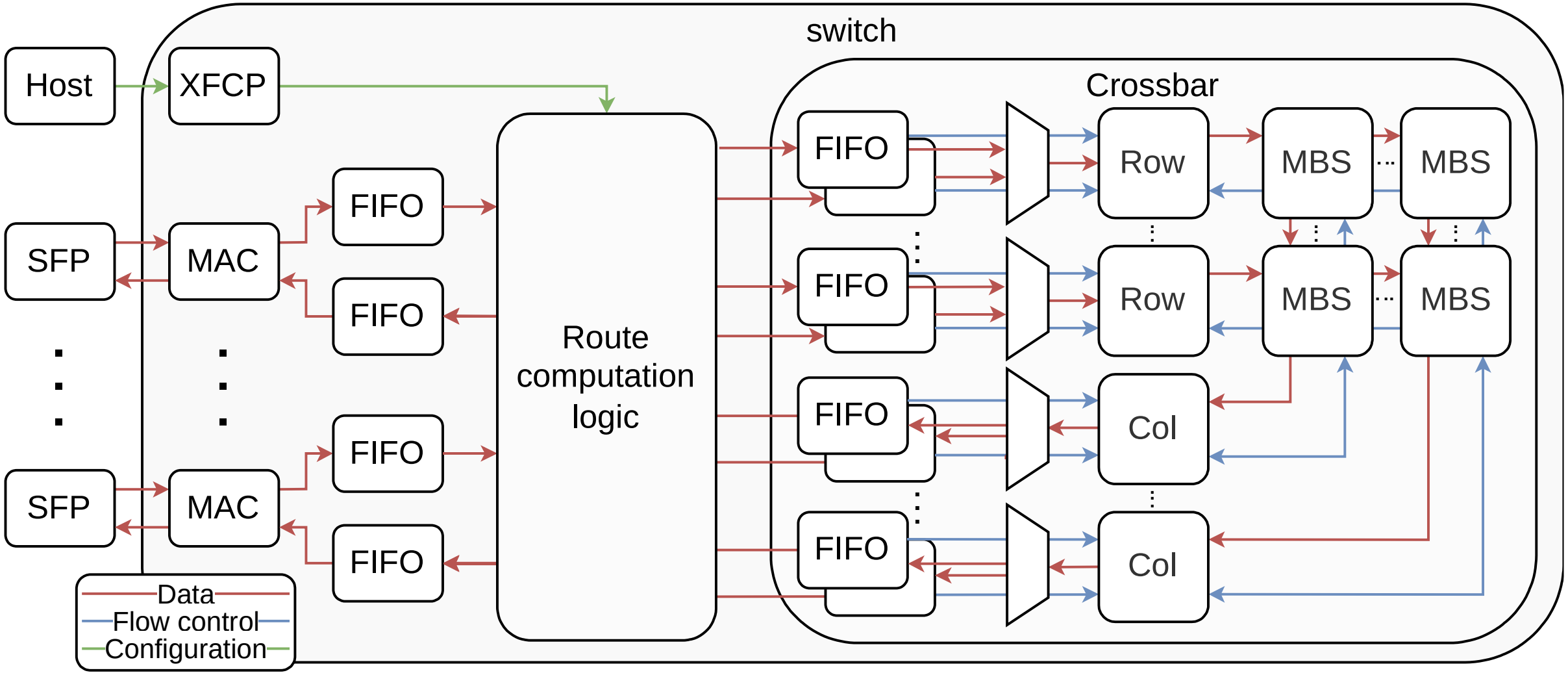
- We were looking for a **flexible and high performance communications platform** in which we would integrate the different TSN capabilities for our use case in IFMIF-DONES
  - High performance NIC implementation: end node implementation
  - The missing element: packet switch
- **‘23 summer stay at University of California, San Diego, with Alex Forencich** for designing and implementing an FPGA-based packet switch



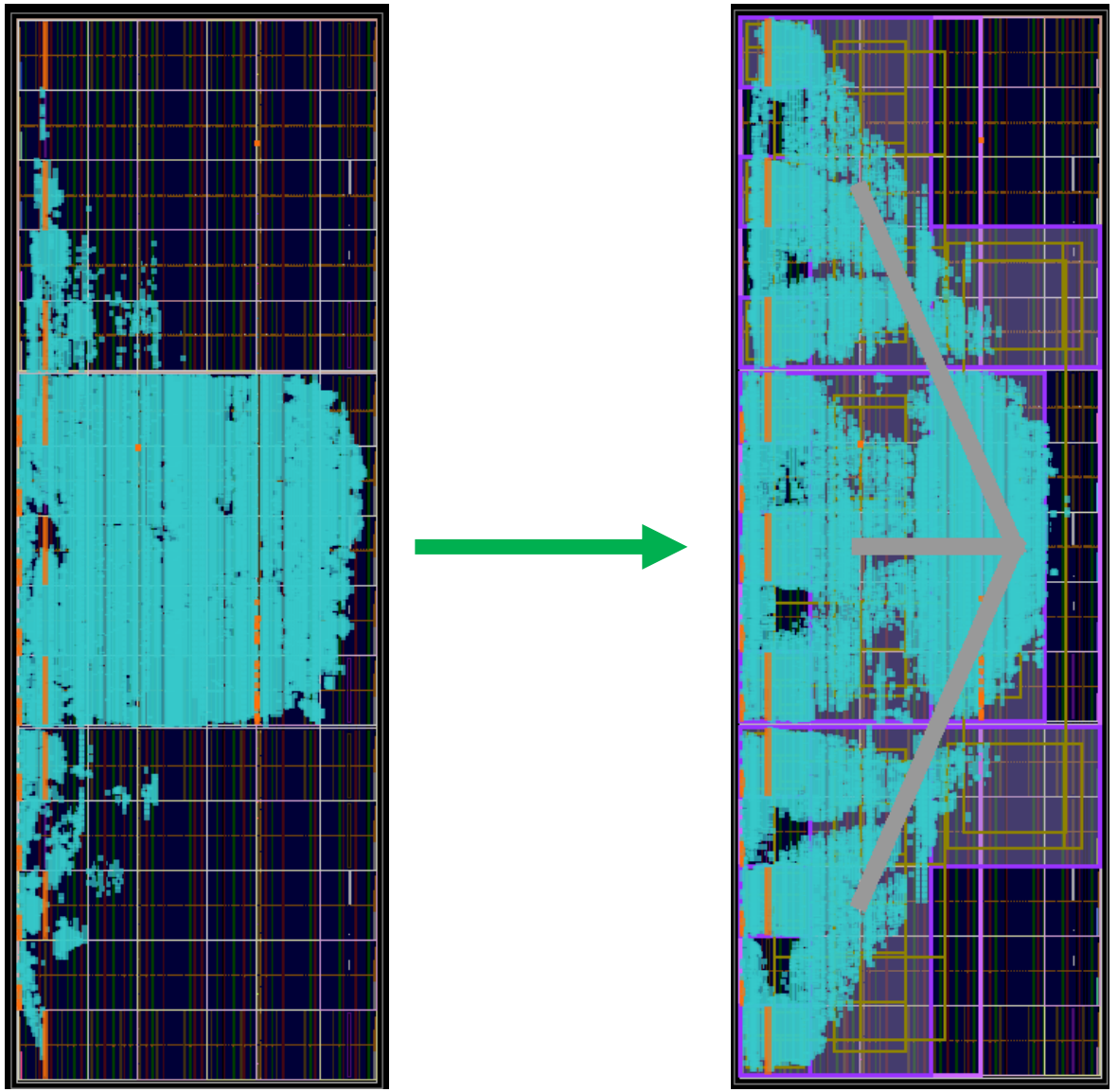
100 Gbps Ethernet packet switch

- Checklist:
  - High performance switch capable of up to 100 Gbps per port
    - Port splitting/rate changes: 1x100G vs 4x1/10/25G
  - Modular design:
    - Easily upgradable design
    - Depending on target platform (number of ports…)
    - Pluggable modules
  - Configurable: depending on desired application
  - Prioritization inside the switch: for running applications such as TSN
- Limited bibliography about FPGA-based packet switch implementations
  - No open-source designs or limited performance
- Two main separated (mostly independent) parts:
  - Crossbar: interconnection between all ports
  - Route computation logic: based on the frame header, get the corresponding output ports

Architecture overview



Some design problems: congestion



Chronological achievements

Start of the project
- Bibliographic revision
- High level discussions

February 2023

Architectural decisions
- HDL simulations of typical switch approximations
- High level discussions

May 2023

Start building the switch
- First HDL

June 2023

First tests
- Setup development platform
- High level simulations
- First packets coming out, padding frames to multiples of 128 bytes! 4x4, 1MBS, 2 SUP, 1024-bit datapath

July 2023

First implementations
- Basic crossbar completed: no VC, no multicast
- 4x4, 1 MBS, 2 SUP, 1024-bit datapath, ~53% throughput (worst case)
- 4x4, 4 MBS, 2 SUP, 1024-bit datapath, 100% throughput
- Simple routing based on destination MAC address

August 2023

Chronological achievements

August 2023

Fighting Vivado
- First, trying to put an 8x8 design with speed up of 2
- Then, trying to put a 9x9 design with speed up of 2

September 2023

Successful implementations
- 3x3, 1 MBS, 2 SUP, 1024-bit datapath, ~71% throughput (worst case)
- 9x9, 9 MBS, 2 SUP, 1024-bit datapath, ~71% throughput (worst case)

October 2023

Architectural changes
- Protocol conversion: AXIS to more convenient bus signals
- Bus width increased to 1536-bit: metadata!
- 3x3, 1 MBS, 3 SUP, 1536-bit datapath, 100% throughput

November 2023

Virtual channels
- Complete first support for virtual channels: parameterizable

December 2023

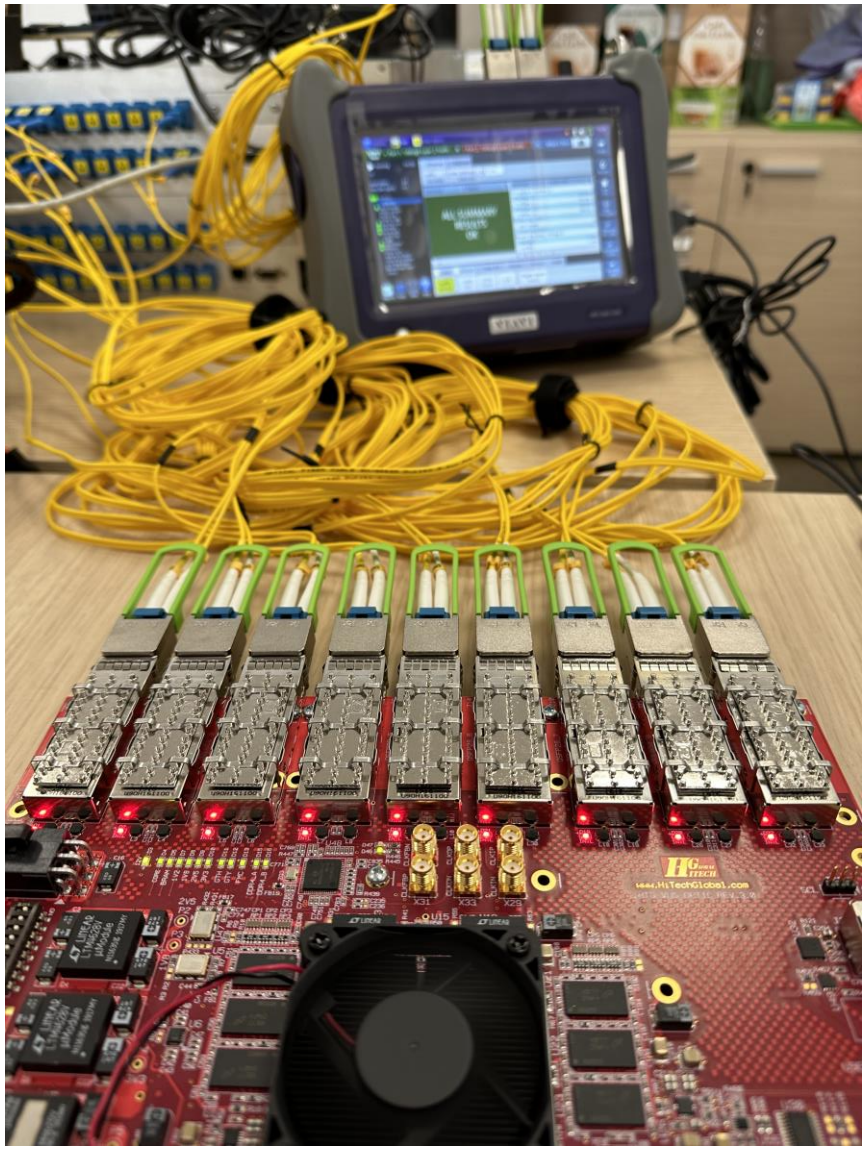
Multicast
- Complete first support for multicast

January - February 2024

Route computation logic
- Added simple UART configuration interface using the XFCP library
- Added “complete” support for Layer 2: dest MAC address learning
- Added very simple support for VLANs

Current status…

- So far, the best design that has been put together:
  - 9-port switch capable of 100 Gbps per port, total aggregated bandwidth of 900 Gbps
  - 4 virtual channels: right now, there is no application taking advantage of them (RR)
  - Jumbo frames are supported up to 9600 B
  - More than 100,000 MAC table entries, theoretically
- The testbed setup is not yet fully operational, but we have everything prepared:
  - Optical switches
  - Network tester
  - Some end-stations for additional traffic generation
- Next: correct bugs, add PTP, VLAN and TSN features
  - **Extend PTP support with WR**



Current status…

- Boards being used: HTG-9200 populated with two variants: VU13P and VU9P
- Possible devices with a more “more convenient” form factor:
  - Arista 7132LB-48Y4C (VU35P)
  - Bittware Terabox 1100L (VU9P)
  - Cisco Nexus 3550-T (VU9P)