The evolution of the Data Acquisition System of KM3NeT

Francesco Benfenati on behalf of the **KM3NeT Collaboration**

INFN-Bologna
Dipartimento di Fisica e Astronomia - Università di Bologna

The KM3NeT neutrino experiment

KM3NeT = Km³ Neutrino Telescope

- **ARCA** (**A**stroparticle **R**esearch with **C**osmics in the **A**byss)

📍Off-shore Sicily, Italy at ~3500m below sea level

Observation of **high energy** (GeV-PeV) **astrophysical neutrino sources**

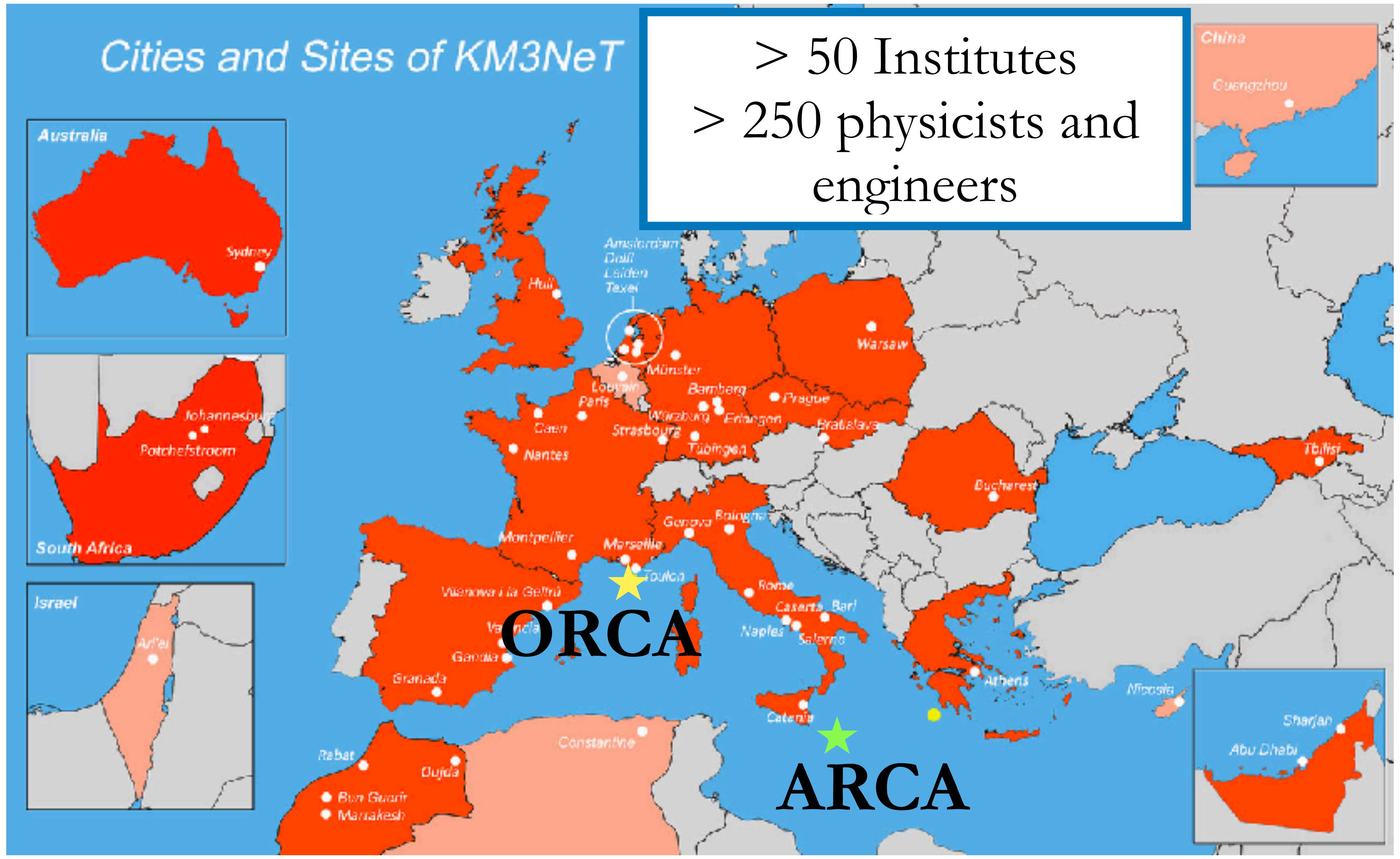
Currently: **ARCA28**

- **ORCA** (**O**scillation **R**esearch with **C**osmics in the **A**byss)

📍Off-shore Toulon, France at ~2500m below sea level

Determination of neutrino **mass hierarchy** and **oscillation** parameters (O(10-100) GeV)

Currently: **ORCA19**

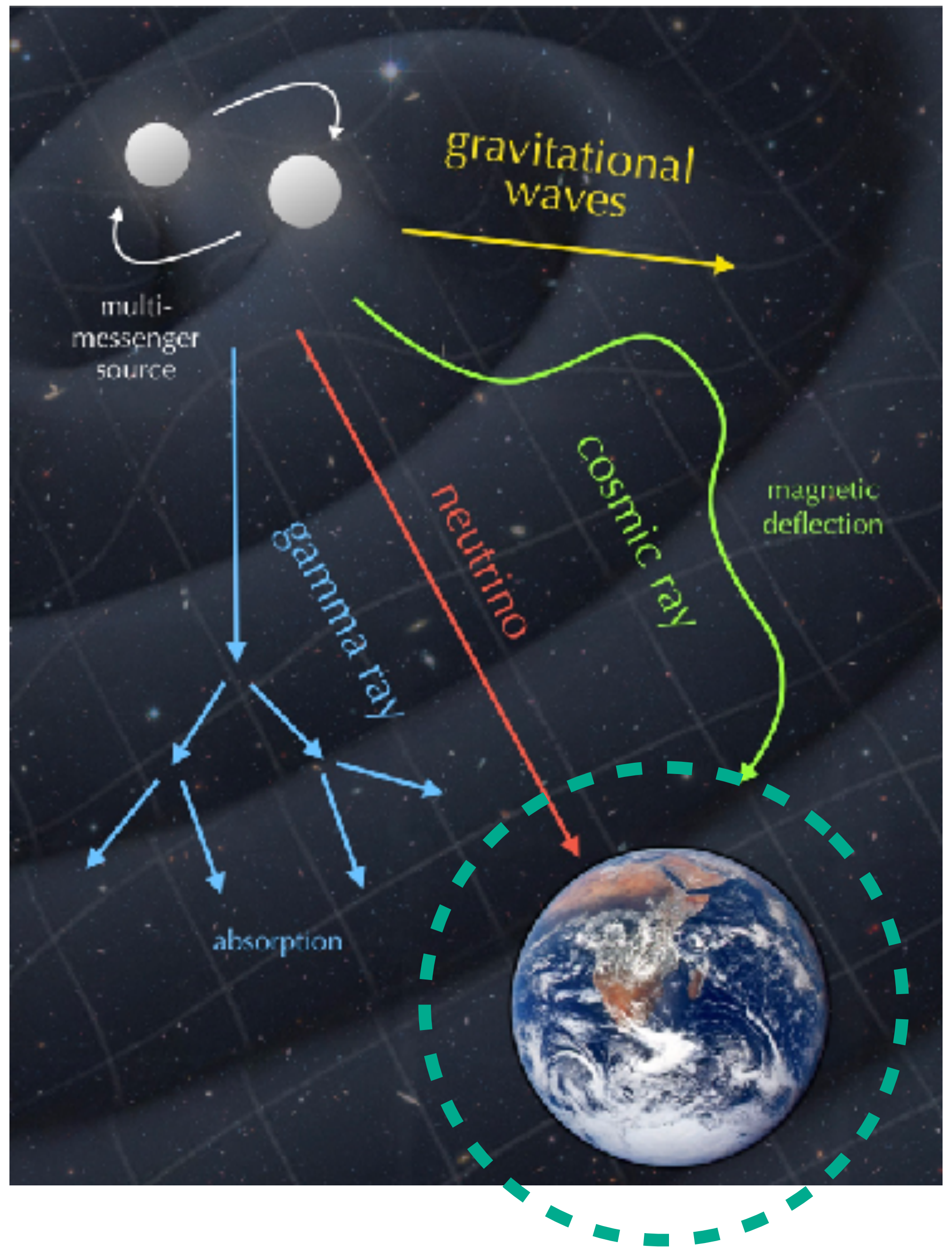


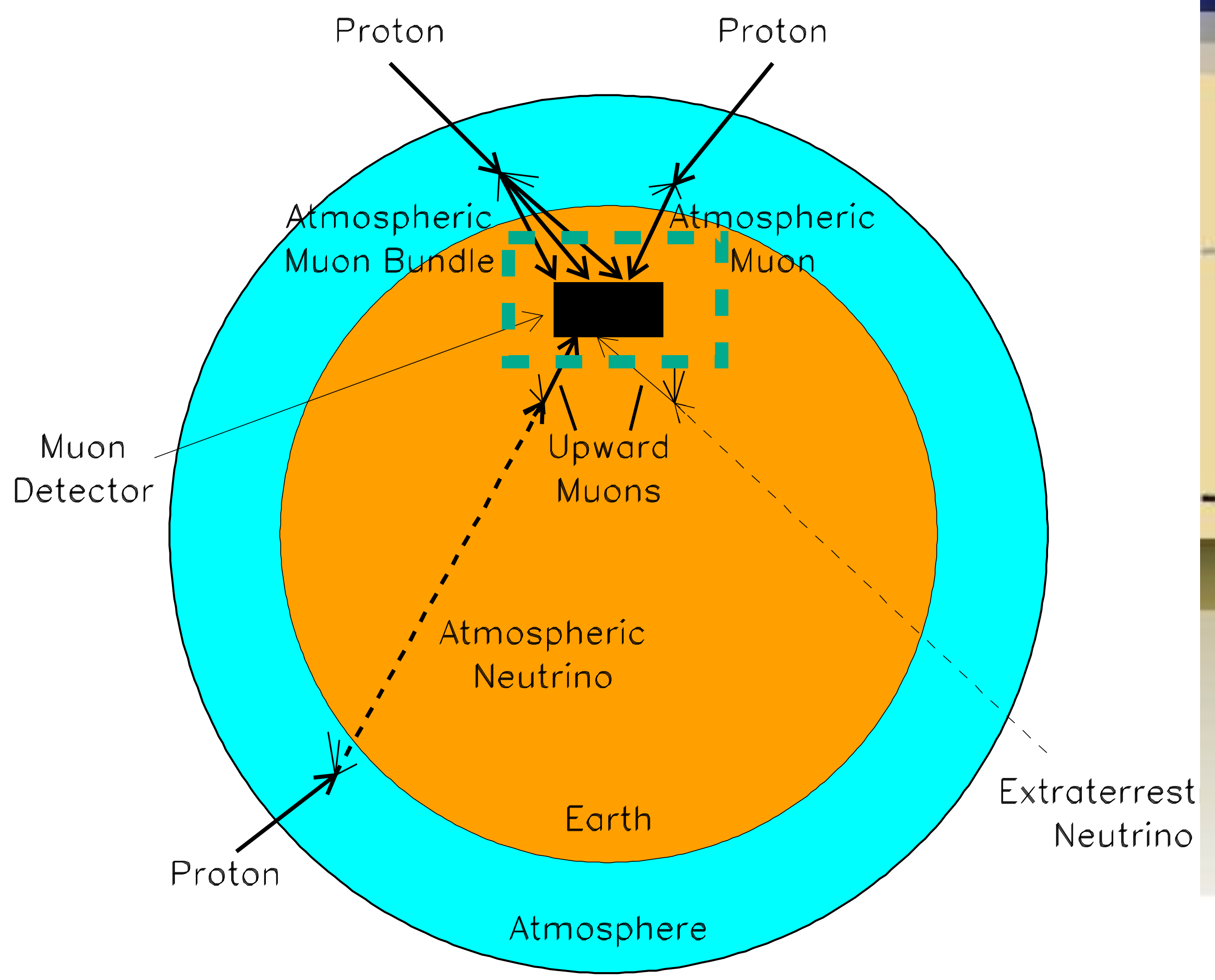
> 50 Institutes
> 250 physicists and engineers

2024 marine campaign starting next week!

Last ARCA marine campaign (2023)

Detection principle

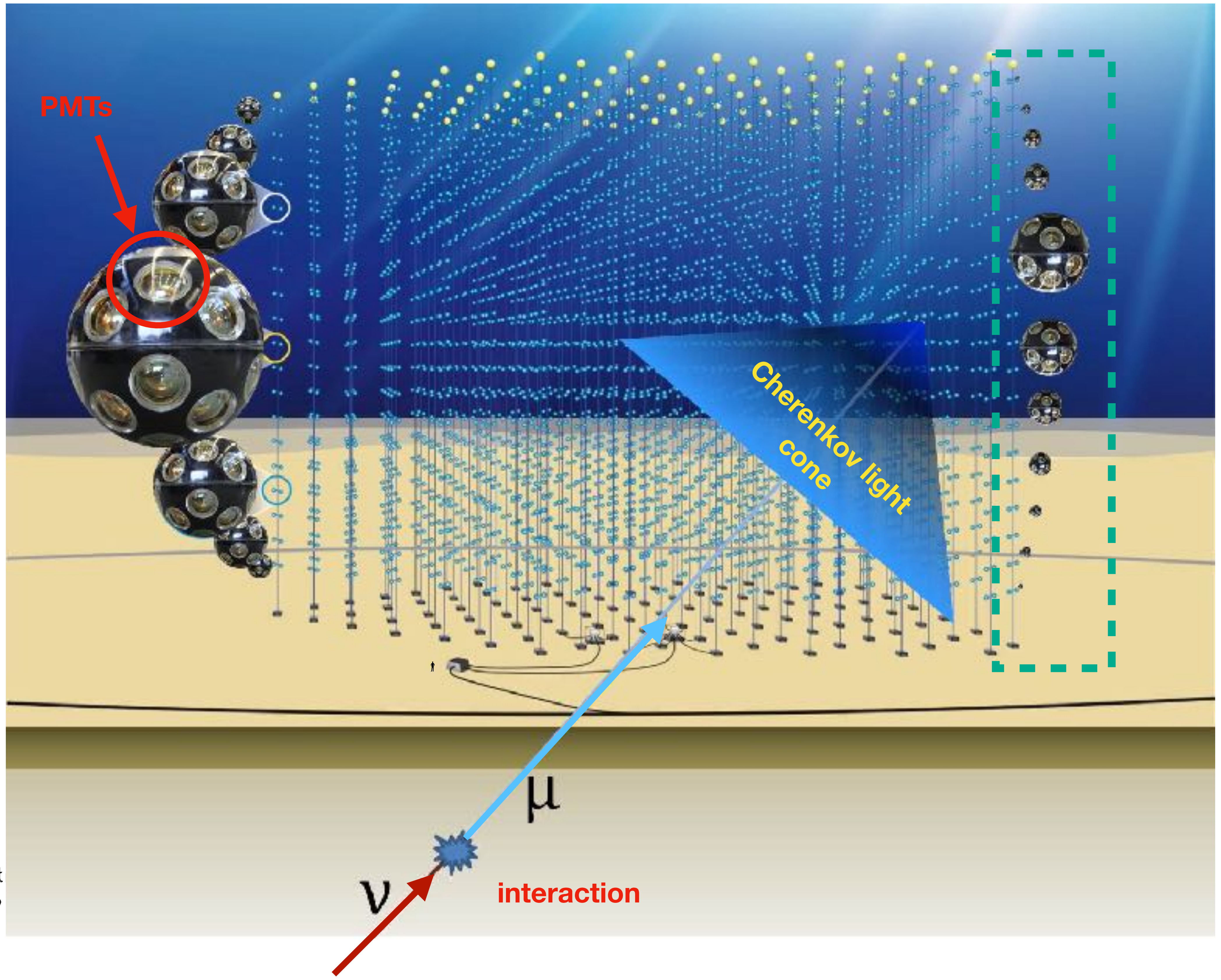


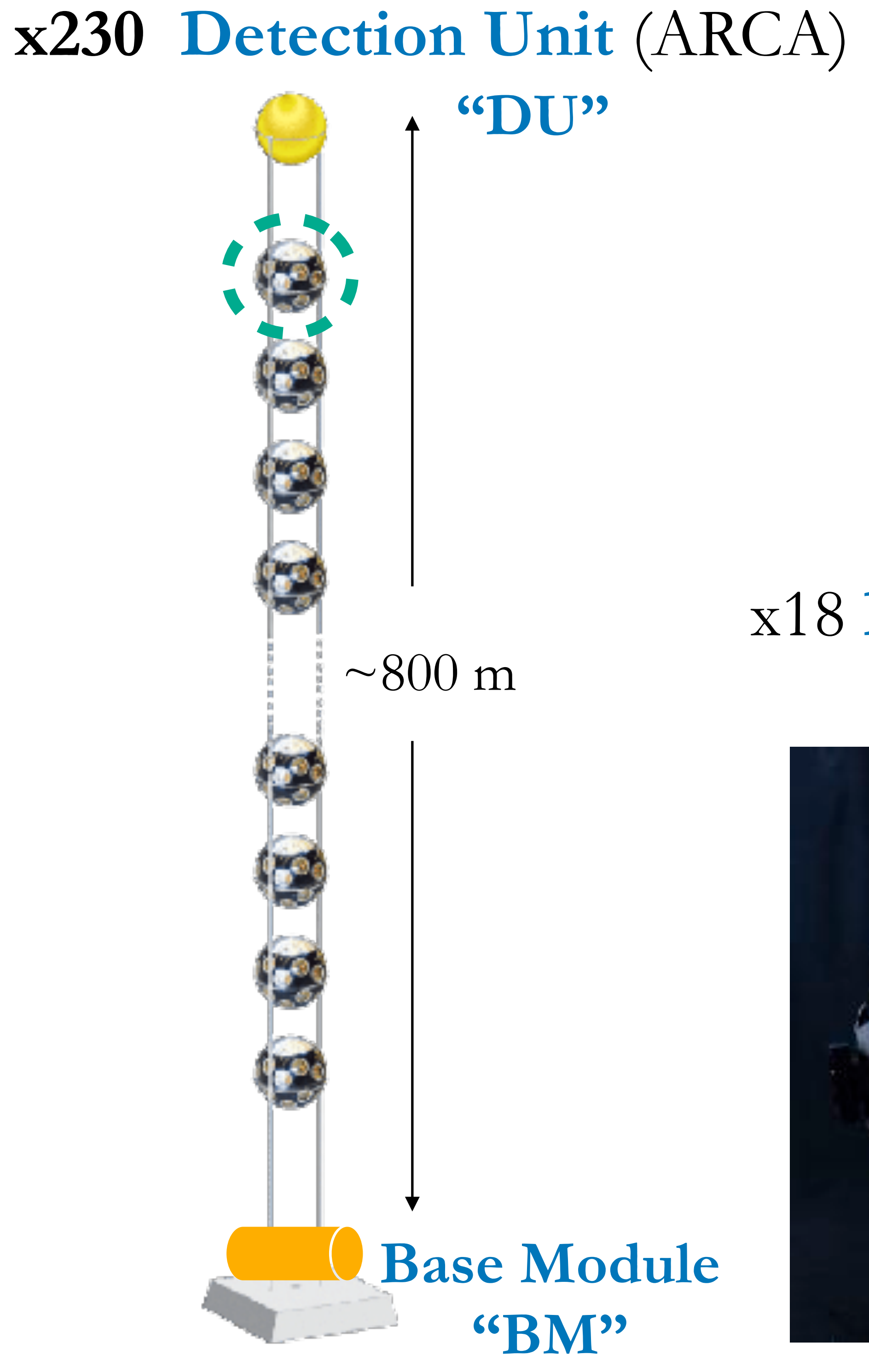


- Reconstruct direction/energy from PMT hits
- Exploit Cherenkov effect in water induced by outgoing leptons

Performances: < 1 ns timing with "White Rabbit" system
< 20 cm positioning accuracy with acoustic pos. system

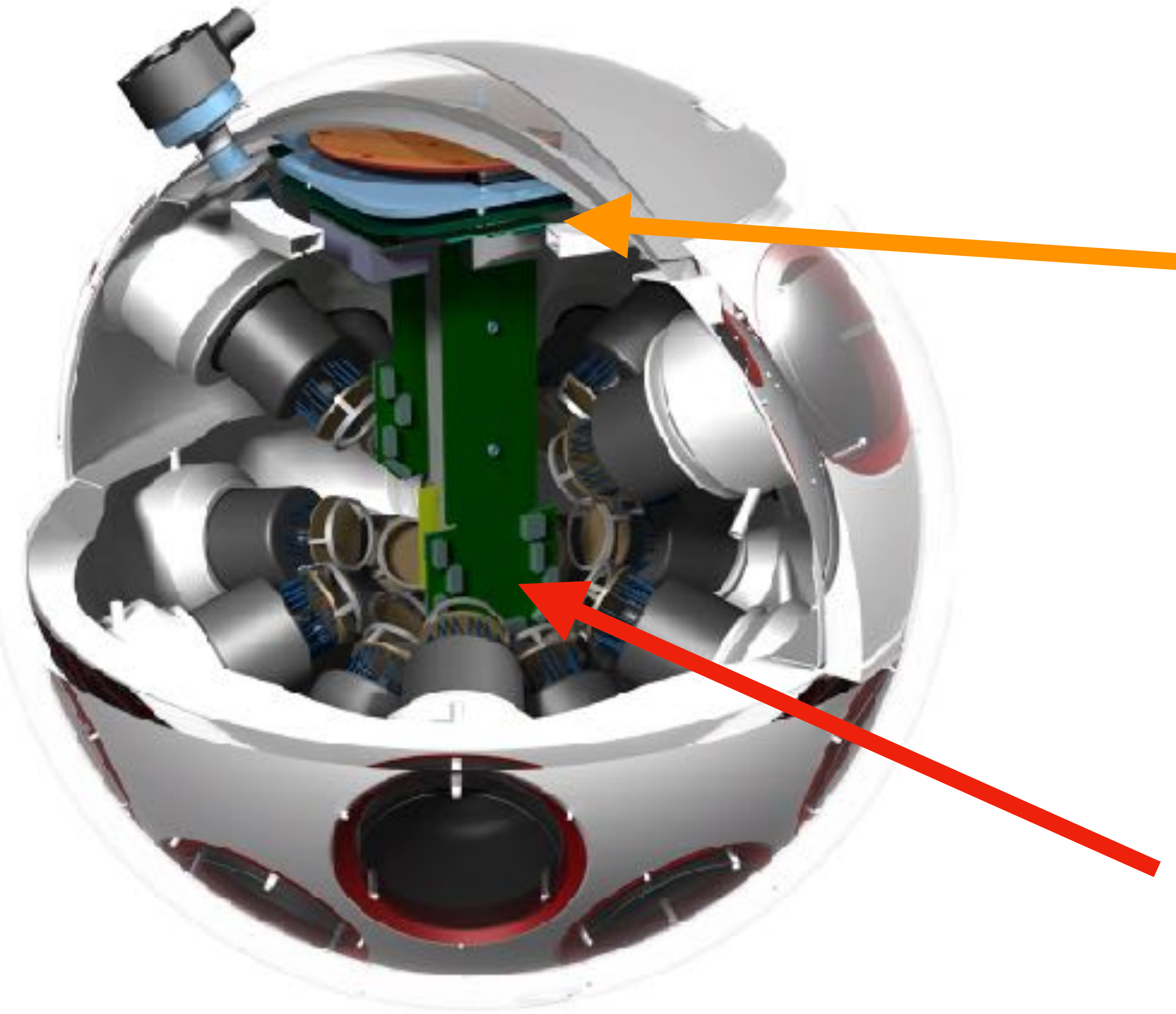
A network of hydrophones and beacons allows, together with DOMs piezo-sensors, to determine the position of DOMs via signal triangulation

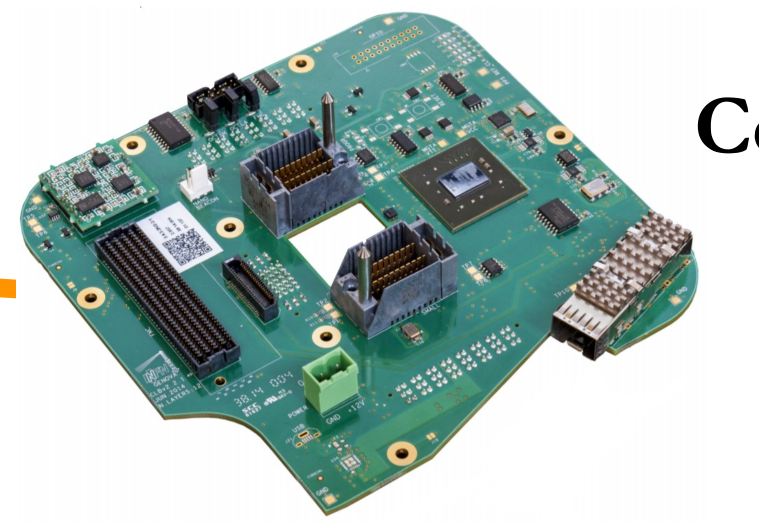




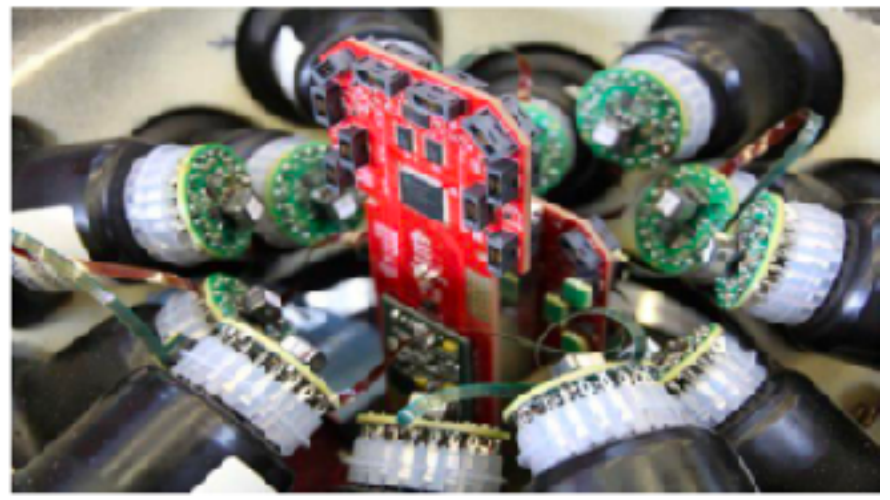
x18 Digital Optical Module
x31 PMTs "DOM"

The Central Logic Board





Central Logic Board (CLB)

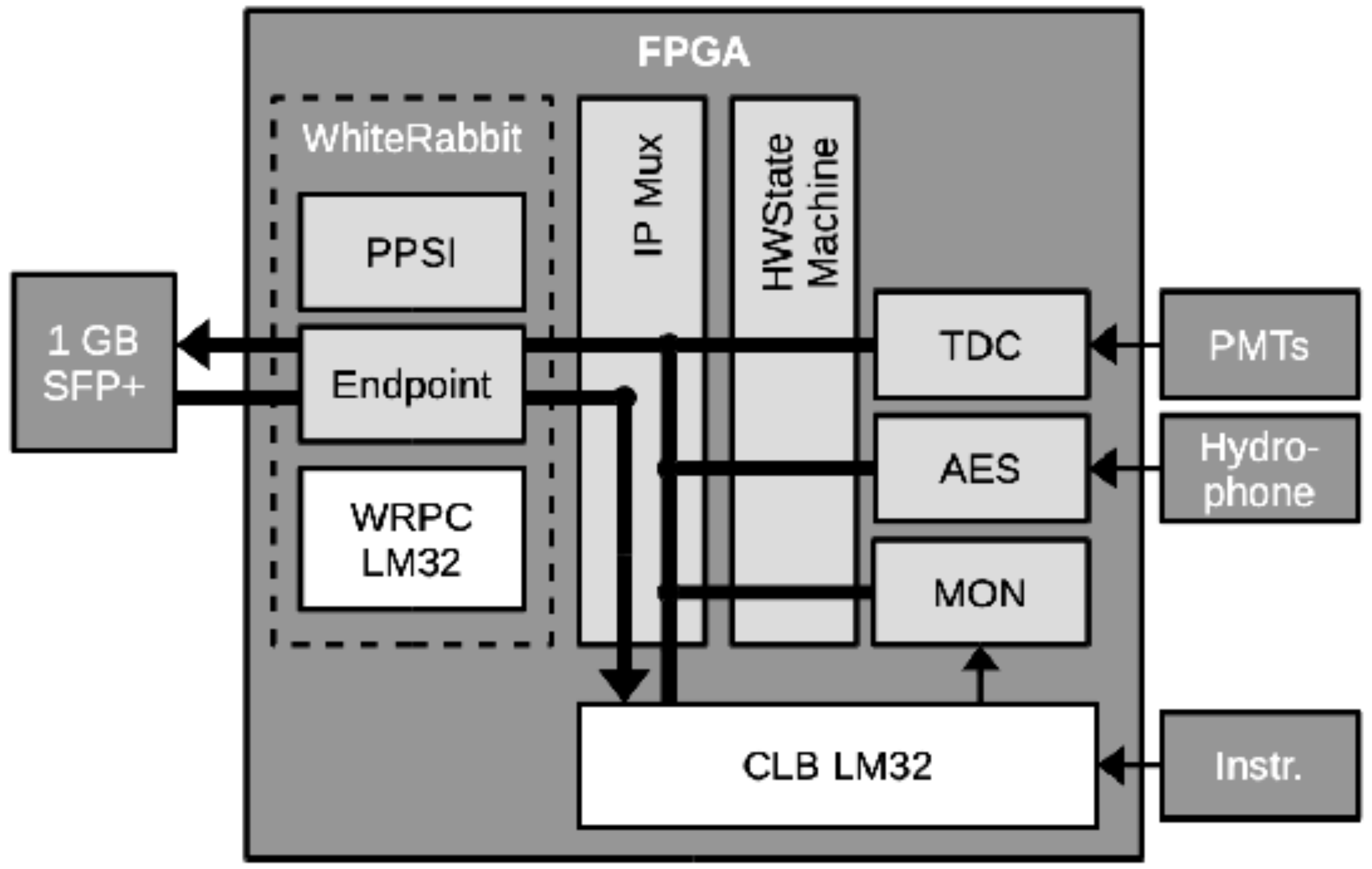


Octopus Large: 19 PMTs and **Small:** 12 PMTs + **piezoelectric** sensor (for acoustic positioning)

Central Logic Board firmware architecture

Two LM32 cores
- **WhiteRabbit** core for timing control
- KM3NeT CLB core for **DAQ control/ instrumentation readout**

Three DAQ modules
- Time to Digital Converter (**TDC**) - from PMTs
- **AES**-standard receiver - from hydrophones
- **MON**itoring - for performance information



High-level diagram of CLB gateware and network data-path

D.A.Q. requirements

KM3NeT

Physical constraints

- Big volume
- Water optical properties (absorption & scattering of blue-green photons ~50-100 m)
- Good angular resolution O(0.1°) for neutrino astronomy

→ Many optical modules O(1000/km^3)

→ Scalable DAQ design

Detector constraints

- Trigger-less streaming readout
- complex DAQ structures in harsh conditions (mandatory: minimal underwater complexity)

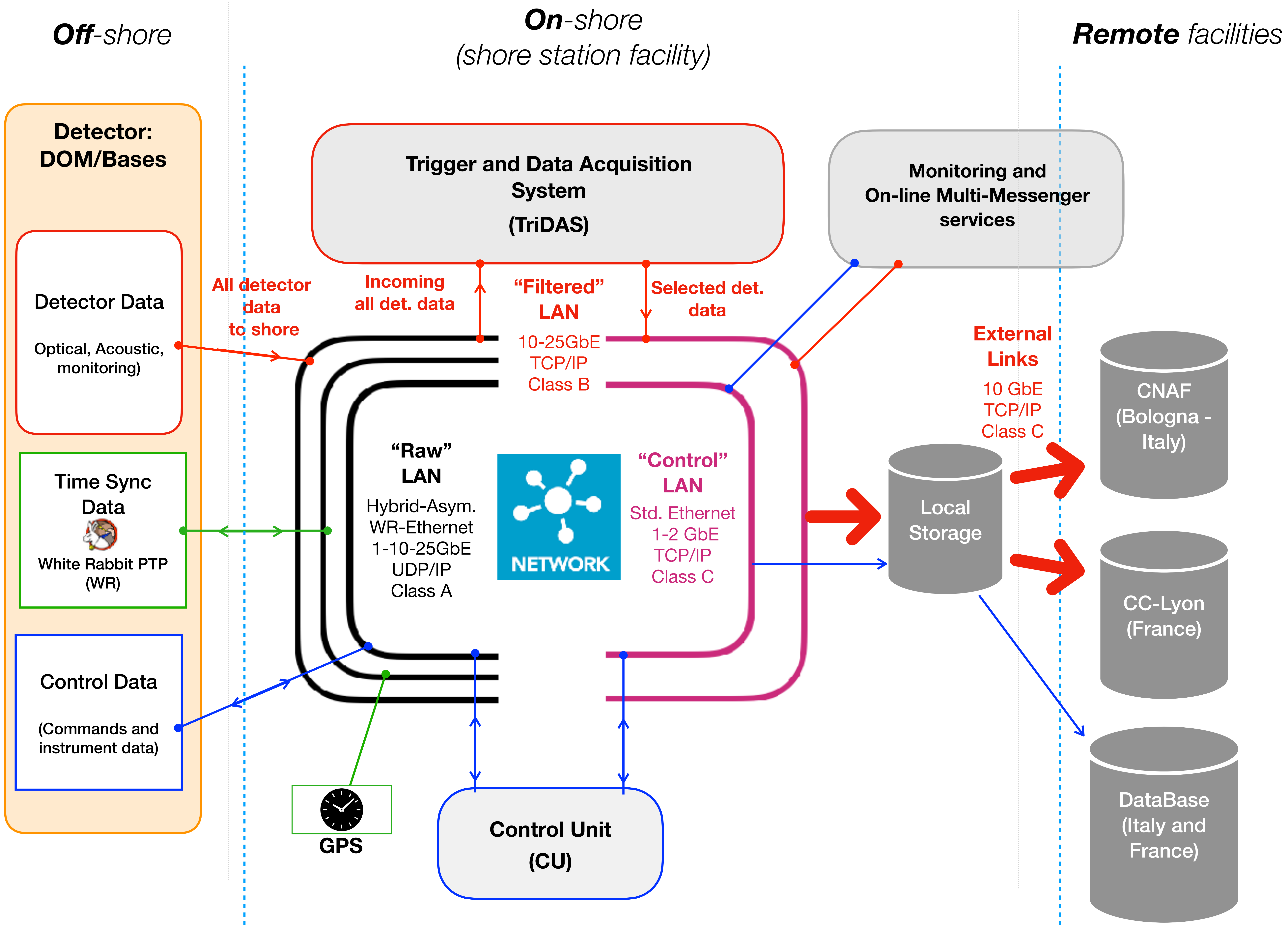
→ **ALL DATA TO SHORE approach**

Drawbacks

- signal-to-noise ratio extremely disfavoured :
- muon rate (atmospheric dominating): O(100) Hz/km^3
- ^{40}K decays (~constant): O(10) kHz/PMT
- Bioluminescence (occasional): O(100) kHz/PMT

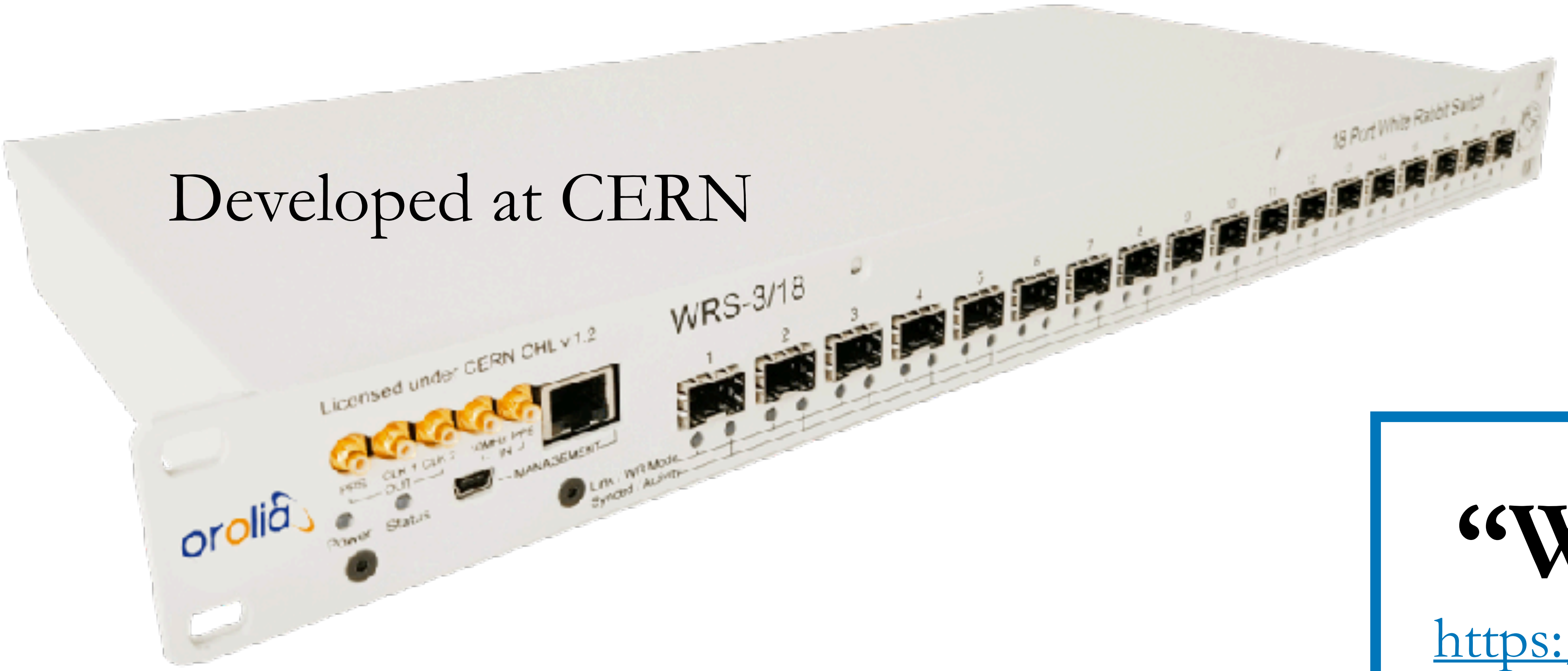
→ High continuous throughput to shore:
→ **large bandwidth switching infrastructure**
→ **strong data reduction**

Network general overview



Timing synchronization

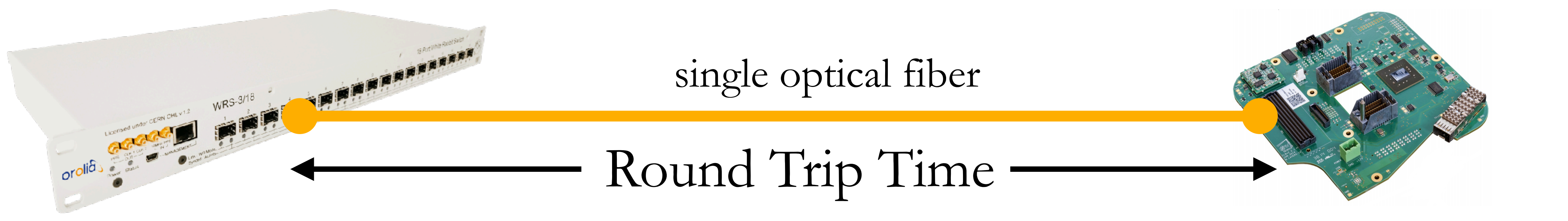
Developed at CERN



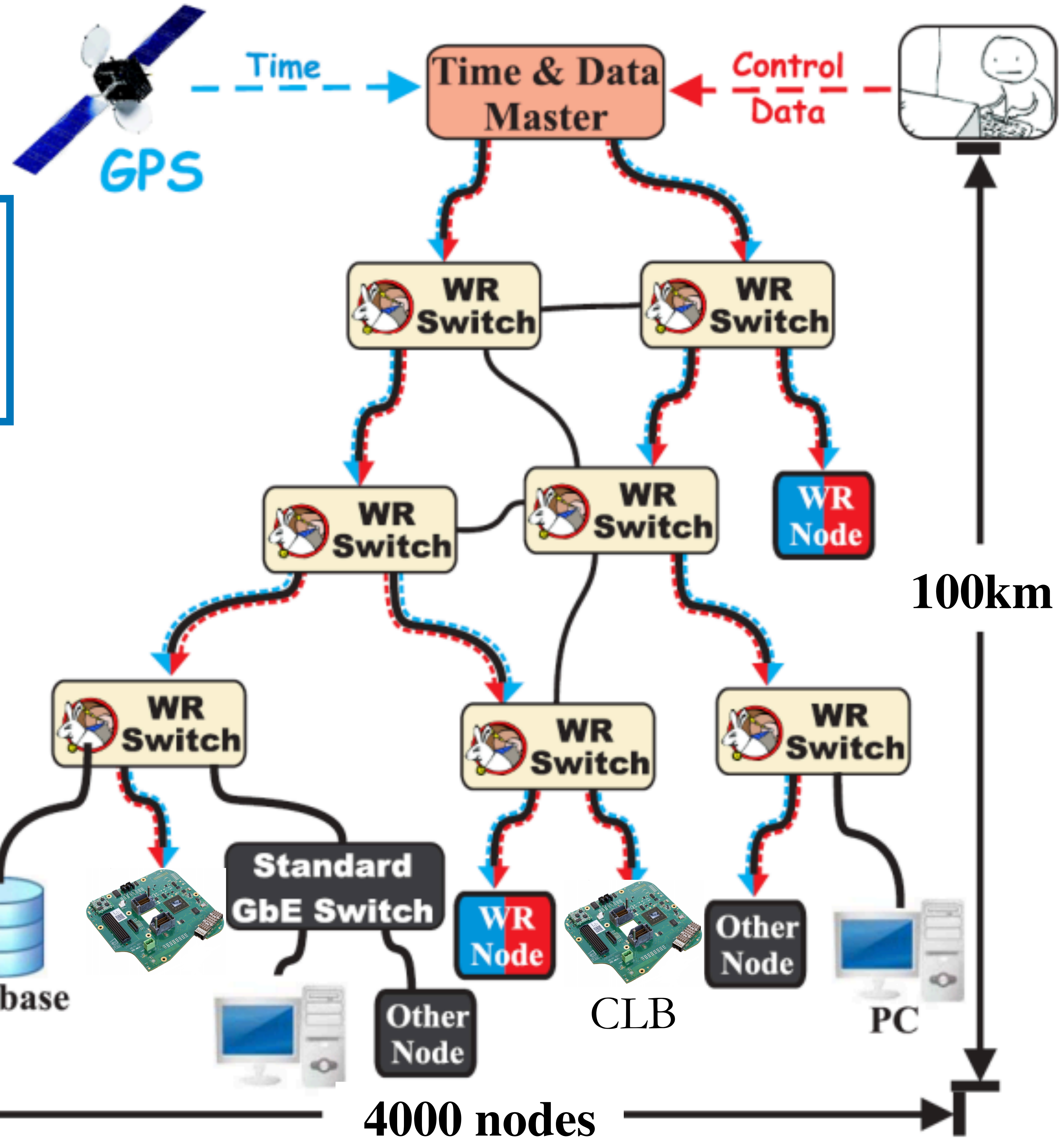
"WRS" = White Rabbit Switch

"White Rabbit"
https://white-rabbit.web.cern.ch/

- Enhanced synchronous ethernet: Ethernet + PTP protocol
- Synchronization: accuracy better than 1 ns (tens of ps of precision)
- Deterministic, reliable, and low latency control-data delivery

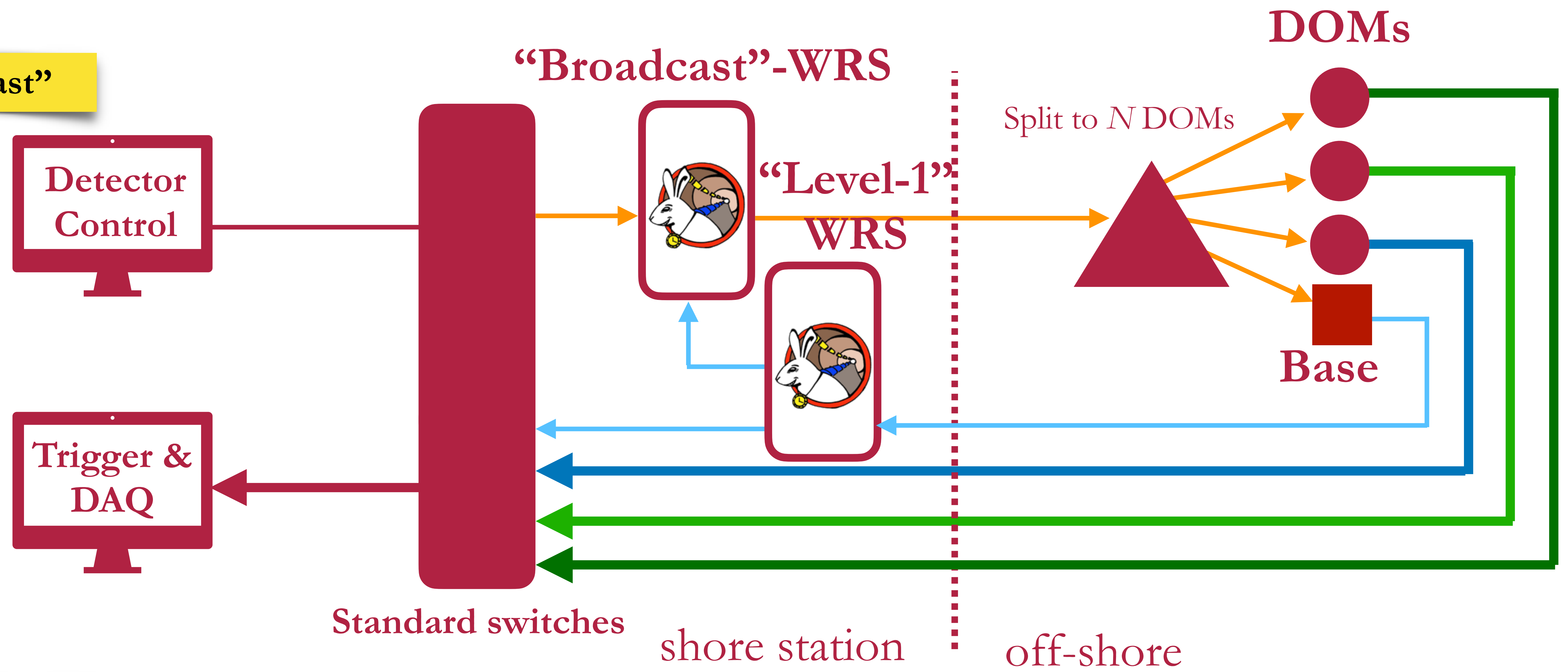


Master (WRS) Slave (CLB/WRS)



Broadcast White Rabbit - topology

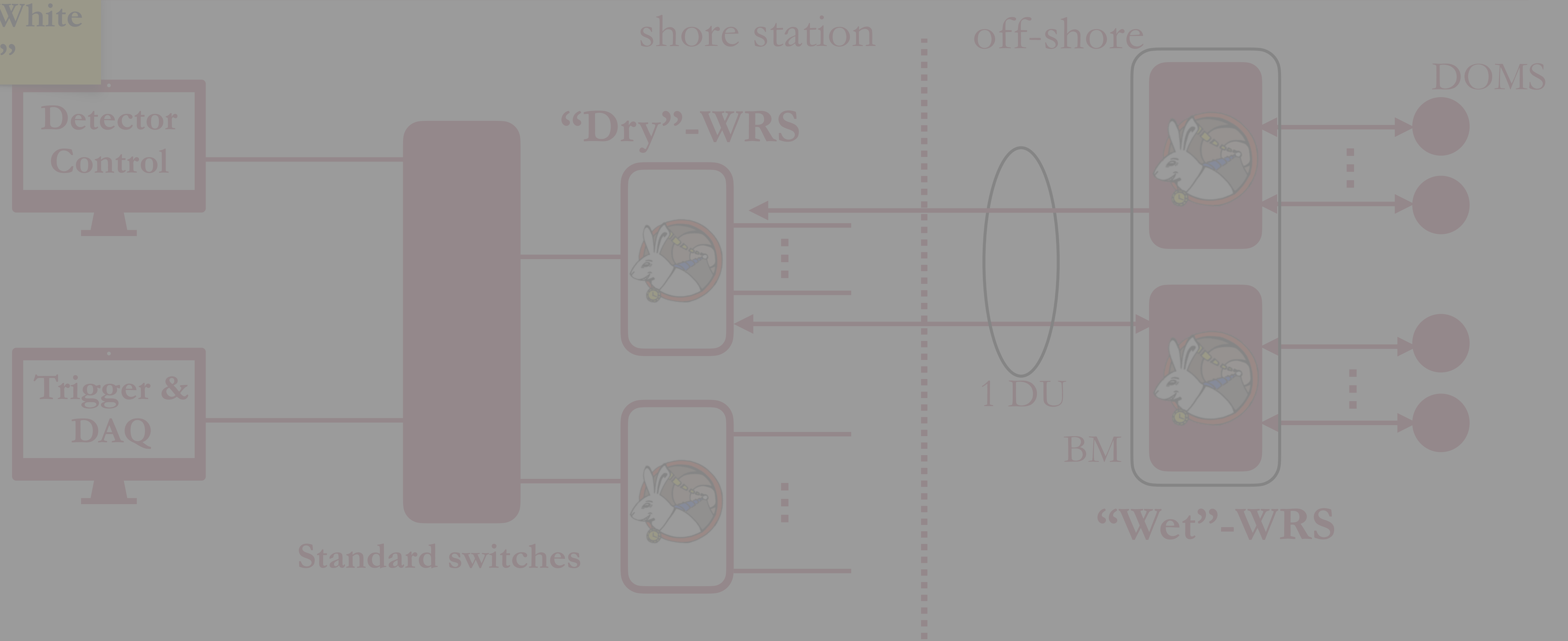
"Broadcast"



"Broadcast" scenario
(current implementation)

- **ARCA:** up to 32 strings
- White Rabbit switch (WRS) **specific customization** required
- **Not scalable** to large number of strings (limited fibers in submarine cables)

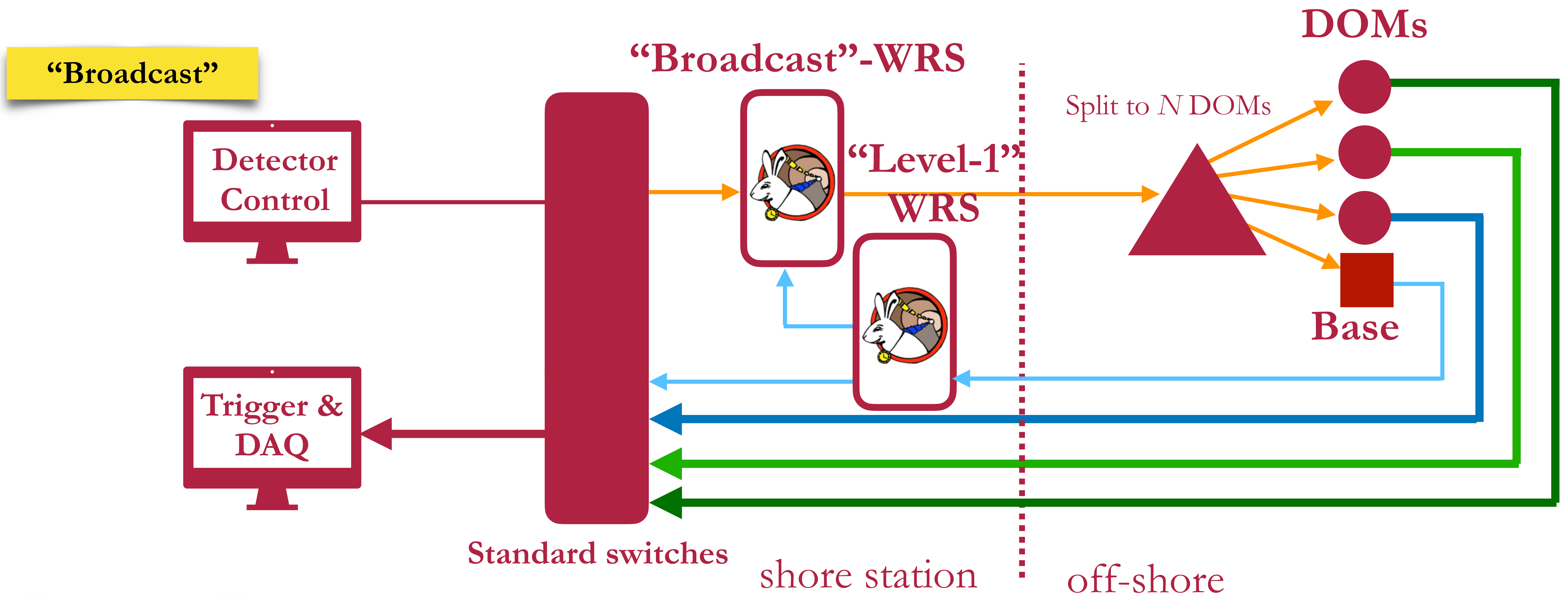
"Standard White Rabbit"



"Standard White Rabbit" scenario

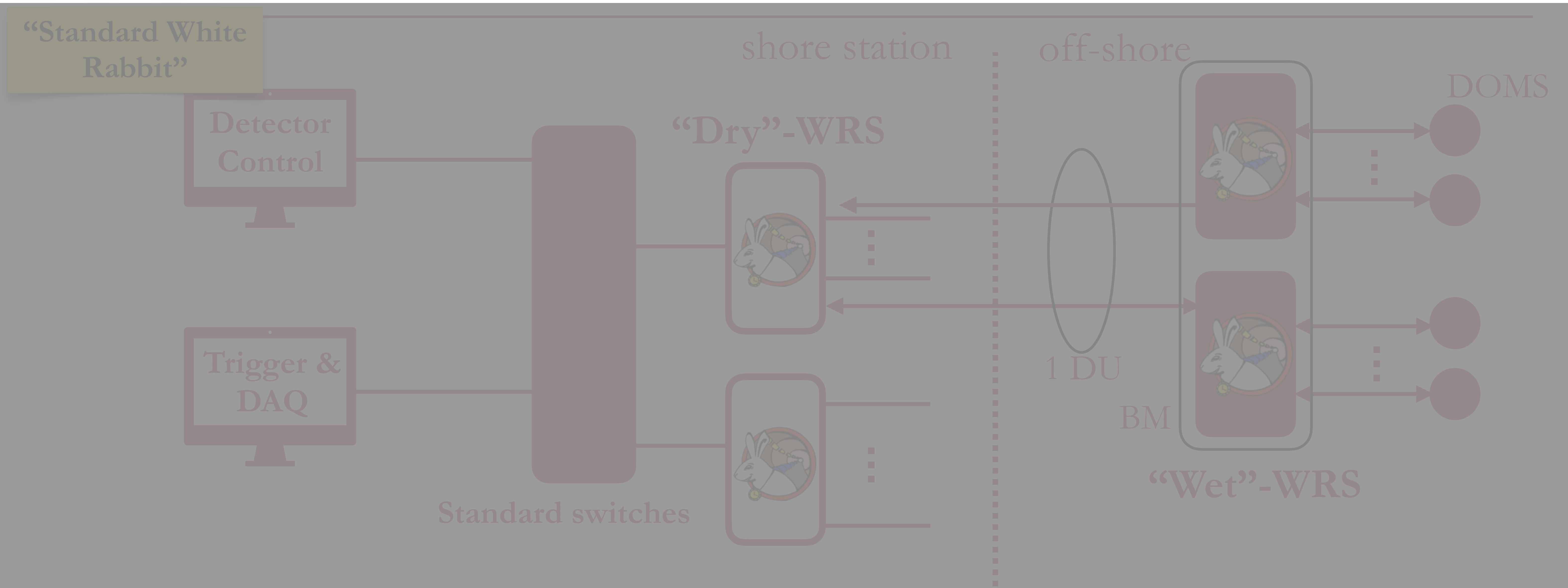
- **Reduction** of a factor 10 in the number of fibers in the submarine cable
- **Standard technology**: easy debug, maintenance and updates
- Requires new software (**"Next-Generation"** firmware for new **CLBs v4/v5**) + hardware

Software Defined Networking



"Broadcast" scenario (current implementation)

- One shared channel to communicate from shore to sea (***broadcast***). One dedicated channel **for each** DOM to send data/replies to shore
- **Software Defined Networking** is needed to prevent fatal conditions due to uncontrolled loops in the switching infrastructure

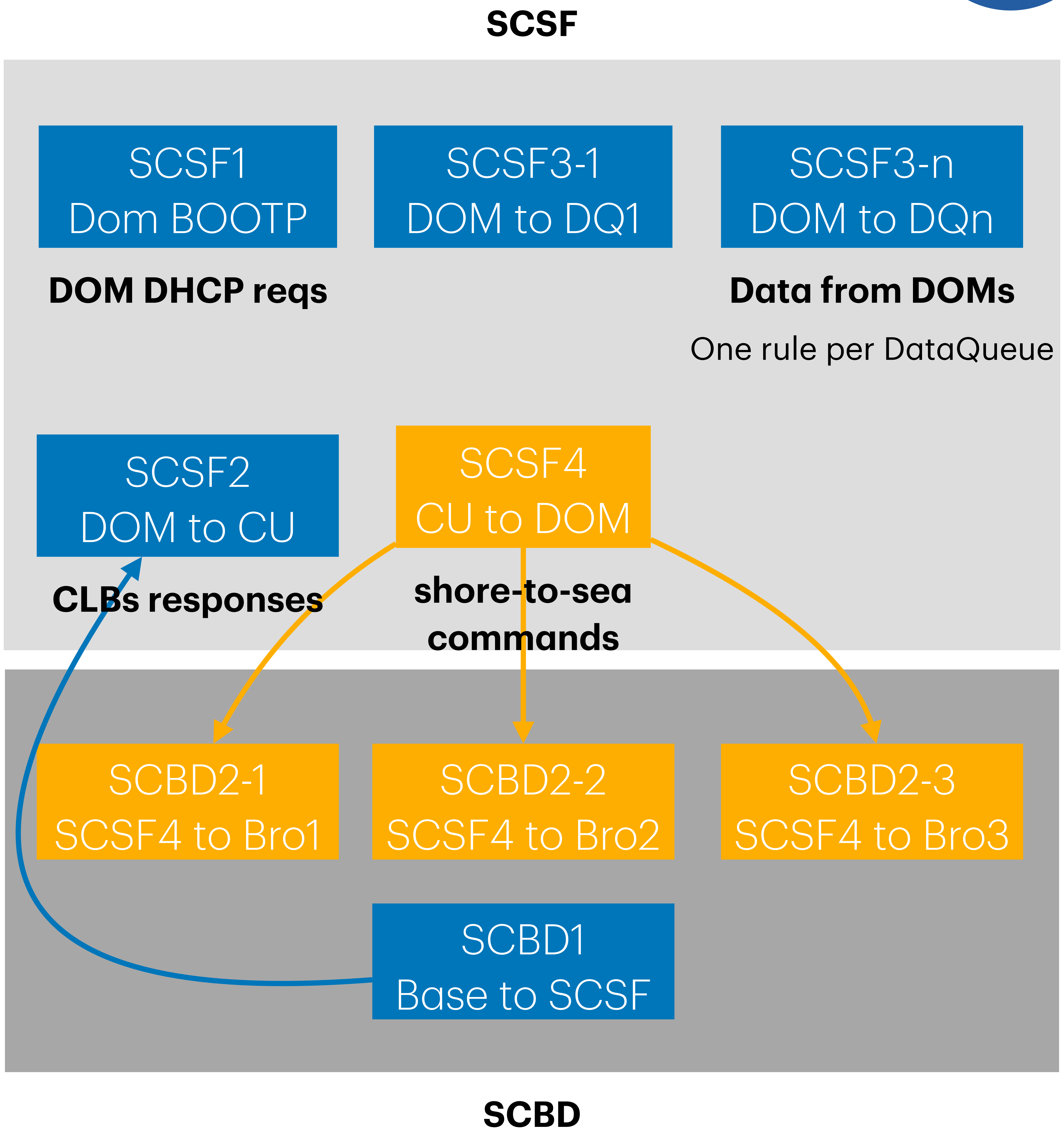


"Standard White Rabbit" scenario

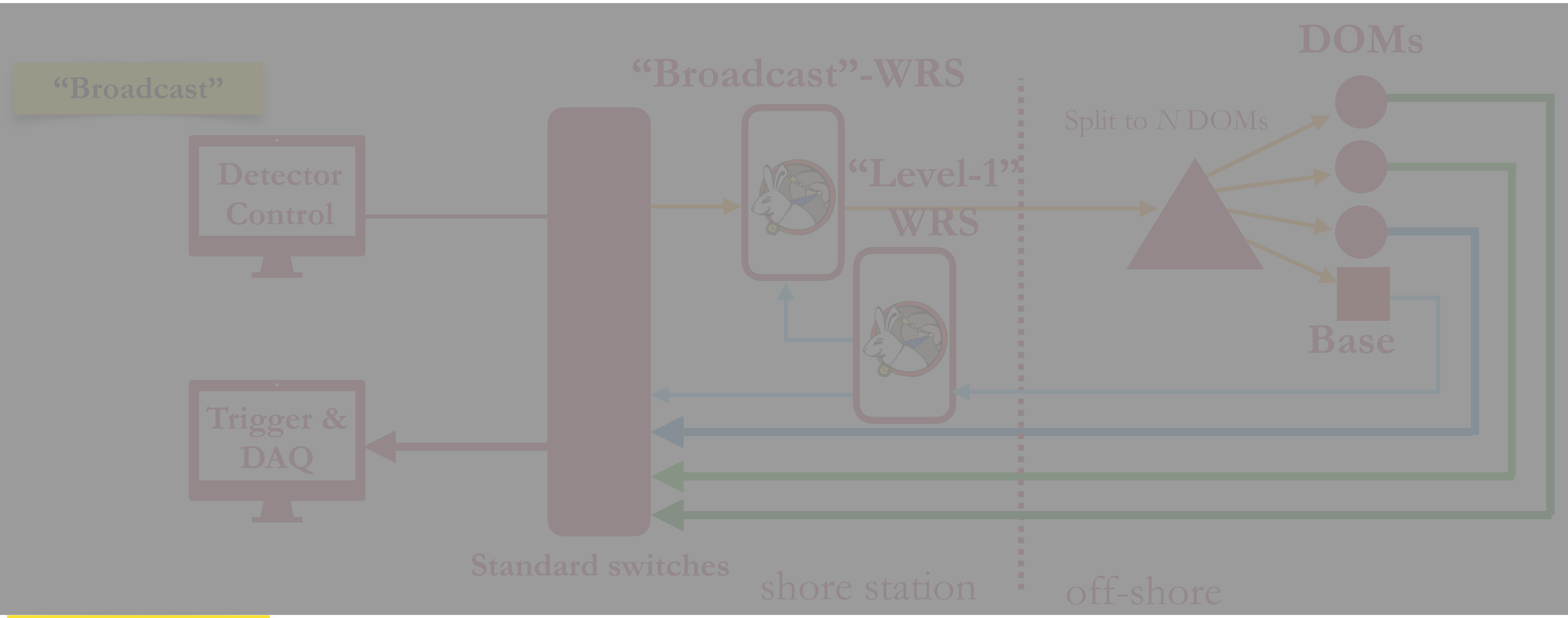
- **Reduction** of a factor 10 in the number of fibers in the submarine cable
- **Standard technology**: easy debug, maintenance and updates
- Requires new software (**"Next-Generation"** firmware for new **CLBs v4/v5**) + hardware

Software Defined Networking

- Asymmetric network topology leads to switch operating in a not-standard configuration: any possible network flow **must be** mapped by an explicit Software Defined Network **rule**
  - if a packet does not match any rule is **discarded**
- Two instances representing the macro group of flows
  - **From Detector to Shore**
  - **From Shore to Detector**

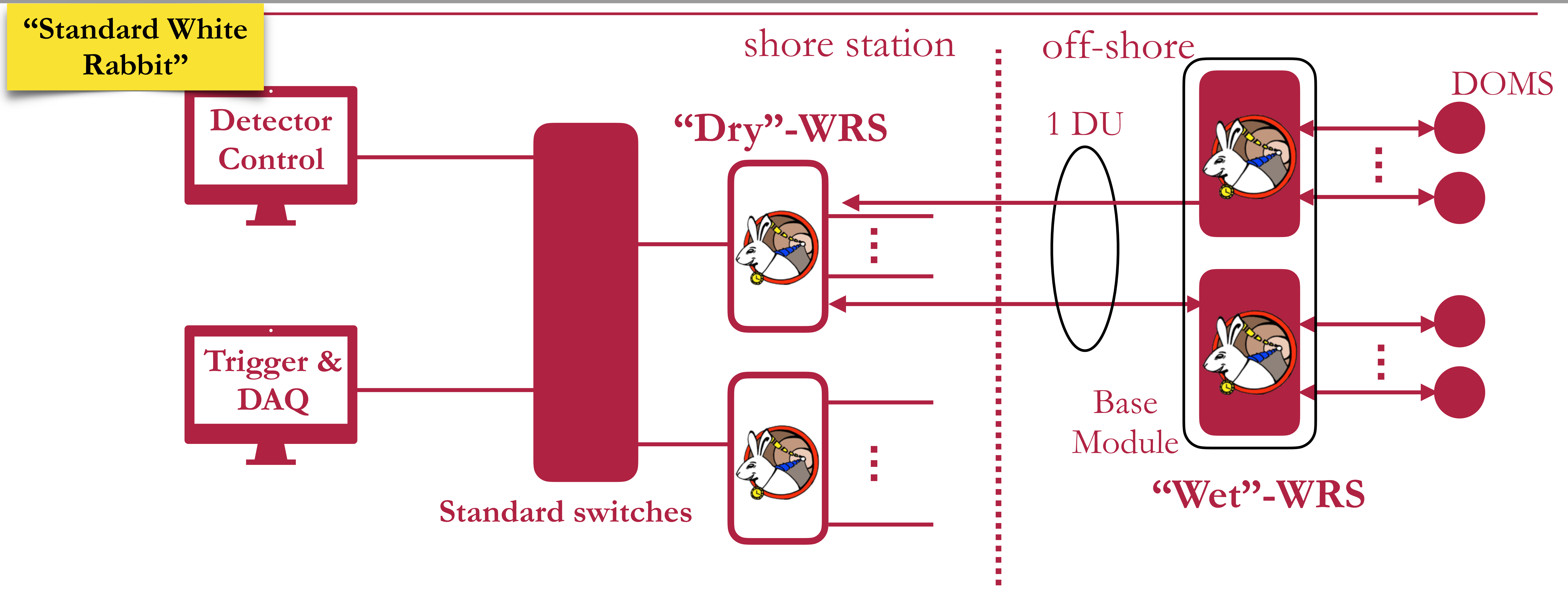


Standard White Rabbit - topology



"Broadcast" scenario (current implementation)

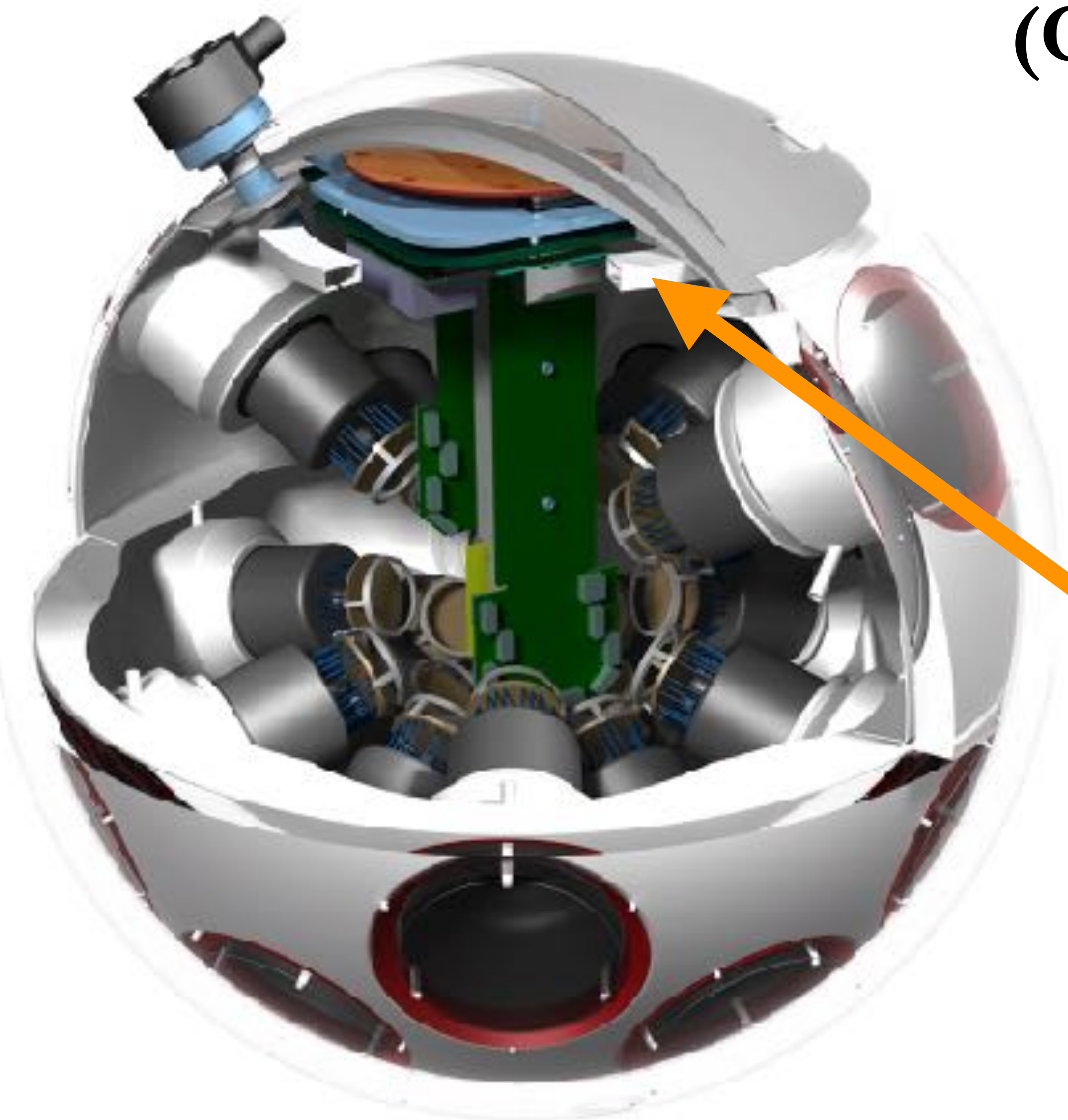
- **ARCA:** up to 32 strings
- White Rabbit switch (WRS) **specific customization** required
- **Not scalable** to large number of strings (limited fibers in submarine cables)



"Standard White Rabbit" scenario

- **ARCA:** beyond 32 strings
- **Reduction** of a factor 10 in the number of fibers in the submarine cable
- **Standard technology**: easy debug, maintenance and updates
- All 1:1 connections

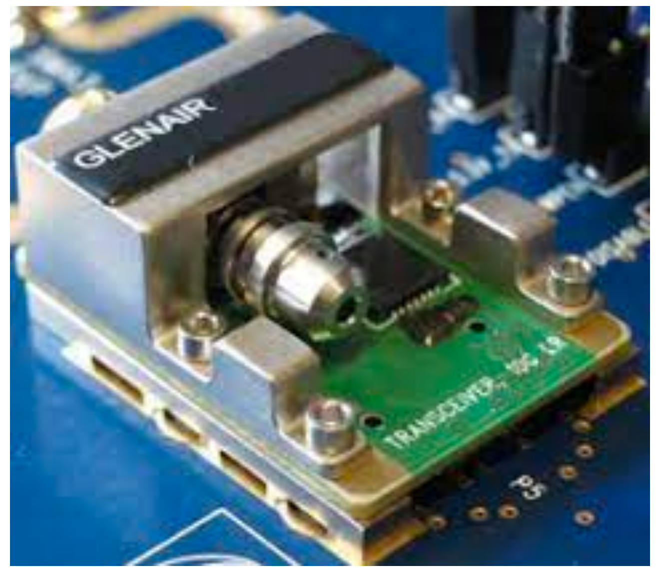
Standard White Rabbit - DOM



Central Logic Board (CLB)

Old v2

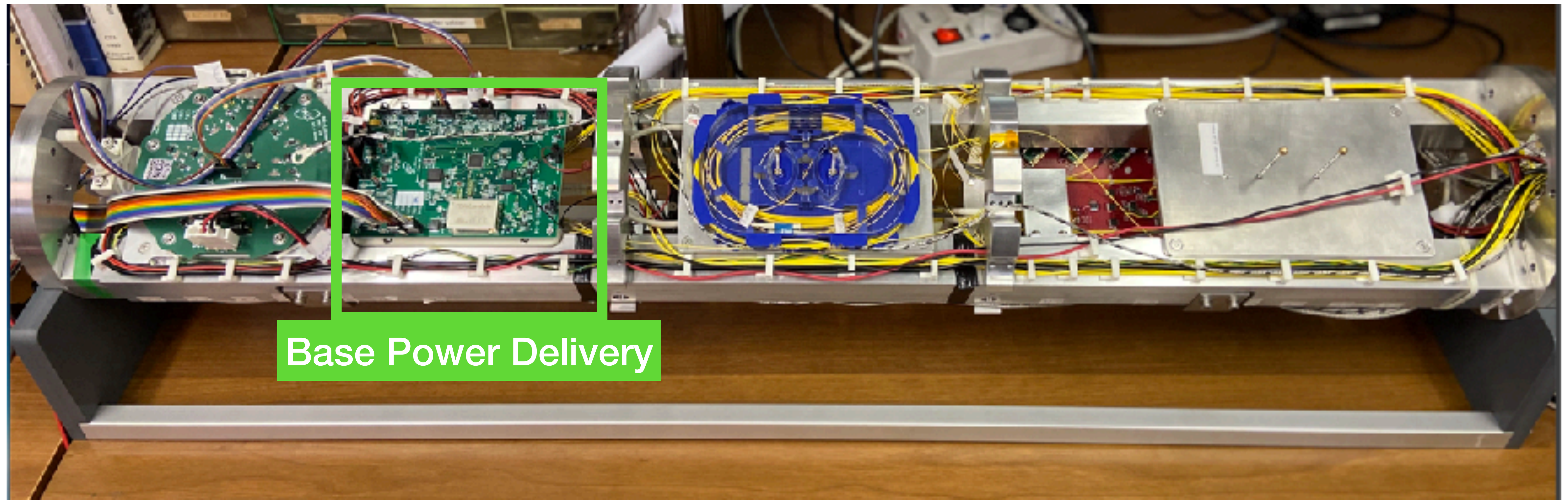
<u>New v4.5</u>
(Designed @IFIC)

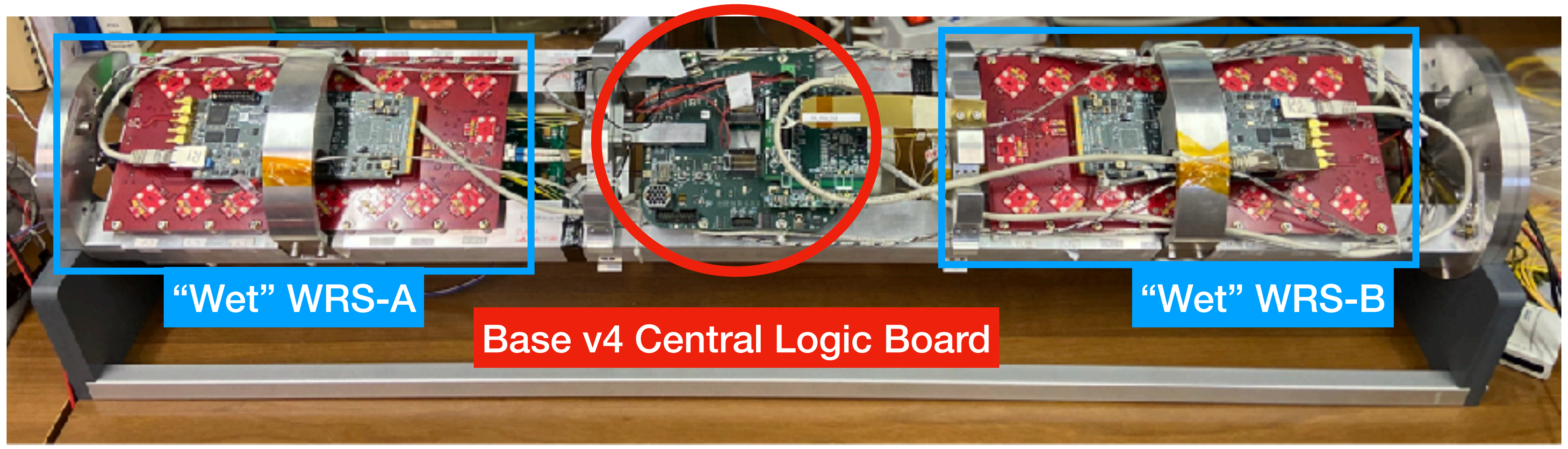


- The only modified component is the **Central Logic Board**
- Main changes: improvement of the clocks routing, new oscillator models → **better phase noise and stability**
- New **Glenair bidirectional transceiver** (higher reliability) and Glenair optical fiber

Standard White Rabbit - Base Module

- **Re-designed form factor**
- **New power boards**
- **Custom White Rabbit Switches**

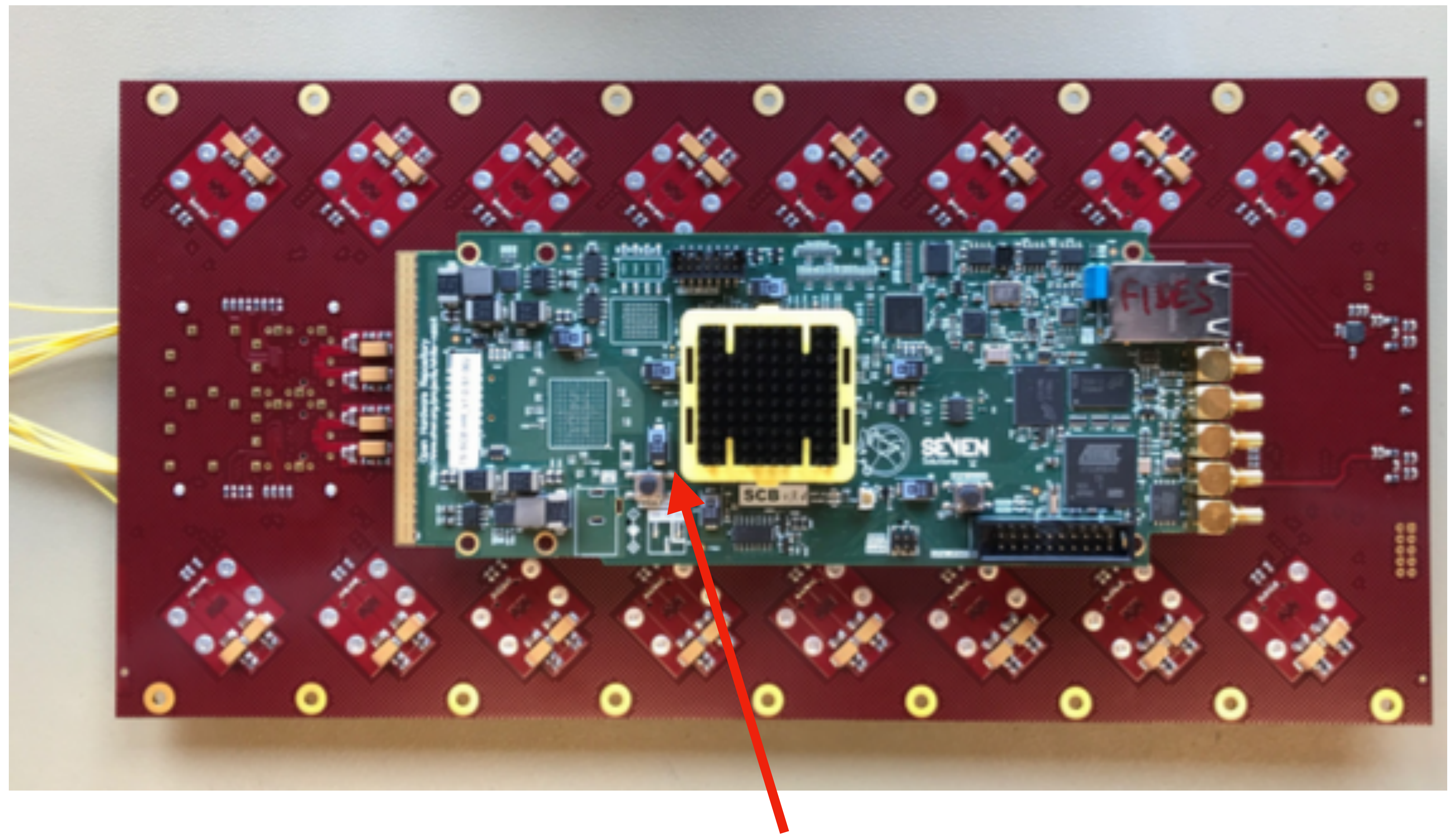




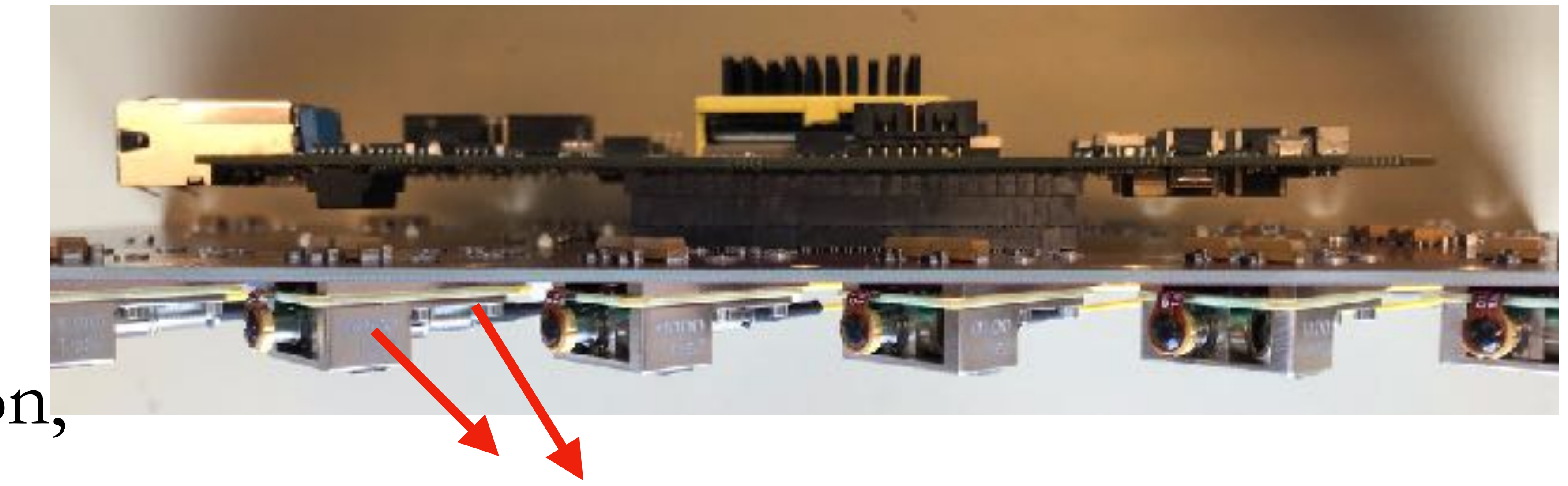
Wet White Rabbit Switch:

- 2 tunable long range SFPs for connection with the on-shore station
- CLB connected to both WWRS (cold **redundancy** applied)
- 23 bidirectional short range transceivers for DOM connections, CLB connection, and inter-WWRS connection (**backup interlink**)

Switching “**Glenair**” backplane KM3NeT (designed @Nikhef)

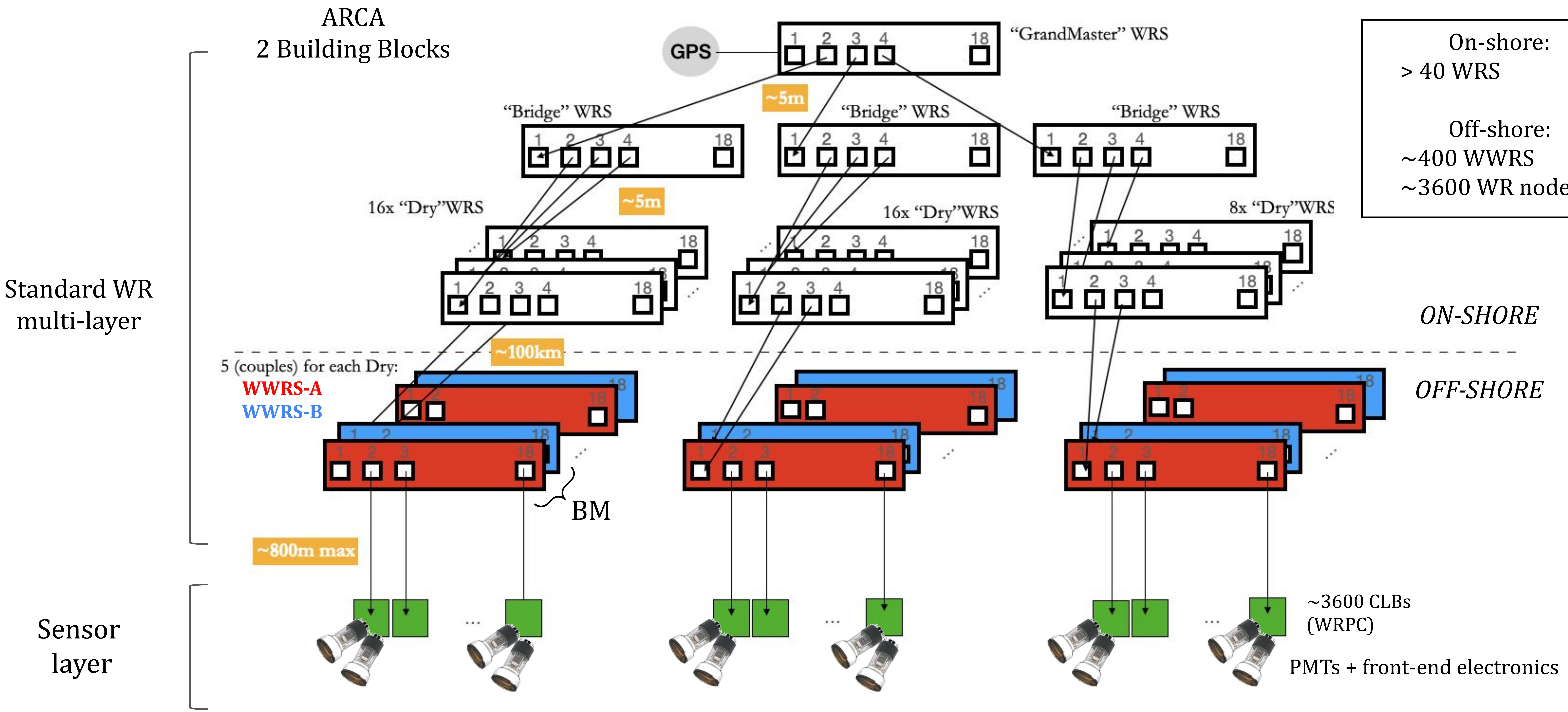


Standard WR Switch Core Board



Custom **Glenair** backplane and Glenair transceivers

KM3NeT White Rabbit network



On-shore:
> 40 WRS

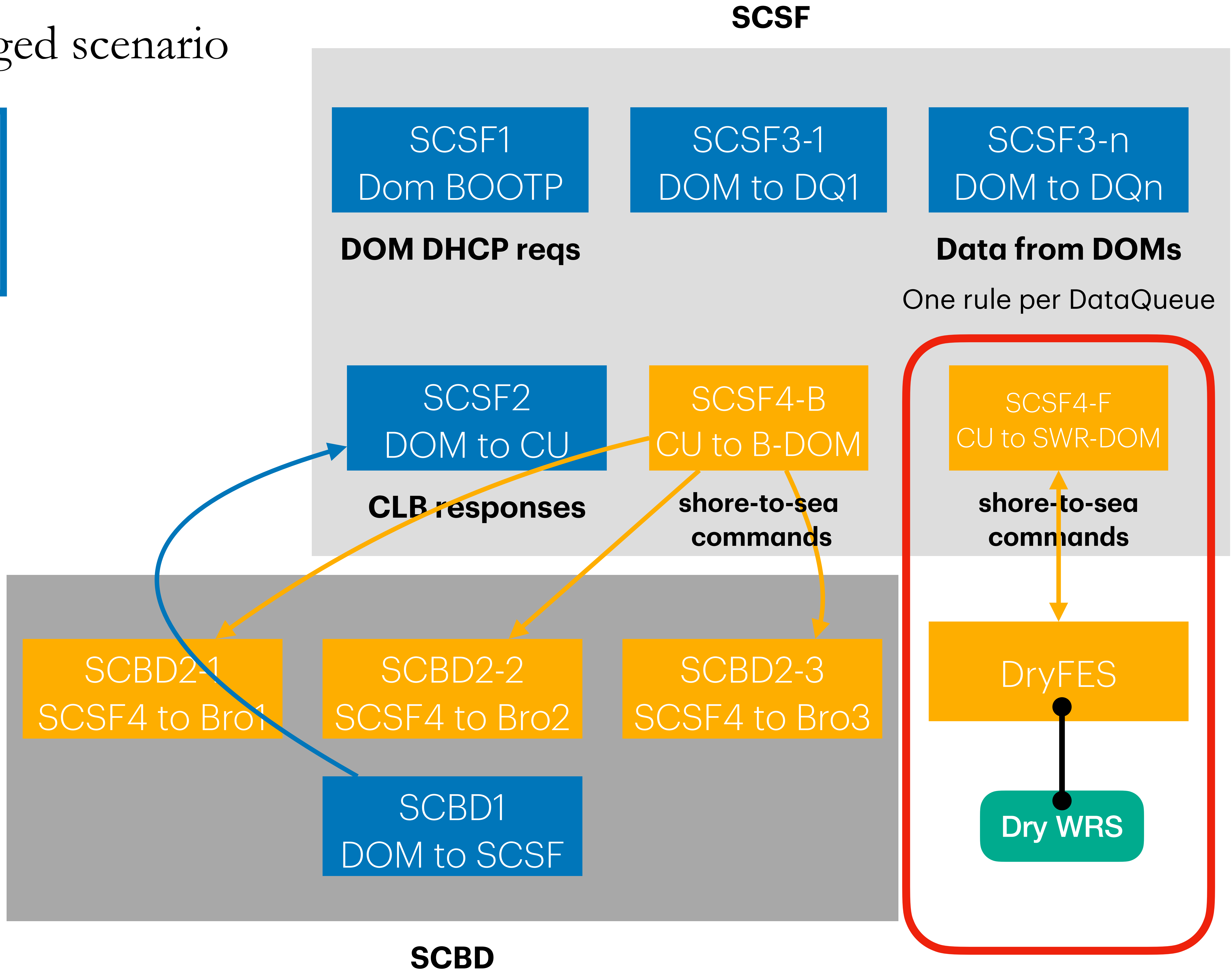
Off-shore:
~400 WWRS
~3600 WR nodes

Software Defined Networking

Broadcast + Standard White Rabbit: merged scenario

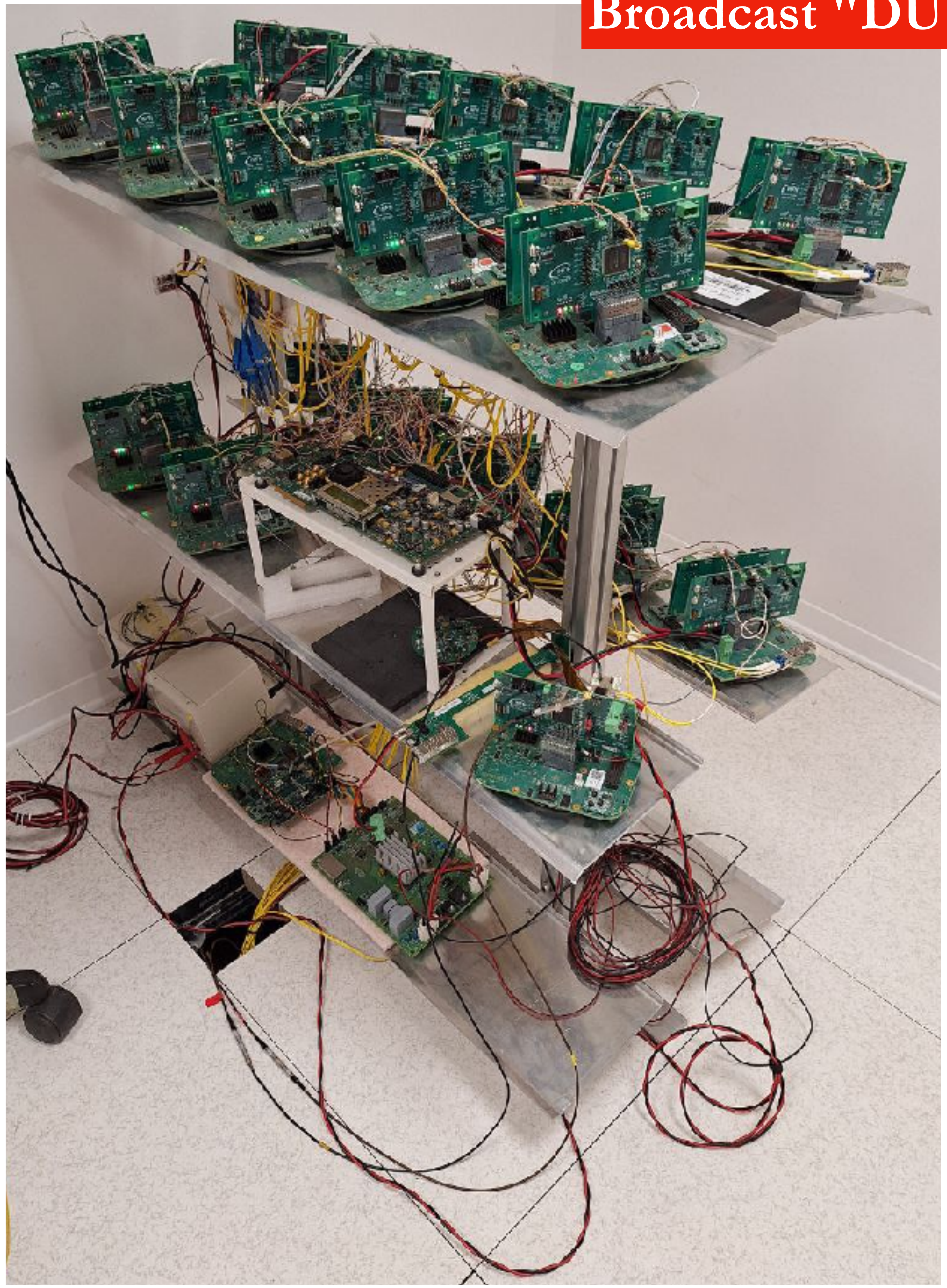
1) Proper routing of the **communication**
2) Proper data **streaming aggregation**

- The SDN implementation has been adapted to include DUs in Standard White Rabbit.
- It required to separate **shore-to-sea** traffic by IP address

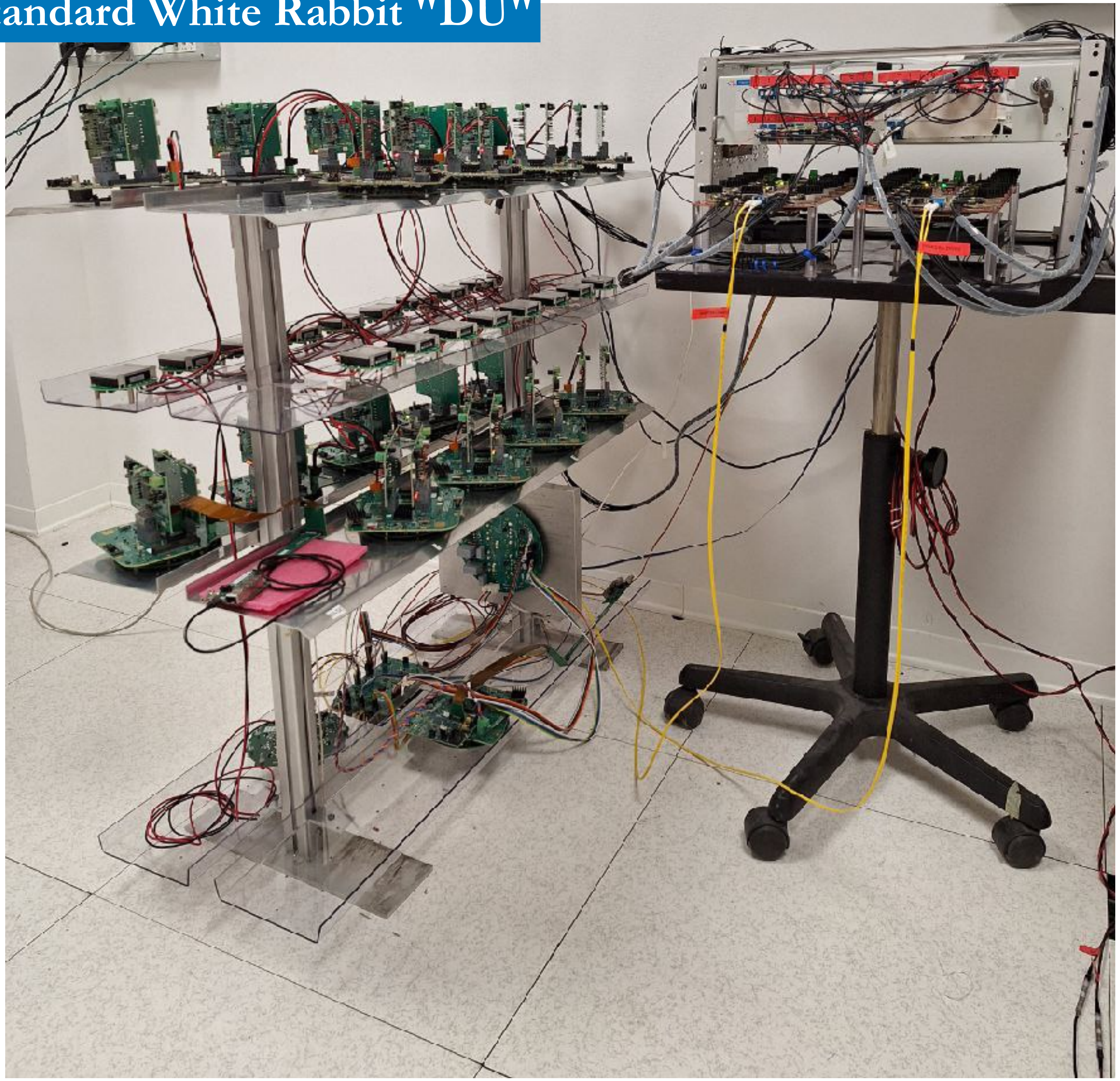


Test environment in Bologna

Broadcast "DU"

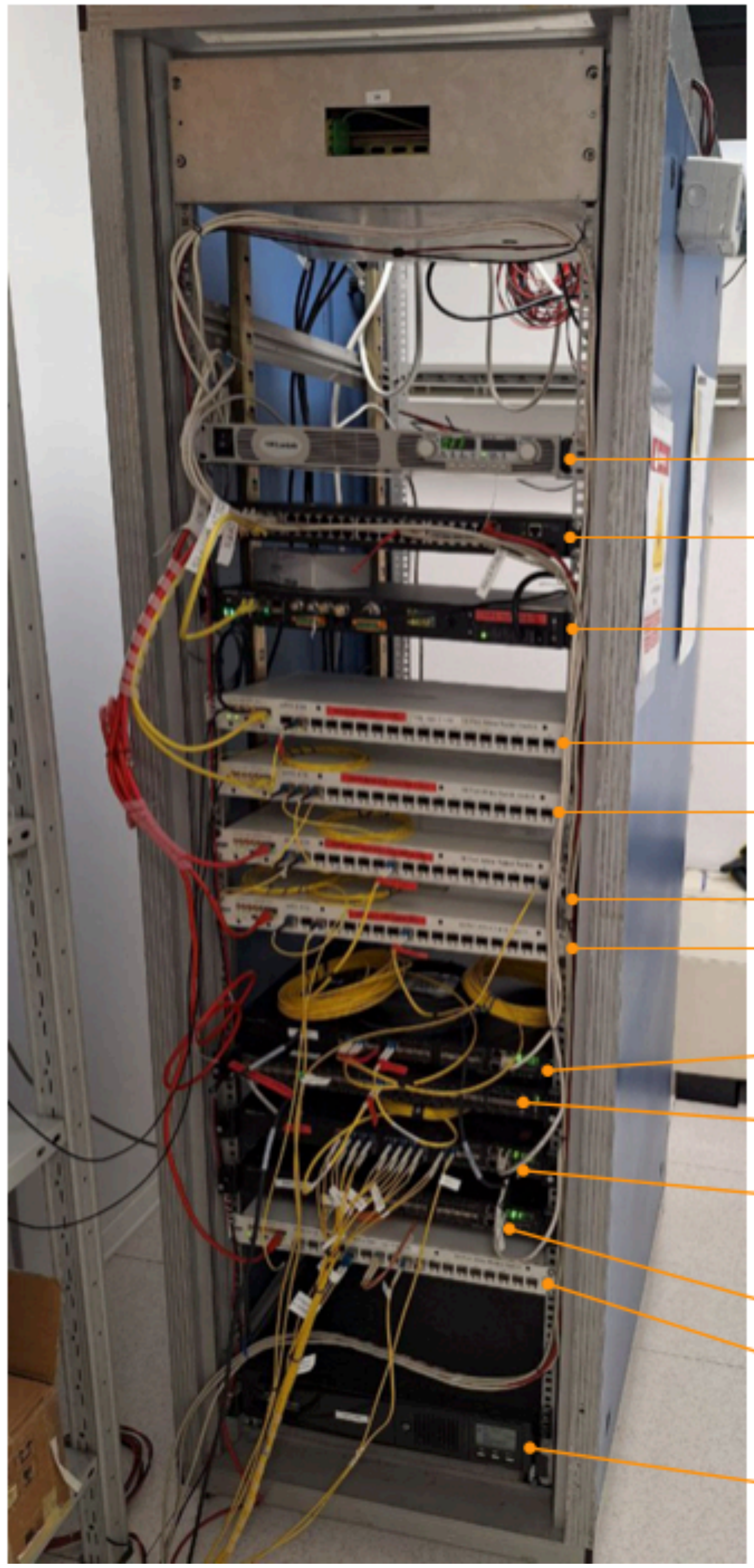


Standard White Rabbit "DU"

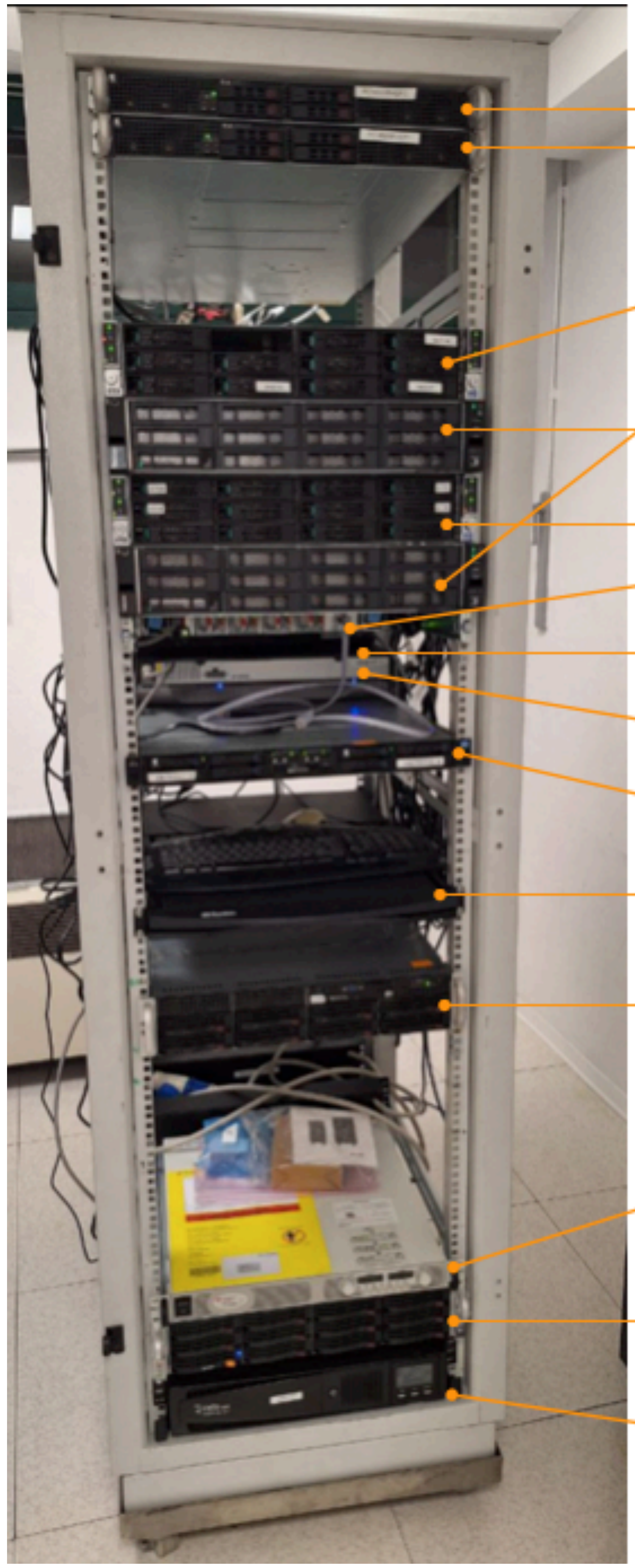


Test environment in Bologna

Rack 1 - networking



Rack 2 - computing



Conclusions

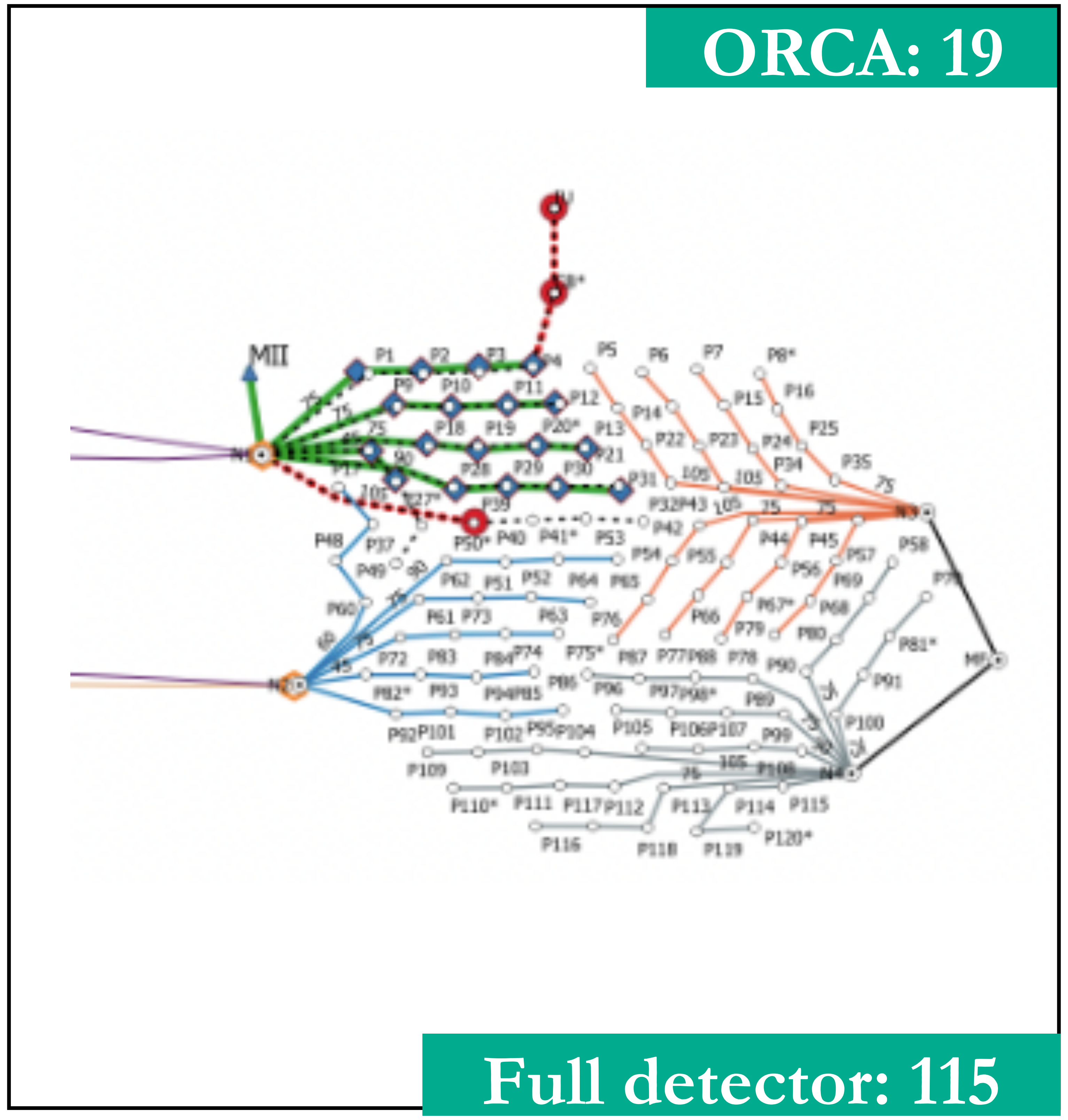
- Coming soon: a much bigger detector with two different architectures running **simultaneously**
- Data Acquisition in the Standard White Rabbit architecture has been successfully tested in Detection Unit integration sites
- Data Acquisition in the hybrid architecture has been successfully tested in Bologna (lab)

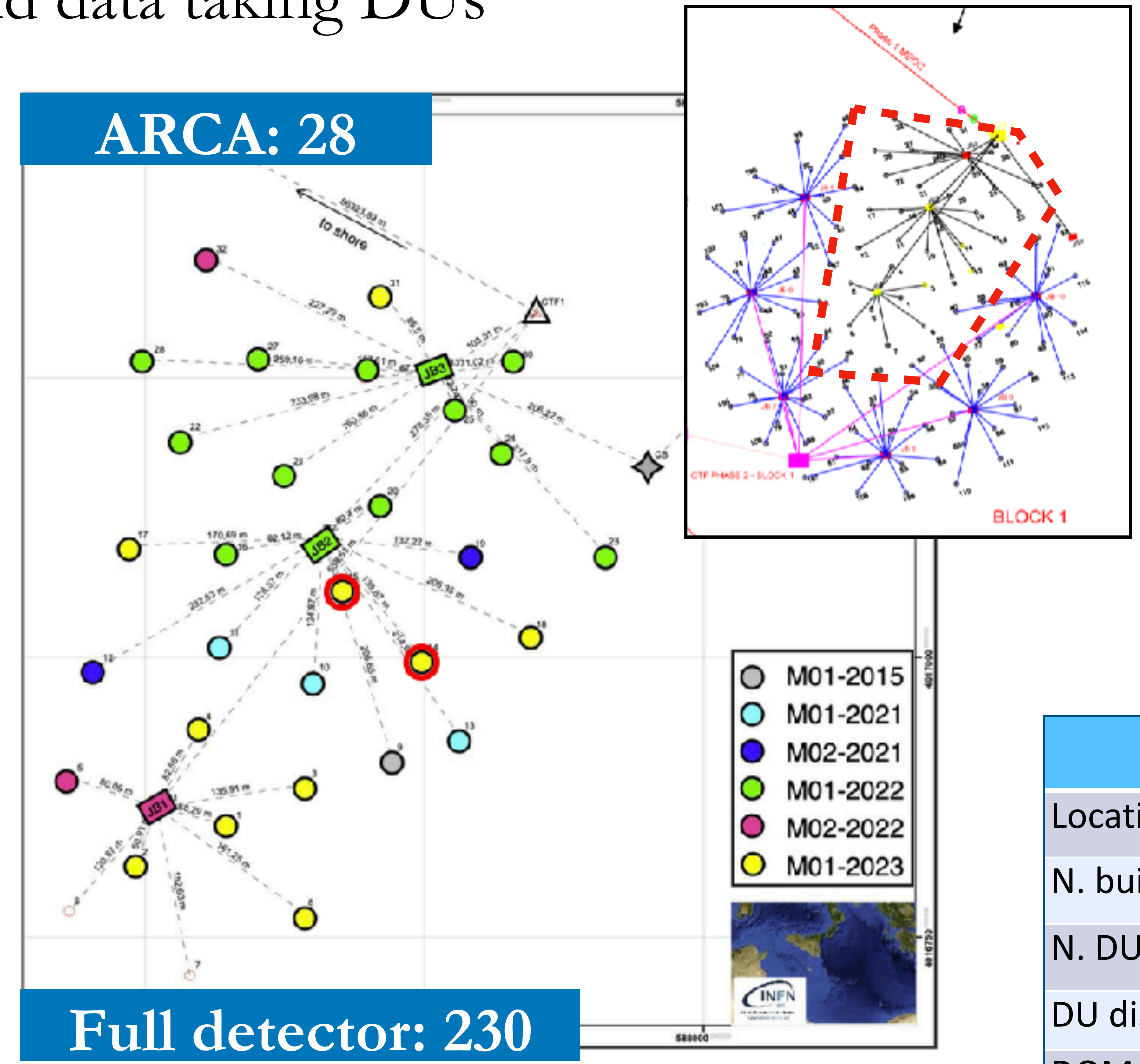
… stay tuned!

Spare slides

KM3NeT Neutrino telescopes

- Current sea floor maps of deployed and data taking DUs

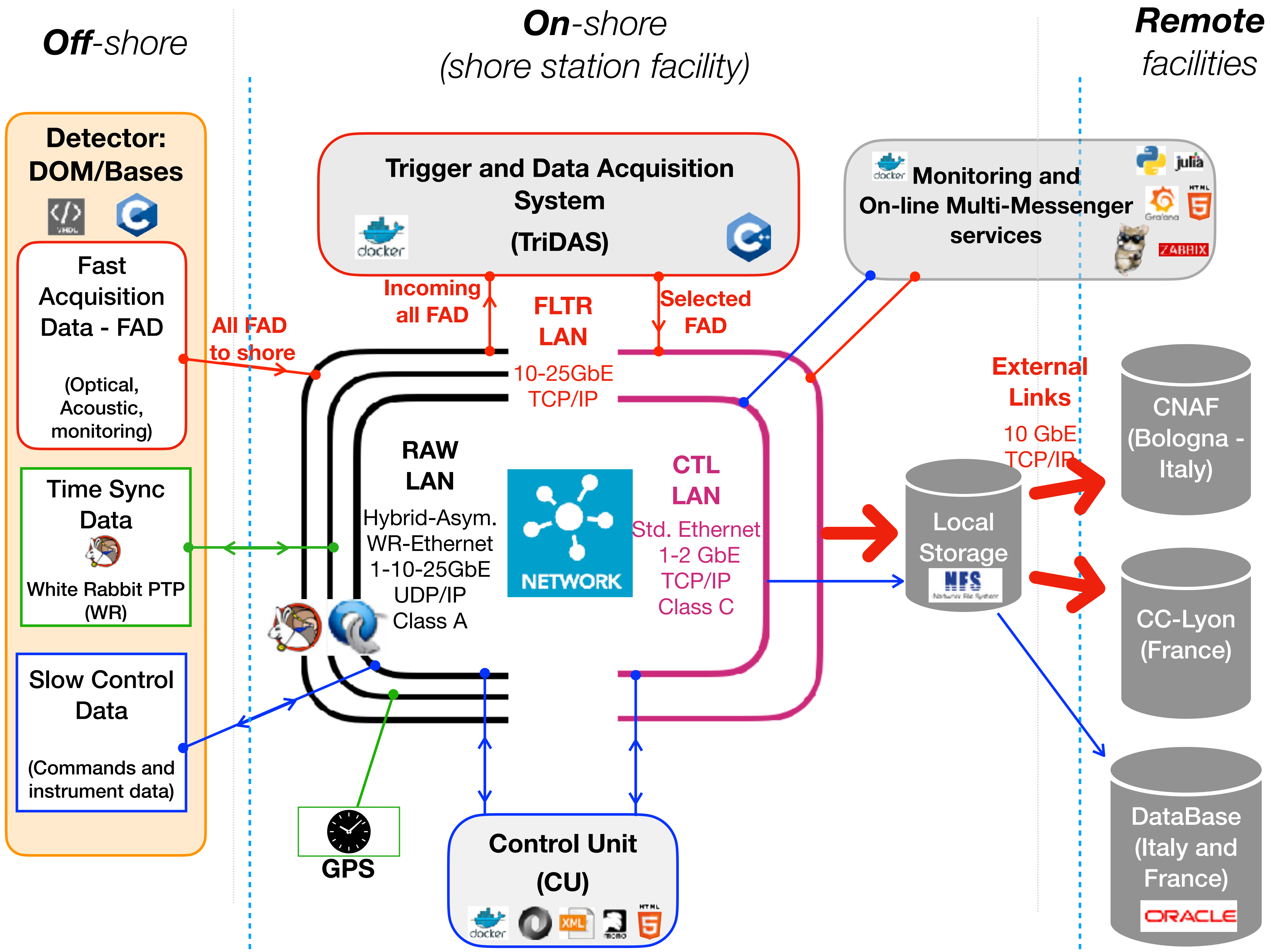


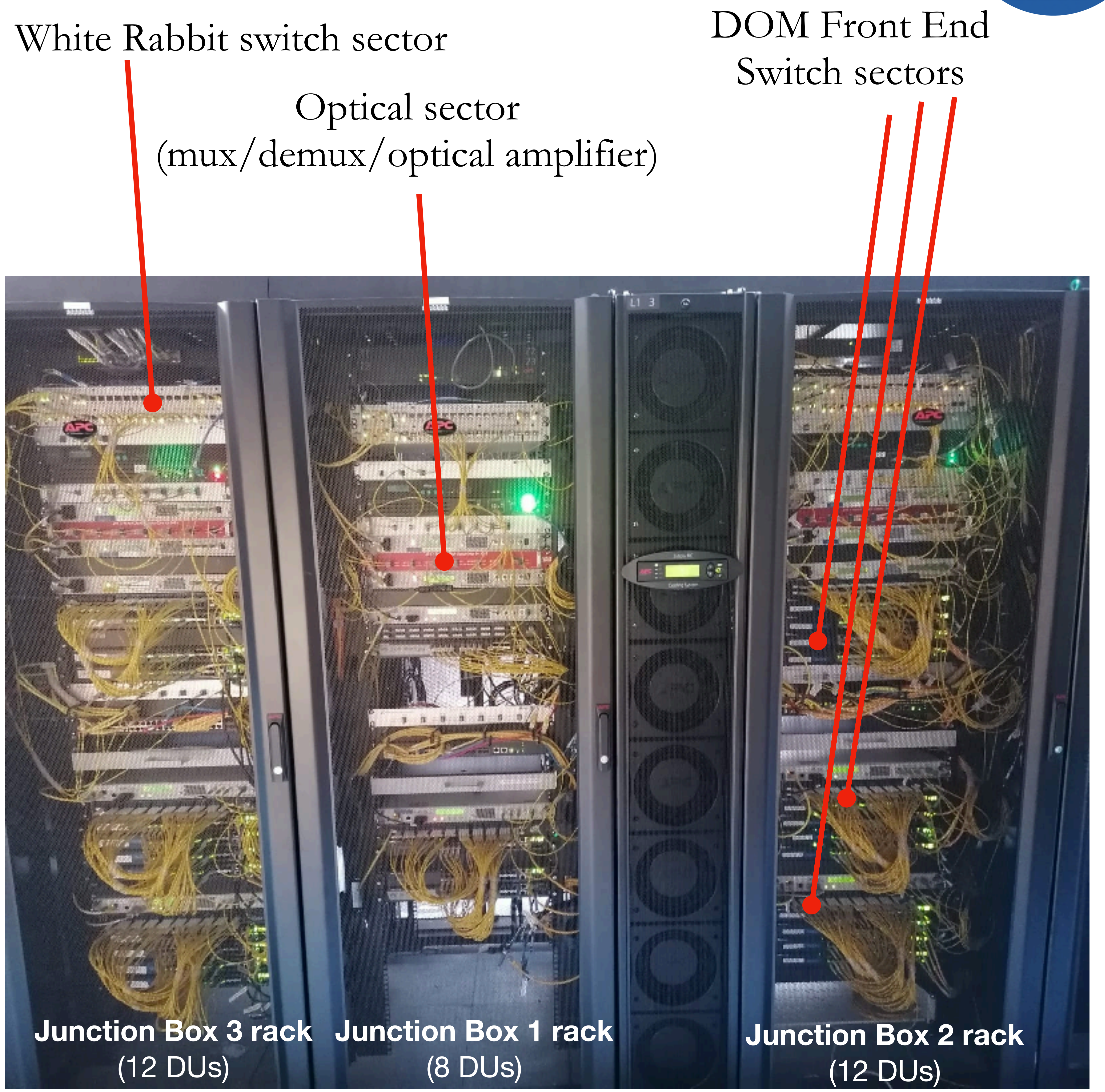


	ARCA	ORCA
Location	Italy	France
N. building blocks	2	1
N. DU per b.b.	115	115
DU distance	90 m	20 m
DOM spacing	36 m	9 m
DU height	~ 800 m	~ 200 m
Instrumented mass (Mton)	2*650	7
Depth	3500 m	2500 m

- Different topologies, same detector concept

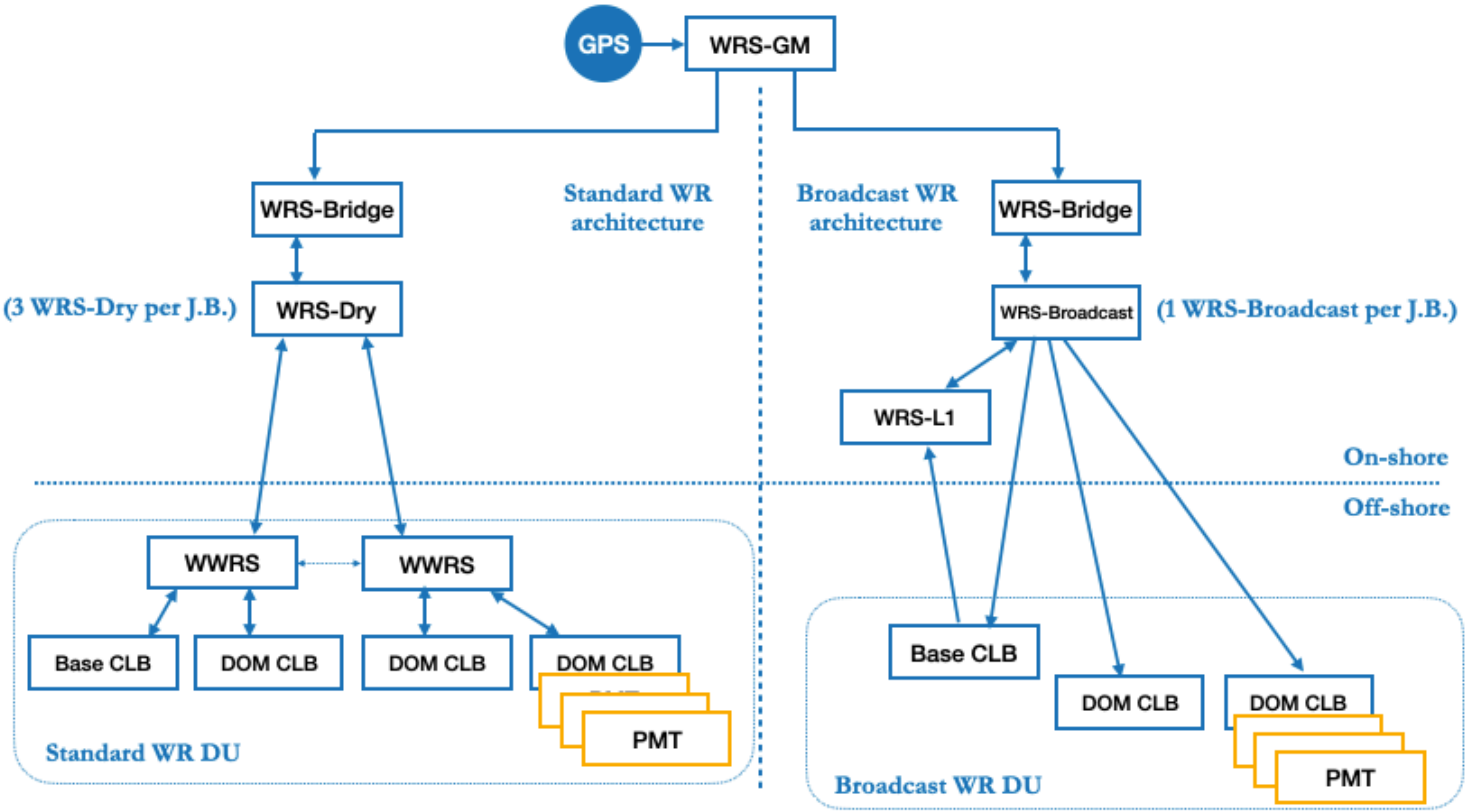
D.A.Q general overview

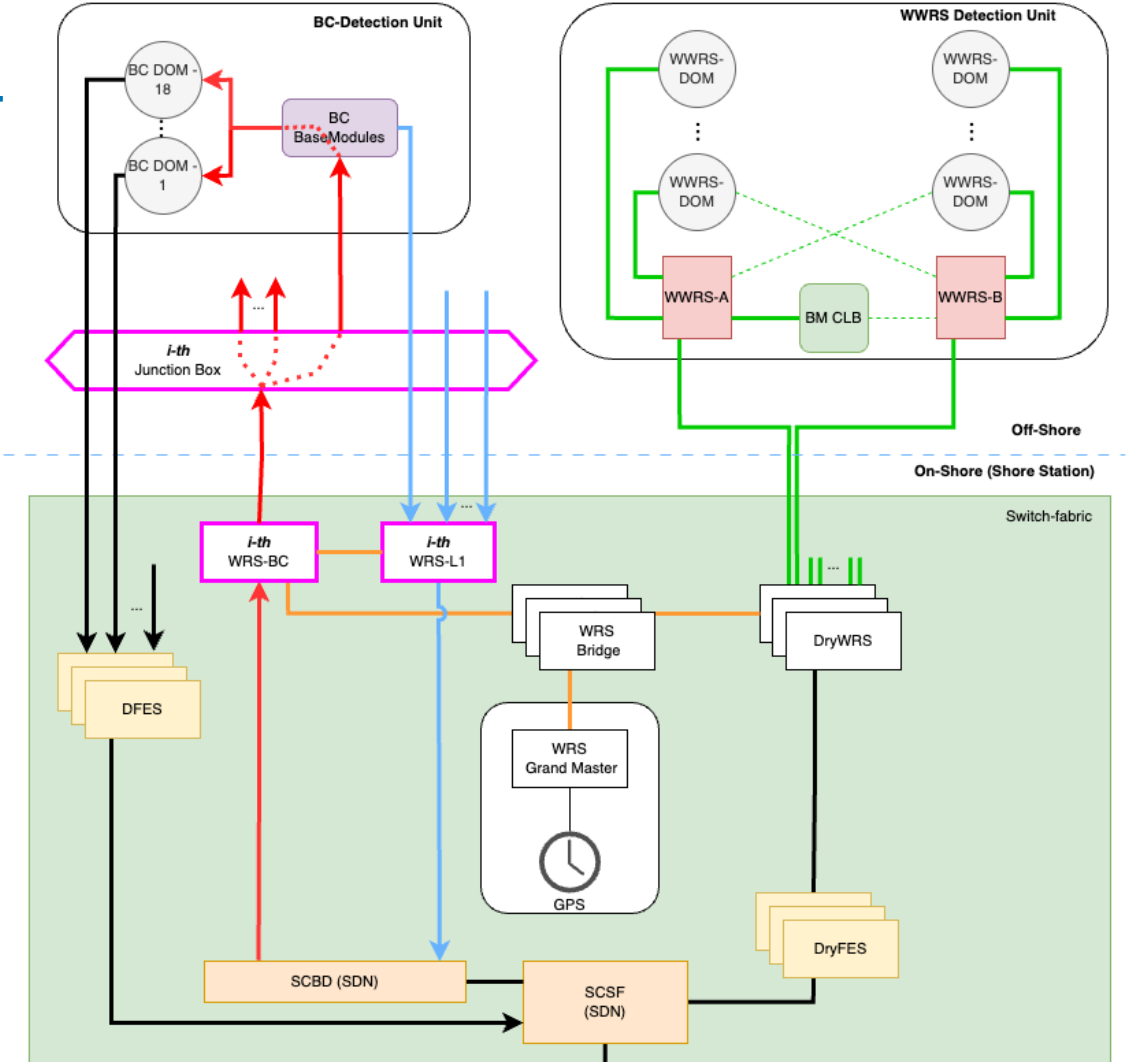




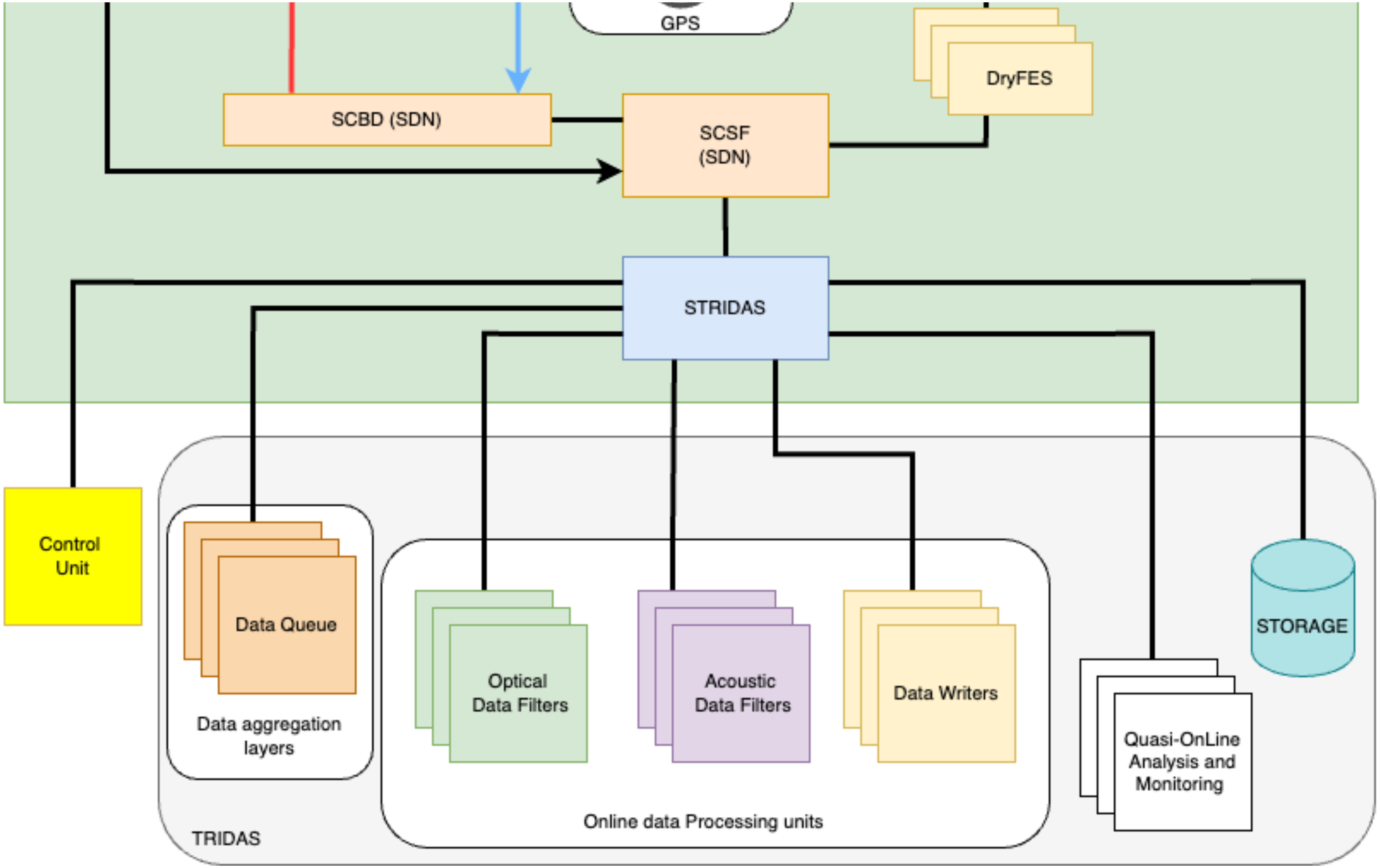
ARCA shore station

KM3NeT White Rabbit network

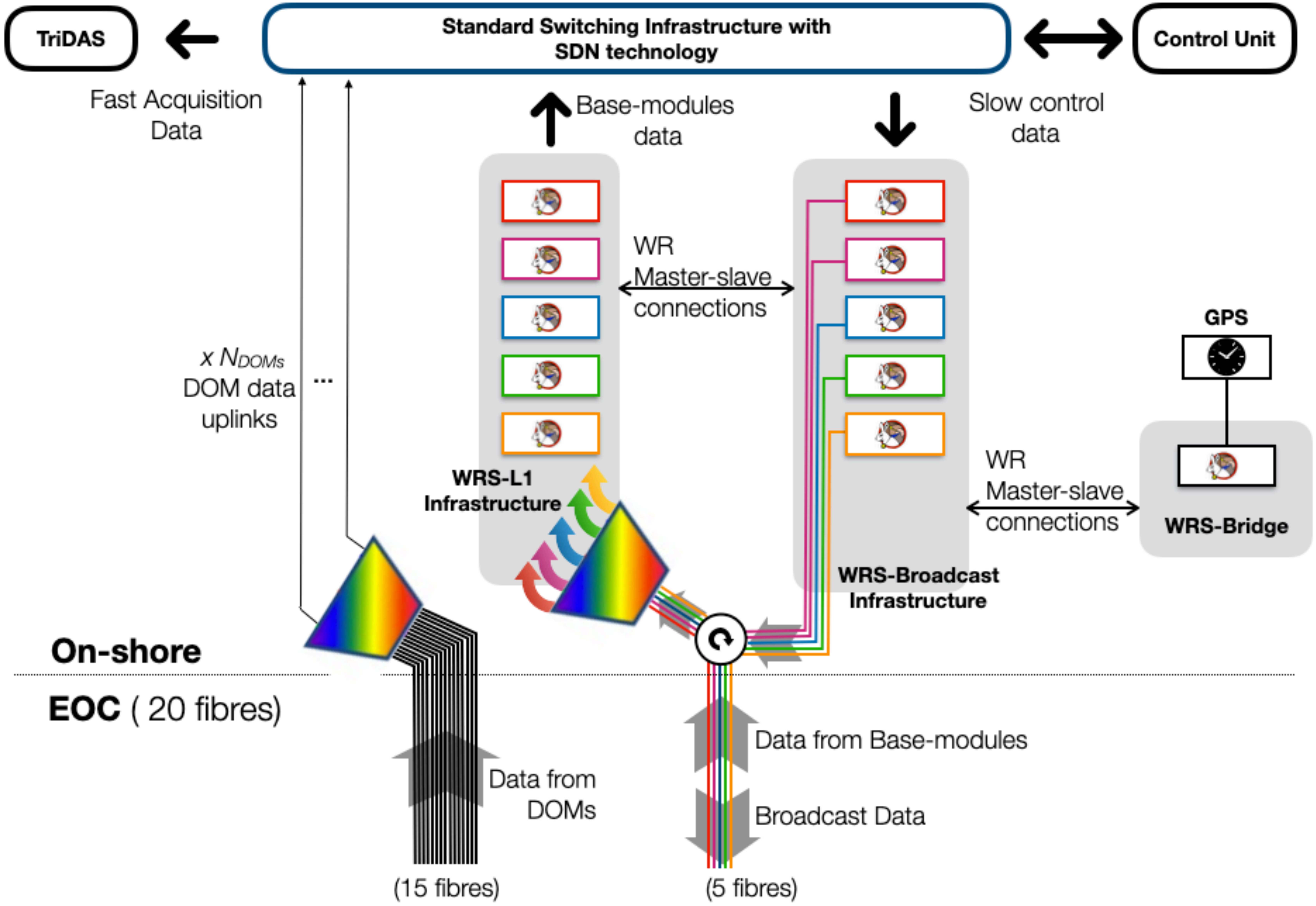




BC-Detection Unit
BC DOM - 18
BC DOM - 1
BC BaseModules
WWRS Detection Unit
WWRS-DOM
WWRS-DOM
WWRS-DOM
WWRS-DOM
WWRS-A
BM CLB
WWRS-B
i-th Junction Box
Off-Shore
On-Shore (Shore Station)
Switch-fabric
i-th WRS-BC
i-th WRS-L1
WRS Bridge
DryWRS
DFES
WRS Grand Master
GPS
DryFES
SCBD (SDN)
SCSF (SDN)

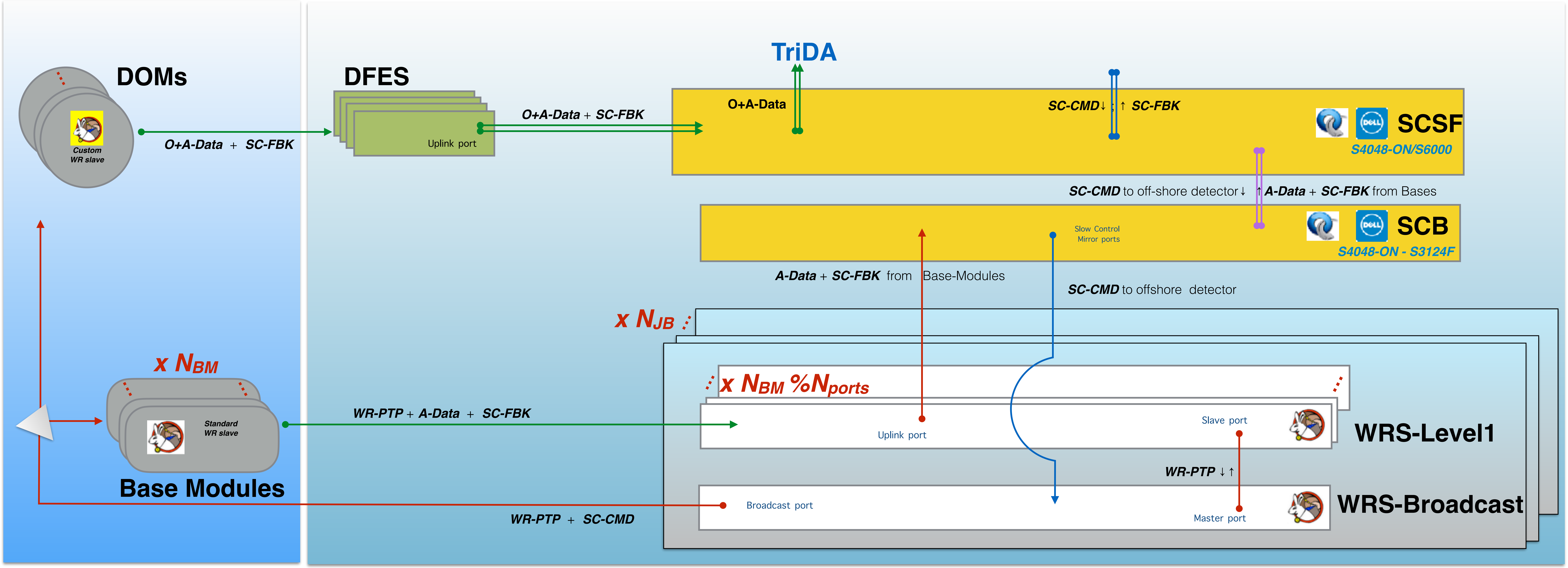

GPS
DryFES
SCBD (SDN)
SCSF
(SDN)
STRIDAS
Control
Unit
Data Queue
Data aggregation
layers
Optical
Data Filters
Acoustic
Data Filters
Data Writers
Online data Processing units
Quasi-OnLine
Analysis and
Monitoring
STORAGE
TRIDAS

Software Defined Networking



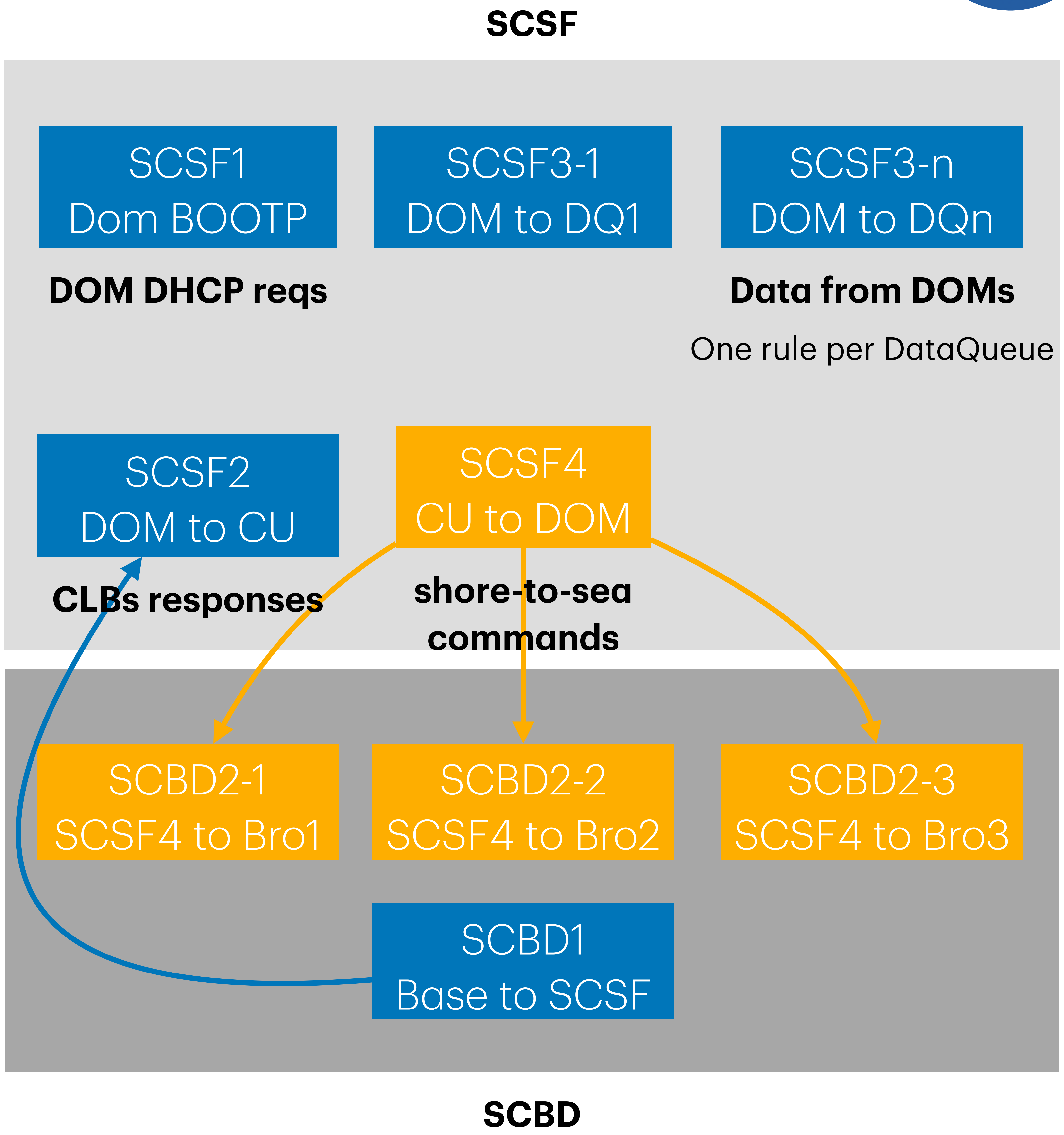
- One shared channel to communicate from shore to sea (the ***broadcast*** channel).
  A dedicated channel for each DOM to reply to shore
- Broadcast intrinsic **asymmetry** violates the basic principles of the ethernet communication.
- Software Defined Networking is needed to prevent fatal conditions due to uncontrolled loops in the switching infrastructure

Software Defined Networks in the Broadcast scenario



Software Defined Networking

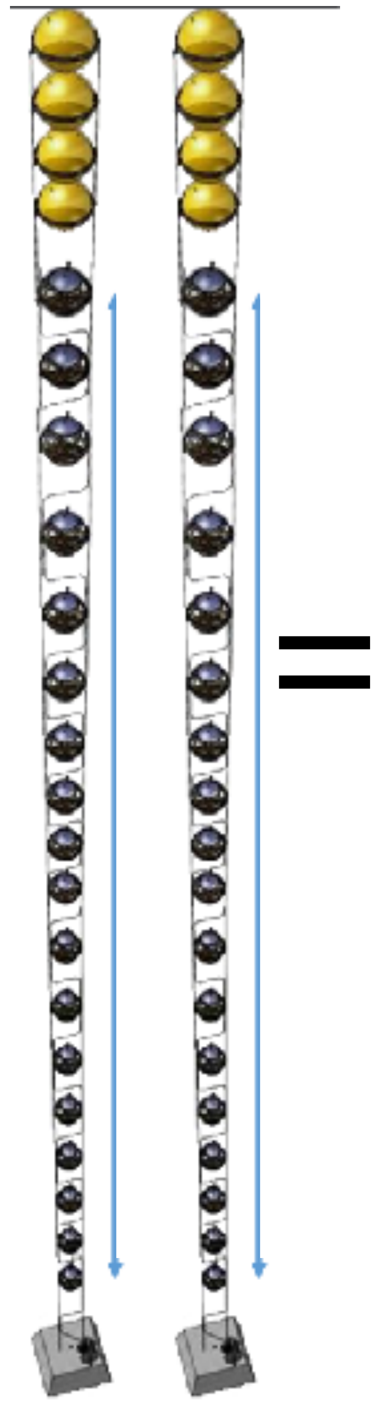
- Asymmetric network topology leads to switch operating in a not-standard configuration: any possible network flow **must be** mapped by an explicit **SDN rule**
  - if a packet does not match any rule is **discarded**
- Two SDN instances representing the macro group of flows
  - **From Detector to Shore**
  - **From Shore to Detector**
- Implementation :
  - JAVA controller (Karaf) implemented on container
  - Python helper script



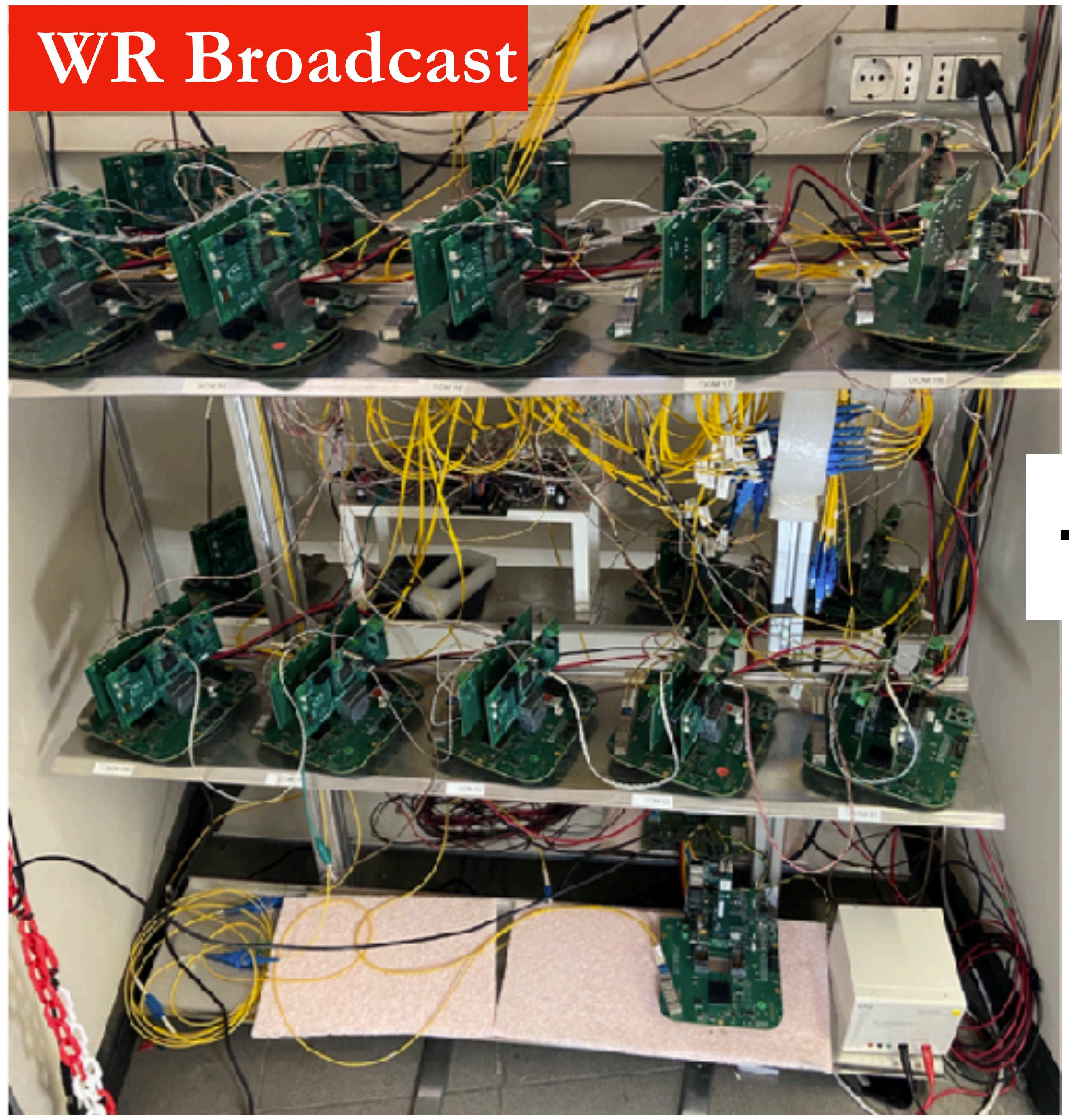
Bologna Common Infrastructure

- **1 DU test bench (Broadcast)** → 18 x CLBs/DOMs **v2** + 1 BM
- **1 DU test bench (Srd. White Rabbit)** →18 x CLBs/DOMs **v4** - 1 x complete BM v4
- ***OctoPAES***: INFN-BO custom electronic boards for **PMT/piezo emulation**
- **Fully-compliant on-shore infrastructure to production sites;**
- **Computing farm for Trigger and DAQ implementation.**

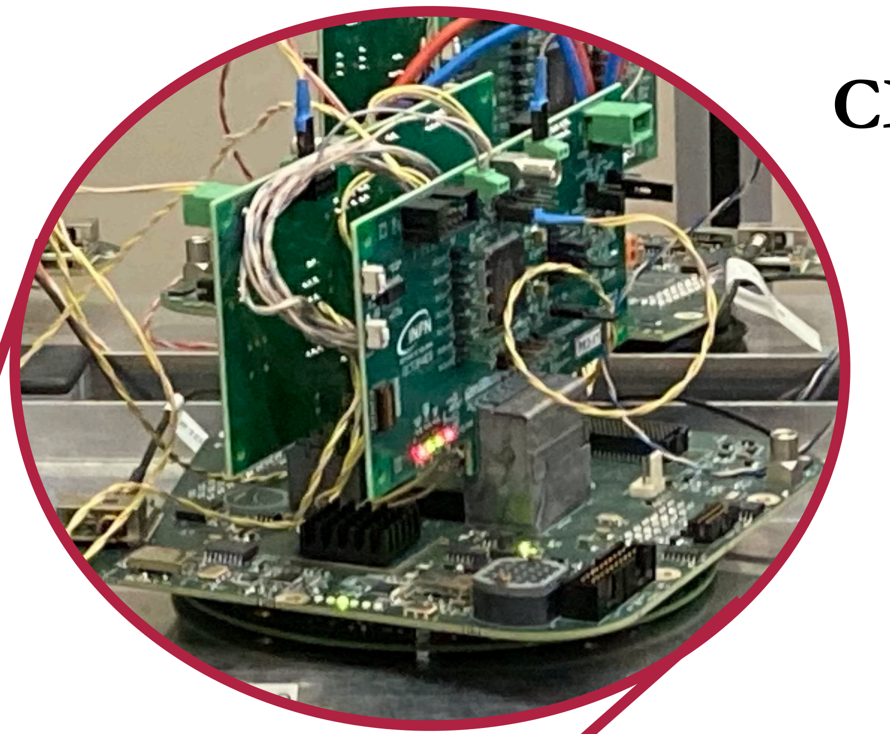
- Test runtime condition of a DU
- Throughputs of various channels (PMTs/Acoustic/Monitoring)
- Effectiveness of NG-Firmware for CLBv4
- Control of DU and BM CLBv4 boards
- Check boards temperatures and power consumptions

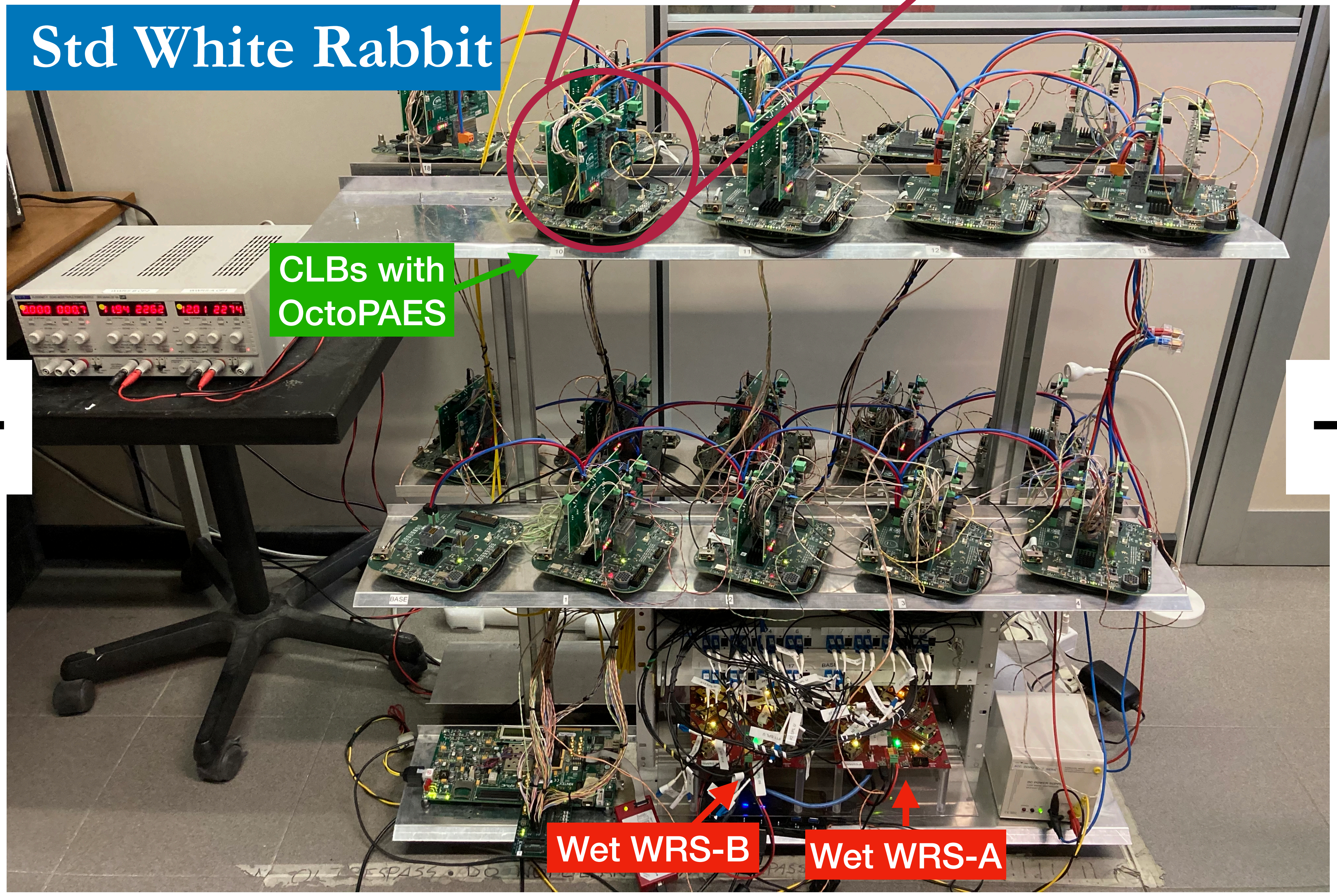


=



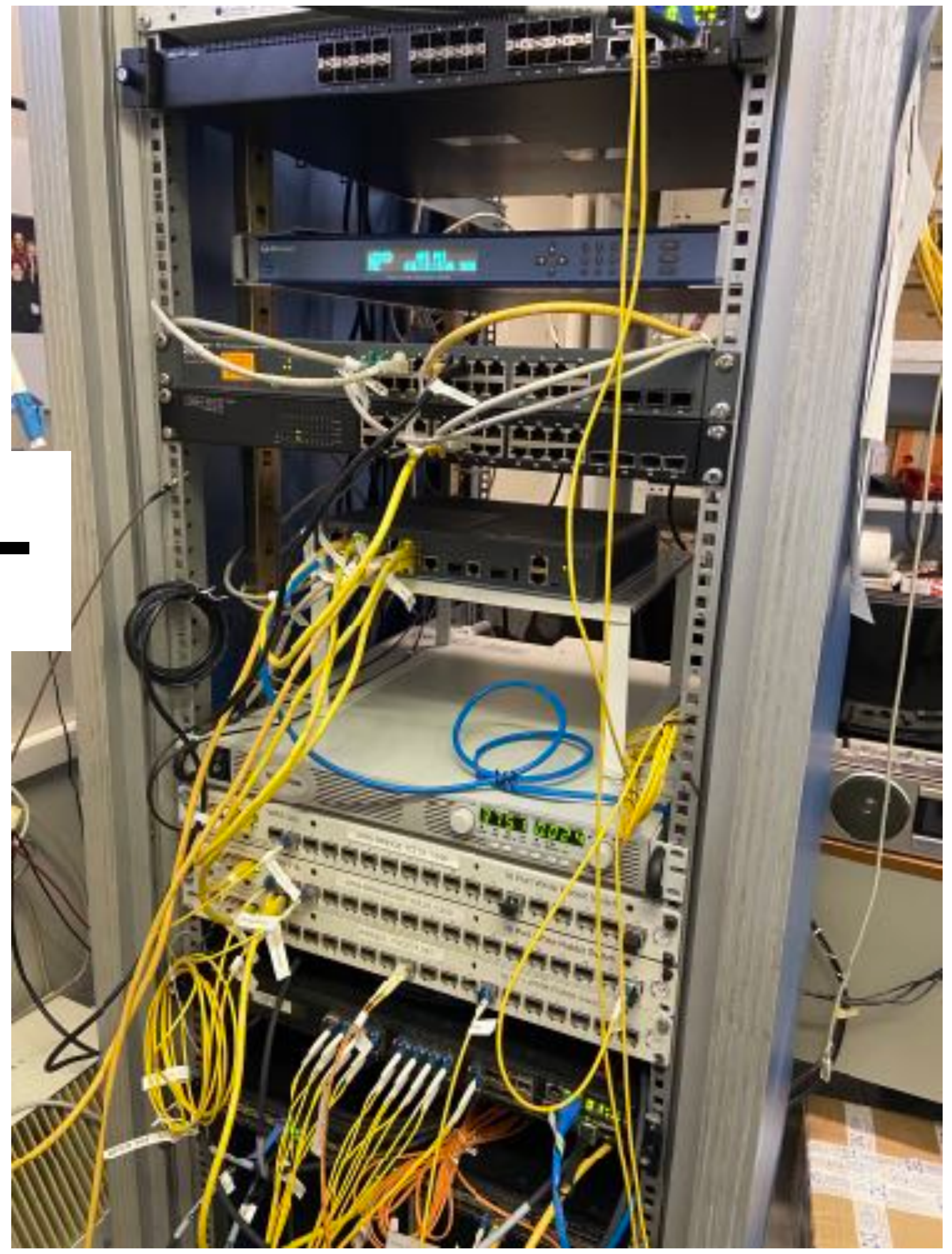
+





CLB v4 with 2 OctoPAES

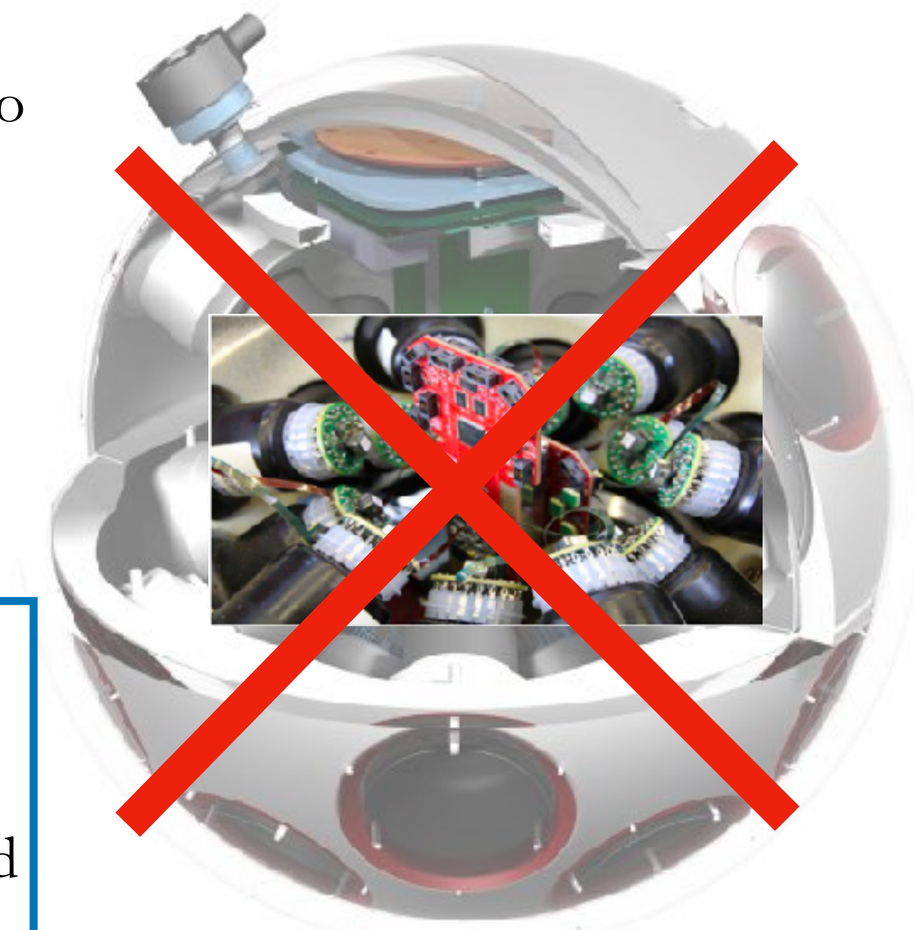
+

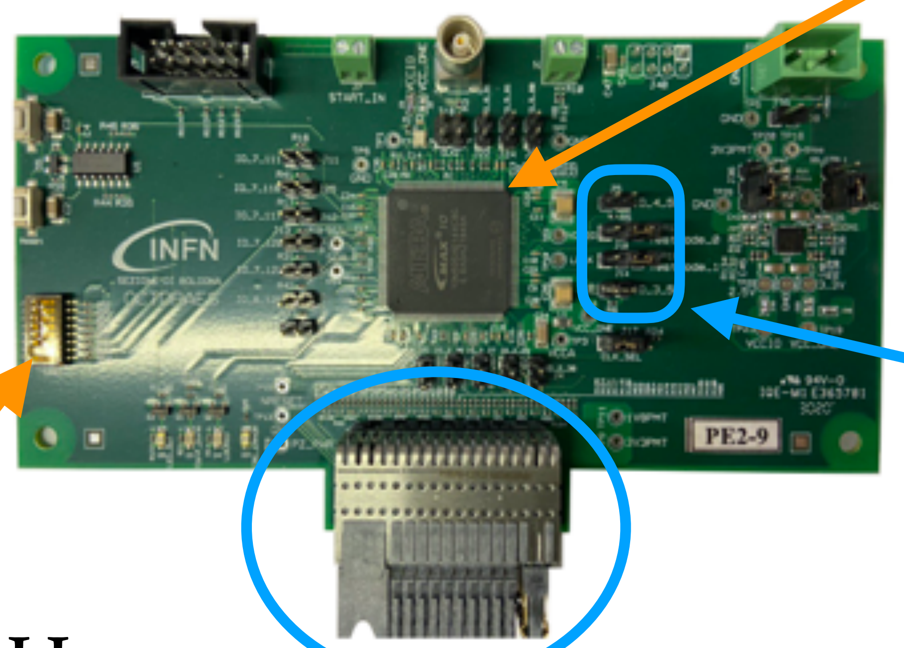


The OctoPAES board

- **PMT** and **Acoustic Emulation** of **Signal**
- **Emulates the presence of PMTs + Octopus boards (Small/Large)+piezo/hydrophone**

No Octopus, PMT bases, PMTs, piezo





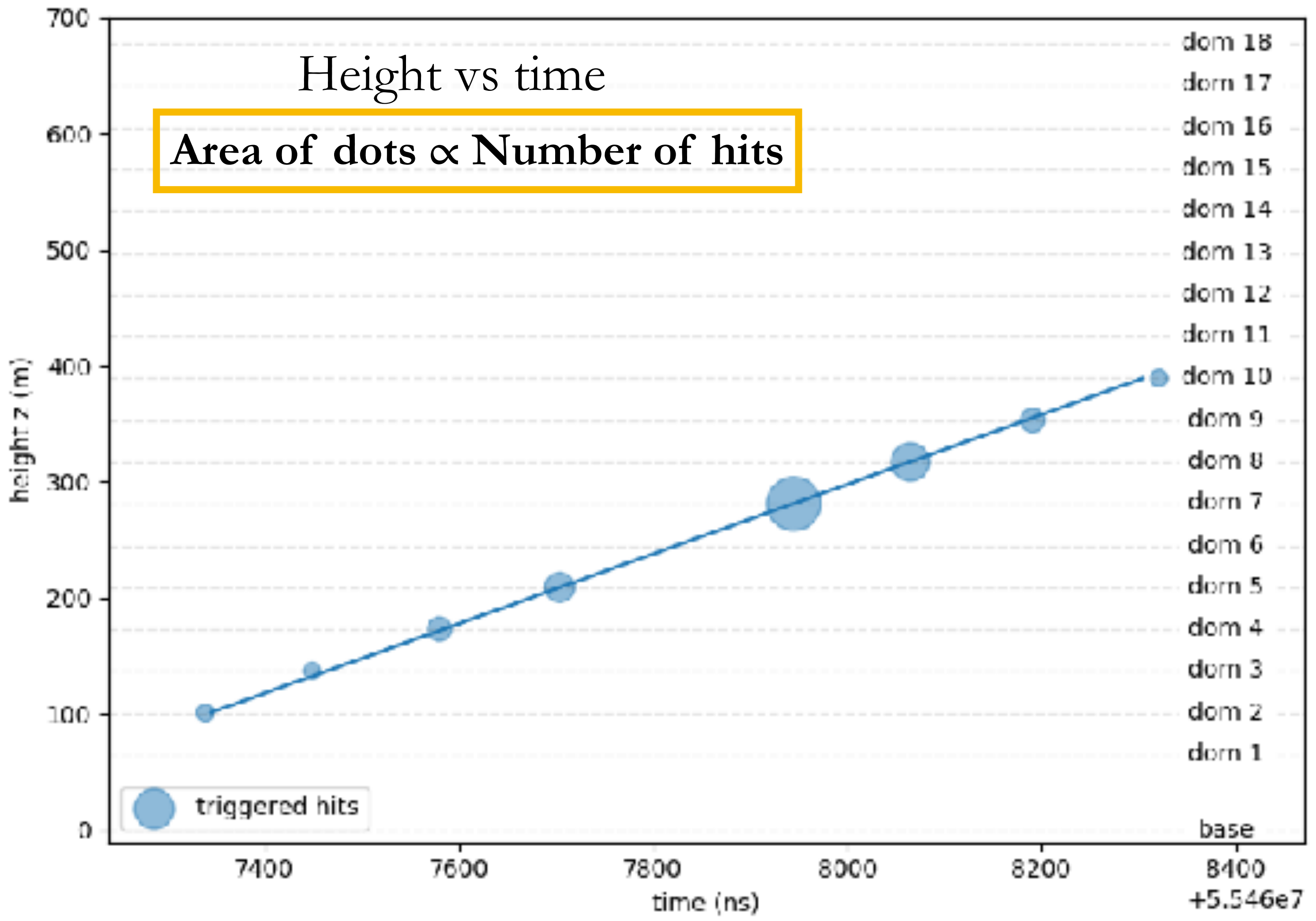
CPLD

jumper for injection frequency selection

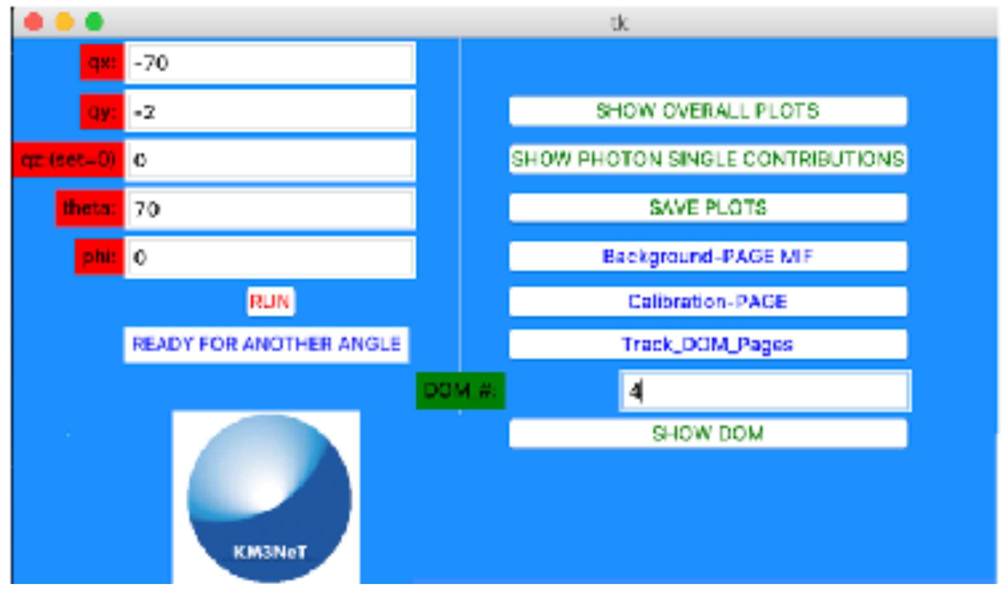
Thanks to the dip switch it can work as **LARGE** or **SMALL**

Specific connector 10M04SCE144I7G

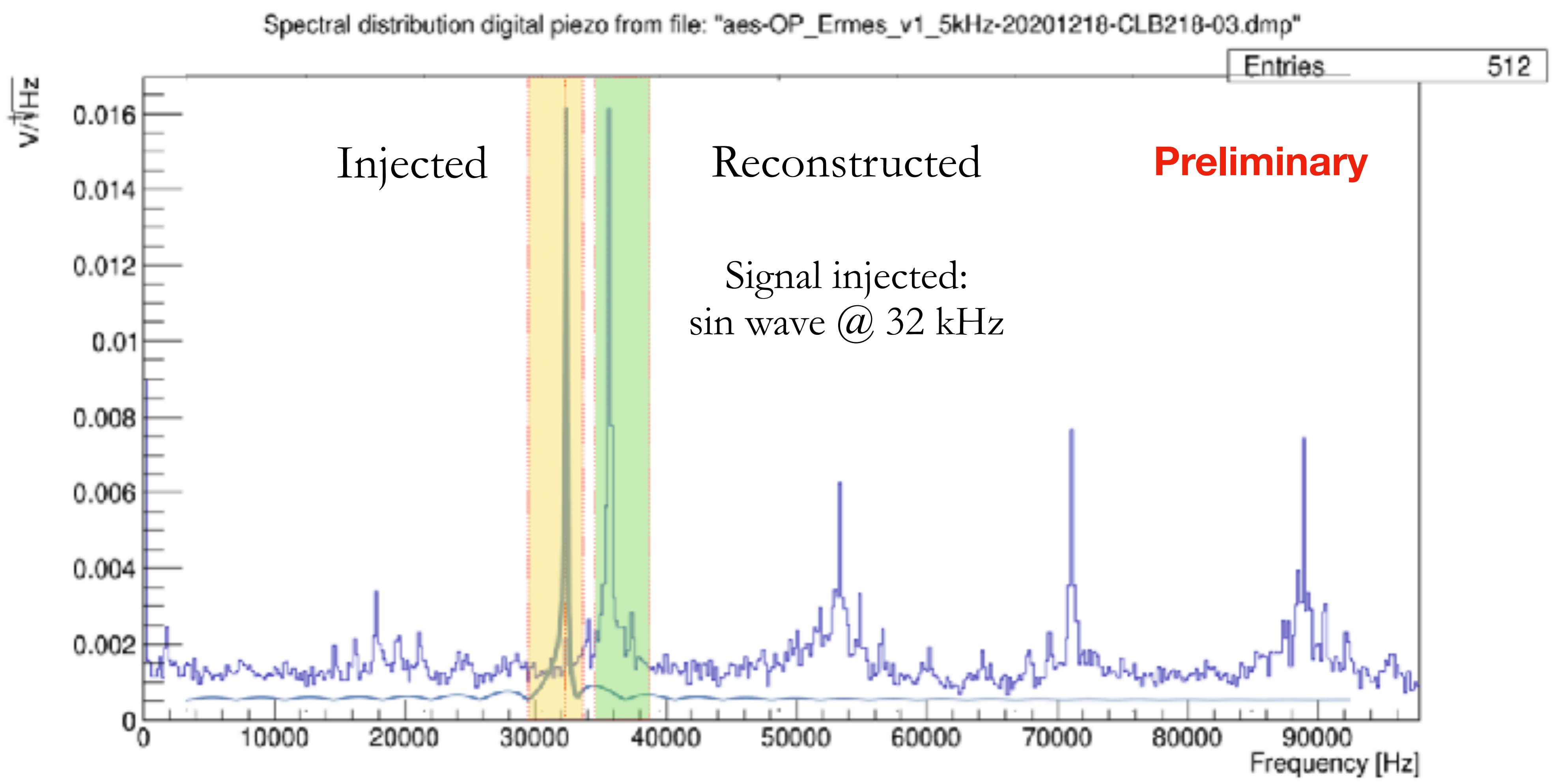
- Hit-time information encoded in a binary file
- Master/Slave with clock distribution in daisy chain + start/stop in parallel
- Signal (@5Hz,10Hz,1kHz,10kHz) + Background (@5kHz)
- Binary file creation automatized with a GUI



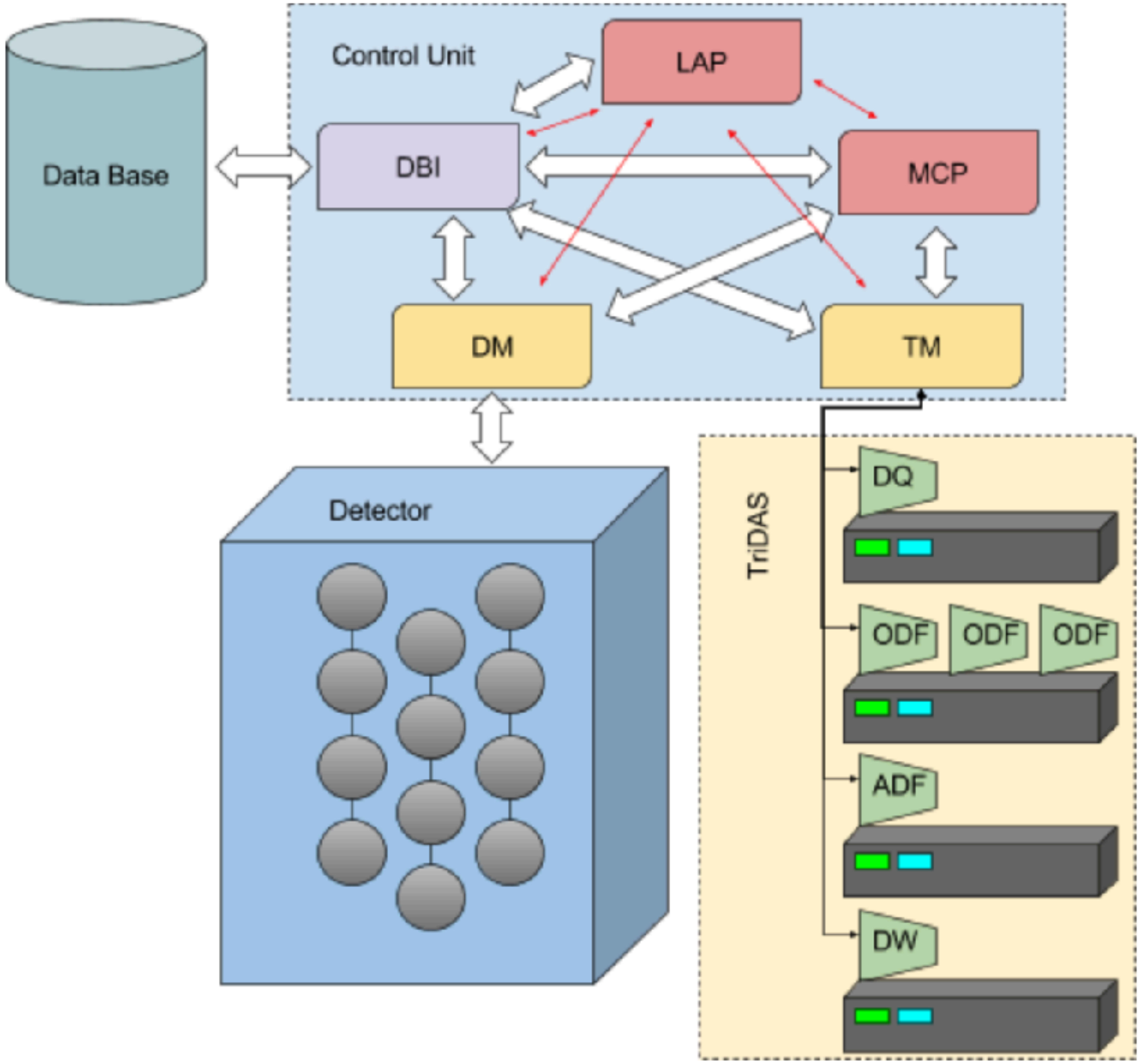
Graphical User Interface



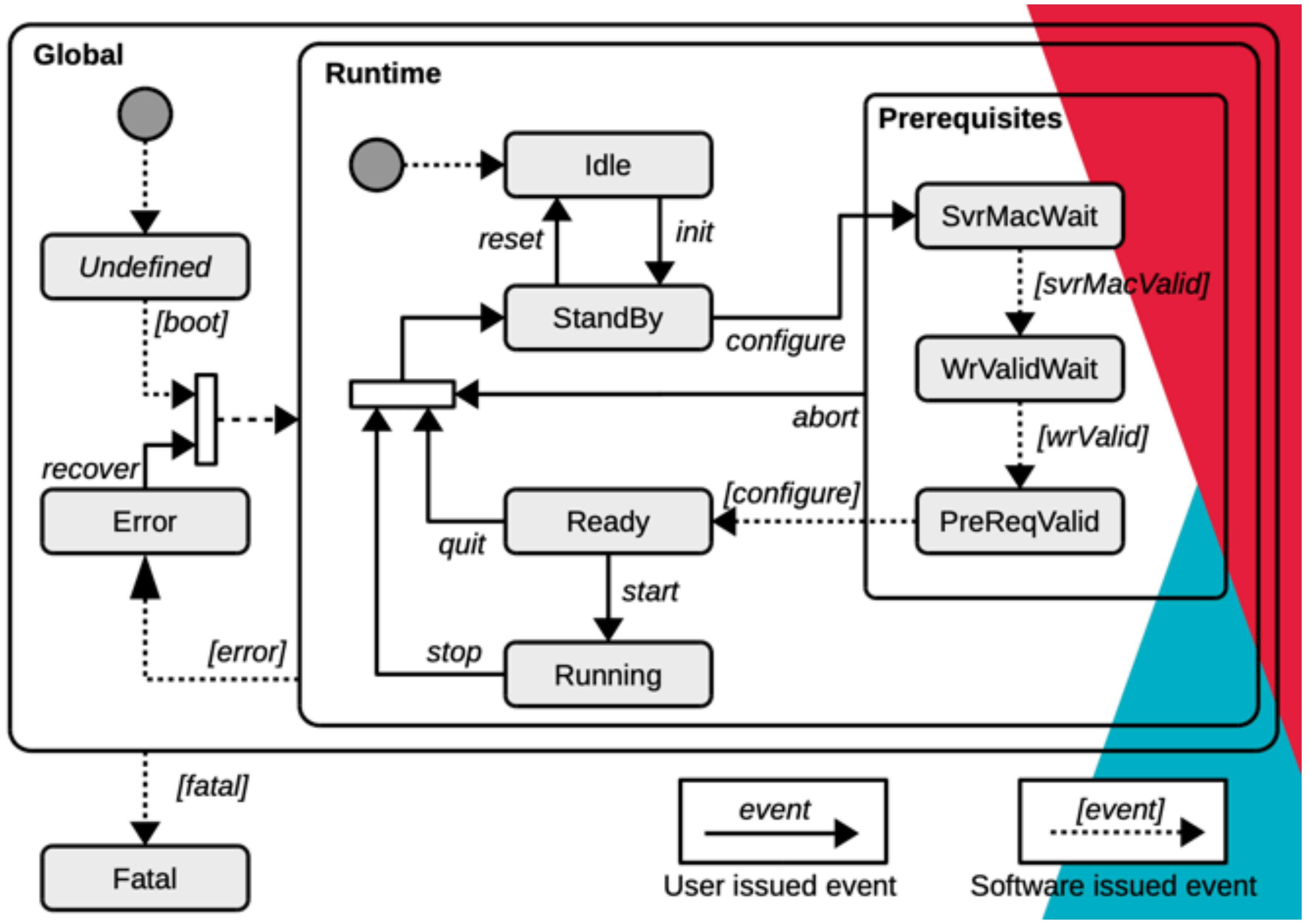
Track physical parameters ($\vec{q}, \theta - \phi$) ⟺ **GUI** ⟺ **LVDS pulses generated by OctoPAES**



The Control Unit

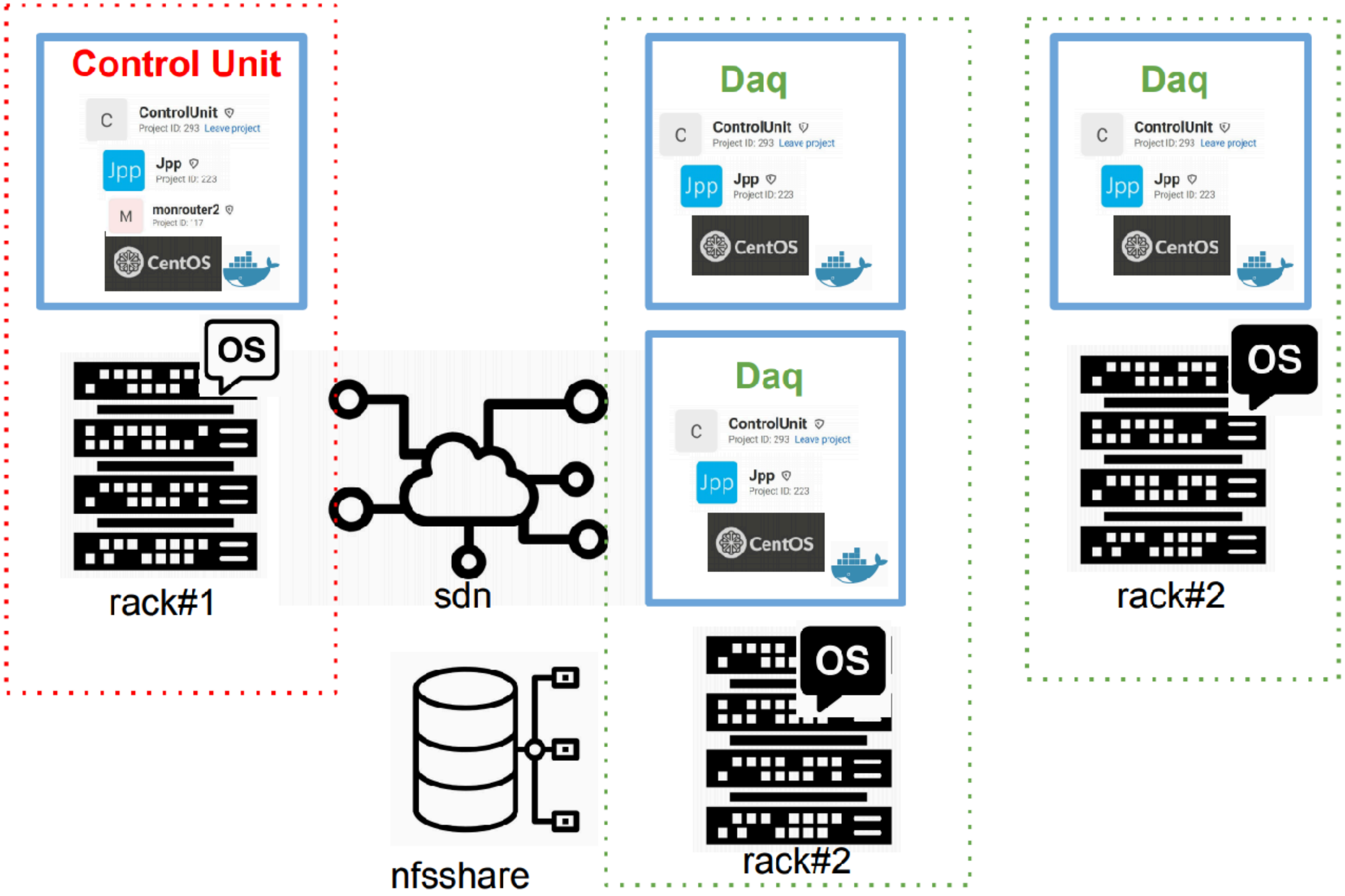


The Control Unit components and their relationships. White and black arrows represent flows of information and/or control signals. Red arrows show the flow of authentication information. The flow of data from the TriDAS to the final storage is not shown.



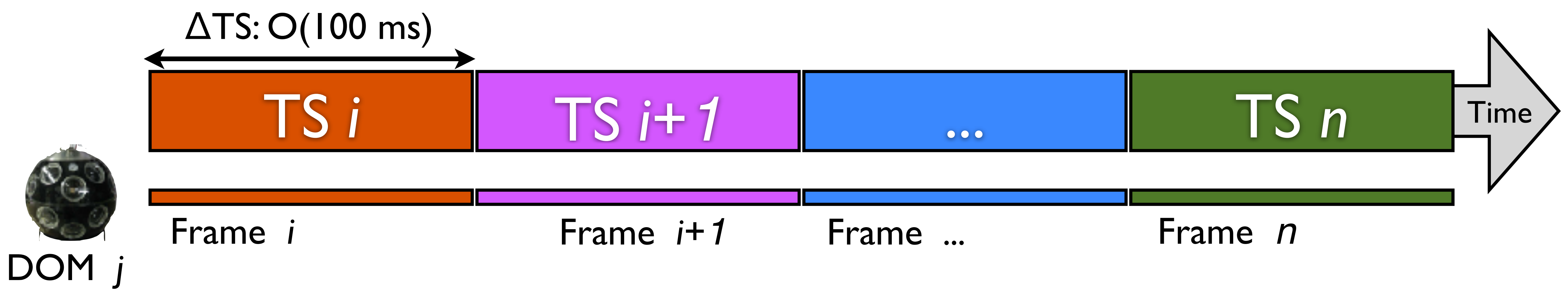
The CU is a collection of (web) services which, via a state machine, drive

- the Detector
- the computing processes
- the interactions with DB for
    - runsetups, calibrations
    - Instruments data logging



Control Unit
ControlUnit
Project ID: 293 Leave project
Jpp
Project ID: 223
monrouter2
CentOS
OS
rack#1
sdn
nfsshare
Daq
ControlUnit
Project ID: 293 Leave project
Jpp
Project ID: 223
CentOS
Daq
ControlUnit
Project ID: 293 Leave project
Jpp
Project ID: 223
CentOS
OS
rack#2
Daq
ControlUnit
Project ID: 293 Leave project
Jpp
Project ID: 223
CentOS
OS
rack#2

- **Timeslice** (TS): it is the abstract subdivision of the continuity in the time-line of the experiment.
- **Frame**: it is the group of information of a certain flavour (TDC, AES, MON) occurred in a DOM during a TS.

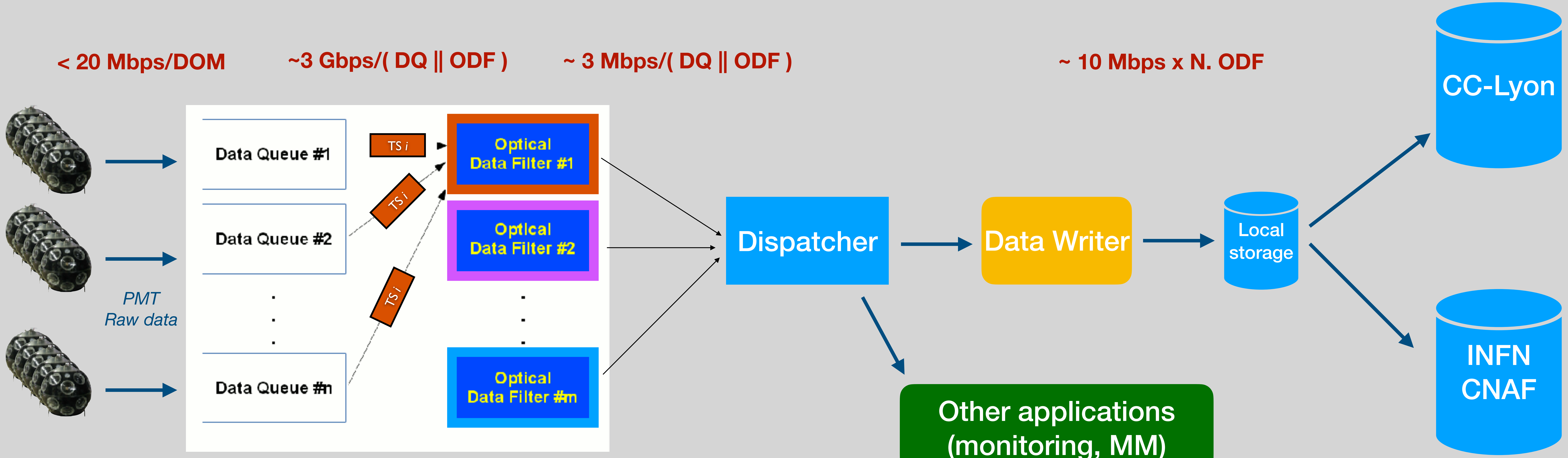


Distributing the computational load

- Each trigger algo applied to one full set of frames of one TS.
- Multiple TSs handled in parallel

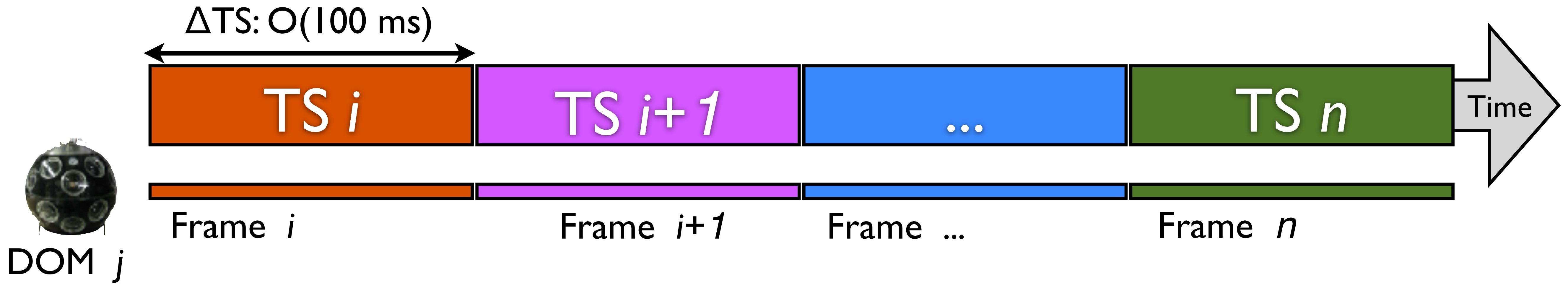
Optical World

A DQ collects data from a sector of DOMs and DU-BMs.
All DQs transfer all their data from a precise Time Slice to the very same oDF.



If needed more *Filtering Power*, just add it!

- **Timeslice** (TS): it is the abstract subdivision of the continuity in the time-line of the experiment.
- **Frame**: it is the group of information of a certain flavour (TDC, AES, MON) occurred in a DOM during a TS.

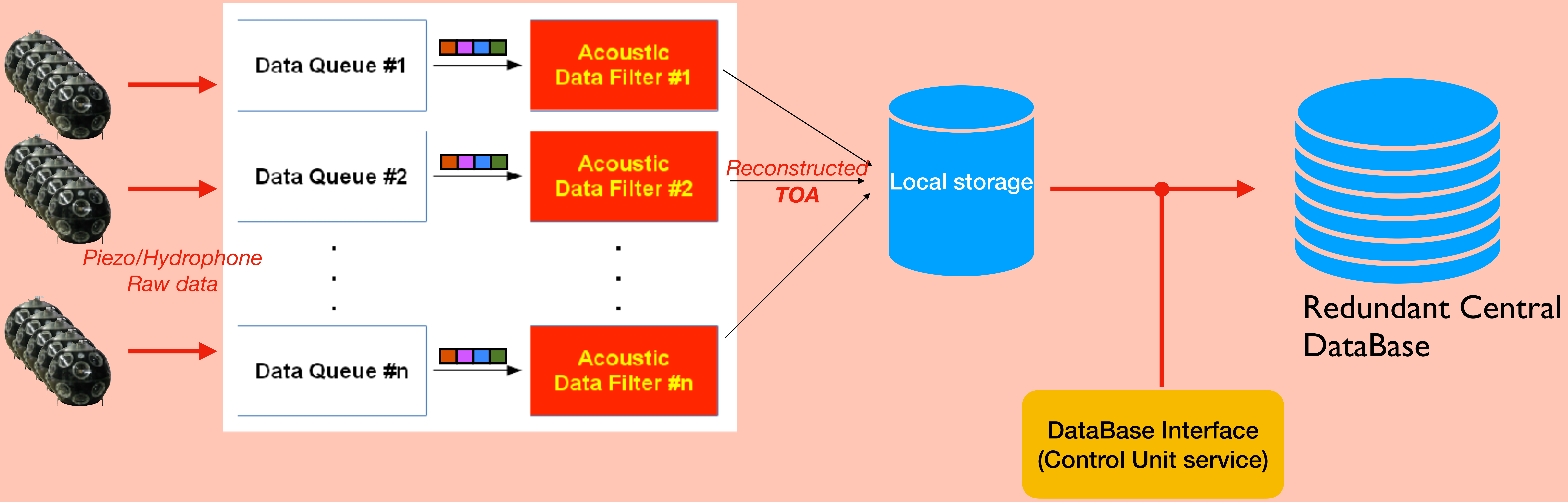


Distributing the computational load

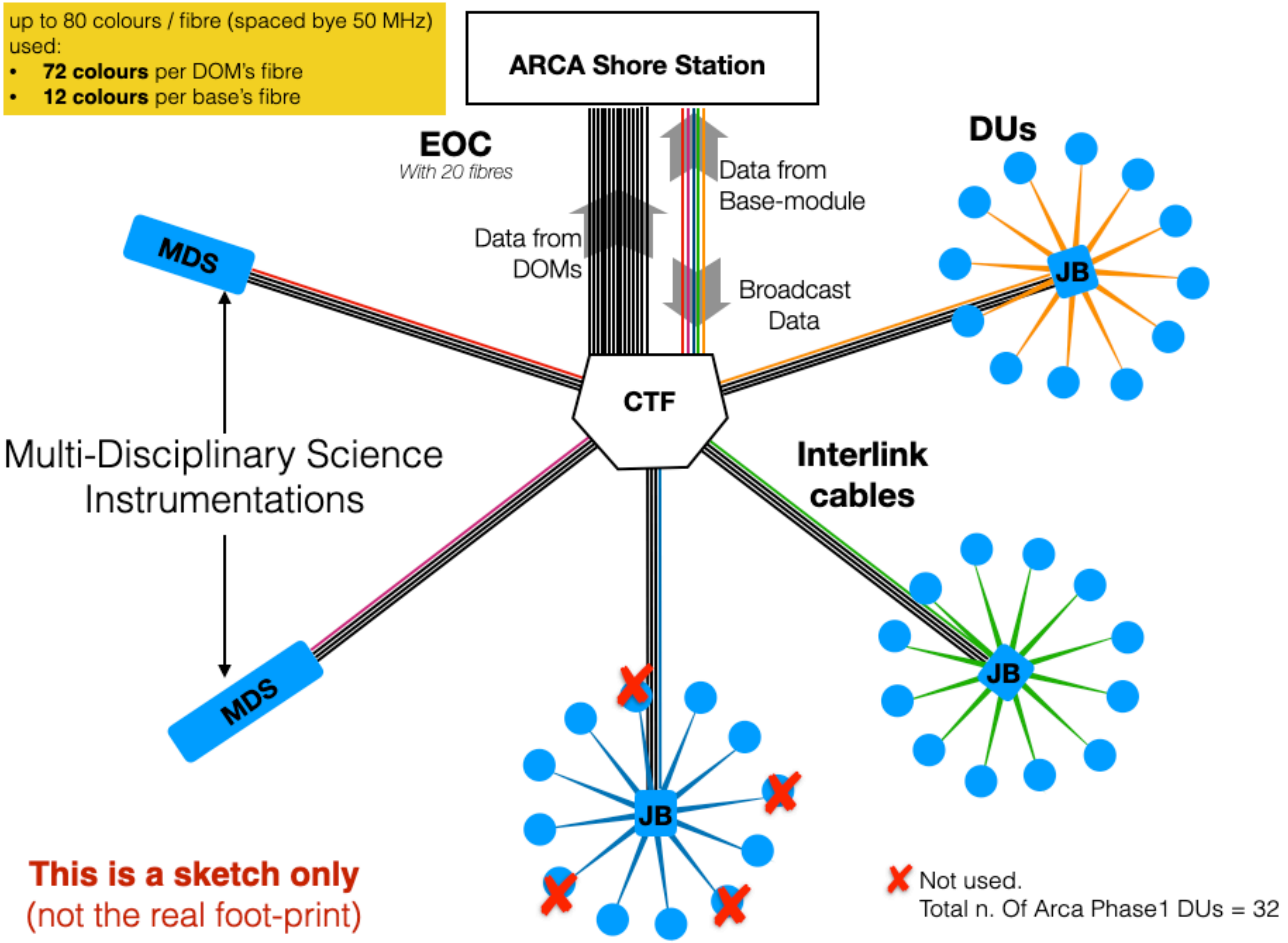
- Each trigger algo applied to one full set of frames of one TS.
- Multiple TSs handled in parallel

Acoustic World

Acoustic data must be sent in a continuous stream, addressing all data from one DQ to a single Acoustic DF.
Independent reconstruction of the *Time Of Arrival* (**TOA**) of acoustic signals from various beacons



32 ARCA DUs only
(ARCA-Phase 1)



This is a sketch only
(not the real foot-print)

Optical data for Physics

Case	n_{DU}	n_{DOMs}	$n_{pmt/DOM}$	ν_{single}/PMT (kHz)	hit size (bit)	$\nu_{trigger}$ (Hz)	Event window (μs)
KM3NeT–Ph1, It	24	18	31	15	50	40	6
KM3NeT–Ph1, *Fr*	7	18	31	15	50	13	6
KM3NeT–1 Block (Ph2, *Fr*)	115	18	31	15	50	220	6
KM3NeT–2 Blocks (Ph2, *It*)	230	18	31	15	50	440	6

Case	DOM thp (Mb/s)	DU thp (Gb/s)	Det thp (Gb/s)	Sel thp (MB/s)	Sel thp (TB/day)	Stored (TB/y)	event size(kB)
KM3NeT–Ph1, It	23.0	0.4	10.0	1.6	0.13	49.0	7.5
KM3NeT–Ph1, *Fr*	23.0	0.4	2.9	0.4	0.03	12.0	2.2
KM3NeT–1 Block (Ph2, *Fr*)	23.0	0.4	48.0	14.0	1.20	440.0	36.0
KM3NeT–2 Blocks (Ph2, *It*)	23.0	0.4	96.0	44.0	3.80	1400.0	72.0

Acoustic data for positioning

Case	Raw Thp/Sensor (Mb/s)	Raw Thp/DU (Mb/s)	Raw Thp/Detector (Gb/s)	TOA (Mb/s)	Positions (Mb/s)	Storage (TB/y)
Phase 1–*It*	13.0	240.0	5.7	0.20	0.08	1.10
Phase 1–*Fr*	13.0	240.0	1.7	0.06	0.02	0.32
1 Block, Ph2 *Fr*	13.0	240.0	27.0	0.94	0.38	5.20
2 Blocks, Ph2 It	13.0	240.0	55.0	1.90	0.75	10.00

```
SamplingRateHz = 195.3 × 10^3;
ResolutionBit = 24;
NChannels = 2;
```

Fast Data Acquisition channels