

DE LA RECHERCHE À L'INDUSTRIE



# POSSIBLE CONTRIBUTIONS OF DEDIP IN ELECTRONICS TO HYPER KAMIOKANDE

**D. Calvet,  
CEA Paris-Saclay**

Saclay, 8 December 2020

[www.cea.fr](http://www.cea.fr)



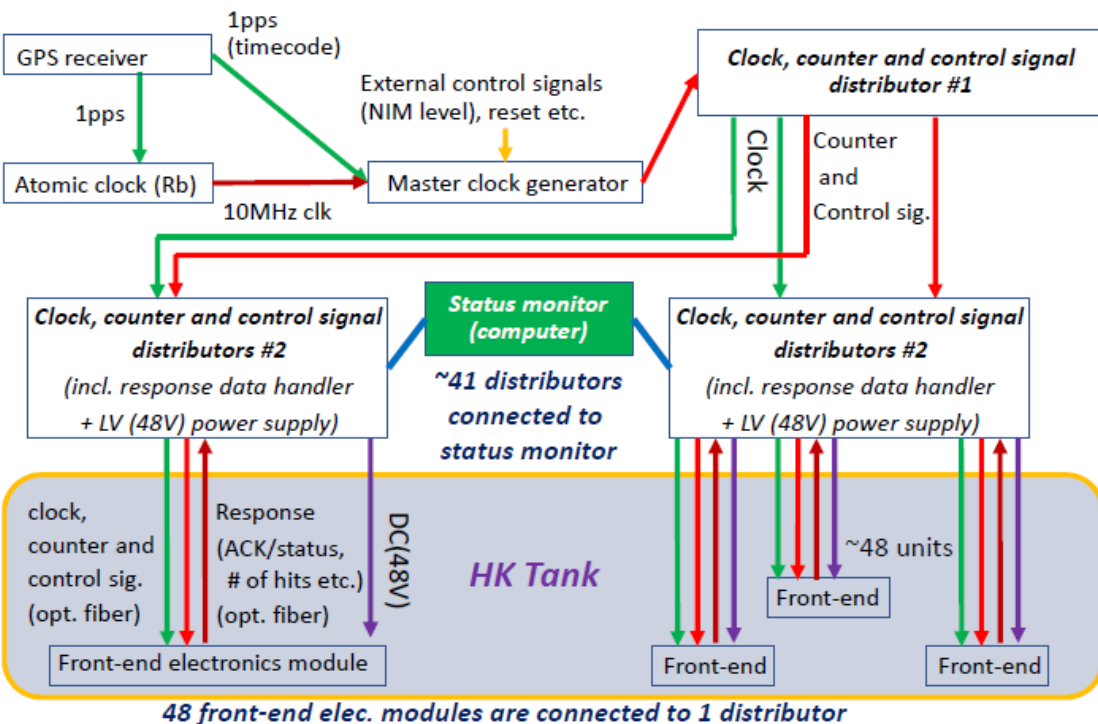
## A new QtC ASIC

- Development of HKROC led by Omega group
  - Currently no involvement of DEDIP. Some discussions on-going for a possible contribution (see e.g. CMS HGROC)

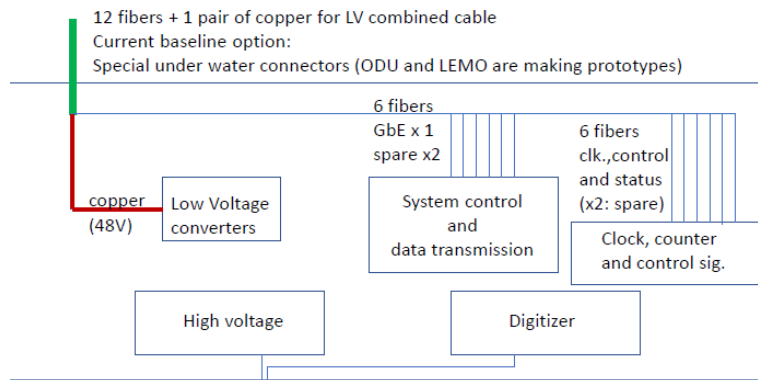
## Clock Distribution System

- On-going effort led by Lpnhe with contributions from INFN and Tokyo University groups
  - Starting discussions on how DEDIP could join this sub-project

## Schematic diagram of the synchronization system



## Schematic diagram of the synchronization system



*Excerpts of presentation made by Hayato-san (mid-2019)*

## Concept

- Underwater front-end (FE) serves 24 PMs. ~2000 FEs total max.
- Distinct networks for DAQ and clock distribution (custom protocol or not)
- Clock distribution by 3-stage tree: 1 master x 41 slave distributors x 48 FEs (for 47232 PMTs)

### ...on Hyper K readout architecture

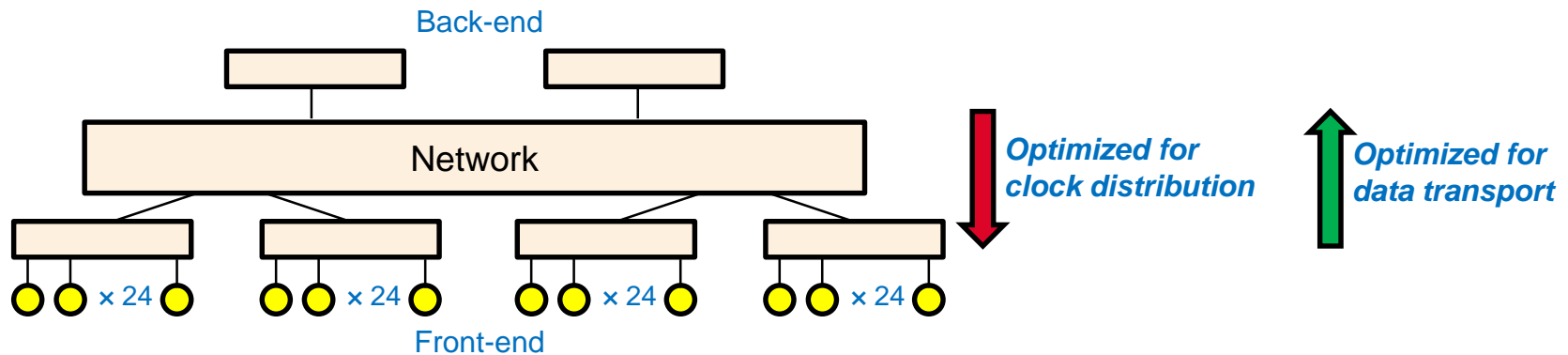
- Separation of clock distribution and DAQ in two distinct networks arguable. Provides independence for development but brings higher cost, complexity, power consumption. Reduces global system availability? (both clock network AND daq network have to be OK)
- Multi-stage tree topology good for clock distribution, seems more robust than cascading a large number of switches in series

### ...on White Rabbit

- Proven technology and will certainly work, but current products on the market not ideally matched: low switch density (18 ports), WR3 switch designed in ~2014 based on Xilinx Virtex 6 almost obsolete; WR4 under development based on a costly Zynq Ultrascale+. Aiming for Ethernet 10G+ but same port count per switch
- Oversized in terms of bandwidth – downstream and also upstream
- Unclear if some functionality are not too much, e.g. dynamic clock phase adjustment
- Strong community of users and CERN support but commercial availability depends on two startup companies, Seven Solutions and CreoTech

## A novel technique for clock distribution

- See previous presentation last month on “Clock-centric” links based on Clock Duty Cycle Modulation – CDCM (full paper at: <https://arxiv.org/abs/2010.14164>)

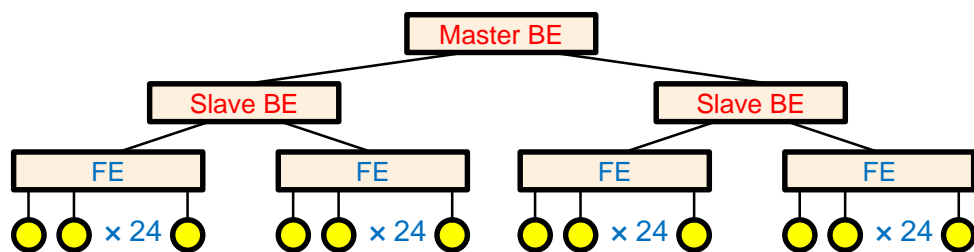
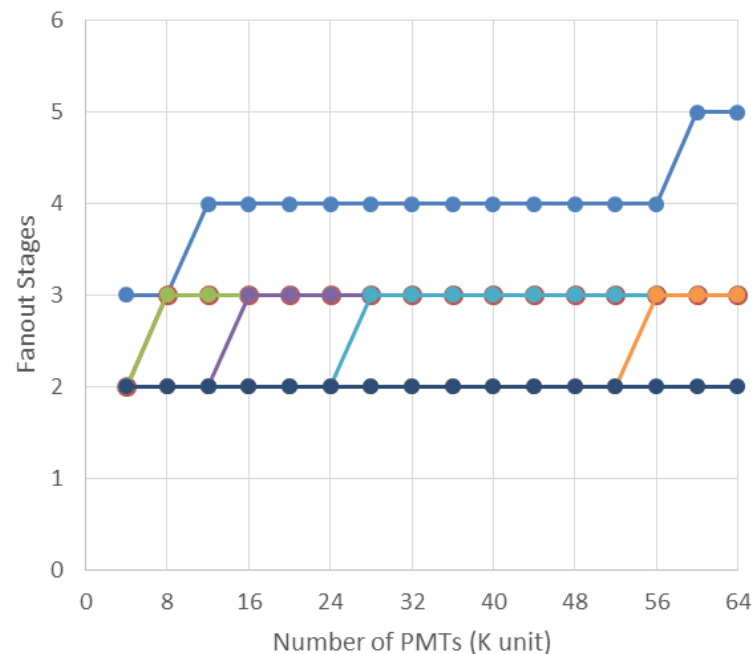
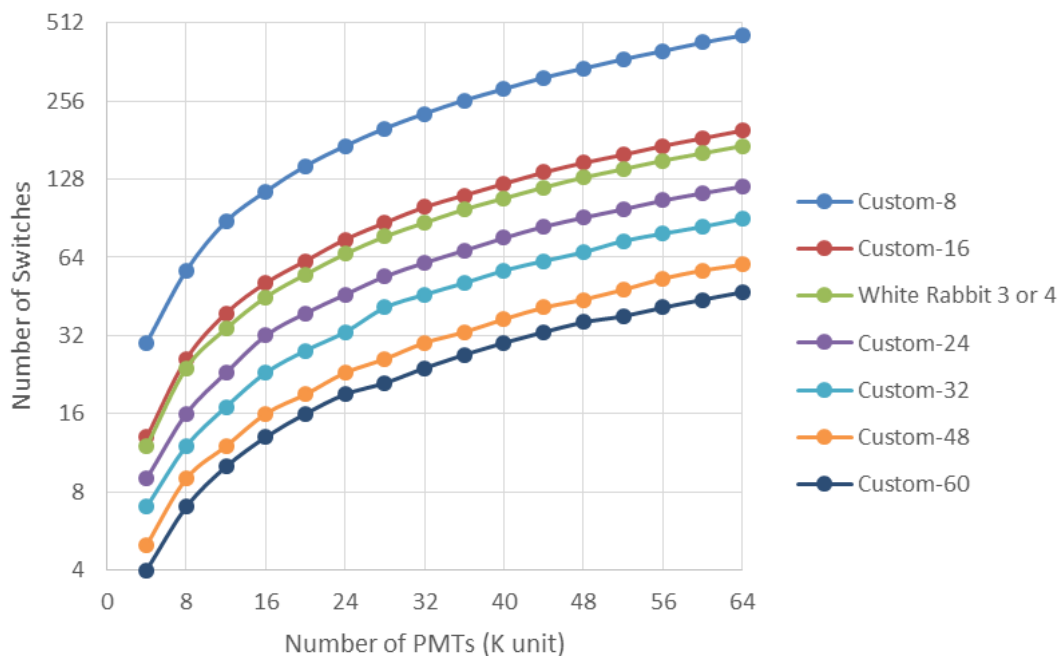


## A novel implementation

- Bandwidth adapted to requirements, i.e. asymmetric, not imposed by a standard like 1G or 10G
- Switch core based on an inexpensive commercial FPGA module
- No superfluous functionality or unnecessary features. Application specific, not a universal solution
- Baseline design dedicated to clock distribution only; upgradable to serve DAQ functions

→ ***Danger to avoid: try to build ourselves a bigger, better, faster, cheaper White Rabbit switch***

# WHAT GRANULARITY FOR THE SWITCHING ELEMENTS?



## Hypothesis and parameters

- Multi stage tree topology
- Switch: N ports; 1 uplink; (N-1) down links
- For models Custom-24/32/48/60 the uplink is not included in the port count

## Graph interpretation

- 48K PMTs scenario: 130 WR3-4 switches in 3-stage tree. With 48 usable ports: 44 units in 2-stage tree
- Other technical considerations needed for optimal custom switch sizing

TABLE XXV. Parameters of the readout design.

Parameter	Hit-only option	Waveform option
Pre-trigger input data rate	5,600 MB/s	23,400 MB/s
Number of RBUs	38	122
Input rate to each RBU	150 MB/s	188 MB/s
Latency provided by RBU (pre-trigger buffer length)	109 s	87 s
Trigger info output rate per RBU	50 MB/s	15 MB/s
TPU data input rate (for 16 TPUs in detector)	117 MB/s	117 MB/s

**HK Design Report, Nov. 30 2018**  
**(46,700 PMTs scenario)**

■ **Hit-only option**

5,600 MB/s / 2000 FE = 2.8 MB/s = 25 Mbit/s

■ **Waveform option**

23,400 MB/s / 2000 FE = 11.7 MB/s = 100 Mbit/s

**HK Technical Note 0005 Apr. 2019 (~55,000 PMTs scenario)**

the data rates from each FEE are expected to be ~8 MB/s for both normal data taking and for a far supernova and for a near supernova ~117 MB/s in the first second and ~189 MB/s for the next 10 s. supernova (this peak rate increases to 235 MB/s depending on PMT location, so buffering on the FEEs exist to handle this momentary saturation in the first second of a near supernova). However

■ **Near supernova worst case data estimate**

235 MB/s x 11 s = 2.6 GB per FE

→ does not seem an issue with 4 GB buffer at FE



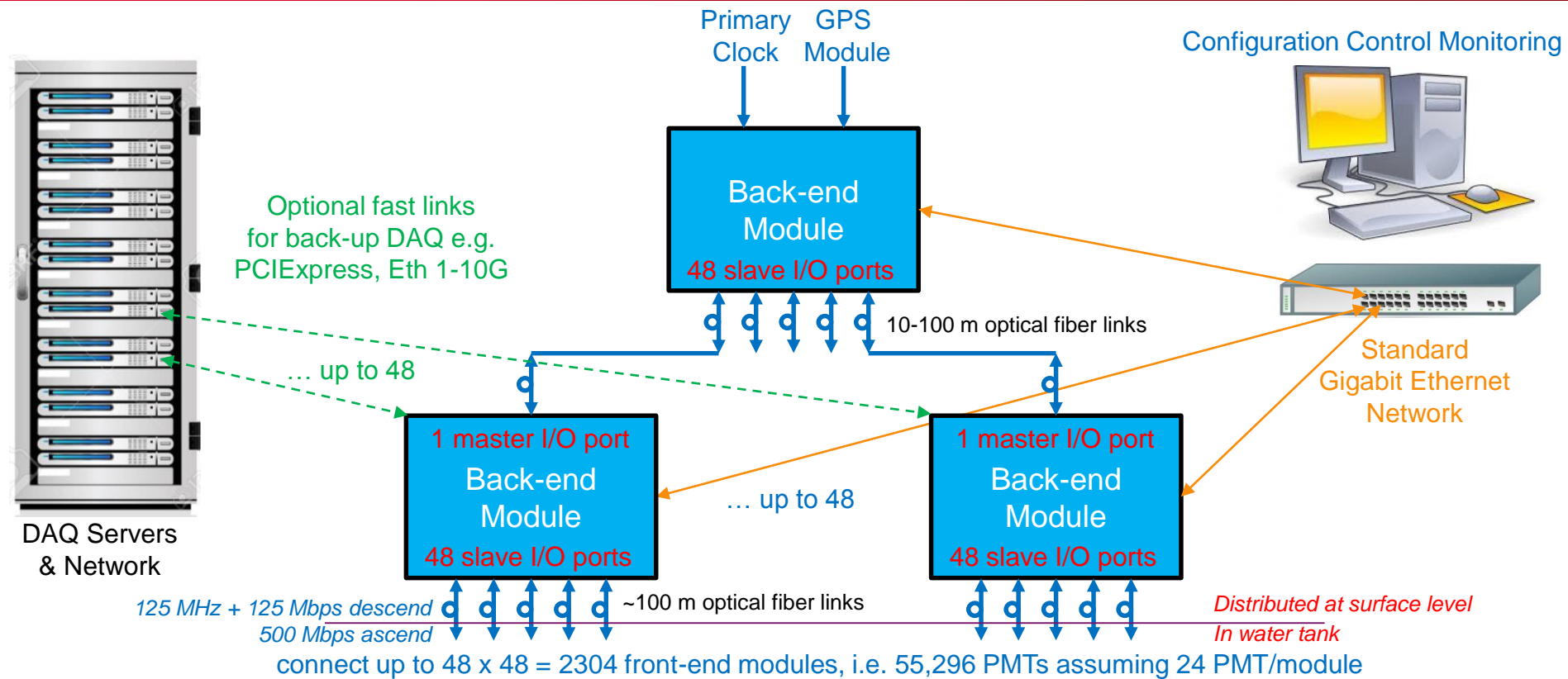
TE803  
Xilinx Zynq Ultrascale+  
4 GB DDR4  
300 € per unit (Qty 1000)

■ **Example scenario with FE equipped with 4 GB buffer and 100 Mbps DAQ link**

8 MB/s (64 Mbps) used for continuous normal data taking => 36 Mbps left for supernovae burst data transfers

Supernova data acquisition time (if limited by FE link speed): 2.6 GB / 36 Mbps = ~10 minutes

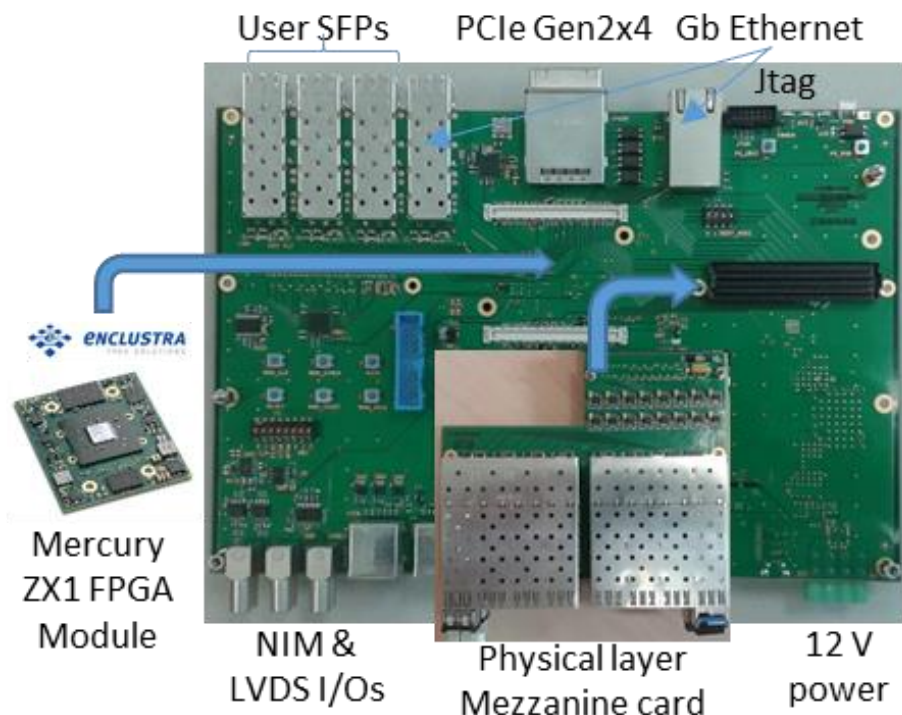
→ 100 Mbps sufficient (?) normal data taking + supernovae rate of 1 every 100 minutes (with 10% dead-time)



## Principles

- A dual stage fanout tree composed of custom back end-modules: 1 root and up to 48 leaves
- An ordinary Gigabit Ethernet network for global configuration, control and monitoring
- An optional fast data link from each leaf back-end module for back-up DAQ





D. Calvet, «Back-End Electronics Based on an Asymmetric Network for Low Background and Medium-Scale Physics Experiments », in IEEE Transactions on Nuclear Science, Vol. 66, N°7, pp. 998-1006, July 2019.

## Trigger and Data Concentrator Module - TDCM

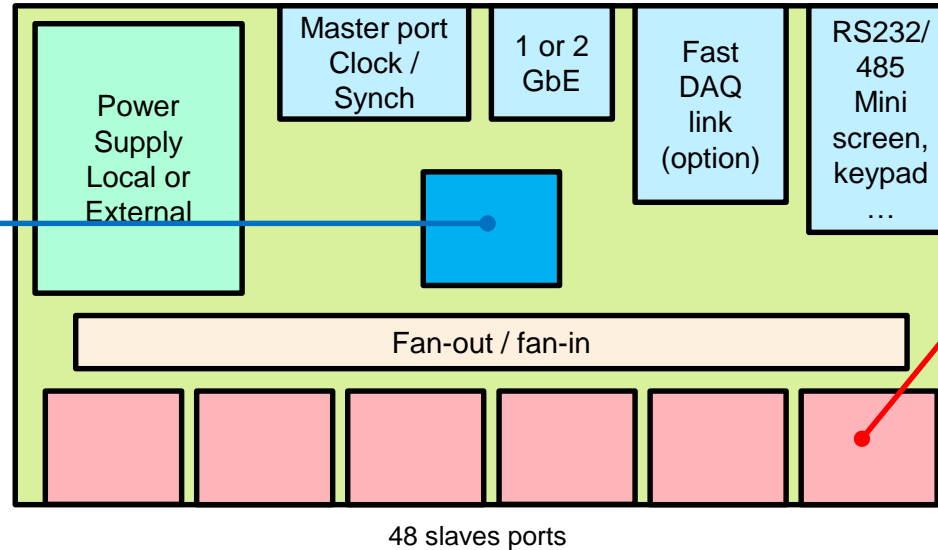
- 1 Master port + up to 32 Slave ports. 1x100 Mbps descend (100 MHz reference clock, trigger, configuration) and 32 x 400 Mbps ascend (detector data, monitoring)
- Low speed link for control & DAQ: Ethernet 1 Gbps RJ45 or GBIC (copper) / SFP (optical)
- High speed links: 1-3 optical links (6.6-10 Gbps) + PCI Express Gen 2 x 4 (untested – not used)
- Actual production cost: 3.8 k€ in 32 ports version equipped with 850 nm transceivers (10 TDCMs)

→ Board now in production for integration with HA-TPCs in T2K near detector upgrade

trenz electronic  
TE803 249 €/u  
TE808 914 €/u



Xilinx Zynq UltraScale+, 2 GByte DDR4  
(4 Gbyte DDR4 avec TE808)



Stackable SFP 2x8



## Main Features

- Based on Zynq Ultrascale+ SoC, Trenz TE803 or TE808
- 1 Master port and 48 slave ports. 8x125 Mbps descend using CDCM (8x500 Mbps without) 48 x 500 Mbps ascend – (1.25 Gbps per port reachable from datasheet, but let's check first...)
- Low speed 1 GbE for control or slow DAQ
- Optional high speed links: Dual/Triple Ethernet 1G, Ethernet 10G, possibly PCIe Gen 2 or 3
- Format options: 6U x 14F (6 per 6U crate) or 1U x 84F (same form factor as WR3/4 switches)
- Estimated cost: 4-5 k€ per unit i.e. 180-220 k€ for 48,000 PMTs (no spare, no contingency)

→ Demonstrator aimed for realistic performance studies, but not planned to be a final design

## Position within HK collaboration and HK WP4 Clock Distribution

- Presently two concepts are subject to R&D:
  - Direct distribution SK-like.
  - Clock embedded into data (clock and data recovery concept)
    - Custom solution
    - White Rabbit

### I propose that a third concept is added:

- Data embedded into clock (modulated clock concept)
  - Custom solution – based on the demonstrator previously described

→ Ready to present this proposal to HK WP4 and HK collaboration.

Proposal to be exposed, discussed and refined with decision makers at IRFU and DEDIP  
- internal scientific council scheduled early February 2021