



Possible use of Infiniband and Ethernet in the CMS DAQ system

IB workshop 14-15 Jan 2013

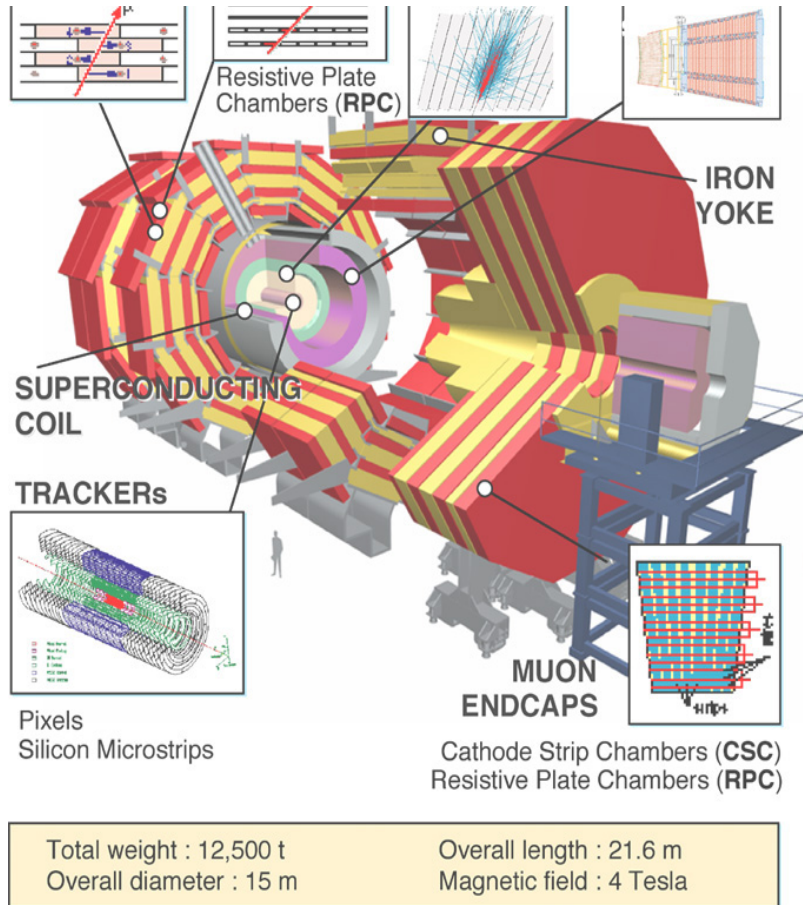
Frans Meijers (CERN-PH)

On behalf of the CMS DAQ group



CMS design parameters and DAQ requirements

Detectors



Detector	Channels	Control	Ev. Data
Pixel	60000000	1 GB	50 (kB)
Tracker	10000000	1 GB	650
Preshower	145000	10 MB	50
ECAL	85000	10 MB	100
HCAL	14000	100 kB	50
Muon DT	200000	10 MB	10
Muon RPC	200000	10 MB	5
Muon CSC	400000	10 MB	90
Trigger		1 GB	16

Average Event size

1 Mbyte

Max LV1 Trigger

100 kHz

Online rejection

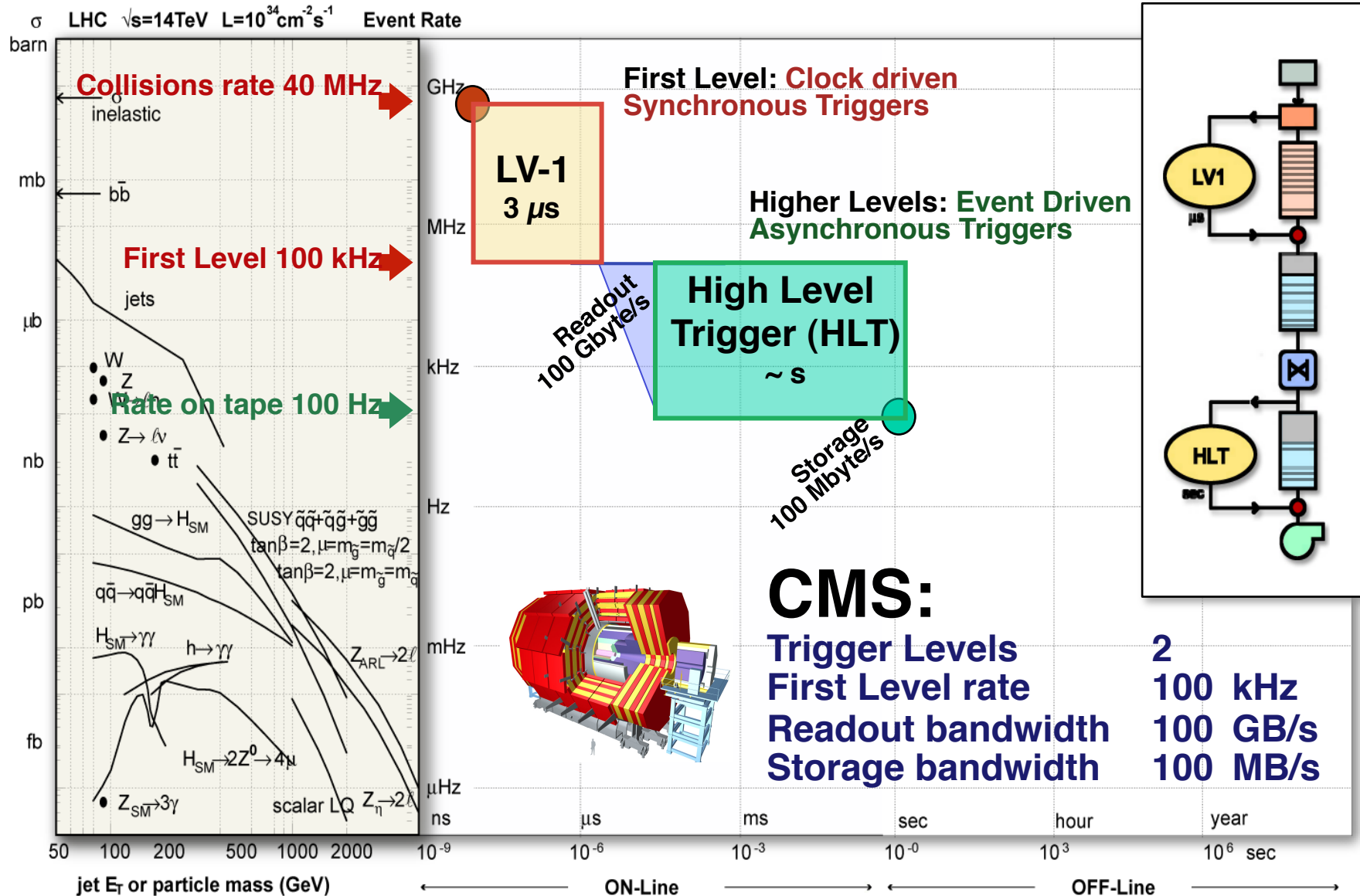
99.999%

System dead time

~ %

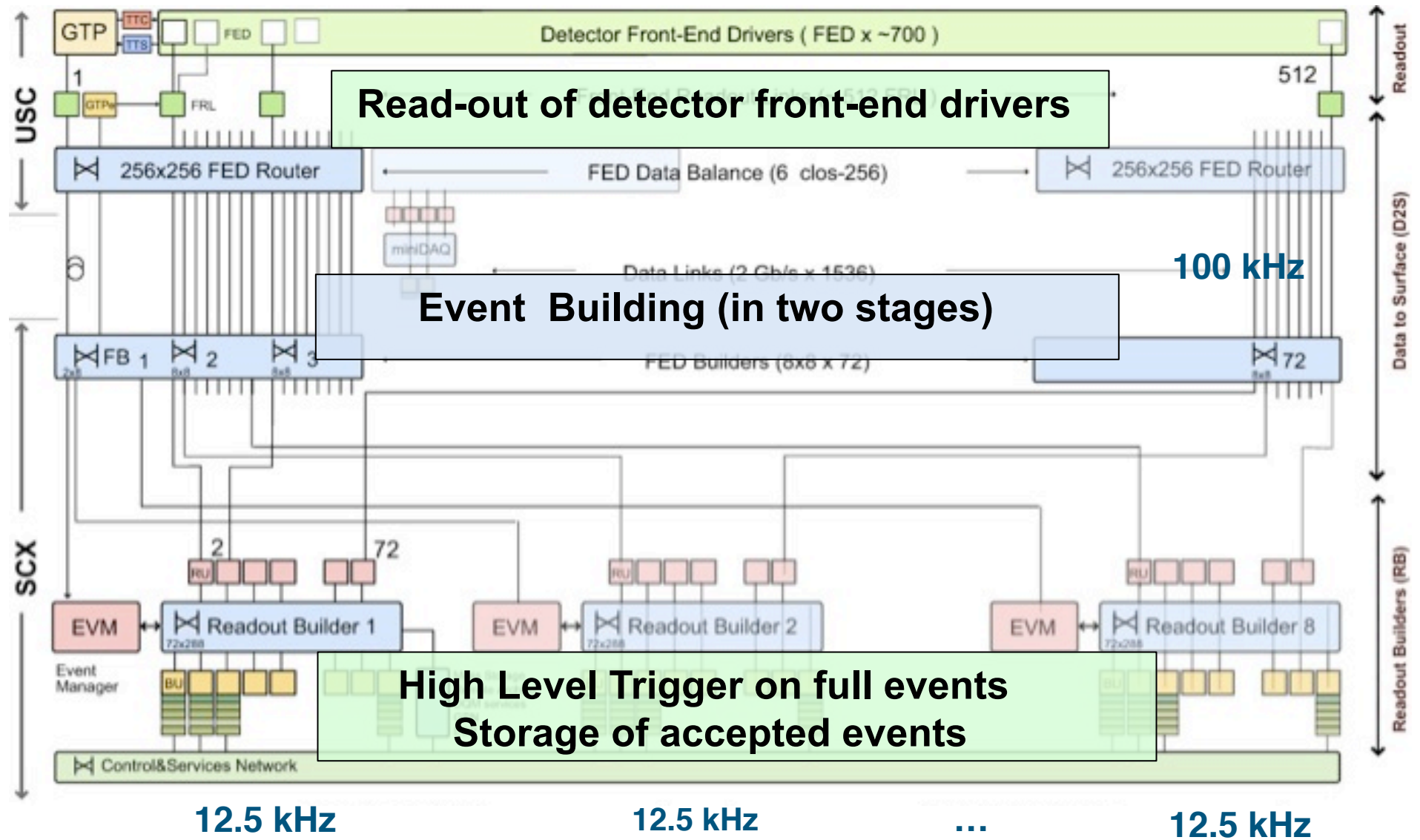


Two Trigger levels



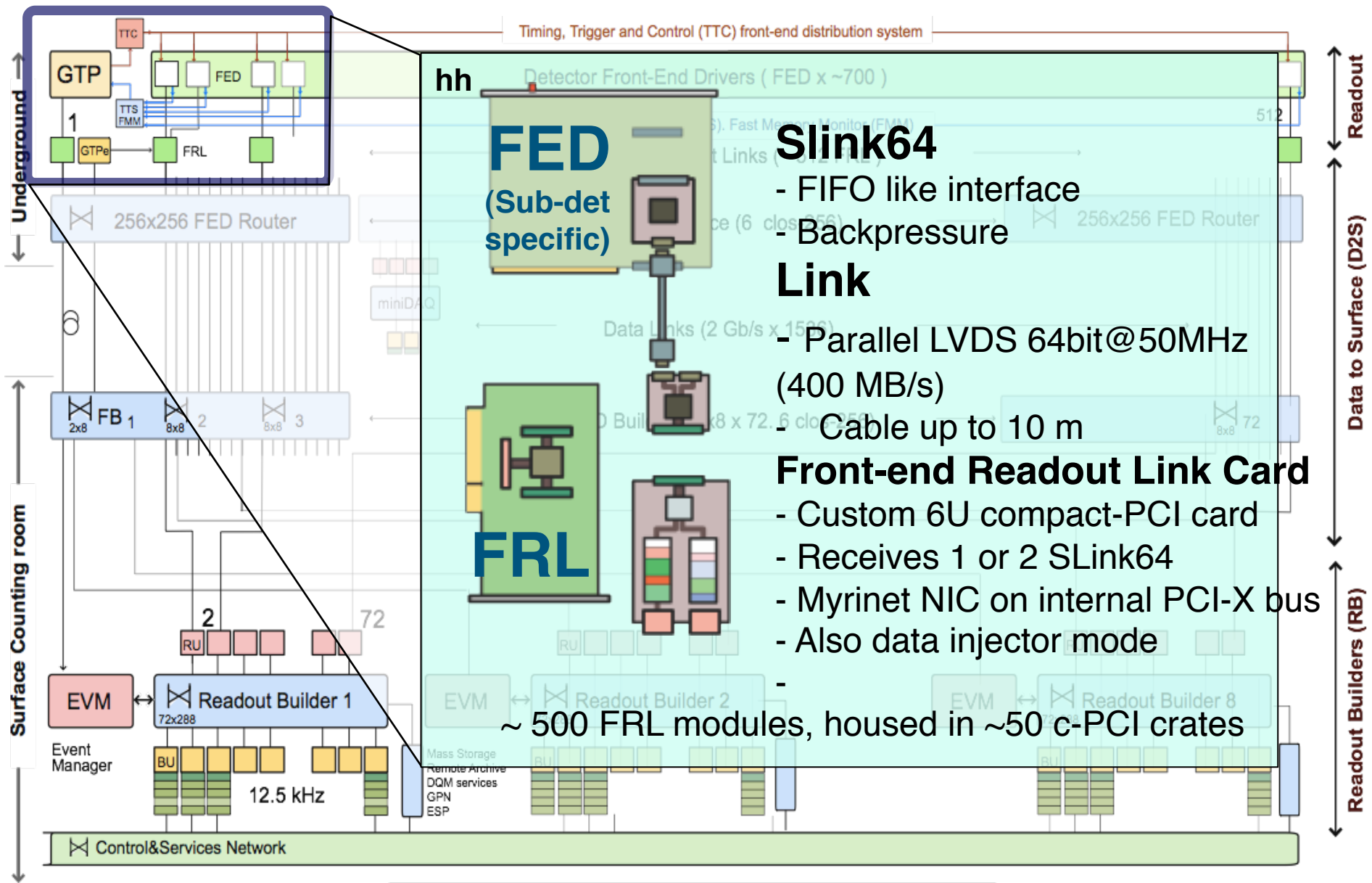


CMS DAQ1 (2008-2012)





Uniform interface - Readout



hh
FED
 (Sub-det specific)

SLink64

- FIFO like interface
- Backpressure

Link

- Parallel LVDS 64bit@50MHz (400 MB/s)
- Cable up to 10 m

Front-end Readout Link Card

- Custom 6U compact-PCI card
- Receives 1 or 2 SLink64
- Myrinet NIC on internal PCI-X bus
- Also data injector mode

~ 500 FRL modules, housed in ~50 c-PCI crates



LHC – CMS outlook

← 5 year life-cycle →

	2011	2012	2013	2014	2015	2016	2017	2018
	7 TeV		LS1		14 TeV		LS2	
lumi 10^{34} /cm ² /s	0.2	0.5				~ 1 x		
events/xing 25/50 ns	4/8	10 / 20				~20 / 40		
Tracker Muons CALO Trigger					new Pixel, more channels complete forward muons new HCAL sensors and electronics uTCA in parallel ('spectator')			

	2018	2019	2020	2021	2022
	LS2	14 TeV		LS3	14 TeV
lumi 10^{34} /cm ² /s		~ 2 x			~ 5 x
events/xing 25/50 ns		~ 40 / 80			~ 100 / 200
Tracker Muons CALO Trigger		uTCA in production			new strips, 5x tracking trigger

One of the possible scenarios being discussed

CONSIDERING
1 MHz R/O and
10 MB event size



DAQ2 for post-LS1

- Replacement cycle
 - PC and network replaced typically each 5 years
- New requirements
 - Some sub-det new back-end electronics in uTCA standard with serial link to cDAQ
 - Some sub-detectors will be replaced which lead to higher data volumes
- Keep “external” boundaries
 - Inputs (custom electronics)
 - About 500 2-4 Gbps “Legacy” FEDs
 - About 10-100 6-10 Gbps New FEDs
 - Output to HLT farm
 - About 500 “Legacy” nodes 1-2 Gbps
 - About 400 new nodes 204 Gbps

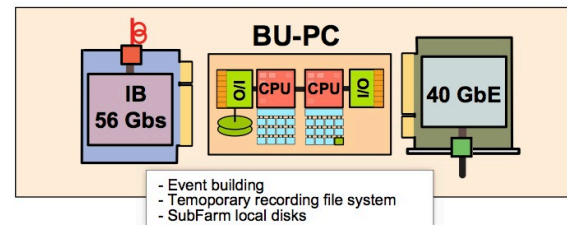
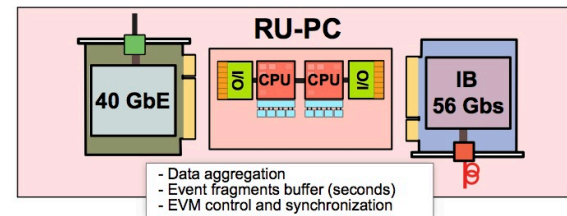
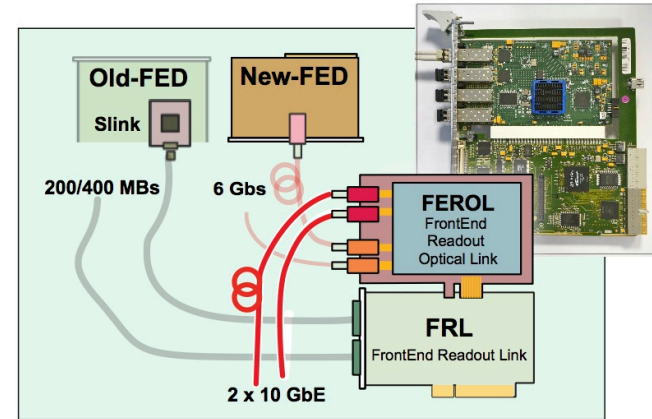
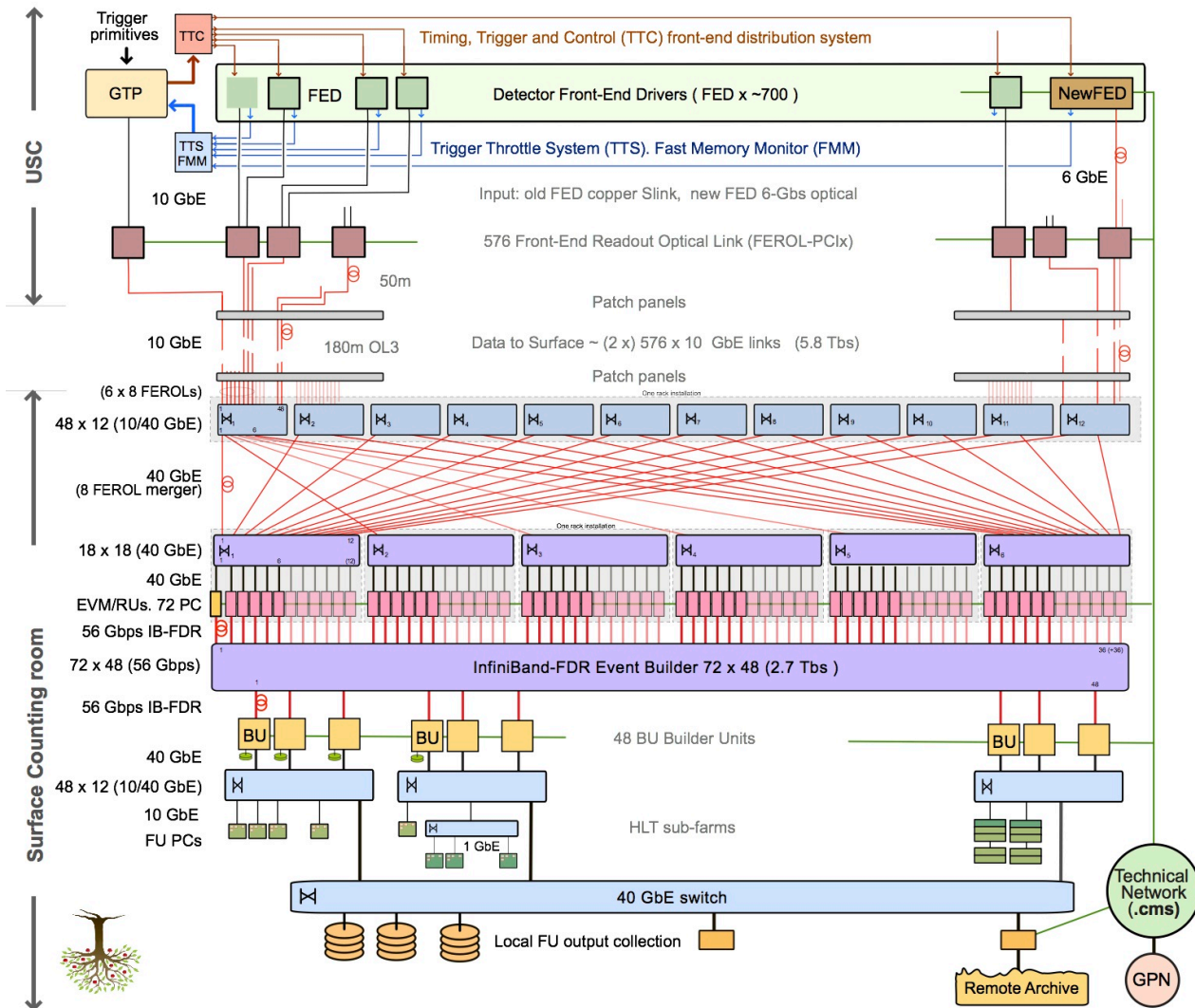


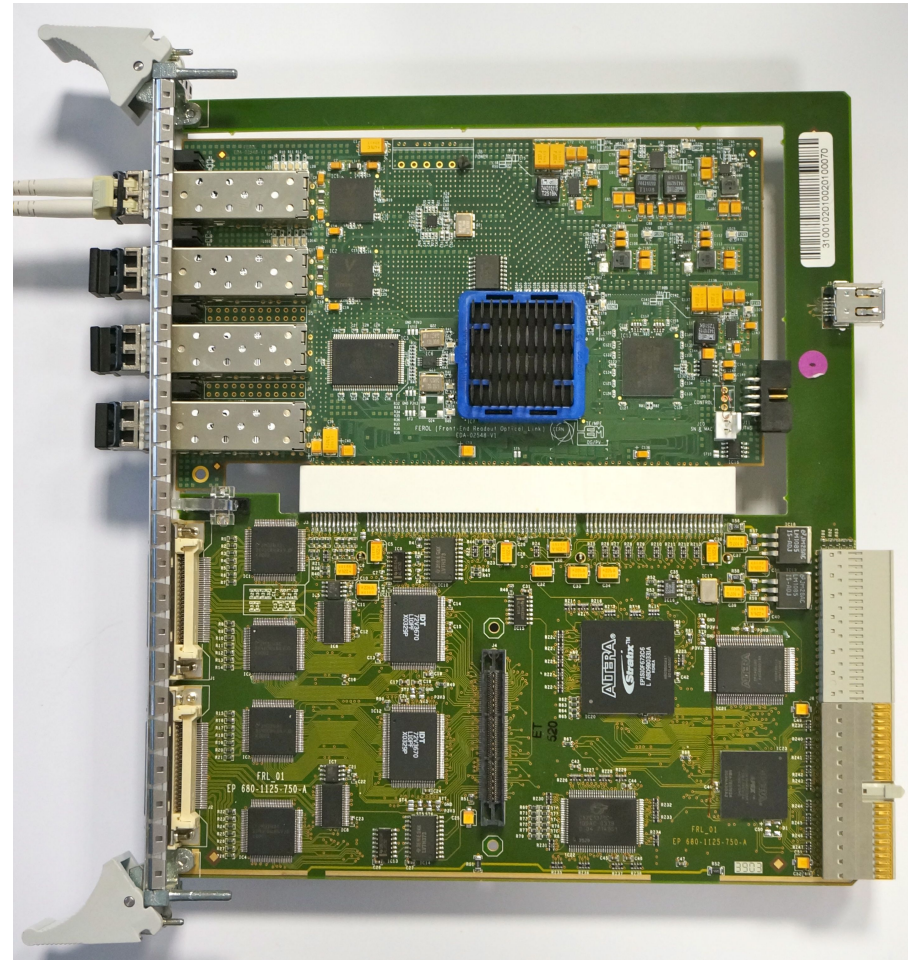
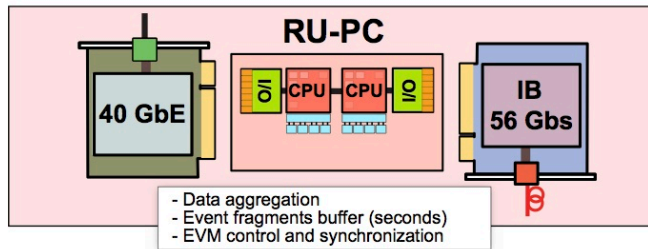
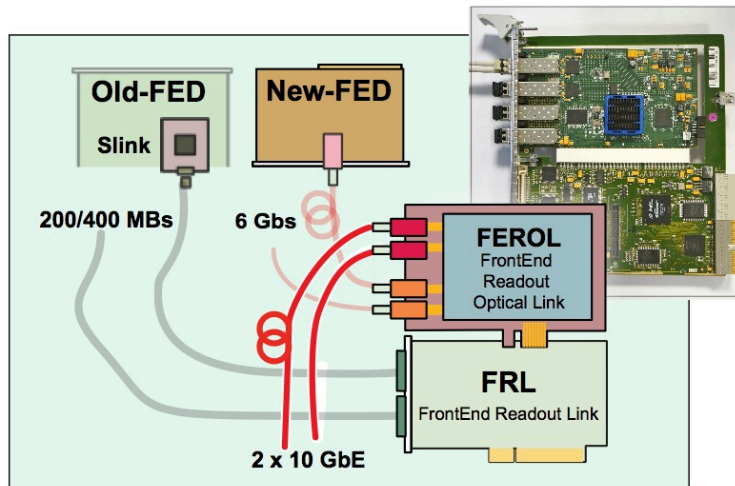
DAQ2 for post-LS1

- DAQ2
 - re-implementation with up-to-date technology
 - Typically 10x less nodes with 10x more performance
 - DAQ1: 2x 2 Gbps Myrinet and 3x1GbE
 - Consider 10 GbE, 40 GbE, IB FDR (56 Gbps)
 - Timescale
 - LS1 in 2013 - 2014
 - Design, evaluation, order, for delivery and installation Q4 2013
 - Switchover DAQ1 to DAQ2 Apr-2014, commissioning, improvements
 - Deployment for physics 2015-16-17

DAQ2

CMS DAQ upgrade for post LS1 (DAQ2)





- FEROL
 - Input: custom protocol
 - Output: 10 GbE serial, reduced TCP/IP sender in FPGA
- Receiver with NIC in PC with standard driver and TCP/IP stack



Transmission Standards	100 Mb Ethernet	1 Gb (1000 Mb) Ethernet	10 Gb Ethernet	40 Gb Ethernet	100 Gb Ethernet
OM1 (62.5/125)	up to 550 meters (SX)	220 meters (SX)	33 meters (SR)	Not supported	Not supported
OM2 (50/125)	up to 550 meters (SX)	550 meters (SX) ^[2]	82 meters (SR) ^[2]	Not supported	Not supported
OM3 (50/125)	up to 550 meters (SX)	550 meters (SX)	300 meters (SR)	100 meters ^[2]	100 meters ^[2]
OM4 (50/125)	up to 550 meters (SX)	1000 meters (SX) ^[2]	550 meters (SR) ^[2]	150 meters ^[2]	150 meters ^[2]

AFBR-79EEPZ

AFBR-79EEPZ QSFP+ eSR4 Pluggable, Parallel Fiber-Optics Module for 40Gigabit and 10Gigabit Ethernet Applications

All Detail

Description

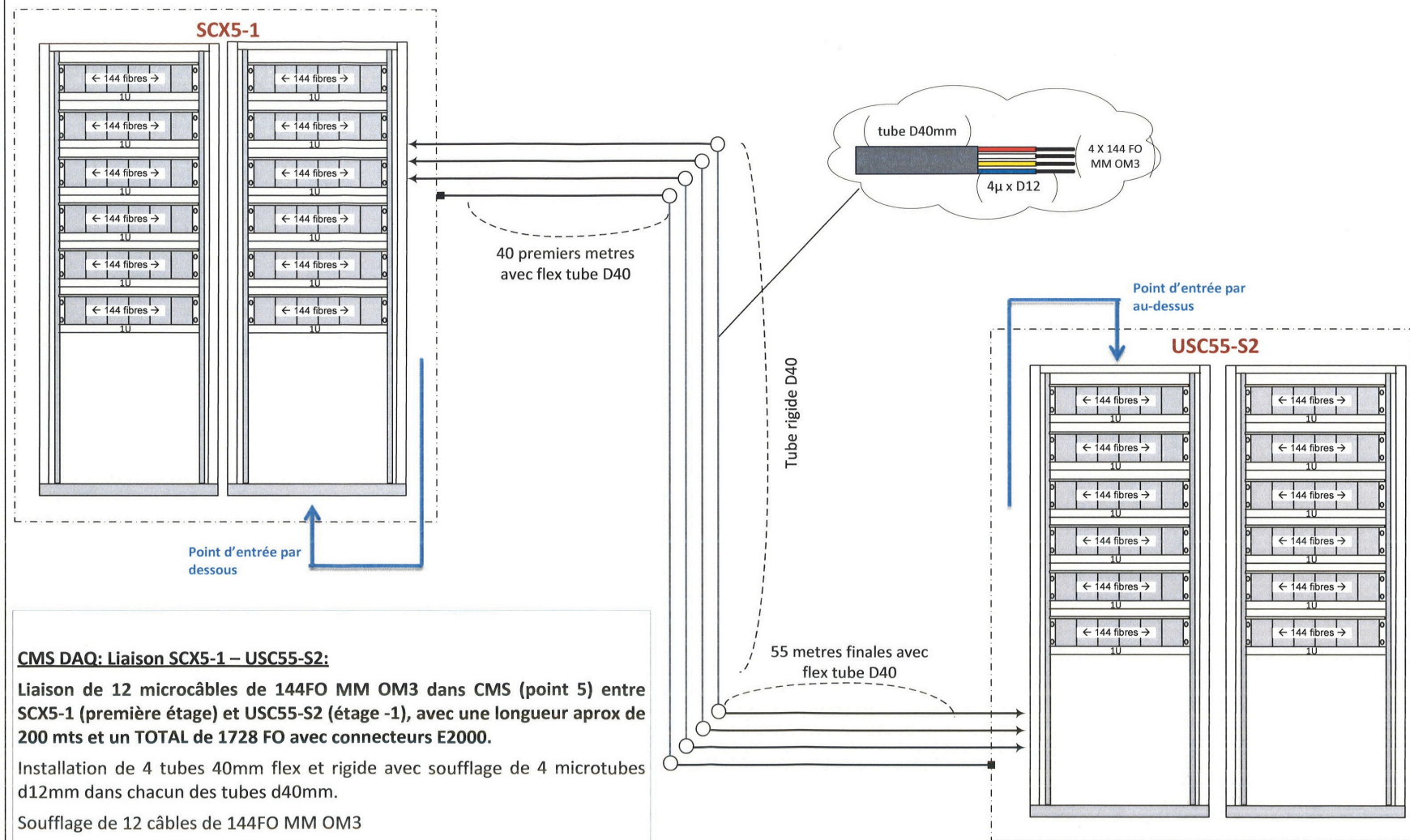
Features

The Avago Technologies AFBR-79EEPZ is a Four-Channel, Pluggable, Parallel, Fiber-Optic QSFP+ Transceiver with integrated push-pull tab for 40 Gigabit Ethernet (40GbE) applications with extended reach up to 300m (400m) using OM3 (OM4) multimode fiber (MMF) and with added capability of inter-operating with IEEE 10GBASE-SR compliant products. This transceiver is a high performance module for short-range multi-lane data communication and interconnect applications. It integrates four data lanes in each direction with each lane operating at 10.3125 Gbps, giving an aggregated bandwidth of 40 Gbps. This transceiver can also be used for high density 10 gigabit Ethernet application. It allows optical interoperability up to 300m (400m) using OM3 (OM4) MMF with any 10 gigabit Ethernet (10GbE) transceivers, compliant to the IEEE 802.3ae 10GBASE-SR specifications, of form factors such as SFP+, XFP and X2. It provides an effective port count of over 100 within 1 RU rack. The push-pull tab facilitates the insertion and extraction of these transceivers in such high density environment. This transceiver is designated as a QSFP+ eSR4 solution, where the letter "e" represents extended reach beyond 100m MMF with interoperability between this QSFP+ transceiver and any IEEE 10GBASE-SR compliant modules.



Lifecycle status:
Active

CMS DAQ: Liaison SCX5 (3562) – USC55 (3524)



CMS DAQ: Liaison SCX5-1 – USC55-S2:

Liaison de 12 microcâbles de 144FO MM OM3 dans CMS (point 5) entre SCX5-1 (première étage) et USC55-S2 (étage -1), avec une longueur approx de 200 mts et un TOTAL de 1728 FO avec connecteurs E2000.

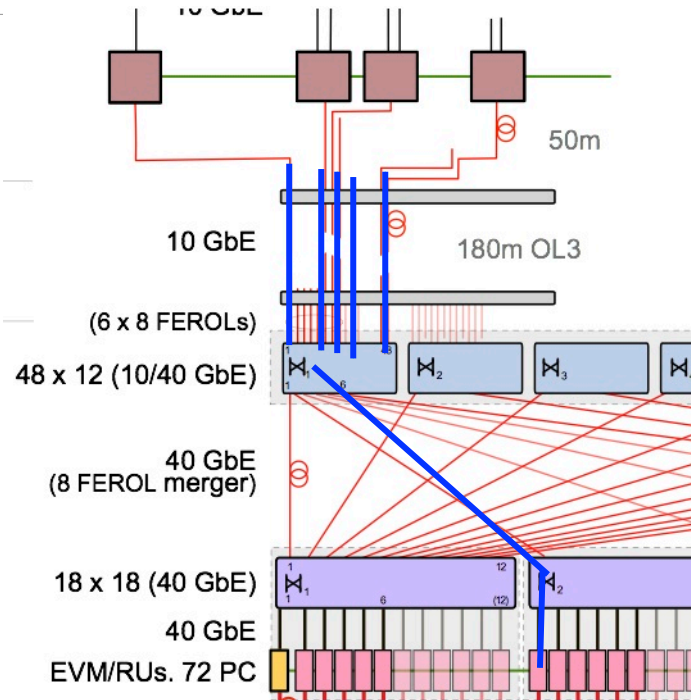
Installation de 4 tubes 40mm flex et rigide avec soufflage de 4 microtubes d12mm dans chacun des tubes d40mm.

Soufflage de 12 câbles de 144FO MM OM3

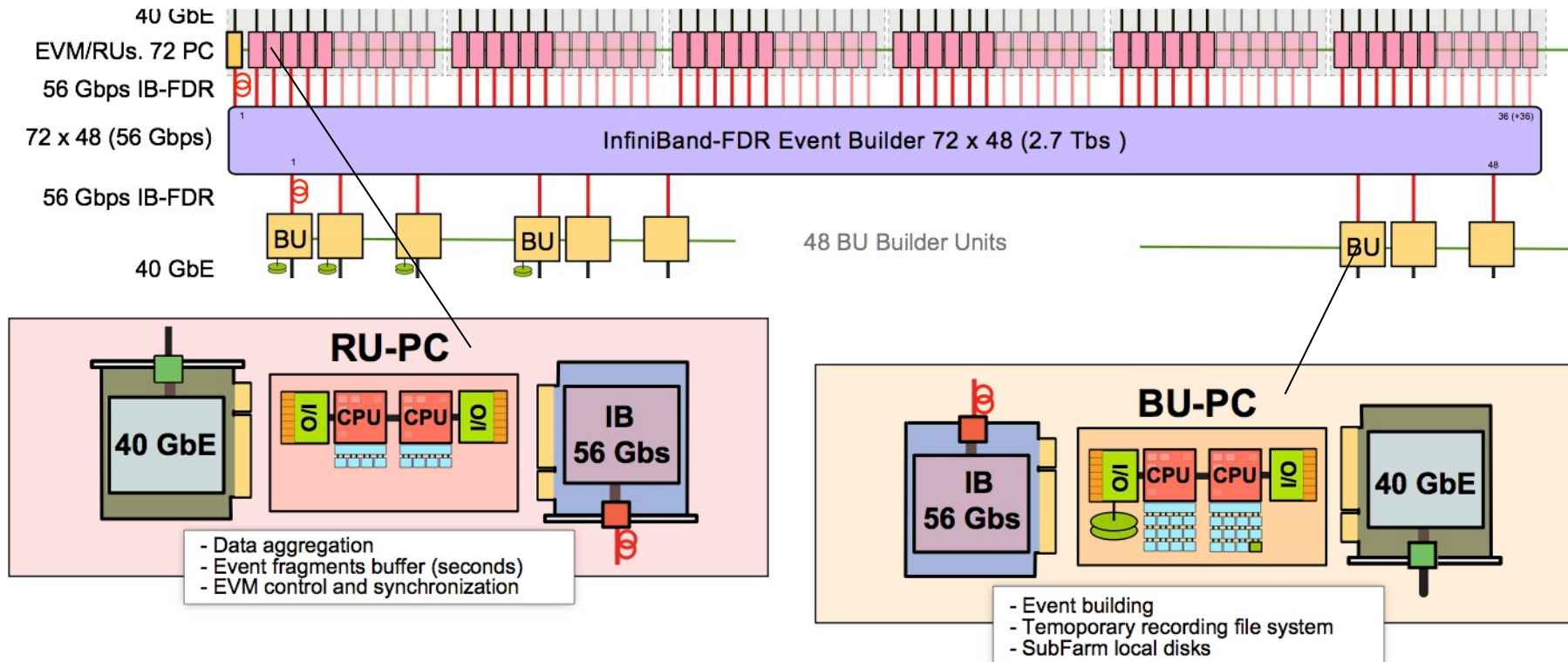
Raccordement de 1728 FO E200 MM OM3



FEROL aggregation



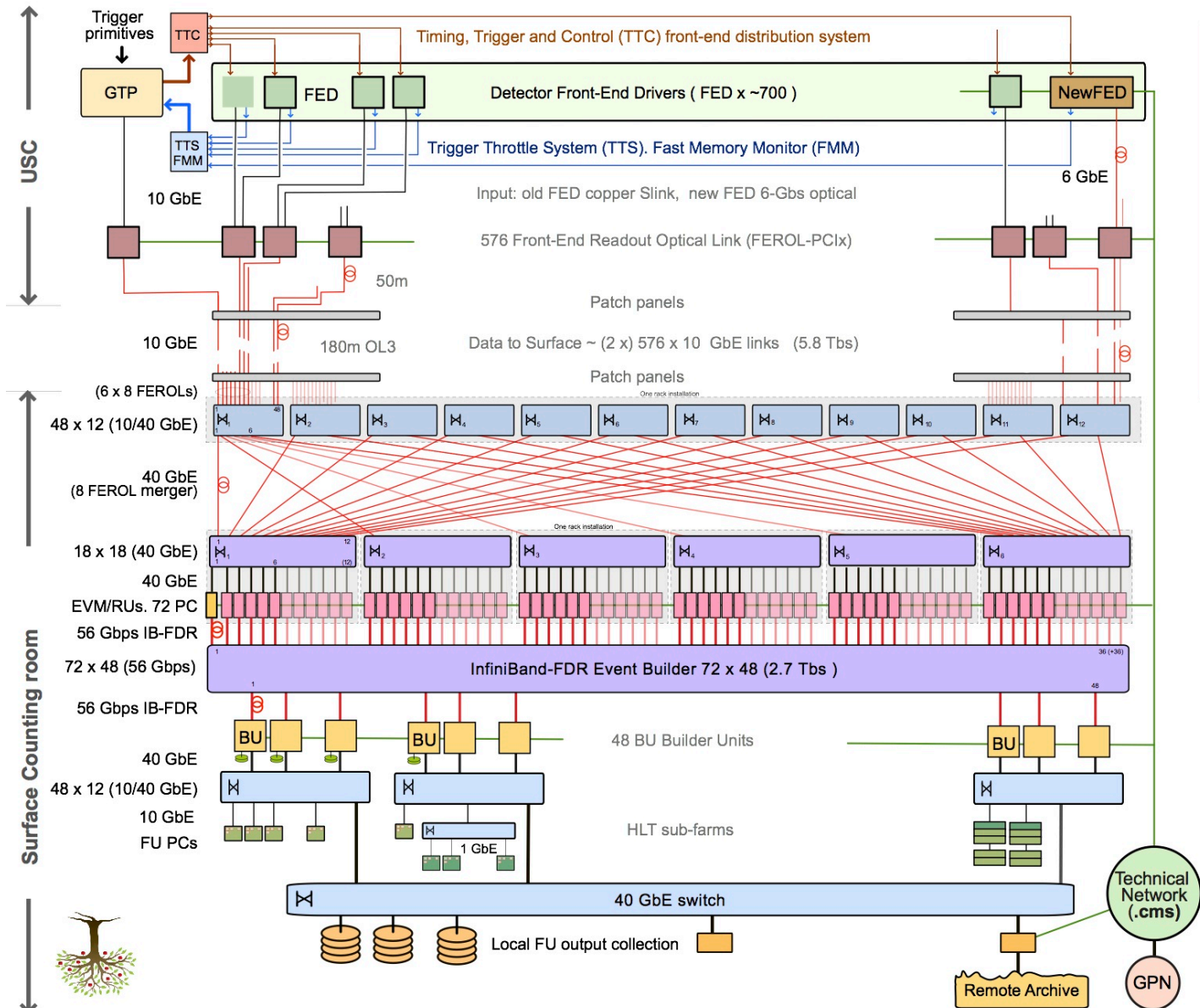
- Aggregation n-to-1, example
 - 16 FEROLs each sending 2 Gbps over 10 GbE link
 - Concentrated in one 40 GbE NIC into PC
 - Reliability and Congestion handled by TCP/IP
- USC – SCX 180m,
 - with OM3 fibres up to 200 m
 - 40 GbE (with 4 lanes 10 Gbps) max. is 150 m – NOT feasible
- Network useful to re-configure when fault with optic, PC, etc



- Performance Scaling with multi-layer switch network?
 - 3 layer Clos
- Implement with “Director” switch or 36-port units?



DAQ2





BACKUP MATERIAL



3 types of Ethernet switches (Mellanox)



SX1016
64 ports 10GbE FSP+



SX1030
36 ports 40GbE QFS



SX102
48 ports 10GbE FSP
12 ports 40GbE QFS

Note: Can use “break-out” cables for 40 GbE to 4x10 GbE (with some restrictions on total number of ports)