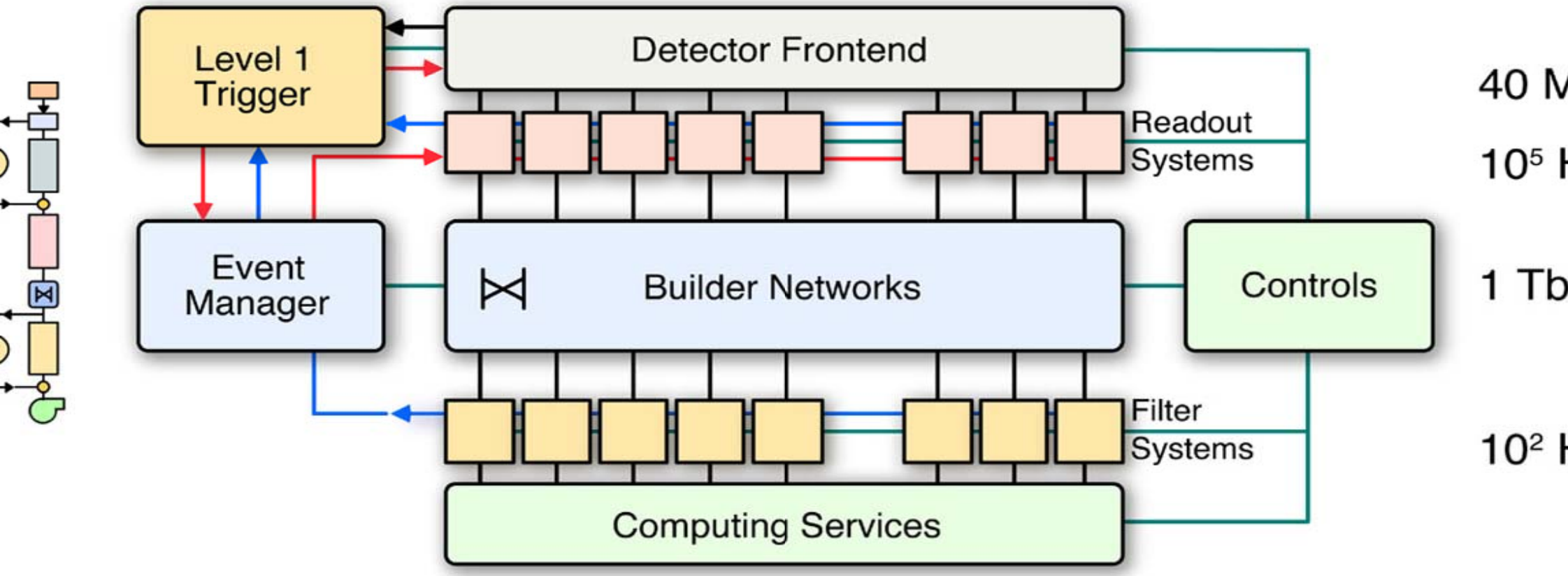
DAQ upgrades at SLHC

S. Cittolin CERN/CMS, 22/03/07

DAQ architecture. TDR-2003
DAQ evolution and upgrades



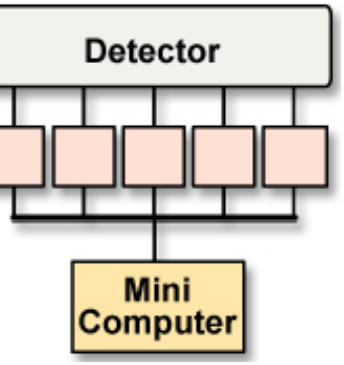
rameters

lision rate		40 MHz
vel-1 Maximum trigger rate		**100 kHz**
erage event size	≈	**1 Mbyte**
ta production	≈	Tbyte/day

TDR design implementation

No. of In-Out units		**512**
Readout network bandwidth	≈	**1 Terabit/s**
Event filter computing power	≈	**10^6 SI95**
No. of PC motherboards	≈	Thousands

(SLHC: x 10)

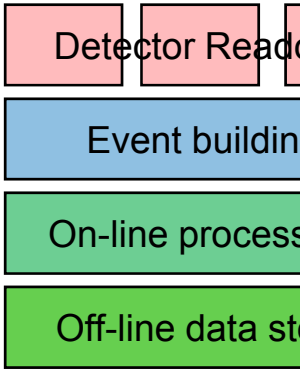


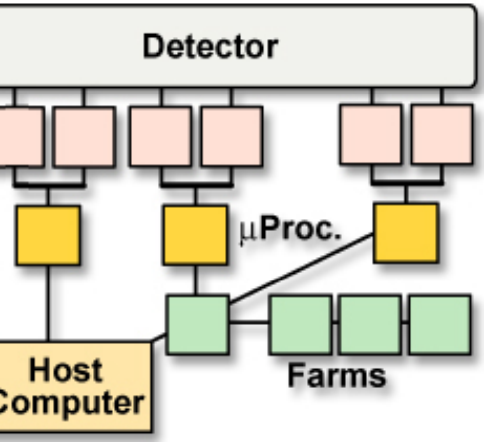
PS:1970-80: Minicomputers

Readout custom design

First standard: CAMAC

- **kByte/s**



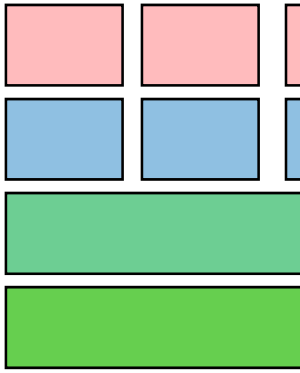


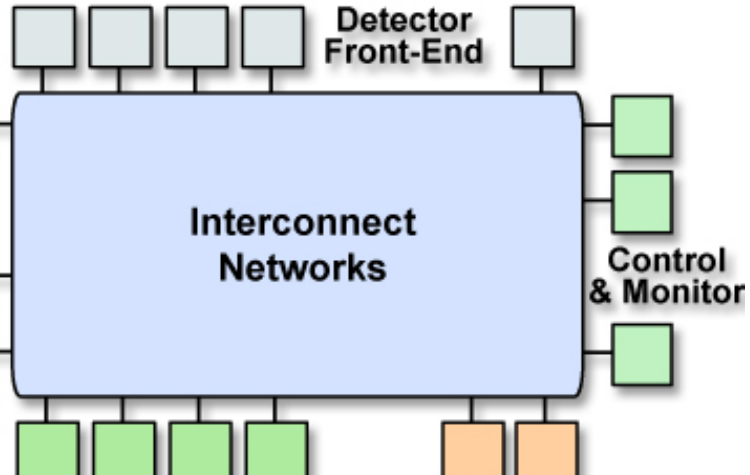
LEP:1980-90: Microprocessors

HEP standards (Fastbus)

Embedded CPU, Industry standards (VME)

- **MByte/s**



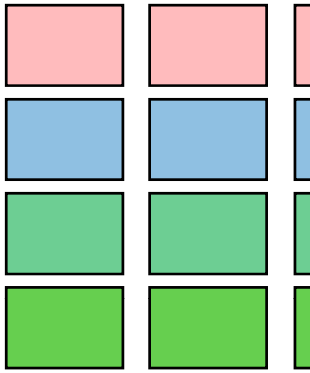


LHC: 200X: Networks/Grids

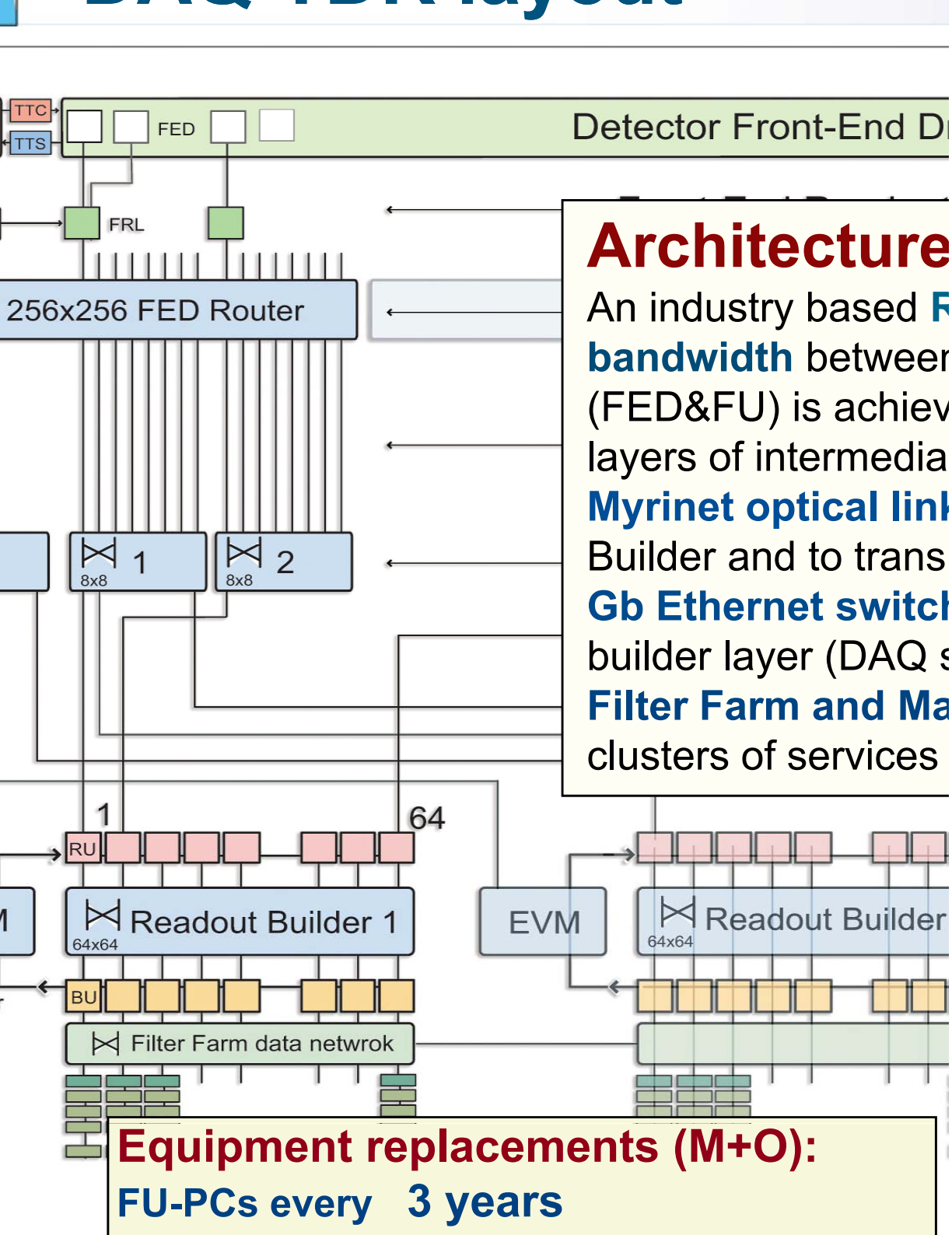
IT commodities, PC, Clusters

Internet, Web, etc.

- **GByte/s**



DAQ TDR layout

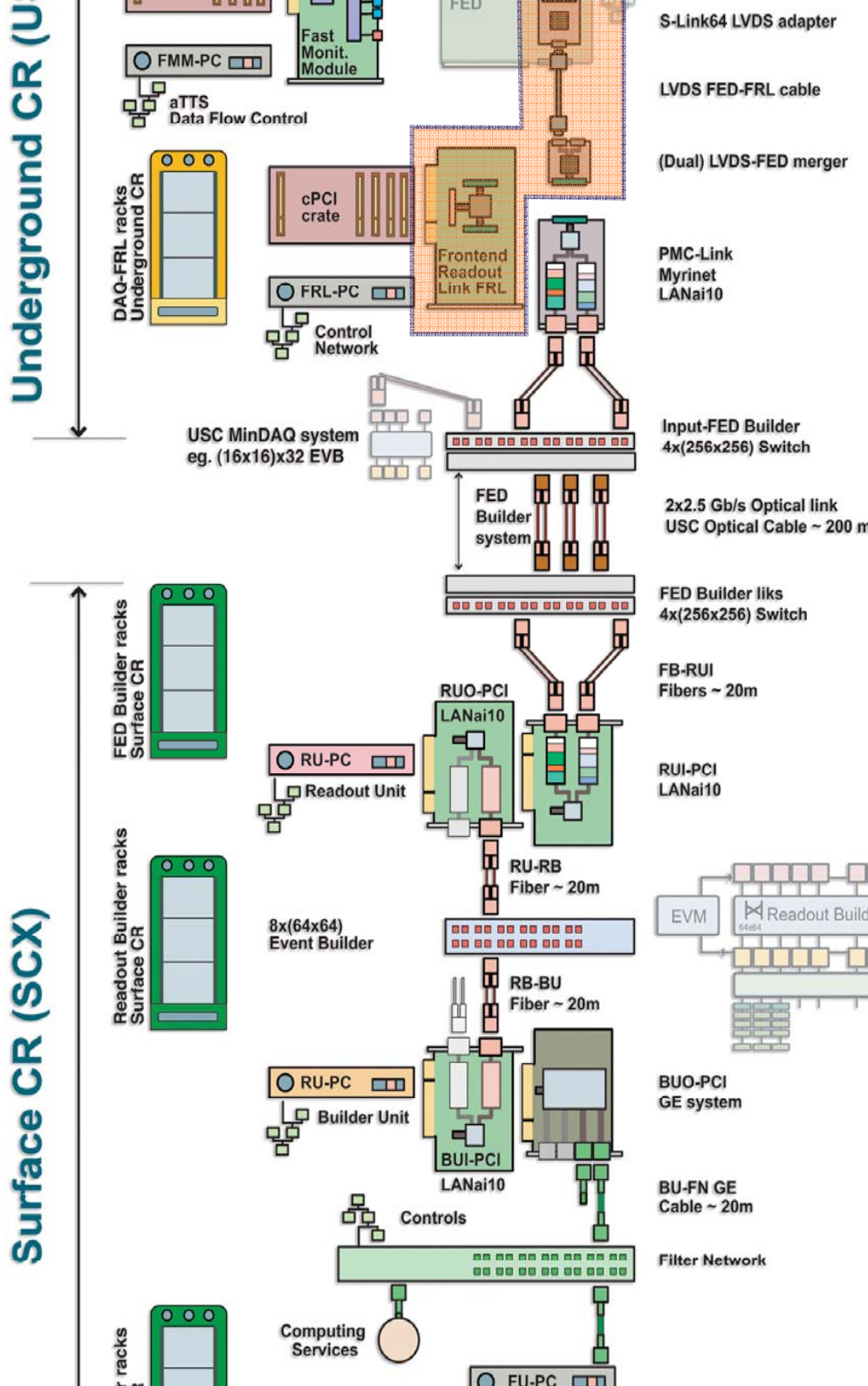


Architecture

An industry based F
bandwidth betweer
(FED&FU) is achiev
layers of intermedia
Myrinet optical link
Builder and to trans
Gb Ethernet switch
builder layer (DAQ s
Filter Farm and Ma
clusters of services

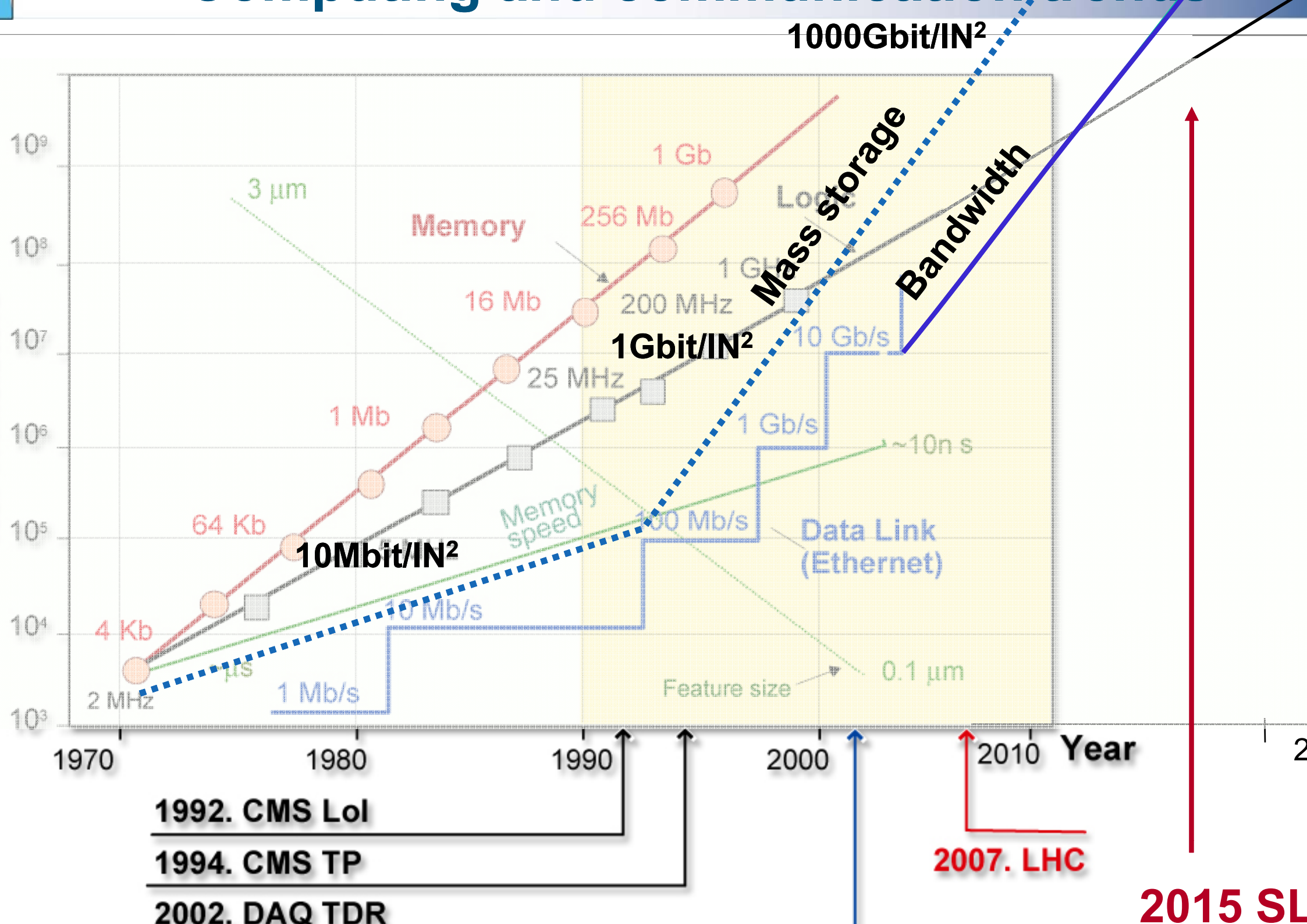
Equipment replacements (M+O):

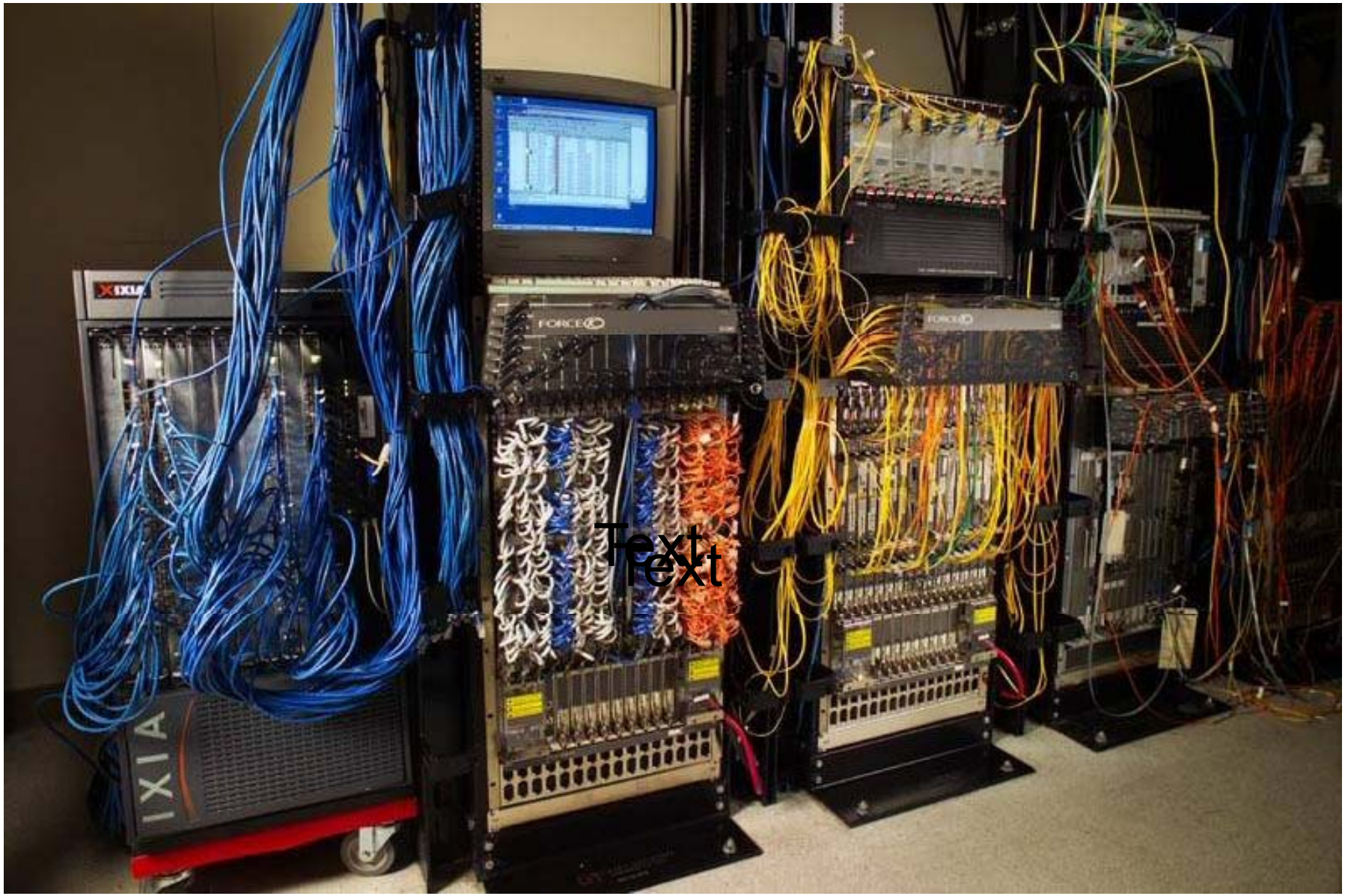
FU-PCs every 3 years



Same level 1 trigger rate, 20 MHz crossing, higher occupancy..
A factor 10 (or more) in readout bandwidth, processing, storage

Factor 10 will come with technology. Architecture has to exploit it
New digitizers needed (and a new detector-DAQ interface)
Need for more flexibility in handling the event data and operating the
experiment

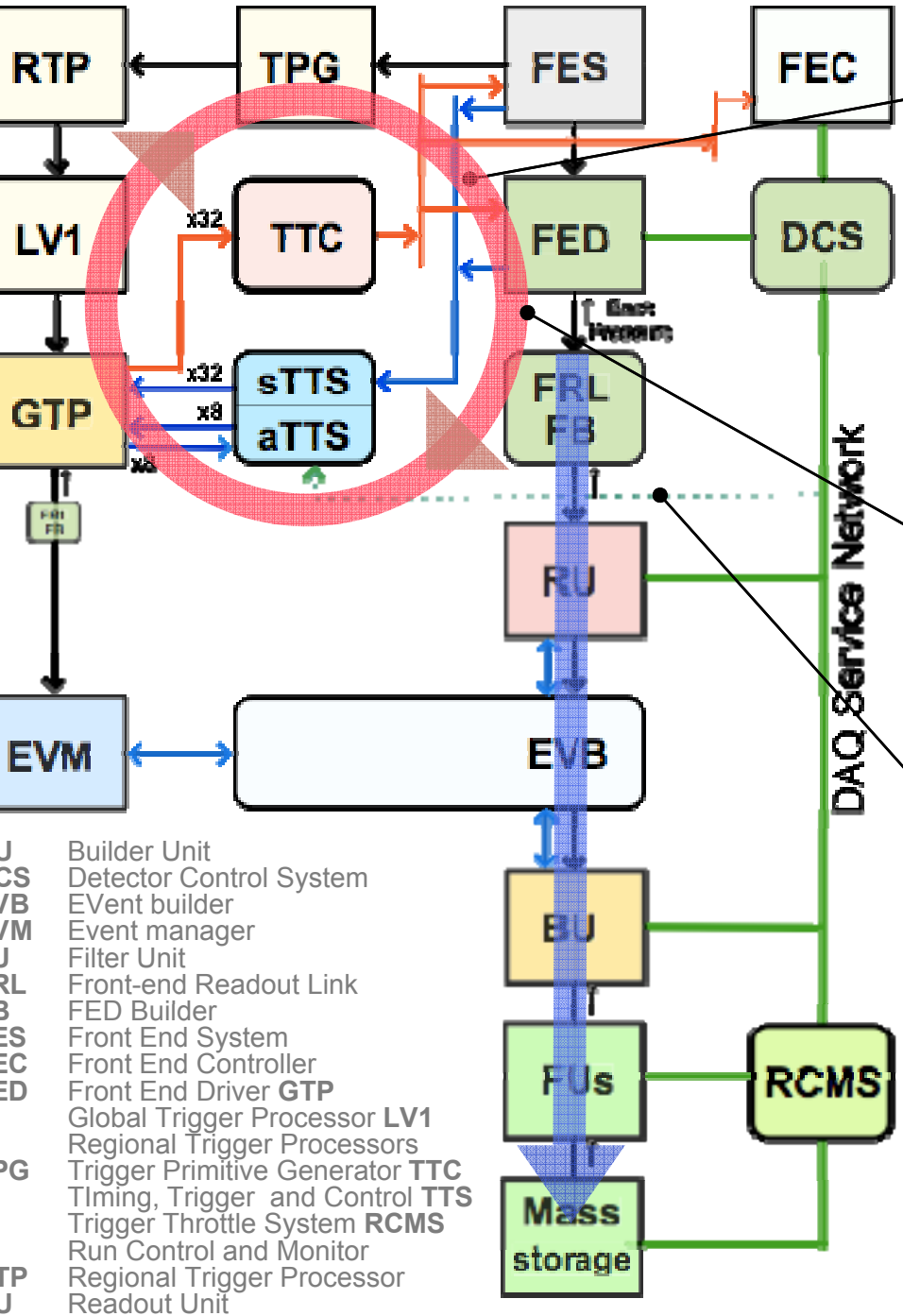

1000Gbit/IN²
Mass storage
Bandwidth
1 Gb
3 µm
Memory
256 Mb
Logic
16 Mb
200 MHz
1Gbit/IN²
10 Gb/s
25 MHz
1 Mb
1 Gb/s
~10n s
64 Kb
Memory speed
100 Mb/s
Data Link (Ethernet)
10Mbit/IN²
10 Mb/s
4 Kb
µs
2 MHz
1 Mb/s
Feature size
0.1 µm
10⁹
10⁸
10⁷
10⁶
10⁵
10⁴
10³
1970
1980
1990
2000
2010
Year
1992. CMS LoI
1994. CMS TP
2002. DAQ TDR
2007. LHC
2015 SL



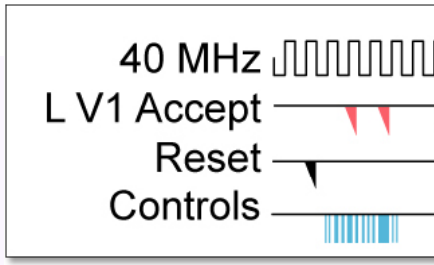
http://www.force10networks.com/products/reports.asp

is a photograph of the configuration used to test the performance and resiliency metrics of the Force10
ale E-Series. During the tests, the Force10 TeraScale E-Series demonstrated **1 billion packets per secc**
hput, making it the world's first Terabit switch / router, and ZERO packet loss hitless fail-over
t speeds. The configuration required 96 Gigabit Ethernet and eight 10 Gigabit Ethernet ports f

LHC-DAQ: trigger loop & data flow



U Builder Unit
CS Detector Control System
VB EVent builder
VM Event manager
U Filter Unit
RL Front-end Readout Link
B FED Builder
ES Front End System
EC Front End Controller
ED Front End Driver **GTP**
Global Trigger Processor **LV1**
Regional Trigger Processors
PG Trigger Primitive Generator **TTC**
TIming, Trigger and Control **TTS**
Trigger Throttle System **RCMS**
Run Control and Monitor
TP Regional Trigger Processor
U Readout Unit

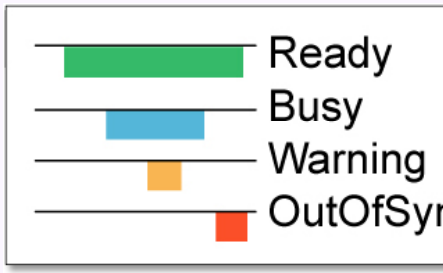


TTC (1 to N)

from GTP to FED/FES

distributed by an optical tree (~1000 leaves)

- LV1 20 MHz rep rate
- Reset, B cmd 1 Mhz Rep rate
- Controls data 1 Mbit/s



TTS (N to 1)

sTTS: from FED to GTP

collected by a FMM tree (~ 700 roots)

aTTS: From DAQ nodes to TTS

via service network (e.g. Ethernet) (~ 1000s nodes)

Transition in few BXs

FRL&DAQ networks (NxN)

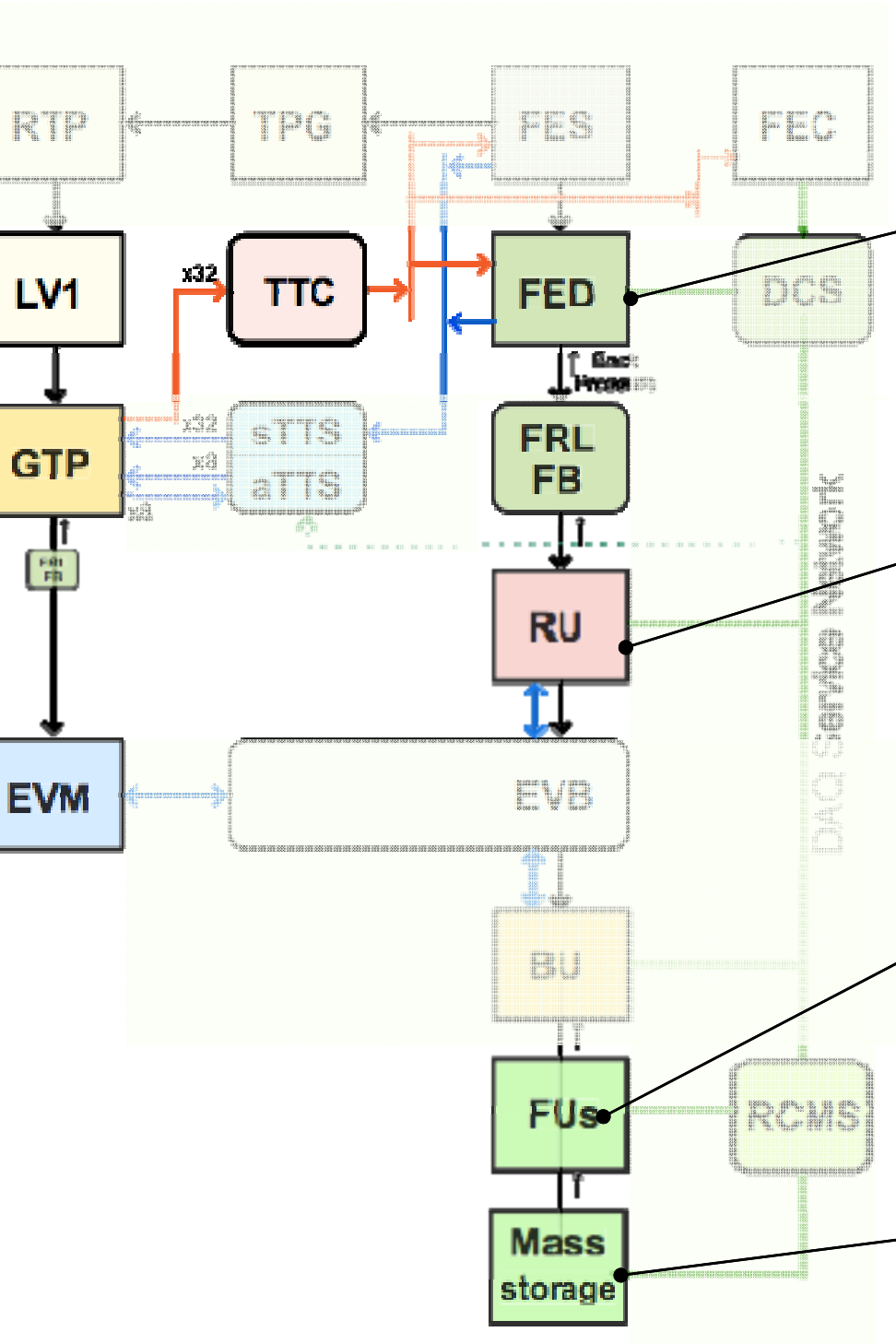
Data to Surface (D2S). FRL-FB

1000 Myrinet links and switches up to 2Tb/s

Readout Builders (RB). RUBU/FU

3000 ports GbEthernet switches 1Tb/s sustained

LHC-DAQ: event description



FED-FRL-RU

TTC rx

Event No. (LOCAL)
Orbit/BX (LOCAL)

FED data fragment

Event No. (LOCAL)
Orbit/BX (LOCAL)
FED data fragment

Event fragments

Readout Unit

from EVM

EVB token
BU destination

FED data fragment

Filter Unit

from GT data

Event No. (LOCAL)
Time (ABSolute)
Run Number
Event Type
HLT info

Full event data

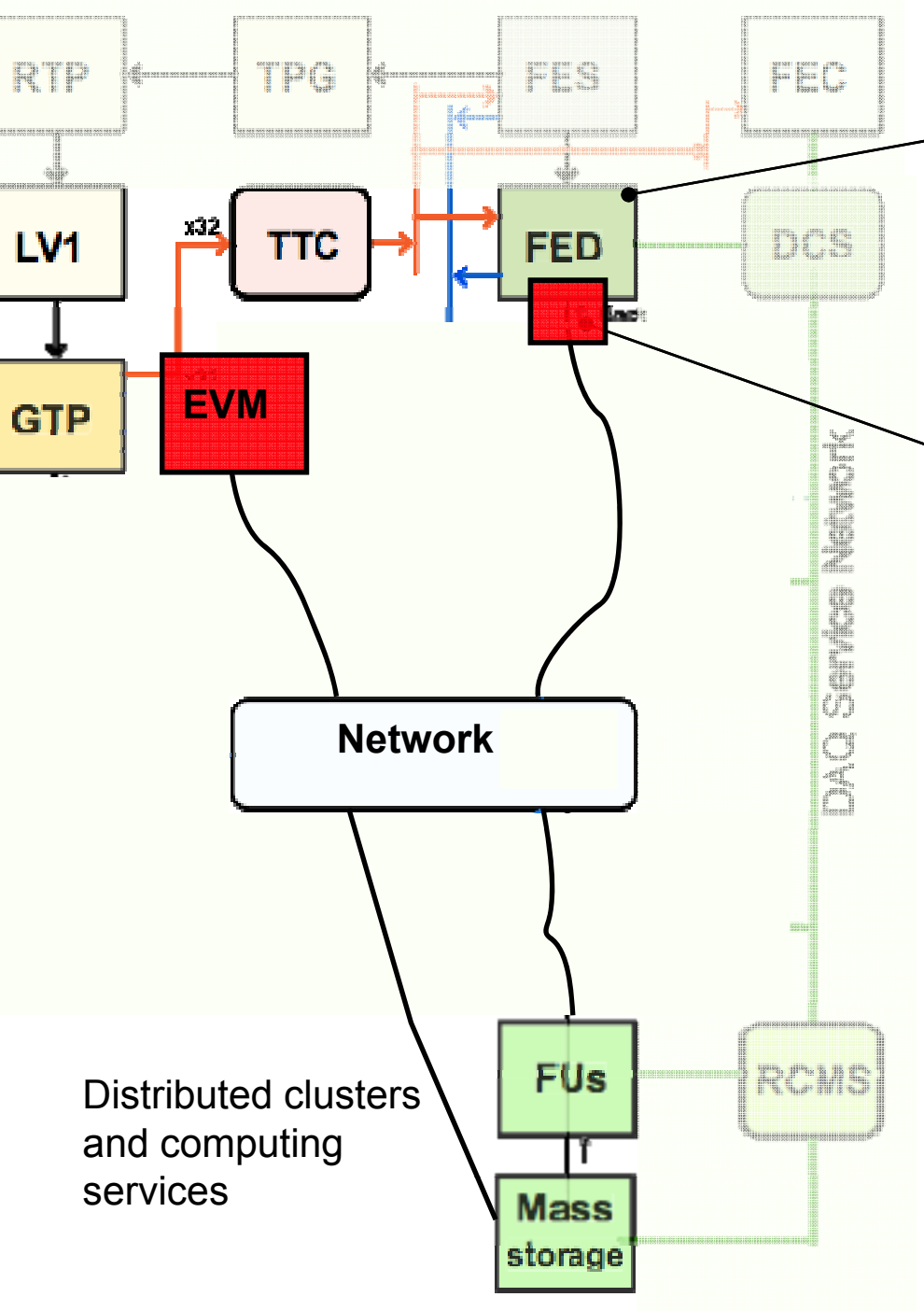
Built events

Storage Mar

Event #(Centr
Time (ABSolu
Run Number
Event Type
HLT info

Full Event Data

SLHC: event handling



Event #(Central)
Time (ABSolute)
Run Number
Event Type
HLT info

Full Event Data

- **TTC**: full event info in each event fragment distributed from the GTP-EVM via TTC

- **FED-FRL** interface e.g. PCI
- FED-FRL system able to handle **high level data communication protocols** (e.g. TCP/IP)
- FRL-DAQ **commercial interface**

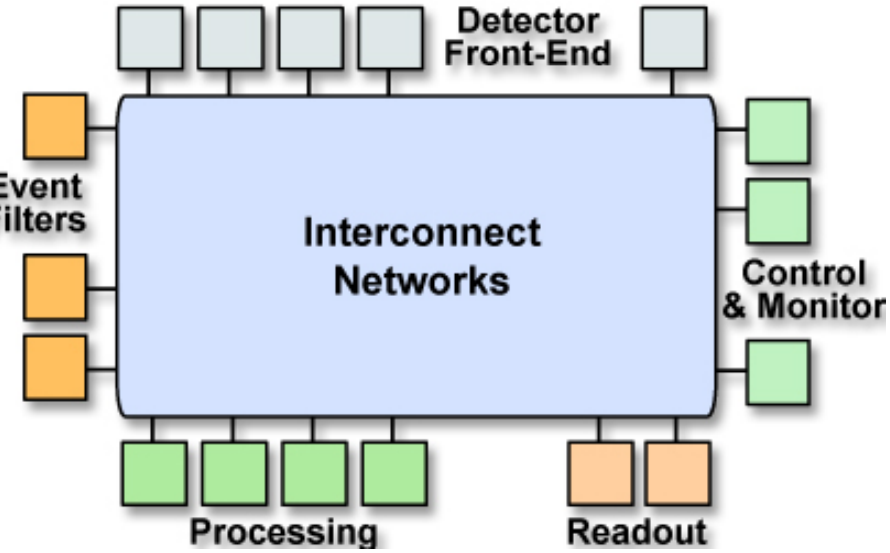
e.g. Ethernet (10 GBE or multi-GBE) to be used fo data readout and FED configuration, control, local DAQ etc.

New FRL design can be based on embedded PC (e.g. ARM like) or advanced FPGA etc.

The network interface can be used in place of VM backplane to configure and control the new FED.

Current FRL-RU systems can be used (with some limitations) for those detector not upgrading their FED design.

Another step toward uniform network architecture:



- All DAQ sub-systems (FED-FRL included) are interfaced to a single **multi-Terabit/s network**. The same network technology (e.g. ethernet) is used to **read and control** all DAQ nodes (FED,FRL, RU, BU, FU, eMS, DQM,...)

- In addition to the Level-1 Accept and to the other synchronous commands, the Trigger&Event Manager-TTC transmits to the FEDs the complete event description such as the **event number, event type, orbit, BX and the event destination address** that is the processing system (CPU, Cluster, TIER..) where the event has to be built and analyzed....
- The event fragment delivery and therefore the **event building will be warranted by the network protocols** and (commercial) network internal resources (buffers, multi-path, network processors, etc.). FED-FU push/pull protocols can be employed...
- Real time buffers of Pbytes temporary storage disks will cover a **real-time interval of days**, allowing to the event selection tasks a better exploitation of the available distributed processing power

ger synchronous data (20 MHz) (~ 100 bit/LV1)

k distribution & Level 1 accept
commands (resets, orbit0, re-sync, switch context etc...)
parameters (e.g. trigger/readout type, Mask, etc.)

nt synchronous data packet (~ 1 Gb/s, few thousands bits/LV1)

t Number	64 bits as before local counters will be used to sychronize and check.
/time and BX No.	64 bits
Number	64 bits
t type	256 bits
t destination(s)	256 bits e.g. list of IP indexes (main EVB + DQMs..)
t configuration	256 bits e.g. information for FU event builder

t fragment will contain the **full event description** etc.. Events can be built in Push or Pull mode. E.g
d 2. will allow many TTC operations be done independently per TTC partition (e.g. **local re-sych, reset, partial readout** etc..)

nchronous control information

guration parameters (delays, registers, contexts, etc.)
tenance commands, auto-test.

itional features? Readout control signal collector (a la TTS)?

replacement. Collect FED status (at LV1 rate from >1000 sources?)
ct local information (e.g. statistics etc.) on demand?
tenance commands, auto-test, statistics etc.

Summary

Q design

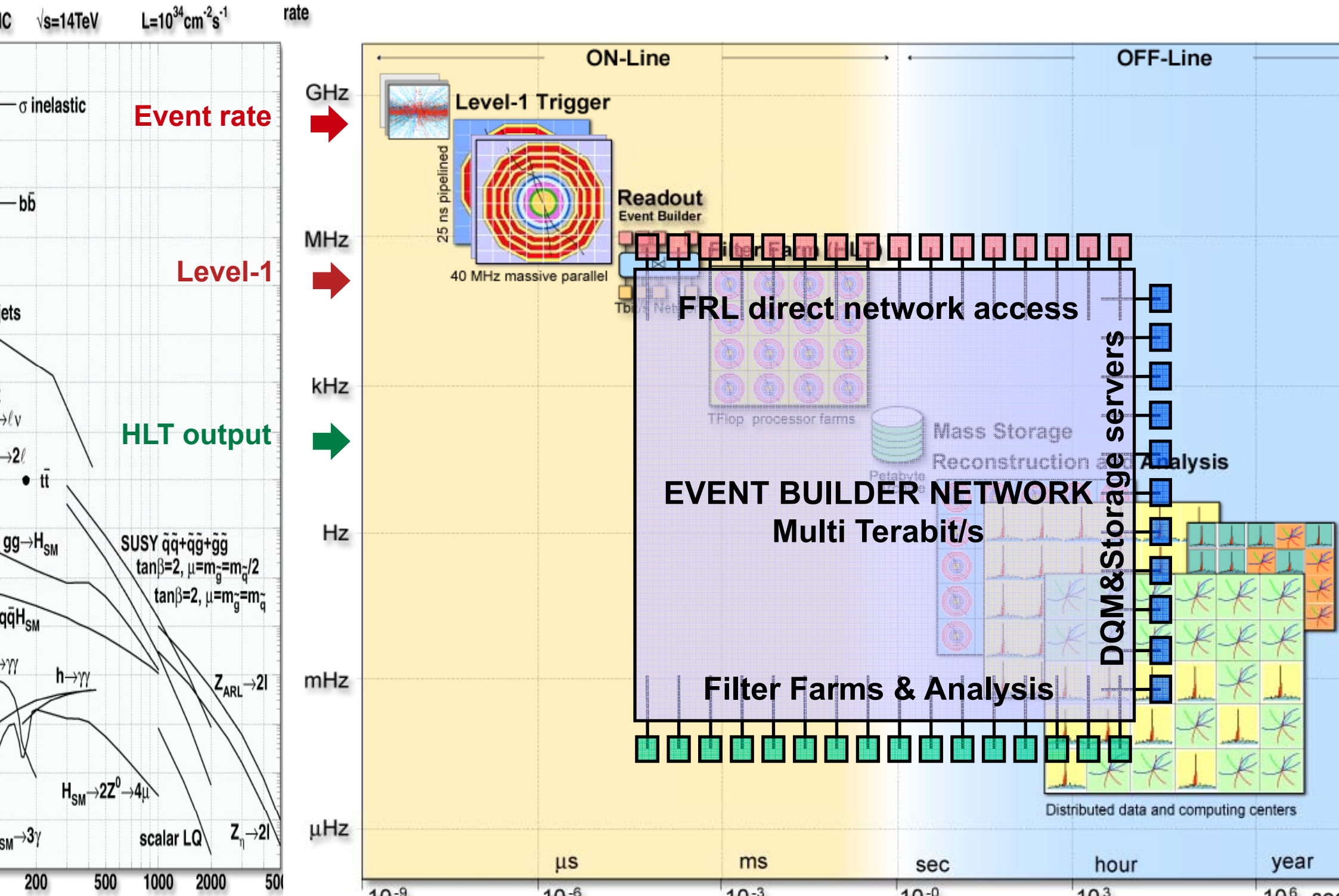
hitecture: will be upgraded enhancing scalability and flexibility and exploiting
naximum (by M+O) the commercial network and computing technologies.
ware: configuring, controlling and monitoring a large set of heterogeneous
distributed computing systems will continue to be a major issue

dware developments

eased performances and additional functionality's are required by the event
handling (use standard protocols, perform selective actions etc.)
new **timing, trigger and event info** distribution system
general front-end DAQ interface (for the new detector readout electronics)
dling high level network protocols via commercial network cards

best DAQ R&D programme for the next years is the
lementation of the 2003 DAQ-TDR

required a permanent contact between SLHC developers and LHC-DAQ
ementors to understand and define the new systems scope, requirements
specifications


√s=14TeV
L=10^34 cm^-2 s^-1
rate
σ inelastic
bb̄
tt̄
Event rate
Level-1
HLT output
gg→H_SM
SUSY q̃q̃+q̃g̃+g̃g̃
tanβ=2, μ=m_g̃=m_q̃/2
tanβ=2, μ=m_g̃=m_q̃
qq̄H_SM
h→γγ
Z_ARL→2l
H_SM→2Z^0→4μ
scalar LQ
Z_η→2l
200
500
1000
2000
GHz
MHz
kHz
Hz
mHz
μHz
ON-Line
OFF-Line
Level-1 Trigger
25 ns pipelined
40 MHz massive parallel
Readout
Event Builder
FRL direct network access
TFlop processor farms
Mass Storage
Reconstruction
Analysis
EVENT BUILDER NETWORK
Multi Terabit/s
Filter Farms & Analysis
DQM&Storage servers
Distributed data and computing centers
μs
ms
sec
hour
year