

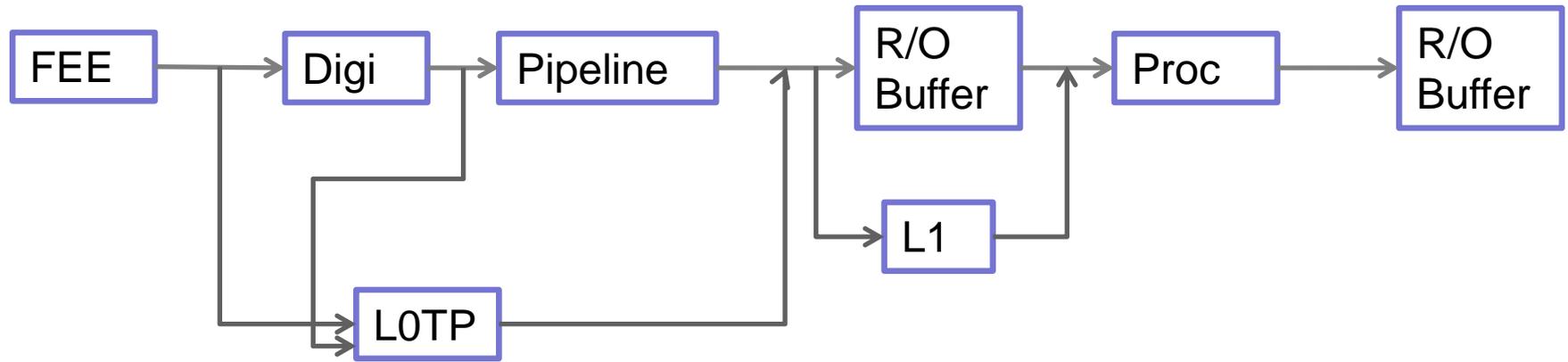
# “Proposta per un Readout Ibrido”

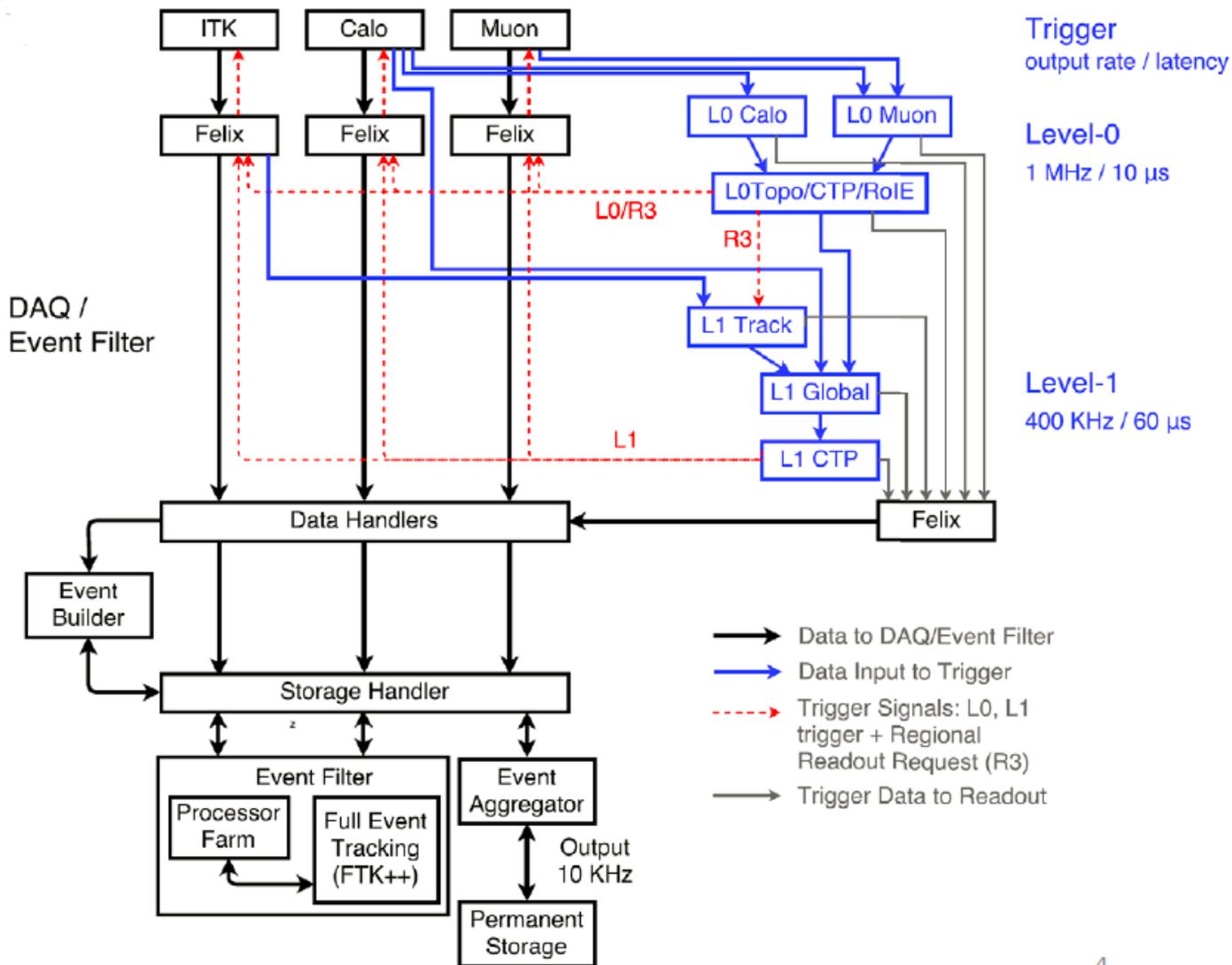
*PRIN Meeting NA62*

*Gianluca Lamanna (Univ. & INFN Pisa),*



# Un sistema di trigger standard

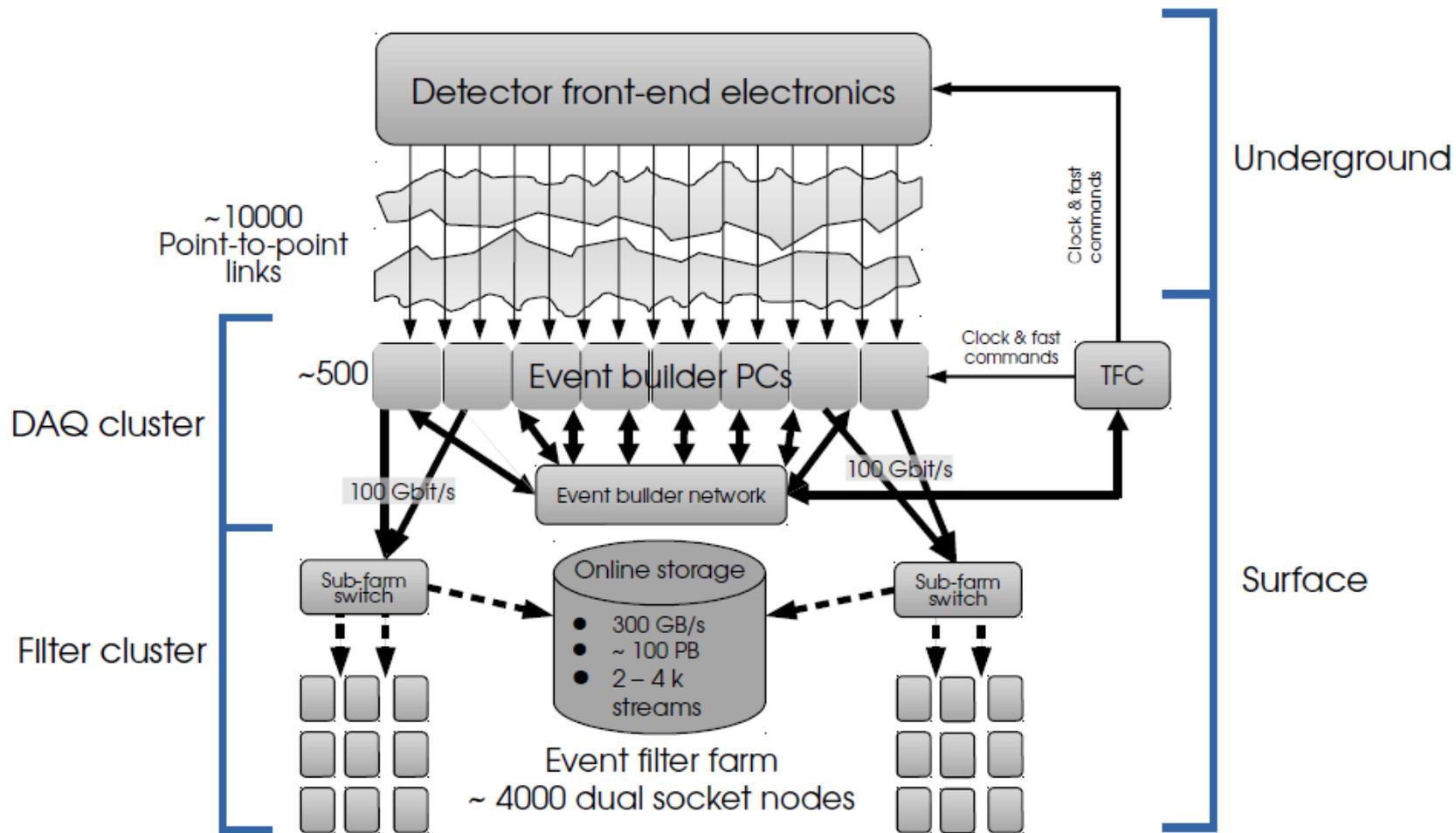




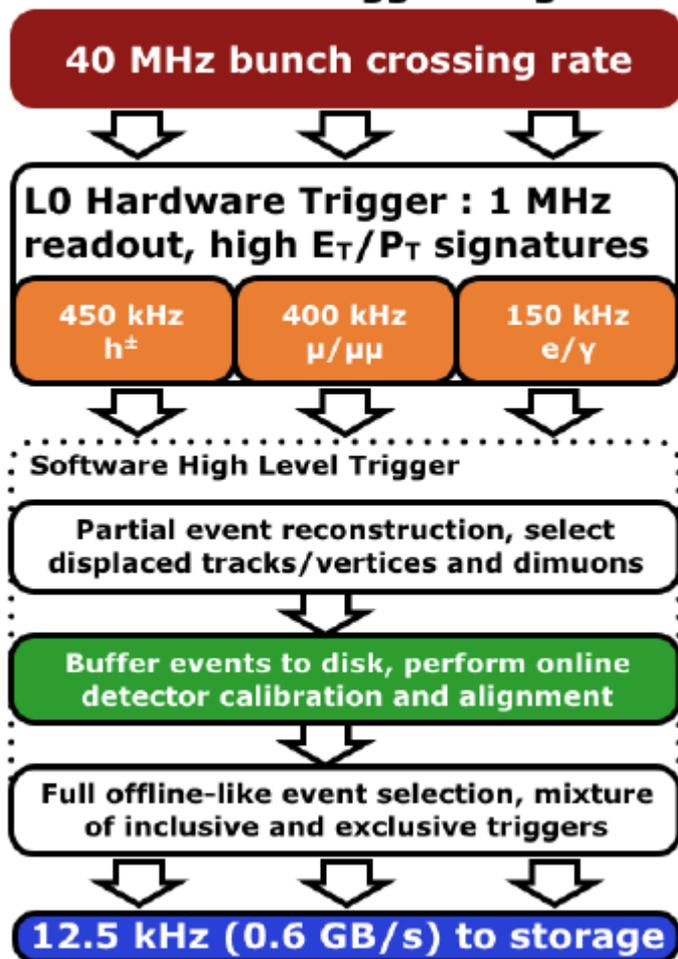
# Perchè non va bene per il futuro

- La sfida è (ovviamente) la selezione di eventi interessanti
  - Timing con il BX sempre fondamentale
  - Indicazioni da Calorimetri e muoni sempre necessarie
  - Essenziale usare anche i vertici
- Problemi
  - Elettronica o detector dedicati al trigger
  - Decisione basata su un subset
  - Trigger basato su quello che ci si aspetta
- Ma nel futuro non si potrà/dovrà:
  - Applicare soglie troppo alte: necessità di mantenere alta l'accettanza
  - Scalare il dataflow con solo la luminosità (energia, pileup, occupancy,...)
  - HW: piccola latenza → piu' parallelismo → più link → più materiale e costo
  - SW: Il tempo di processamento non scala linearmente con la luminosità, la complessità degli eventi diventa rapidamente dominante.

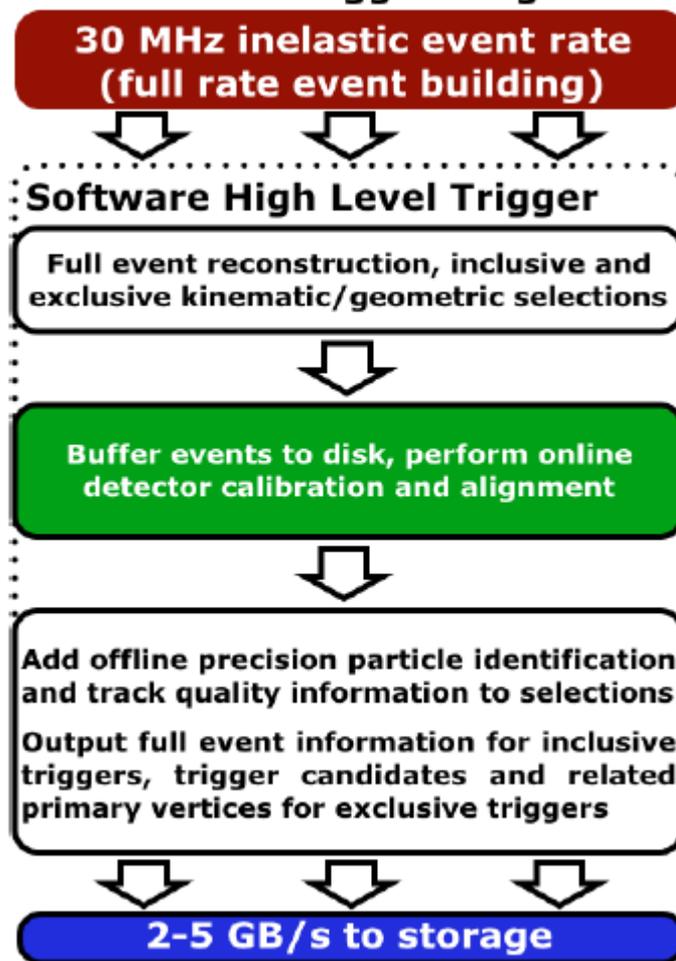
# Triggerless readout



## LHCb Run 2 Trigger Diagram



## LHCb Run 3 Trigger Diagram



# Perchè non va bene per il futuro

## Network:

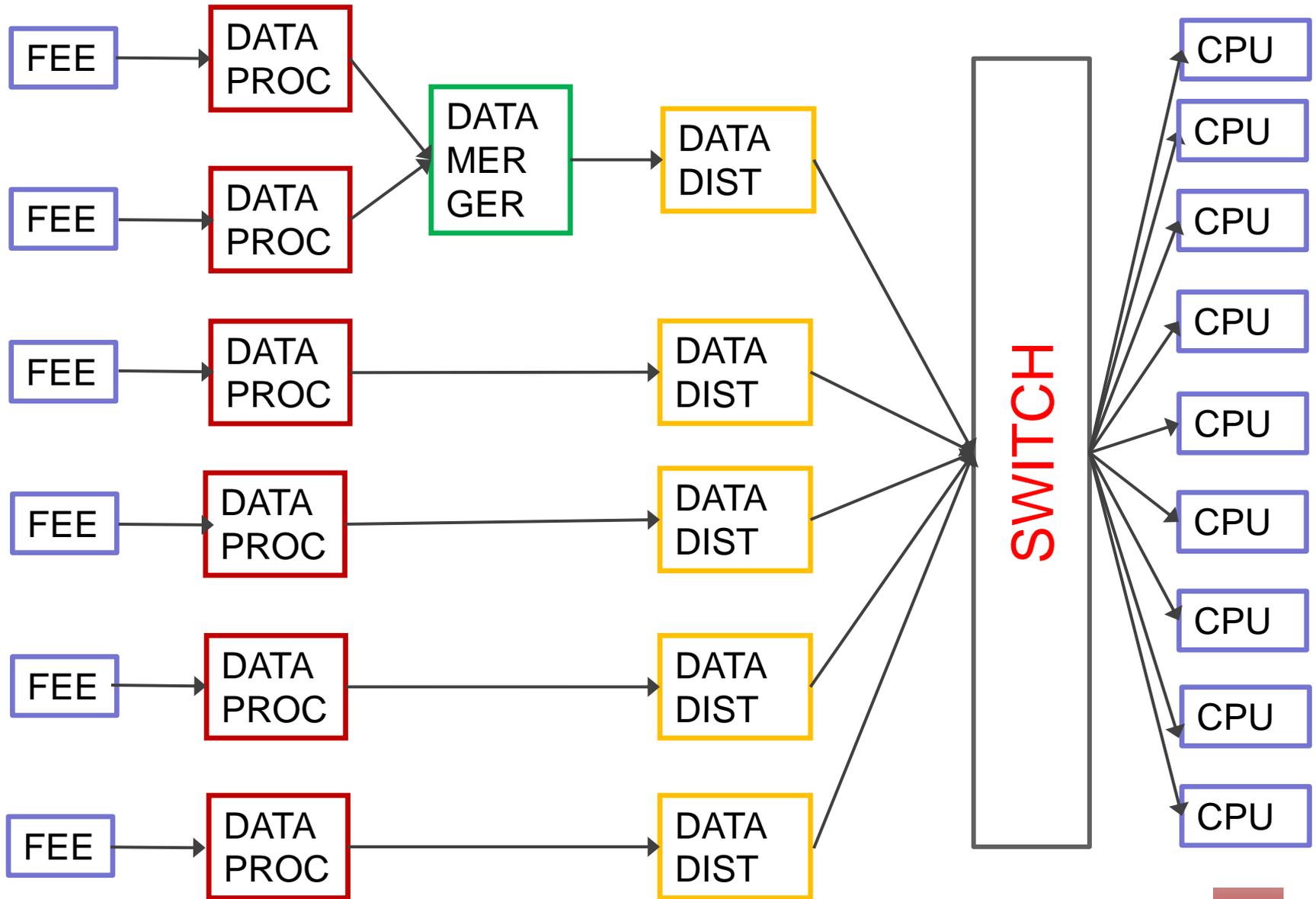
- Circa 10000 links GBT → 40 TB/s
- Readout con PCI40 card:
  - 100 Gb/s DMA
  - Readout max 48 links GBT a 4.48 Gb/s
  - Limite dal bus: circa 100 Gb/s (teorico 128 Gb/s)
  - Circa 20 link per PC → 500 PC di readout
- Switch readout network: 500 links da 100 Gb/s !!! (non così semplice... forse si' nel 2020)

## Throughput di calcolo

- Circa 4000 nodi
- Con pesanti assunzioni sullo scaling dal calcolo attuale:
- La derivata delle performance dei single core potrebbe diminuire
  - miglioramenti sono dovuti essenzialmente ai processi costruttivi e ai miglioramenti dell'architettura
  - Limitazioni dovute essenzialmente alla potenza alla latenza delle memorie e alla velocità dell'I/O
- Possiamo assumere che i computer miglioreranno di un fattore x1000 nel 2020? → difficile!
  - Processamento parallelo
  - DRAM nel dye
  - Dischi a stato solido

- I dati non sanno “dove andare”: fat-tree topology
- Serve un supervisor intelligente (Event Manager) per decidere quale readout PC deve buildare un evento
- I PC di readout sono differenti dai PC di calcolo:
  - Il numero e le caratteristiche dei PC di readout sono decisi dalle limitazioni dei link, il numero e le caratteristiche dei PC di filtering sono decisi dalle necessità computazionali
- Il sistema è scalabile: ma lo scaling è costosissimo (difficile da prevedere)

# Streaming Readout Concept



## ■ DATA PROC:

- Pre-processamento dei dati: compressione, formattazione, timestamping, timeframing, channel ID, ...
- Elaborazione più raffinata: “feature extraction”, ToT, ampiezza, carica totale → possibile riduzione di banda
- Stream solo di dati validi (zero suppression)
- Link veloci: GBT, Apenet, etc etc

## ■ DATA MERGER:

- Unisce due sorgenti omologhe o logicamente connesse

## ■ DATA DIST:

- Instrada il frammento al PC corretto
- Gateway tra link interno (GBT, ...) e rete di calcolo (Infiniband, Ethernet,...)

## ■ PC:

- Raccoglie I dati inerenti ad una timewindow
- Analizza: accetta e salva, rifiuta e cancella