



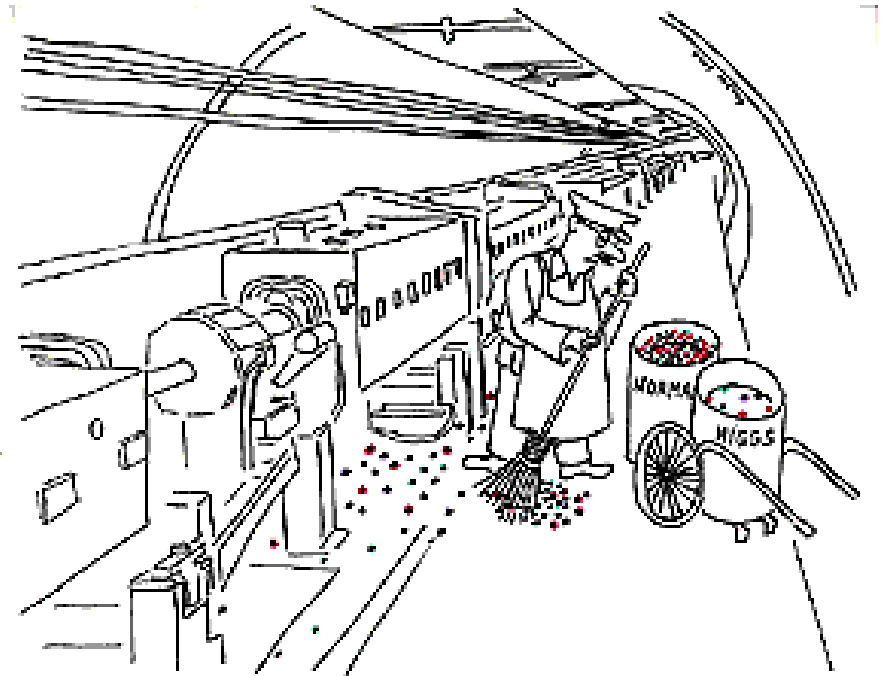
Acquisition de données



J.P. Cachemiche
Centre de Physique des Particules de Marseille

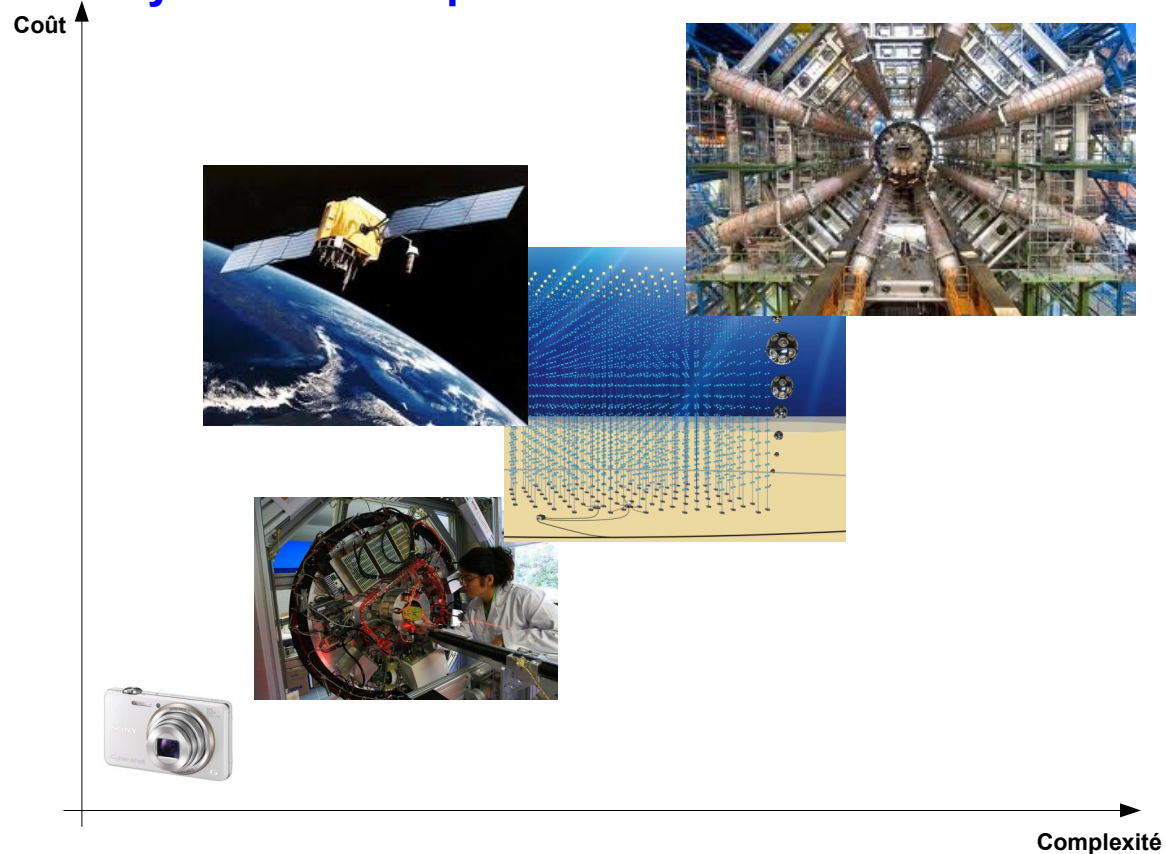
Plan

- Introduction
- Concepts de base
- Front-end
- Readout
- Trigger
- Distribution temporelle
- Contrôle et monitoring



Systemes d'acquisition

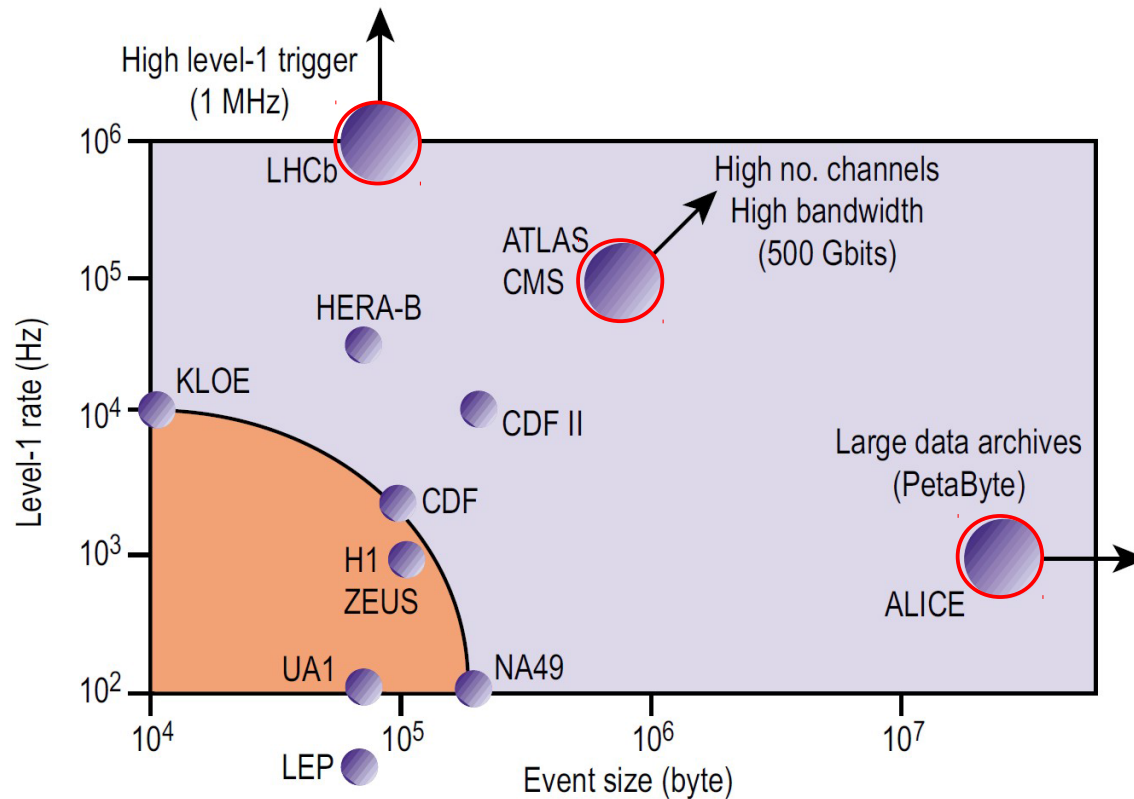
Variabilité des systemes d'acquisition



- Leur structure peut varier considérablement en fonction des performances recherchées, de l'ampleur de la mesure et des contraintes d'exploitation

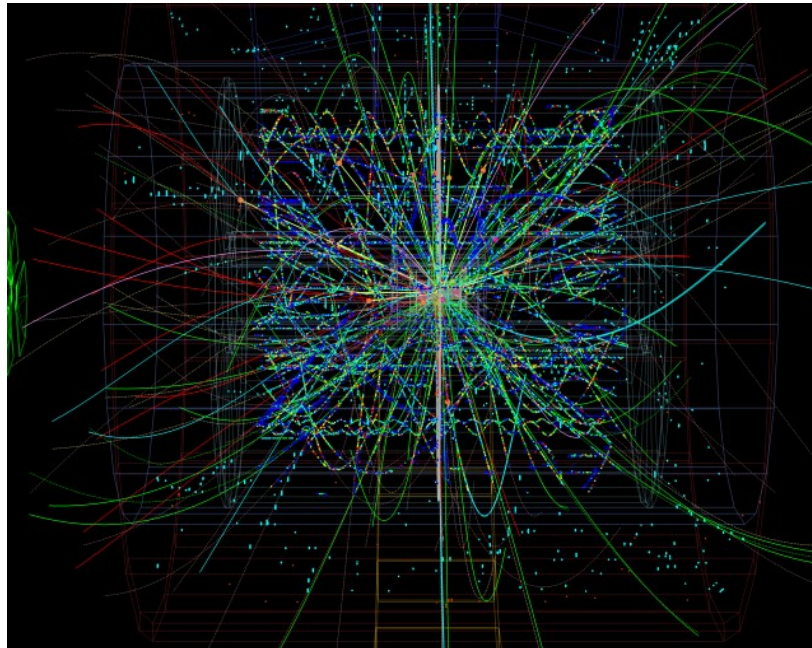
Systemes d'acquisition

Dans des domaines d'applications similaires, les contraintes peuvent être très différentes



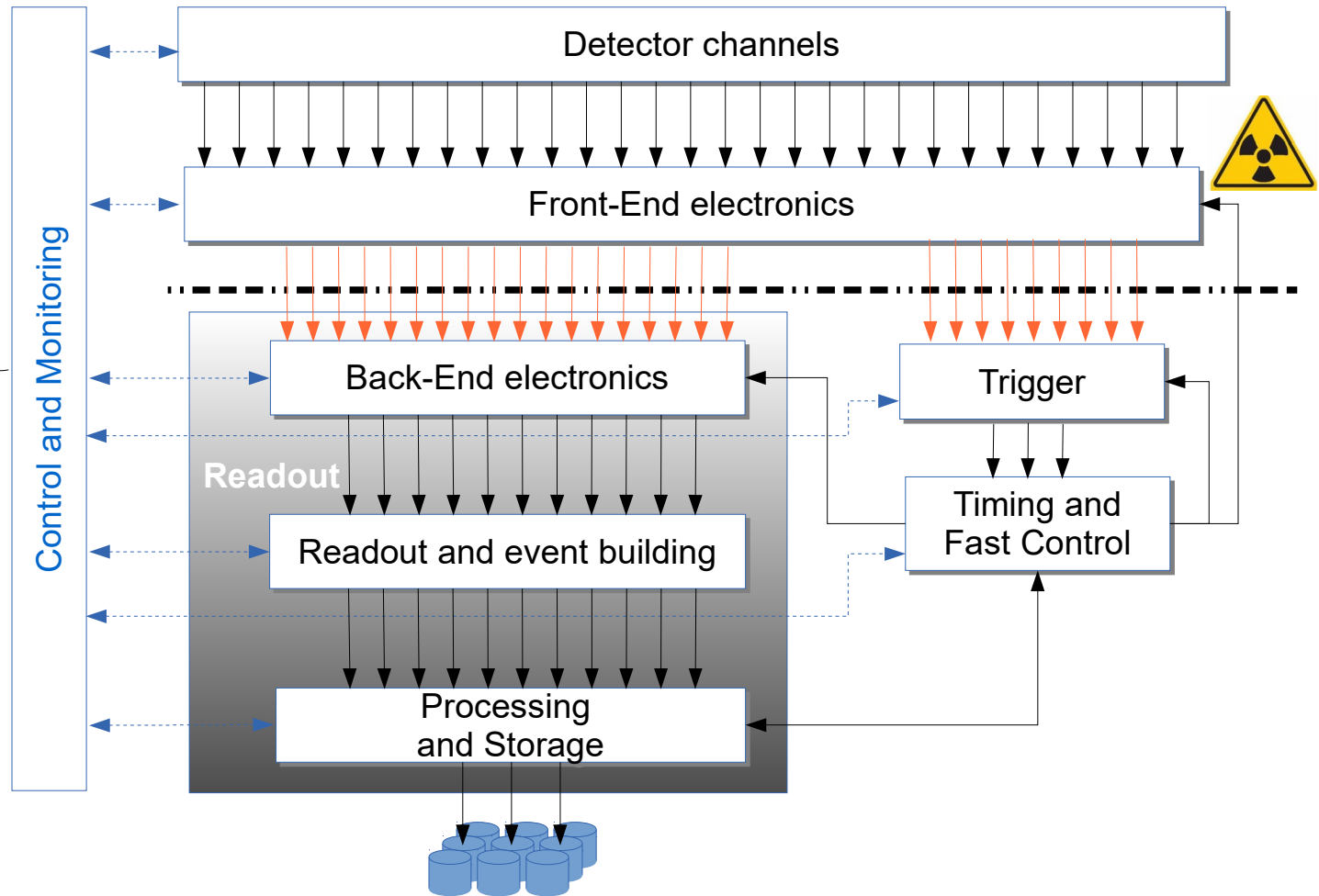
Systemes d'acquisition

- On va expliquer les problematiques de quelques uns des plus plus puissants systemes d'acquisition actuels : les DAQ des experiences ATLAS, CMS, ALICE et LHCb du Large Hadron Collider au CERN ;
- Les architectures decrites sont donc specifiques des contraintes rencontrees dans ces systemes mais peuvent etre rencontrees sur tout type d'acquisition.

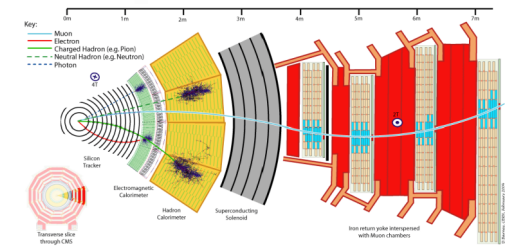
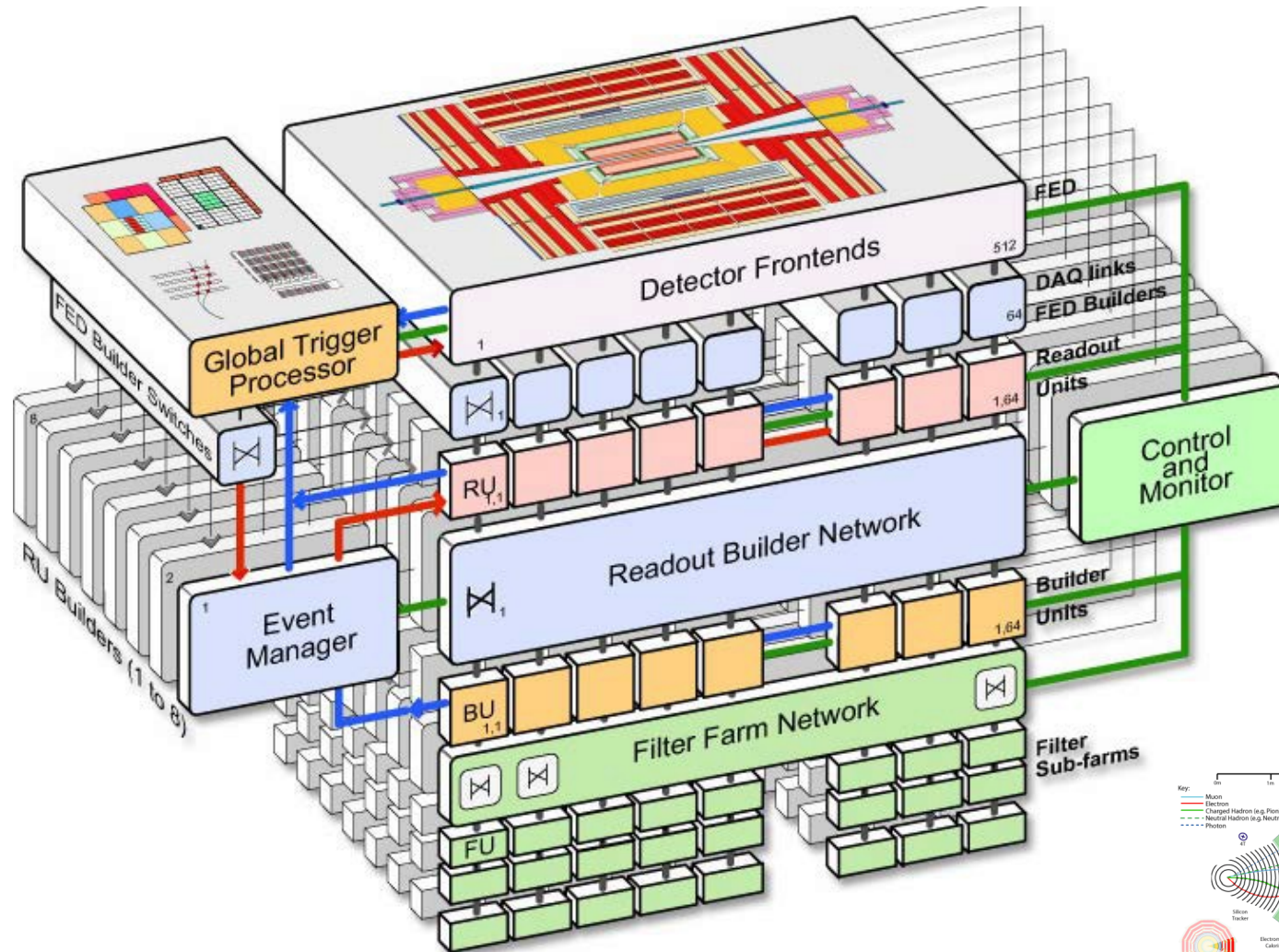


Concepts de base

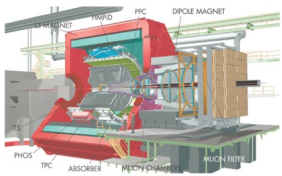
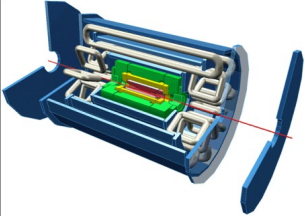
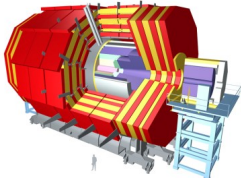
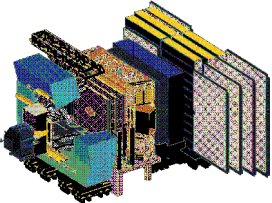
Éléments d'un système d'acquisition



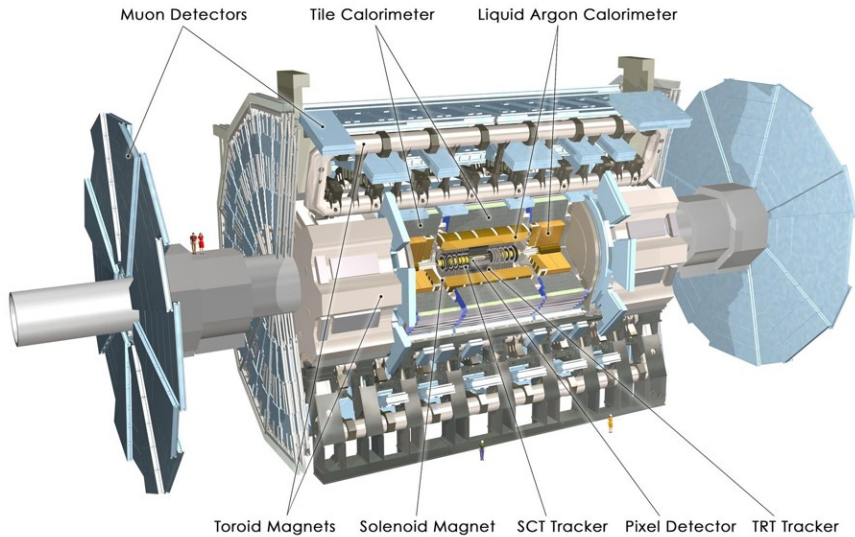
Exemple : le système d'acquisition de CMS



Caractéristiques des détecteurs LHC

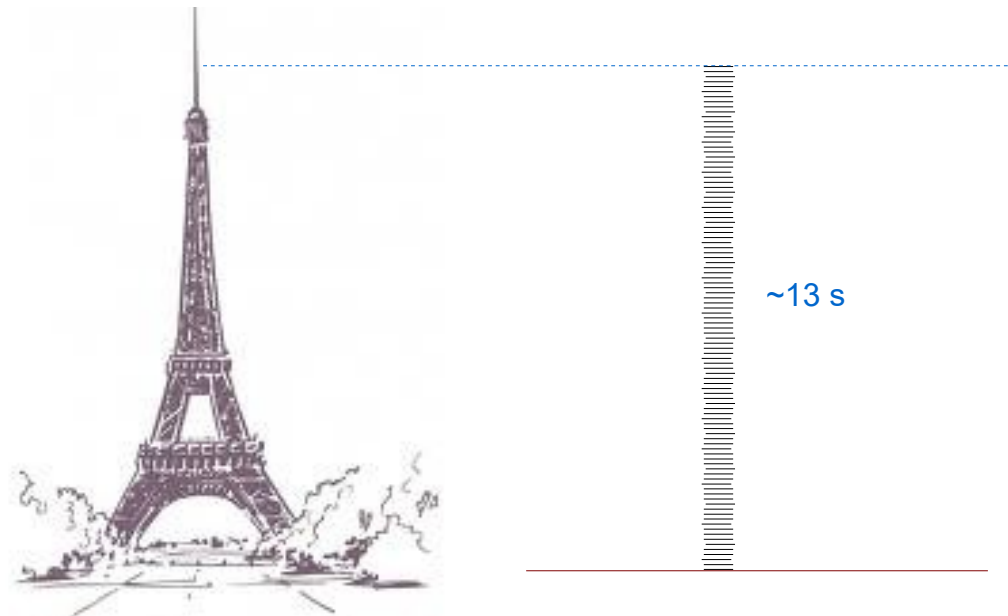
		Nombre de canaux	Taux de collisions	Données générées par un Bunch Crossing
ALICE		~ 60 millions	PbPb : 125 μ s p-p : 1.5 μ s	Pb-Pb : 86 MB p-p : 2.5 MB
ATLAS		~100 millions	p-p : 25 ns	p-p : 1.6 MB
CMS		~70 millions	p-p : 25 ns	p-p : 1 MB
LHCb		~1 million	p-p : 25 ns	p-p : 75 kB

Le challenge

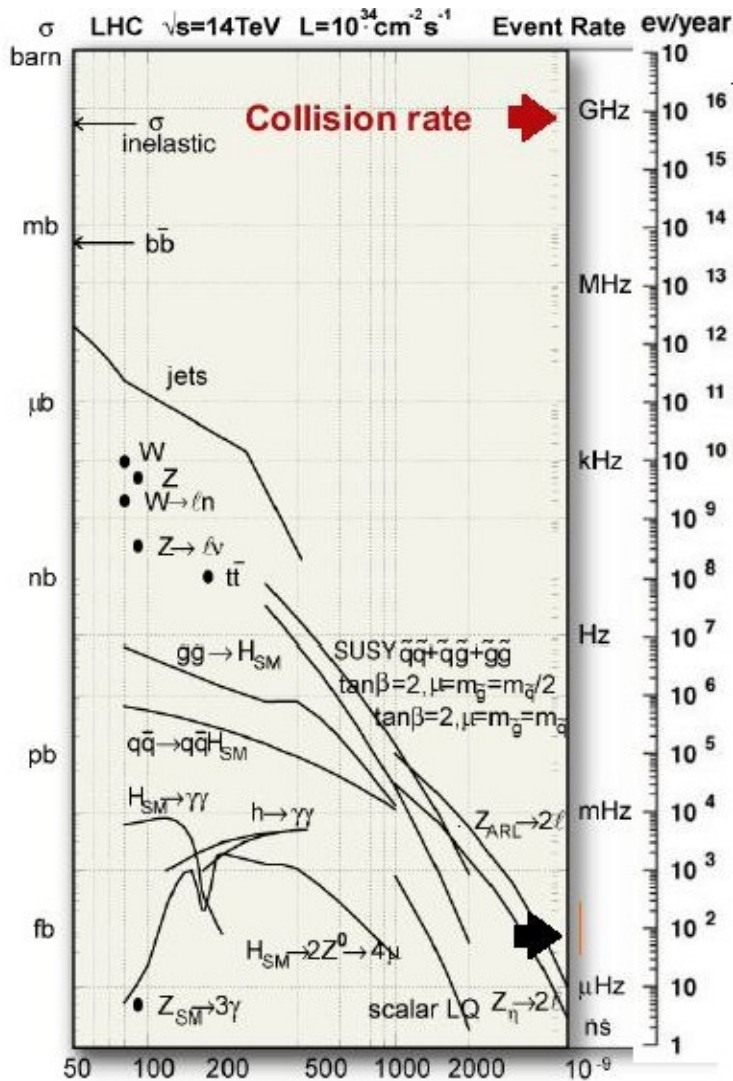


Peut-on stocker tous les événements ?

- Exemple détecteur Atlas
 - Collisions proton-proton :
taille d'un événement = 1.6 Mbytes
 - $1.6 \cdot 10^6 * 40 \cdot 10^6 = 64 \text{ Tbytes par seconde}$
 - ~16 000 DVD par seconde



Réduction du nombre d'événements

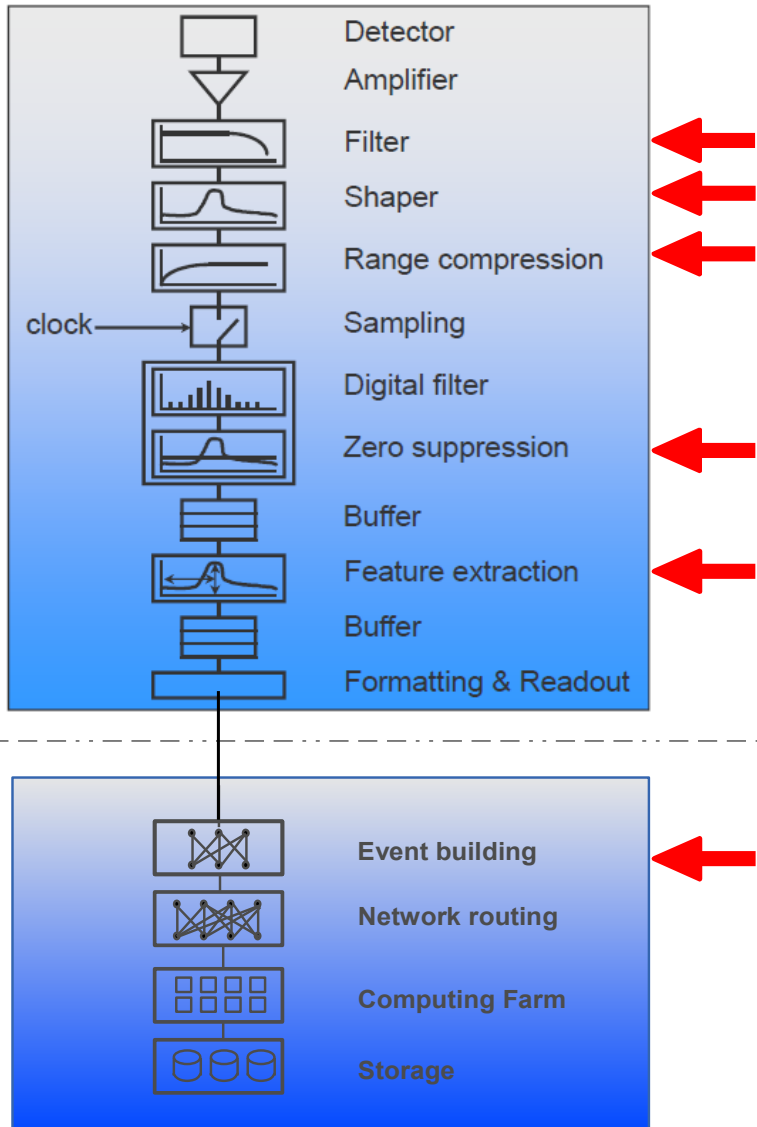


La plupart des collisions sont non pertinentes !

- Quelques chiffres pour le LHC
 - Bunch crossings : $4 \cdot 10^7$ Hz
 - Collisions de protons : 10^9 Hz
 - Production de nouvelles particules : 10^5 Hz (quelques milliers par an)
 - 1 événement intéressant sur 10^{14} collisions !

- On filtre les événements inutiles
 - **Rôle du trigger et de l'event filter**

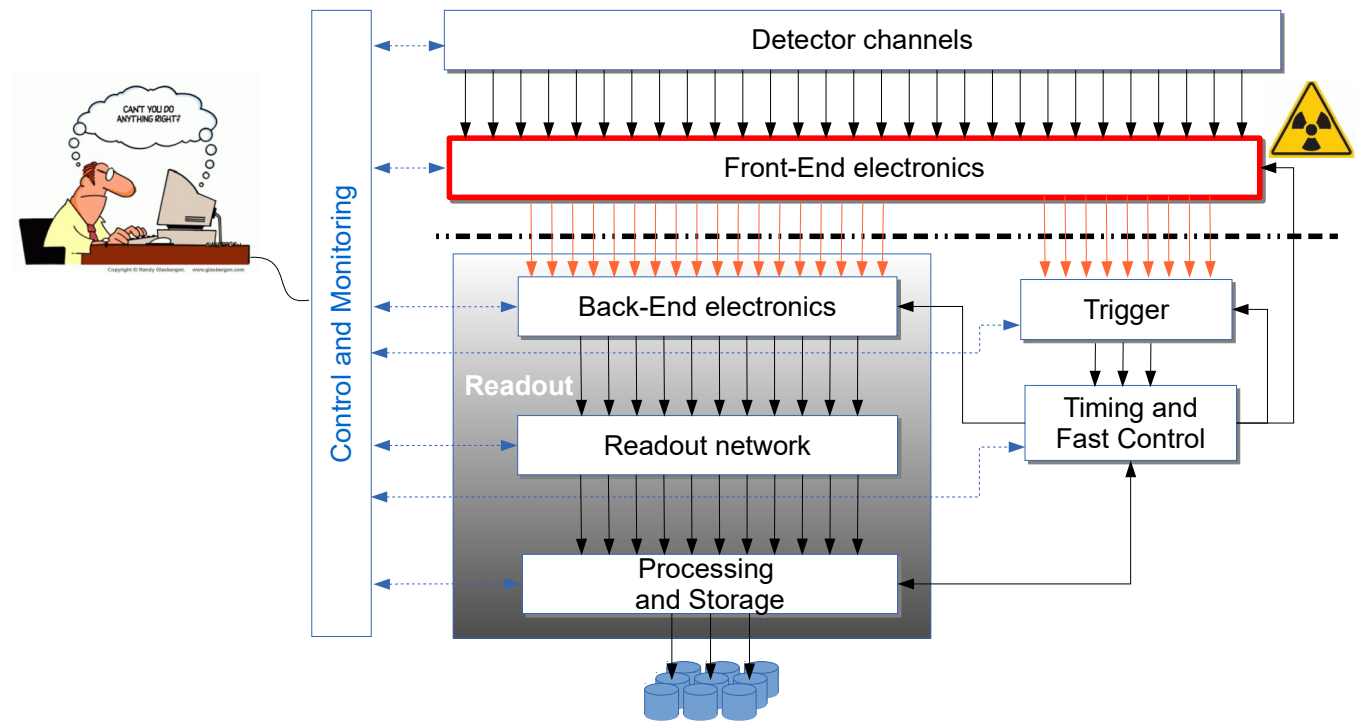
Réduction du volume des événements



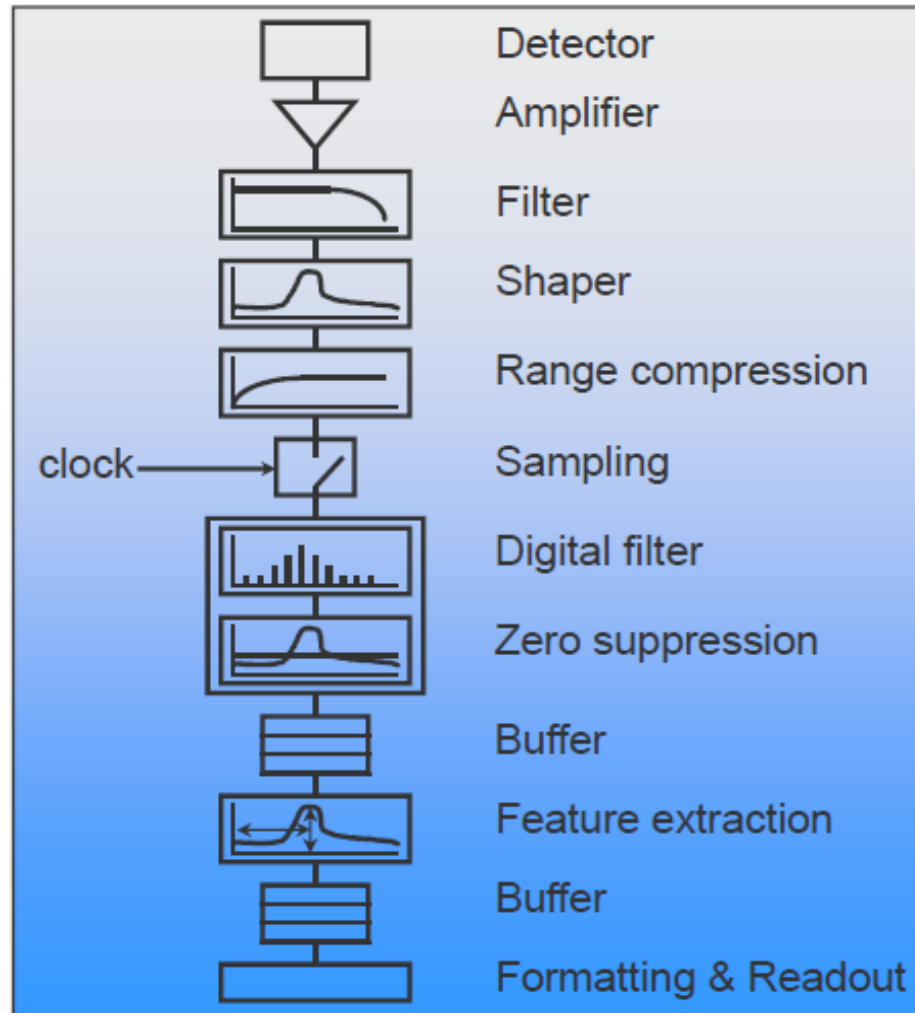
Les données sont compressées tout au long de la chaîne :

- Au niveau des Front-Ends
 - Compression analogique avec pertes
 - Au niveau des Back-Ends
 - Compression numérique sans pertes selon algorithmes dépendant du type de données
 - Réduction algorithmique et stockage
- **Rôle du Readout**

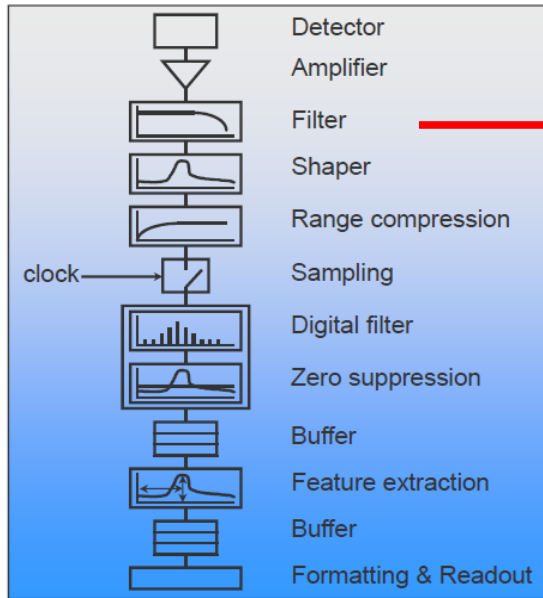
Electronique Front-End



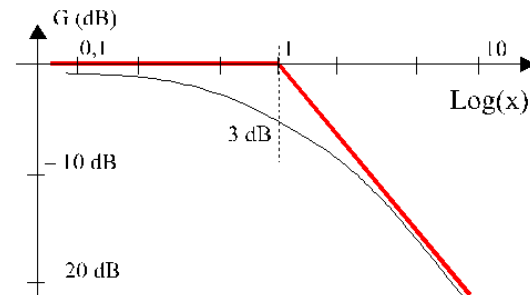
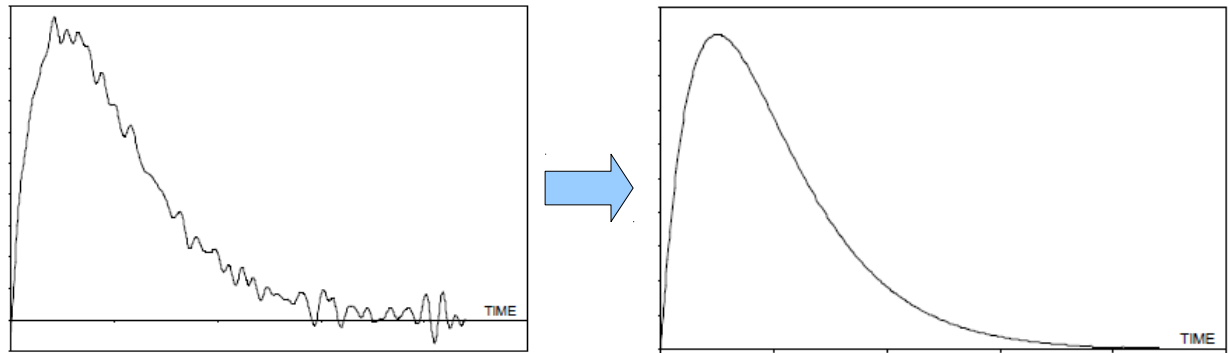
Canal d'acquisition typique



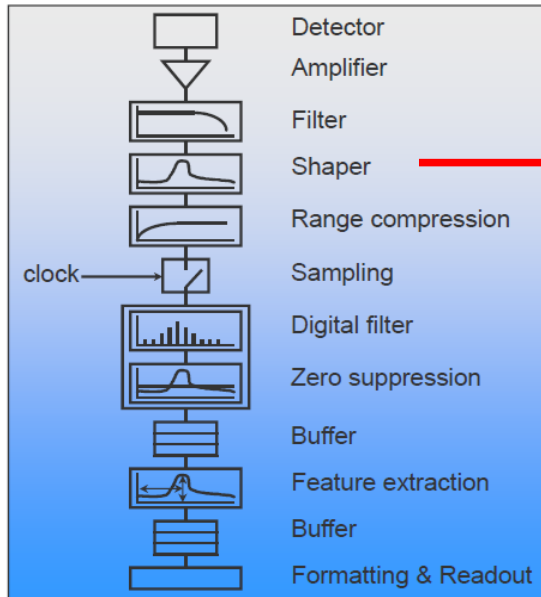
Détails d'un canal d'acquisition



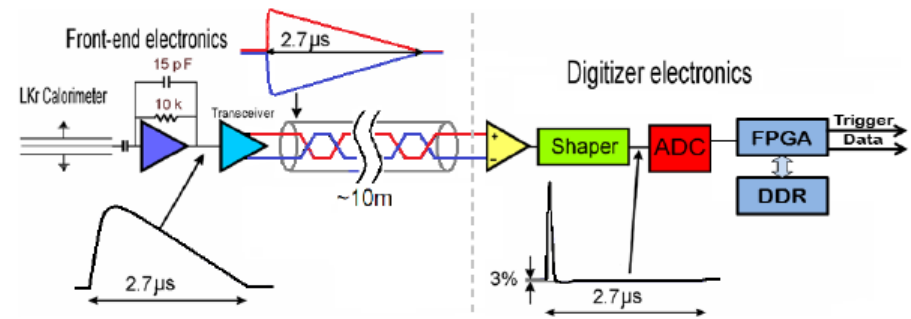
• Optimisation du rapport signal sur bruit par filtrage



Détails d'un canal d'acquisition

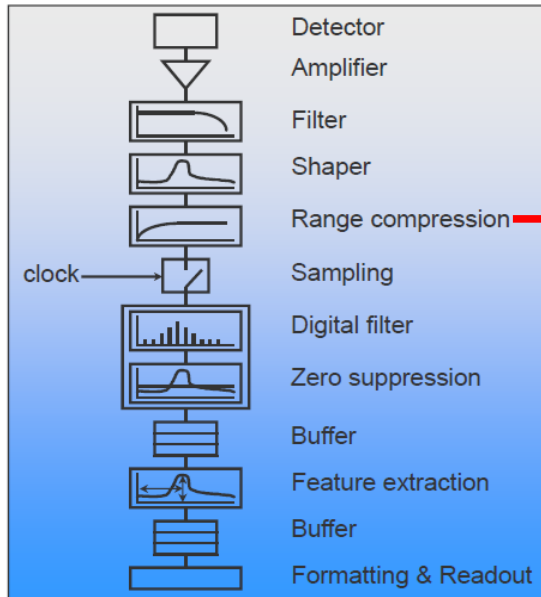


- Le « Shaper » optimise la forme du signal de façon à mieux observer :
 - le signal minimum qu'on souhaite détecter
 - l'amplitude du signal
 - l'instant d'arrivée
 - parfois une combinaison des précédents
- Peut varier suivant ce qu'on cherche à observer :
 - présence ou non (hit)
 - énergie
 - temps

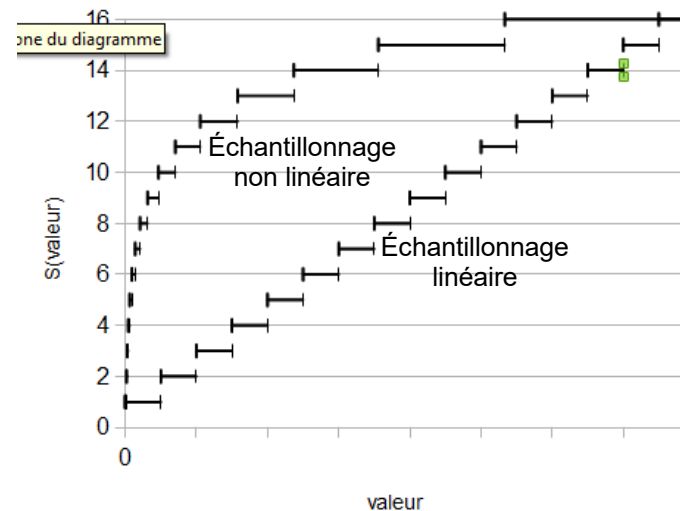


Lkr readout chain – NA52 experiment

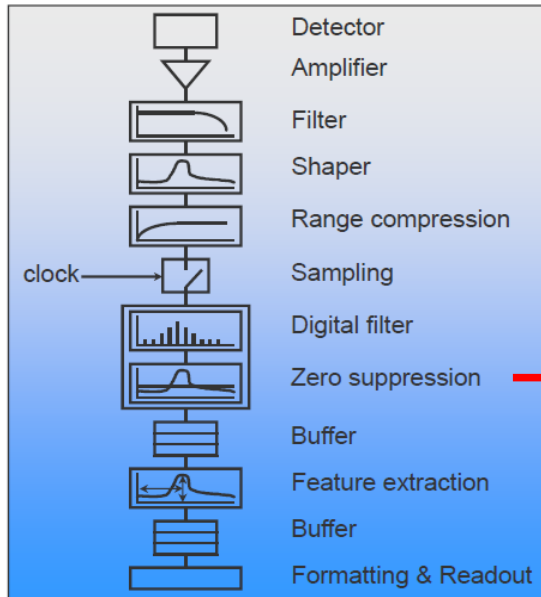
Détails d'un canal d'acquisition



- L'échantillonnage avec « range compression » :
 - L'amplitude du pas de quantification n'est pas la même pour les grandes valeurs du signal que pour les faibles valeurs
 - Permet d'avoir une meilleure résolution sur les faibles amplitudes sans augmenter le nombre de bits du mot encodé

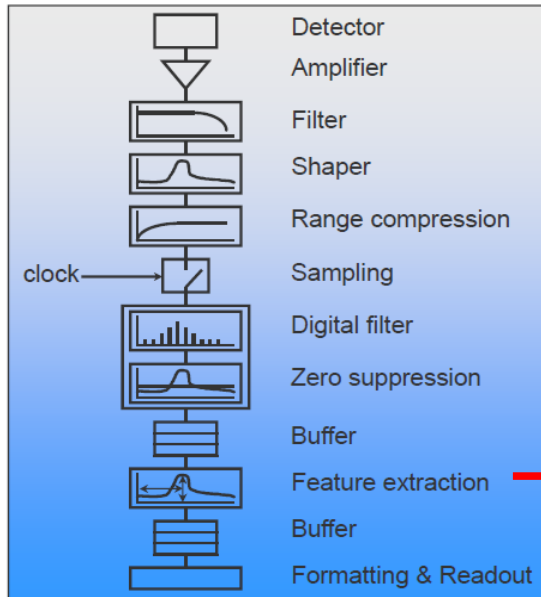


Détails d'un canal d'acquisition

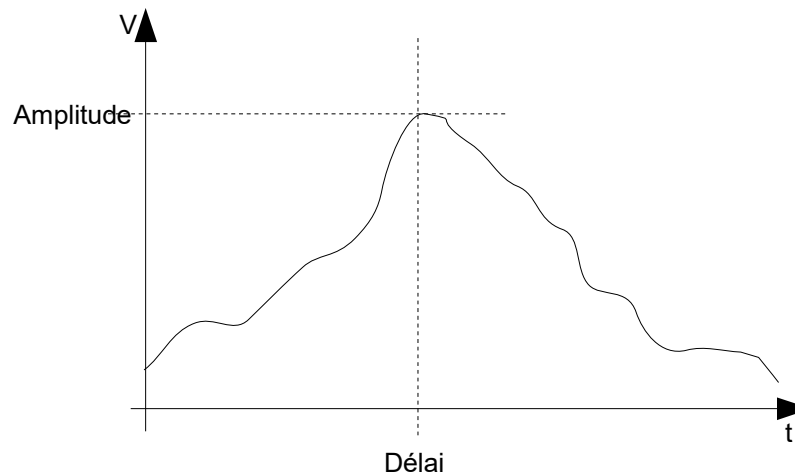


- La « suppression de zéro » consiste à maximiser le nombre de valeurs nulles du signal pour le compresser plus facilement.
- Au sens large toute opération de compression

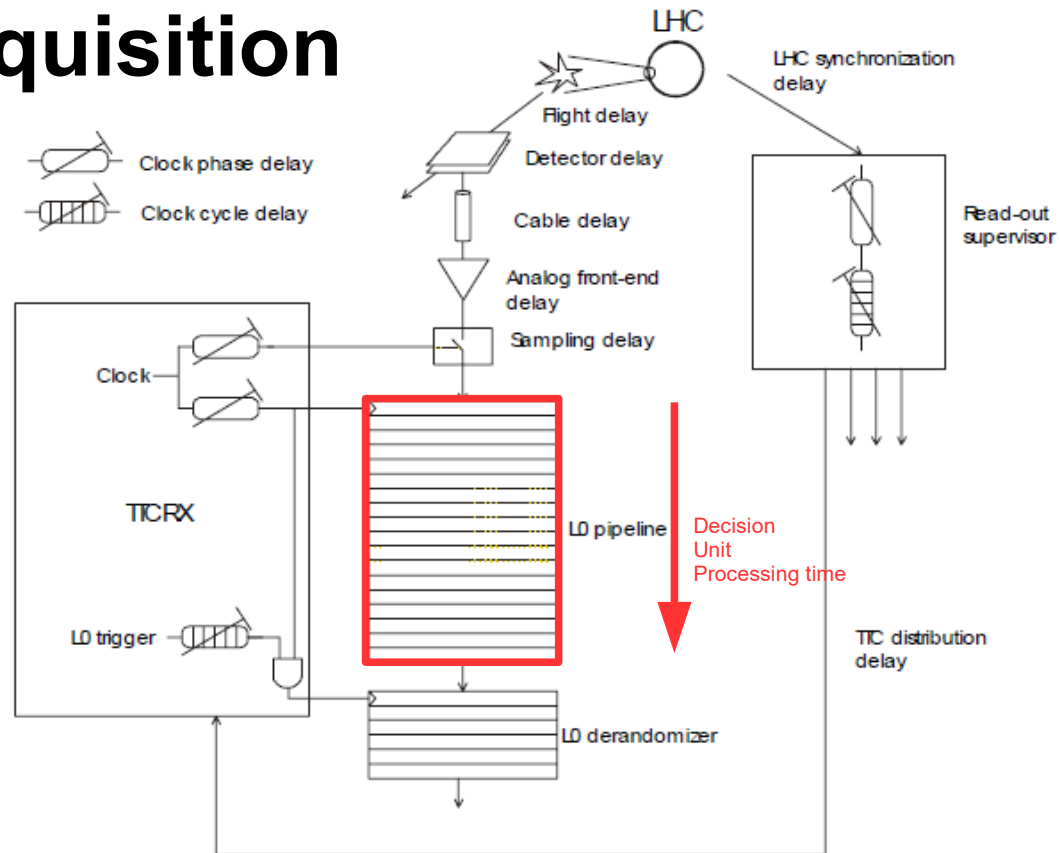
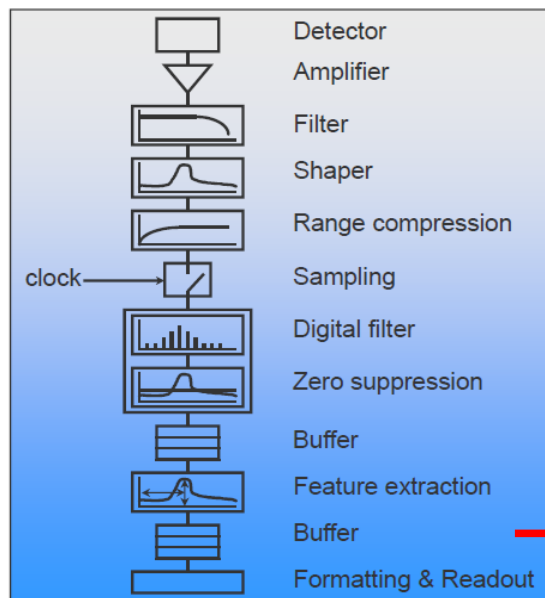
Détails d'un canal d'acquisition



- Exemple de « Feature extraction »
 - Réponse transformée en 2 paramètres simples : délai et amplitude

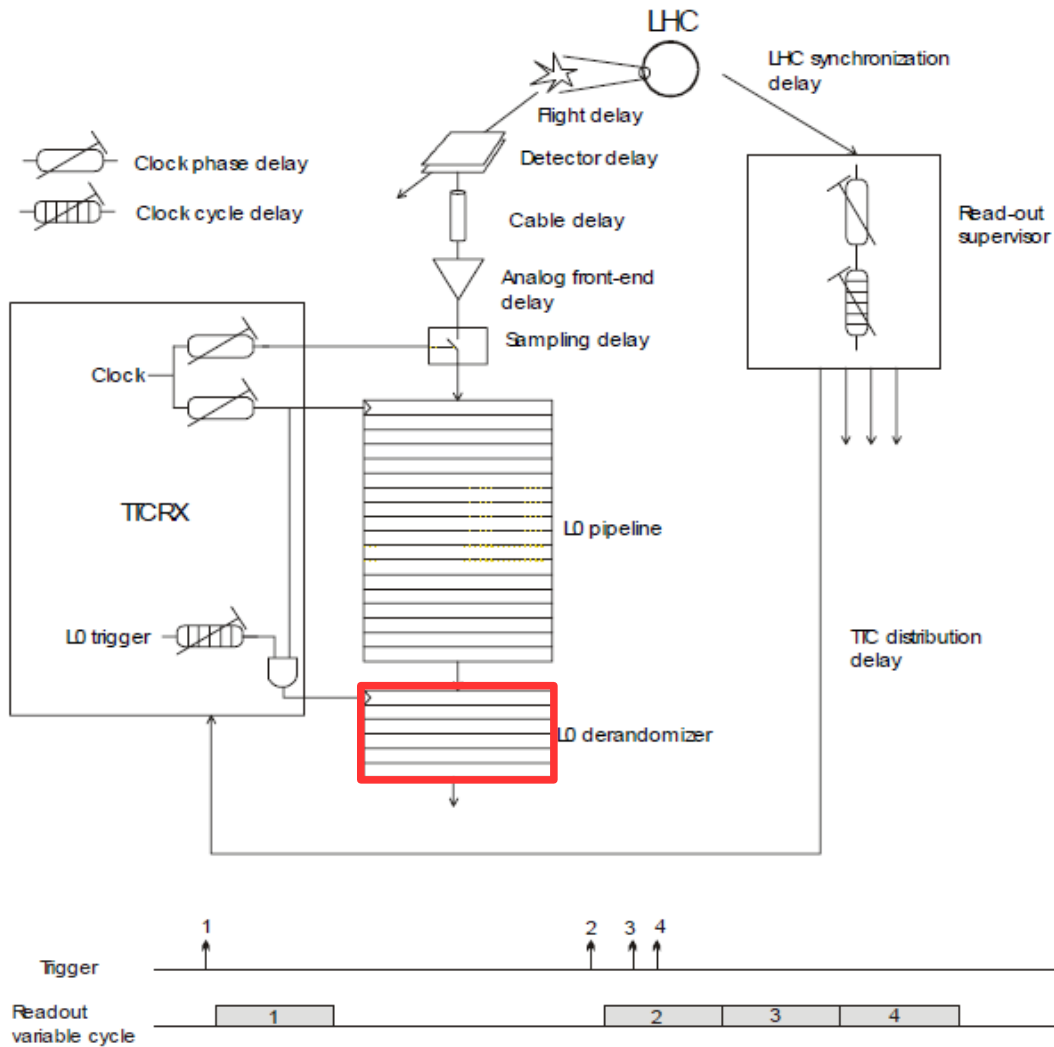


Détails d'un canal d'acquisition



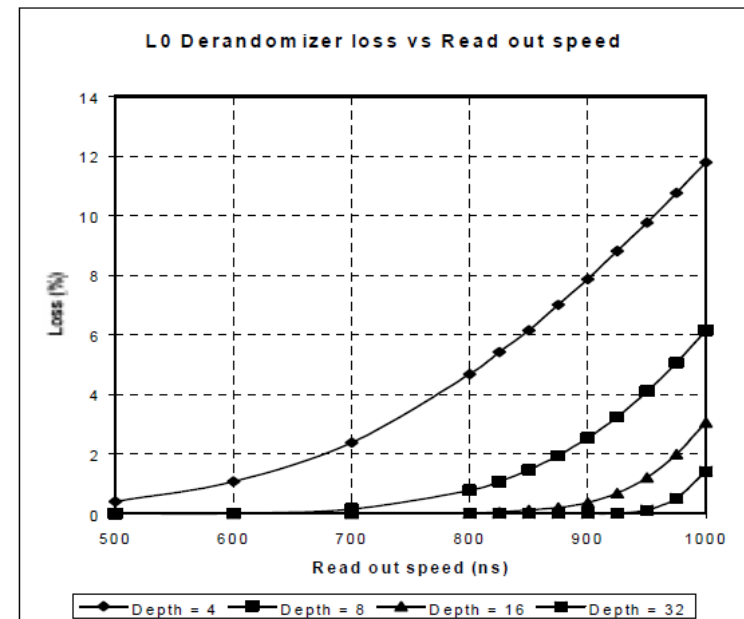
- **Bufferisation** nécessaire en attente de la décision du trigger
- A la fin du buffer les données sont soit gardées soit jetées en fonction du résultat de l'unité de décision
→ Signal L0 Accept

Derandomizer



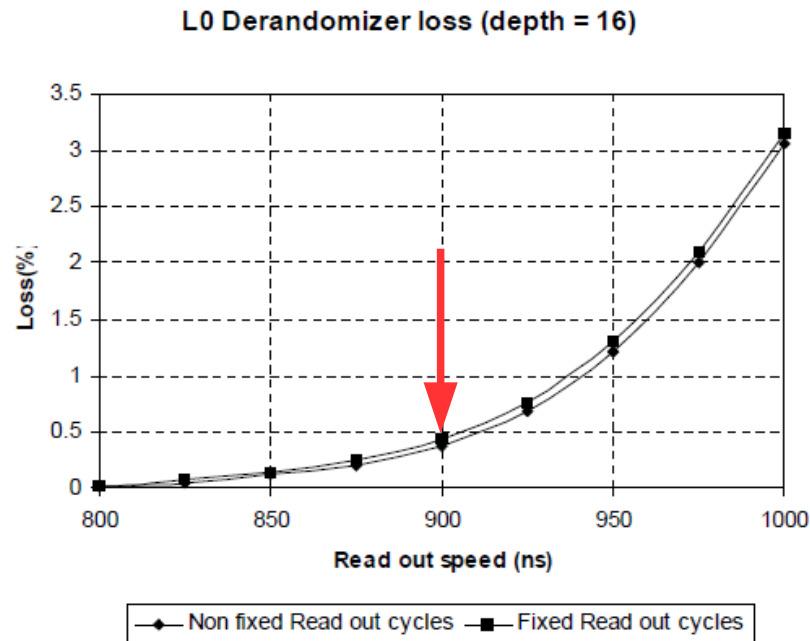
Permet de stocker les événements en attendant qu'ils soient transmis sur le canal de transmission

Doit être dimensionné de façon à ne pas perdre d'événements
 → Optimisé par simulation



Perte d'événements

Il est acceptable (mais pas souhaitable) de perdre des événements !

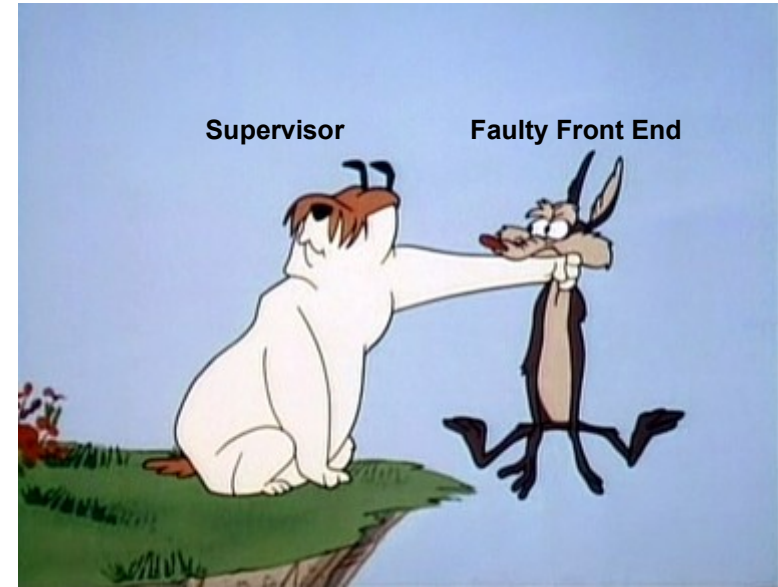


Compromis au niveau de la profondeur du derandomizer car la taille des mémoires doit être faible dans les front-ends

→ Problèmes de SEU

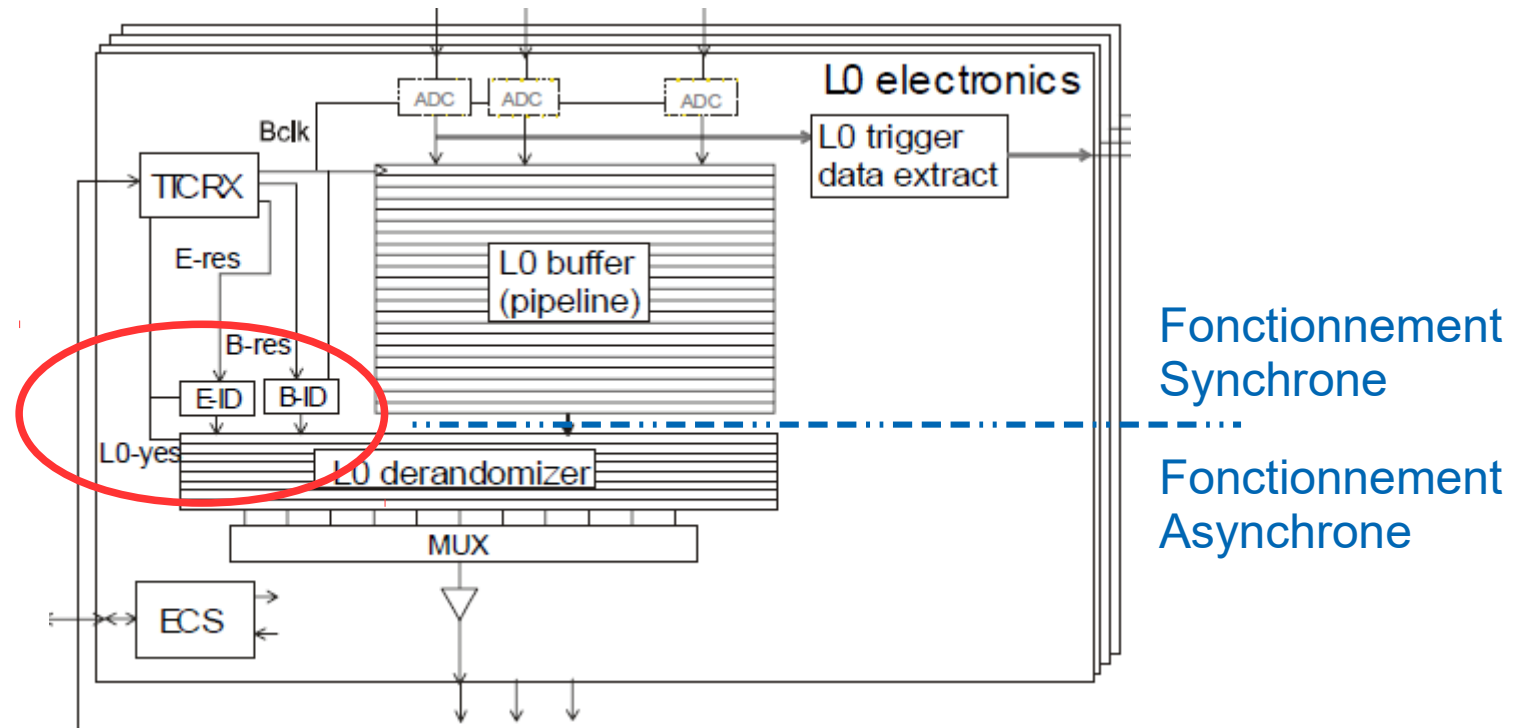
Throttling

- Si les derandomizers d'un Front-End sont prêts de déborder, un signal de « throttle » est émis
- Pas d'action individuelle sur le Front-End
- Le superviseur diminue le nombre d'événements acceptés par l'ensemble des détecteurs jusqu'à obtention d'un niveau raisonnable sur le Front-End en cause
- Identique à une perte d'évènements mais pas de désynchronisation du système



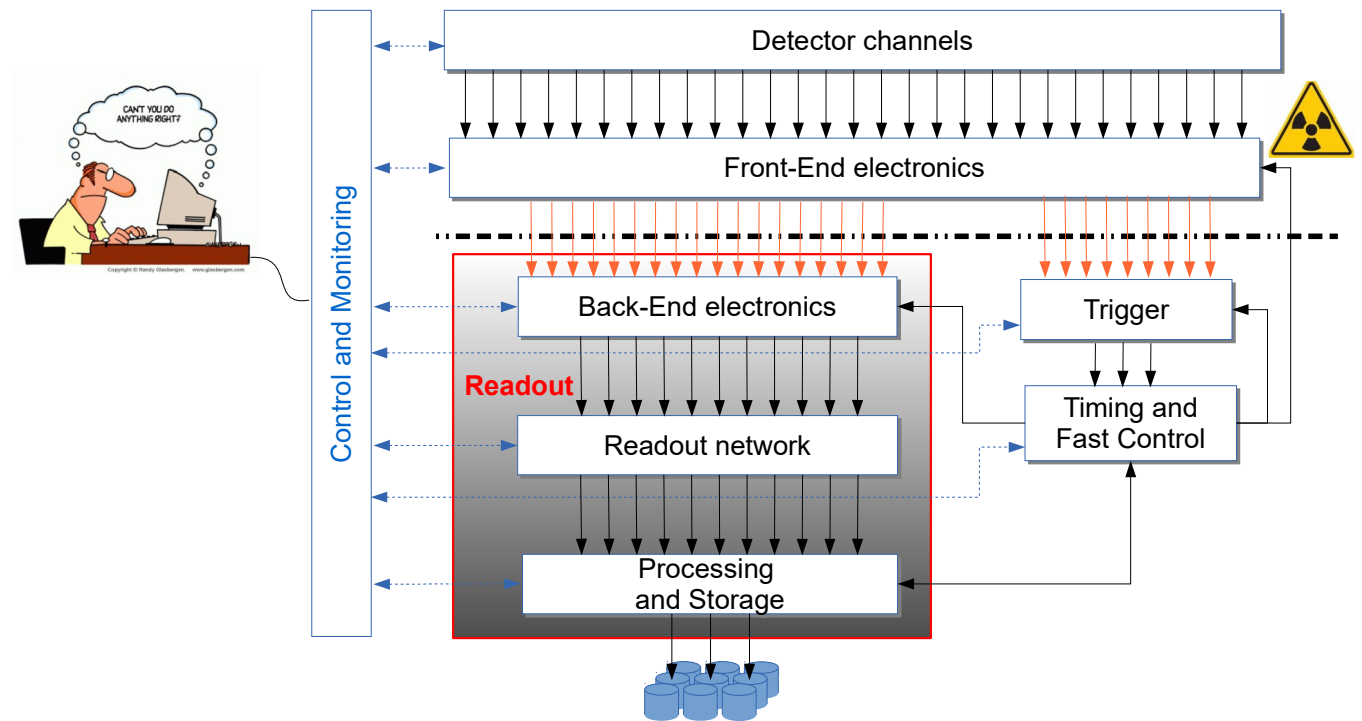
Tagging

- Les événements sont étiquetés car ils ne sont plus consécutifs afin de les reconnaître dans la phase d'assemblage



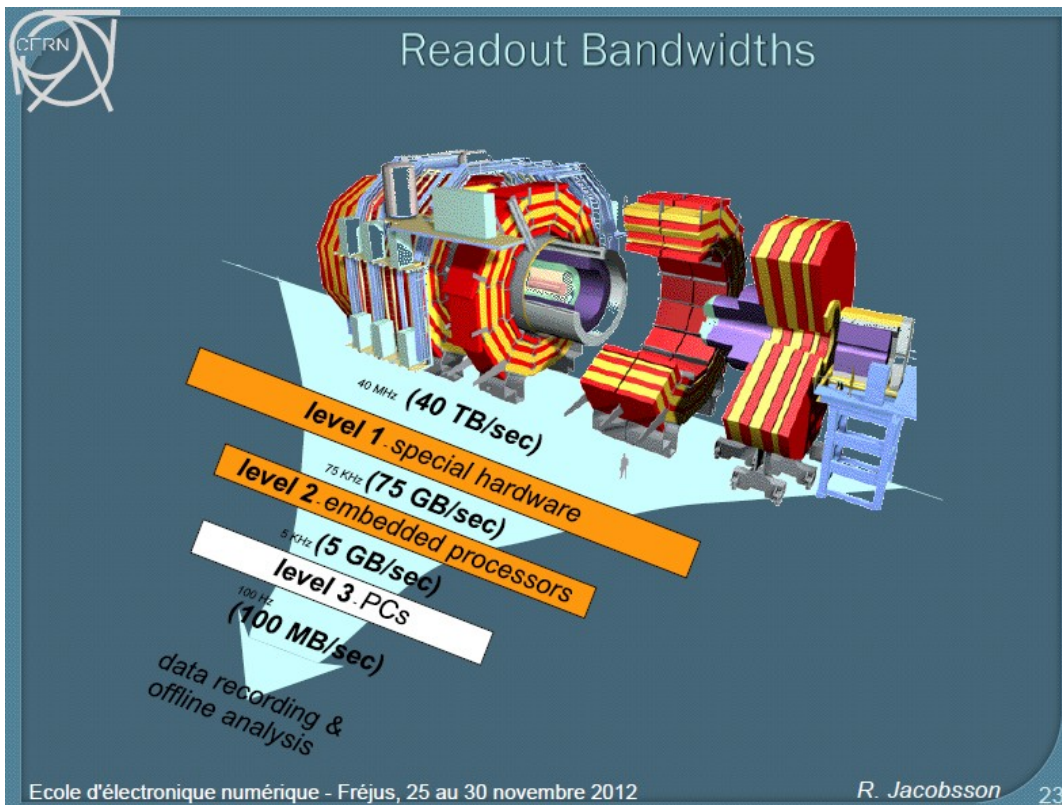
- Mode asynchrone

Systeme de Readout



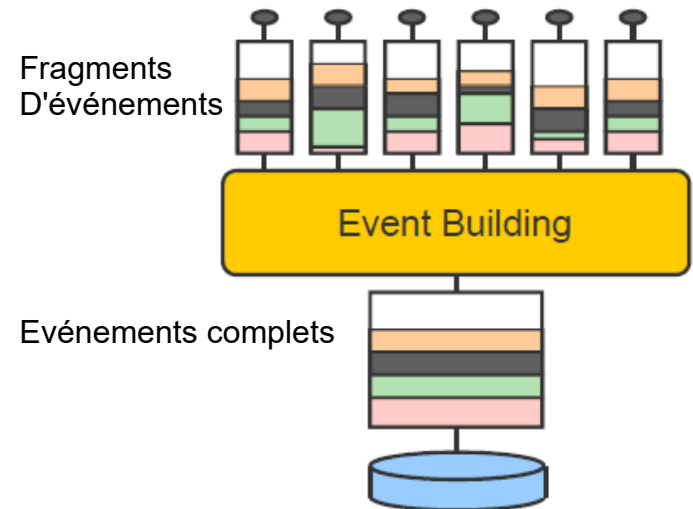
Réduction du flux de données

- La réduction du nombre d'événements à conserver commence dans l'électronique Front-End
- Elle se poursuit du côté backend :
 - Les données restantes sont compressées
 - Plusieurs niveaux d'électronique peuvent être nécessaires
- La réduction se poursuit et se termine dans les fermes



Notion d'event building

- Chaque sous-détecteur envoie au système back-end une **fraction d'événement**
- Tous les événements appartenant à une même collision doivent être rassemblés et agrégés par l'électronique back-end
- Tous les événements appartenant à une même collision doivent être routés à travers le réseau vers un ordinateur unique de la ferme
 - Cet ordinateur verra alors la totalité d'un même événement

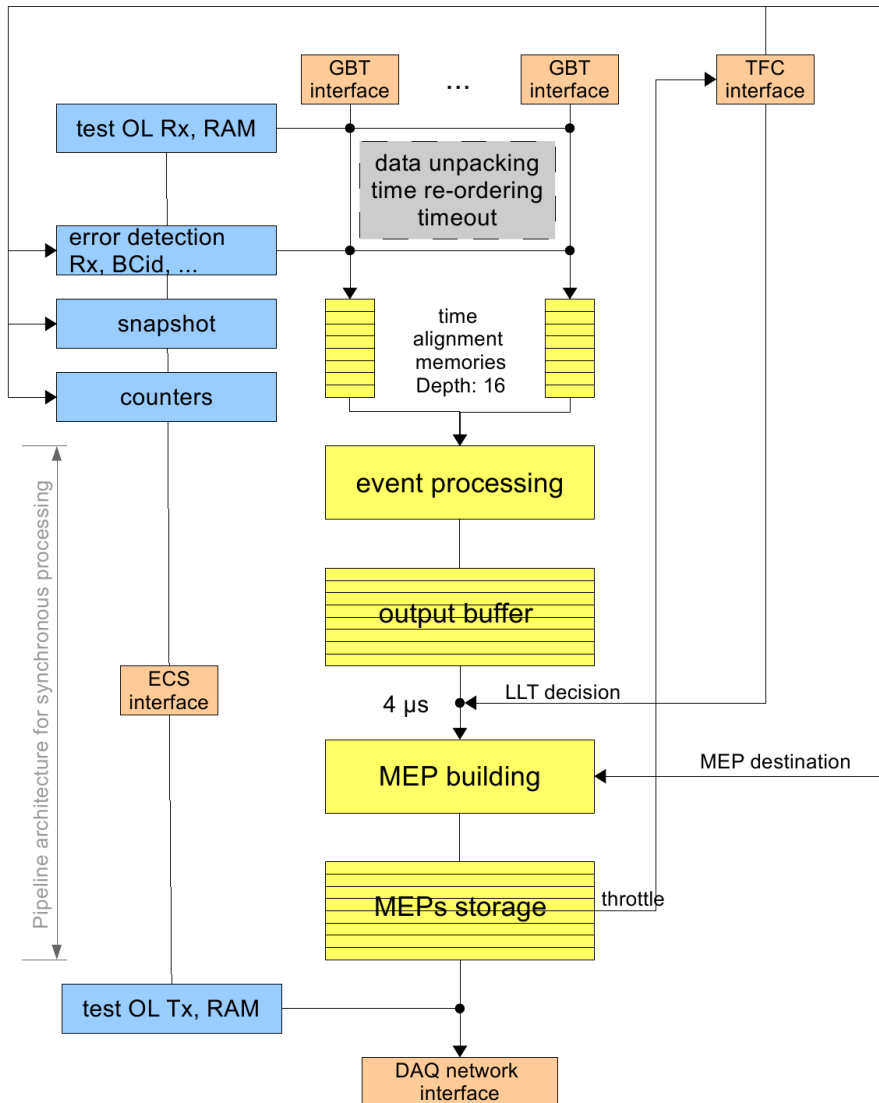


L'ensemble de ces opérations s'appelle **Event Building**.

Elles sont réparties sur l'électronique back-end et le réseau

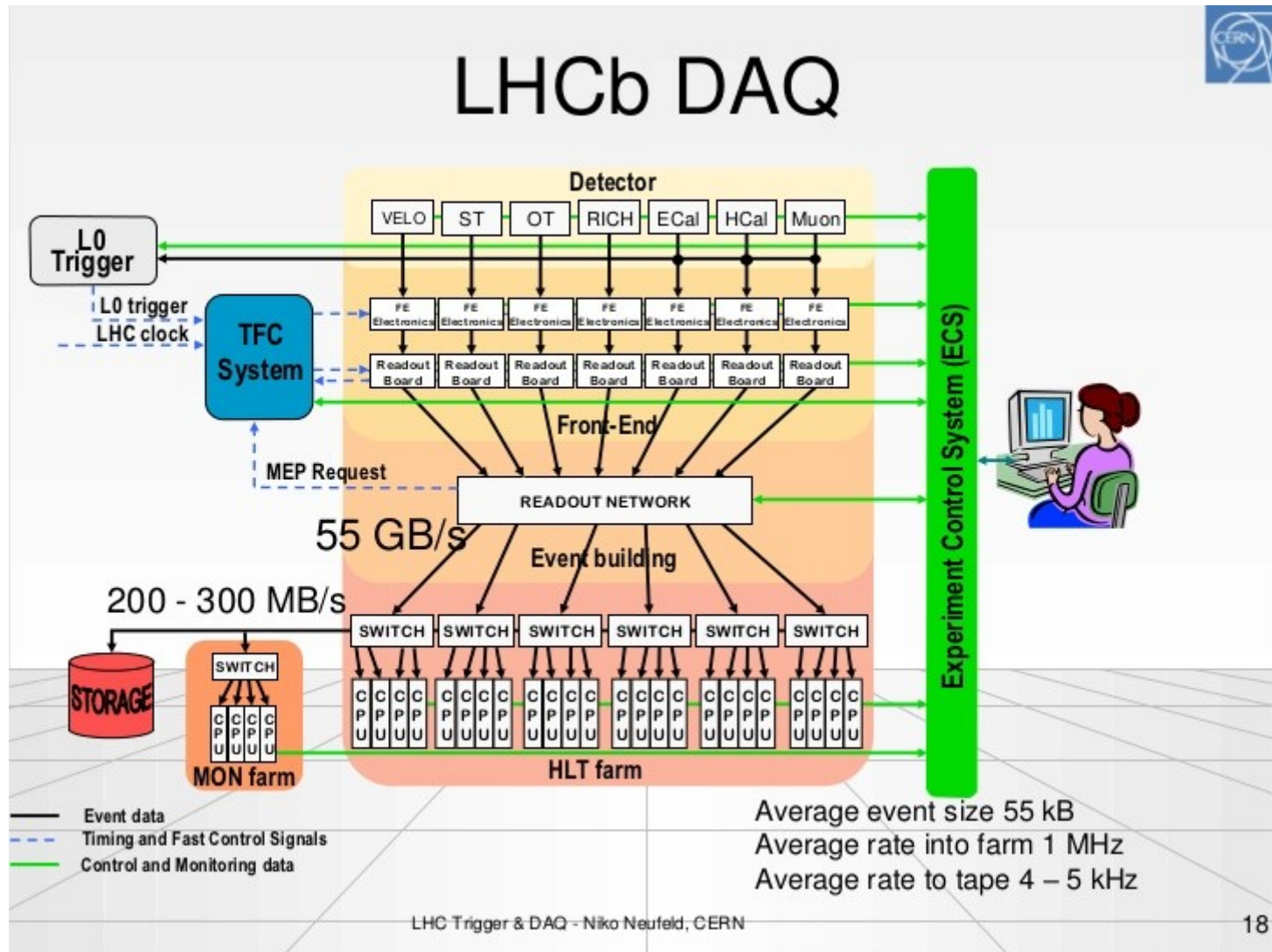
Electronique back-end

bunch crossing identifier and fast commands



- Ces fragments événements arrivent de façon irrégulière :
 - Ils doivent être mis en temps à l'arrivée.
 - Fonctionnement totalement asynchrone
- Des paquets sont formés avec les fragments d'événement appartenant à une même collision.
- Transmission sur un standard de transmission commercial, par exemple Gigabit Ethernet
- MEP : Multi Event Packet. Optimisation des données par rapport à la taille des trames Ethernet
- Chaque MEP doit être envoyé vers un ordinateur disponible de la ferme.

Réseau Event building



Stratégies d'envoi des données dans le réseau

Plusieurs stratégies de gestion du réseau :

Mode « **Push** » :

- Les données sont poussées dans les switches du réseau.
- La source doit connaître la destination d'un CPU disponible

→ Avantages :

- simple pour les cartes back-end
- moins de fibres

→ Inconvénients :

- pas de relecture possible en cas d'erreur
- risque de congestion du réseau
- switches doivent contenir de la mémoire → chers

• Mode « **Pull** »

- Les données sont demandées par les destinations

→ Avantages :

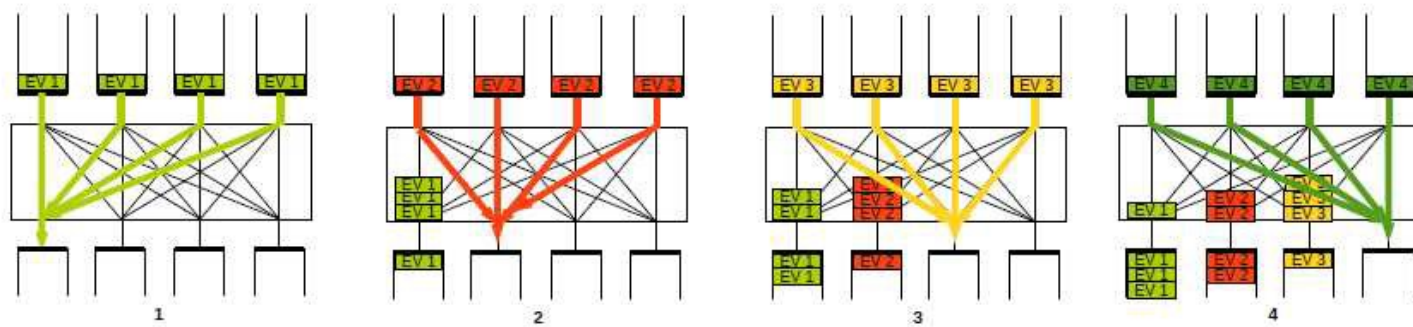
- relecture possible
- switches bon marché

→ Inconvénients :

- l'électronique back-end doit contenir une grande quantité de mémoire
- plus de fibres

Stratégies de readout

Mode Push

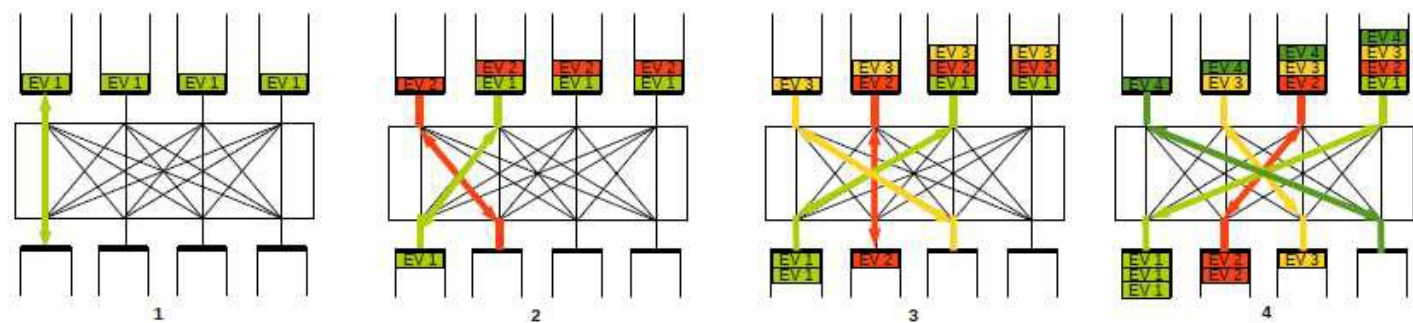


Electronique back-end

Switches

Ferme de calcul

Mode Pull



Electronique back-end

Switches

Ferme de calcul

Fermes de calcul

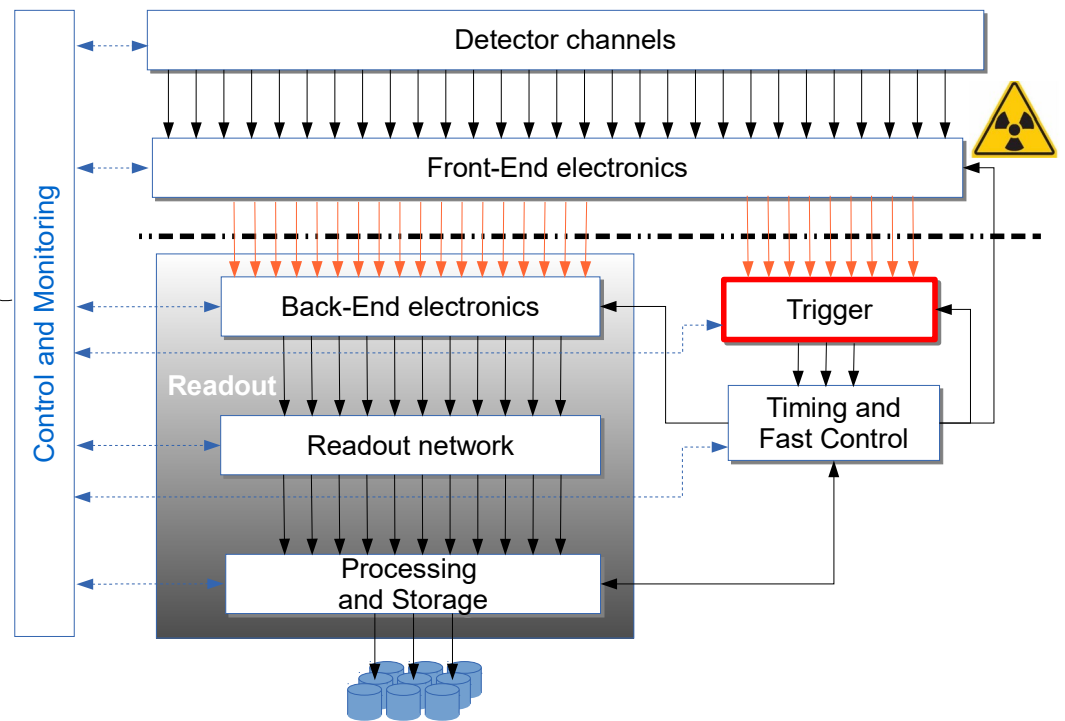
Ordres de grandeur



	ALICE	ATLAS	CMS	LHCb
# cores (+ hyperthreading)	2700	17000	13200	15500
# servers (mainboards)		~ 2000	~ 1300	1574
total available cooling power	~ 500	~ 820	800	525
total available rack-space (Us)	~ 2000	2400	~ 3600	2200
CPU type(s)	AMD Opteron, Intel 54xx, Intel 56xx	Intel 54xx, Intel 56xx	Intel 54xx, Intel 56xx Intel E5-2670	Intel 5450, Intel 5650, AMD 6220

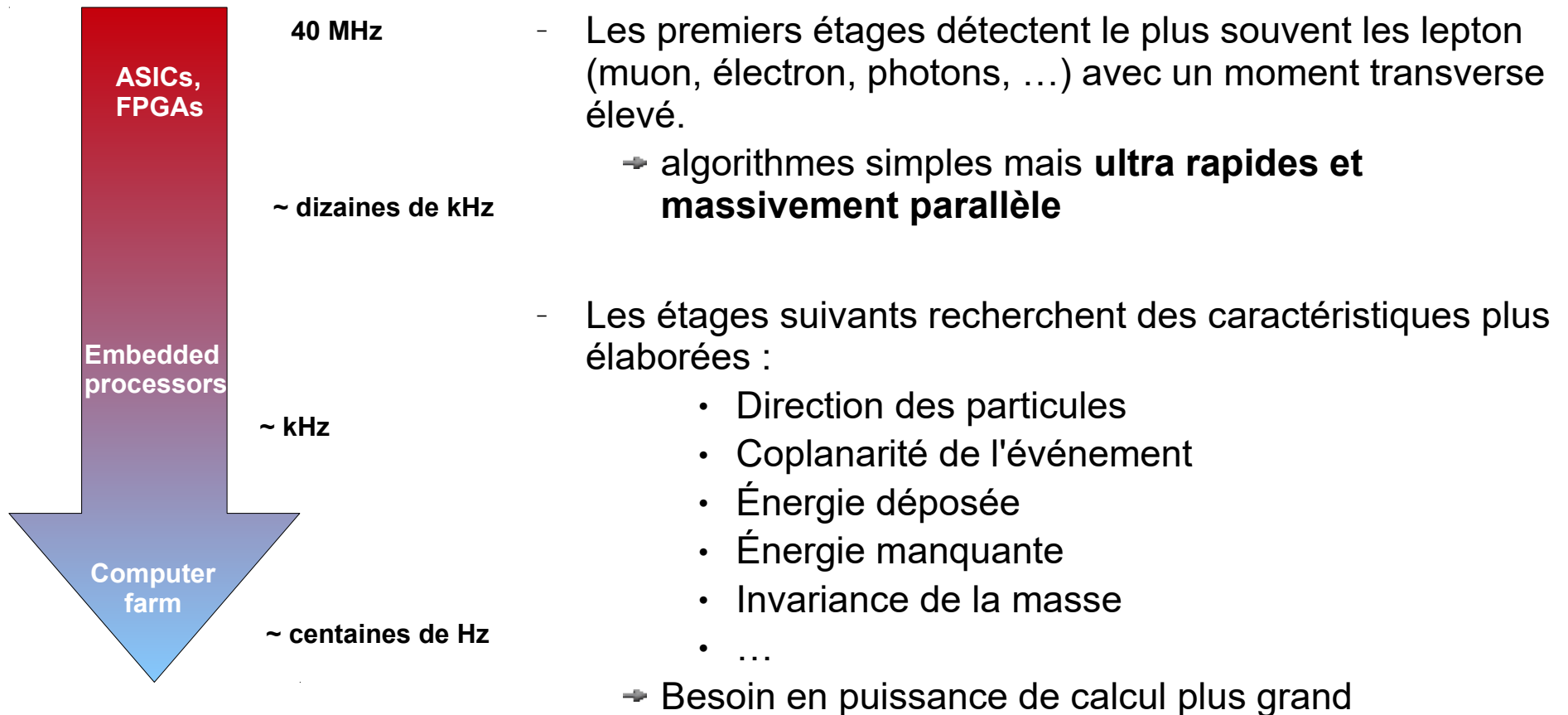
Source Niko Neufeld, CERN

Trigger



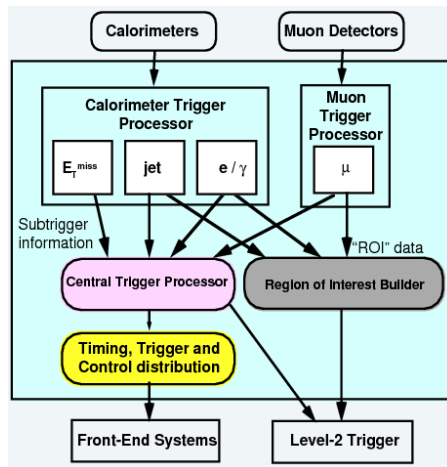
Tri des événements

Plusieurs niveaux pour filtrer : Low Level Triggers (LLT)+ High Level Triggers (HLT)



Niveaux de trigger

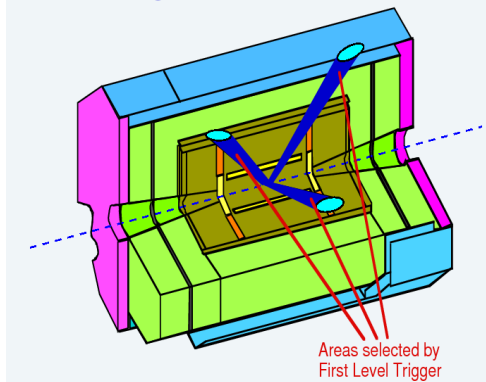
Exemple d'ATLAS



- 3 niveaux de trigger

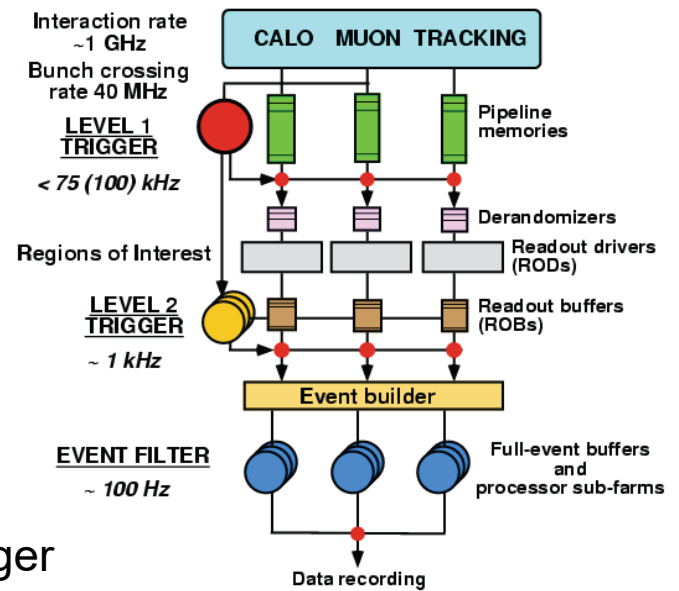
- **Niveau 1** : recherche de candidats avec un haut P_T : muons, electrons/photons, hadrons et jets.
 - 3000 Gbits de données d'entrées par seconde
 - Taux d'événements réduit à ~75-100 kHz
- **Niveau 2** : le niveau 1 sélectionne des régions d'intérêt. Seules les données de ces régions sont envoyées au Niveau 2
 - Réduction du volume d'information à transmettreAnalyse plus fine en utilisant la granularité totale des détecteurs + informations du inner tracker
 - Taux d'événements réduit à ~1 kHz

Regions of Interest (Rol)

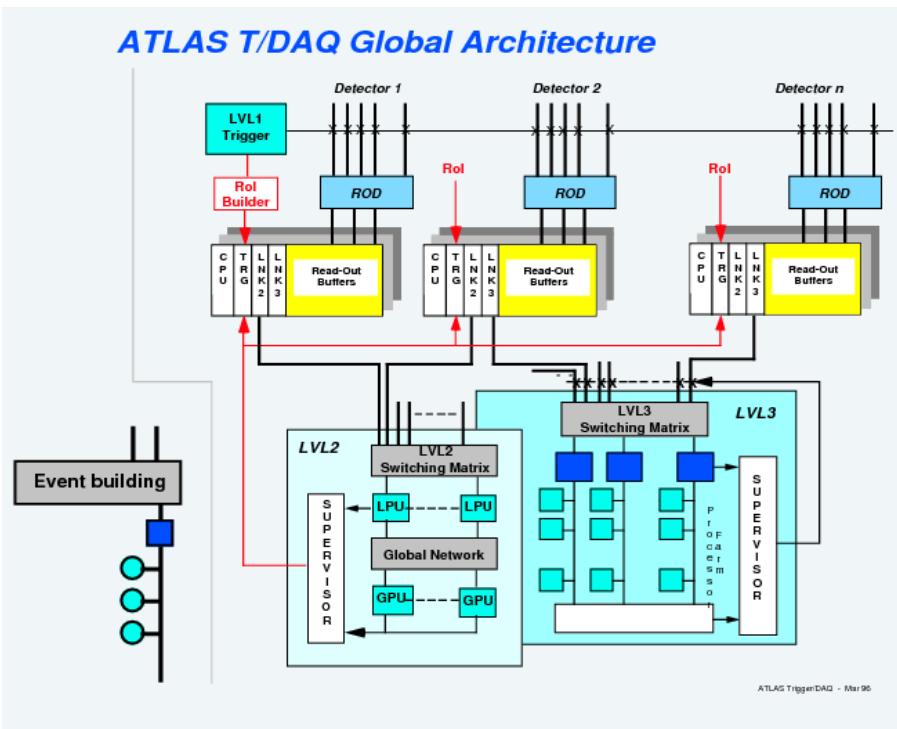


Niveaux de trigger

Exemple d'ATLAS



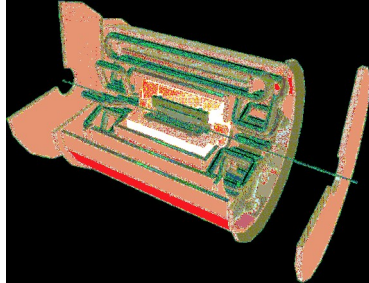
ATLAS T/DAQ Global Architecture



- 3 niveaux de trigger
 - **Niveau 3** HLT (High Level Trigger): event building puis recherche par les fermes de calcul sur la totalité de l'événement
 - algorithmes similaires à l'analyse de donnée, mais en temps réel
 - temps de calcul ~ 1 seconde par événement
- Taux d'événements conservés : ~ 100 Hz

Niveaux de trigger

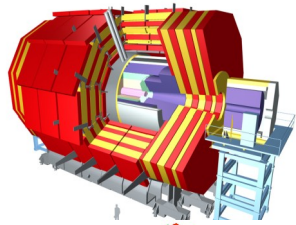
ATLAS



3 niveaux de trigger	Niveau 1	Niveau 2	HLT
Latence	2.5 μ s	~ms	~ s
Débit	100 Gbytes/s	10 Gbytes/s	1.6 Gbytes/s
Taux d'evts	70 kHz	~ kHz	100 Hz

Taille d'événement :
1.6 Mbyte

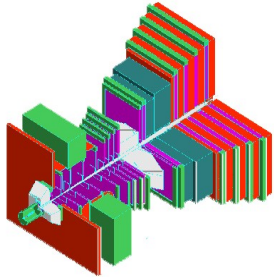
CMS



2 niveaux de trigger	Niveau 1	HLT
Latence	3 μ s	~ s
Débit	100 Gbytes/s	100 Mbytes/s
Taux d'evts	100 kHz	100 Hz

Taille d'événement :
1 Mbyte

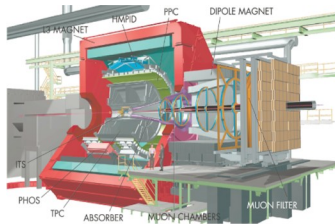
LHCb



3 niveaux de trigger	Niveau 0	HLT1	HLT2
Latence	4 μ s	58ms	~ s
Débit	10 Gbytes/s	3.5 Gbytes/s	350 Mbytes/s
Taux d'evts	1 MHz	50 kHz	5 kHz

Taille d'événement :
75 kbytes

ALICE

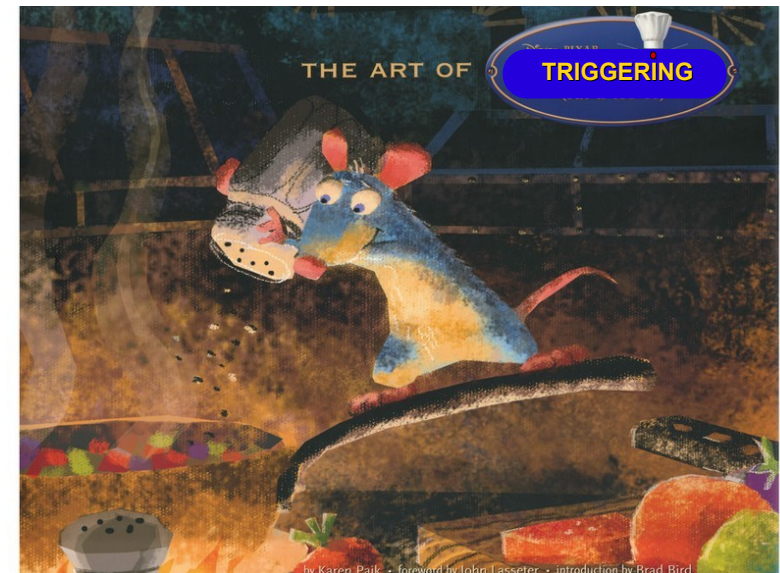


4 niveaux de trigger	Niveau 0	Niveau 1	Niveau 2	HLT
Latence	1.2 μ s	6.5 μ s	88 μ s	~ s
Débit			5 Gbytes/s	1.2 Gbytes/s
Taux d'evts	4 kHz		500 Hz	50 Hz

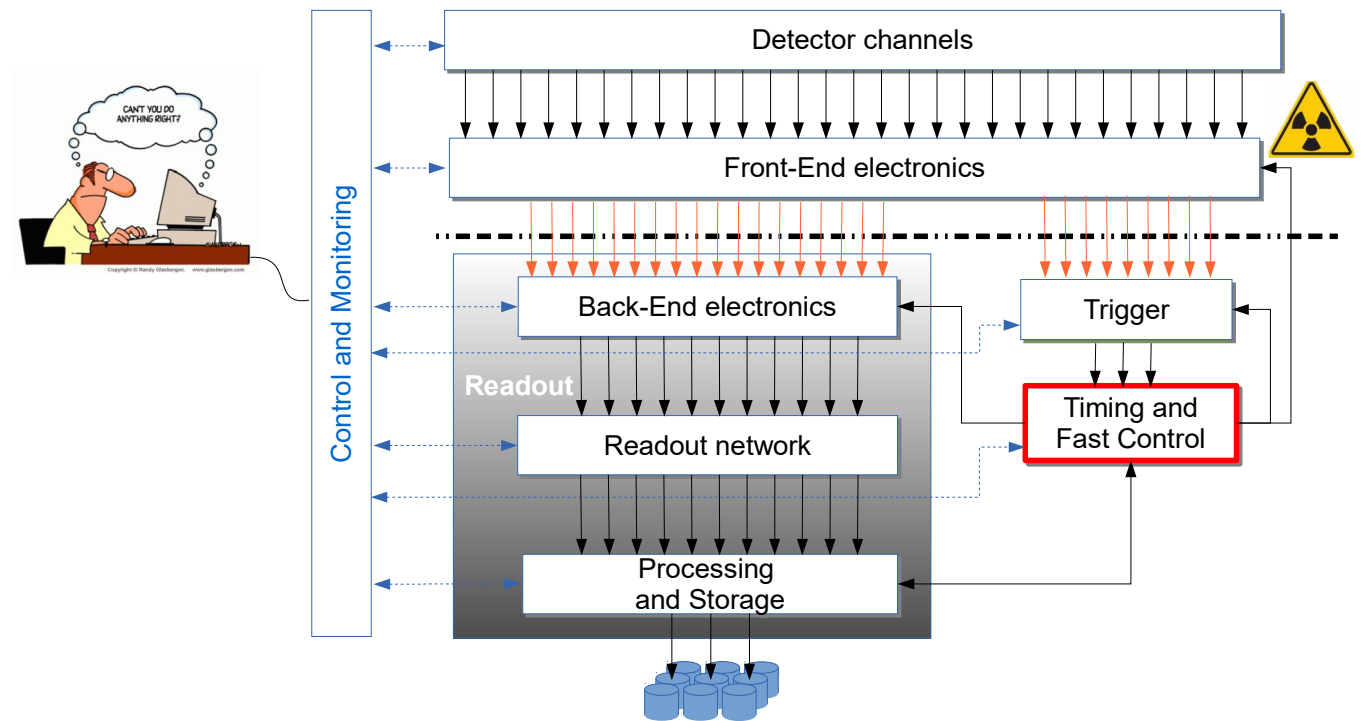
Taille d'événement :
Pb-Pb 86 Mbytes
p-p 2.5 Mbytes

Exigences d'un bon trigger

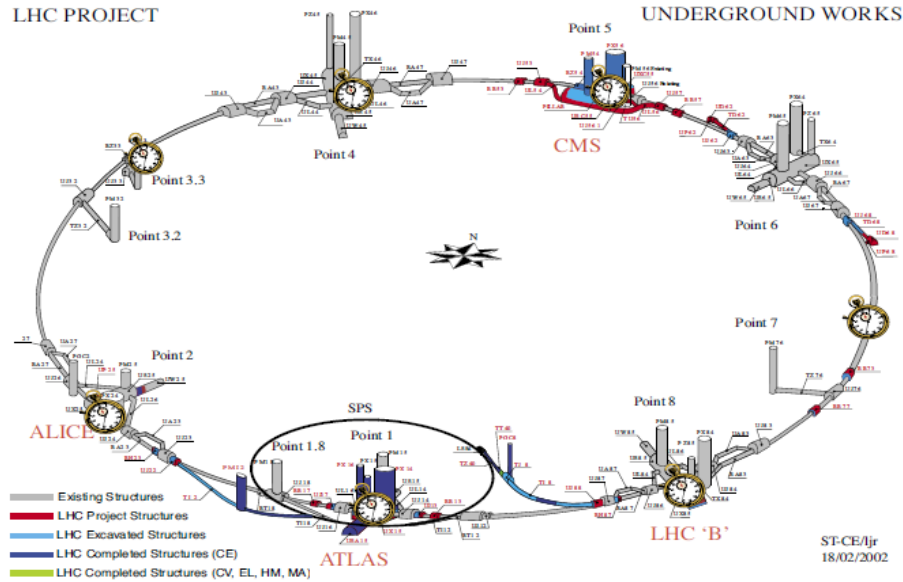
- Efficacité la plus grande possible
 - Un événement rejeté est définitivement perdu
 - Peu sensible au bruit
 - Dépourvu de biais de mesure
- Taux de sélection le plus bas possible
- Très rapide
 - Pour diminuer la taille mémoire dans les Front-Ends
- Flexible
- Economique (!)
- Fiable



Supervision temporelle

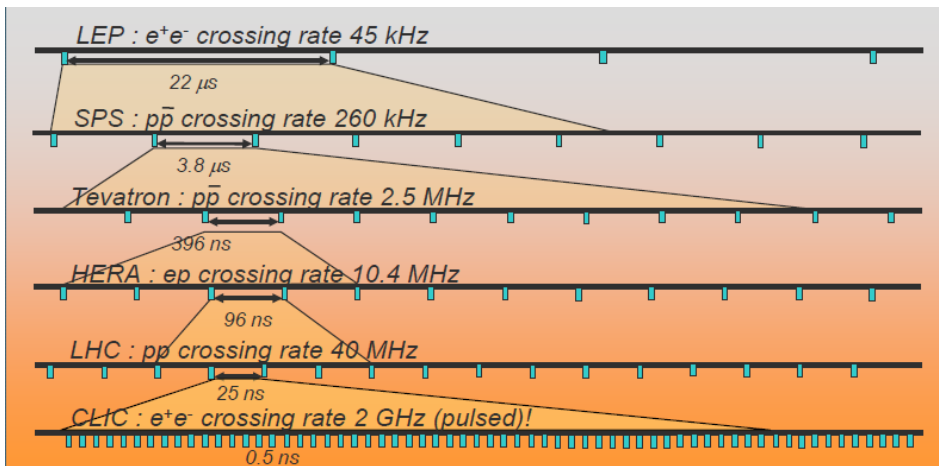


Synchronization



Globale

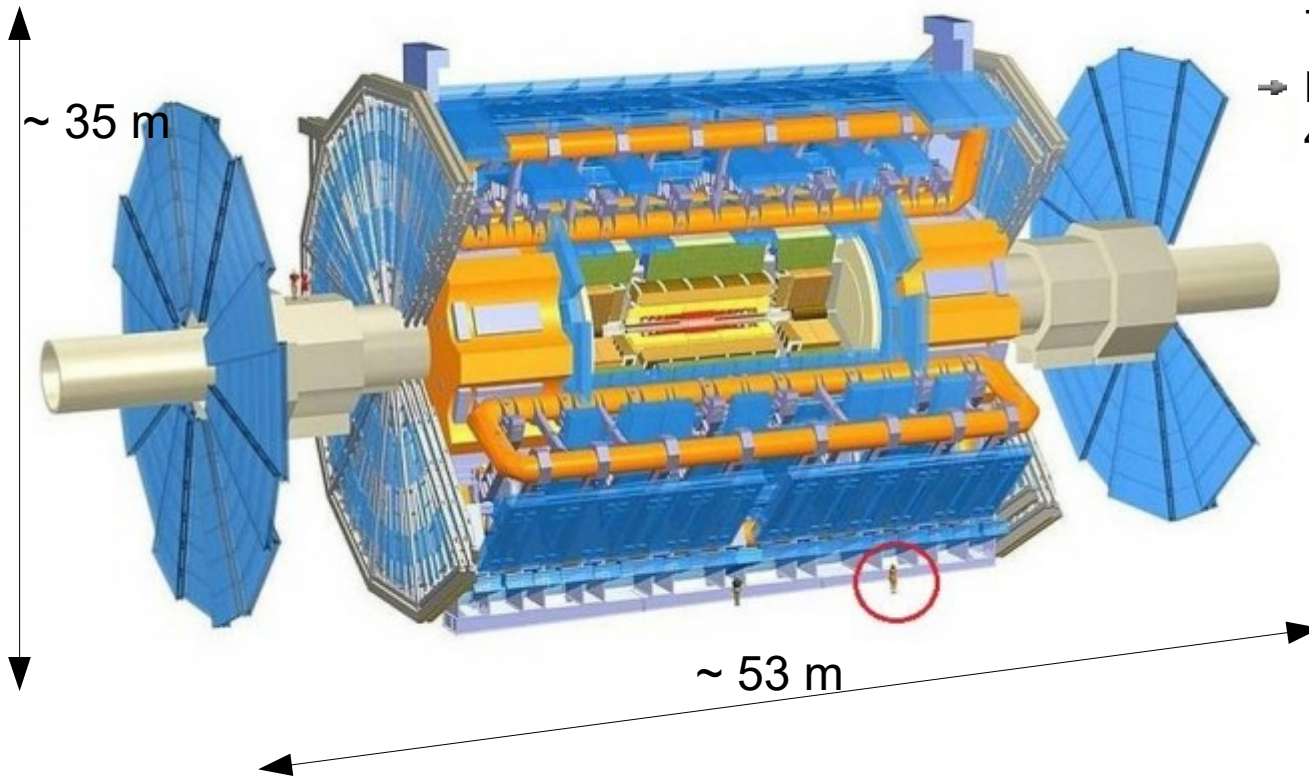
- Au niveau de l'accélérateur tous les éléments doivent être synchronisés sur une grande distance
 - Distribution sur plusieurs km de fibre
- Et avec une grande précision
 - Sur le LHC jitter maximal : ~ 8 ps RMS



Synchronization

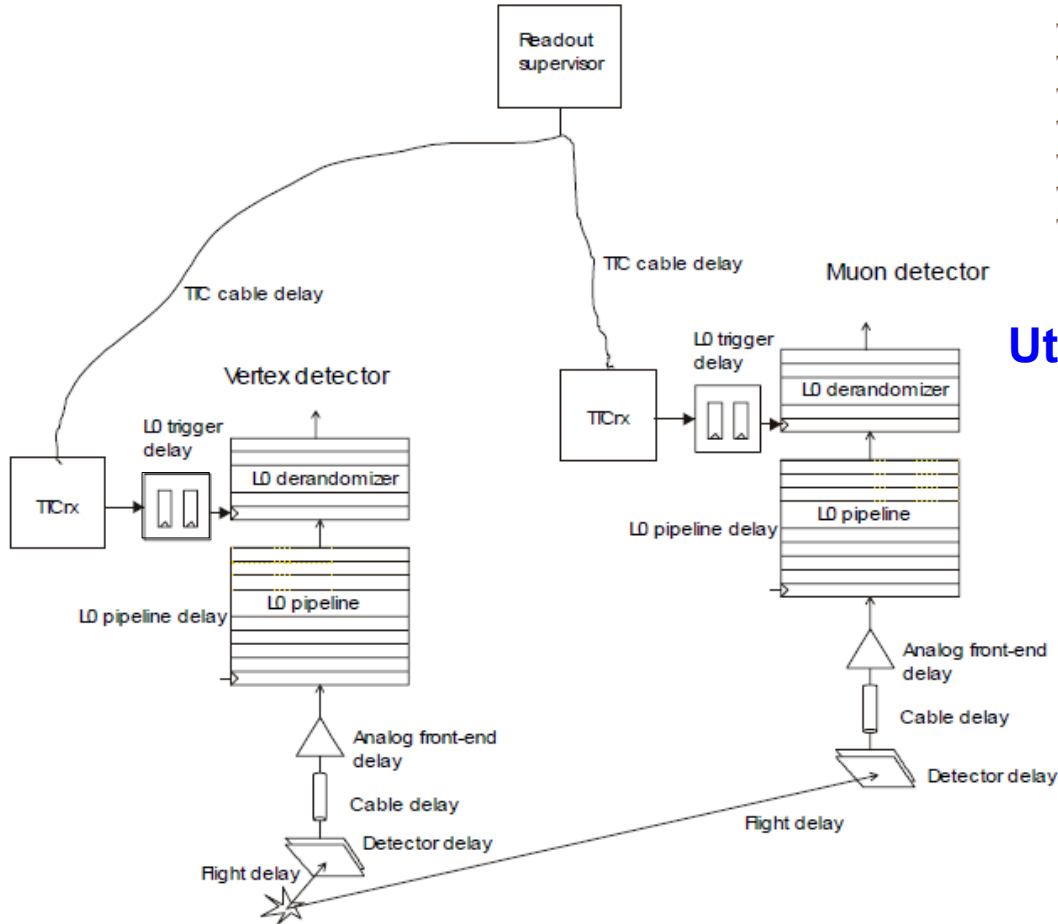
Locale

ATLAS



- Au niveau de l'expérience également
 - Temps de vol d'une particule 7.5 m / cycle d'horloge
 - Délais dans les cables 4 m / cycle d'horloge

Principe d'alignement

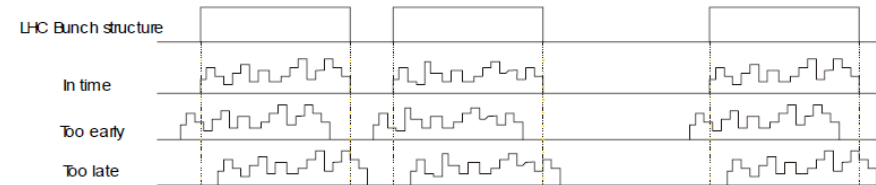


Éléments à prendre en compte

T_{flight} :	Flight time from interaction point
$T_{detector}$:	Detector delay
$T_{detector_cable}$:	Cable delay from detector
T_{analog} :	Delay in analog front-end
$T_{l0_pipeline}$:	L0 pipeline latency
$T_{l0_pipe_delay}$:	Delay from L0 trigger from TTCrx to data latched in L0 derandomizer
T_{ttrcx} :	Delay in TTCrx (programmable)
T_{ttc_cable} :	Delay in optical TTC distribution

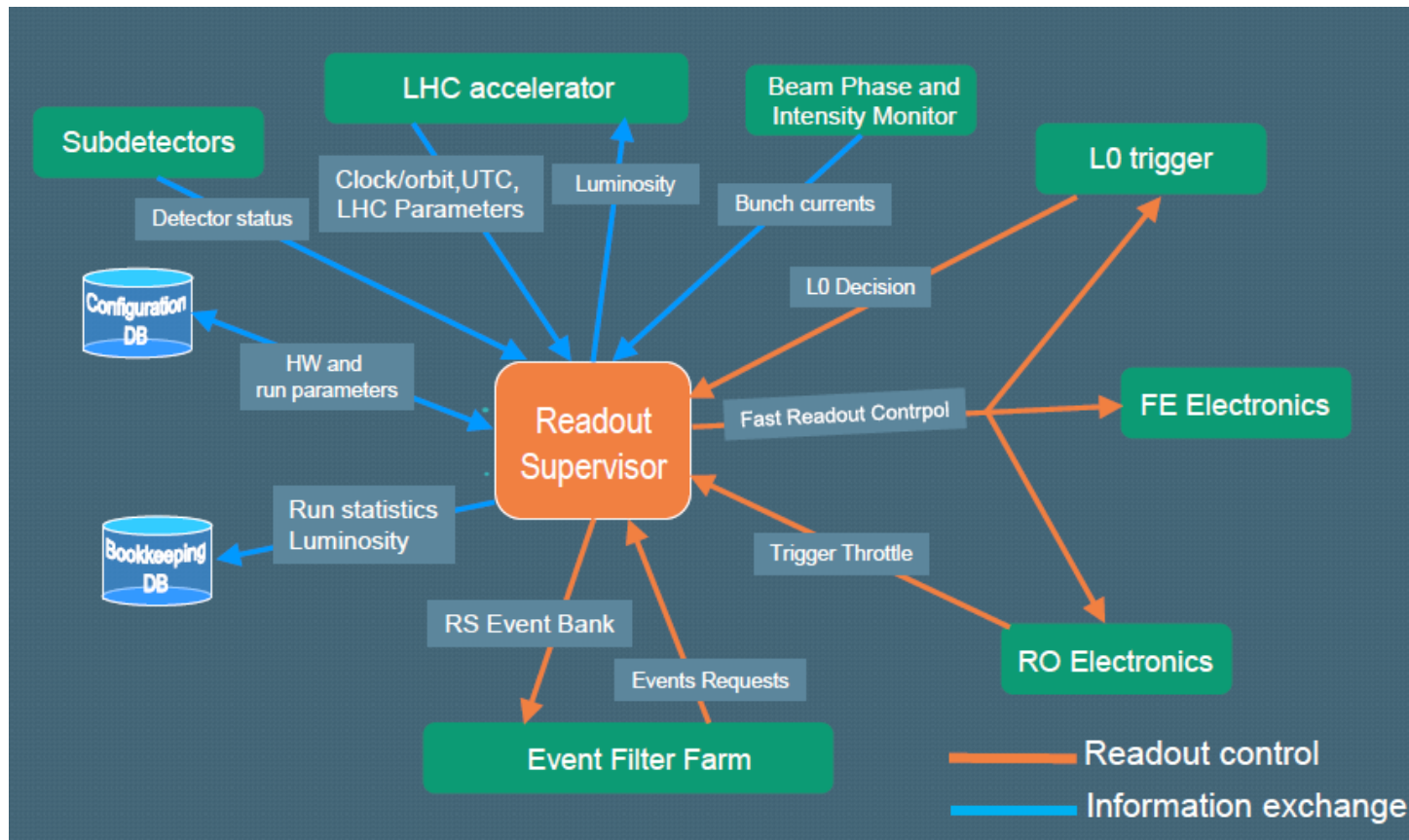
Utilisation de lignes à retard locales

- Alignement par rapport à une particule facilement identifiable
ex : cosmique
- Ou bien par rapport au faisceau



Timing and Fast Control supervisor

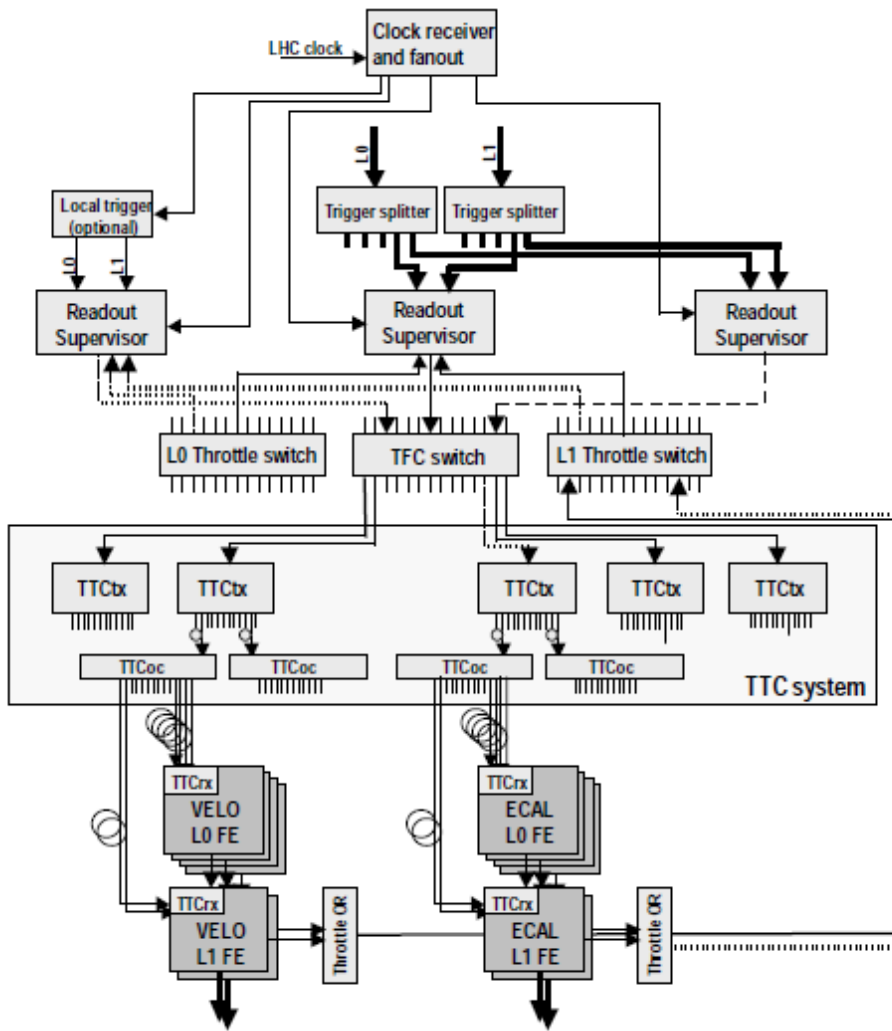
- Assure la distribution temporelle précise
- Assure la distribution des triggers aux Front-Ends et systèmes de Readouts
 - Véritable chef d'orchestre du système



Source R. Jacobsson - CERN

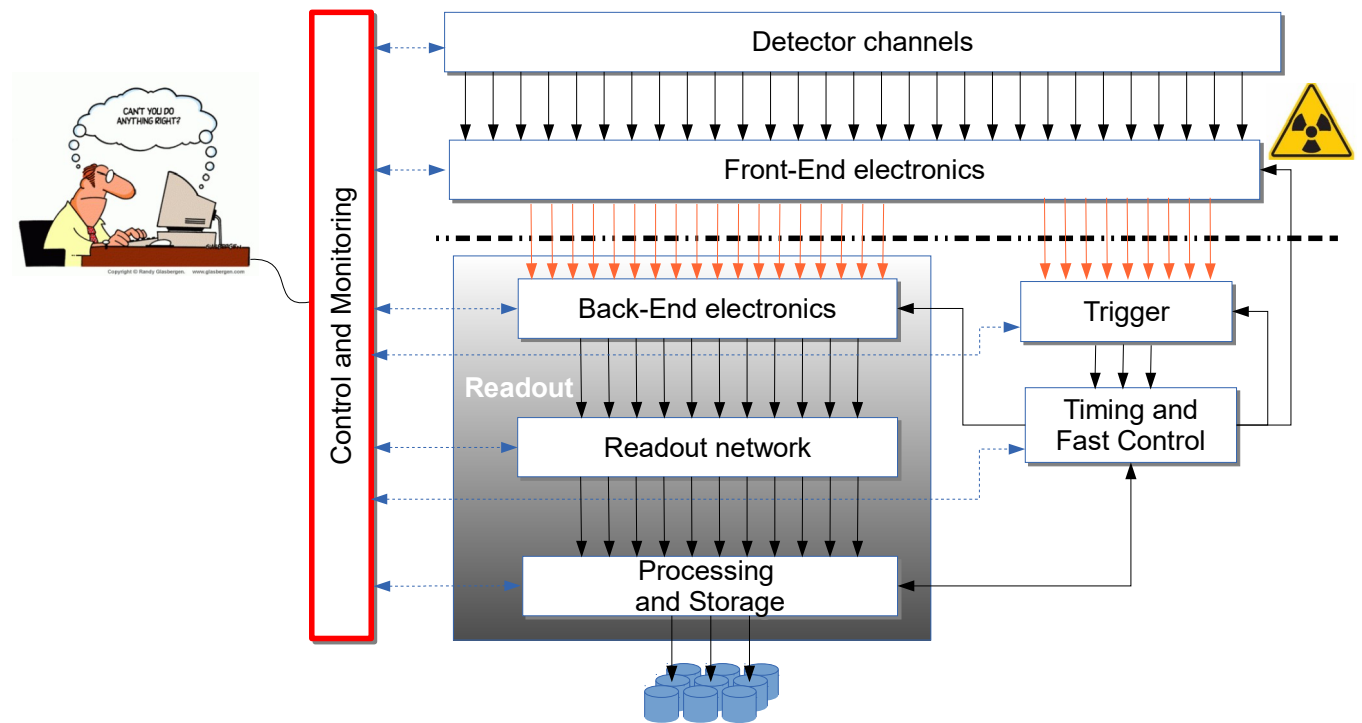
Timing and Fast Control supervisor

- Distribution des horloges et triggers sur une arborescence optique
- Possibilité de partitionnement
 - Permet des fonctionnements indépendants des sous-systèmes
 - Très utile durant les phases de commissioning



Source R. Jacobsson - CERN

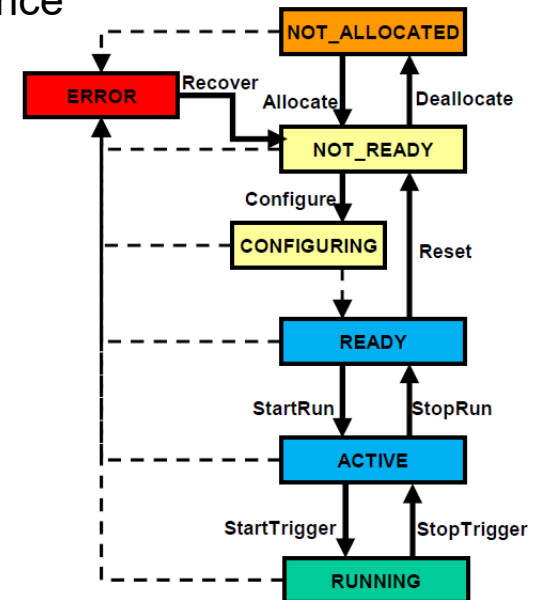
Contrôle et monitoring



Experiment Control System

Assure la supervision globale de l'expérience :

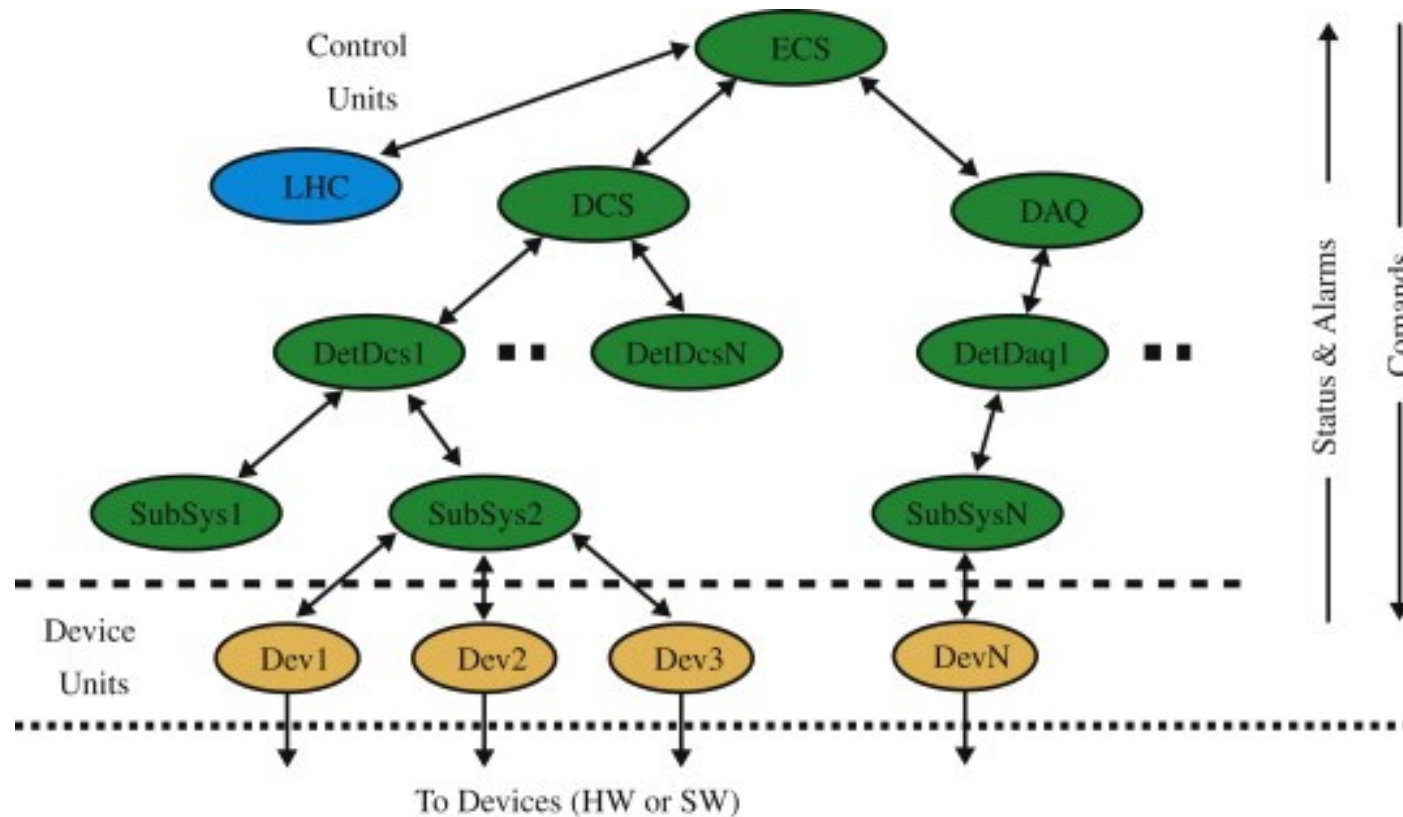
- Pilotage de l'acquisition de données et du trigger :
 - Chargement des paramètres, type de RUN, activation (START) ou désactivation (STOP) de l'expérience, ...
 - Partitionnement
 - ➔ Permet de faire tourner certaines parties de l'expérience indépendamment
 - Détection et récupération d'erreurs
 - Surveillance du système et des flux de données
- Pilotage des **sous-systèmes**
 - Gaz, HV, LV, températures, ...
- Pilotage des **infrastructures de l'expérience**
 - Refroidissement, ventilation, distribution électrique, ...
- Interaction avec les **éléments externes à l'expérience**
 - Aimant, accélérateur, systèmes de sécurité, etc.



Experiment Control System

Système hiérarchique :

- Plusieurs milliers d'information à surveiller



Panneaux de supervision

- Contrôle
 - Synthétisation des résultats à haut niveau
 - Détails sur demande

LHCb: TOP

System: LHCb | State: **READY** | Auto Pilot: OFF

Run Number: 95443 | Activity: CCESCAN

Run Start Time: 08-Jul-2011 12:41:13 | Trigger Configuration: Calibration_Scan_PrescaleL0

Run Duration: 000:05:58 | Time Alignment: TAE half window 0, L0 Gap

Nr. Events: 0 | Max Nr. Events: Run limited to 1000 Events

Step Nr: 0 | To Go: 60 | Automated Run with Steps: Step Run with 66 Steps | Start at: 6

L0 Rate: 0.00 Hz | HLT Rate: 0.00 Hz | Dead Time: 0.00 %

Data Destination: Castor | Data Type: CALIBRATION11 | File: aqarea/lhcb/data/2011/RAW/FULL/LHCb/CALIBRATION11/95443

Sub-Detectors: TDET, VELOA, VELOC, TT, IT, OTA, OTC, RICH1, RICH2, PRS, DT_ALLOCATE, ECAL, HCAL, MUONA, MUONC

Trigger Components: LODU, TCALO, TMUA, TMUC, TPU

Sub-System States:

Sub-System	State
DCS	READY
DAI	READY
DAQ	READY
RunInfo	READY
TFC	NOT_READY
HLT	ERROR
Storage	NOT_READY
Monitoring	READY
Reconstruction	NOT_ALLOCATED
Calibration	NOT_READY
LHCb_HV_TOP	READY

HC22CROC TOP

Device: HC22CROC | State: **READY**

Control: Slot Power, State, Control

Crate Configuration: Sel. FEB, Hdr. FEB, Mode Clock, Mode LO/Ch B

General Control: Mode Clock, Mode LO/Ch B, Clock Settings (Fine, Coarse), FE Synchronisation (FEB, CROC)

Optical Measuring Mode: State, Physics Test, Left/Right All Fibres

Subfarm1: System1-Manager3

System: Subfarm1 | State: **NOT_READY**

Sub-System States:

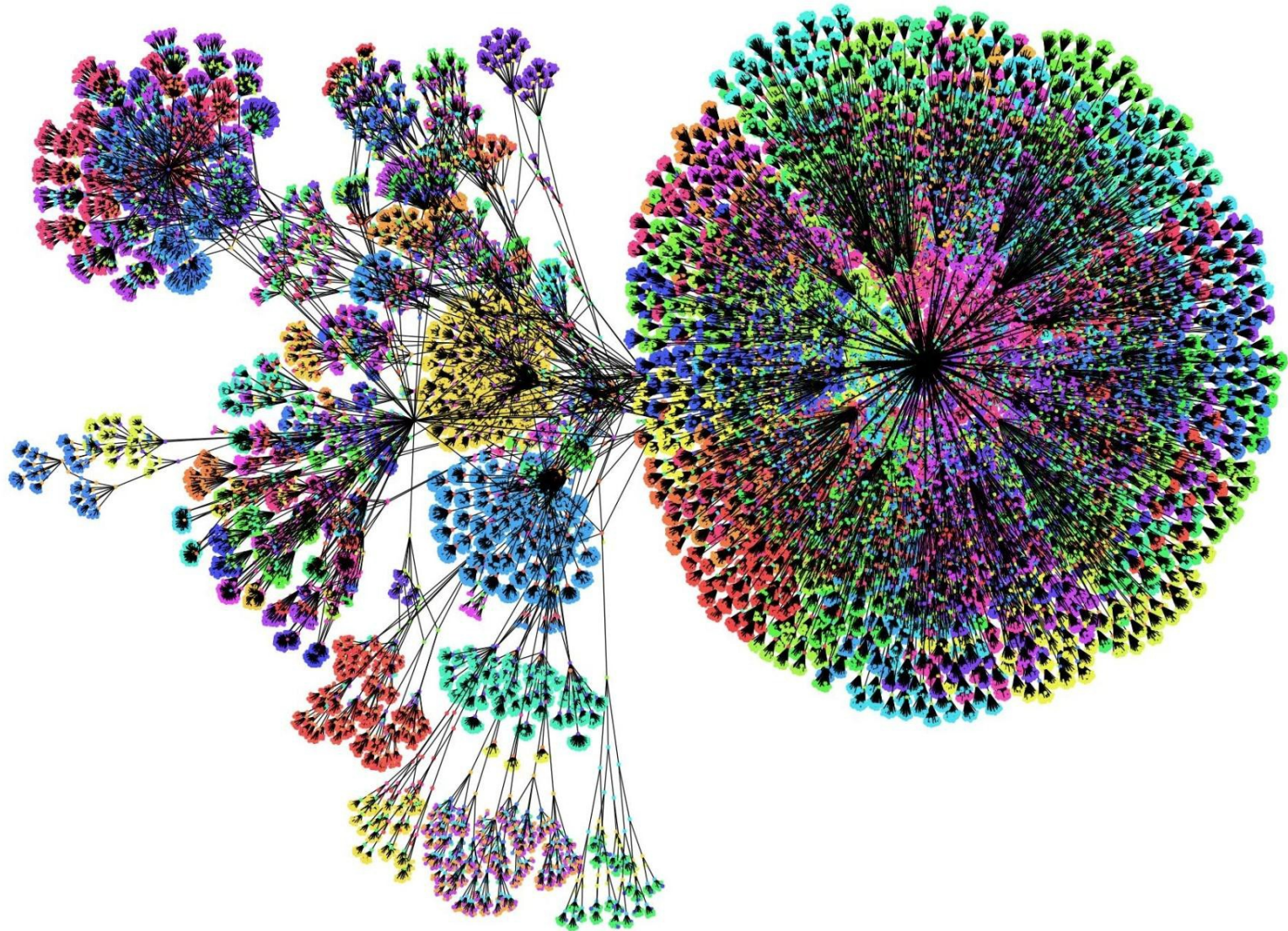
Node	State
Node001	RUNNING
Node002	RUNNING
Node003	RUNNING
Node004	RUNNING
Node005	NOT_READY
Subfarm1Coordinat	READY

Monitoring Data:

- trigger rate: 2.000000
- # processed events: 1229
- # B -> u+ u- X: 527
- # B -> e+ e- X: 59
- # rejected events: 643
- status: processing

available histograms: ECALhitmap, HCALhitmap, eventtype, gEfficiency, VELOefficiency

Contrôle et monitoring



CMS Control and monitoring

Conclusion

Complexité considérable

- Essentiel d'avoir un système de monitoring très performant

Besoin de flexibilité

- Durée de vie d'une expérience = ~ 10 à 15 ans
- Tendances à rendre le maximum de fonctions programmables
 - Utilisation de FPGAs, même dans les Front-Ends
 - Migration de certaines fonctions dans les fermes de calcul

Besoin de robustesse

- Opération 24h/24 durant des périodes très longues

Compromis performance/coûts toujours un challenge

- Trouver le meilleur équilibre entre développement spécifiques et usage des composants du commerce, tout en limitant les risques
- Requier des études prospectives permanentes