

Trigger et acquisition de données

Pierre-Yves Duval
(CPPM)



Plan

Systeme d'acquisition de donnees (DAQ)

Notion d'assembleur d'evenements (event builder)

Notion de selectionneur (trigger)

Specifications et elements de conception d'un systeme DAQ

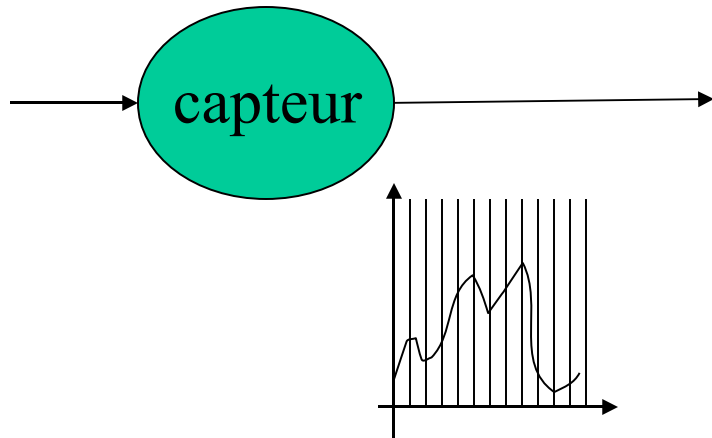
Les bancs de test en physique nucleaire et des particules

Contrôle et commande des systemes DAQ

Le systeme trigger DAQ d'une grande experience sur accelerateur

Acquisition de données simple

Point de mesure analogique (courant)

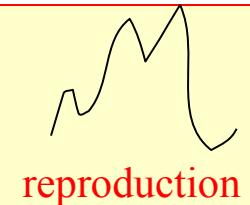


Conversion numérique

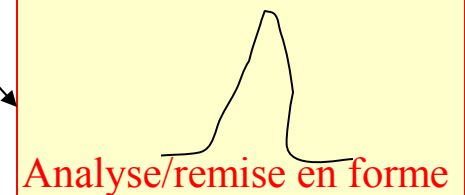


Signal d'échantillonnage

```
00011001
00011010
00011110
00011000
00001111
```



OK ou KO
sélection



Total=1236577
Stockage/archivage

Besoin de plus de puissance si:

- mesure à plus haute fréquence
- plus de points de mesures
- traitements plus complexes

Terminologie:un évènement

Evènement

Ensemble des valeurs des mesures effectuées pendant **un phénomène physique** qui s'est déroulé à **une date particulière dans le temps**.

Un sous ensemble des données d'un évènement est un **fragment d'évènement**

Exemples:

Accélérateur : données associées à une collision

Astroparticules : données associées à une gerbe ou à la trajectoire d'un muon

Temps mort

Temps pendant lequel des **données sont perdues ou jetées** parce que le système d'acquisition est bloqué ou en surcharge.

Dimensionnement et temps mort

Spécifications => dimensionnement:

Flux d'entrée :

volume des données d'entrée

loi de distribution de leur arrivée: périodique, aléatoire, rafale
(bande passante moyenne, taille des tampons)

Traitement à effectuer: durée et distribution des durées (puissance)

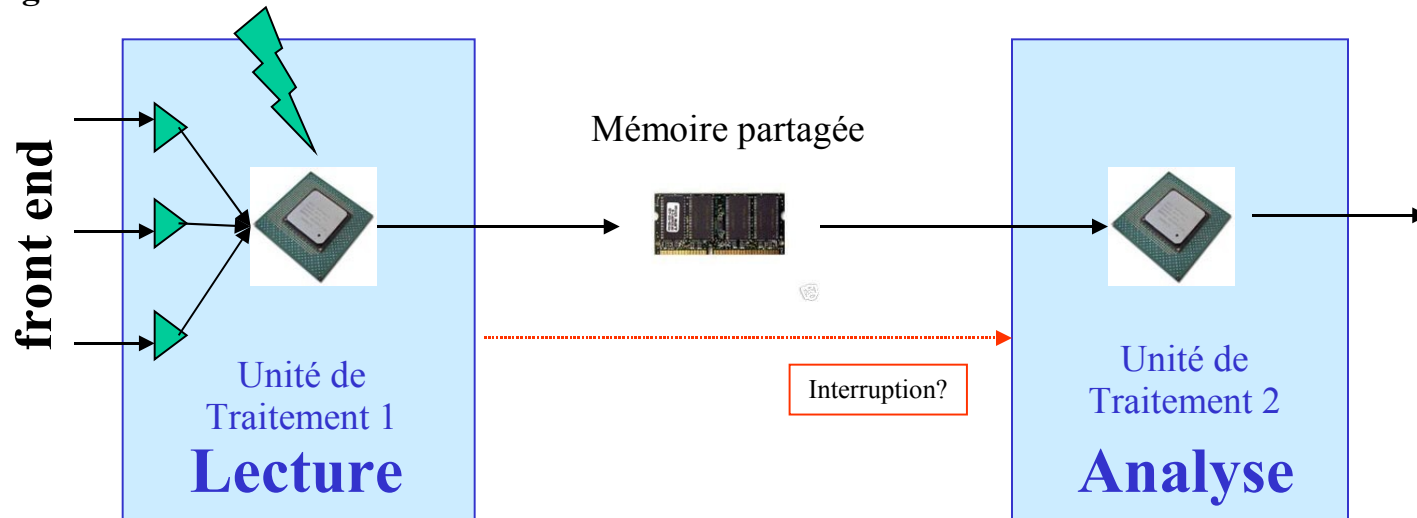
Volatilité des données: Durée de disponibilité dans le frontal (réactivité)

Performance => Temps mort:

Taux de perte acceptable.

Blocs fonctionnels de base d'une chaîne DAQ

Signal de déclenchement de la lecture

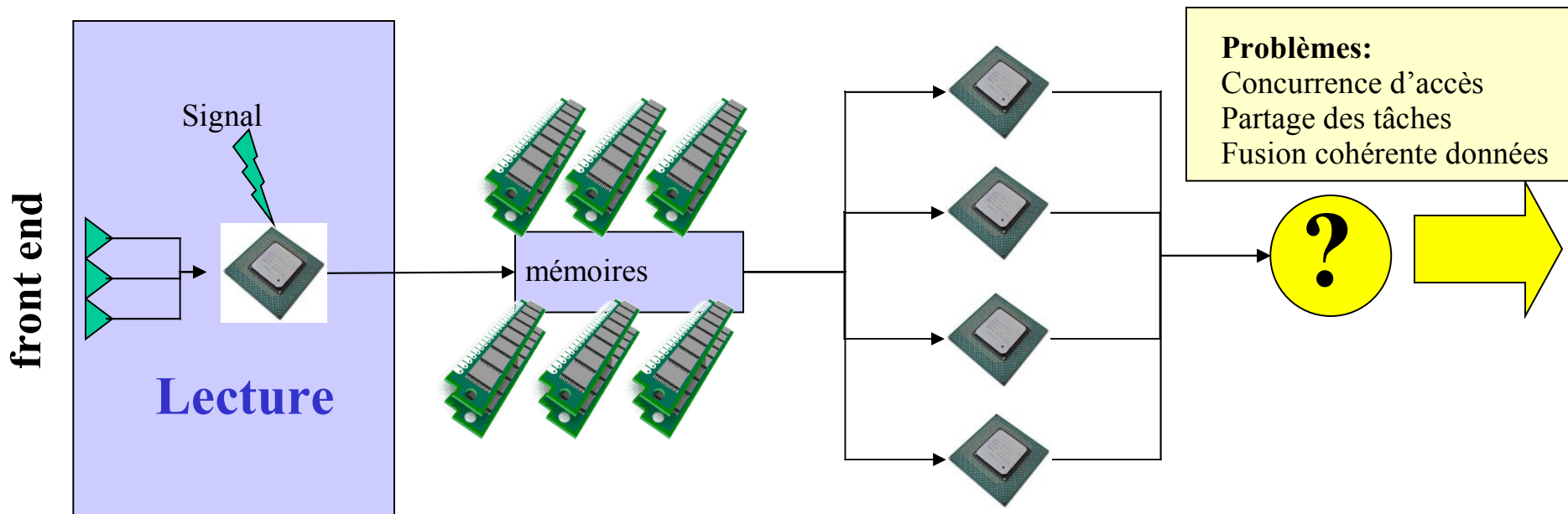


Types d'activation:
Mode interruption
Mode polling temporisé
Polling pur

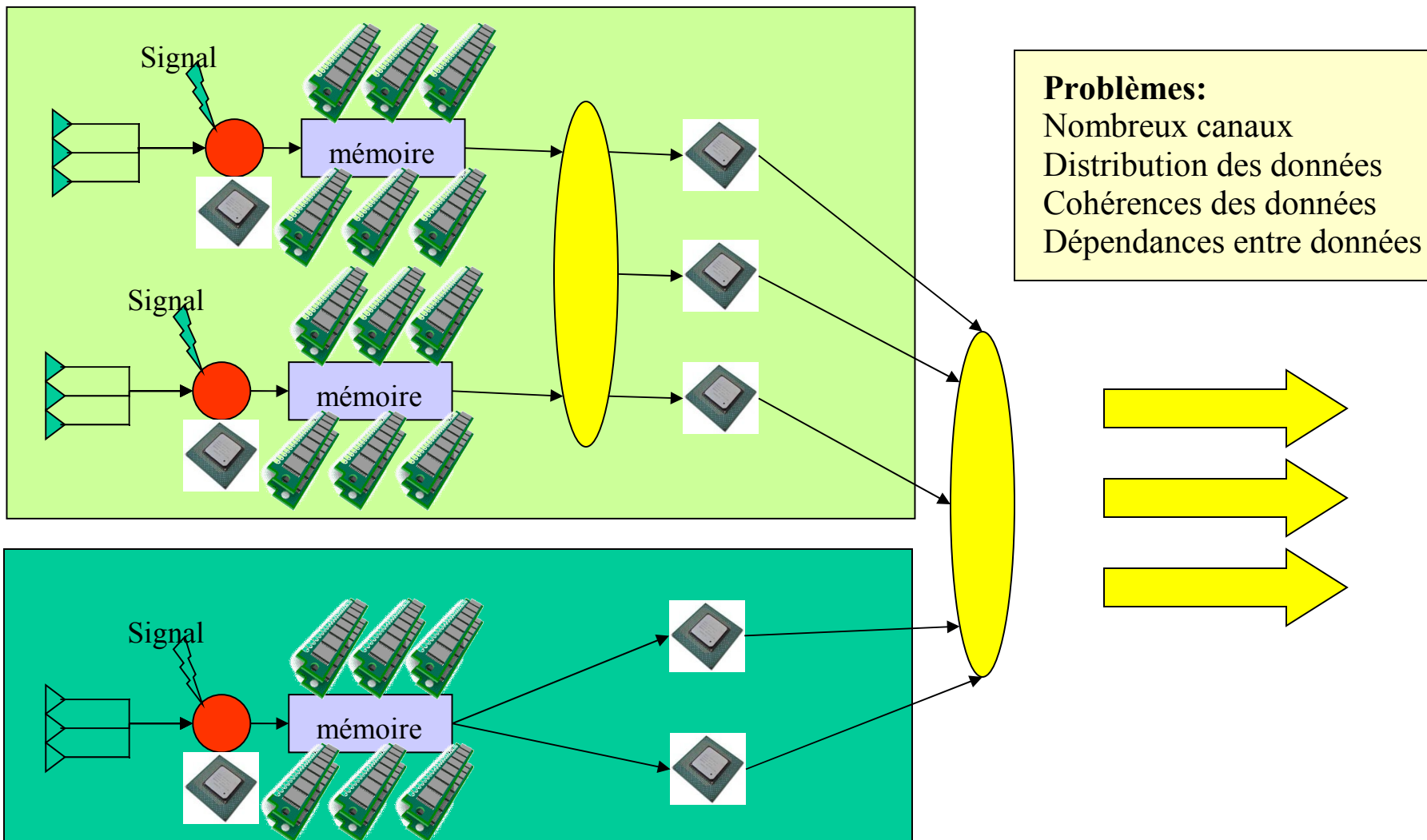
Blocs fonctionnels de base d'une chaîne DAQ

Si flux régulier Et limité par les temps de traitement Alors
augmenter les unités travaillant en parallèle
(exemple moteur DMA pour la lecture, multiprocesseur pour l'analyse)

Si flux irrégulier avec des rafales Alors
augmenter la capacité des tampons (buffers) pour absorber les rafales



Blocs fonctionnels de base d'une chaîne DAQ



Notion d'assembleur d'évènements (Event builder)

Problème: quand il y a plusieurs canaux:

- répartis géographiquement
- des chaînes de traitement de durées variables en parallèle

les délais d'acheminement des données temporellement corrélées (évènements du même évènement de physique) sont très variables selon les canaux.

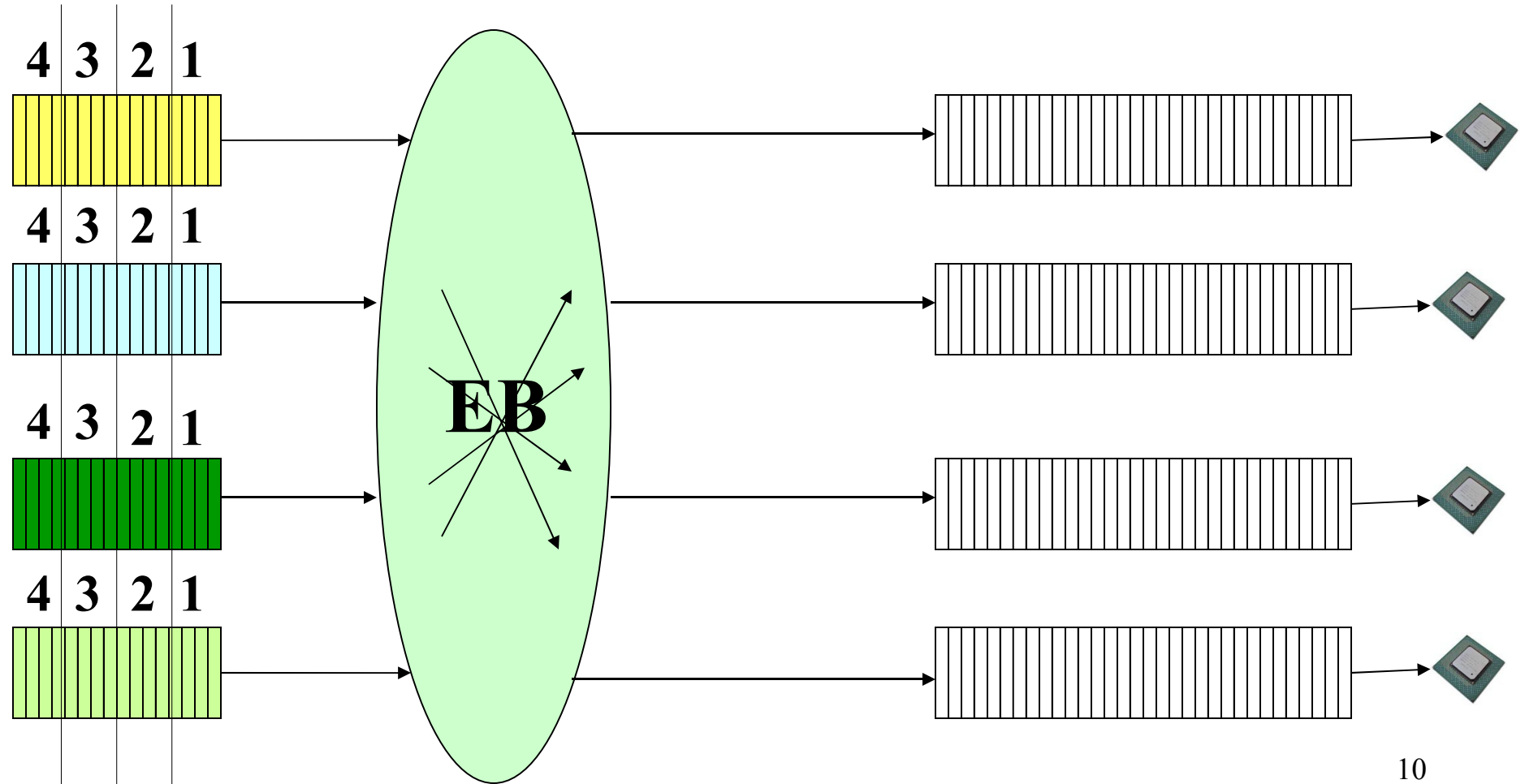
L'event builder collecte et assemble les fragments d'évènements venant de différents canaux en un fragment plus gros de données **appartenant au même évènement**. Le dernier assemblage construit l'évènement complet.

Deux fonctions:

- **assemblage**
- **routage vers plusieurs unités de traitement parallèles**

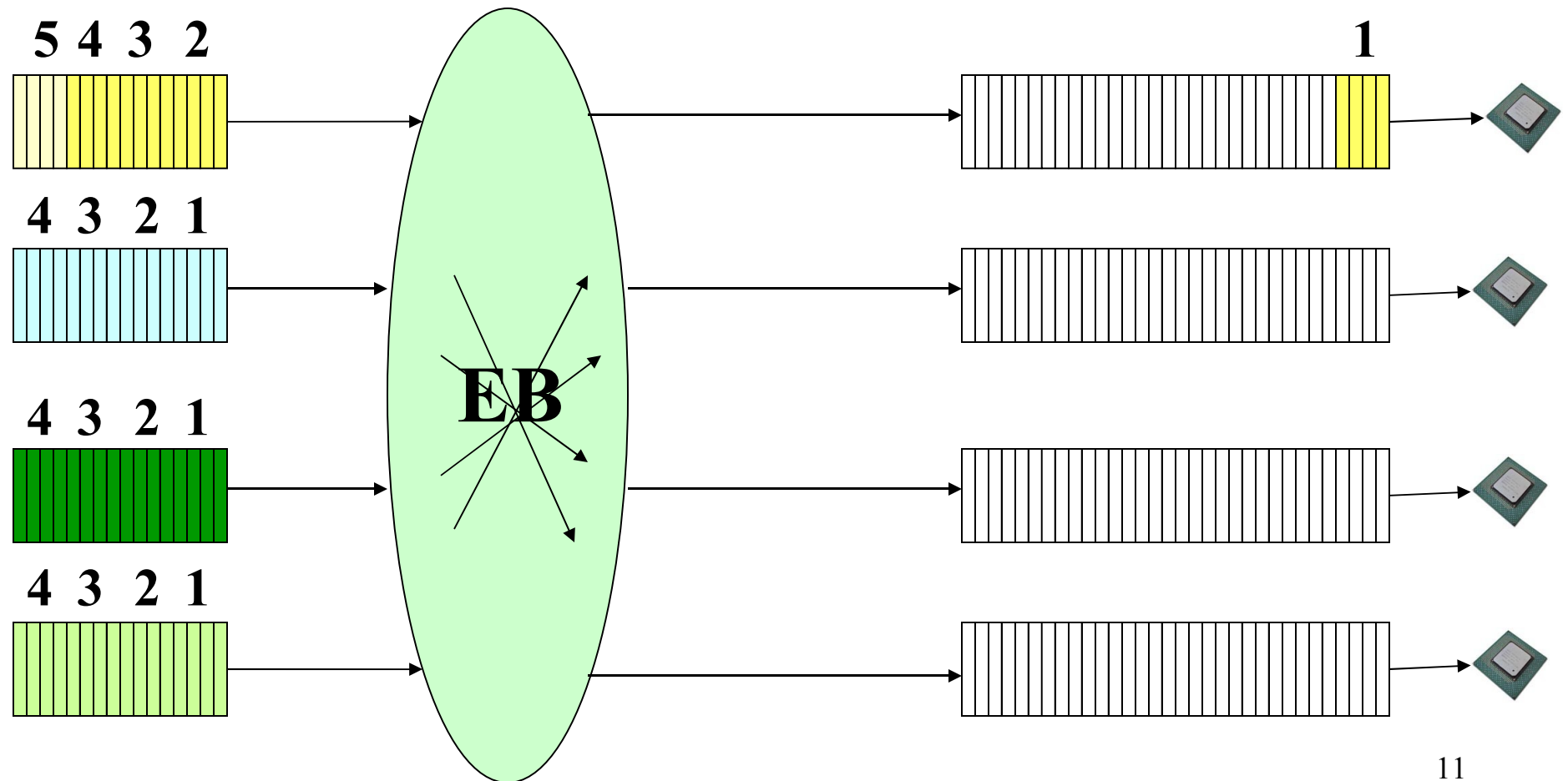
Event builder

Fragments en attente (la logique FIFO temporisée permet de resynchroniser les flux)



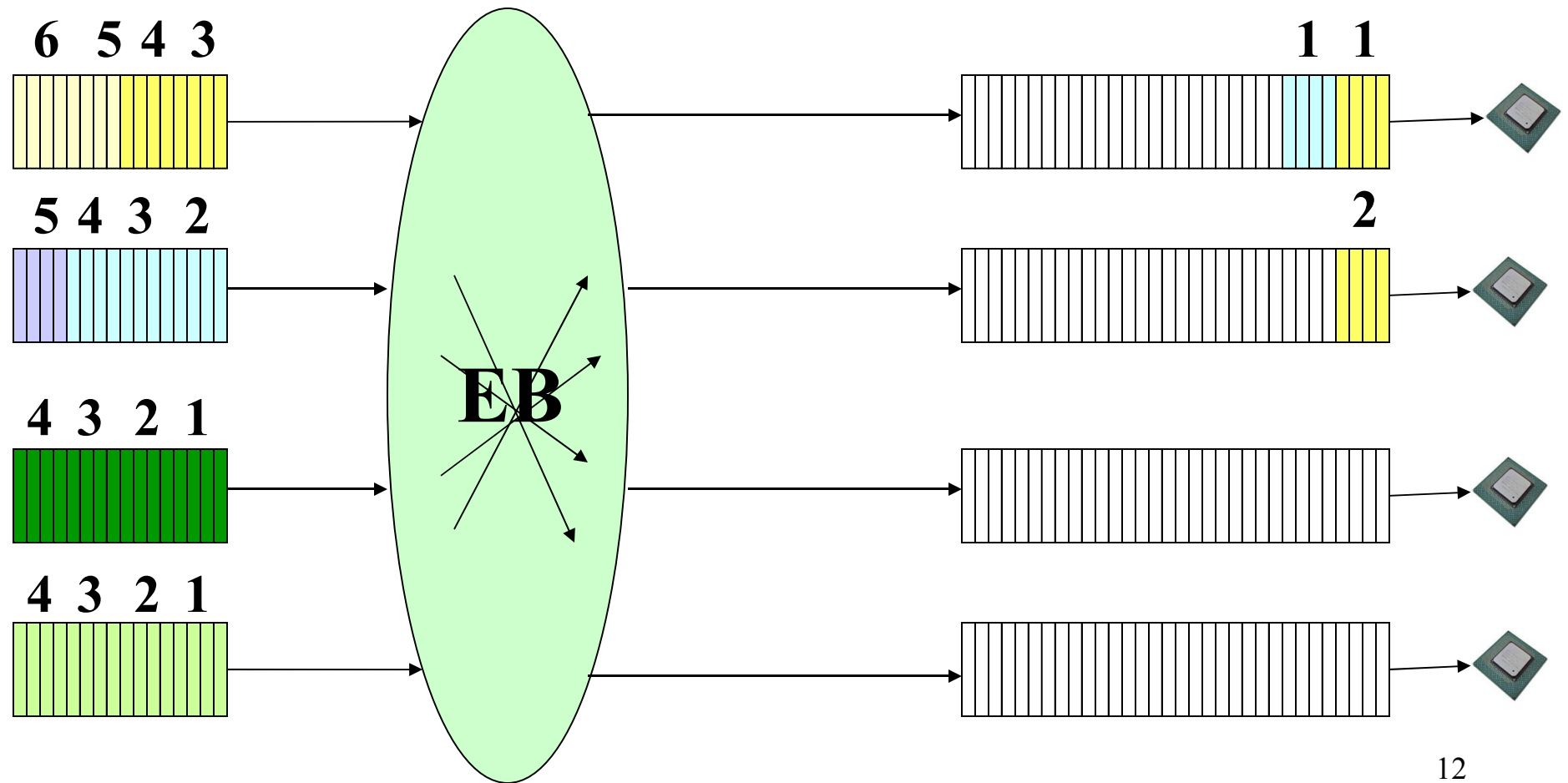
Evt 1 disponible --> affectation à une destination

Event builder



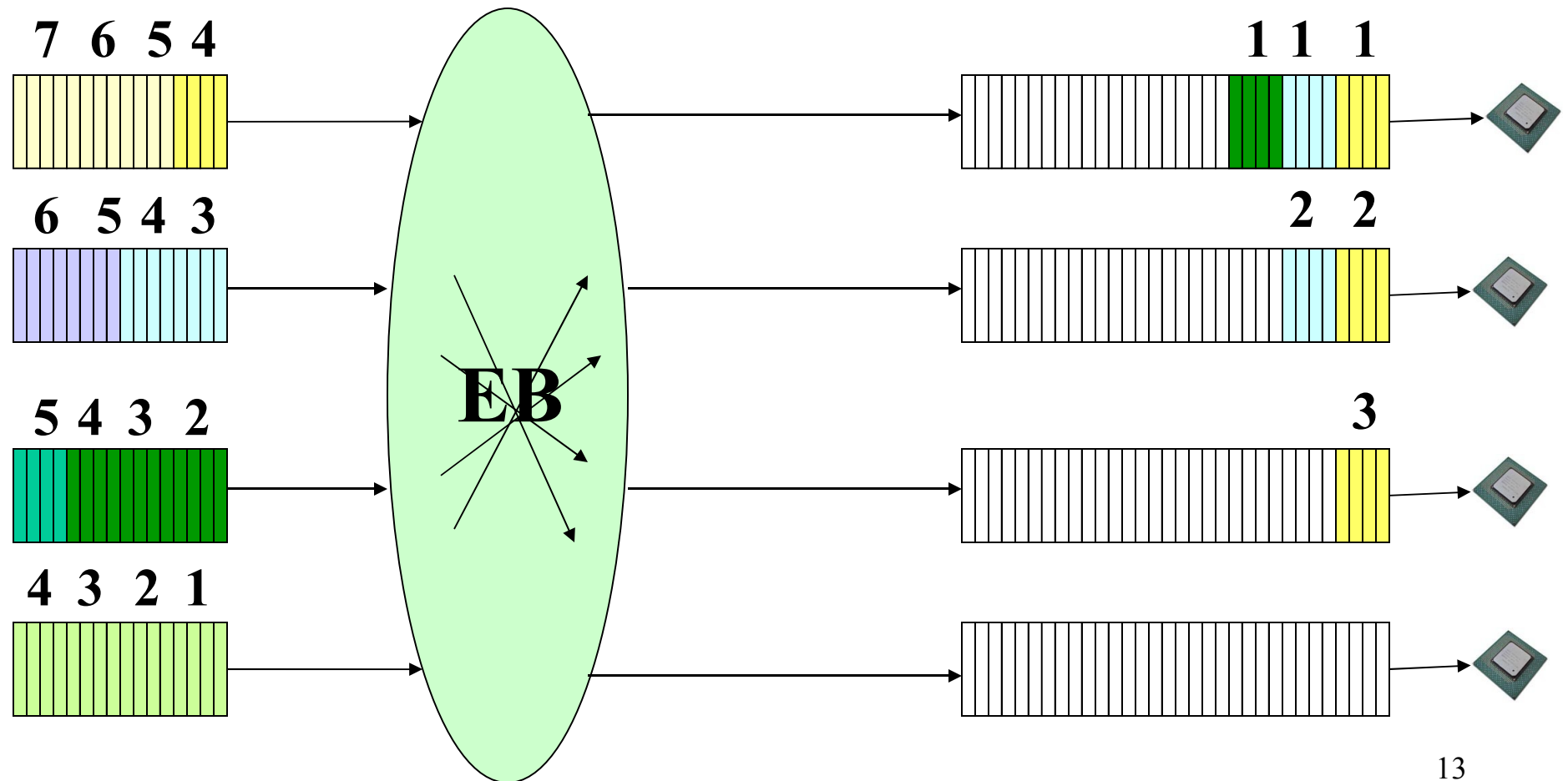
Evt 2 disponible --> affectation à une destination

Event builder



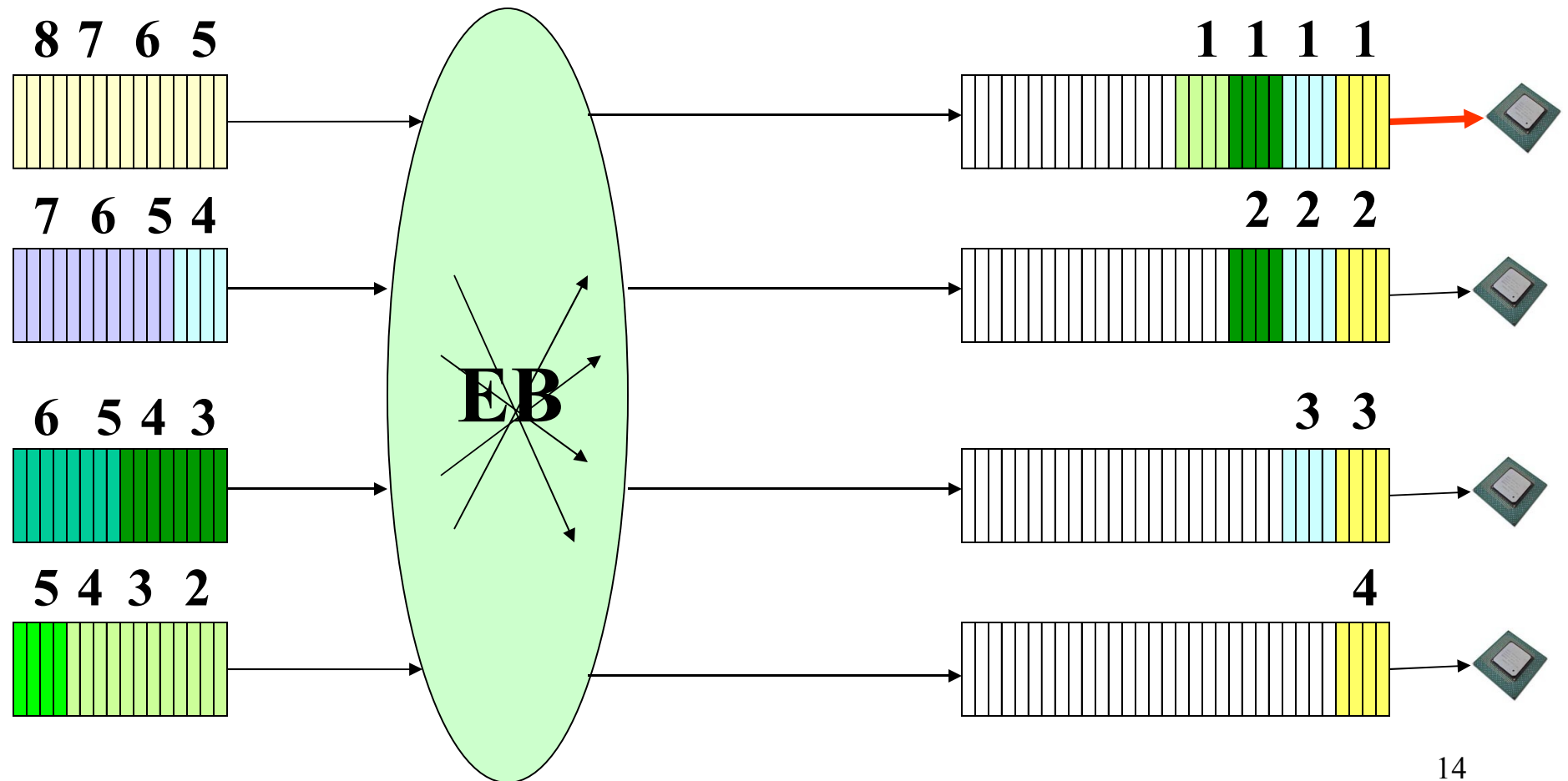
Evt 3 disponible --> affectation à une destination

Event builder



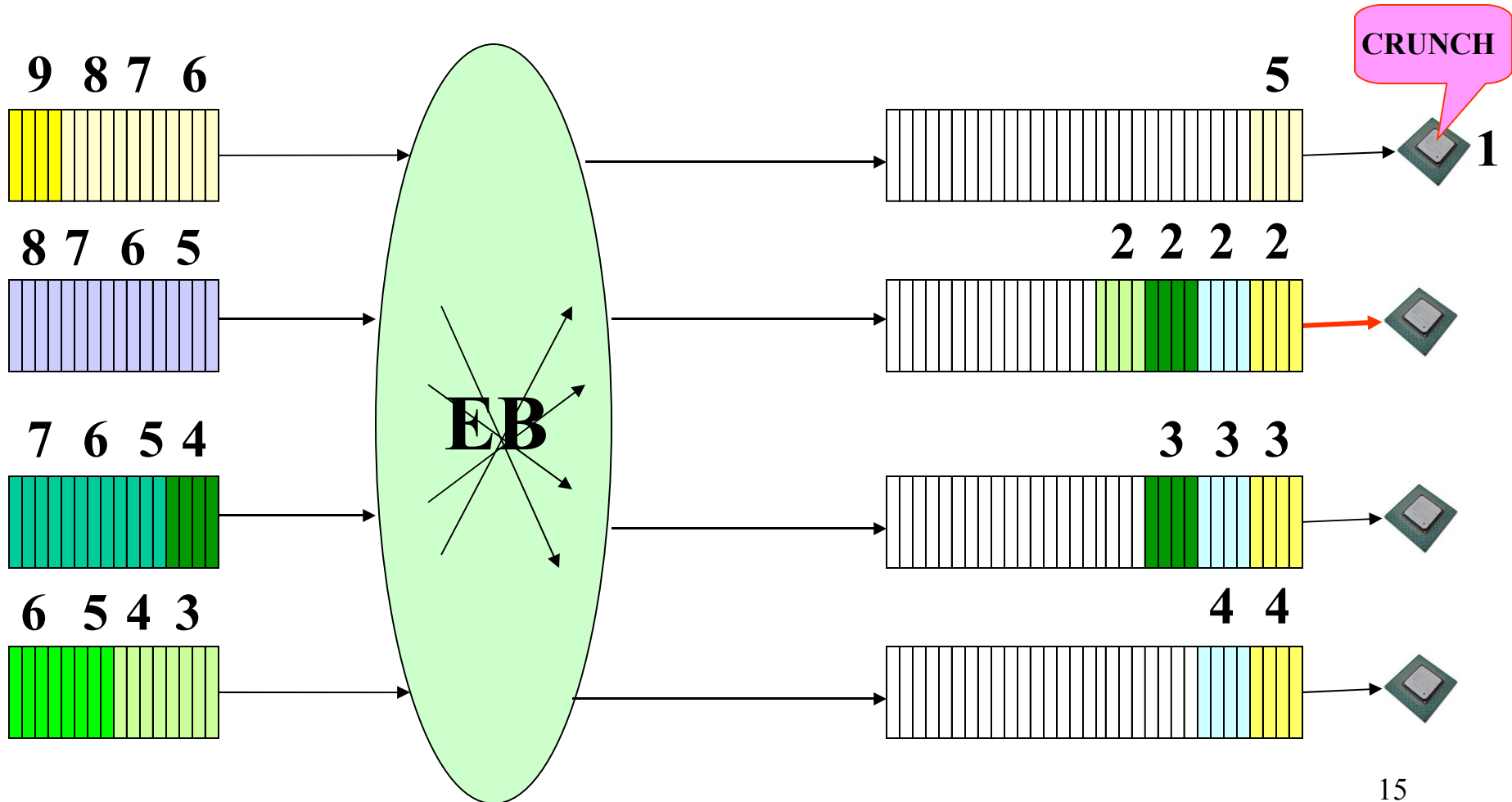
Evt 4 disponible --> affectation à une destination

Event builder



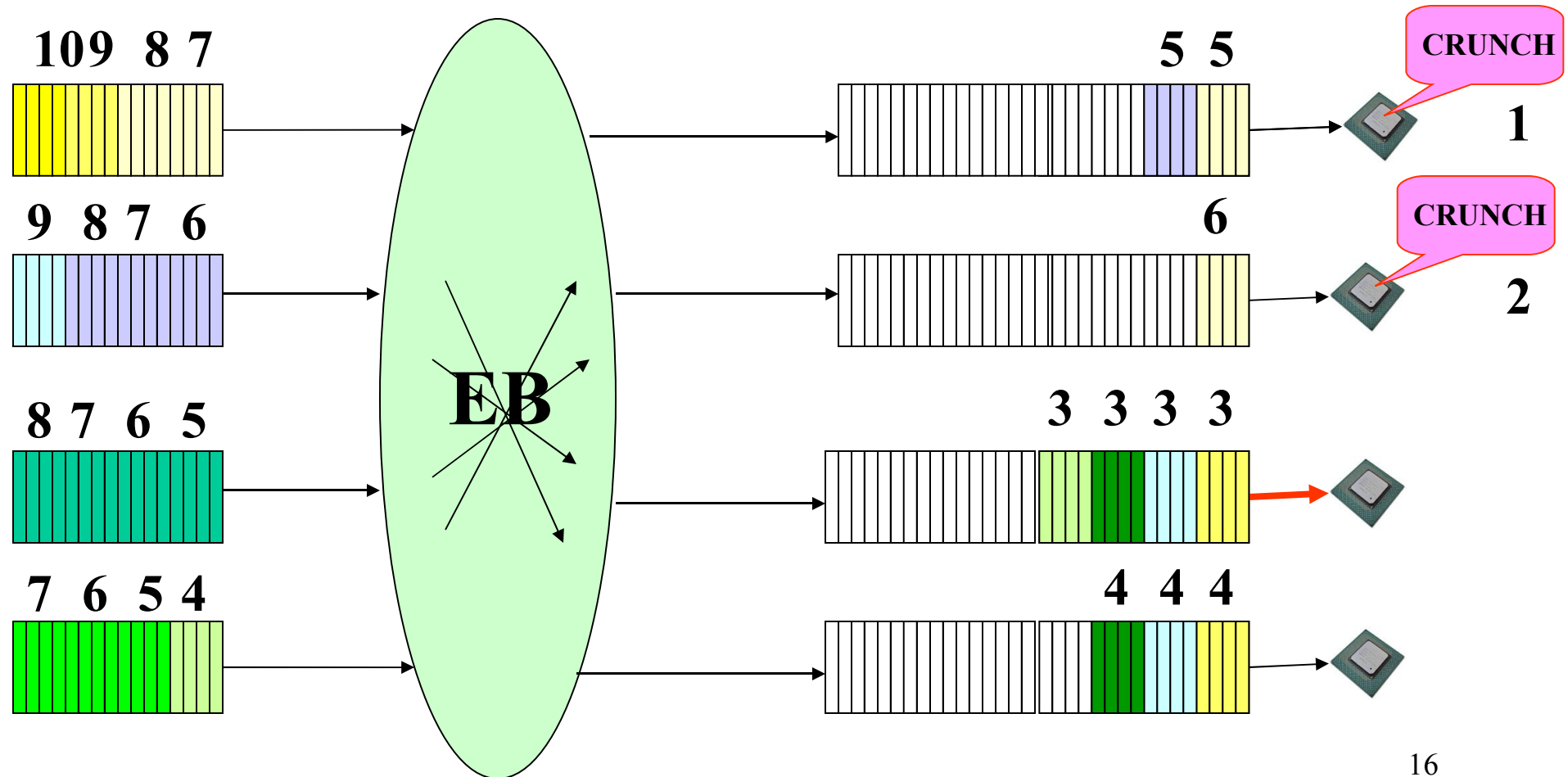
Evt 5 disponible --> affectation à une destination

Event builder



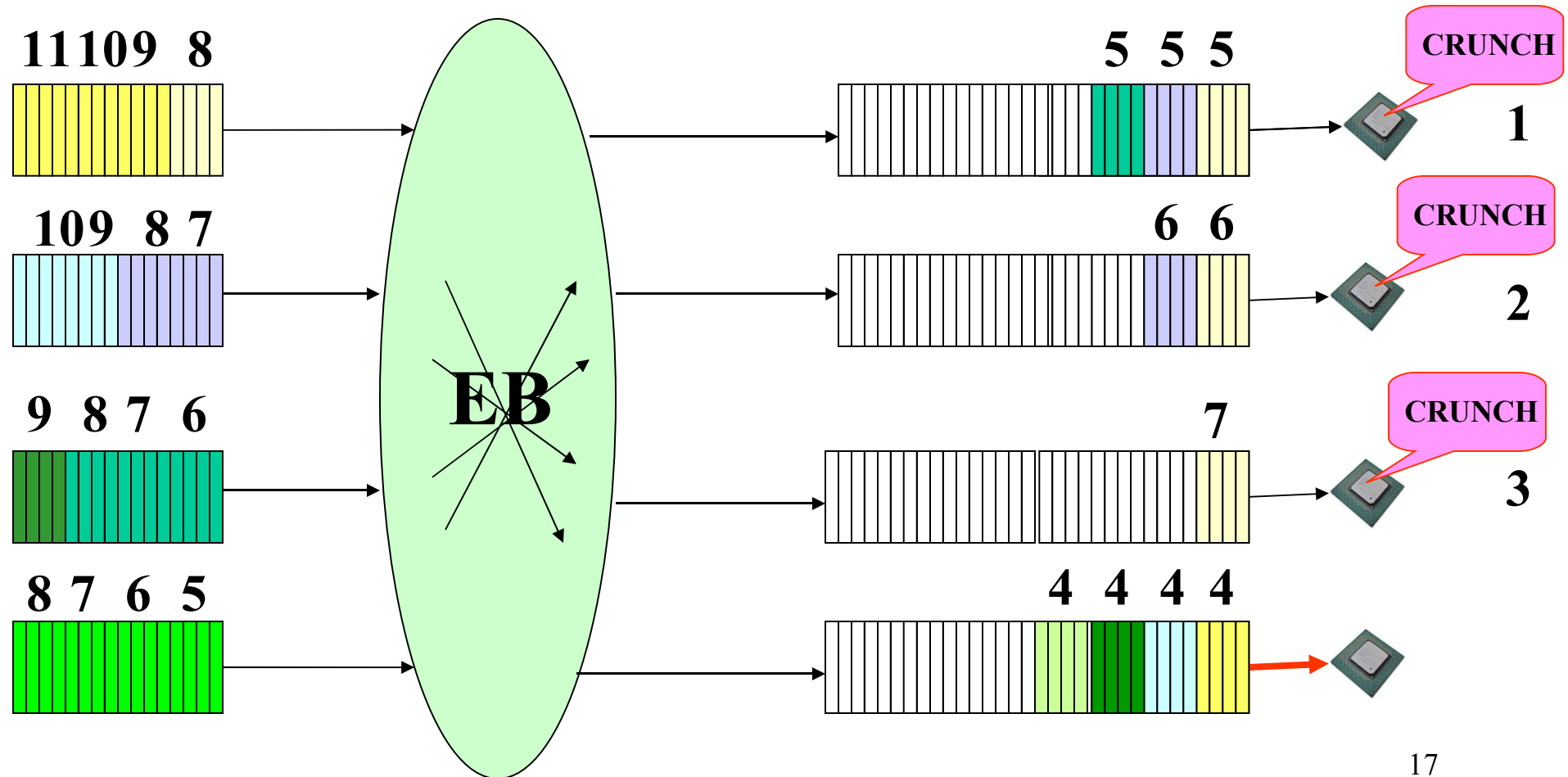
Evt 6 disponible --> affectation à une destination

Event builder



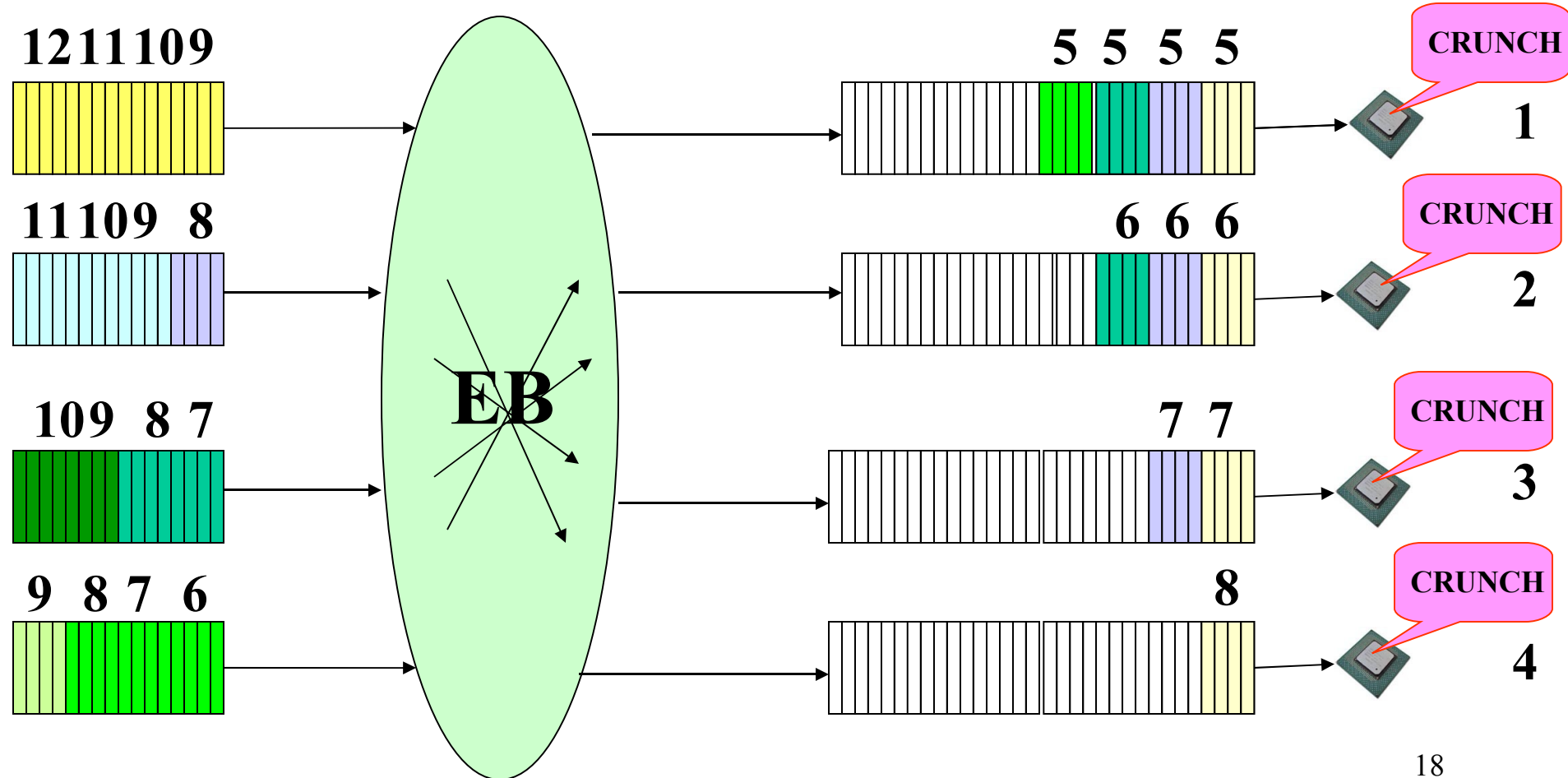
Evt 7 disponible --> affectation à une destination

Event builder



Evt 8 disponible --> affectation à une destination

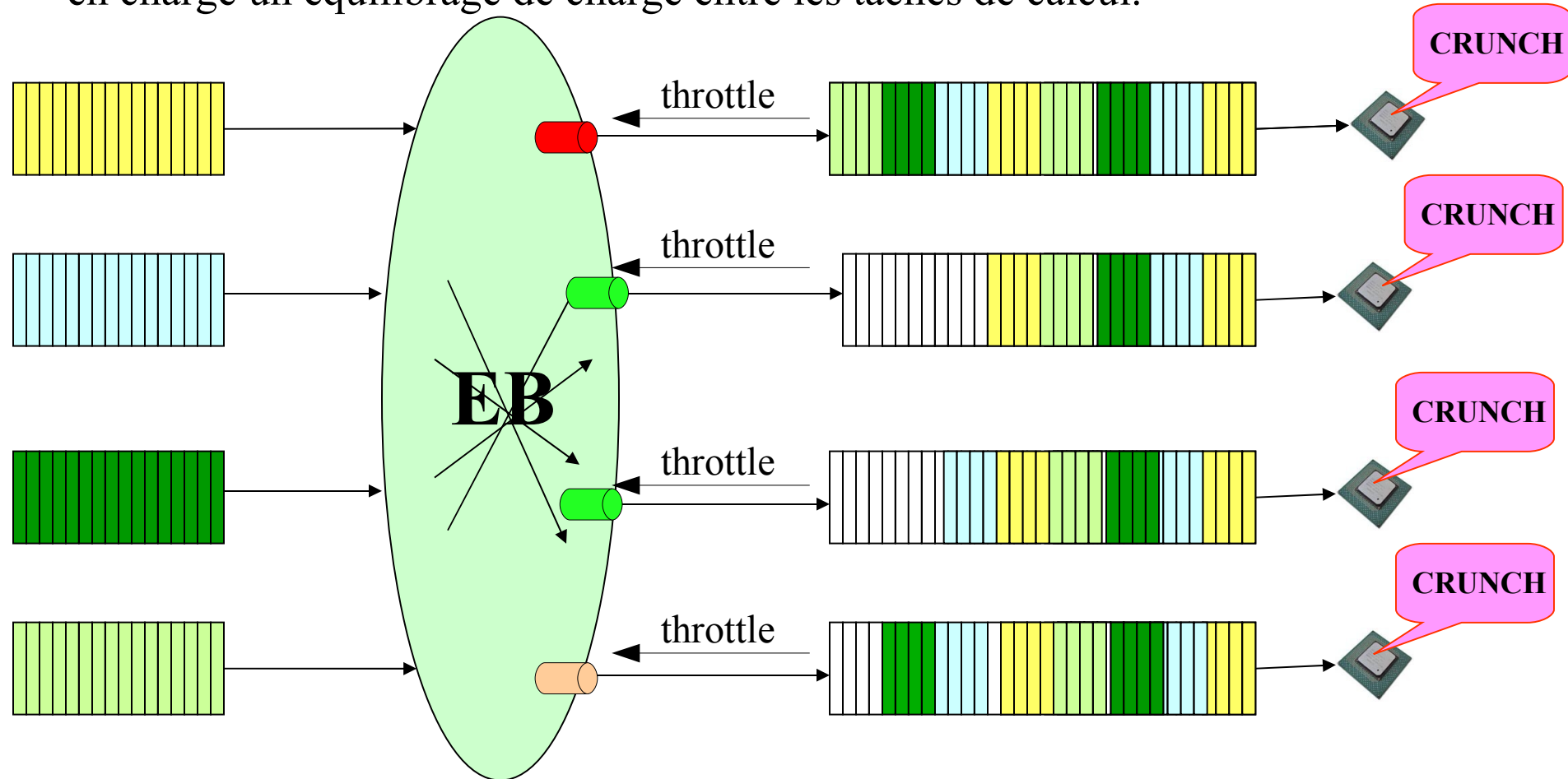
Event builder



Evt 9 disponible --> affectation à une destination

Event builder

Une politique intelligente d'allocation des événements aux destinataires doit assurer en charge un équilibrage de charge entre les tâches de calcul.



On supprime temporairement de la distribution les tâches dont la file d'attente est proche de la saturation jusqu'à ce que cette saturation se résorbe

Notion de sélectionneur (trigger)

Un sélectionneur (trigger) est une unité de traitement qui

sélectionne en ligne (au vol)

les données potentiellement intéressantes pour la physique sur la base de:

- corrélation temporelles ou géographique
- détection de seuil
- forme/reconnaissance de signal

Il assure:

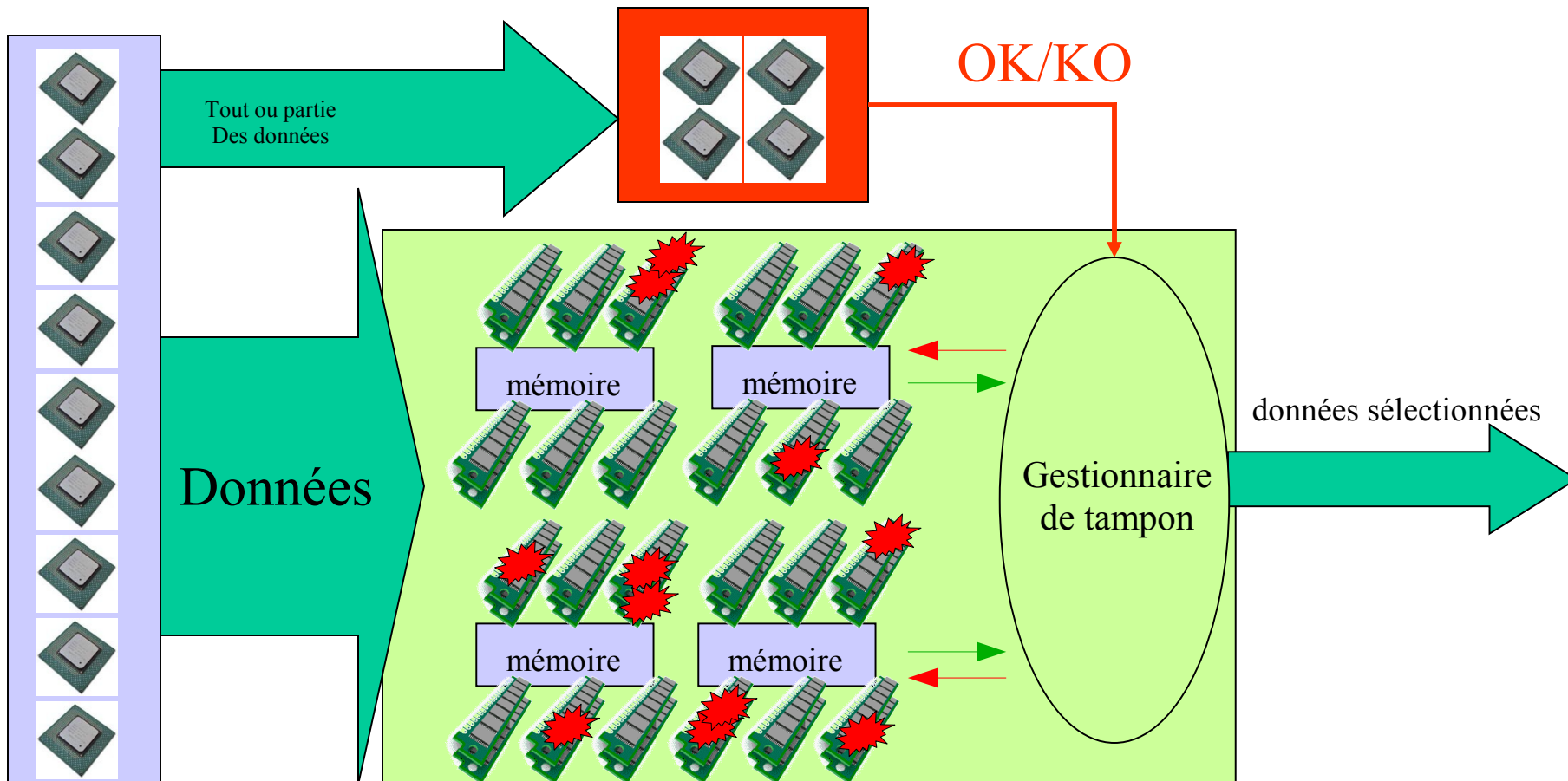
- l'**envoi du signal** de sélection/rejet à d'autres unités de traitement
- l'**ajout de marques** (données d'information ou d'estampillage)

L'objectif est de ne garder et analyser toutes mais que les données potentiellement intéressantes pour une physique particulière (particules et énergies particulières)

Le trigger peut-être implanté en matériel ou en logiciel (rapidité/complexité)

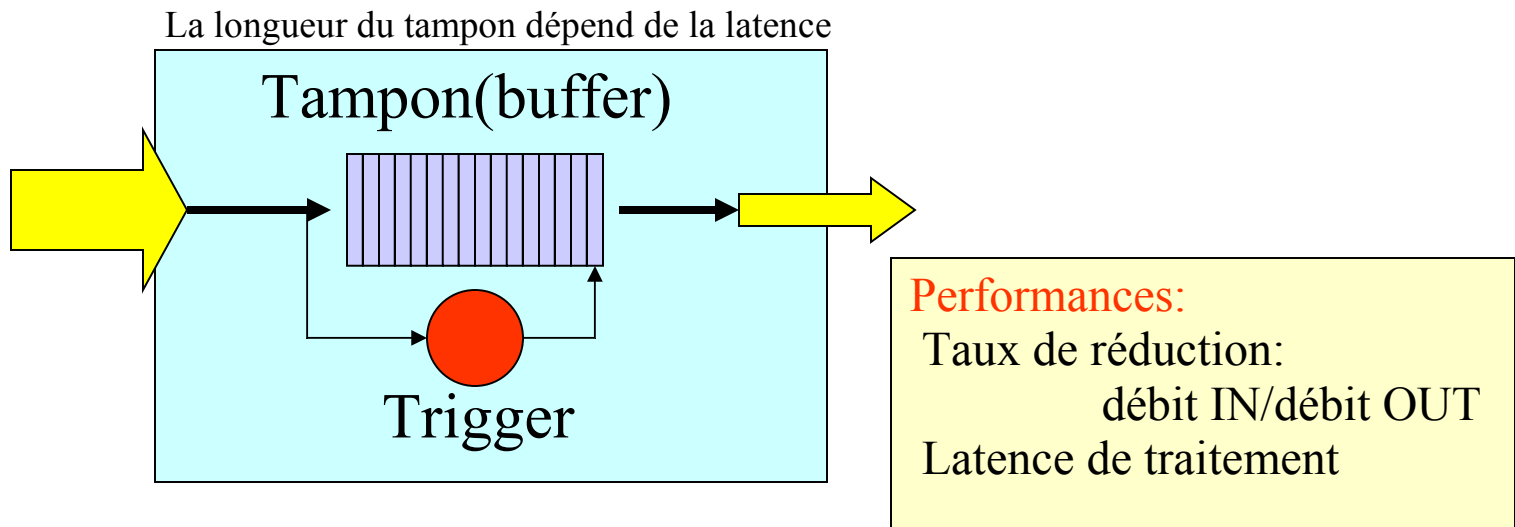
Notion de trigger

Un trigger a un **temps limité** pour prendre une décision: **Sélection/Rejet**.



Notion de sélectionneur (trigger)

Bloc de spécification d'un grand système trigger DAQ:



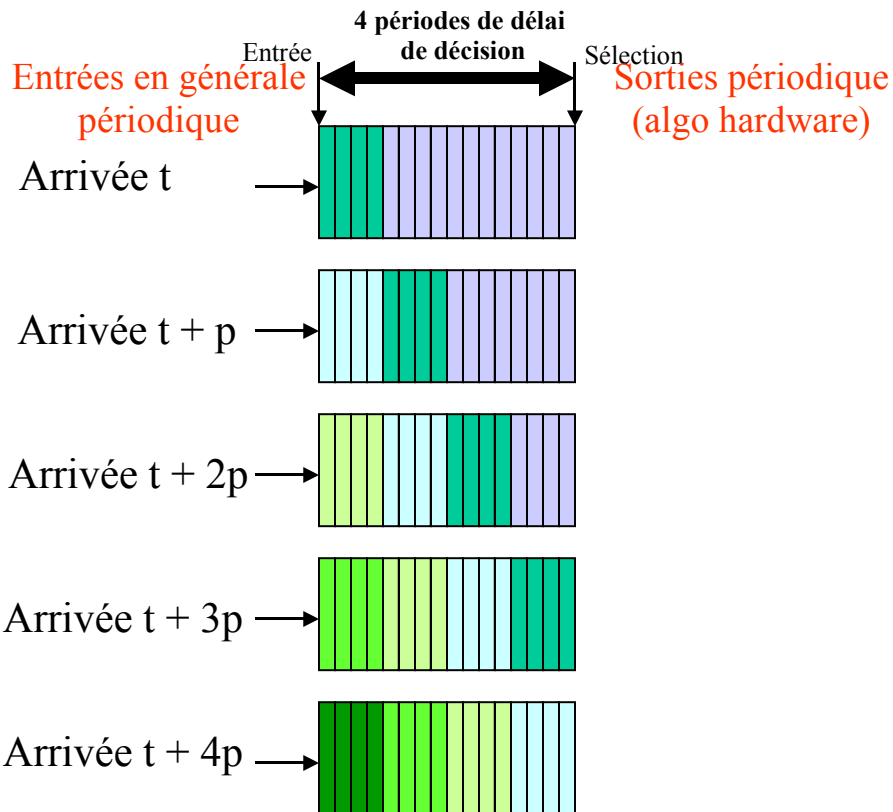
Bloc synchrone ou asynchrone:

Synchrone: le flux d'entrée des données est régulier et le traitement du trigger pulsé par la même horloge avec un délais fixes (électronique numérique)
(sortie toujours asynchrone)

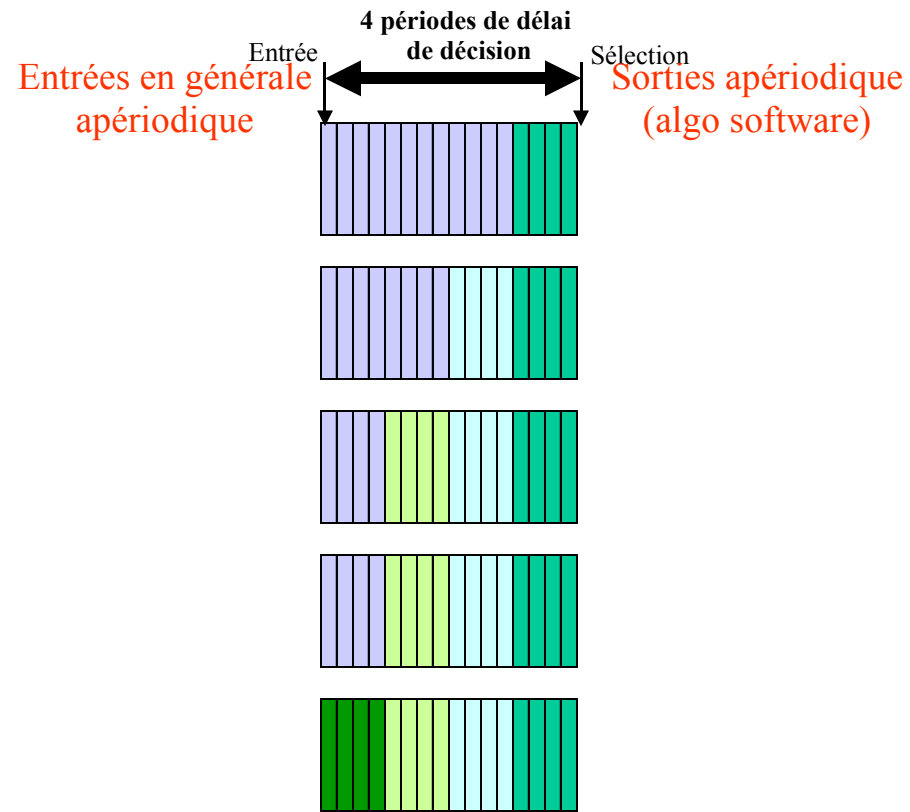
Asynchrone: le flux d'entrée est irrégulier (statistique) et le traitement du trigger variable selon le contenu des données (software)

Notion de sélectionneur (trigger)

Tampon synchrone



Tampon asynchrone

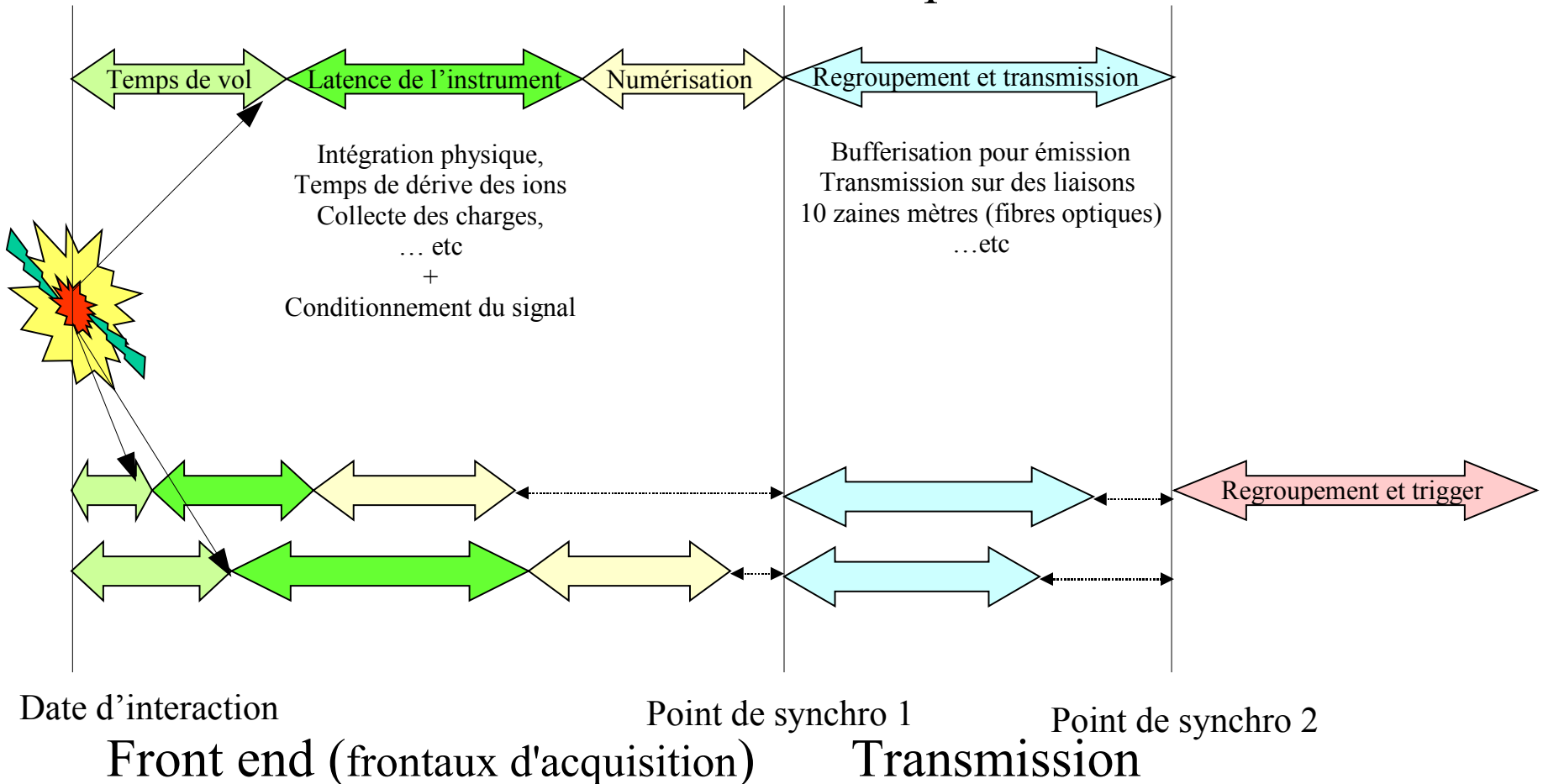


Tampon plein=> risque de perte

Mise en temps - synchronisation

Délais très variable dans les canaux d'acquisition de l'électronique frontale: Mise en temps

Échelle du temps



Identification des données datation - horloge

Deux approches:

- numéro séquence (associé à un numéro de collision sur accélérateur)
- estampille temporelle (horloge commune synchronisée)

Difficulté:

La distribution en parallèle à des composants géographiquement répartis sur une grande échelle d'un signal d'horloge très précis (ns)

Systeme spécifique souvent associé à la diffusion de messages de contrôle courts (une horloge est un cas particulier signal/message)

Éléments de spécifications

Données:

- Nombre de canaux en entrée
- Débit d'entrée sur les canaux
- Lois d'arrivée (périodique ou non, paramètres)

Traitements

- Taux de perte accepté (temps mort)
- Complexité des calculs
- Durée (statistique, stabilité)

Budget disponible

Performances

Contraintes

TDAQ
système

Contraintes d'usages:

- Contraintes d'environnement (hostile ou normal : radiations, satellite, sous-marin ...)
- Contraintes d'exploitation(continue, saisonnière)

Contraintes techniques:

- Évolutions prévues (extensibilité)
- Exigence d'utiliser certains composants ou de respecter certaines normes

Éléments/guides de conception

Choisir les technologies en partant des moins chères:

composants standards du marchés grand public

composants standards de marchés spécialisés avec des spécifications proches

composants maison

Adopter les standards

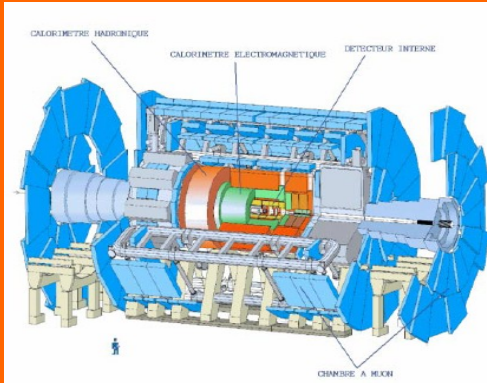
Faire un maximum de chose en logiciel (évolutivité, durée du cycle de réalisation)

Minimiser les développements spécifiques matériel (préférence firmware (FPGA))

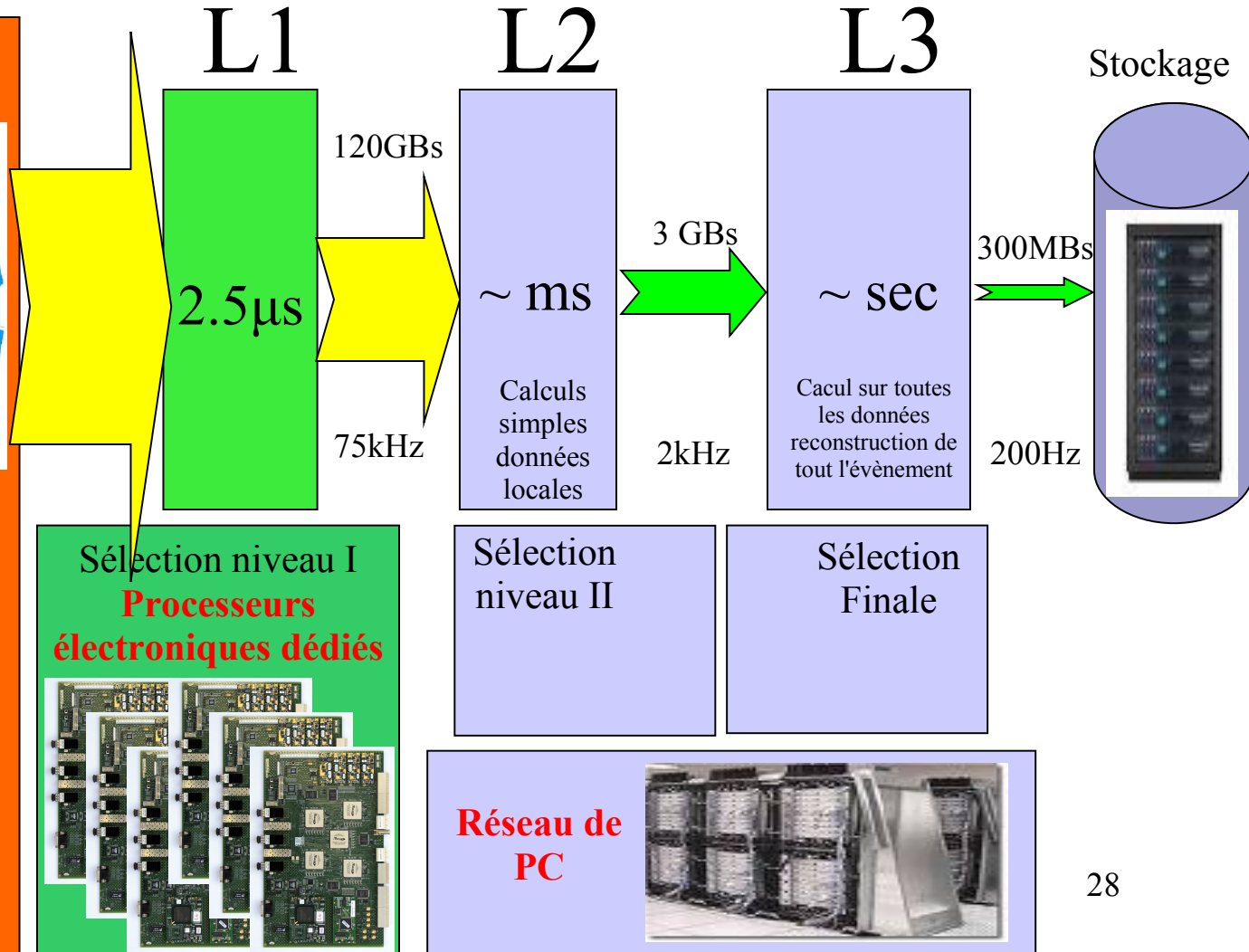
Fractionner le système en sous blocs pour localiser les développements spécifiques/critiques

Architecture d'une grande expérience (ATLAS)

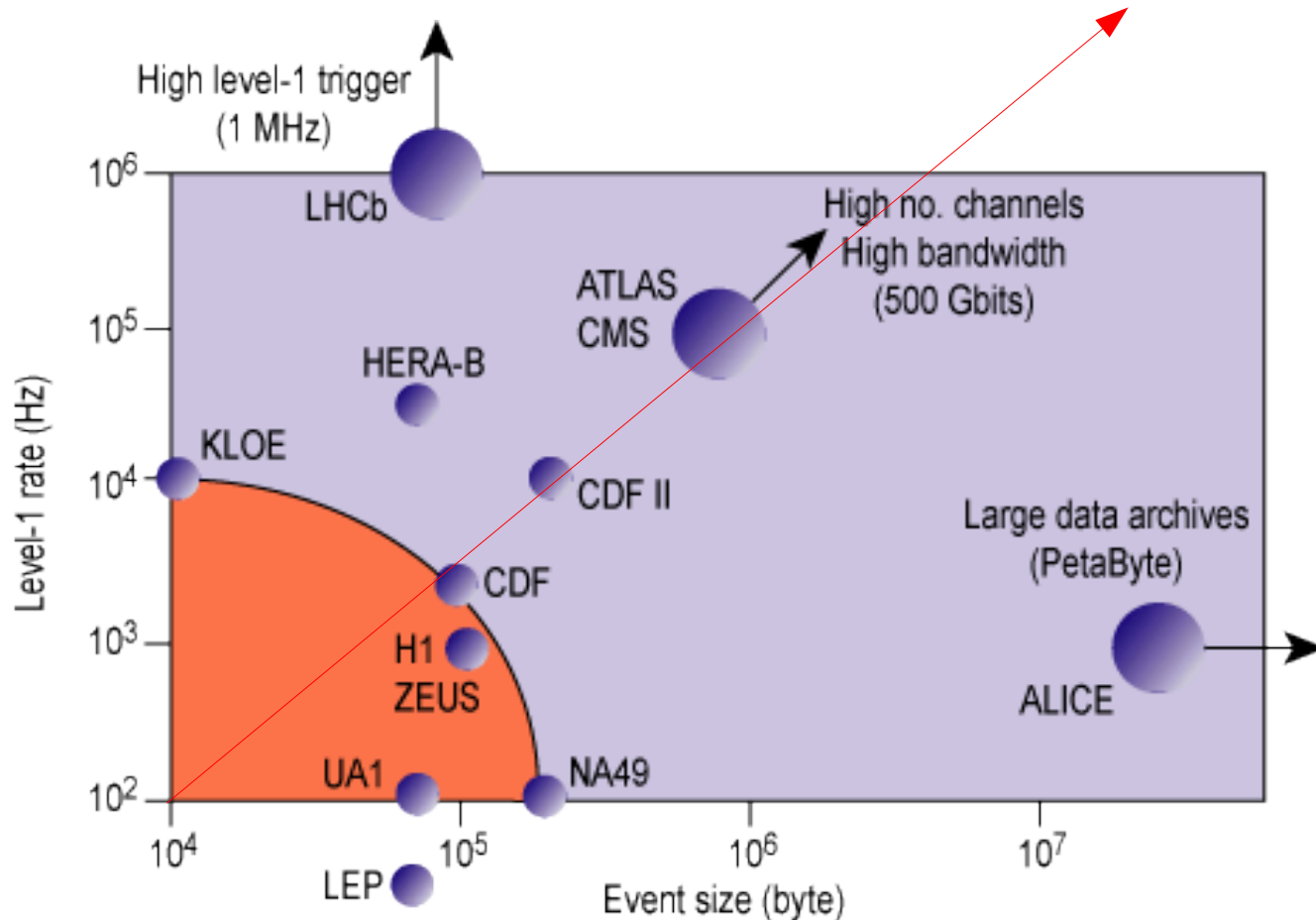
LHC=40MHz



Electronique frontale
et numérisation
(front end)



Performances des grandes expériences



Les bancs de test en physique nucléaire et des particules

Objectifs des bancs de test
Fonctions à fournir aux utilisateurs
Modes de fonctionnement

Les bancs de test en physique nucléaire et des particules: objectifs

- 1-Étude d'un morceau de détecteur
- 2-Étude d'un procédé de détection ou validation
- 3- Étude de prototype de chaîne d'acquisition

Test physique: élément de détection soumis à une source ou en faisceau

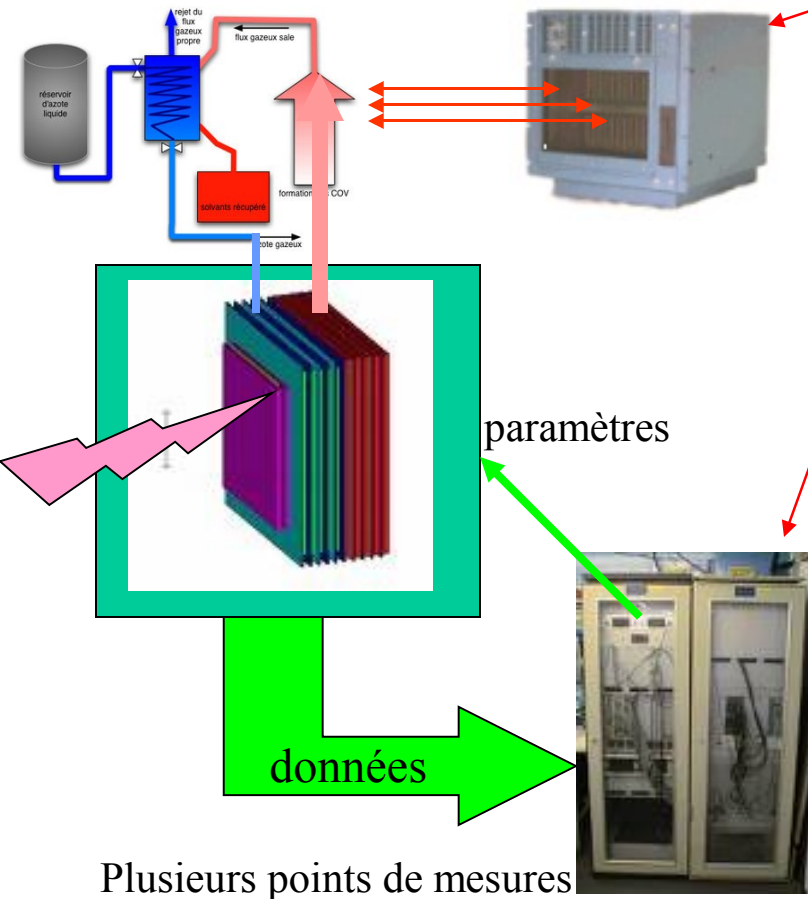
Test technique: chaîne de composants d'acquisition lisant des données prédéfinies

Cela peut-être une mini expérience en soit: exemple les tests en faisceau des sous-détecteurs pour le LHC.

Acquisition de données:

banc de test

contrôle d'environnement/détecteur
(Slow control)



Contrôle d'acquisition **slow control**
(run control)

monitoring



Données sélectionnées



Archivage

Acquisition de données: Calibration

Mode de fonctionnement particulier: la calibration

Caractérisation de l'instrument de mesure: procédé + chaîne DAQ

- bruit de fond (à vide)
- stimulations connues à l'entrée de la chaîne de mesure
 - électrique (sensibilité, linéarité, réponse en fréquence, rampe, échelons temporisés, influence réciproques)
 - optique (laser)

En calibration toutes les données doivent être prises et cela conduit sur les gros détecteurs à des volumes très importants. Ces données sont archivées et exploitées pour corriger les mesures (off-line).

Trigger et acquisition de données

Exemple: l'expérience LHCb

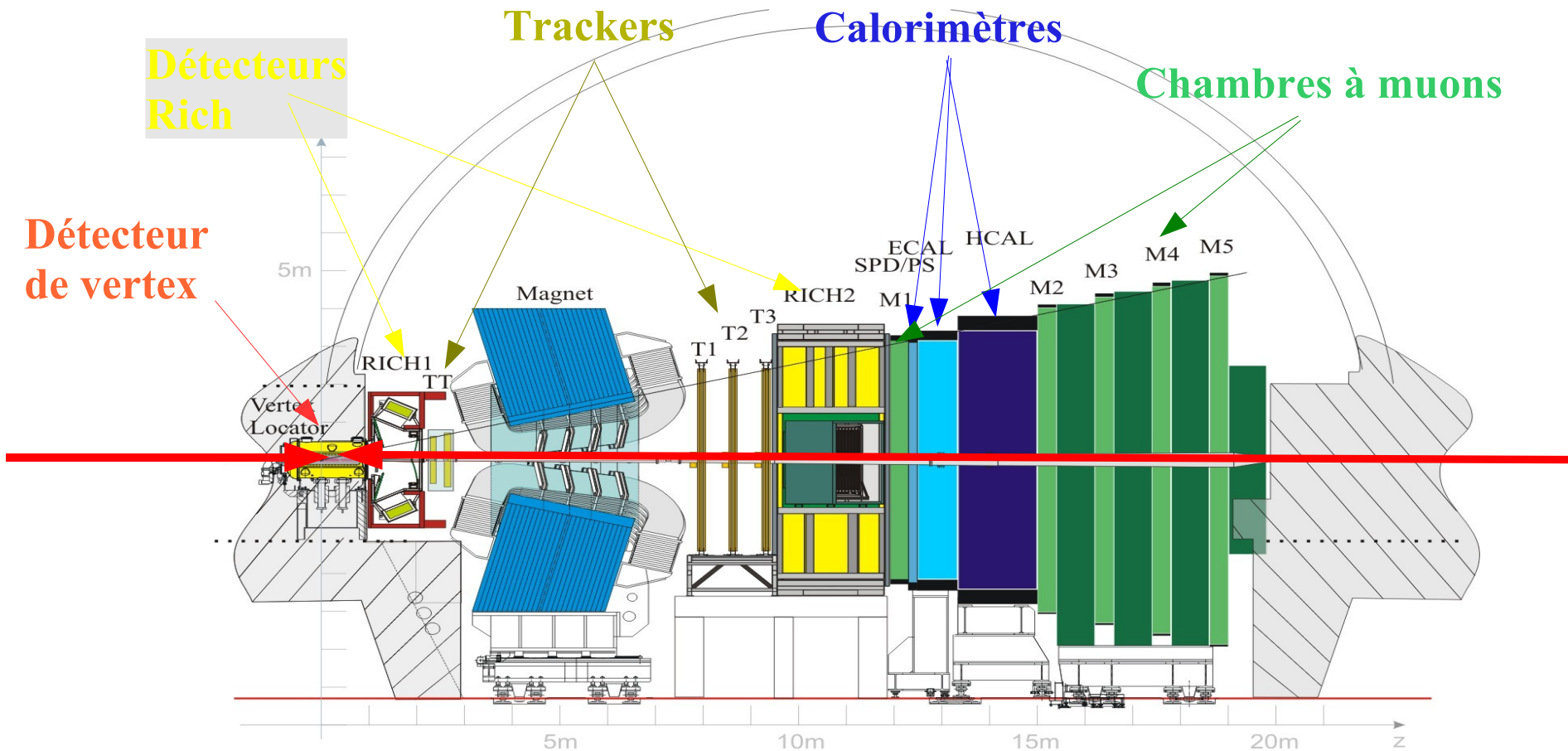
Pierre-Yves Duval (CPPM)





L'expérience LHCb

Physique des mésons B pour l'étude de la violation de CP





Fonctions du système de sélection et d'acquisition

Extraction des données du détecteur

- lire des mesures de capteurs après numérisation **1,2 millions de canaux**

Sélection de celles qui sont intéressantes pour les physiciens

(**100kHz bb_pairs mais interactions intéressantes de l'ordre du Hz**)

- Trouver des critères de sélection et les appliquer au flux de données reçues

Formatage et stockage de ces données

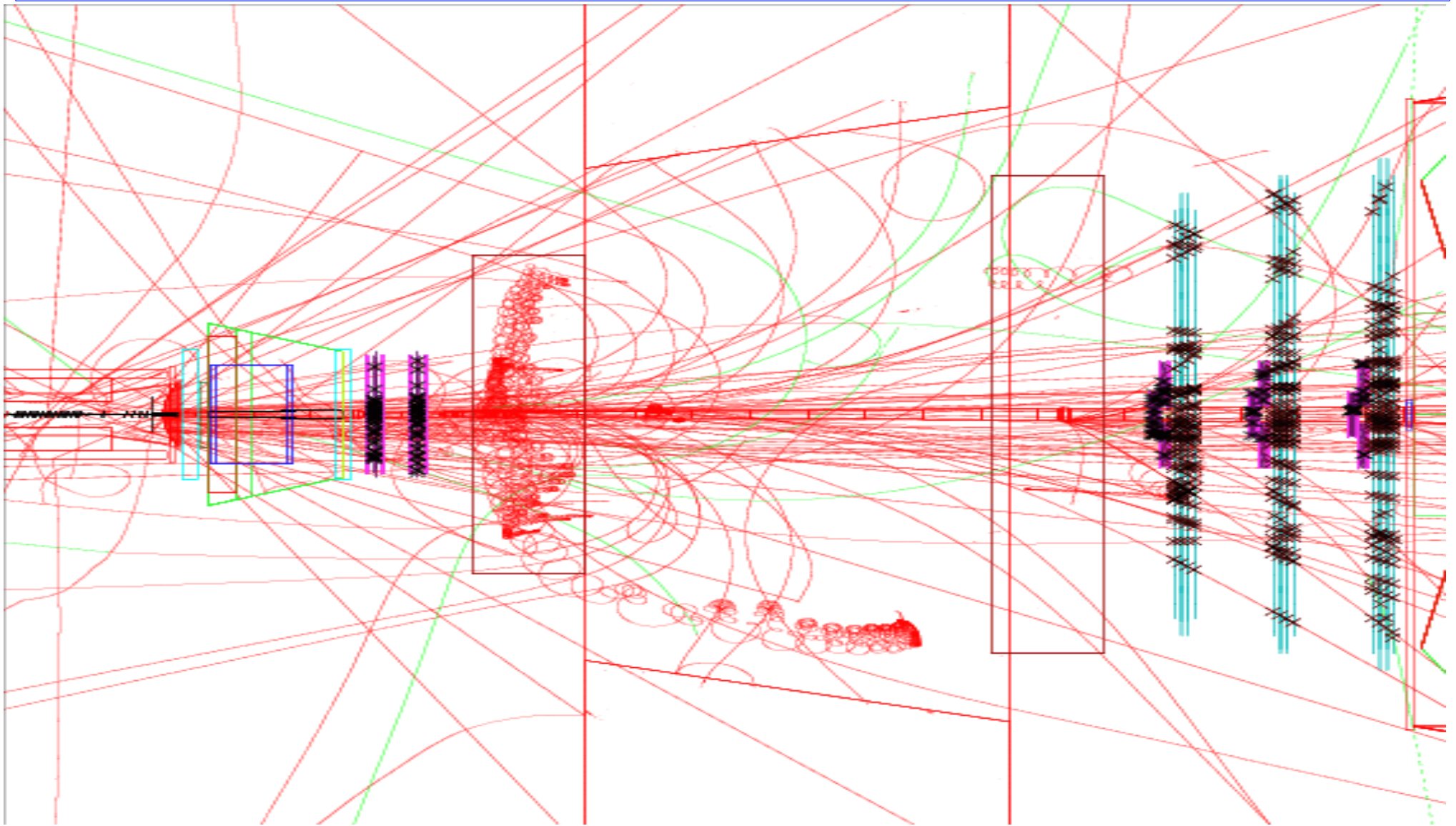
Ces fonctions sont implantées avec une combinaison de:

- composants électroniques numériques (cablés)
- composants informatiques (exécutant des programmes)
- inter-connexions ou réseaux plus ou moins spécialisés

Tendance générale à utiliser de plus en plus de composants commerciaux standards (COTS)



Fonctions du système de sélection et d'acquisition





Spécifications des flux et délais

Modèle de réduction par étage:

Niveau 1 (L0+L1)

Niveau 2 (L2)

Niveau 3 (HLT)

40 MHz
1.2 million canaux

1 Mhz
40000 canaux
50 Gb/s

40 KHz
4 GB/s

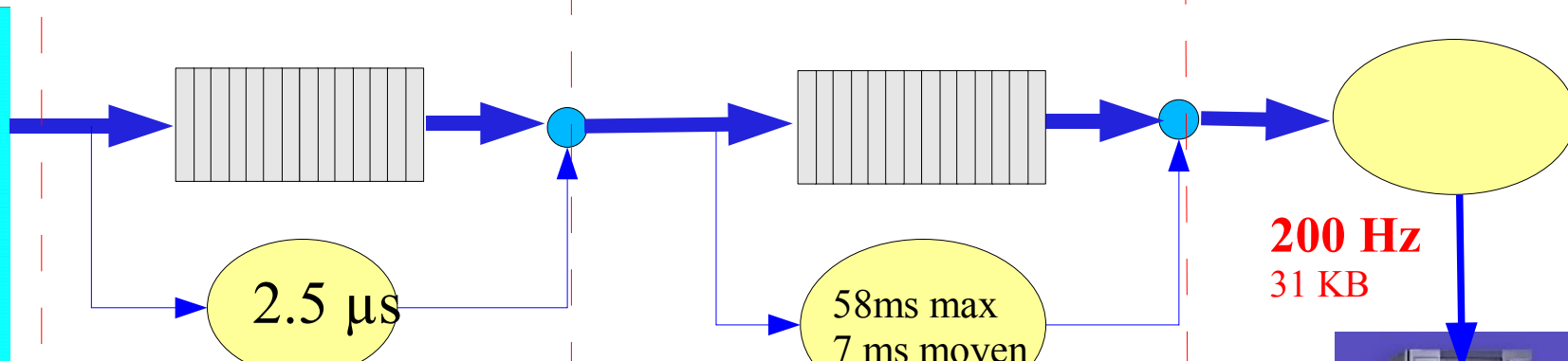
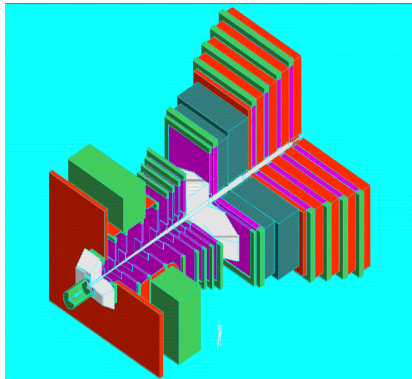
200 Hz
31 KB

2.5 μ s

58ms max
7 ms moyen

Synchrone à 40MHz
Matériel

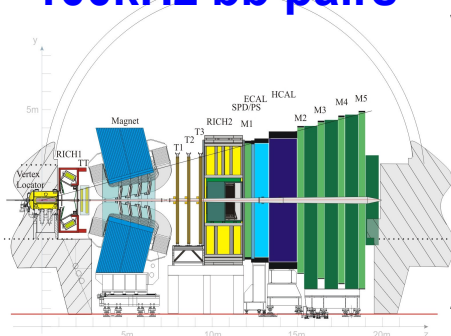
Assynchrone
Ferme de calcul





Vue générale de l'acquisition

Collisions visibles
12 MHz
100kHz bb pairs



1.2 million canaux

Trigger Matériel (L0)

“hauts pt et énergie déposée” calorimètre & chambres à muons
rejet des évènements surchargés

Trigger Logiciel (High Level Trigger)

HLT niveau II: produits partiels de désintégration de B

HLT niveau III: reconstruction totale des traces de désintégration de B

L0+L1

2.5 μ s
Trigger L0
et
compression

Lecture du détecteur
1 MHz

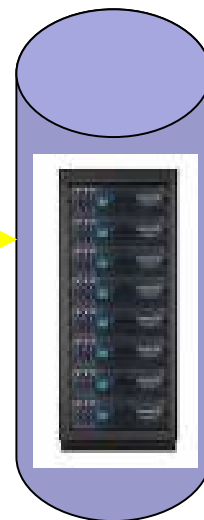
~50Gbs

HLT(L2+L3)

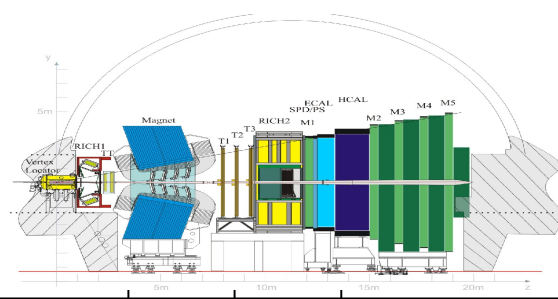
~ ms Simple processing on <u>local</u> data	~ sec Global processing full event reconstruction
--	--

2kHz

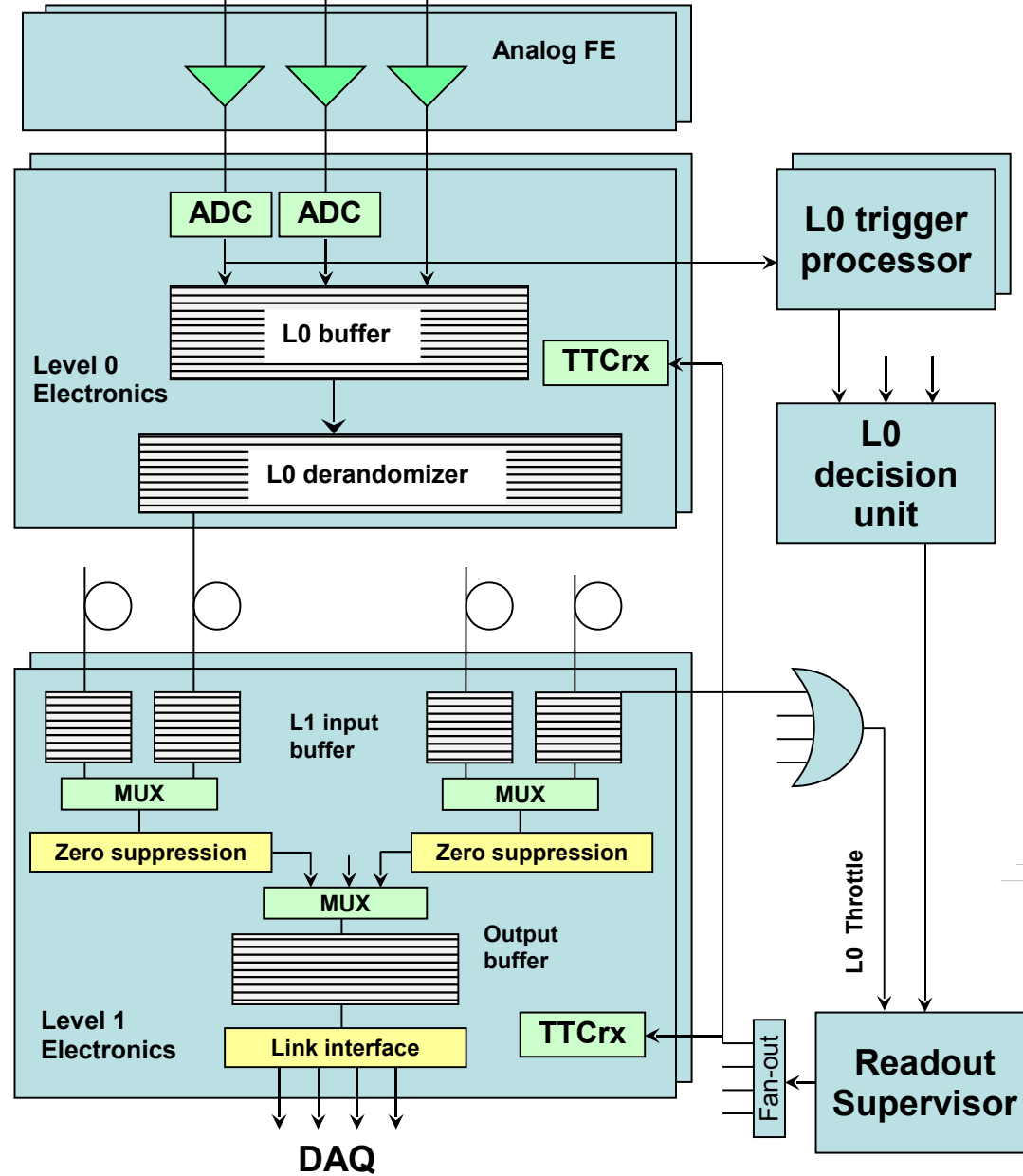
Stockage



Event building



Architecture fonctionnelle de sélection L0 et L1



Sélection au niveau L0

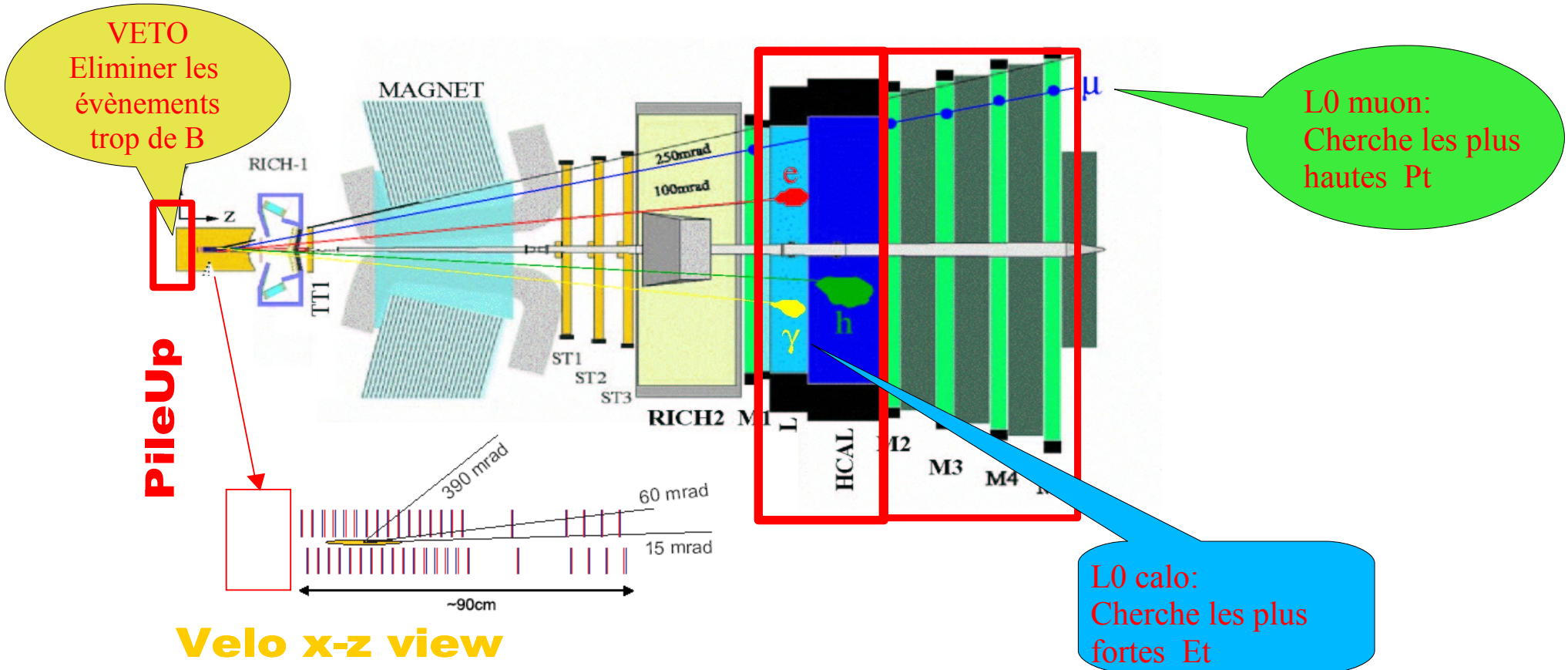
Données de 3 détecteurs avec 3 L0 triggers locaux: calorimètres, muons et veto d'empilement

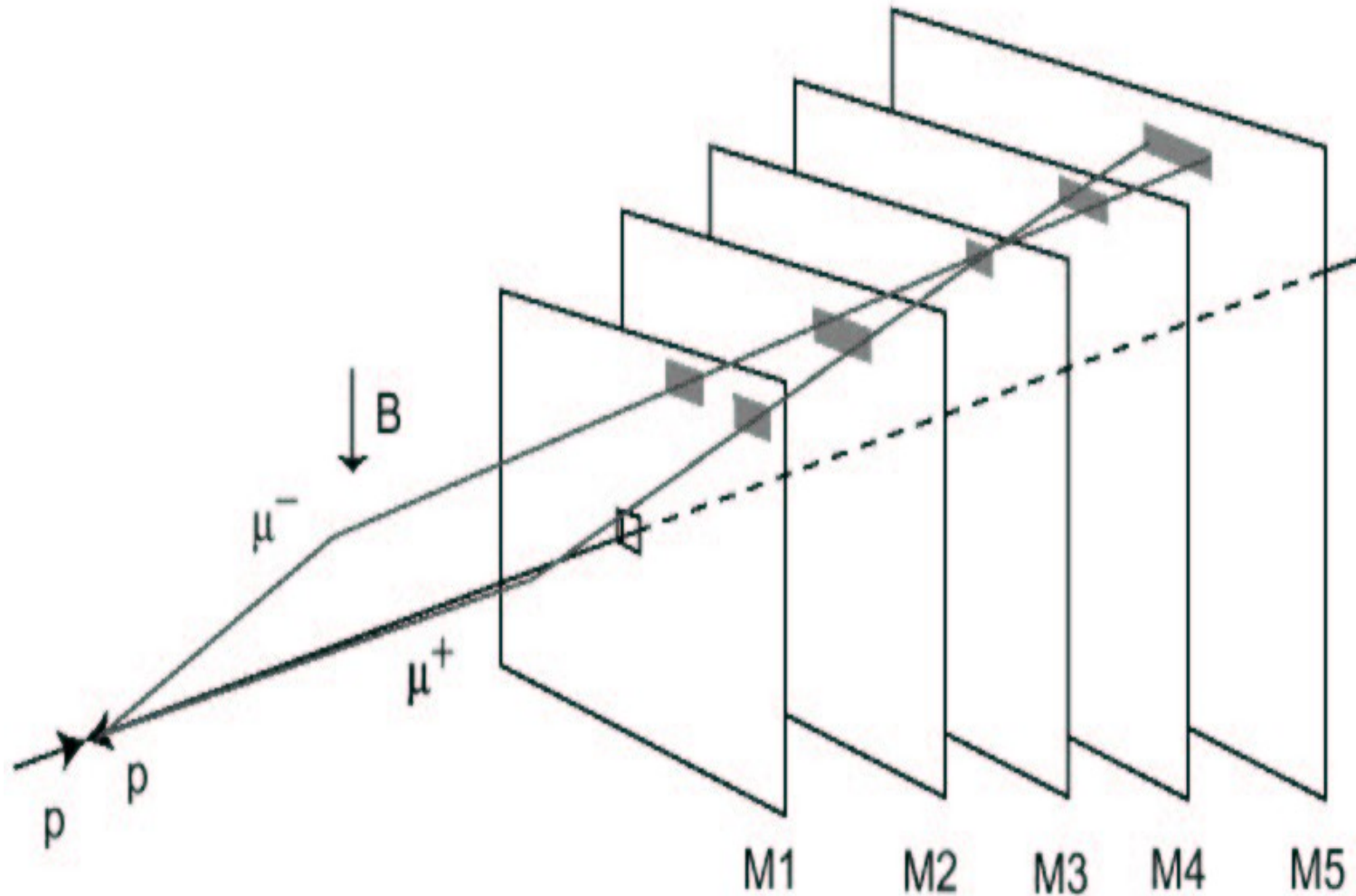
Décision finale par l'unité de décision L0 (règles paramétrables):

Exemple:

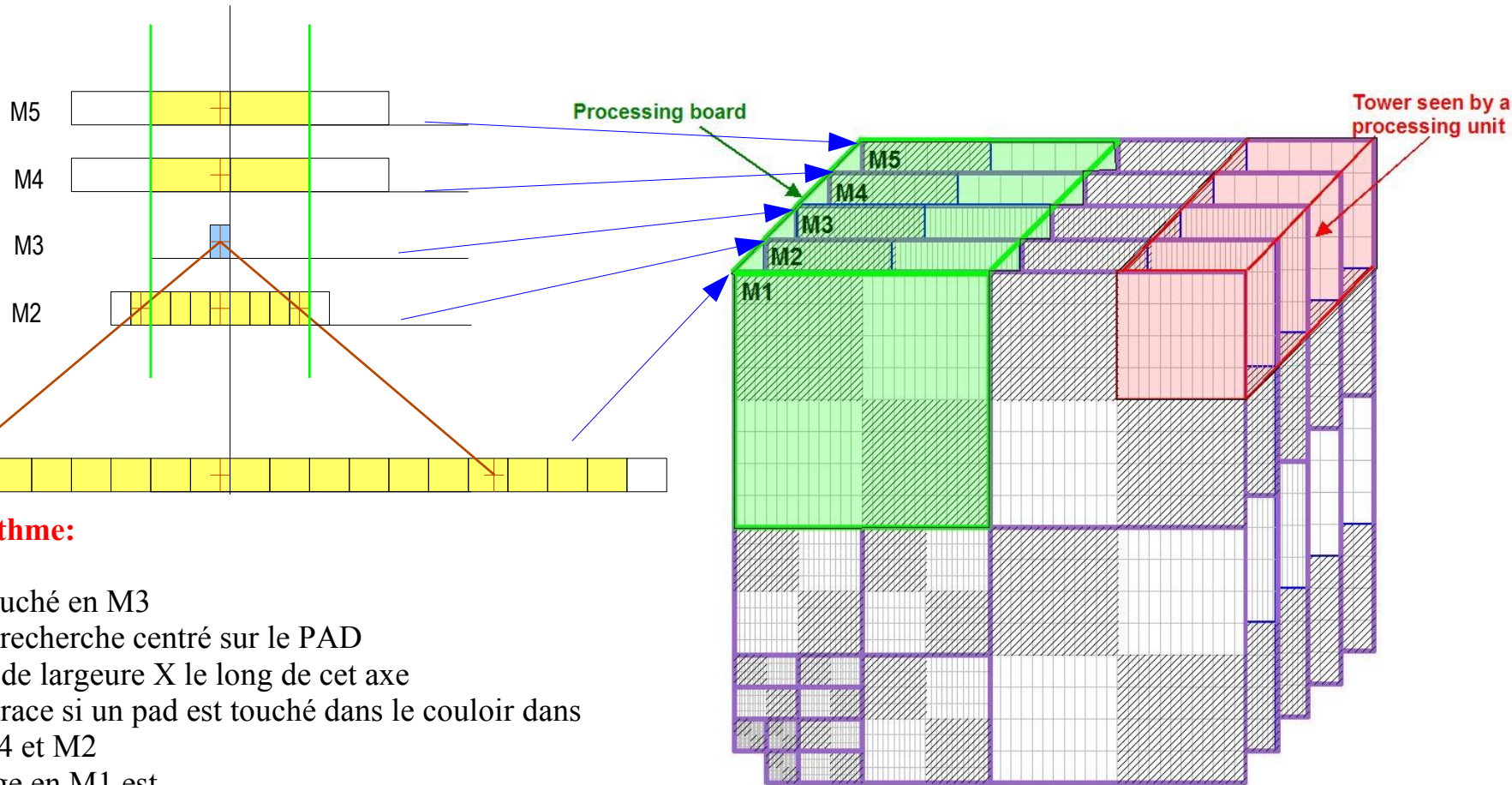
SI (candidats avec E_T OU $P_T >$ seuil) ET (une seul interaction BB)

ALORS selection de l'évènement





Les trajectoires des muons et les pads de détection



Principe de l'algorithme:

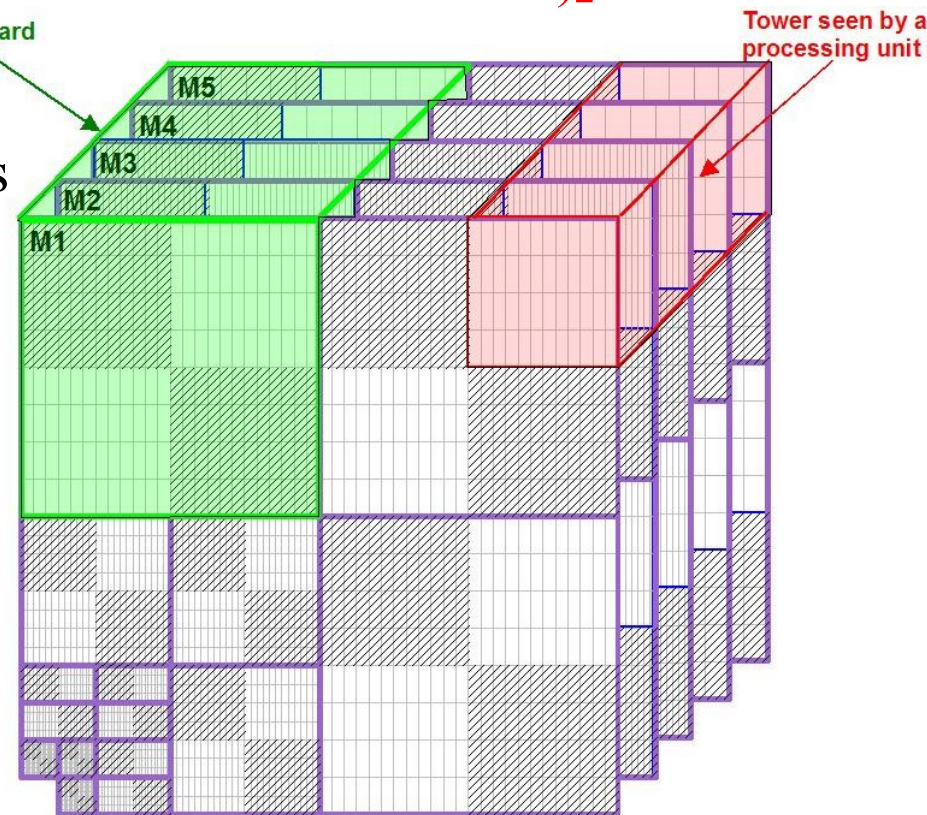
- 1- Trouver un pad touché en M3
- 2- Définir un axe de recherche centré sur le PAD
- 3- Ouvrir un couloir de largeur X le long de cet axe
- 4- Sélectionner une trace si un pad est touché dans le couloir dans les plans M5 et M4 et M2
- 5- Le point de passage en M1 est extrapolé en suivant la droite partant de M3 et passant par le pad touché de M2
- 6- Ce point dans M1 donne l'angle de la trace par rapport au faisceau donc P_T (impulsion transverse)

Algorithme de sélection segmenté **par tour (une tour=une unité de calcul):**

- toutes les tours sont analysées en parallèle
- recherche d'impacts en coïncidence dans 4 plans
- calcul géométrique basé sur les angles des trajectoires pour déduire les impulsions
- sélection des deux meilleures impulsions transverses

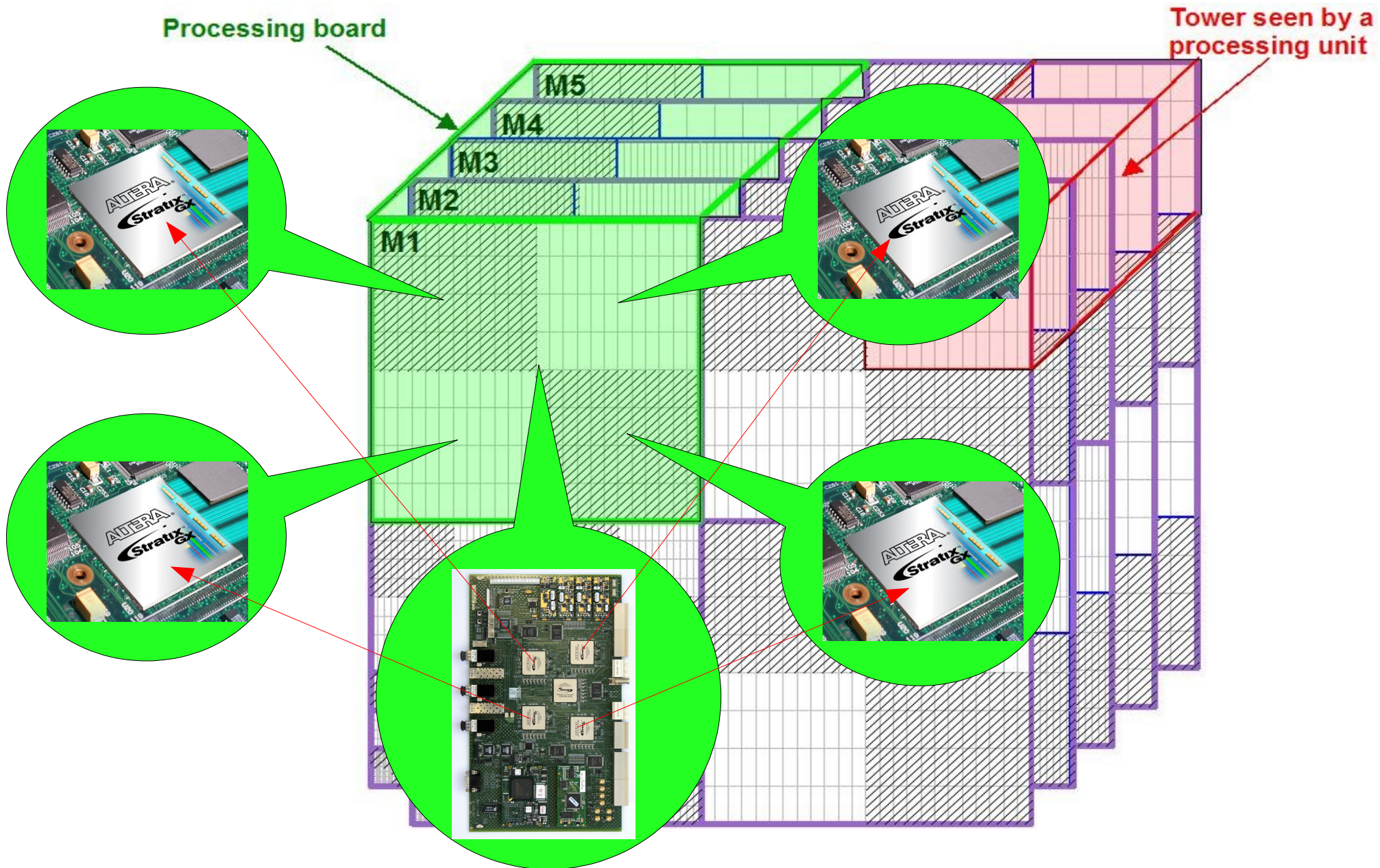
L'analyse des trajectoires traversant plusieurs tours implique des **échanges de données entre les unités de calcul** analysant une tour.

Processing board

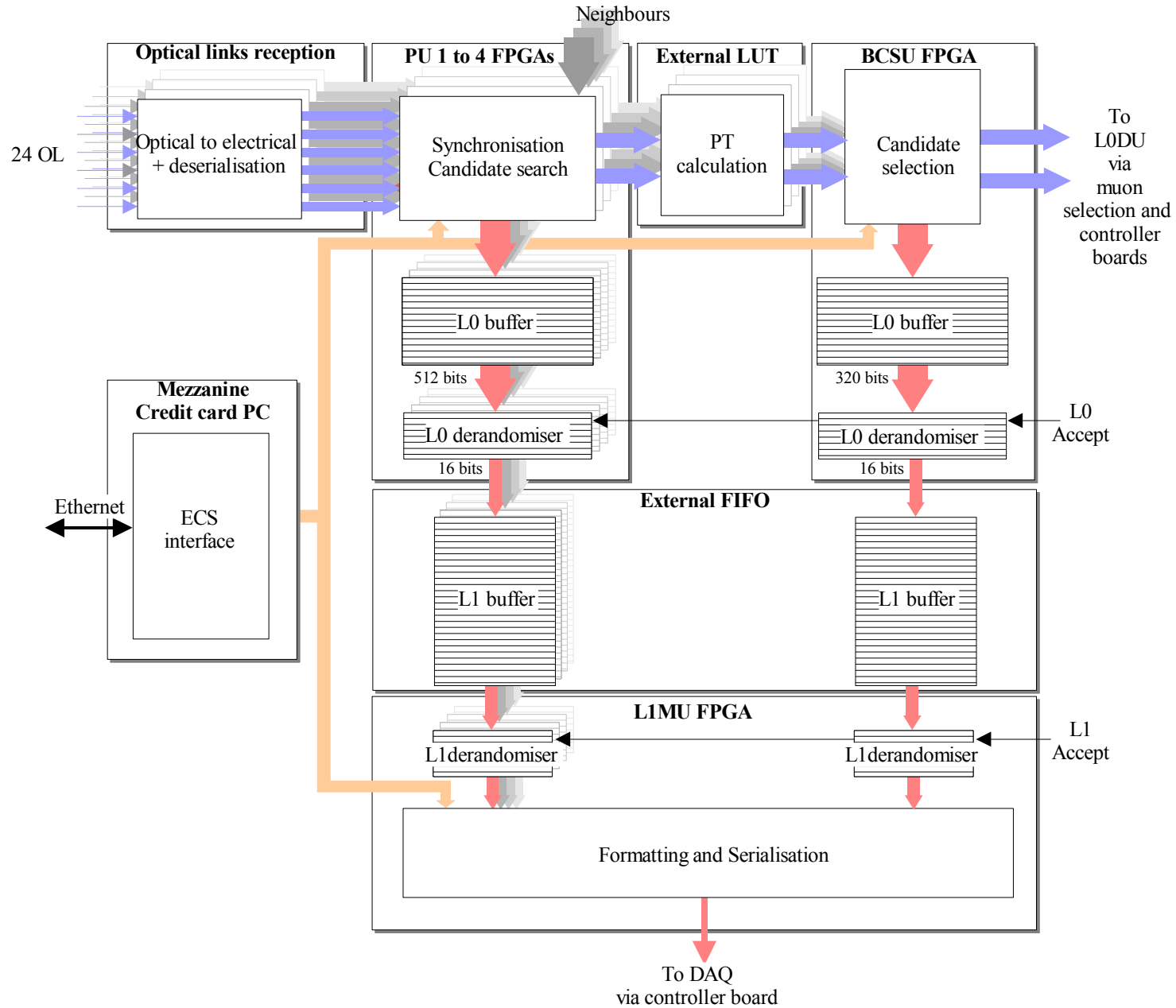


Chaque tour contient la même quantité de données mais couvre une surface plus faible à proximité de l'axe du faisceau car la densité de capteurs (pad) est plus forte.

Trigger L0 allocation matérielle



Sélection au niveau L0: carte de sélection





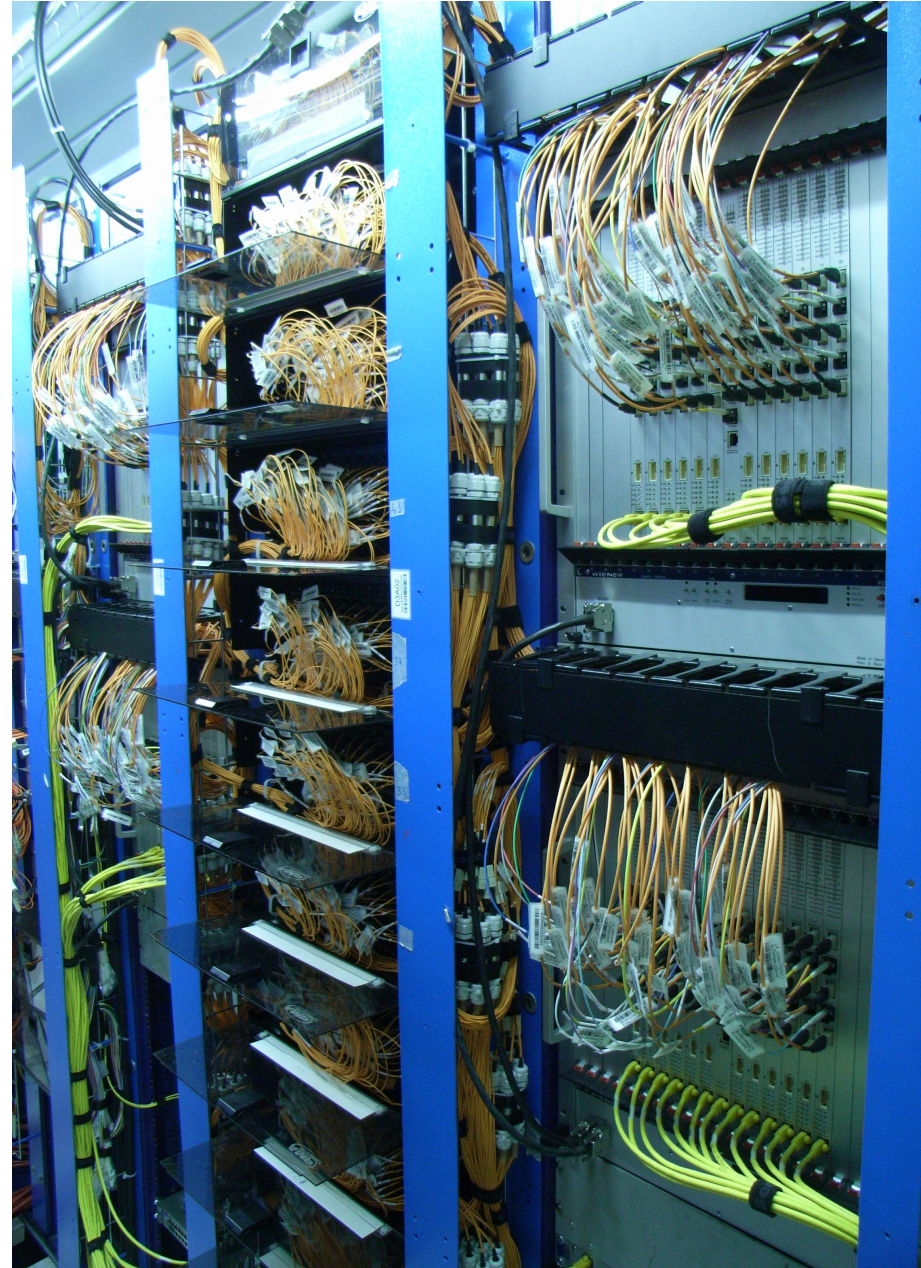
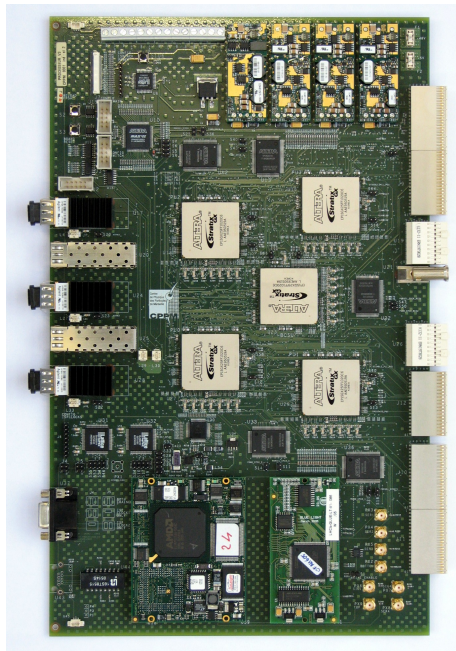
Sélection L0 technologie

Sélection synchrone, massivement parallèle en mode pipeline faite par des unités de calcul (FPGA).

Latence $1.2\mu\text{s}$

1248 optical fibers @ 1.6 Gbps en entrée
112 en sortie

52 cartes de calcul dans 4 chassis.

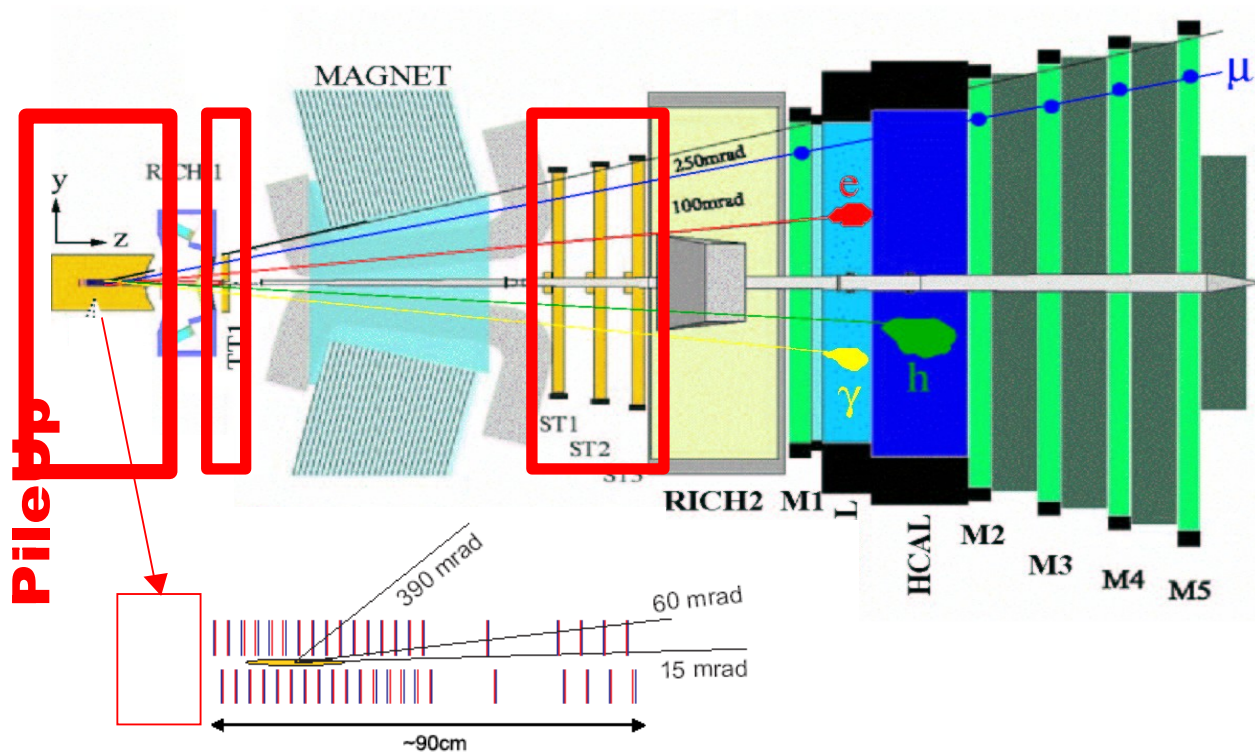


Sélection au L2

Objectif L2(40 Khz)

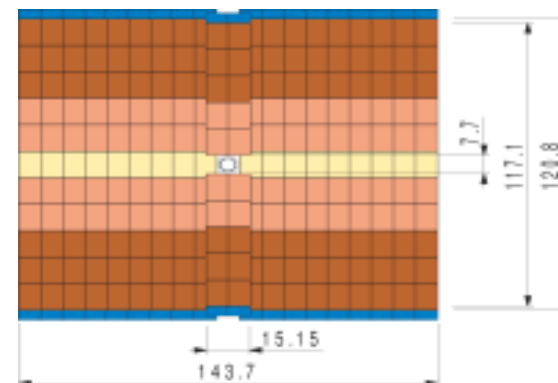
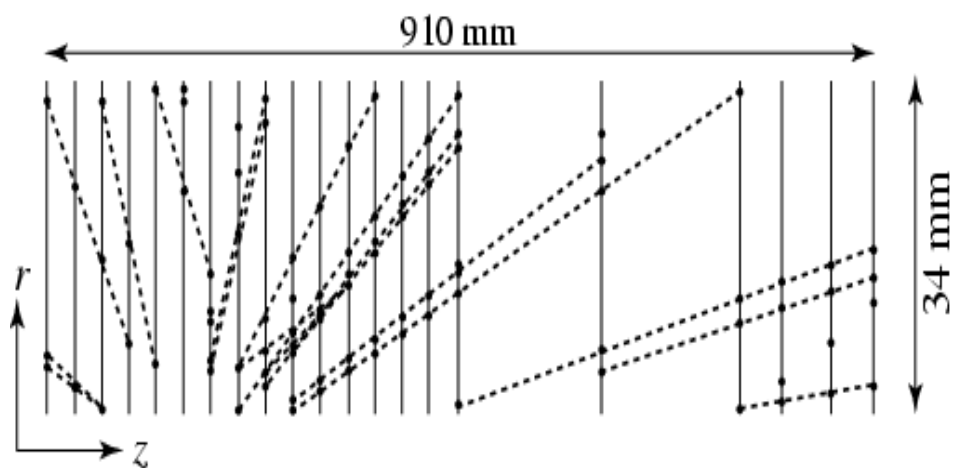
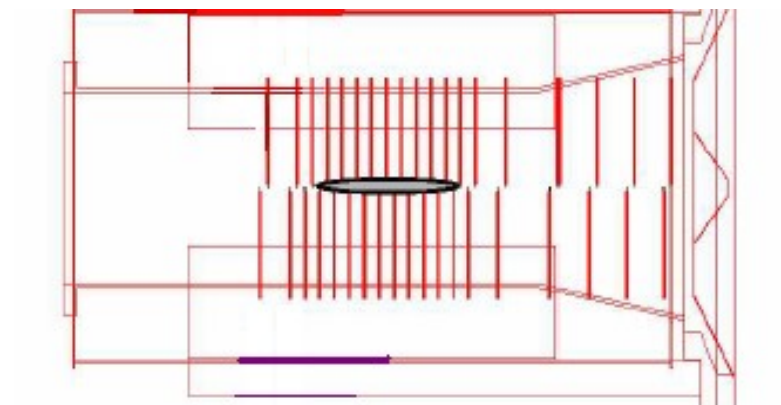
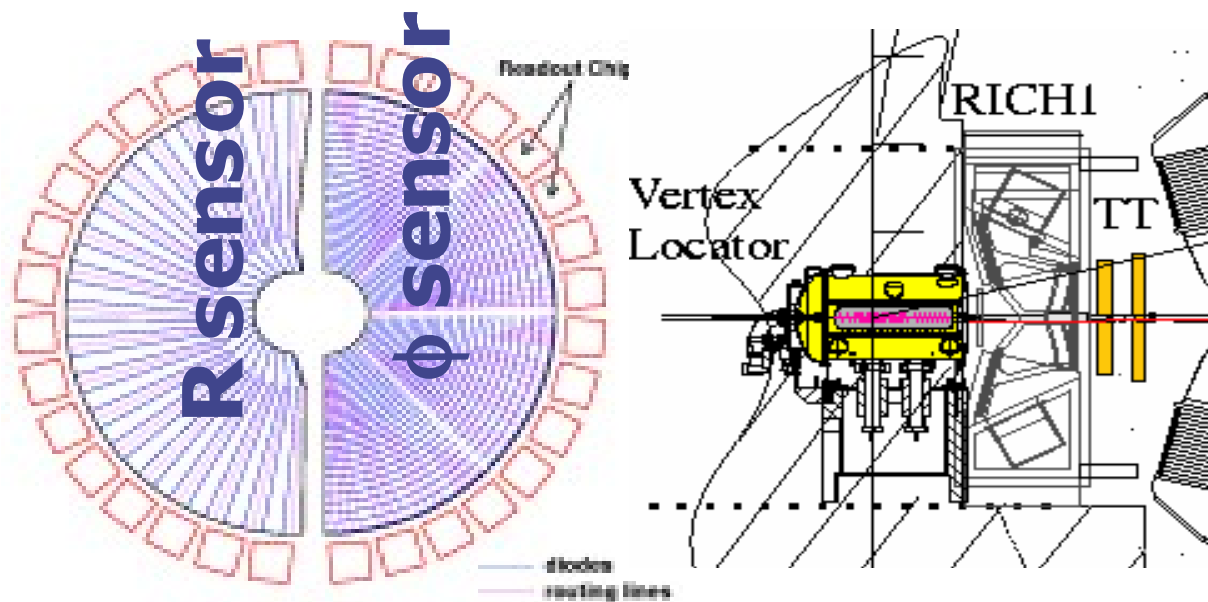
- reconstruction de trace (VELO + TT)
- recherche de vertex déplacé
- corrélation avec les particules identifiées au L0

SI trace issue d'un vertex déplacé pointe vers une particule identifiée au L0 ALORS on sélectionne



Velo x-z view

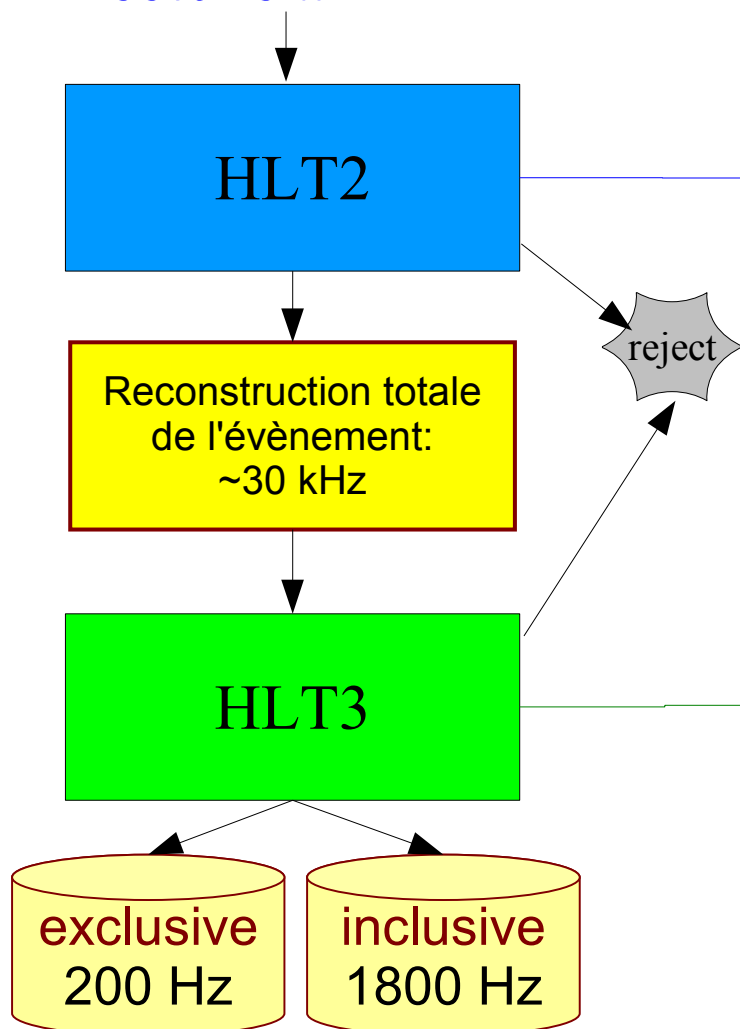
Sélection au L2





High Level Trigger (L2+L3)

Lecture à 1 MHz



– HLT première étape (sélection L2):

- Confirmer la décision L0 avec les trackers
- Reconstruction dans le région d'intérêt
- Sélection par signatures simples (pt, IP, ..)

→ increase fraction of $b\bar{b}$

– HLT seconde étape (sélection L3):

- **exclusive signal selections**
 - full B analysis (relaxed offline cuts)
 - **inclusive streams**
 - trigger on clear signatures
 - gives unbiased B sample
- selection of interesting B-decays



Technologie du High Level Trigger

Architecture commune partagée par L2 et L3 (HLT)

Sélection asynchrone et en série (un événement=1 processus)

flux d'entrée:

L2 volume de 4,8kB à 1.1Mhz

L3 volume 38KB à 40kHz

Algorithmes exécutés sur une **ferme de calcul commune**

~1500 CPU répartis dans ~50 sous-fermes (code partagé avec le off-line)

Routage effectué par un **Méga switch IP** du commerce 1280 ports

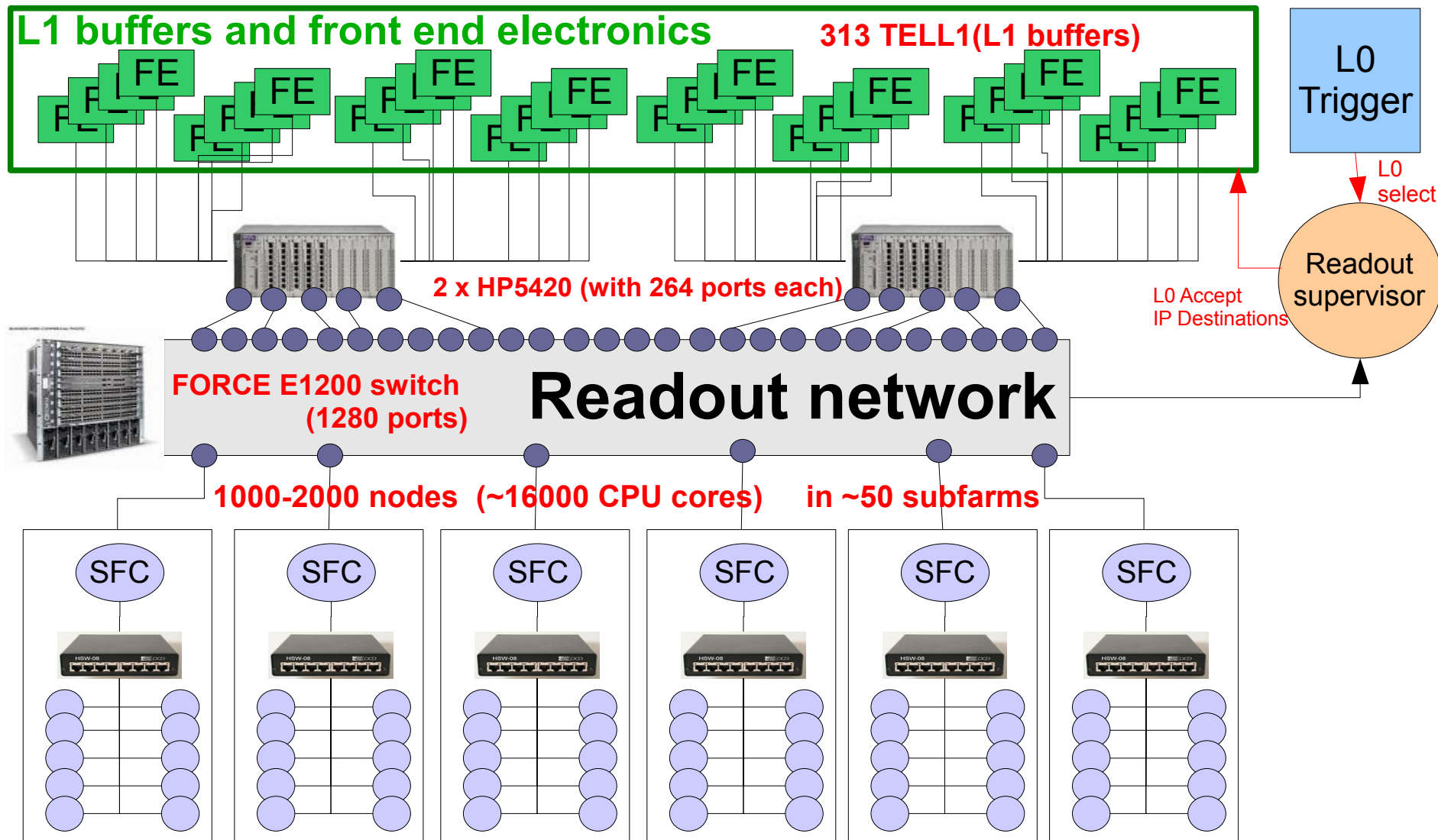
Empaquetage des données dans des paquets IPV4 avec regroupement ~25evts

Concentration amont par 2 gros switch Gigabit ETHERNET de 264 ports.

Utilisation de liaisons cuivre Gigabit ETHERNET



LHCb: DAQ overview





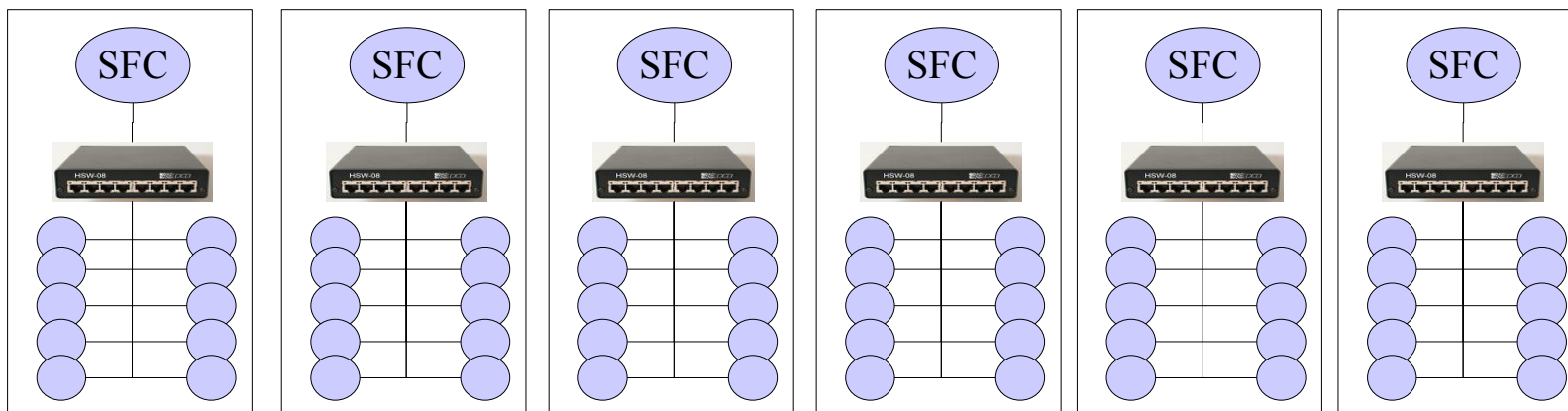
Organisation en sous fermes

Chaque sous-ferme dispose d'une SFC (switch farm interface):

- **assemble** les évènements à partir des fragments contenus dans chaque paquet
- **alloue et distribue** les évènements à chaque machine consommatrice en équilibrant la charge
- **reçoit les décisions de sélection** et en conséquence envoie:
 - les données retenues par le HLT vers le système de stockage
 - signale les saturations/désaturations au read-out superviseur (throttling)
- **sécurité** limite à 58ms maxi le délai de décision L1

Dans chaque nœud de temps de changement de contexte est déterminant pour les performances (traitements L1).
Sous Linux 2.4 et 2.5 de l'ordre de 10μ sec (négligeable par rapport à la latence maxi du L1 58 ms)

Répartition CPU: L1=55% HLT=25% reconstruction=20%





Distribution des signaux et horloges

Les sous détecteurs sont traitées en parallèle et dans chaque sous détecteur des zones différentes sont traitées en parallèle dans des milliers de cartes électroniques chacune contenant un fragment des données.

Comment identifier l'évènement (inter-action) à laquelle chaque fragment de données appartient à son entrée dans le système trigger DAQ (entrée L0)?

Comment diffuser sur plusieurs mètres des signaux indiquant que les fragments d'un évènement donnés doivent être jetés ou gardés?

Dans LHCb l'identification des évènements se fait par le numéro de l'inter-action selon un cycle de 3564 valeurs (le cycle machine).



Distribution des signaux et horloges

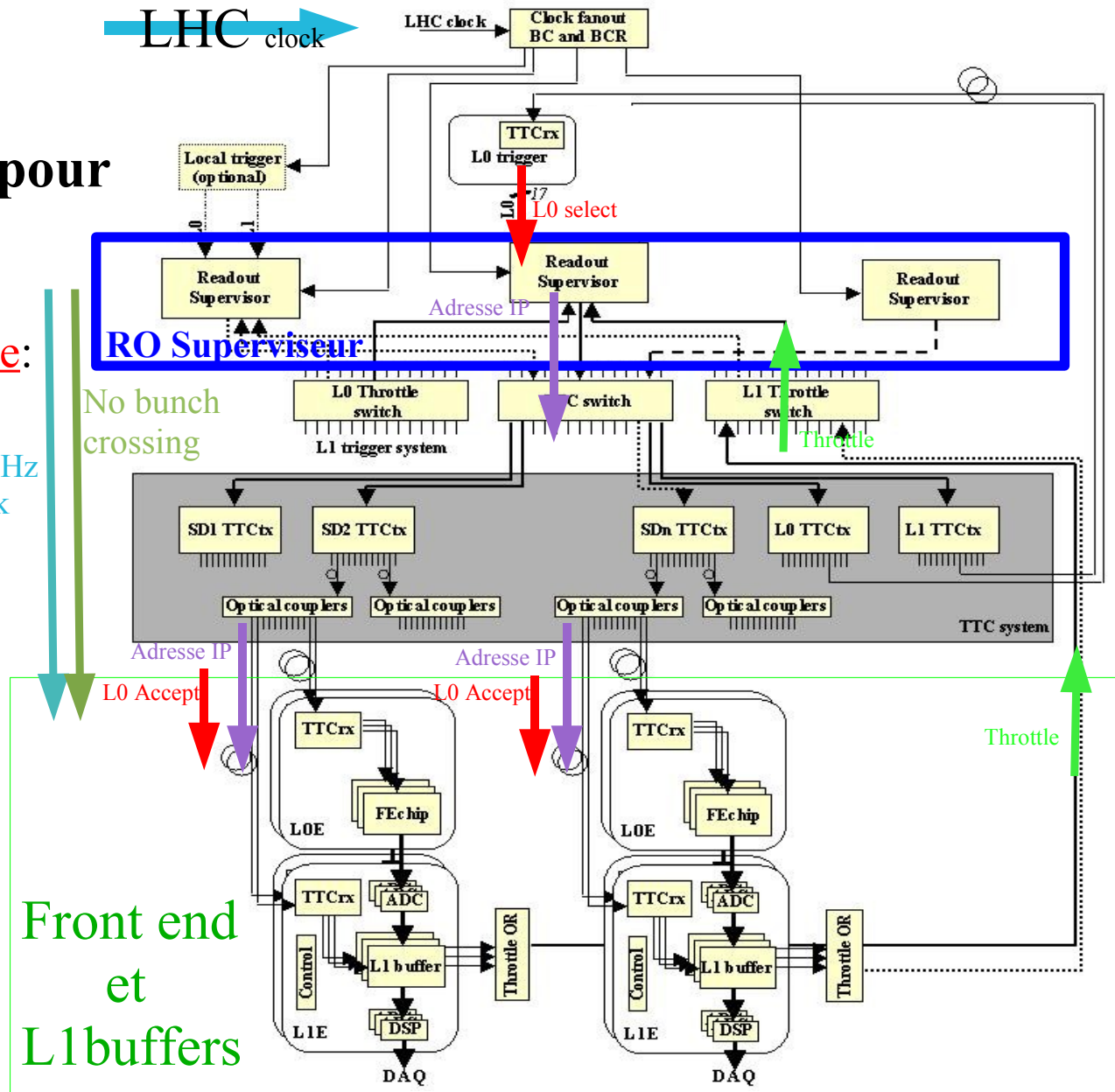
Utilisation du système TTC pour les expériences LHC

Réseau de diffusion sur fibre optique:

- horloge
- message diffusés
 - no évènement, bunch
 - reset compteurs
 - etc ...
- messages adressés

3 composants:

- TTCtx
- TTCRx (ASIC)
- Etoiles optiques de diffusion





Conclusions

Le système de sélection et d'acquisition de LHCb est conçu sans temps mort pour réduire le flux des données d'une fréquence de **40MHz à 200Hz**.

Ce DAQ est représentatif de l'évolution des grands systèmes trigger DAQ actuels:

- il utilise un maximum de **composants standards du marché**
- il utilise une grande ferme de calcul (~1500 PC) partagée pour les sélection L2 et L3
- le code tournant en ligne est très proche du code hors ligne
- les sélections de premier niveau sont faites avec des processeurs cablés spécialisés, reprogrammables, implantés avec des FPGA
- la distribution d'évènements est organisée autour d'un Méga switch IP 1280 ports

Il est conçu de façon modulaire pour permettre facilement:

- son **extension** au cours du temps
- le remplacement des composants pour bénéficier de l'**évolution rapide** des technologies informatiques et réseaux du marché.