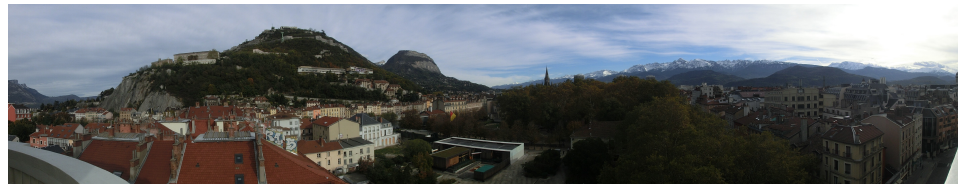


Enregistrer et analyser pour découvrir

Catherine Biscarat

biscarat@lpsc.in2p3.fr

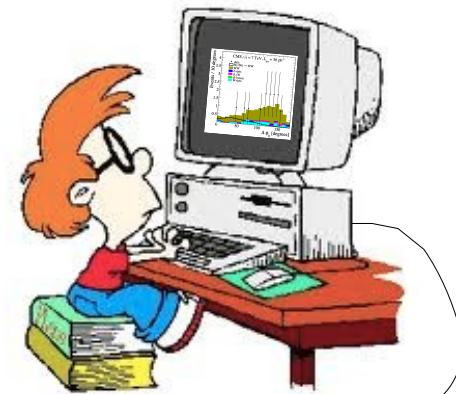
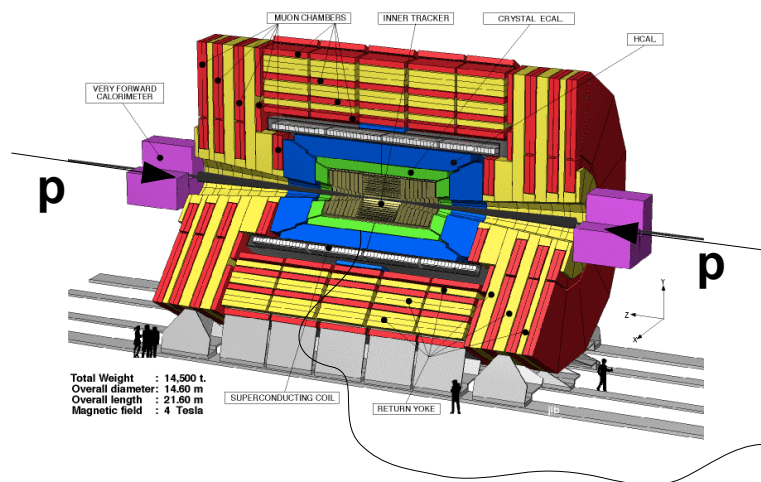
Laboratoire de Physique Subatomique et de Cosmologie de Grenoble



Rencontres de physique de l'infiniment grand à l'infiniment petit,
le lundi 22 juillet 2013

Ensemble, aujourd'hui

- Cadre de la physique des particules
- Problématiques liées à la prise de données
 - Traitement des données (computing)
 - Choix/tri des événements
 - Acheminement des données au physicien (analyse)
 - La grille de calcul du LHC



Petite intro sur l'oratrice

- Auparavant : expérimentaliste en physique des particules, sur collisionneur



Détecteur aux EU, à Chicago, collisionneur ppbar Tevatron à $\sqrt{s} = 2$ TeV
- calorimétrie, production d'événements simulés, recherche SUSY

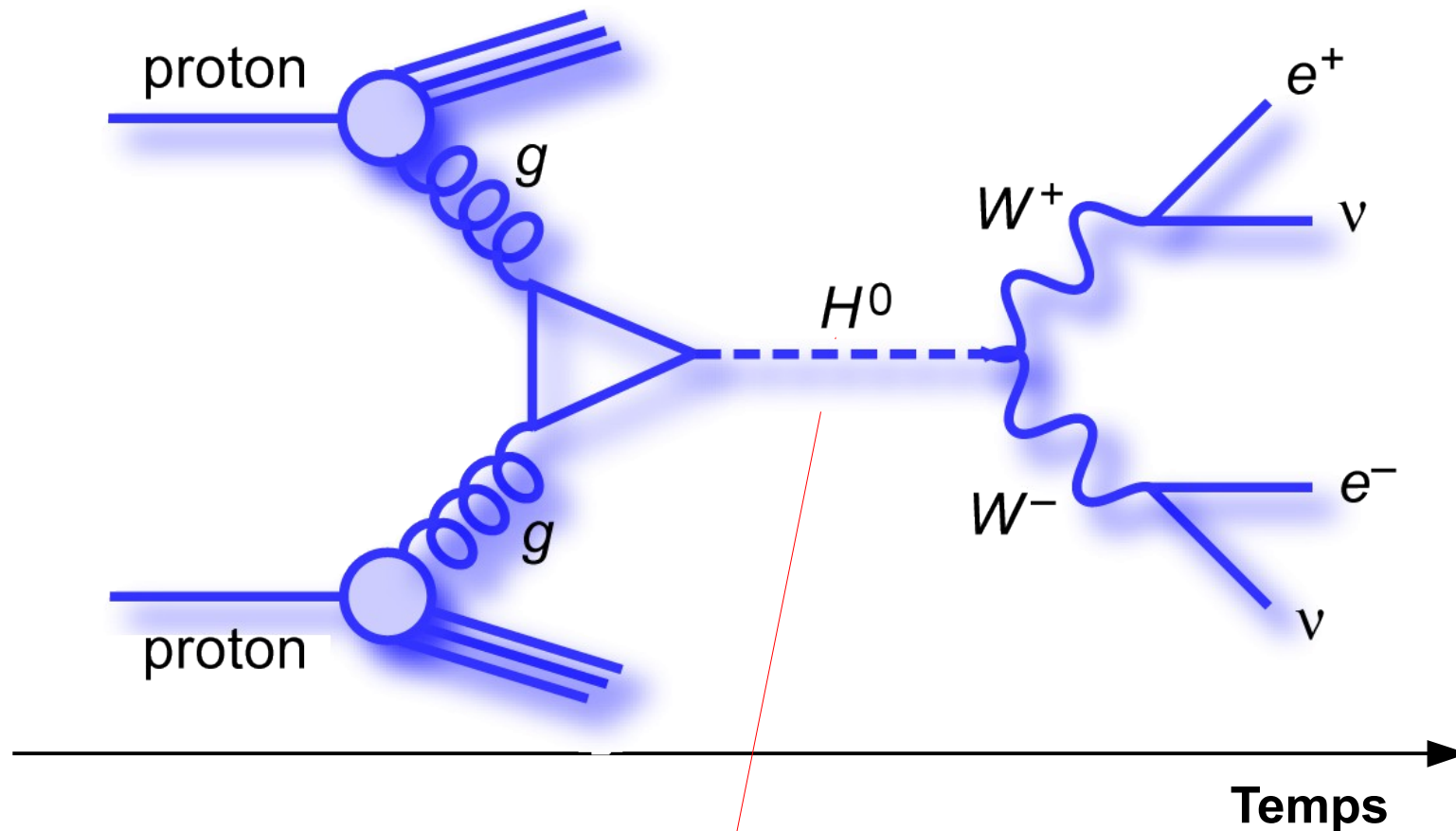


Détecteur au CERN, à Genève, collisionneur pp LHC à $\sqrt{s} = 14$ TeV
- calorimétrie, recherche de nouvelles particules
- responsable des activités de calcul ATLAS dans un centre de calcul majeur

- Ingénieur en informatique à l'IN2P3/CNRS
Grilles de calcul (WLCG et grille légère grenobloise CIGRI)

Comment chercher le Higgs ?

Si le boson de Higgs existe, alors il peut être produit dans une collision de deux protons de grande énergie :



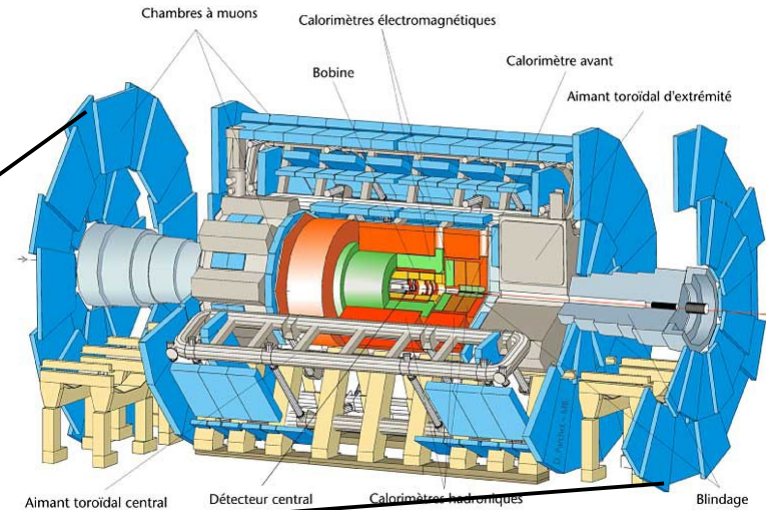
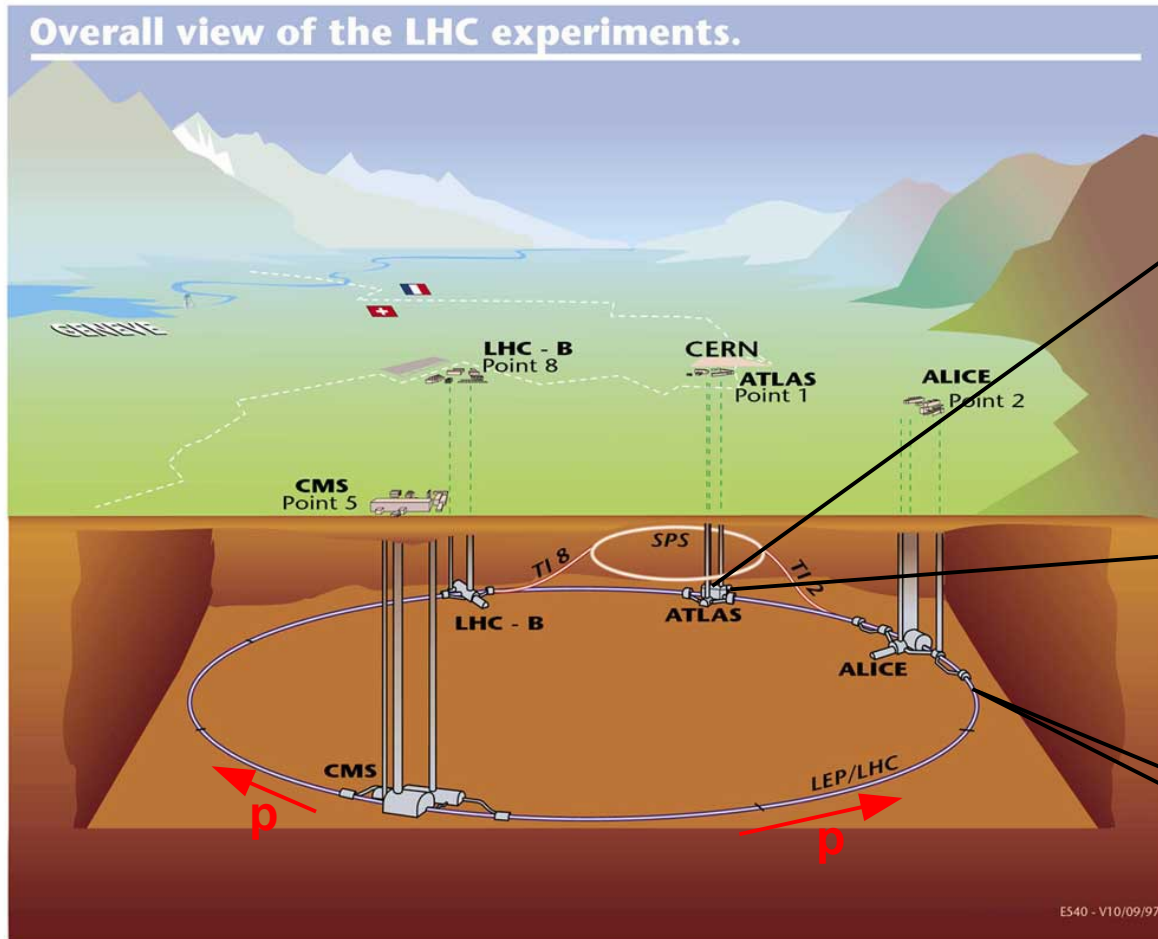
deux protons entrent en collision

deux gluons (constituants des protons) "fusionnent" pour créer un boson de Higgs

le Higgs se désintègre immédiatement en une paire de bosons W (qui se désintègrent immédiatement à leur tour)

les particules stables dans l'état final (ici e , ν) atteignent le détecteur

L'appareillage en pratique



Taux de production des événements

LHC

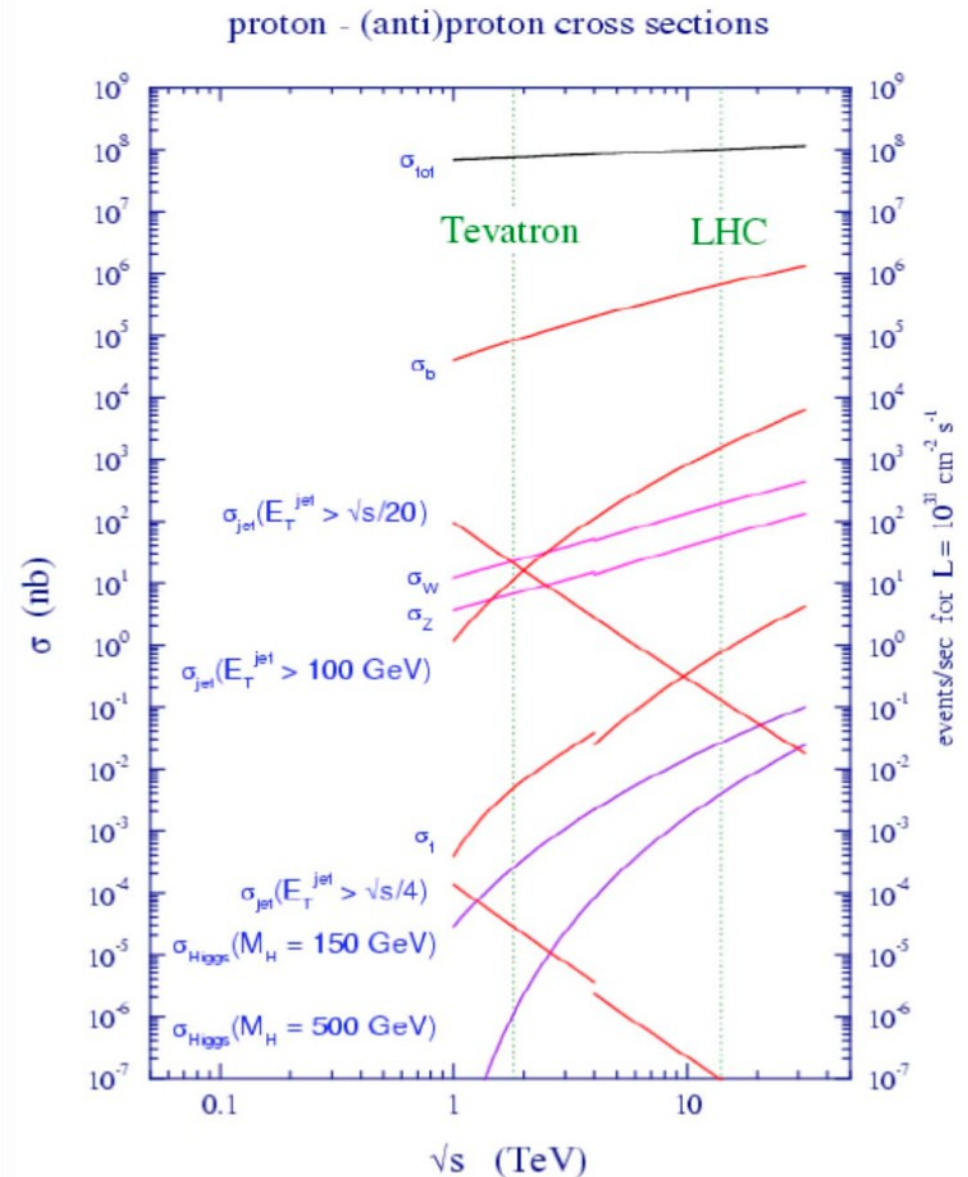
- croisement de faisceaux : 40 MHz
- Taille d'un événement : 1,6 MB

- Difficile à transmettre
- Coûteux à stocker
- Long à analyser

Mais tous les événements ne nous intéressent pas de la même façon

Il faut :

- ne pas louper les événements rares (type Higgs)
Sinon : perte définitive
- collecter une part d'événements bien « connus » par ailleurs
Vérification des mesures



Quelques ordres de grandeur



- Les données accessibles (produites): les **Chutes du Niagara** (1.5 million gpm).
- 40 millions de croisements de paquets de protons par seconde
 - Qui correspondraient à 100 000 CDs écrits par seconde (4 x terre-lune/an)

Quelques ordres de grandeur



- Ce que nous pouvons nous permettre d'écrire (bande) : **lance à incendie** (100 gpm).
- nous choisissons et stockons ~200 événements par seconde,
 - soit 27 CDs écrits par minute (1 expérience).

Quelques ordres de grandeur



- Ce que nous publions : **quelques gouttes** !
- Soit, quelques poignées d'événements.

Tri en ligne des événements

- Trois niveaux de **déclenchement**
- **Temps de décision** de plus en plus grand
- Événement de plus en plus **complet**

Niveau 1

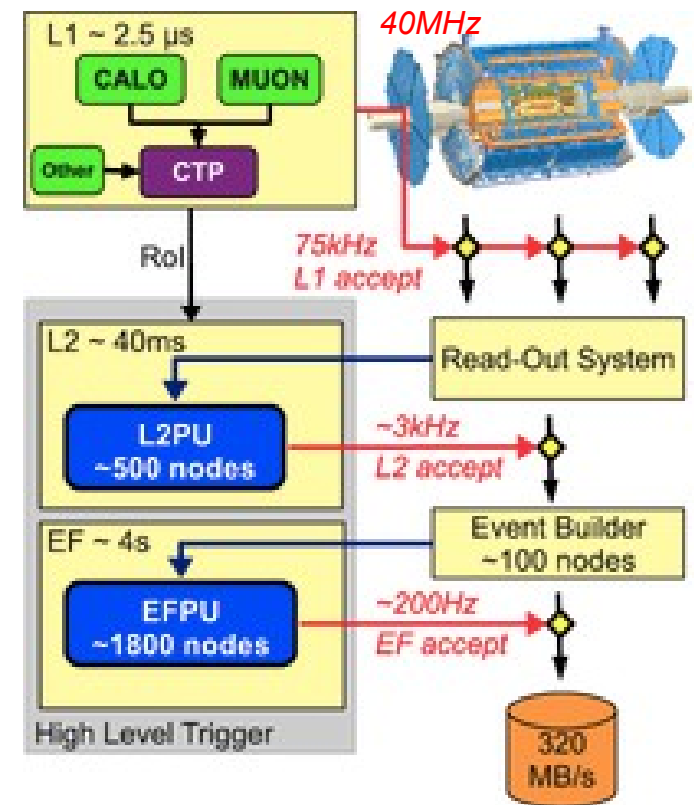
- Circuits électroniques dédiés (FPGA)
- Calorimètres et détecteurs de muons (une partie seulement de l'information)

Niveau 2

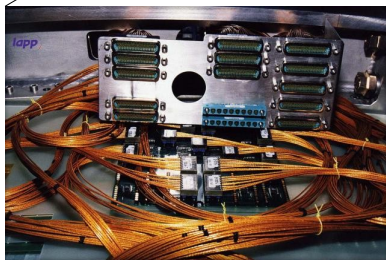
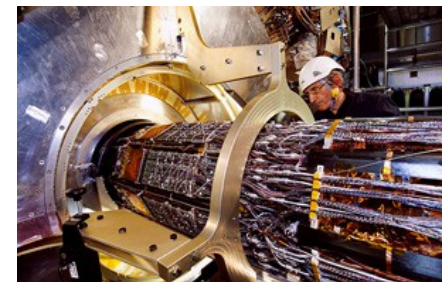
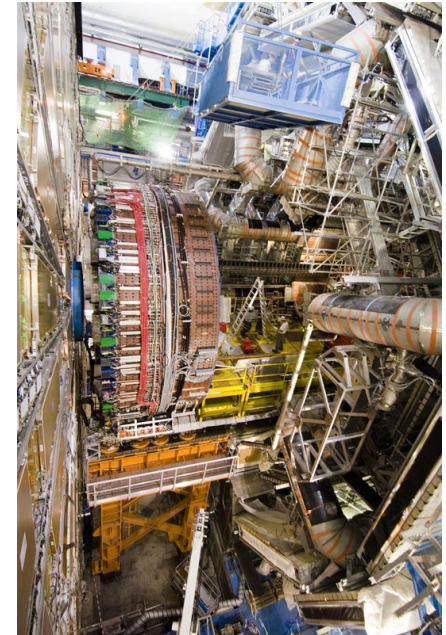
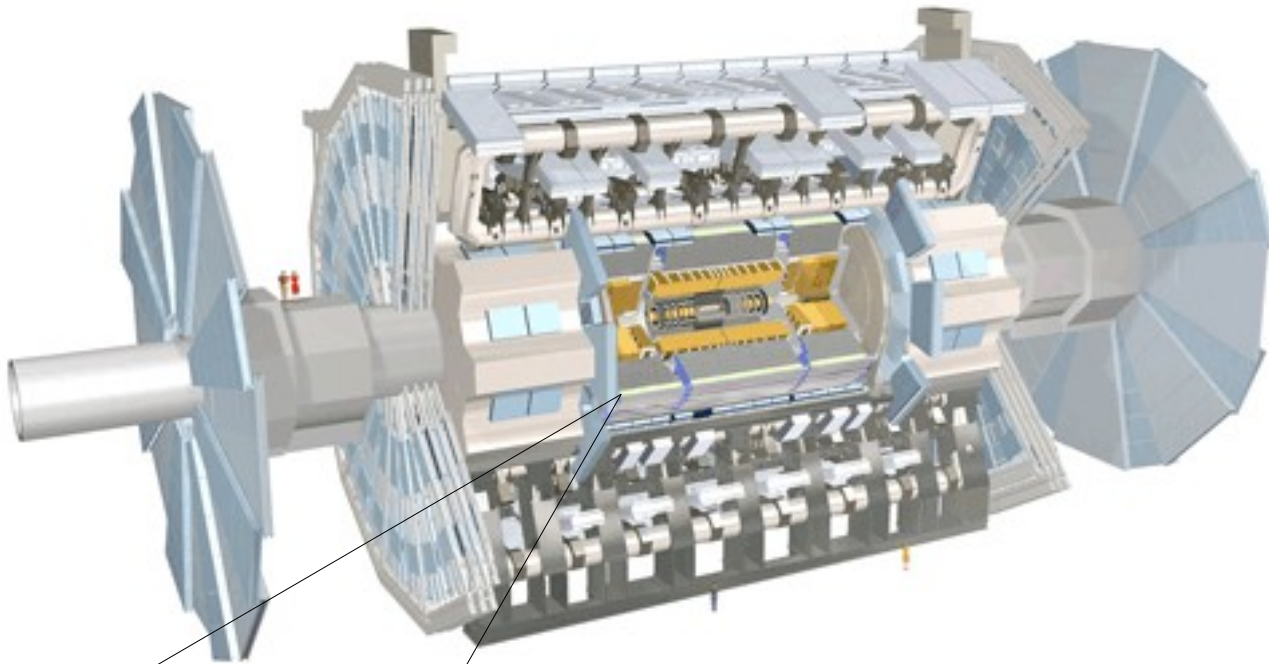
- Événement complet dans régions d'intérêt identifiées au niveau 1
- Algorithmes spécialement rapides

Niveau 3

- Événement complet
- Algorithmes raffinés, de type “analyse”



Les données brutes (RAW)



```
101100 101011 010001  
110111 001011 001100  
100001 111100 100110  
110101 110011 100101  
001010 101000 001010  
111001 100101 000011  
010111 001001 010100  
100010 010100 101111  
100100 101001 001010  
000010 100101 111001
```

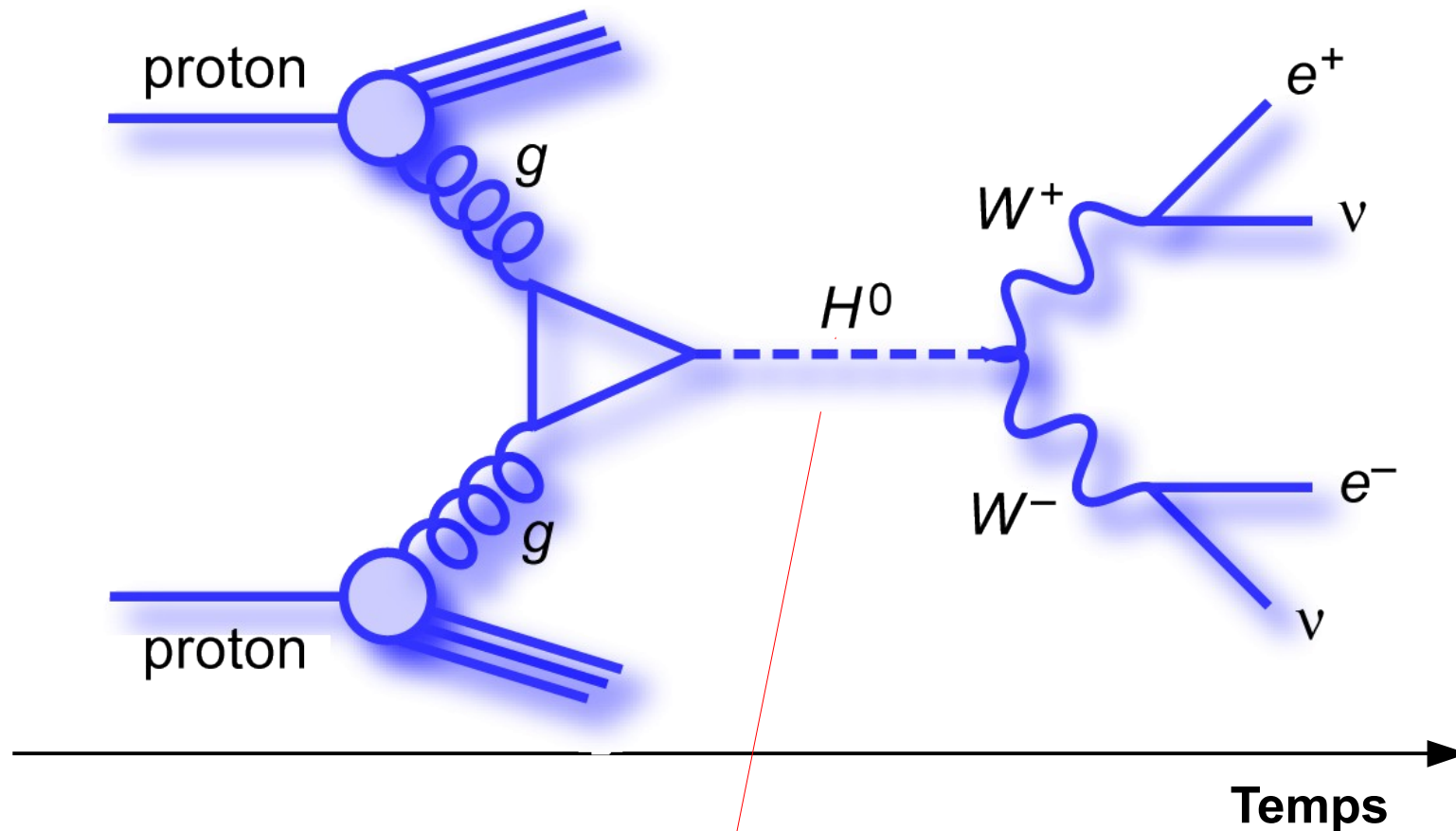
Pixels → oui/non
Calo → tensions en "Volt"

Stockées sur bandes
magnétiques au CERN



Comment chercher le Higgs ?

Si le boson de Higgs existe, alors il peut être produit dans une collision de deux protons de grande énergie :



deux protons entrent en collision

deux gluons (constituants des protons) "fusionnent" pour créer un boson de Higgs

le Higgs se désintègre immédiatement en une paire de bosons W (qui se désintègrent immédiatement à leur tour)

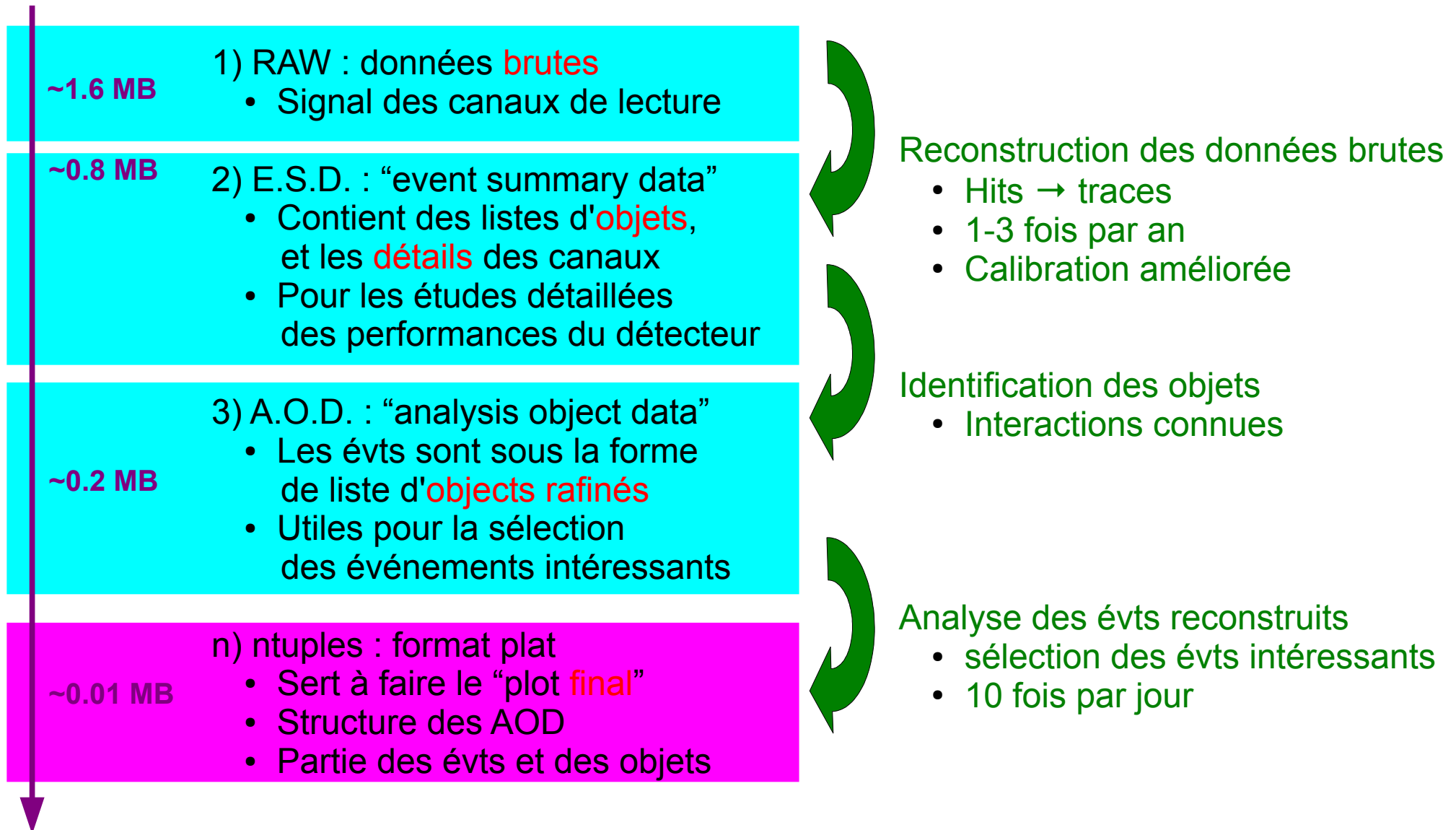
les particules stables dans l'état final (ici e, ν) atteignent le détecteur

Vers l'analyse des données

Taille/évt

Format de données

Type de traitement



Vers l'analyse des données

Taille/évt

Format de données

Type de calcul

~1.6 MB

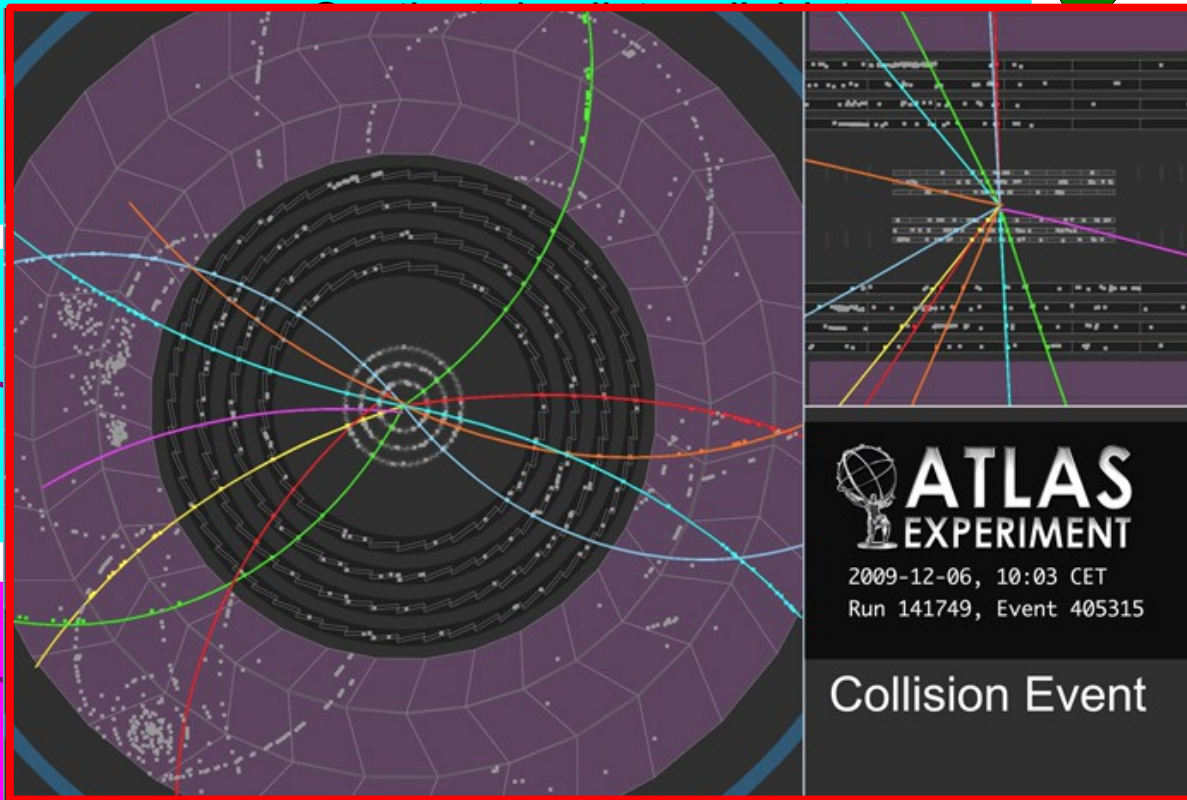
- 1) RAW : données brutes
 - Signal des canaux de lecture

~0.8 MB

- 2) E.S.D. : "event summary data"

Reconstruction des données brutes

- Hits → traces



Identification des objets

- Interactions connues

analyse des évt reconstruits

- sélection des évt intéressants

<http://atlas.web.cern.ch/Atlas/public/EVTDISPLAY/events.html>

Vers l'analyse des données

Taille/évt

Format de données

Type de calcul

~1.6 MB

- 1) RAW : données brutes
 - Signal des canaux de lecture

~0.8 MB

- 2) E.S.D. : "event summary data"
 - Contient des listes d'objets

Reconstruction des données brutes

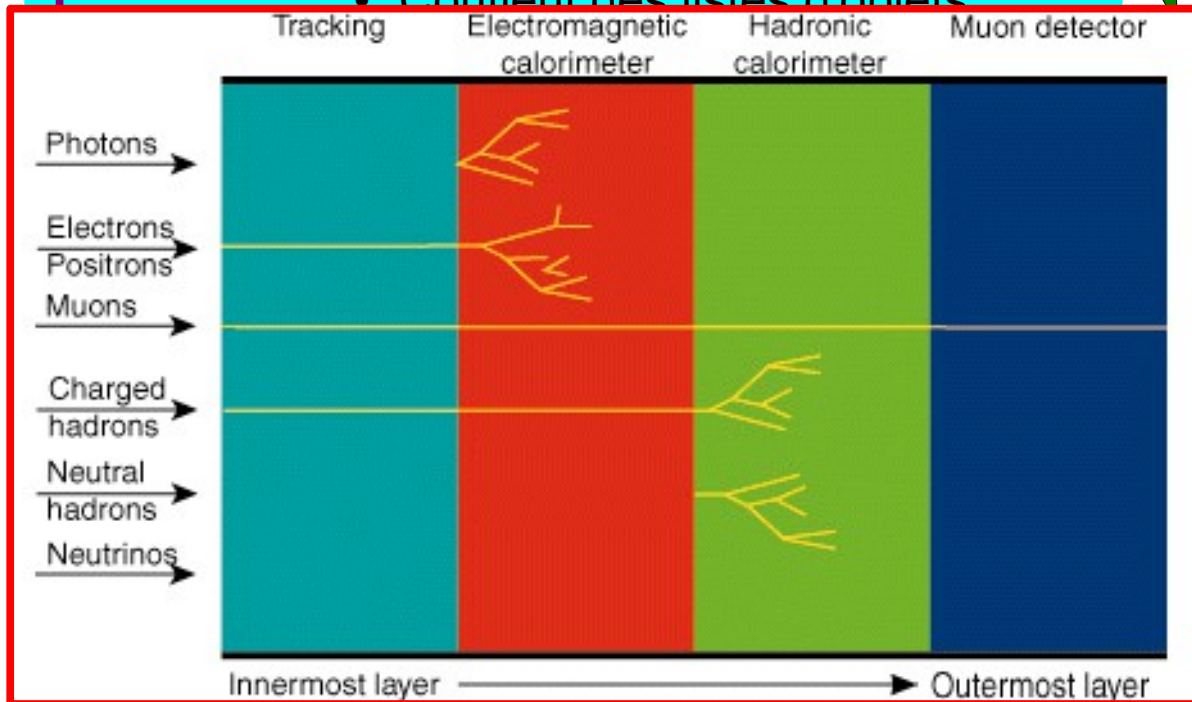
- Hits → traces

Identification des objets

- Interactions connues

Analyse des évtS reconstruits

- sélection des évtS intéressants



- Partie des évtS et des objets

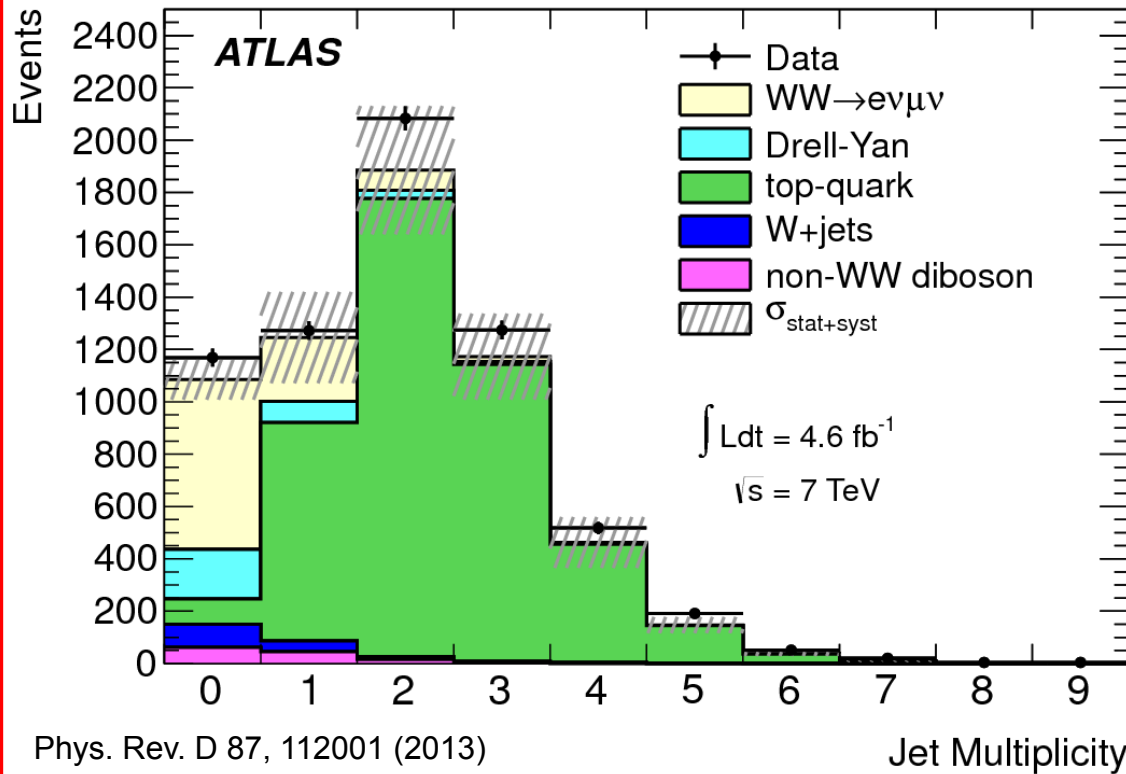
Vers l'analyse des données

Taille/évt

Format de données

Type de calcul

1) RAW : données brutes



Reconstruction des données brutes

- Hits \rightarrow traces

Identification des objets

- Interactions connues

Analyse des évt reconstruits

- sélection des évt intéressants

~0.01 MB

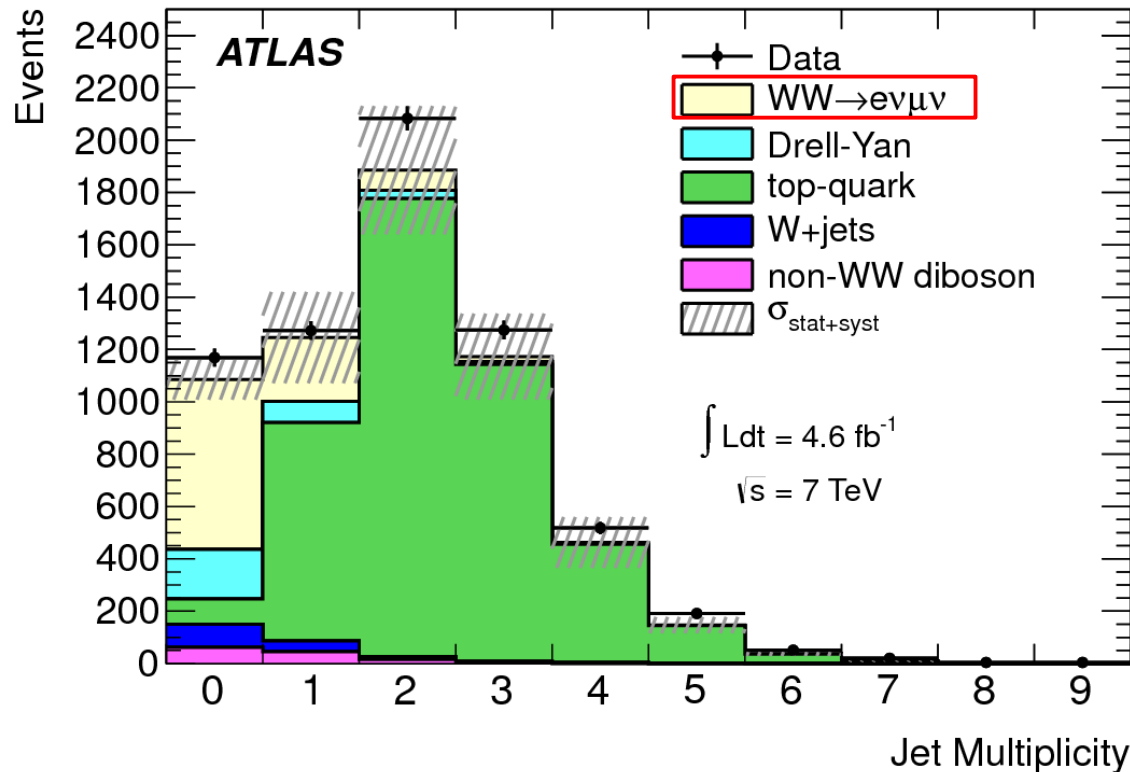
- Sert à faire le "plot final"
- Structure des AOD
- Partie des évt et des objets

Interpréter les données

En d'autres termes :

Confronter les données à un modèle (le MS)

- **Accord** données/simulation avec le MS (compréhension du détecteur)
- Trouver des **déviations** (découverte de nouveaux phénomènes)



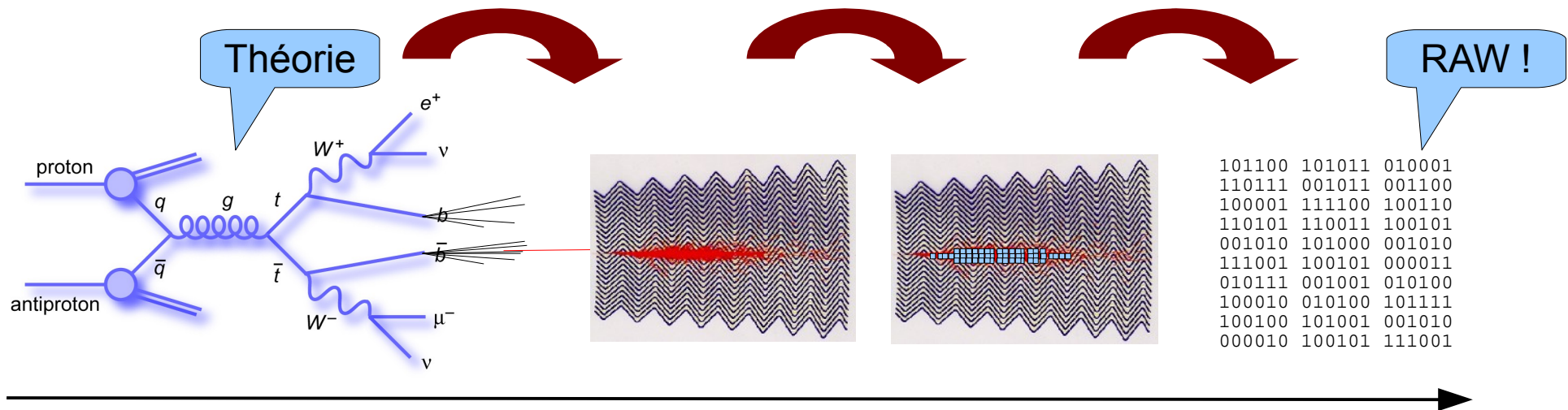
Exploitation d'une expérience : impossible sans simulation !!

La simulation

Simuler quoi au juste ? **Les données brutes !**

Trois ingrédients :

- 1) Modéliser la “physique” (collisions, processus)
- 2) Modéliser l'interaction des particules dans le détecteur
- 3) Modéliser les signaux transmis par le détecteur

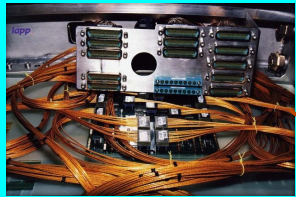


Vers l'analyse des données

Traitement centralisé

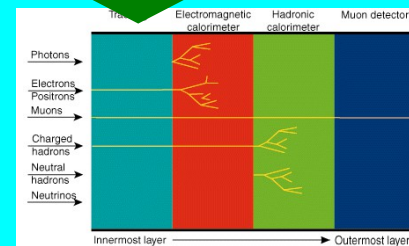
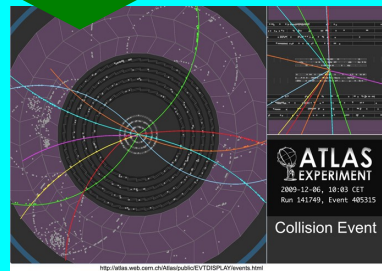
Reconstruction
des données brutes

Identification des objets
et sélection d'état final



```

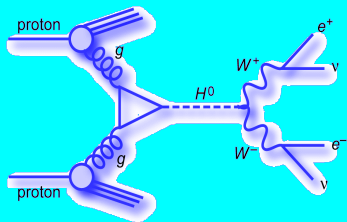
101100 010001
110111 001100
111100 100110
110101 110011
001010 001010
100101 000011
010111 010100
    
```



groupe/individu

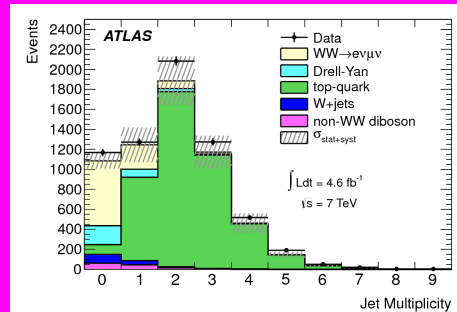
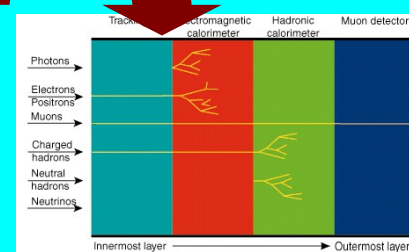
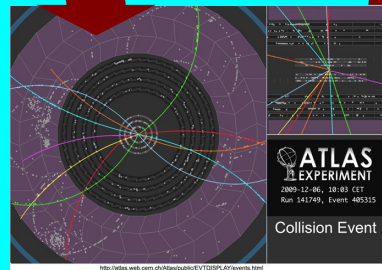
Analyse finale
(n fois / jour)

Simulation des événements

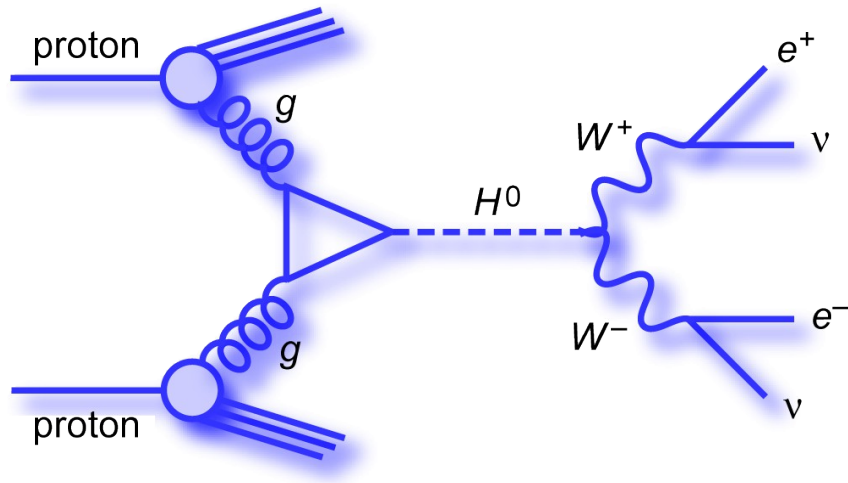


```

101100 010001
110111 001100
111100 100110
110101 110011
001010 001010
100101 000011
010111 010100
    
```



Recherche d'un processus rare

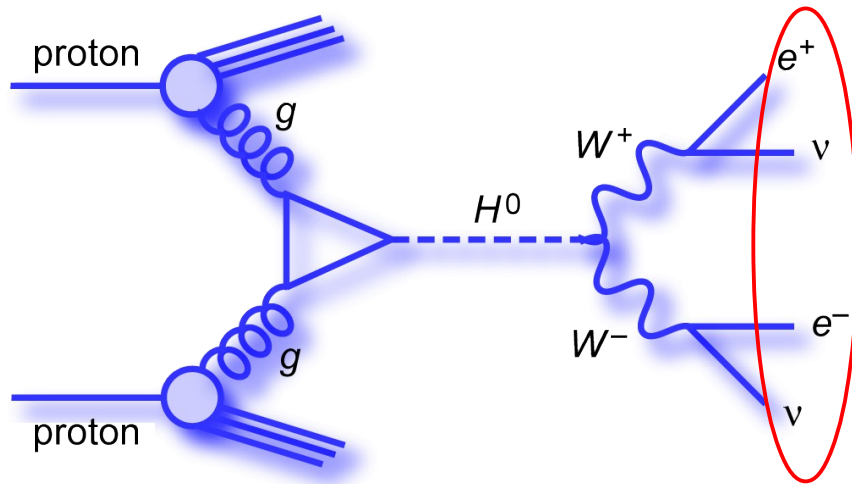


Nous avons déjà discuté cette chaîne de réaction.

Pour la recherche du Higgs elle constitue **“le signal”**.

Rappelons-nous que seuls les particules stables de l'état final atteignent le détecteur.

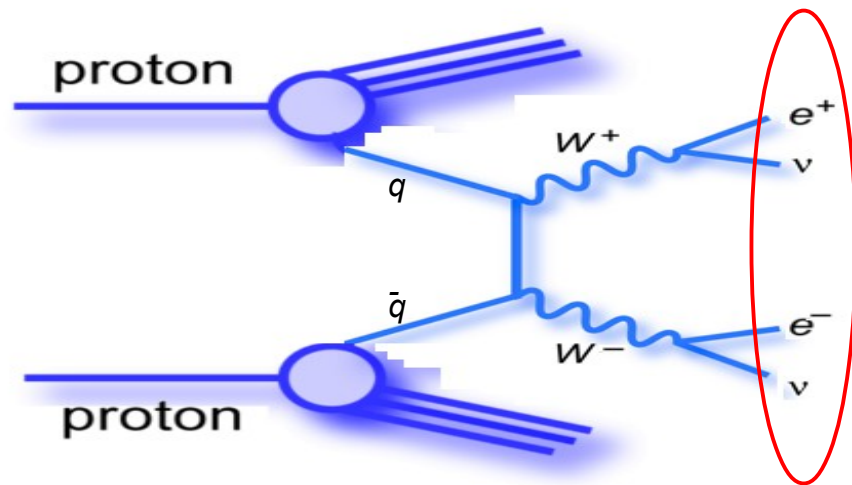
Recherche d'un processus rare



Nous avons déjà discuté cette chaîne de réaction.

Pour la recherche du Higgs elle constitue **“le signal”**.

Rappelons-nous que seuls les particules stables de l'état final atteignent le détecteur.

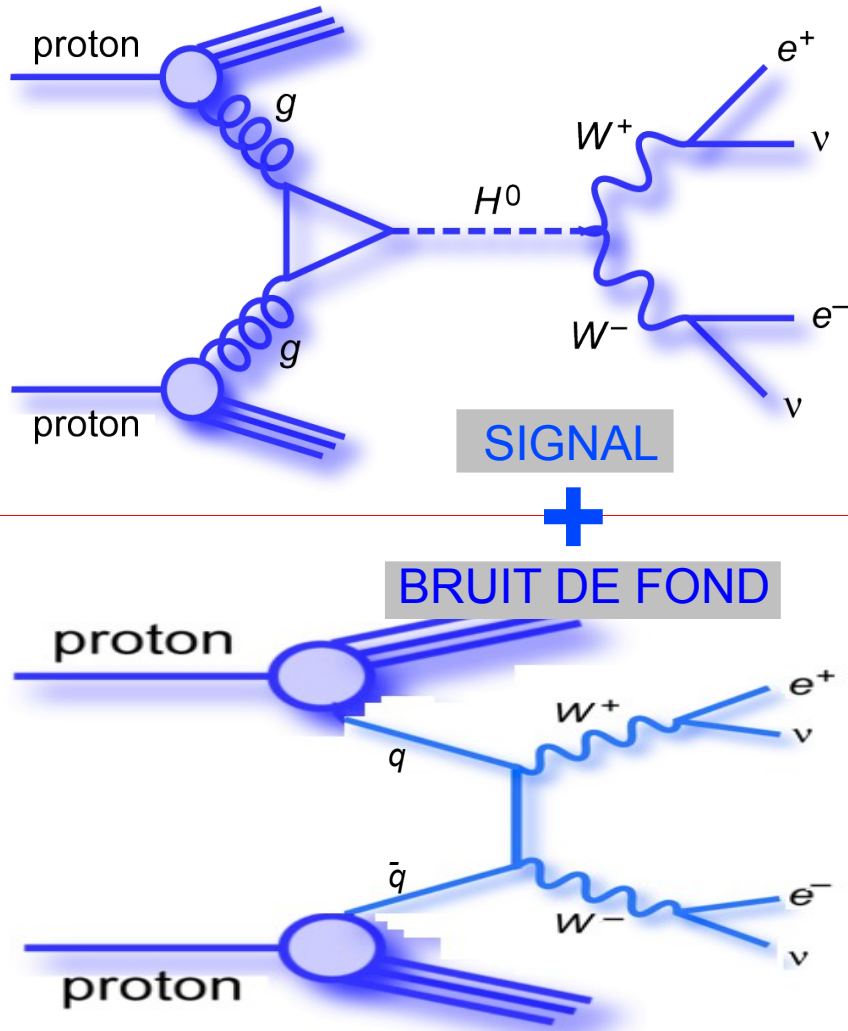


Malheureusement il existe d'autres chaînes qui donnent le même état final – et qui sont possibles même si le Higgs n'existe pas !!

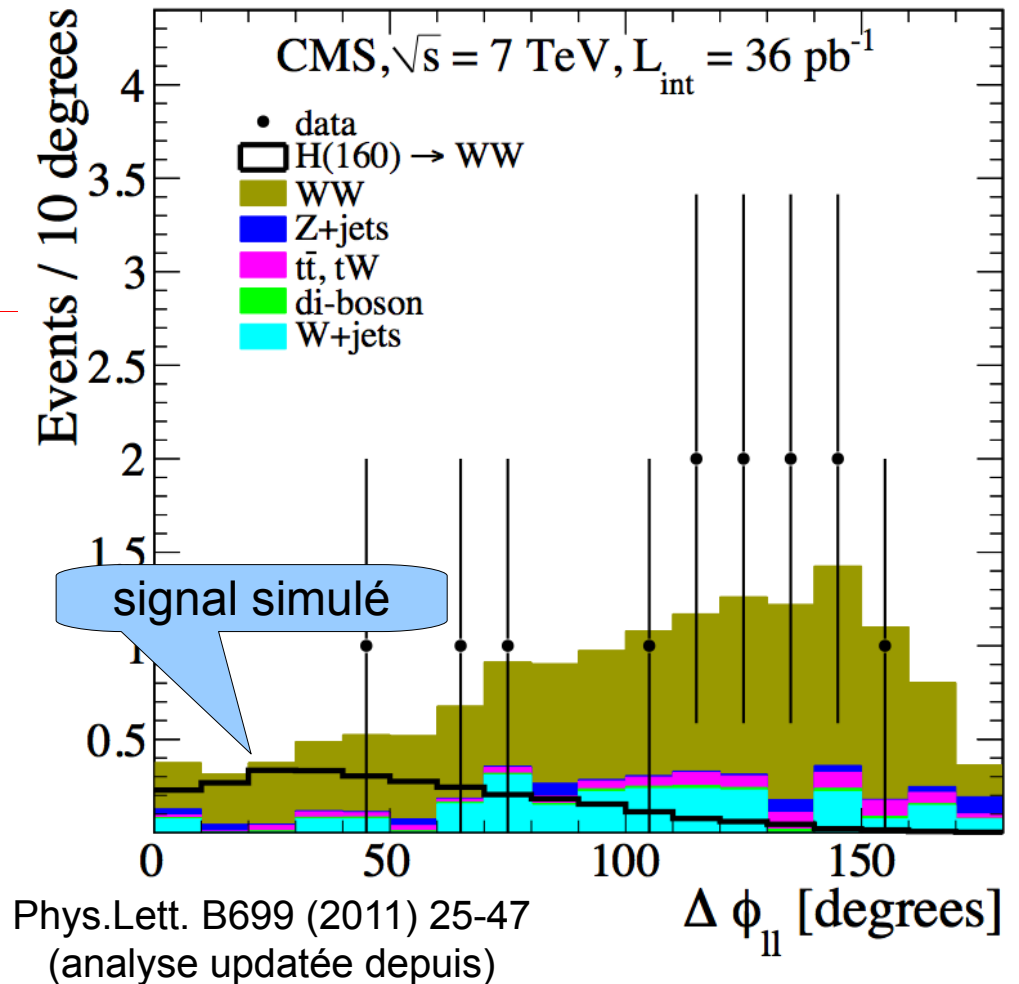
Pour la recherche du Higgs elles constituent **“un bruit de fond”**.

Pour rendre les choses encore pire : ce type de réaction est **beaucoup plus abondant que le signal**.

“Signal” et “bruit de fond”



Ce lot de données réelles (la réalité) est-il plus compatible avec l'hypothèse FOND SEUL ou bien FOND+SIGNAL ?



Quantités de données au LHC

A vos calculettes !

- * 100 collisions enregistrées par sec.
 - * 1-2 MB de données par collision
- plusieurs PB de données produites par an et pour un détecteur LHC

Soit **600 000 films DVD produits par un détecteur en un an**

1 MB
1 photo digitale
1 dictionnaire
1 CD = 650 MB

1 GB = 1000 MB
5GB = 1 film DVD

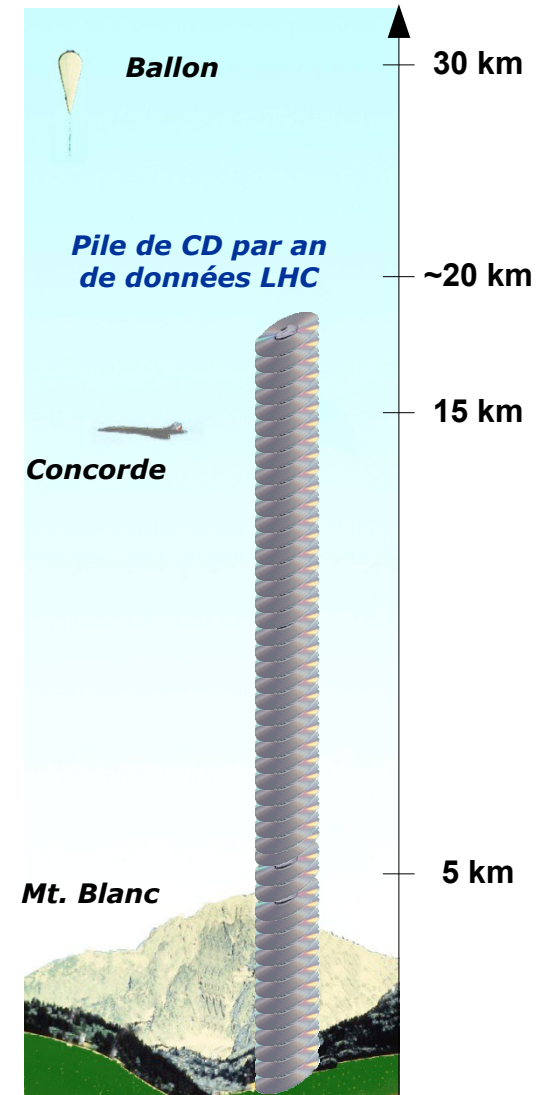
1TB = 1000 GB
Livres produits par an

1PB = 1000 TB
Production par an de
1 expérience LHC

1 EB = 1000 PB
Production mondiale
d'information en 1 an



Enorme **quantité de données**
à traiter et à stocker;
aucun centre de calcul
ne peut le faire seul.



Innovation : la grille de calcul

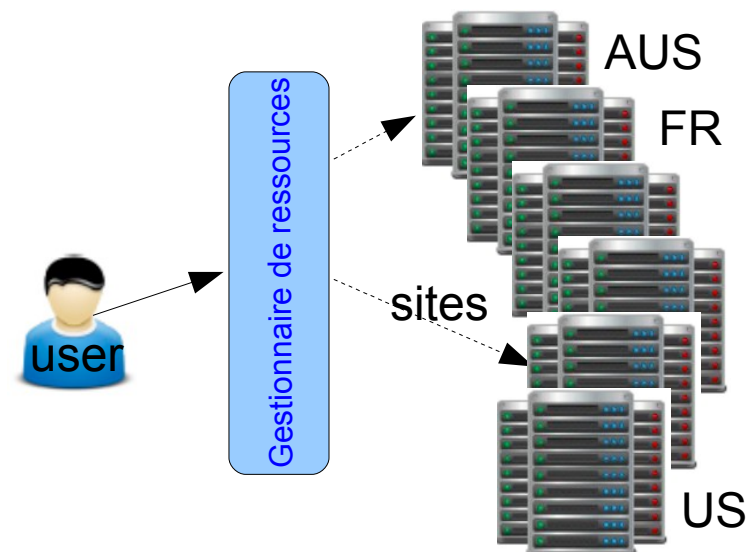
Une nécessité

- Enorme **quantité de données** à traiter et stocker
- Utilisateurs **distribués partout** dans le monde
- Introduction de la **grille de calcul**
 - Mutualisation de ressources de plusieurs unités pour un but commun
 - Terme pris du réseau électrique (*electric power grid*) où un appareil électrique est alimenté quelque soit où il est branché

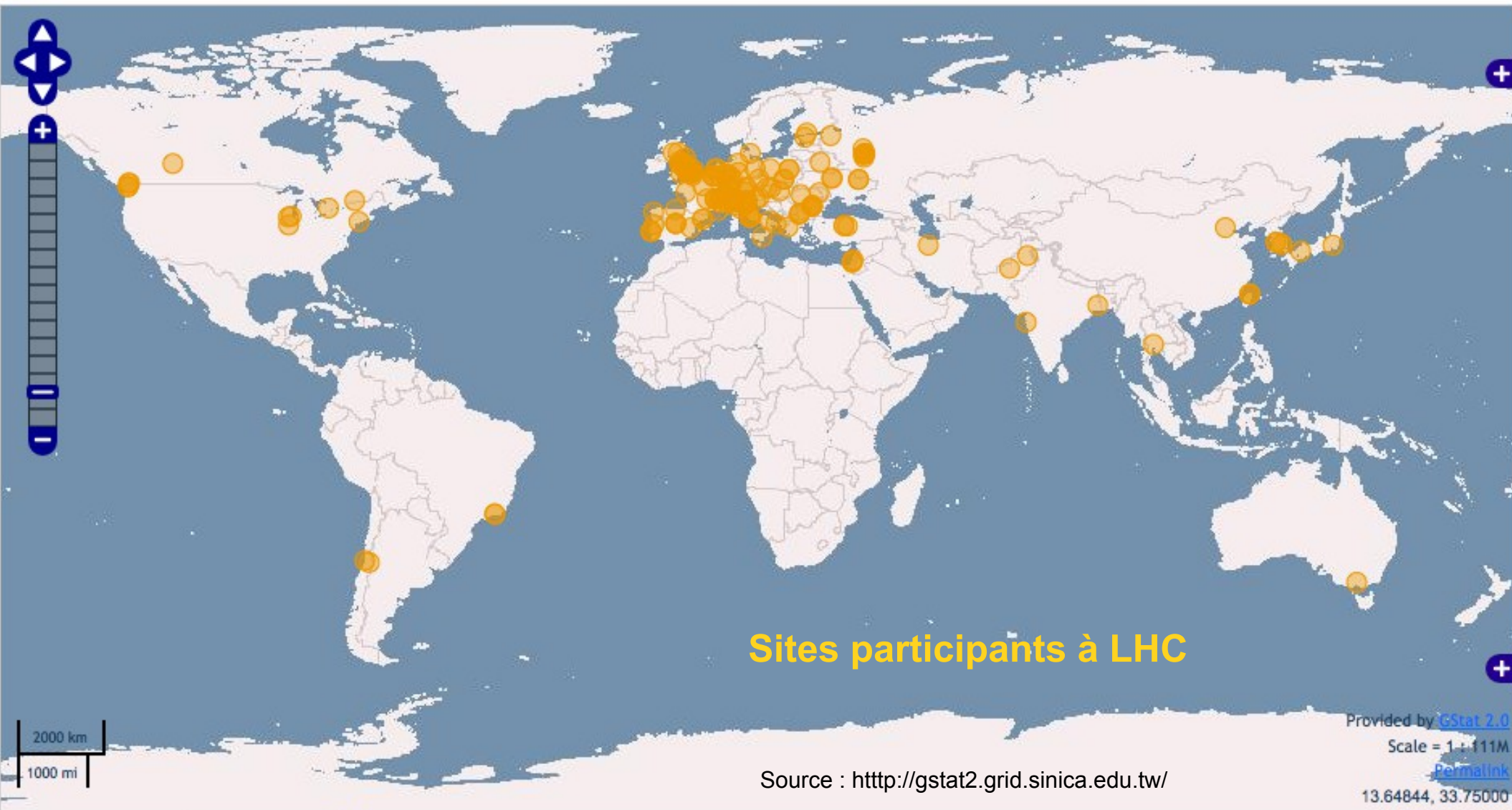


Worldwide LHC Computing Grid (WLCG)

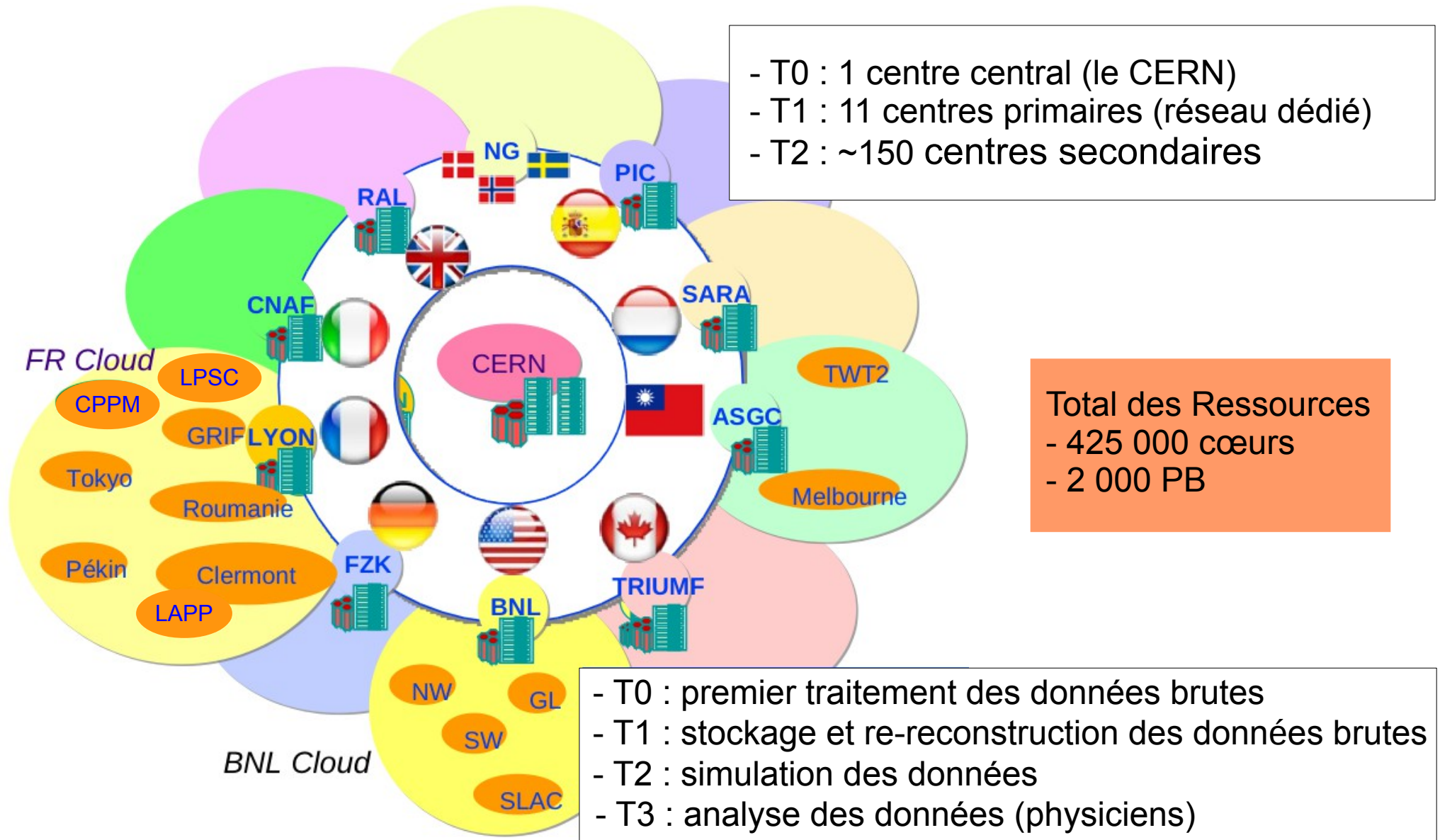
- 150 centres de calcul et de stockage
- 35 pays
- 4 expériences LHC
- **réseau privé** (LHCOPN)
- première mondiale



Géographie des sites



Hiérarchie de centres de calcul



Architecture de grille

Les grilles sont complexes.

Le *Middleware*

le **liant** : définit **la saveur de la grille**

- WLCG : expériences LHC
- EGI : *European Grid Infrastructure*
- OSG : *Open Science Grid*
- GRISBI : biologie, santé (FR)
- CIMENT : universités de Grenoble
- ...

Les ressources

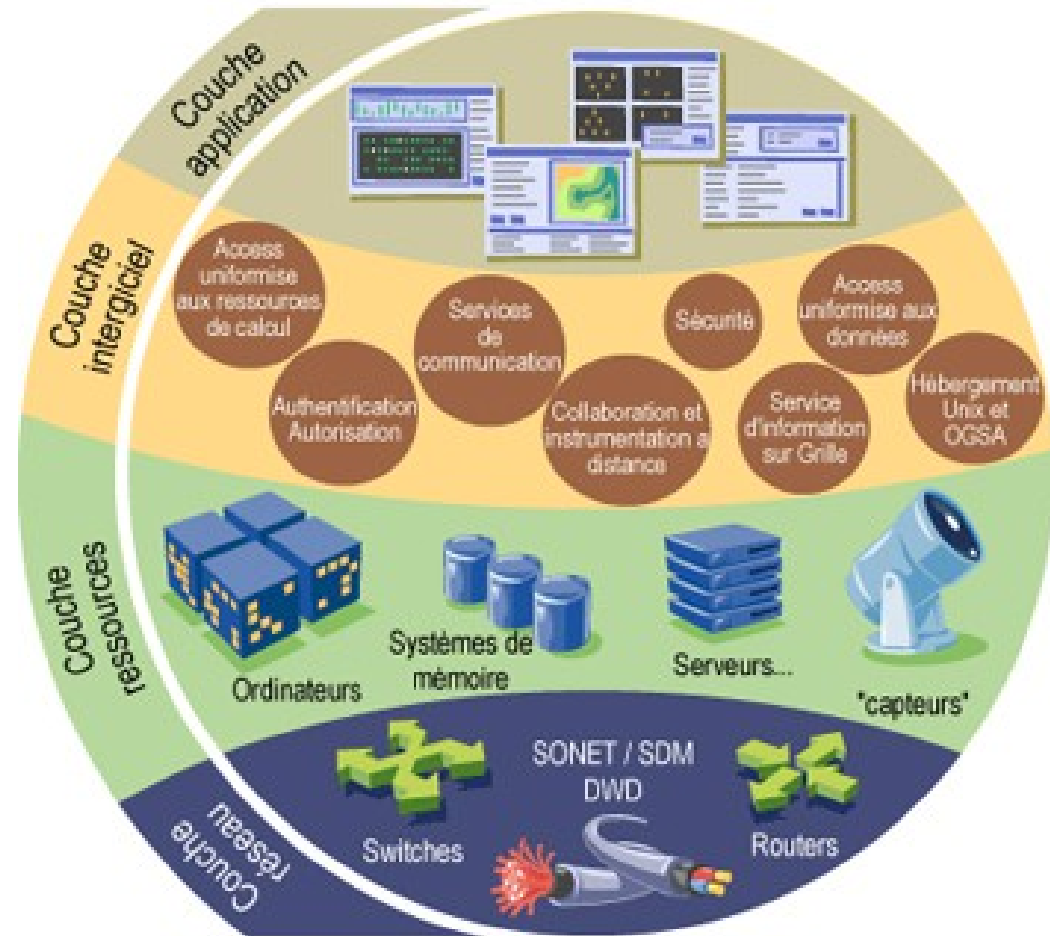
la partie **classique** d'un centre de calcul

- Calcul & stockage

Le réseau

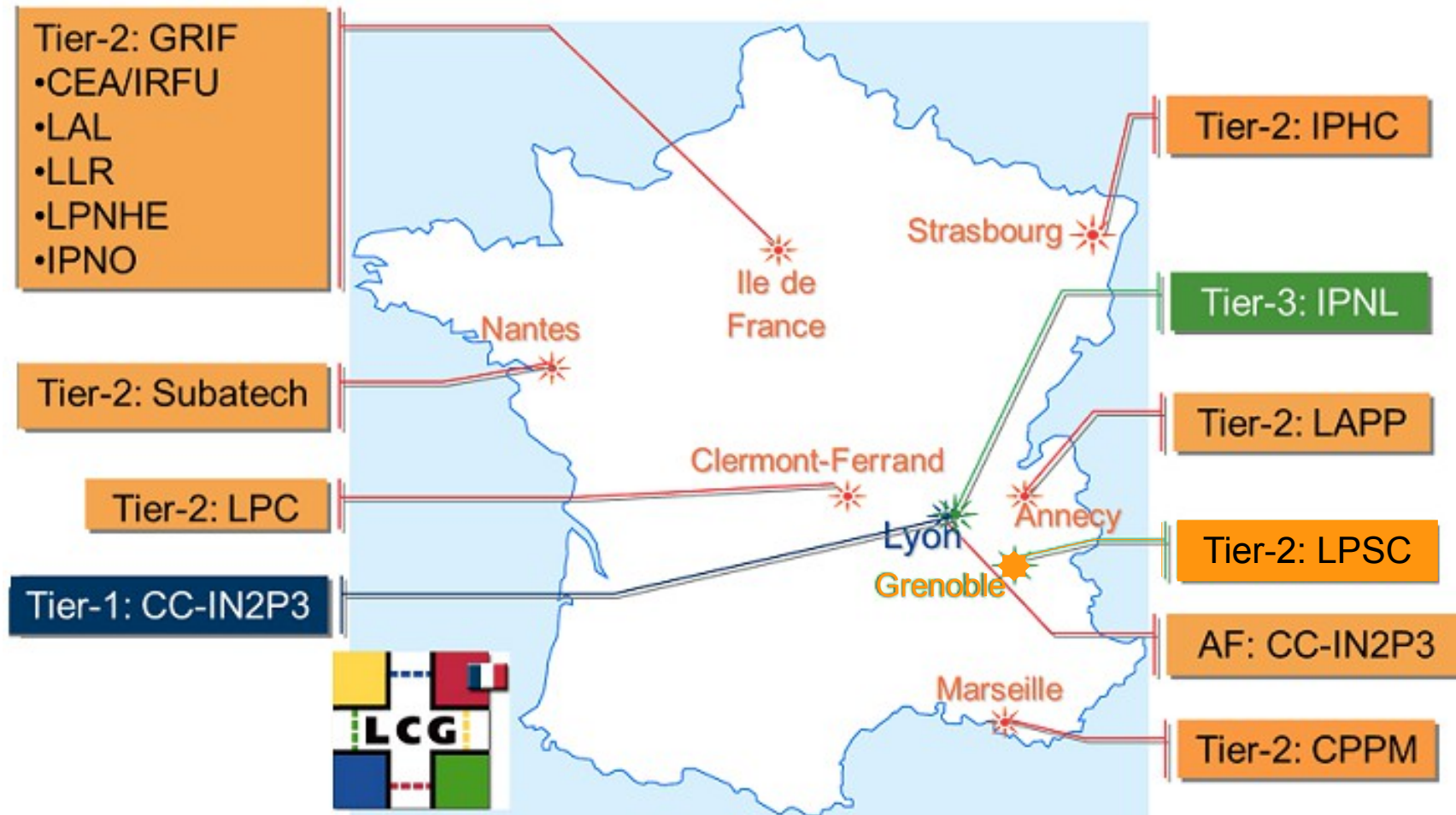
la **moelle épinière** des grilles

- 1-10 Gb/s (soit 2 CD/s !)
- 100 Gb/s est disponible



Ref : gridCafe

Centres de calcul WLCG en France



Le Tier-1 français



Centre de Calcul de l'IN2P3, à Lyon

- ~100 ingénieurs
- Site majeur des expériences LHC (4 fois T1 et T2)
- Seul site WLCG si complet dans le monde
- Ouvert sur d'autres disciplines (sciences de la vie, ...)



Grille

- Site grille complet, leader en FR (services centralisés)
- Assure ~10% des ressources de WLCG

Ressources

- 18 000 tâches de calcul simultanées
- Flot de plus de 100 000 tâches / jour
- 30 PB de stockage (disque et bande)
- LHC ~ 2/3 du centre

Accords CERN-sites

- Rythme 24/7
- Disponibilité > 98%



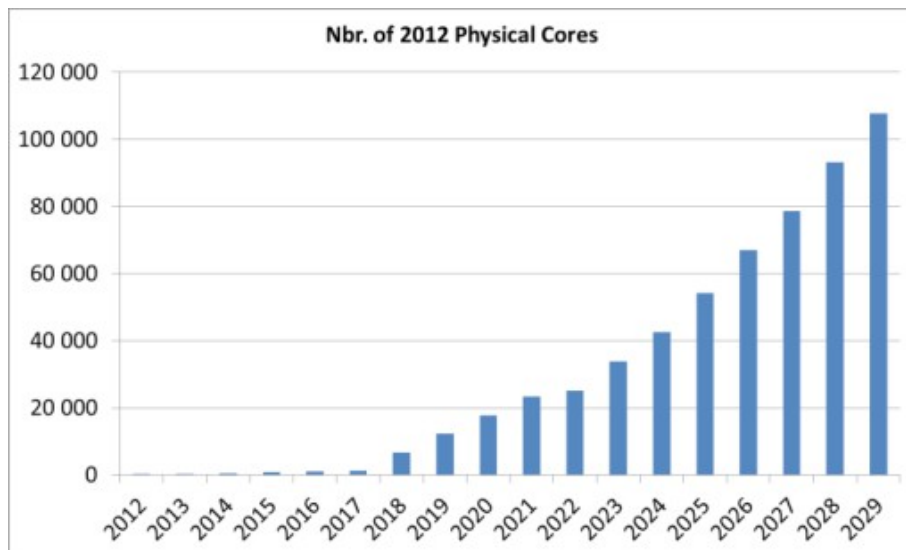
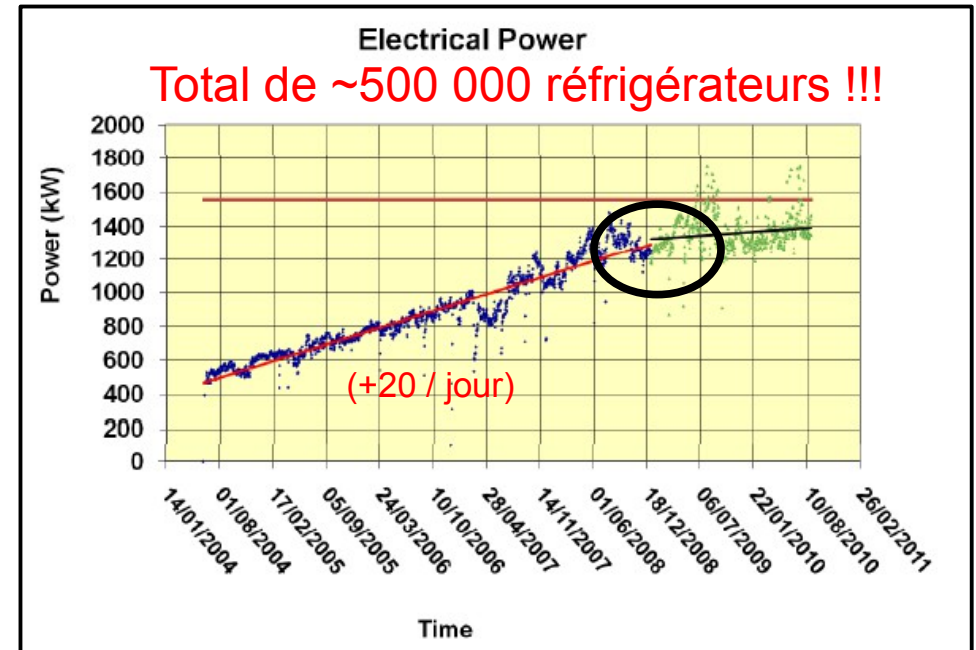
Le Tier-1 français

Puissance électrique

- **Difficulté majeure** dans les centres de calcul

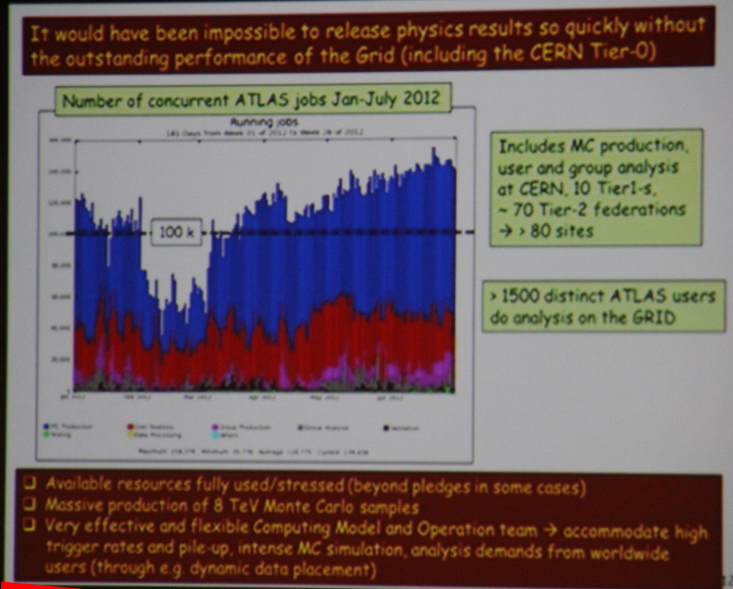
Nouvelle salle informatique

- 2011/2015/2019: 0,6/1,5/3,2 MW
- **Atout majeur** pour les projets scientifiques dans notre discipline (par exemple upgrade du LHC, LSST)



« Computing enables physics »

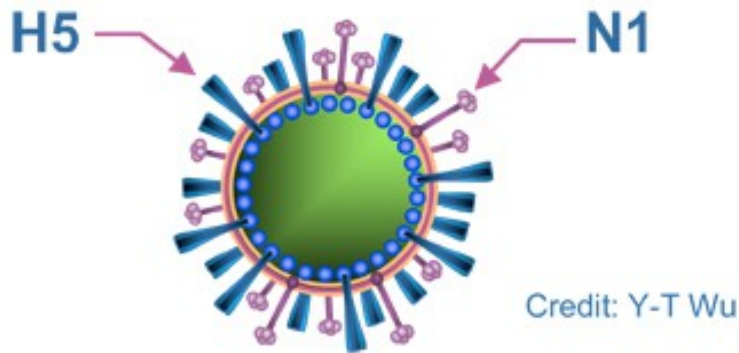
Photography: C. Biscarat



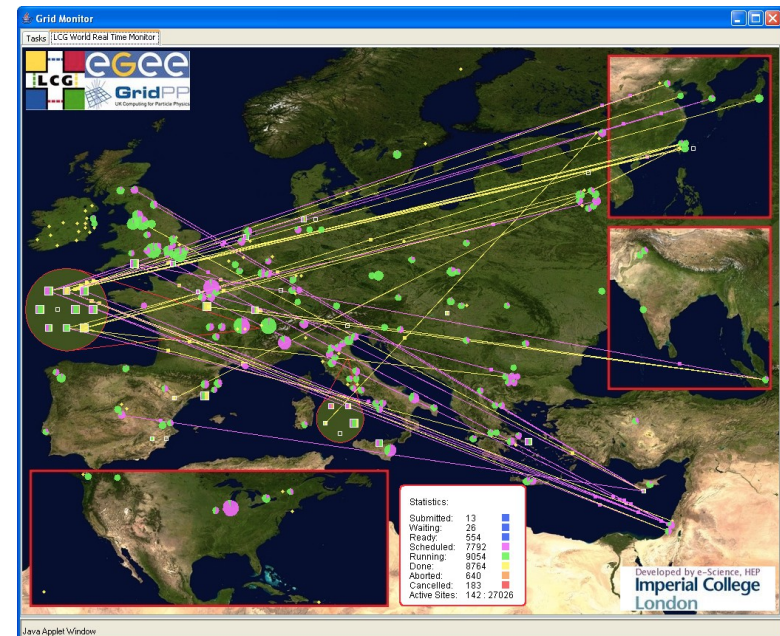
CERN seminar,
July 4th 2012,
retransmitted at
ICHEP (Melbourne)

Grippe aviaire

- Calculs de probabilité d'accrochage sur les sites actifs du virus pour le bloquer
- Accélération avec utilisation de la Grille de calcul Européenne EGEE (EGI)
- Avril 2005 :
 - 300 000 calculs de probabilités
 - **100 années** de calcul sur un ordinateur



Sites participants



Et maintenant ?

Derrière le **SUCCÈS** de la grille

- ▶ Complexité importante
- ▶ Efforts humains importants



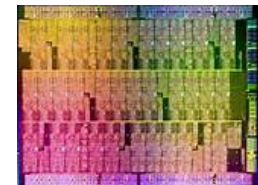
Pendant que nous développons la grille

- ▶ D'autres disciplines traitent maintenant de grandes masses de données
- ▶ Apparition de protocoles/processus standards
- ▶ Naissance des technologies de « clouds »
- ▶ Evolution technologique vers de nouvelles architectures



amazon

facebook



Notre expérience de la grille → simplification des opérations

Nouvelles idées sur le marché → intégration dans le calcul pour le LHC

Au-delà de la physique des particules

Un modèle de calcul très particulier

- ▶ Depuis toujours : de très nombreux petits événements indépendants
- ▶ La grille de calcul est idéale (**High Throughput Computing**)

Besoin en super-calculateurs

- ▶ Un calcul doit être traité sur des centaines de cœurs à la fois
 - ▶ Communication extrêmement rapide entre les nœuds
 - ▶ Grande mémoire disponible
- ▶ QCD sur réseau, astroparticule
- ▶ HPC (**High Performance Computing**)
 - ▶ Pyramide de Tiers-0/1/2 en France
 - ▶ Machine Curie : 92 000 cœurs, 15 PB de stockage
 - ▶ Sismologie, mathématiques, chimie, mécanique des fluides...

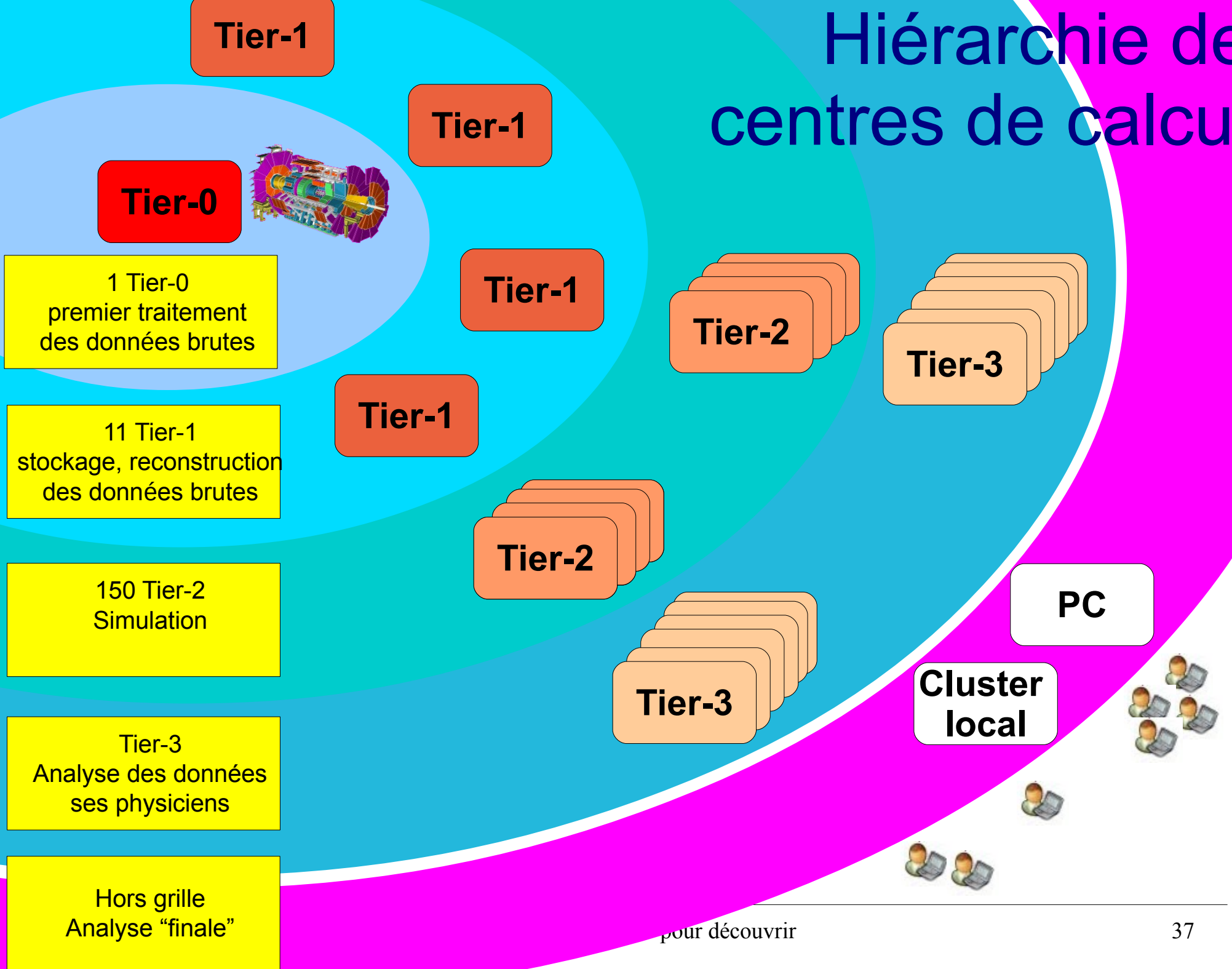


Résumé – points forts

- Recherche de processus très **rare**s cachés dans un flux de collisions banales
 - Besoin de simulations détaillées (détecteurs très précis)
 - **Enorme besoin en informatique** (stockage, calcul, transfert, logistique)
 - **Un papier et un crayon ne suffisent pas à analyser ça !**
- Développement de **la grille de calcul pour le LHC**
 - Pionnière, mondiale
- **« Computing enables physics »**
 - Les résultats sont là - toute **cette chaîne est performante !**
 - **Découverte du Higgs annoncée il y a un an**
 - Richesse d'autres résultats de physique
- **Retombées** : recyclage de nos techniques pour d'autres disciplines et le public
 - **Impact dans la qualité de vie (médecine, ...)**
- Les infrastructures de calcul sont lourdes
 - la France est très bien placée !
 - **Les développements continuent**

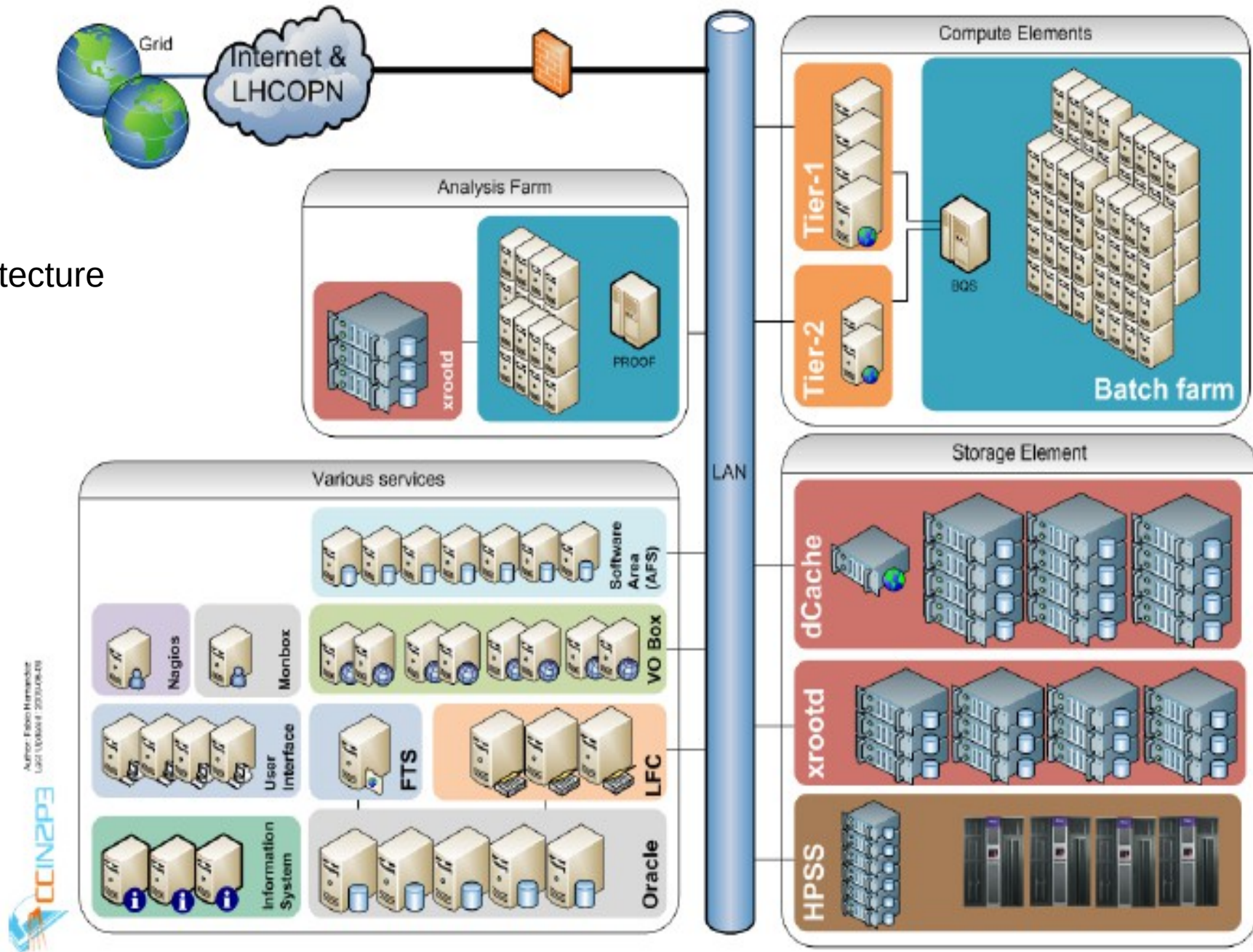
Back-up

Hiérarchie de centres de calcul

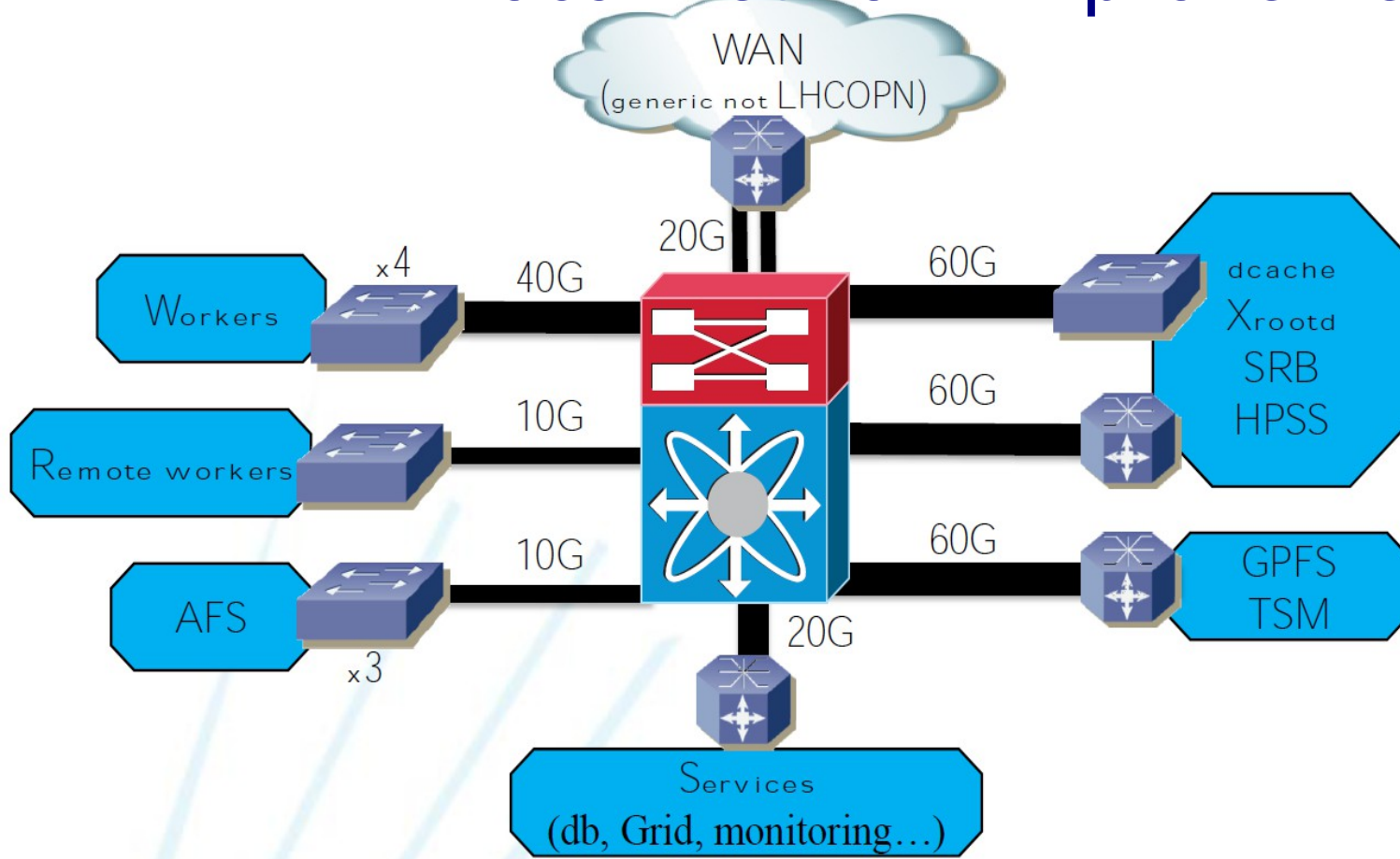


CC-IN2P3 site overview

- Shared architecture



Local network improvement



- Increased core capacity
- Isolate services
- Remove bottlenecks
- Shortened paths

Area	Old bandwidth	New bandwidth
AFS	30G <i>shared</i>	30G
GPFS TSM	40G <i>shared</i>	60G
Dcache Xrootd srb HPSS	40G <i>shared</i>	120G
Workers	40G <i>shared</i>	170G
WAN	10G	20G

Operations at CC-IN2P3

- MOU requirements for T1 centres:

Memorandum of Understanding between CERN and the institutions participating in WLCG

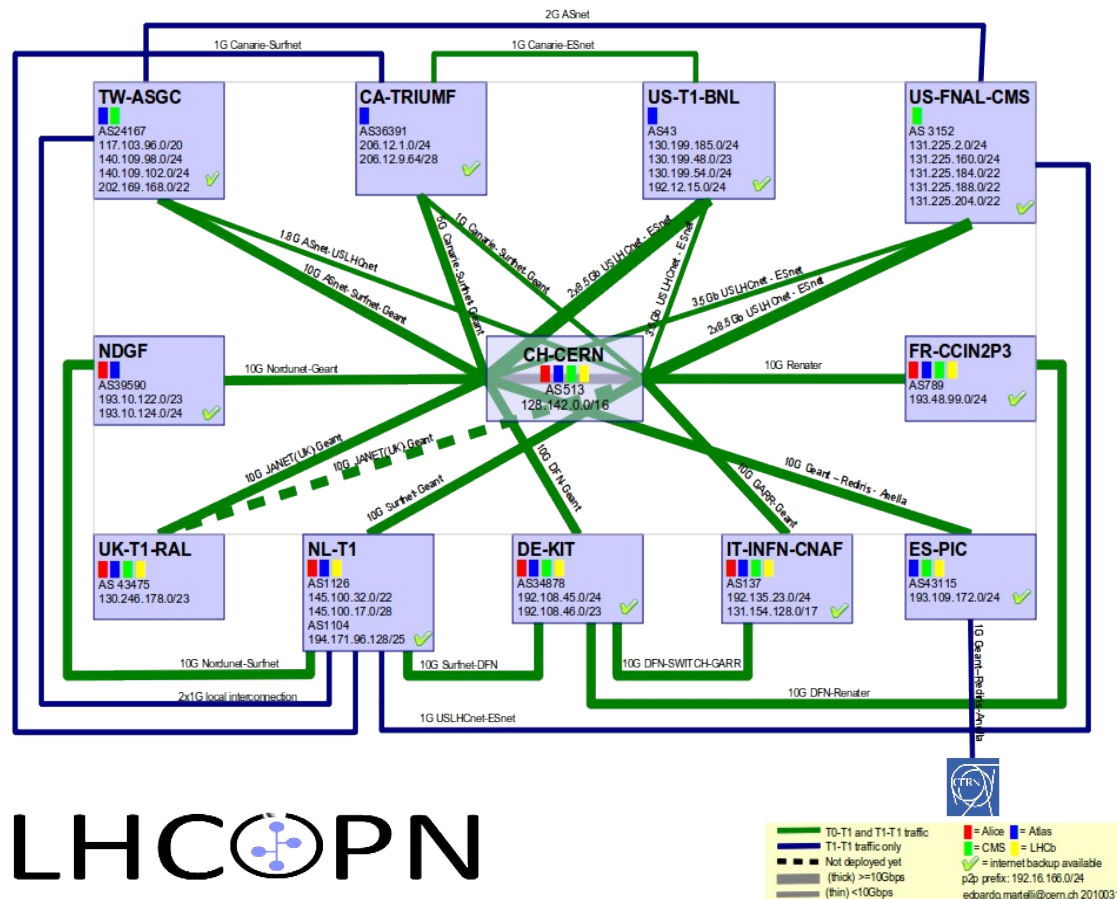
(CERN-C-RRB-2005-1/
Rev. April 2009)

<i>Service</i>	<i>Maximum delay in responding to operational problems</i>			<i>Average availability⁵ measured on an annual basis</i>	
	<i>Service interruption</i>	<i>Degradation of the capacity of the service by more than 50%</i>	<i>Degradation of the capacity of the service by more than 20%</i>	<i>During accelerator operation</i>	<i>At all other times</i>
Acceptance of data from the Tier-0 Centre during accelerator operation	12 hours	12 hours	24 hours	99%	n/a
Networking service to the Tier-0 Centre during accelerator operation	12 hours	24 hours	48 hours	98%	n/a
Data-intensive analysis services, including networking to Tier-0, Tier-1 Centres outwith accelerator operation	24 hours	48 hours	48 hours	n/a	98%
All other services ⁶ – prime service hours ⁹	2 hour	2 hour	4 hours	98%	98%
All other services ⁶ – outwith prime service hours ⁹	24 hours	48 hours	48 hours	97%	97%

The response times in the above table refer only to the maximum delay before action is taken to repair the problem. The mean time to repair is also a very important factor that is only covered in this table indirectly through the availability targets. All of these parameters will require an adequate level of staffing of the services, including on-call coverage outside of prime shift.

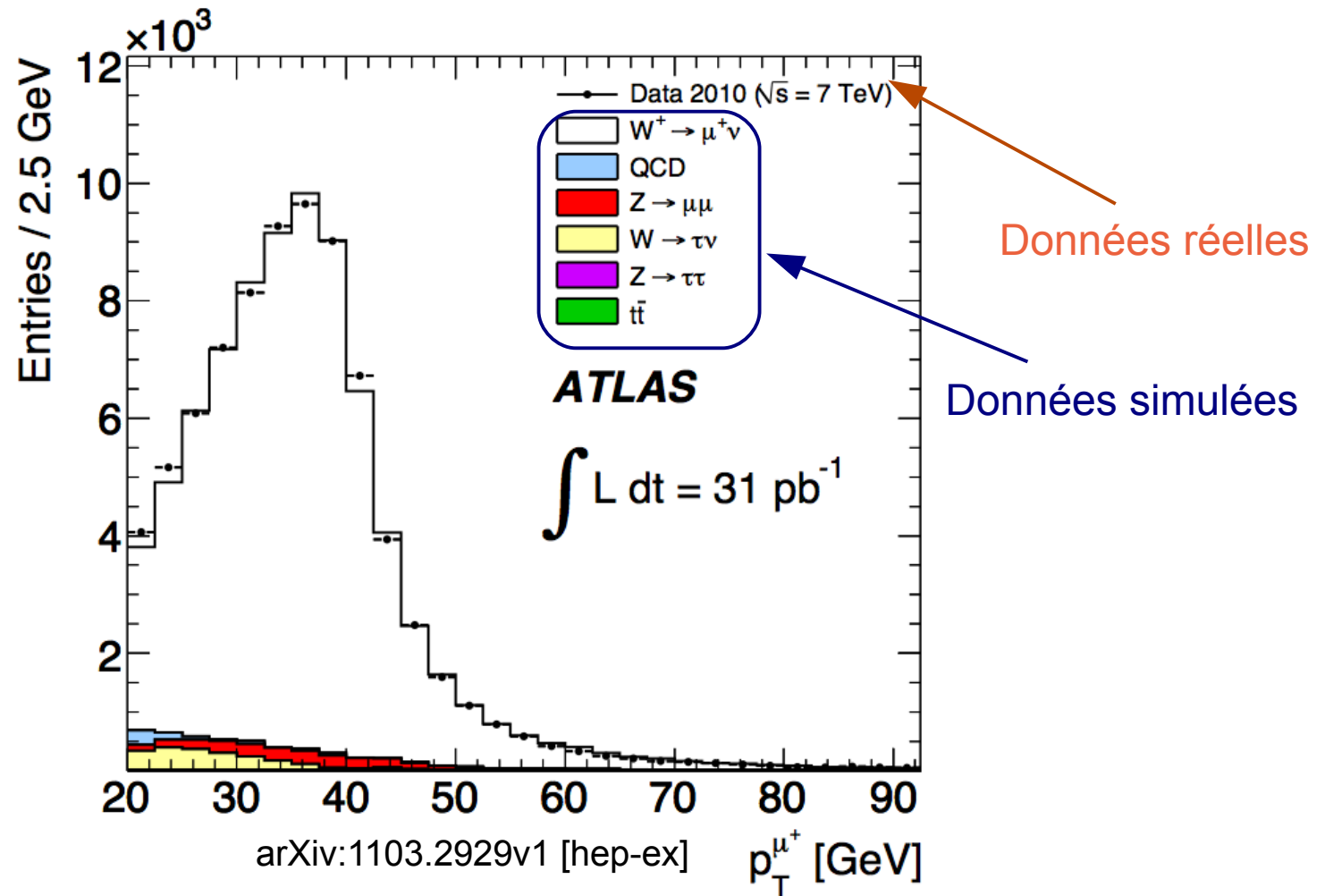
Réseaux

- Réseaux performants
- *LHC Optical Private Network*, privé et parallèle aux autres réseaux (RENATER pour la France, GEANT pour l'europe)
- T0->CC (T1) : nominal ~225 MB/s



Les chandelles standards

Sélection d'événements "W" dans les données 2010 de ATLAS.

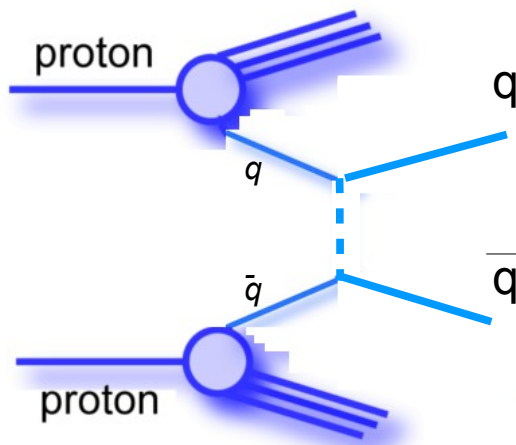


Recherche d'un processus rare : “signal” et “bruit de fond”

Nous avons déjà discuté cette chaîne de réaction.

Pour la recherche du Higgs elle constitue “le signal”.

Rappelons-nous que seuls les particules stables de l'état final atteignent le détecteur.



De mauvaises identifications dans le détecteur peuvent donner le même état final apparent – et qui sont possibles même si le Higgs n'existe pas !!

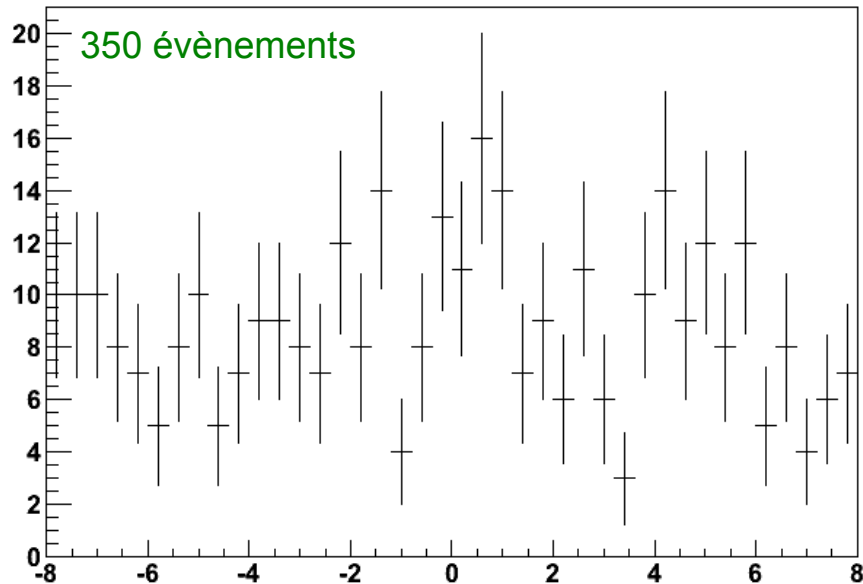
Pour la recherche du Higgs elles constituent un autre “bruit de fond”.

Temps

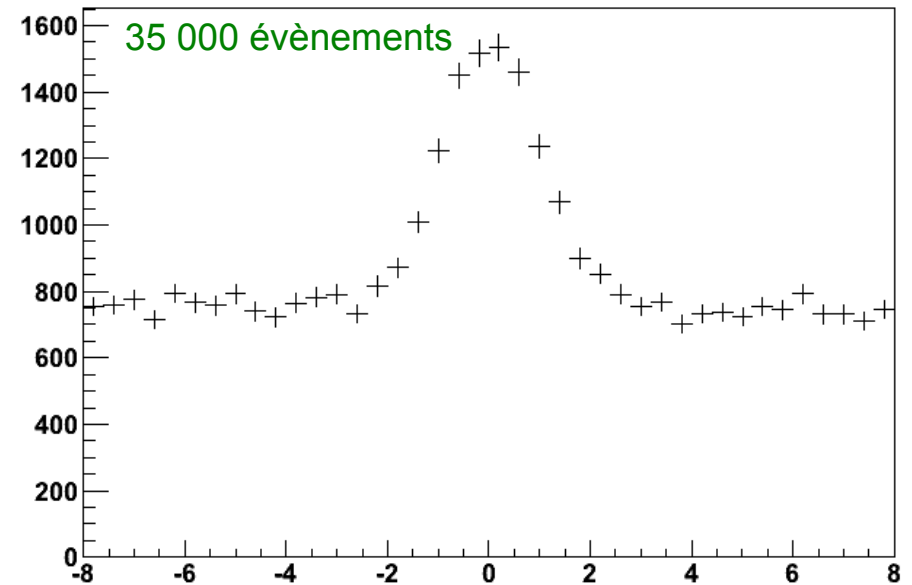
Interlude sur la statistique

Les deux graphes en bas montrent la même distribution, à gauche avec peu de données, à droite avec 100 fois plus de données :

histogram



histogram



Avec peu de données la situation est bien moins claire :

- est-ce tout simplement un spectre plat ?
- ou est-ce qu'il y a un pic quelque part ?

Avec beaucoup de données nous voyons clairement la structure : un spectre plat (dû au bruit de fond), plus un pic (dû au Higgs).

Trouver une nouvelle bestiole

Le chien



L'écureuil



Le renard

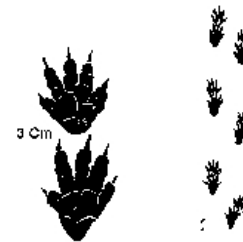


Hey ! Regarde, il y a 3 doigts !
C'est une nouvelle bestiole ?

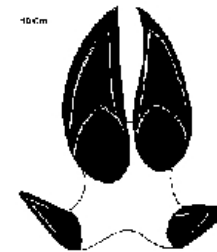


© AGSE - Droits réservés - www.scoutorama.org

Le hérisson



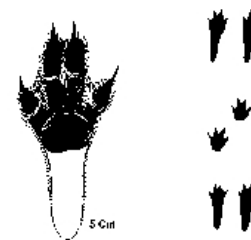
Le sanglier



1. Vieux mâle, au pas rapide.
2. Laie en marche lente.



Le lièvre

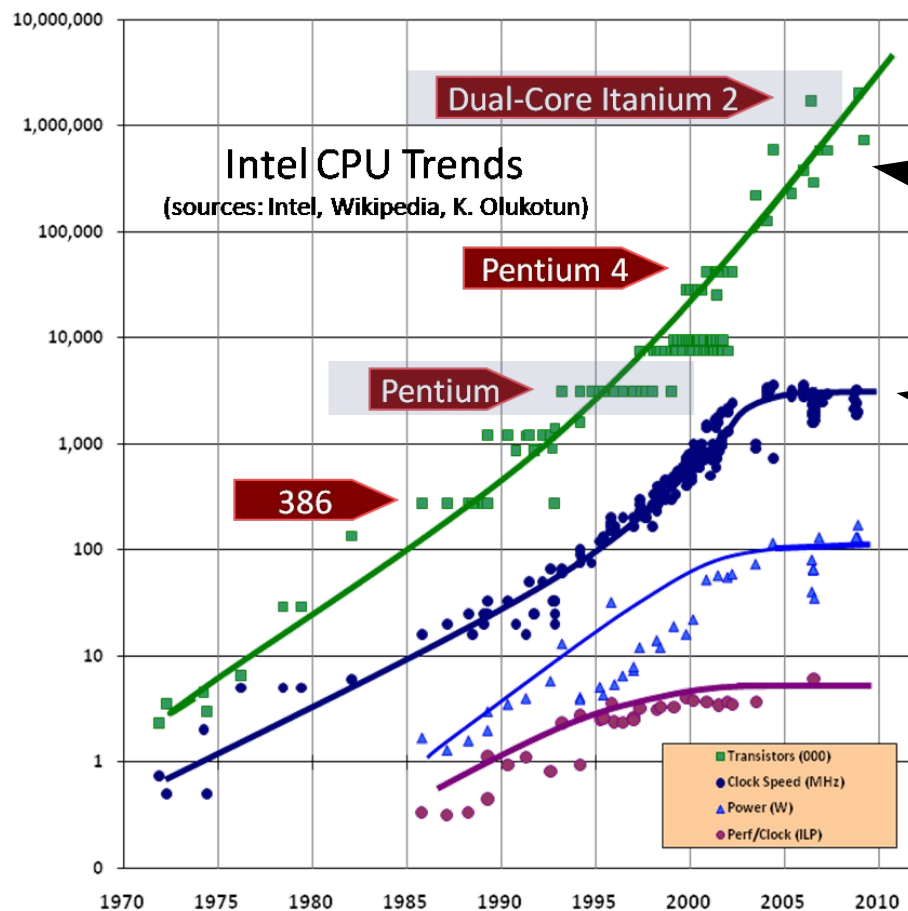


Le chevreuil



2005 : « The free lunch is over »

- Les performances des CPUs conventionnelles **ne progressent plus**
- Si les besoins en HEP augmentent, il faut **s'adapter aux nouvelles architectures** : multi-cores, many-cores, GPU (console de jeux)



Multi-cores **actuelles**
• **Limitation mémoire et IO par coeur**

Fréquence d'horloge atteint un plateau

CIMENT = mésocentre grenoblois de calcul intensif



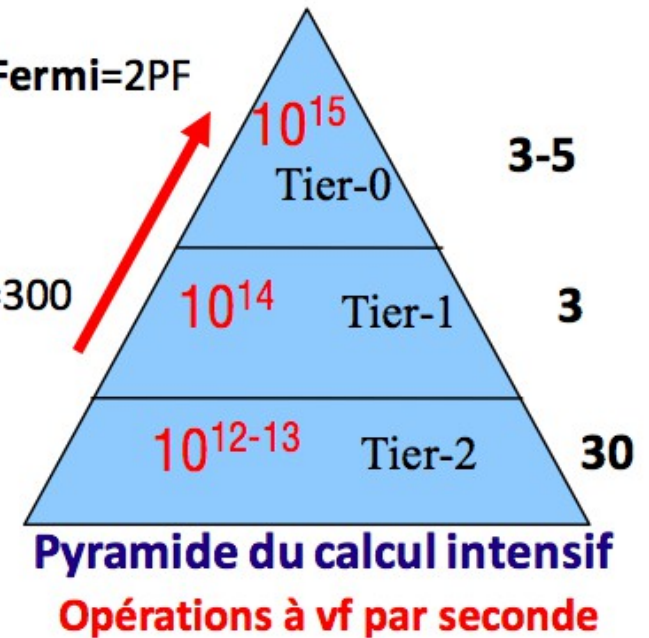
~1 Pflops/s sur une plateforme
Jugene=1PF ; Hermit=1PF ; Curie=2PF ; Fermi=2PF



~100 Tflops/s sur une plateforme
Babel = 139 ; Jade=267 ; Titane-CPU+GPU=300



32 Tflops/s répartis sur **10**
plateformes de 96 à 576 cœurs
Accessibles en mode grille (CiGRI)



Des plateformes de calcul intensif avec un accès souple

Un réseau d'ingénieurs et de chercheurs diffusant expertise et formation

Permettant d'expérimenter les outils et méthodes de calcul haute performance

Favorisant le passage vers les centres nationaux et européens

Une organisation en pôles : regroupement d'équipes de recherche, de laboratoires, qui ont développé une culture de mutualisation et de partage des ressources.

CIMENT/MaiMoSiNE crée du lien entre les pôles, et stimule les échanges d'expertise.