



10 avril 2012

Gestion des grandes masses de données et physique des hautes énergies



Dominique Boutigny

Crédits CC-IN2P3 :
Laurent Caillat-Vallet
Jean-Yves Nief
Pierre Girard
Mattieu Puel





LHC



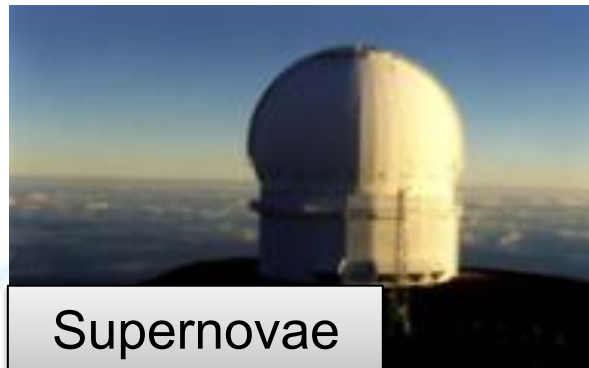
HESS



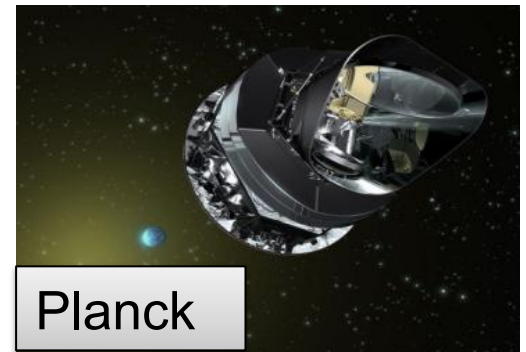
Auger



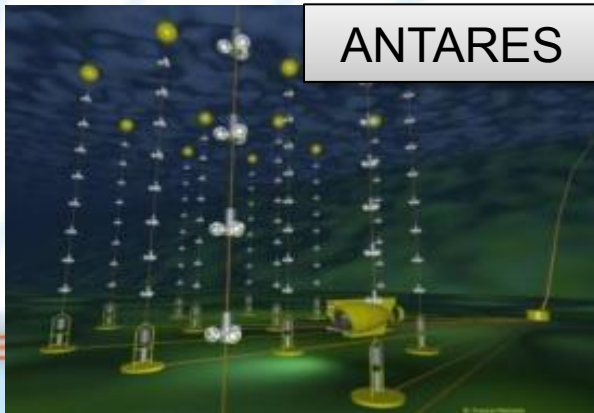
AMS



Supernovae



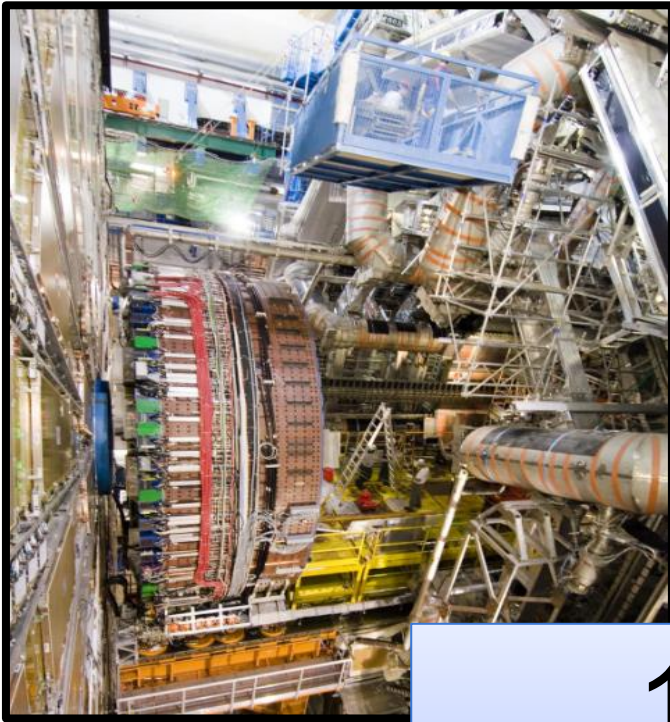
Planck



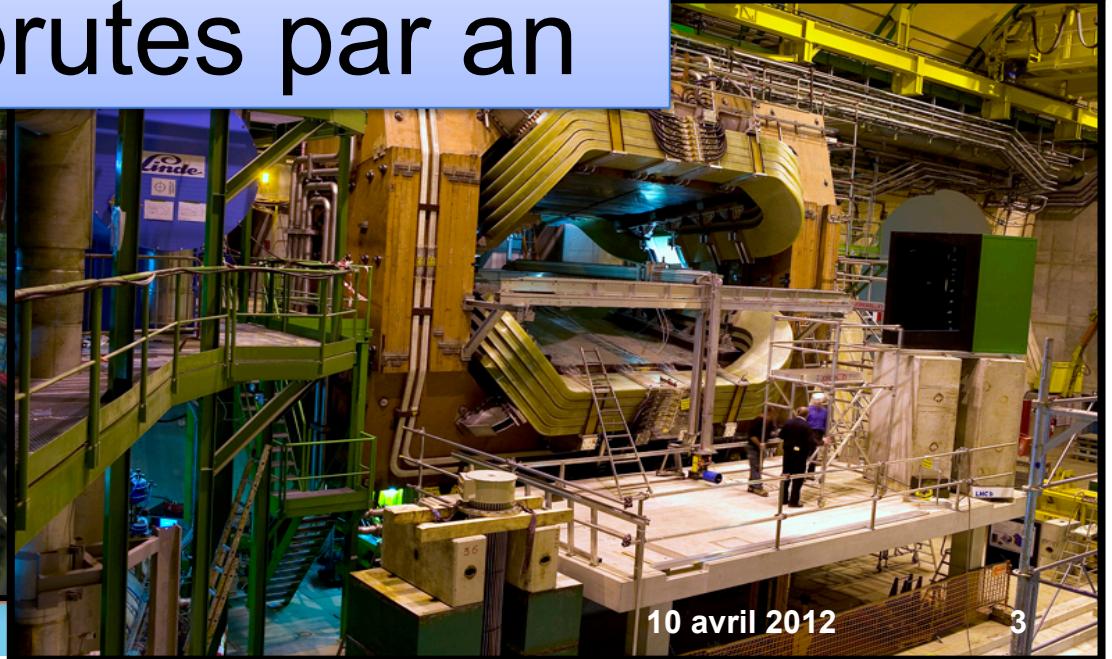
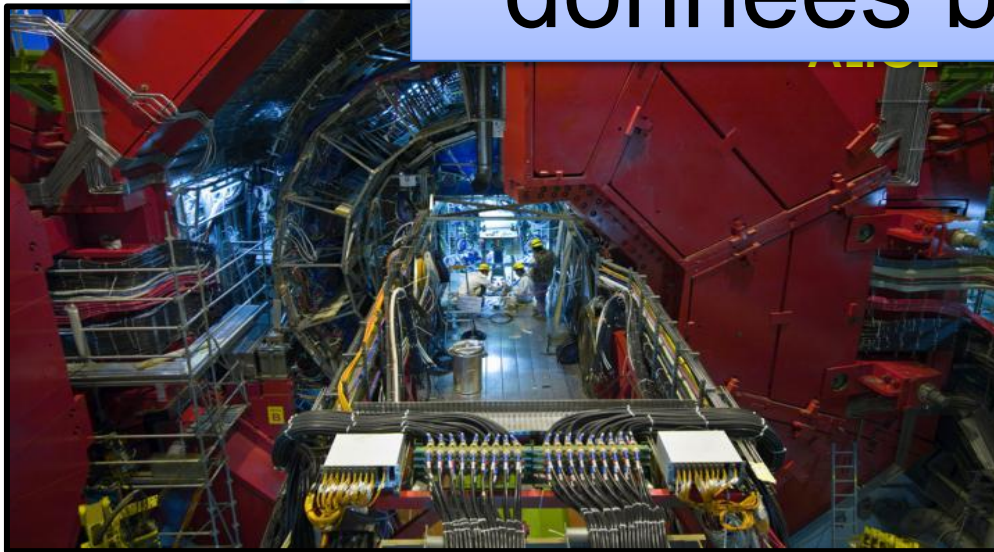
ANTARES



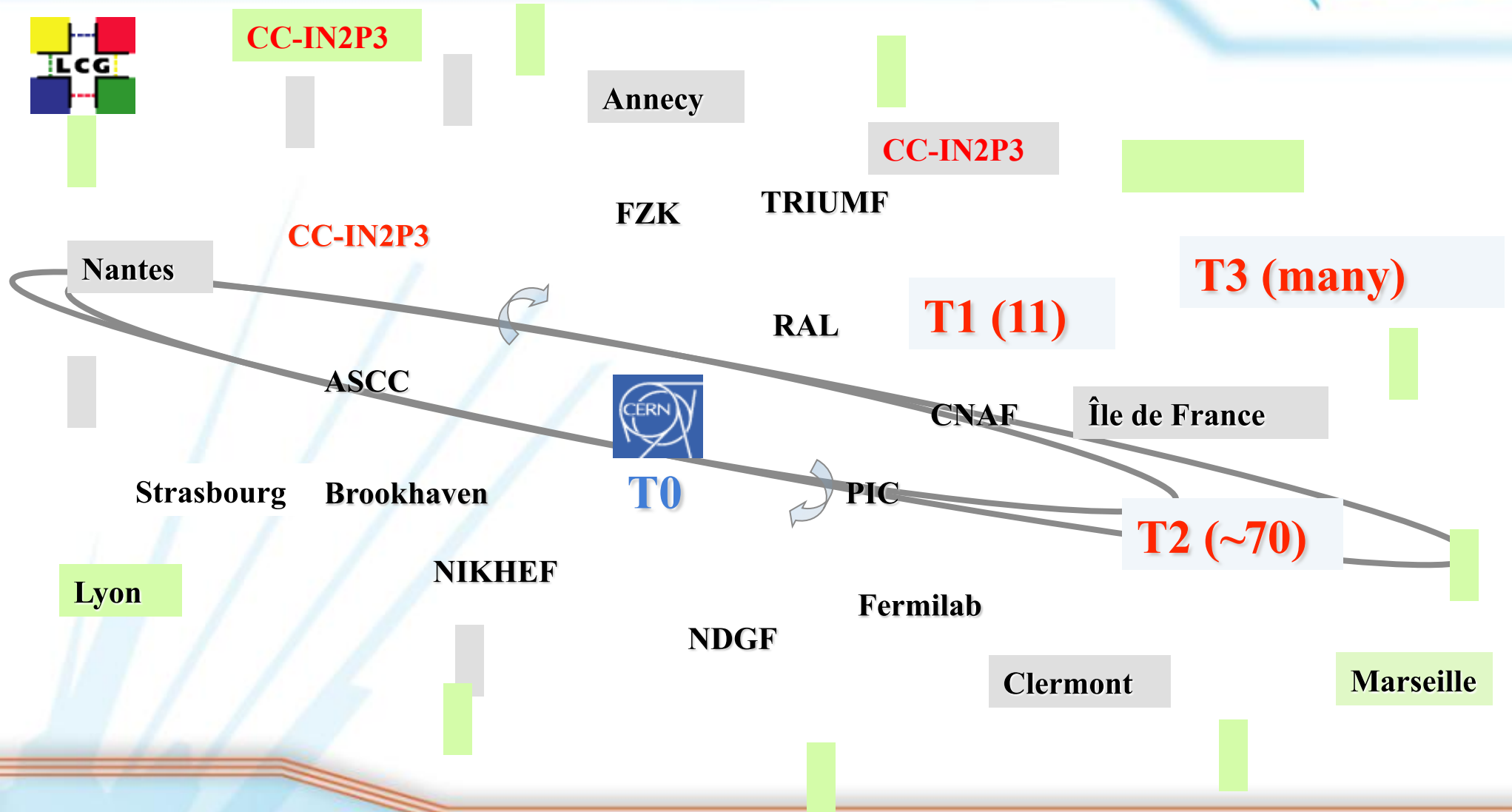
Embryogenèse



15 Petaoctets de données brutes par an



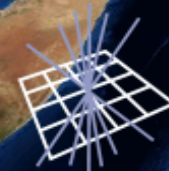
Une architecture de Grille globale pour le LHC



Scheduled = 19327
Running = 17317

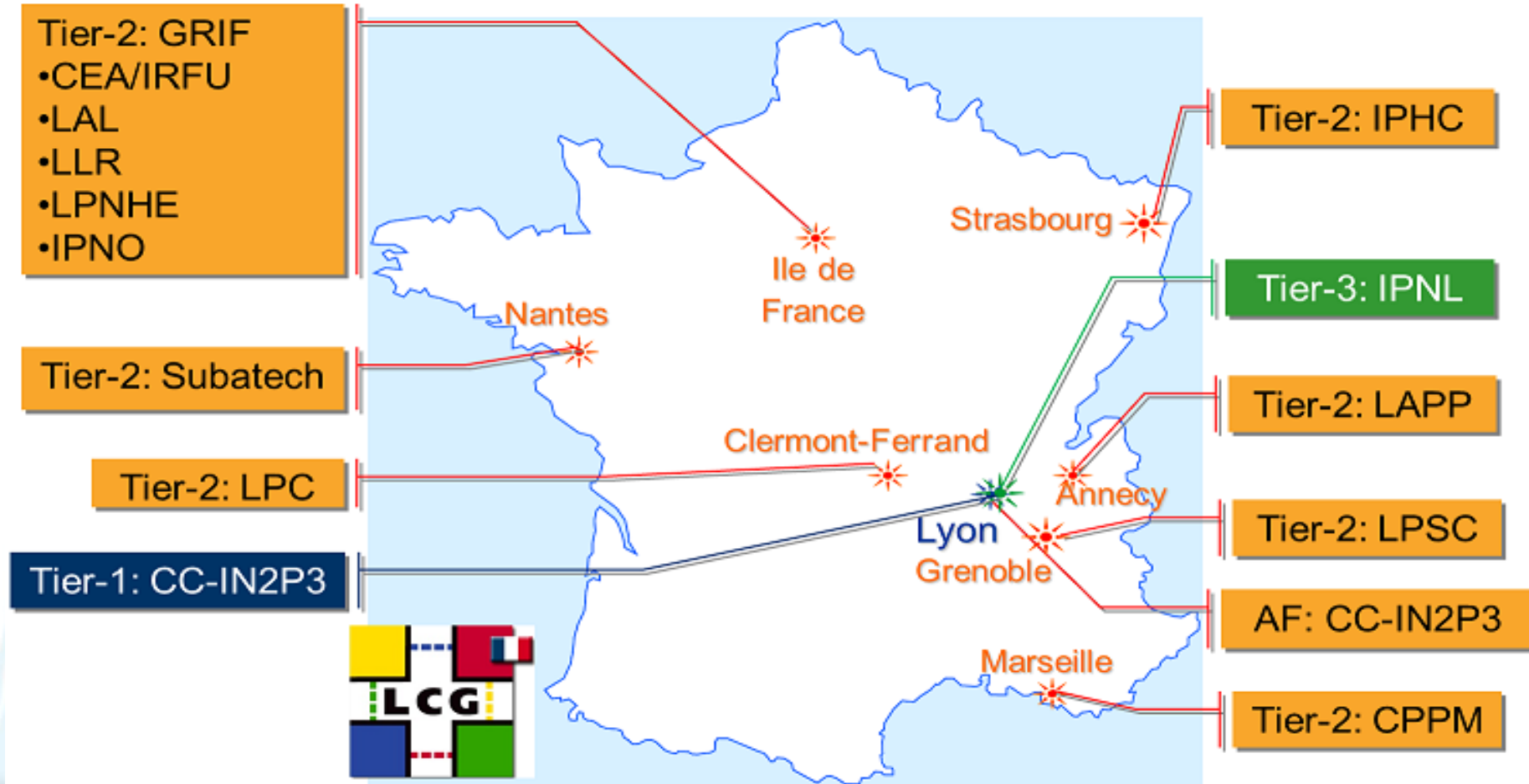
~60 000 CPU – 70 Po de
stockage – 5500 utilisateurs

16:19:25 UTC





Sites LCG-France



Traitement de données en physique HEP



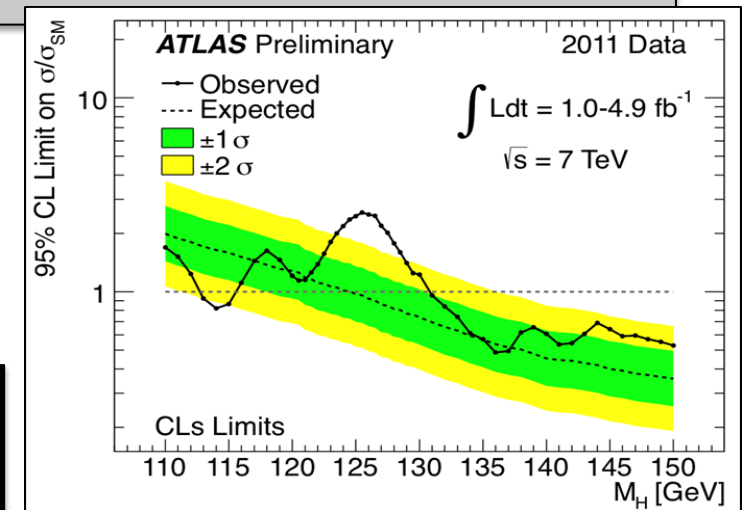
Les collisions de particules sont indépendantes les unes des autres

- Traitement de données : "Embarrassingly parallel problem"
- Chaque CPU (core / thread) traite un ensemble de données
- Les analyses de physique consistent à filtrer les données afin d'effectuer un traitement statistique

Il faut passer au crible des millions / milliards d'enregistrements d'interactions afin de localiser les quelques "événements" intéressants

Au niveau informatique il s'agit essentiellement d'un problème d'accès aux données

Le comportement des détecteurs est modélisé avec une très grande précision → Simulation



→ CPU
→ Accès aux données



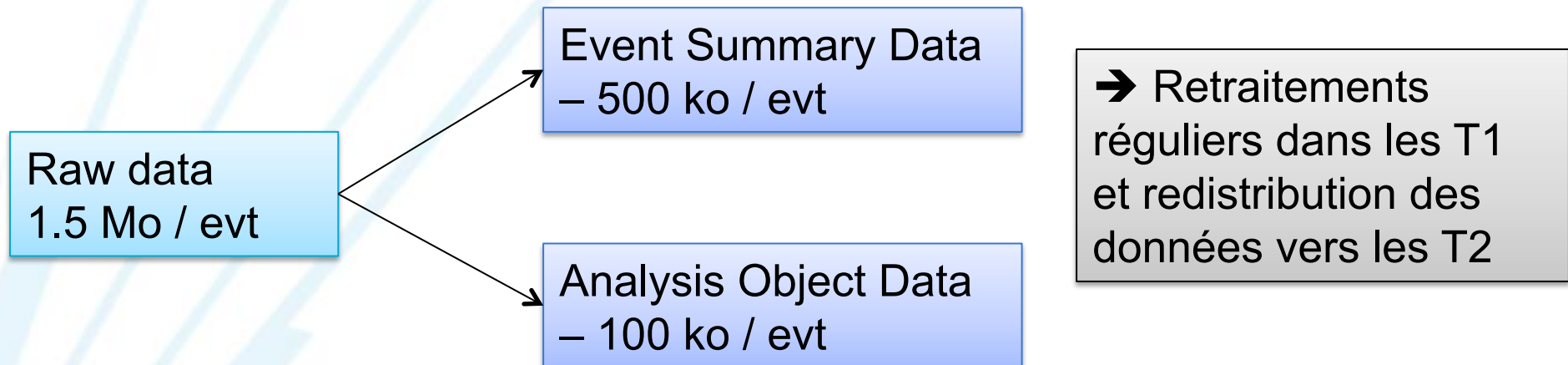
Traitement de données en physique HEP



Caractéristiques des données d'une expérience LHC (ATLAS) :

- Taille : 1.6 Mo / interaction enregistrée
- Taux d'enregistrement ~200 Hz
- ➔ ~320 Mo/s et **par expérience**

Traitement local au CERN puis distribution dans 11 centres T1



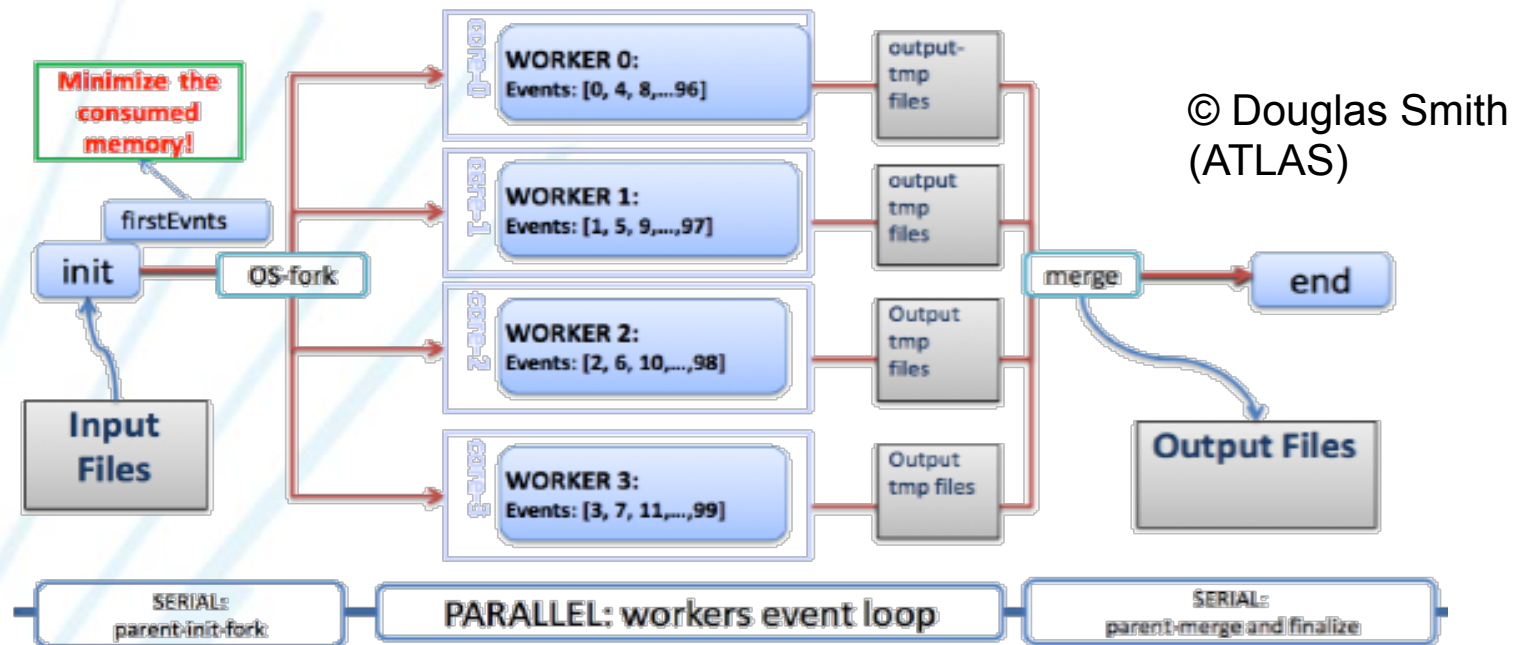
Vers une évolution du modèle de calcul



On a vécu des dizaines années avec des codes de reconstruction et d'analyse purement séquentiels → aucune exploitation du parallélisme

Changement radical : Exploitation de l'ensemble des cœurs

Demande des expériences de disposer de machines "whole node"



Gain important sur l'utilisation de la mémoire



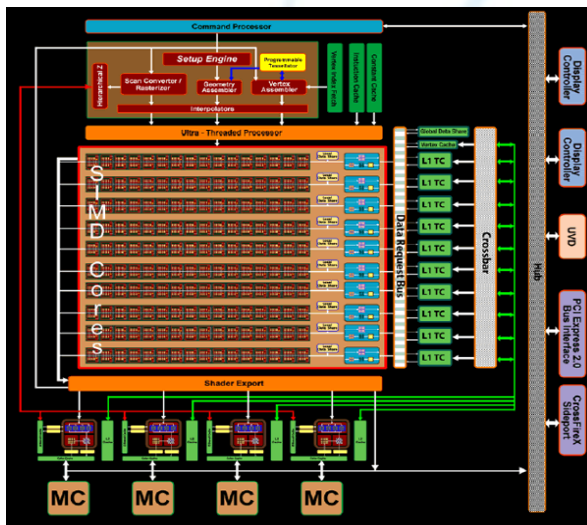
Parallélisation



La parallélisation du code se heurte à des difficultés techniques importante mais sera incontournable avec l'arrivée des machines "manycores"

➔ Effort de longue haleine – Typiquement tout le code est à ré-écrire

Des initiatives en ce sens sont lancées dans les collaborations

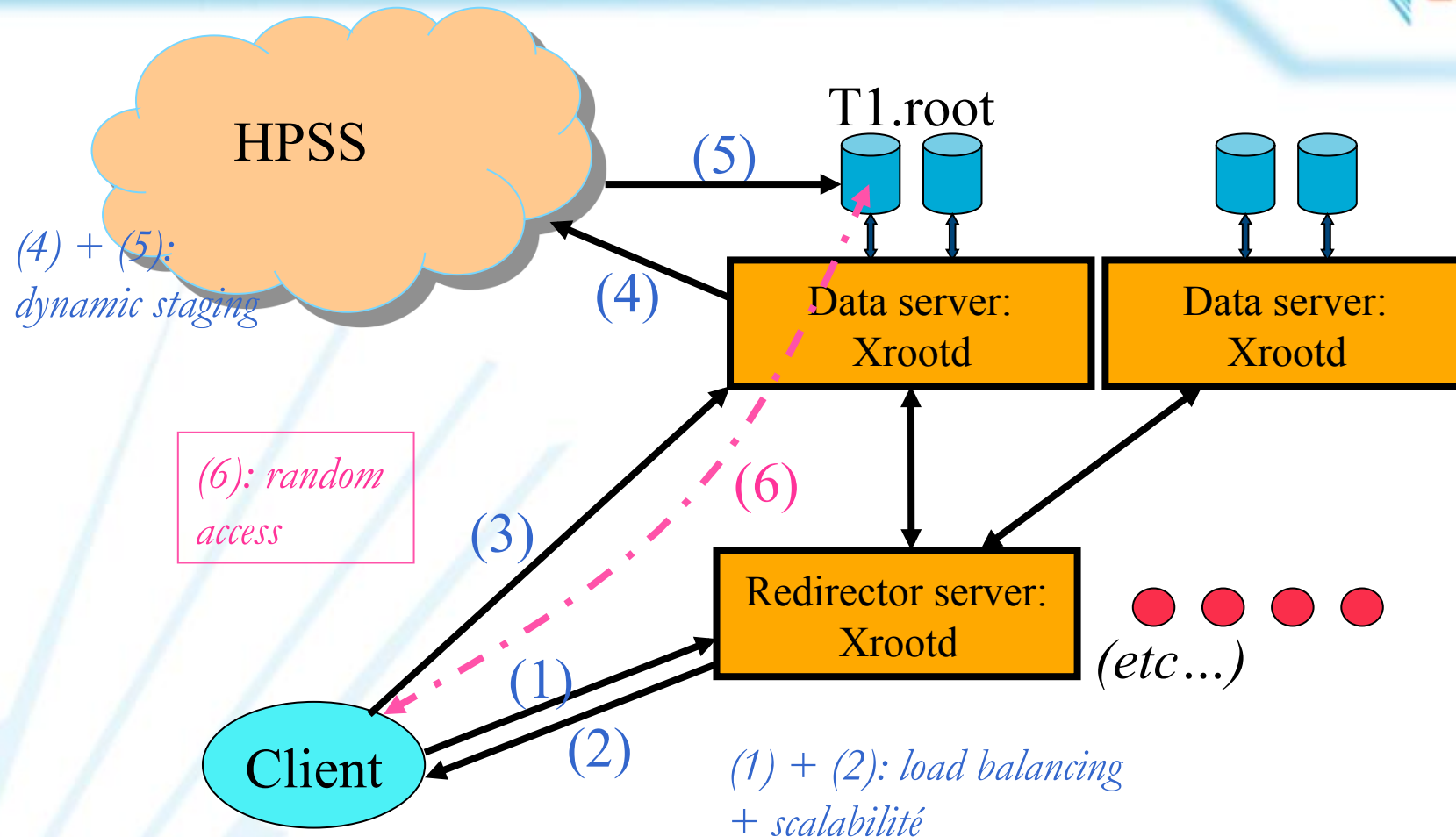


L'utilisation des GPU pour traiter certaines parties du code commence à être testée par la physique des particules, la physique des astroparticules

- ➔ Traitement d'images
- ➔ QCD sur réseau

Plateforme GPU au CC-IN2P3 dans le cadre d'un partenariat avec DELL

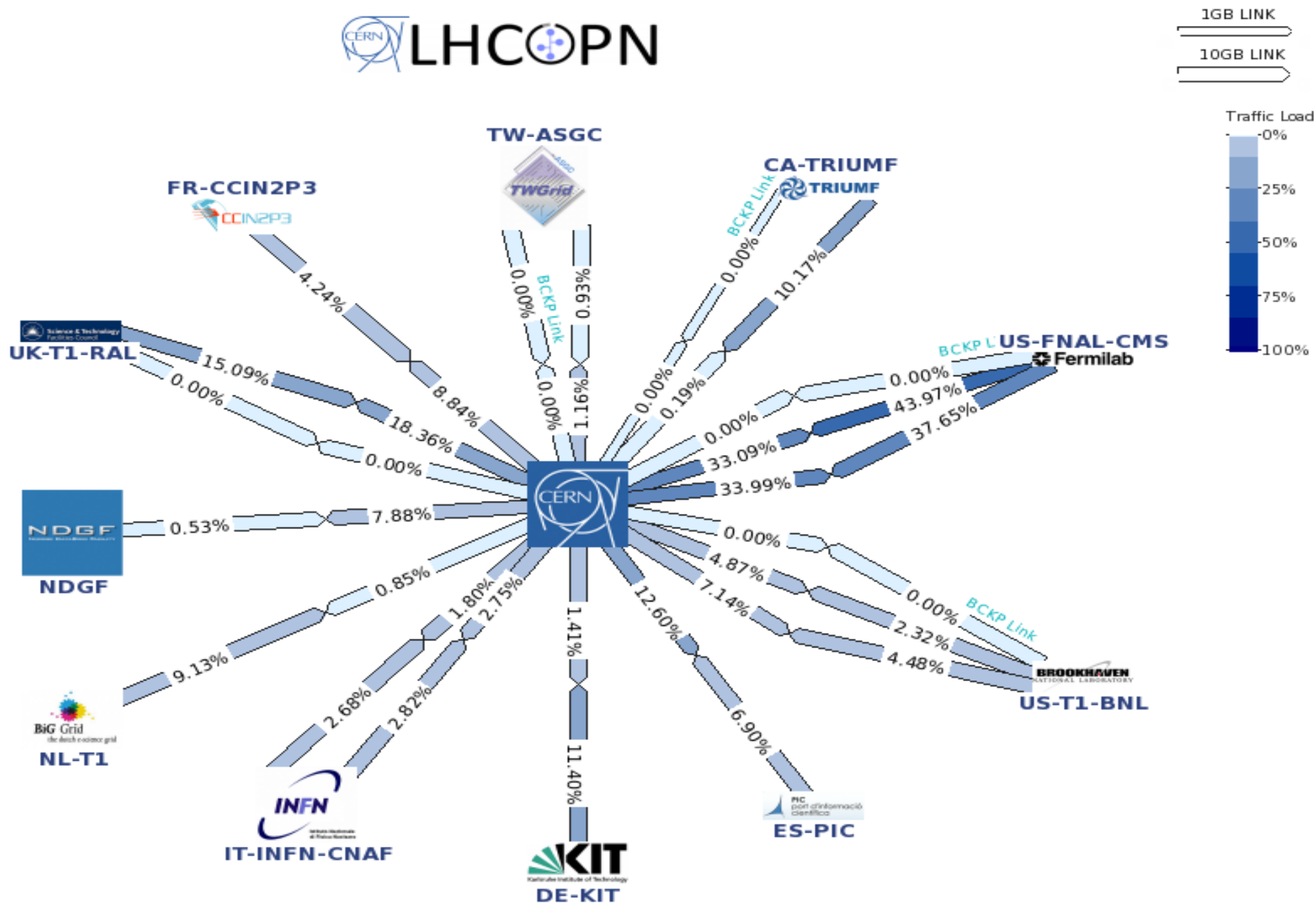
Exemple d'accès aux données : Xrootd



/hpss/in2p3.fr/T1.root ?

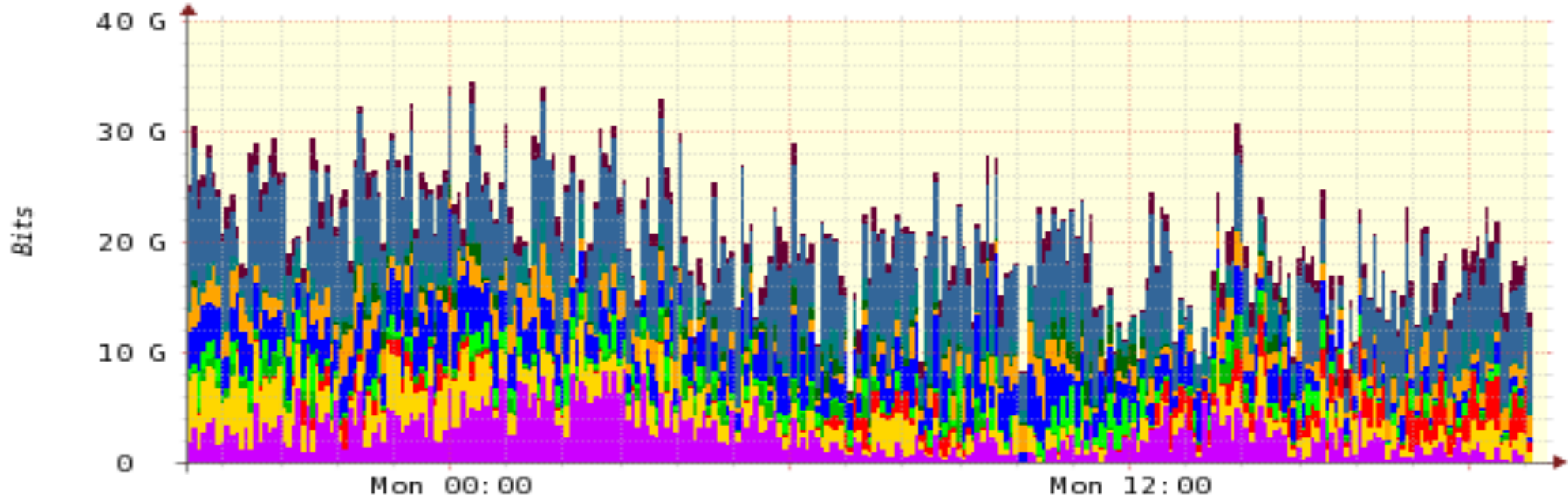
Systeme similaire : SRM / DCache

CERN LHCOPN



LHCOPN TOTAL Traffic (CERN -> Tiers1)

RRDTOOL / TOBI OETIKER

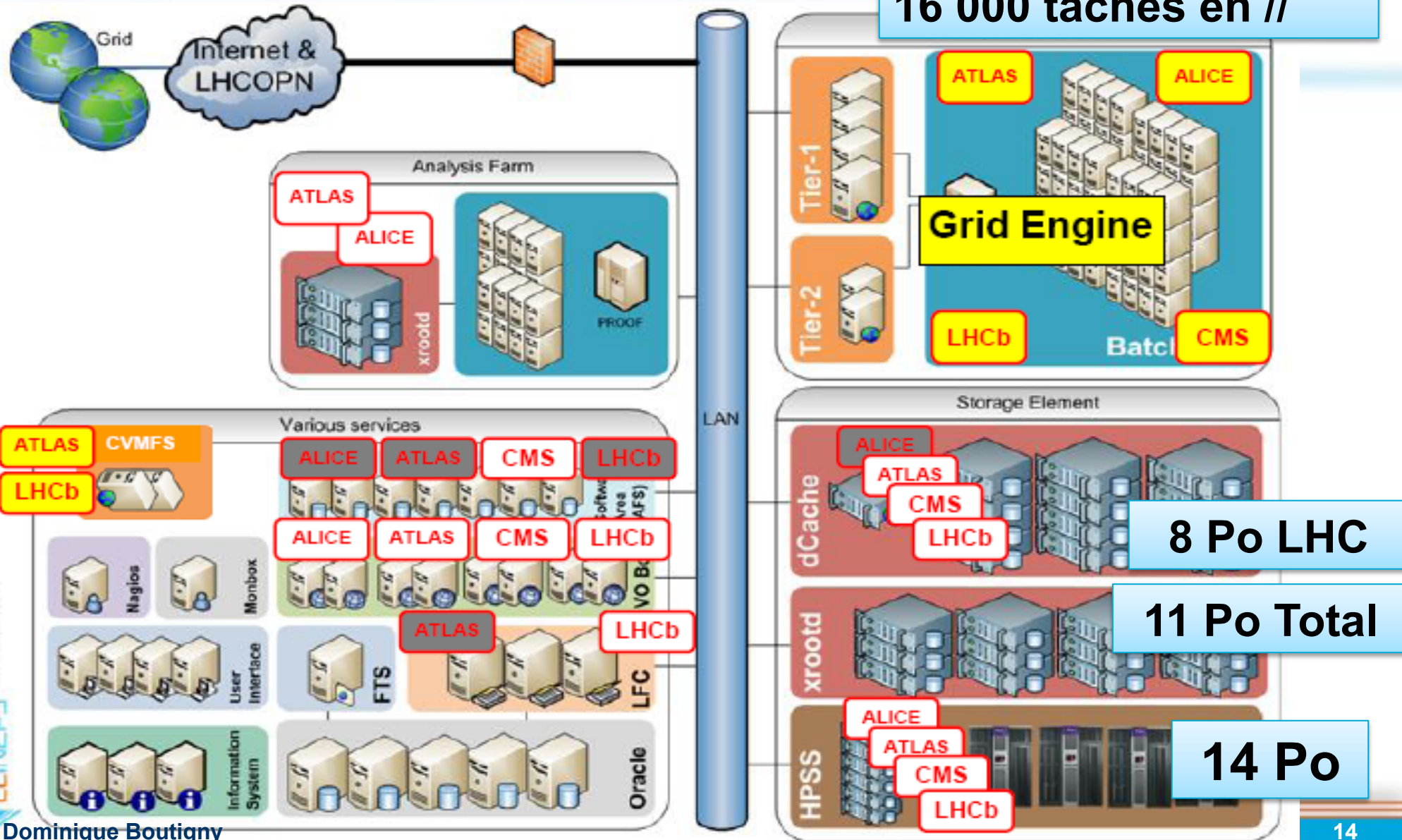


	Avg	Max	Peak
To DE-KIT	2.86G	8.49G	141.
To IN2P3	2.14G	7.41G	884.
To NDGF	919.64M	5.02G	787.
To NLT1	551.77M	3.60G	85.
To ASGC	886.18M	3.78G	115.
To CNAF	2.81G	9.29G	455.
To RAL	1.56G	4.85G	1.
To TRIUMF	424.71M	2.24G	18.
To BNL	1.27G	5.33G	1.
To FNAL	5.92G	12.08G	6.
To ES-PIC	1.09G	3.83G	1.

Total to Tiers1 Avg: 20.38G Max: 34.43G Curr: 13.49G
 Last update: Mon Mar 19 2012 19:19:31

Ressources du CC-IN2P3

16 000 tâches en //

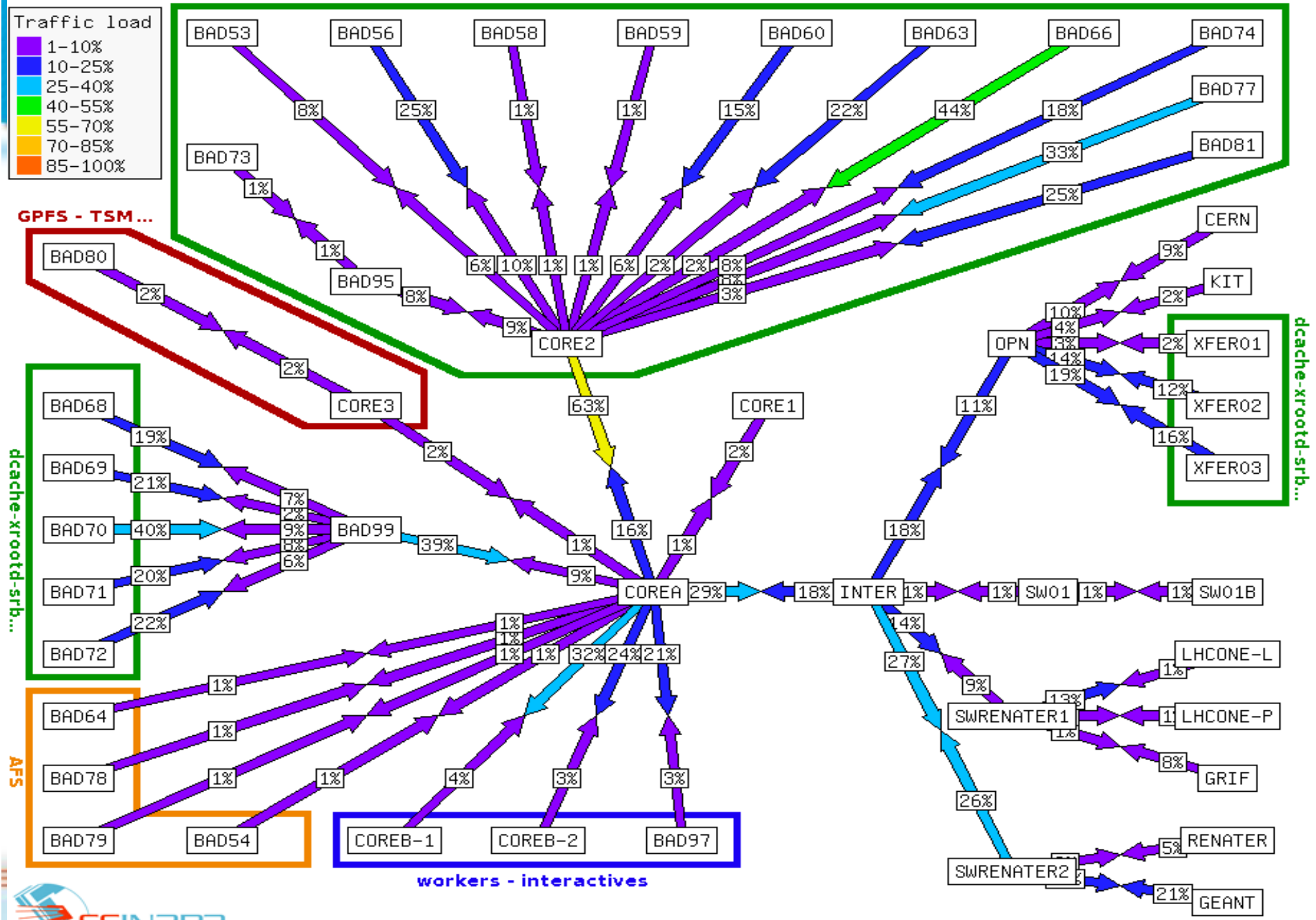
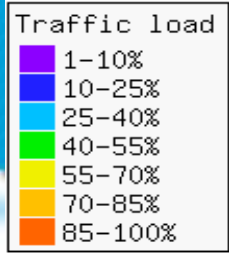


Autor: Fabrice Remondet
Last Updated: 2009-04-20



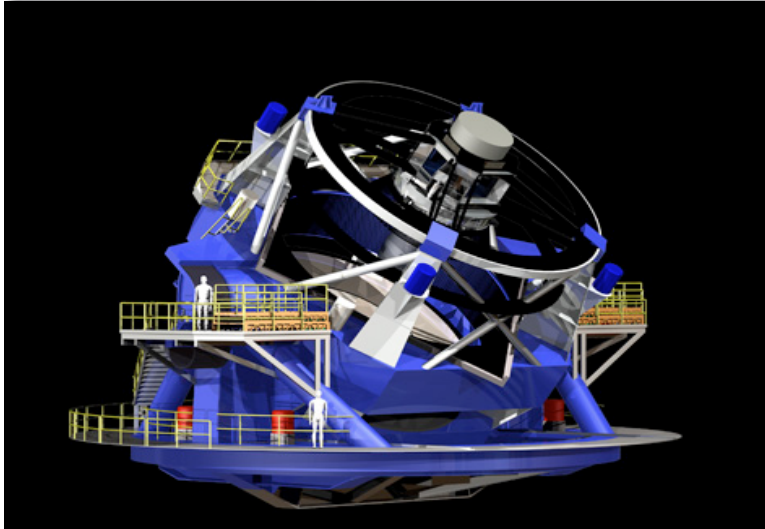
Dominique Boutigny

acsis - TSM - HPSS - dcache - xrootd - srb ...



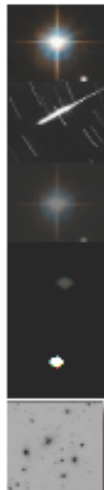


Les futures expériences d'astroparticules



LSST : *Large Synoptic Survey Telescope*
3.2 Gpixels – 1 image toutes les 15s
➔ 15 à 30 To de données chaque nuit

L'ensemble du ciel est sondé 1000 fois en 10 ans



Database Table	Billions of rows	Column width in kilobytes	Total Terabytes
Object	37.0	3.2	109.0
Moving Object	0.006	0.5	0.003
Source	4920.0	0.8	3594.0
Forced Source	32100.0	0.04	71.0
Difference Image Source	196.0	0.4	1079.0
CCD Exposure	16.0	0.6	0.6

Crédit : Jeff Kantor – LSST Corporation

Stockage des données dans LSST

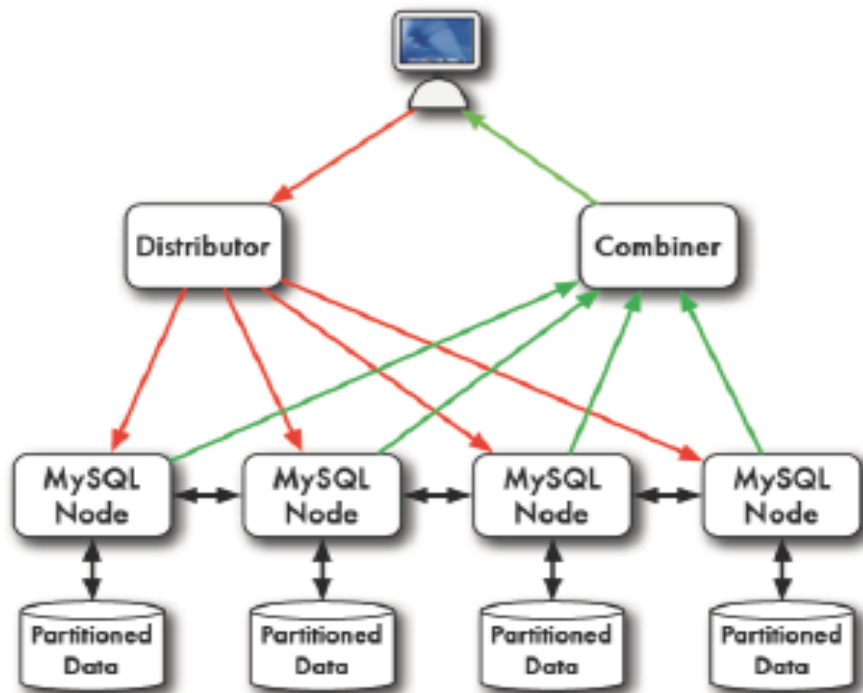


Solution de base : qserv

SciDB

<http://dev.lsstcorp.org/trac/wiki/dbQservOverview>

<http://www.scidb.org/>



Base de données Open Source
spécifiquement conçue pour le
stockage de masses de données
scientifiques

Modèle de données basé sur des
tableaux multidimensionnels

Vise un passage à l'échelle jusqu'à
plusieurs centaines de Po sur 1000
nœuds



Virtualisation



La virtualisation permet de dé-corréler l'infrastructure matérielle (réelle) de l'infrastructure présentée à l'utilisateur (virtuelle)

Sur un serveur multi-cœur, il est possible d'instancier plusieurs machines virtuelles ayant des caractéristiques données :

- Système d'exploitation
- Environnement logiciel
- Éventuellement environnement d'analyse

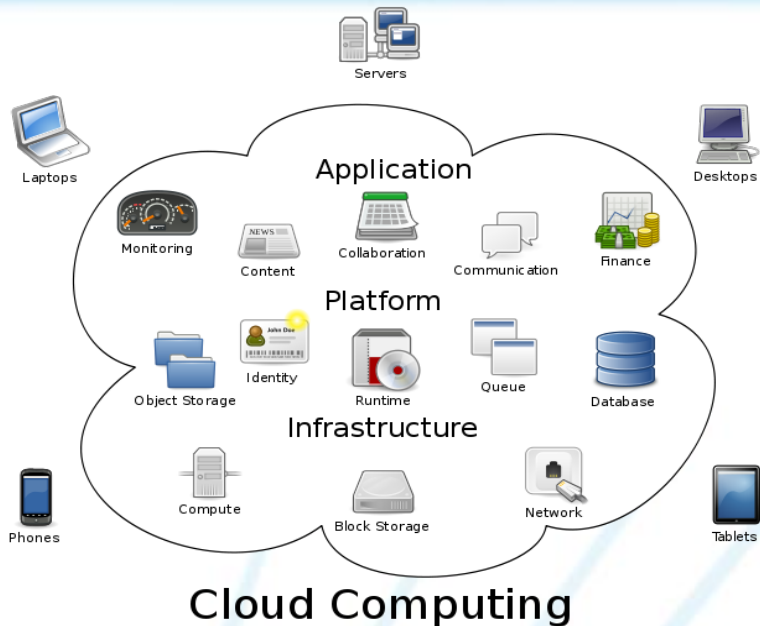
La virtualisation va prendre une place de plus en plus importante dans les centres de traitement de données

- S'affranchir au maximum des spécificités des sites pour créer des infrastructures adaptées aux expériences

Cloud : IaaS

➔ Introduction de la virtualisation dans les architectures de grille

Vers une infrastructure de Cloud



IaaS : Avoir accès à la demande à un grand nombre de ressources virtualisées

PaaS : Un ensemble d'outils pour développer, déployer et gérer des applications

Applications

SaaS : L'application elle-même est disponible à travers un navigateur

Hiérarchie naturelle : IaaS – PaaS – SaaS

Les grilles vont évoluer pour intégrer de plus en plus de services virtualisés

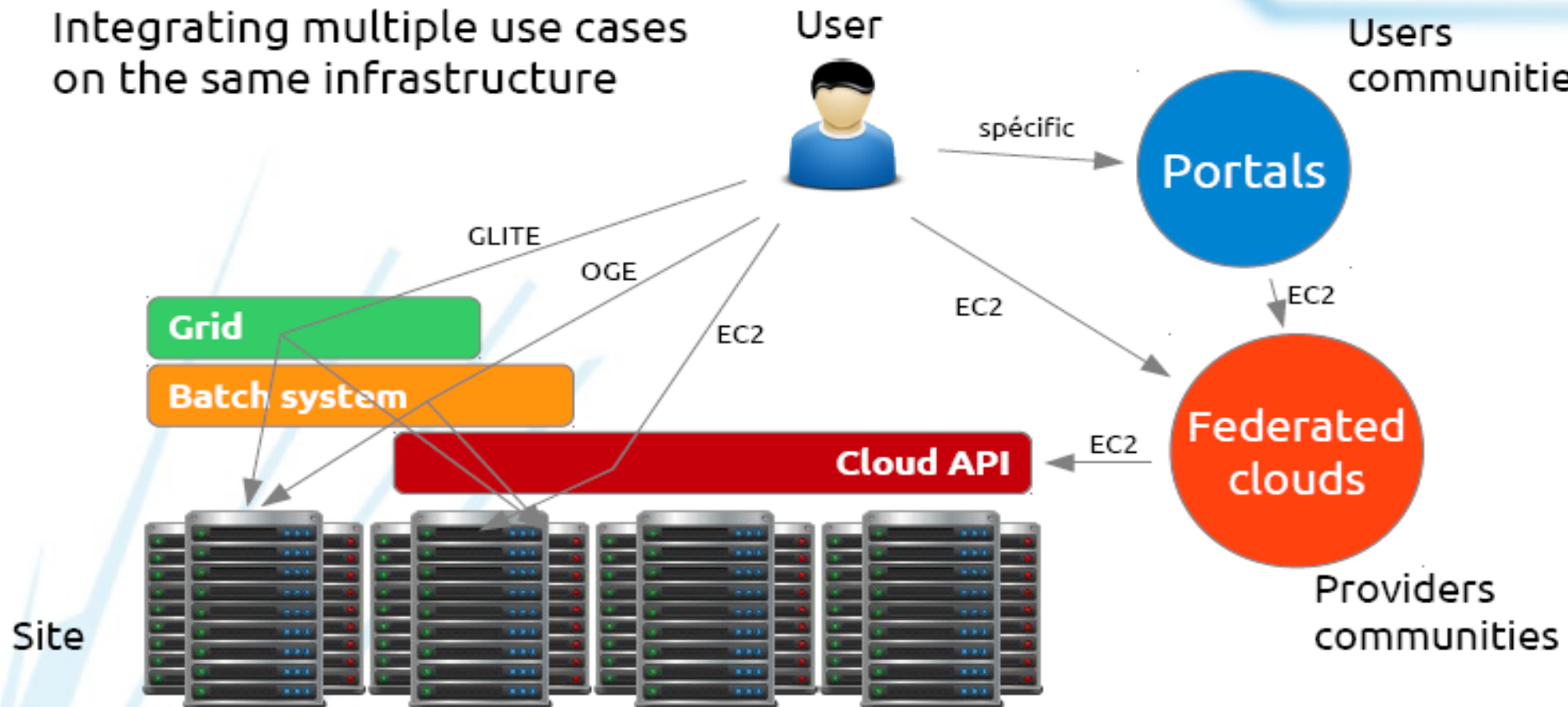
→ Initiative France-Grilles pour la mise en œuvre d'un cloud académique au CC-IN2P3

Prototype de cloud académique au CC-IN2P3 via France Grilles



Integrating multiple use cases on the same infrastructure

Users communities



Investissement initial : 320 k€

Ouverture du service IaaS en bêta-test d'ici l'été

Le CERN est engagé dans un effort très important pour structurer ses ressources informatiques sous forme de clouds

- 🌐 July 2012
 - 🌐 10% resources in CC managed by IaaS system (1000 nodes) including installation, configuration, etc.
- 🌐 January 2013
 - 🌐 50 % resources managed (over 3 sites)
- 🌐 July 2013
 - 🌐 95% resources managed
- 🌐 Initial focus on service/server consolidation use cases

Borrowed from Ian Bird (CERN)

<http://indico.cern.ch/materialDisplay.py?sessionId=4&materialId=0&confId=155066>

10 avril 2012



Nouvelle salle machine du CC-IN2P3

Juin 2010



Avril 2011



7.5 M€ financée dans le cadre du
CPER 2007-2013

Dominique Boutigny

Nouvelle salle machine du CC-IN2P3



Capacité:

Design innovant et modulaire

2011
40 racks
600 kW

2012
80 racks
1 MW

2015
125 racks
1.5 MW

2019
240 racks
3.2 MW

En plus des 1 MW existant pour l'ancienne salle

Cette nouvelle salle est un atout majeur pour les projets scientifiques de la physique des hautes énergies

- Traitement de la moitié des données du LSST
- Jeu complet de données en Europe

Amenée également à jouer un rôle important dans le cadre de l'ouverture multidisciplinaire du CC-IN2P3