

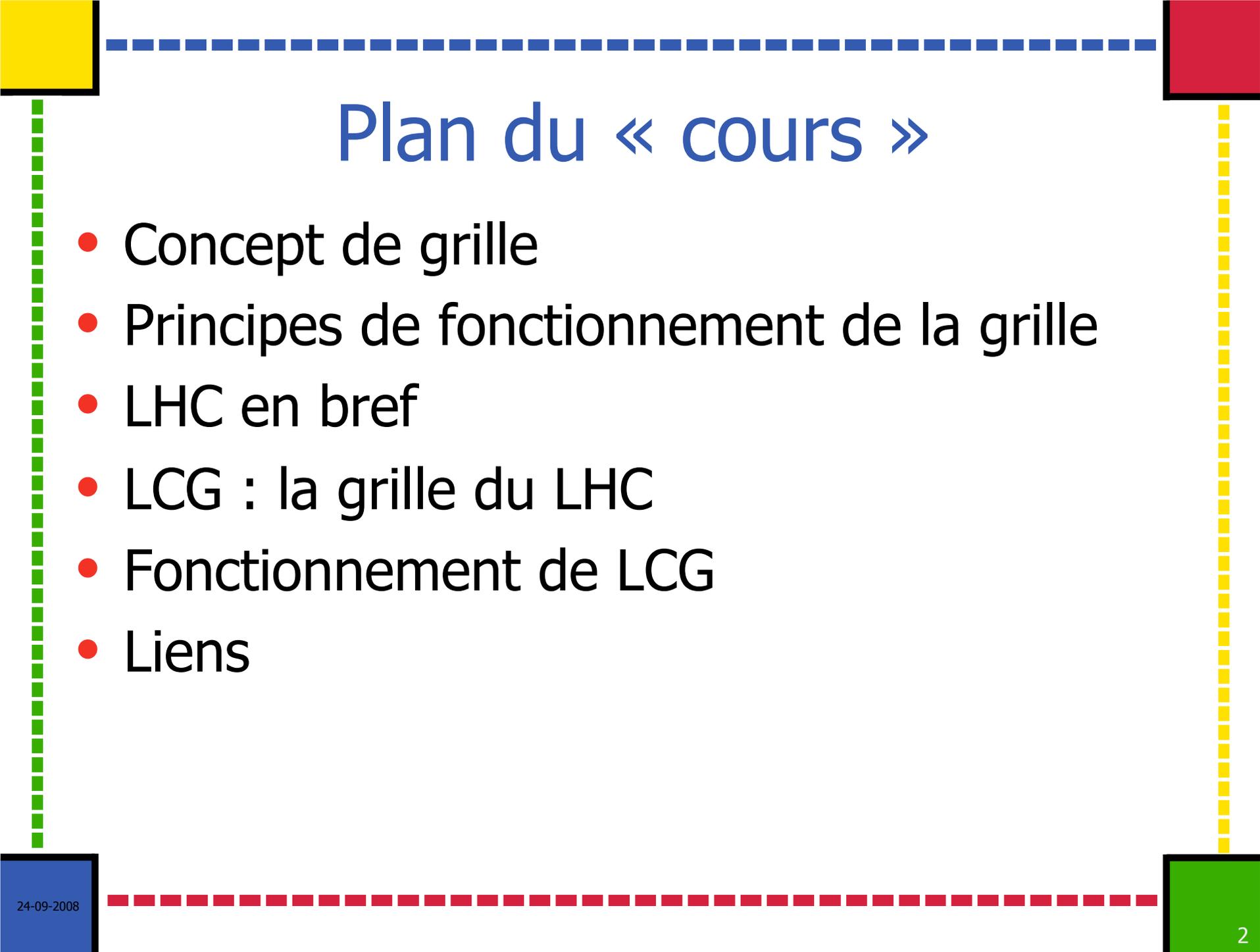
La grille de calcul du LHC

Laurent APHECETCHE

24 Septembre 2008

JC2008 - Seignosse





Plan du « cours »

- Concept de grille
- Principes de fonctionnement de la grille
- LHC en bref
- LCG : la grille du LHC
- Fonctionnement de LCG
- Liens

1. CONCEPT DE GRILLE : DU REVE A LA REALITE

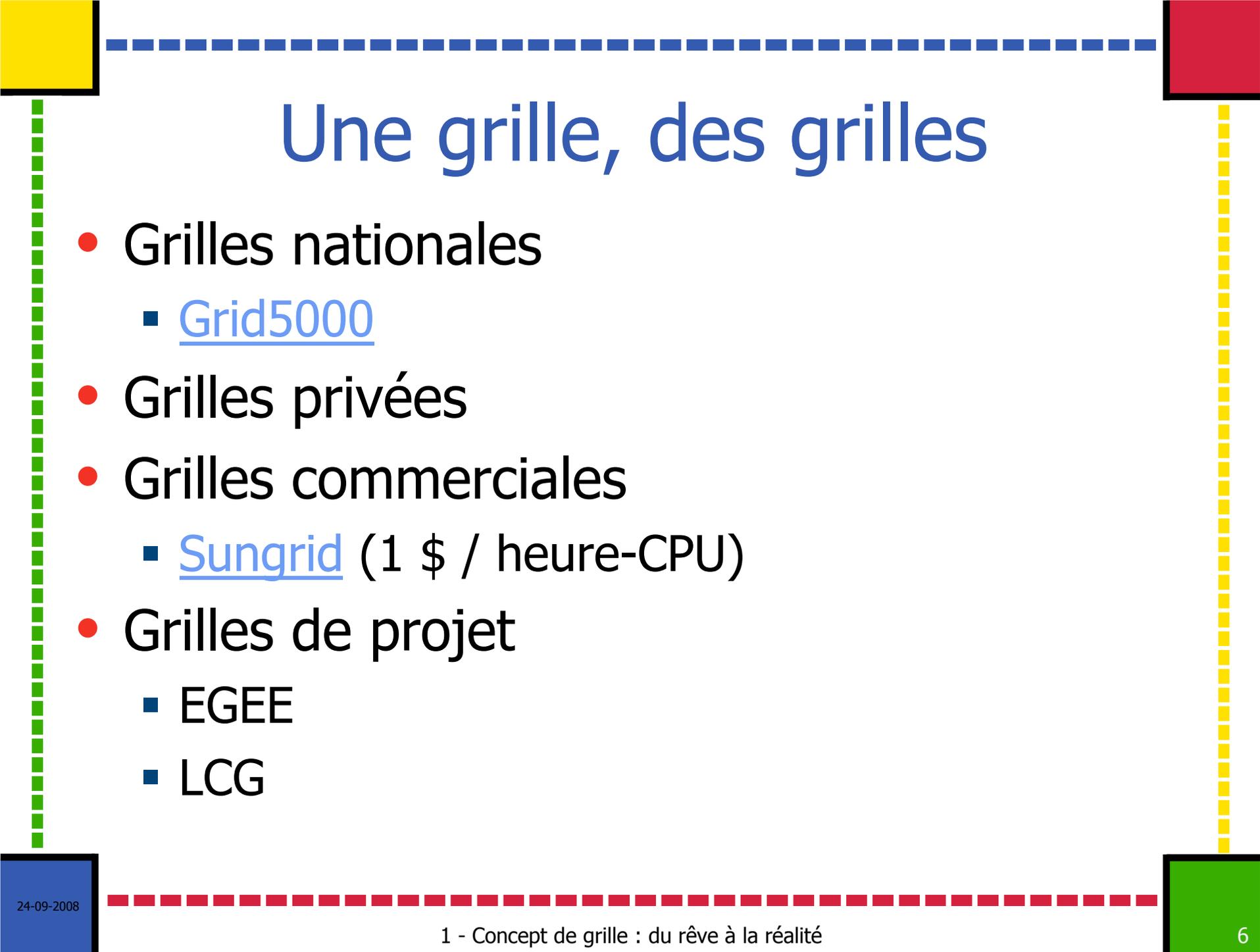


L'ordinateur mondial



L'esprit de la loi de Moore

- L'originale :
 - (1965) The complexity for minimum component costs has increased at a rate of roughly a **factor of two** per year
 - (1975) One year -> two years
- On entend beaucoup de variantes (plus ou moins fausses...)
- Ce que cela veut dire :
 - croissance exponentielle de la capacité du matériel informatique (CPU, disque, réseau)

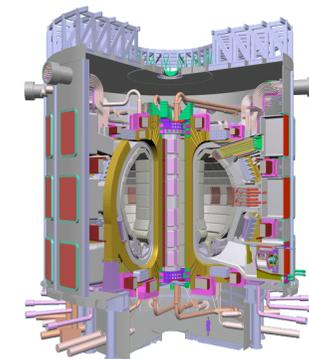
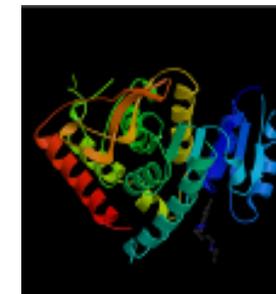
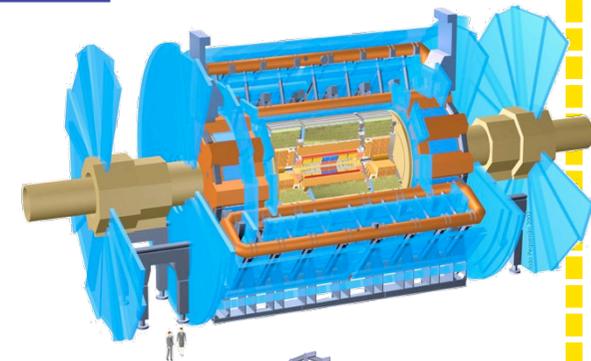
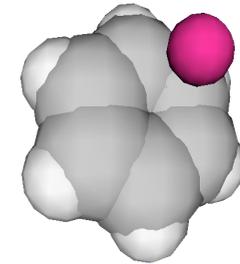


Une grille, des grilles

- Grilles nationales
 - [Grid5000](#)
- Grilles privées
- Grilles commerciales
 - [Sungrid](#) (1 \$ / heure-CPU)
- Grilles de projet
 - EGEE
 - LCG

Quelles applications ?

- Hautes énergies
- Astrophysique
- Sciences de la vie
- Sciences de la terre
- Fusion
- ...

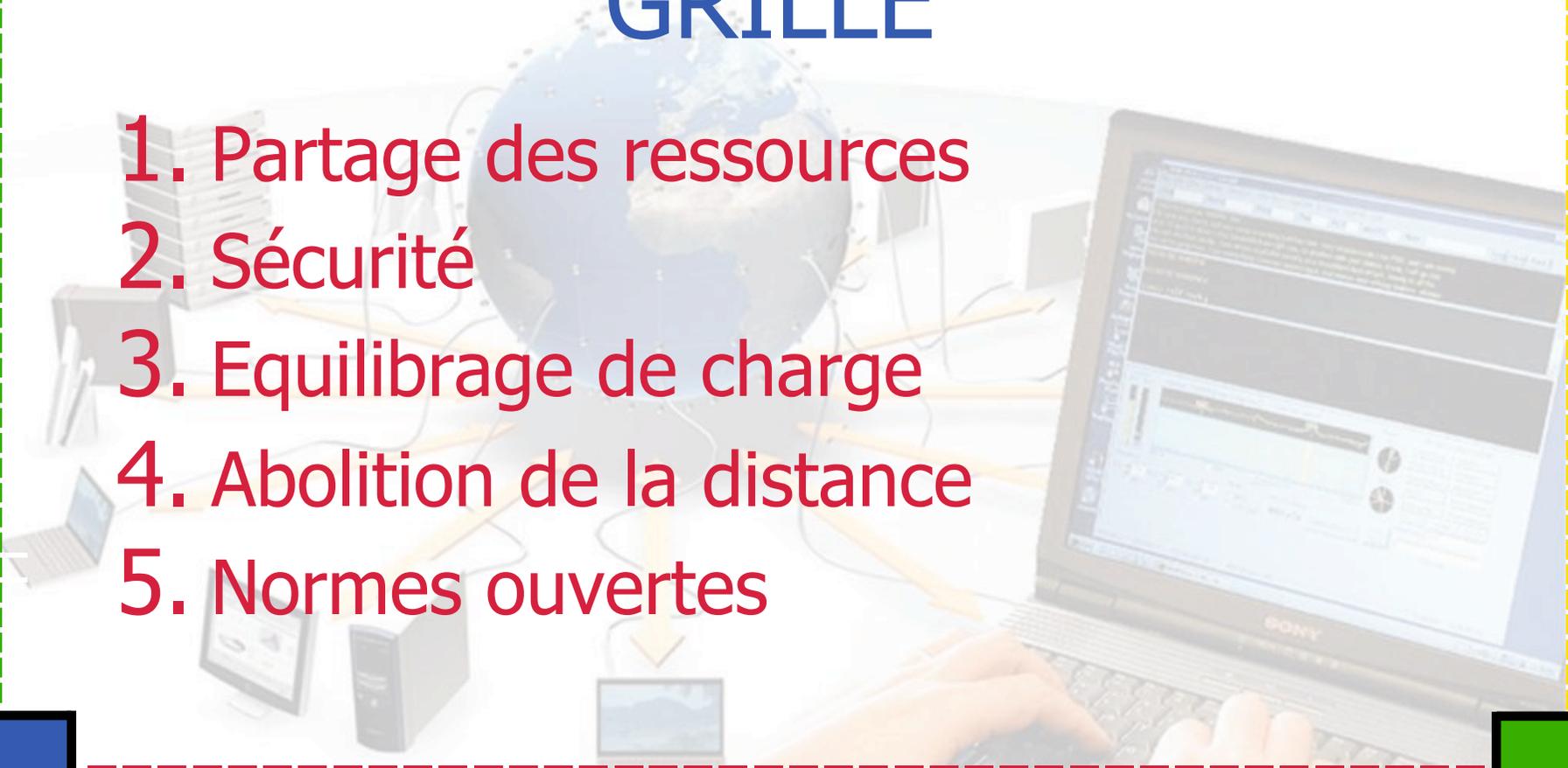


Organisations

- Organisation Virtuelle (VO)
 - Communauté d'institutions et/ou d'individus partageant les mêmes intérêts et buts scientifiques.
 - Pour HEP, 1 VO = 1 expérience
- Réelles
 - EGEE-III
 - OSG

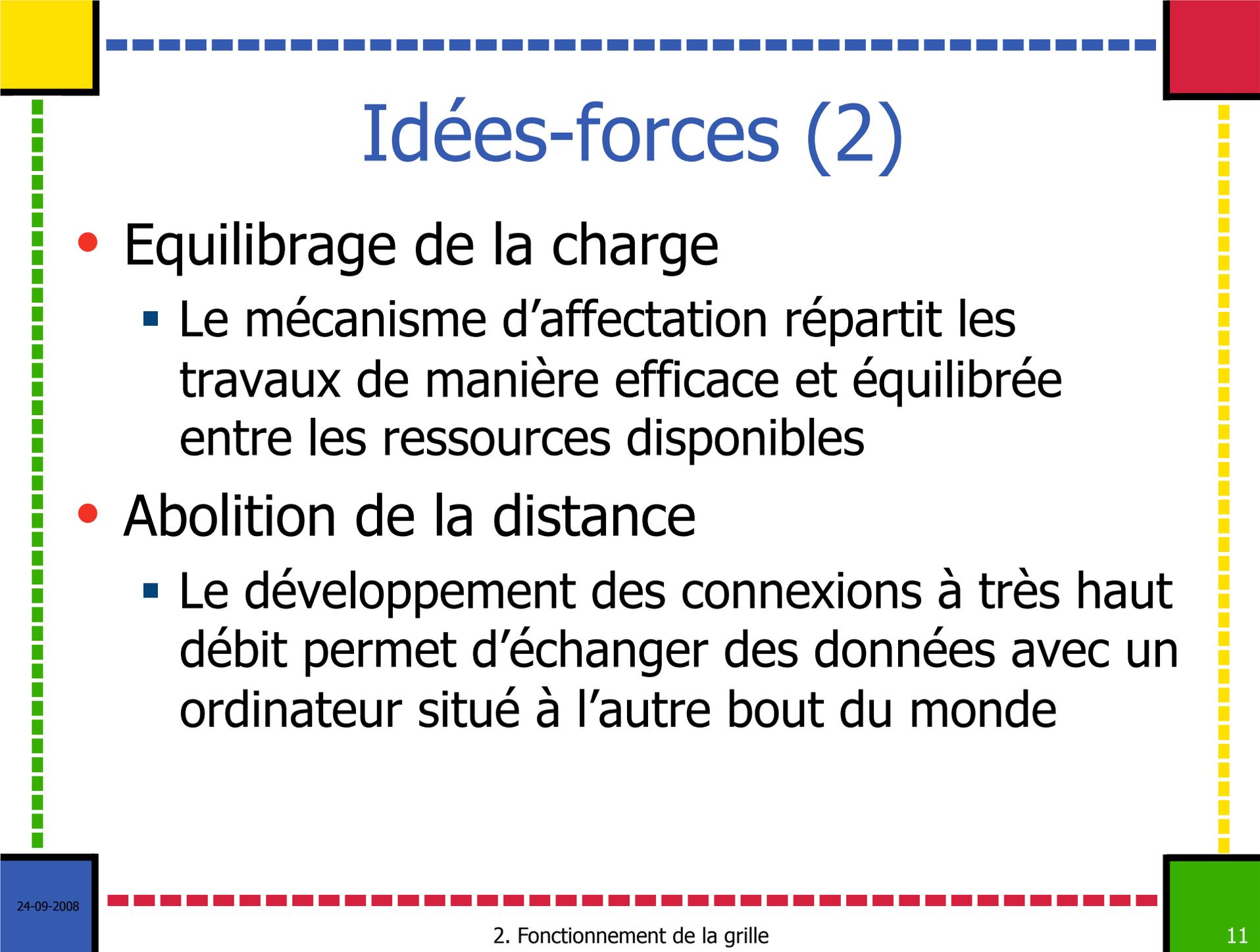


2. PRINCIPES DE FONCTIONNEMENT DE LA GRILLE

1. Partage des ressources
 2. Sécurité
 3. Equilibrage de charge
 4. Abolition de la distance
 5. Normes ouvertes
- 

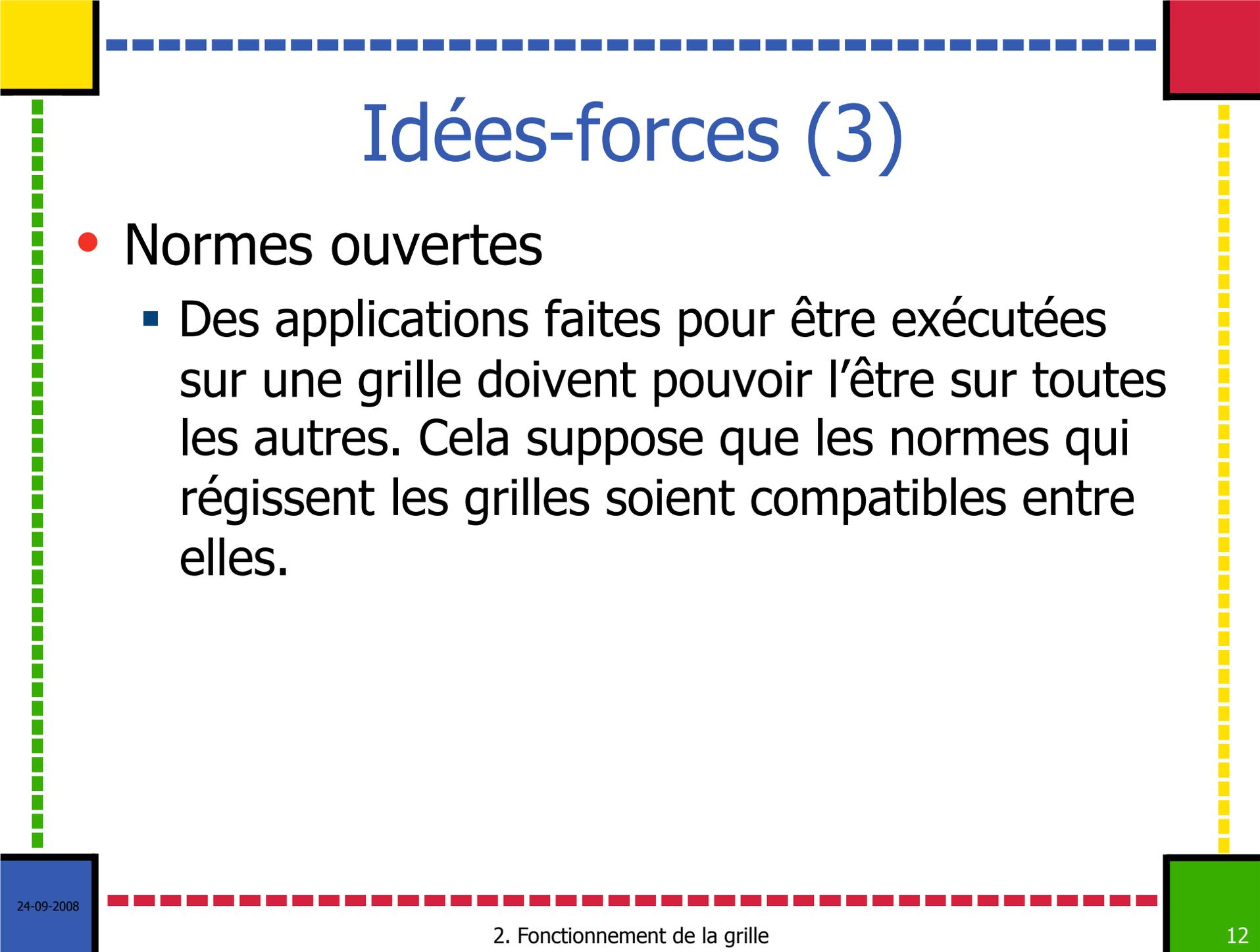
Idées-forces (1)

- Partage des ressources
 - Plusieurs ordinateurs, situés partout dans le monde, mettent leur puissance de calcul et de stockage en commun.
 - C'est le principe de base de la grille
- Sécurité
 - Les fournisseurs et les utilisateurs de la puissance de calcul, qui ignorent leur identité réciproque, doivent pouvoir échanger des données en toute sécurité



Idées-forces (2)

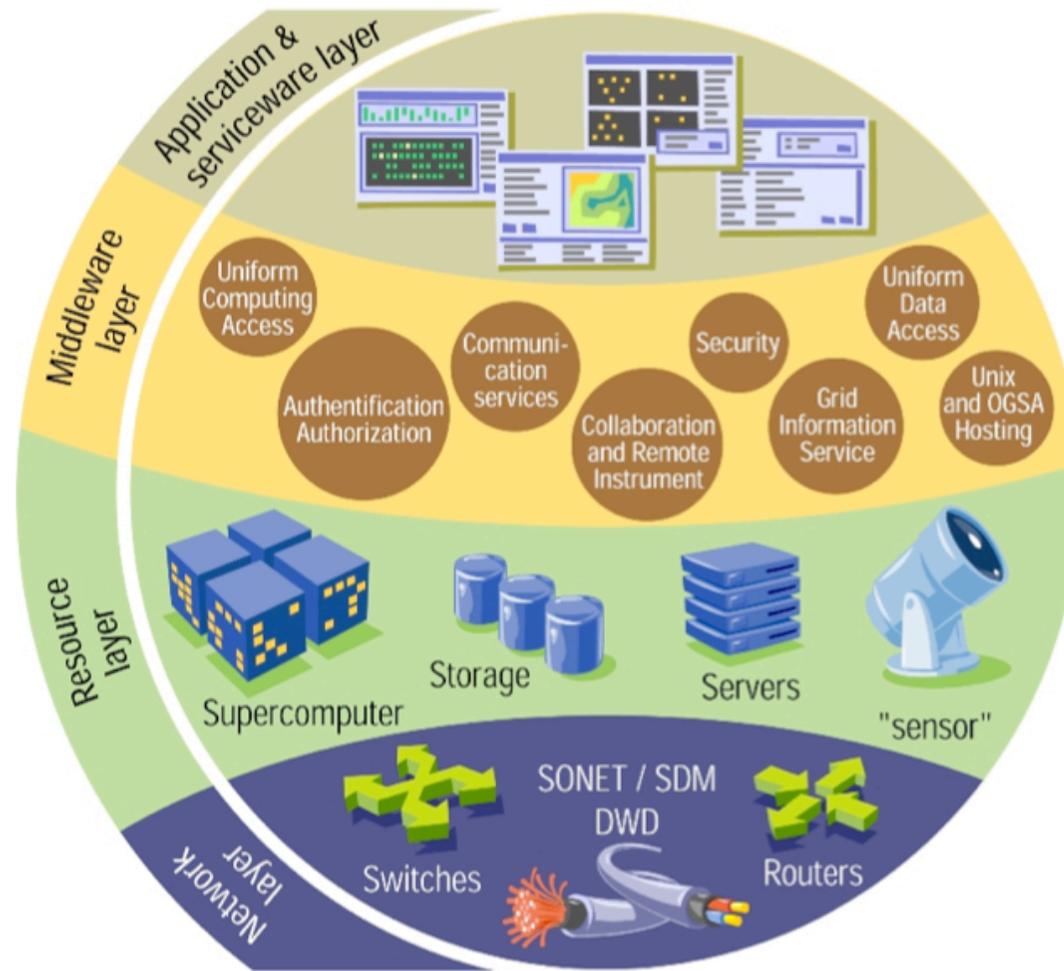
- Equilibrage de la charge
 - Le mécanisme d'affectation répartit les travaux de manière efficace et équilibrée entre les ressources disponibles
- Abolition de la distance
 - Le développement des connexions à très haut débit permet d'échanger des données avec un ordinateur situé à l'autre bout du monde



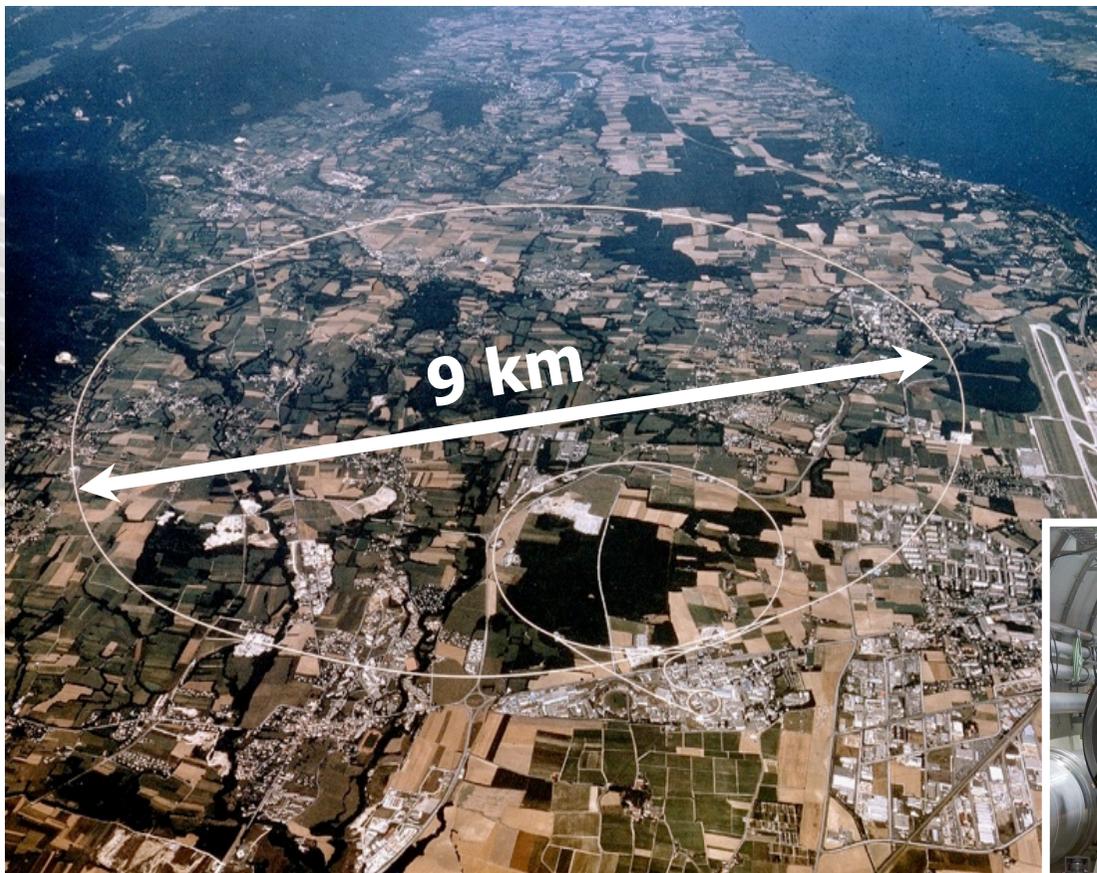
Idées-forces (3)

- Normes ouvertes
 - Des applications faites pour être exécutées sur une grille doivent pouvoir l'être sur toutes les autres. Cela suppose que les normes qui régissent les grilles soient compatibles entre elles.

Les briques de base



3. LHC EN BREF



M,G,T,Peta... octets



LHC data

- 40 million collisions per second
- After filtering, 100 collisions of interest per second
- A Megabyte of data digitised for each collision
= recording rate of 0.1 Gigabytes/sec
- 10^{10} collisions recorded each year
= 10 Petabytes/year of data

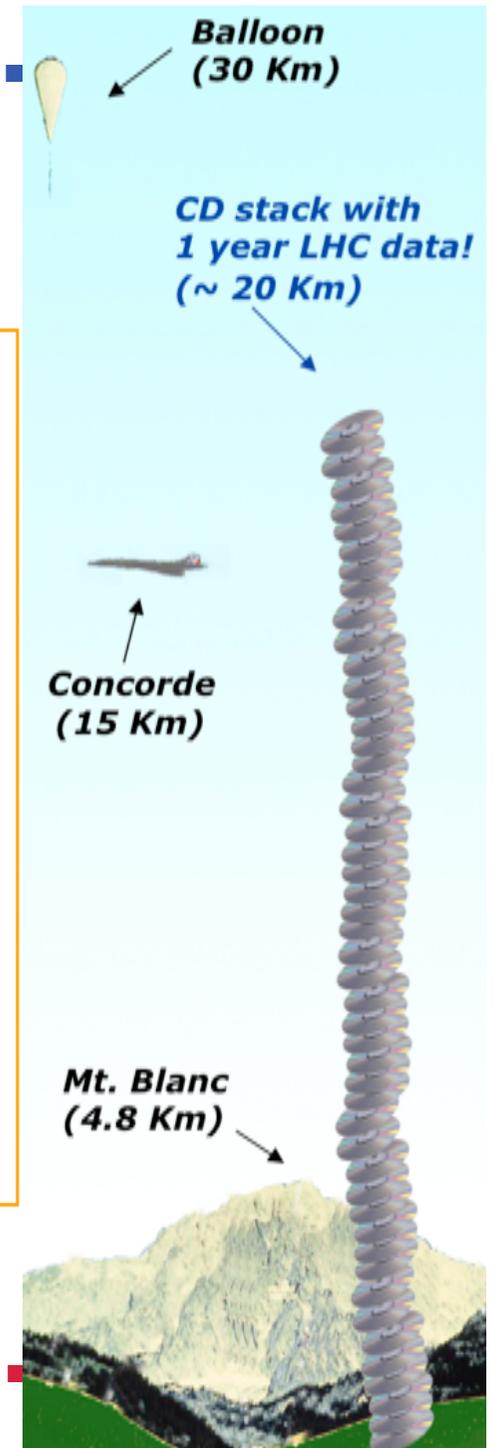
1 Megabyte (1MB)
A digital photo

1 Gigabyte (1GB)
= 1000MB
A DVD movie

1 Terabyte (1TB)
= 1000GB
World annual book production

1 Petabyte (1PB)
= 1000TB
Annual production of one LHC experiment

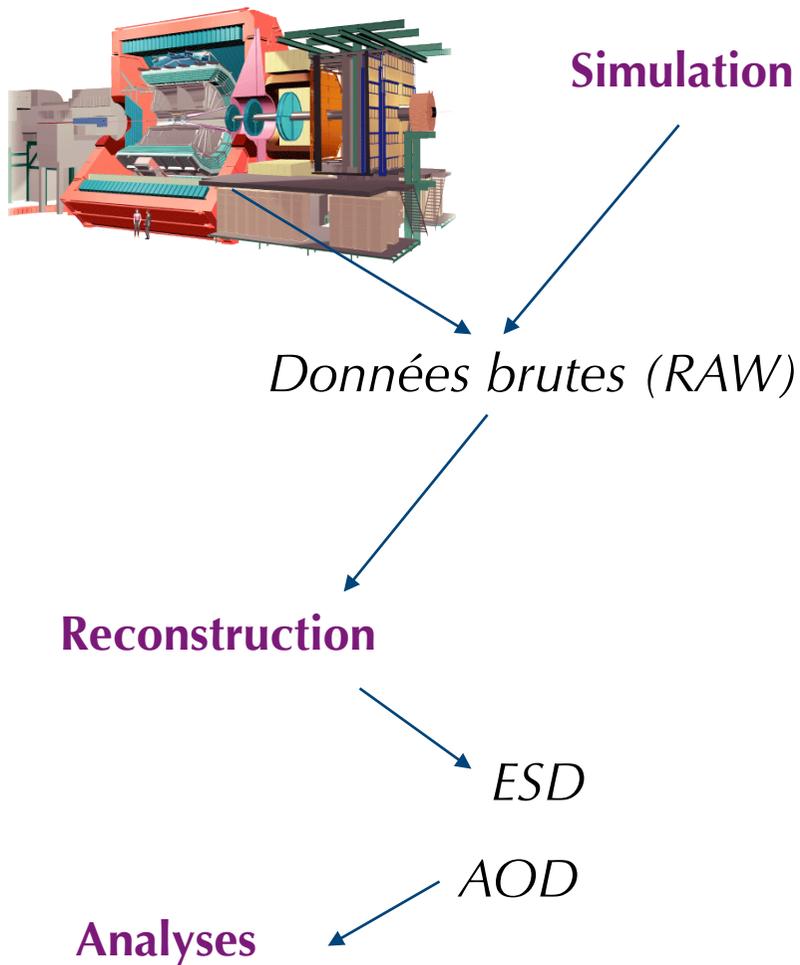
1 Exabyte (1EB)
= 1000 PB
World annual information production



Les unités : kSI2K, TB, GB/s

- Unité puissance processeur
 - 1 PC (ou Mac) actuel \sim kSI2K
 - Spec.org
- Unité stockage
 - Tera-octets = 1000 Giga
- Unité débit réseau
 - (dizaine de) Gigabit/seconde

Du capteur à la publication

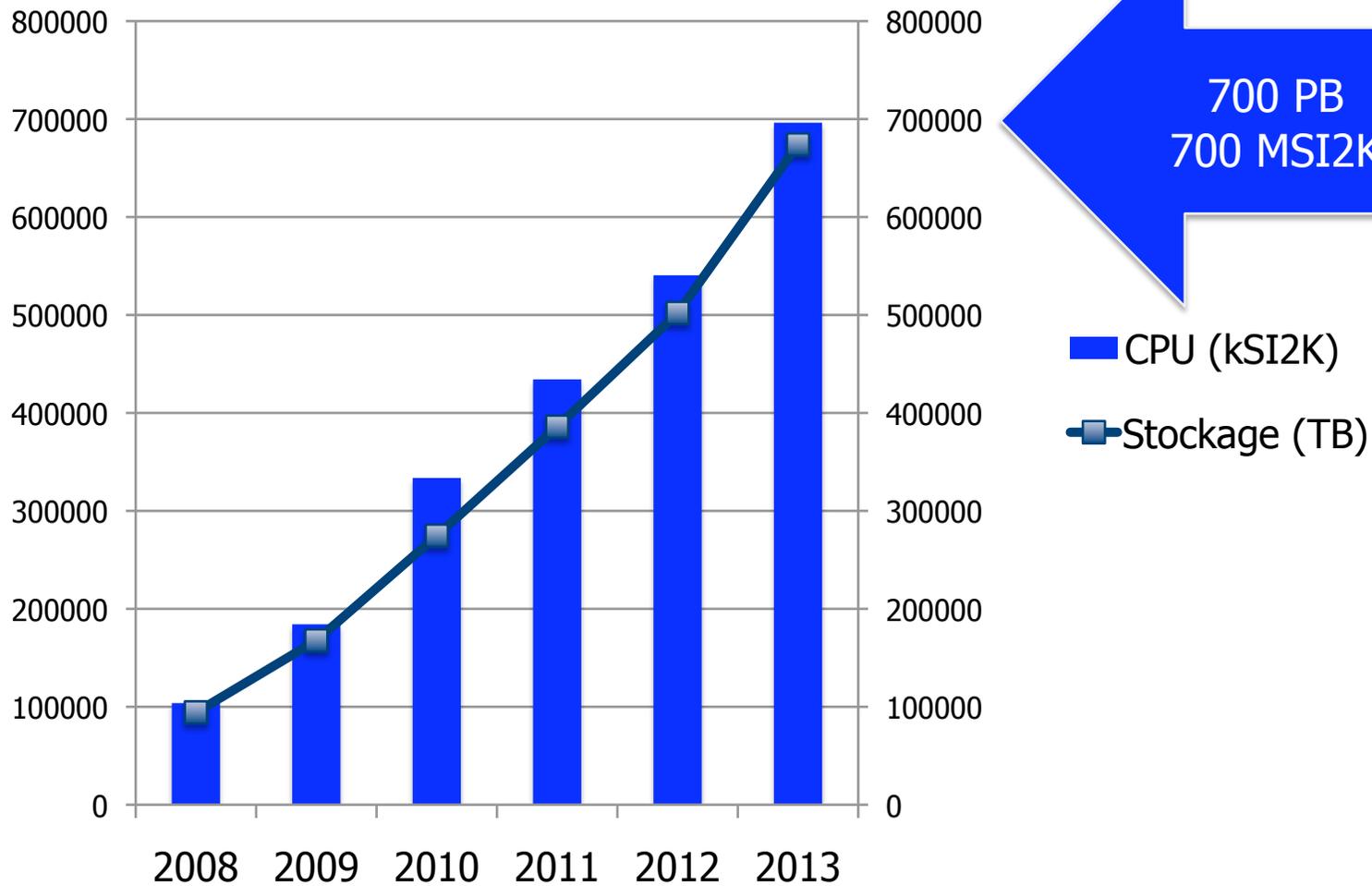


- Générer des particules, en utilisant des codes théoriques
- « Transporter » ces particules dans le détecteur, et déterminer leurs interactions avec la matière
- Simuler la réponse du détecteur à ces interactions

Des signaux électriques enregistrés par le détecteur aux informations physiques pertinentes : énergies, temps de vol, impulsions, identification des particules...

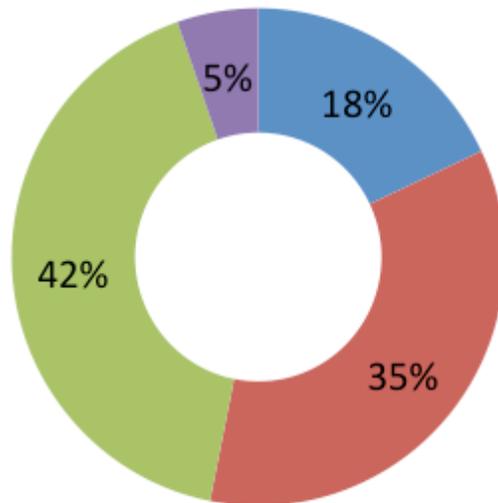
Le cœur du métier...

L'évolution des besoins



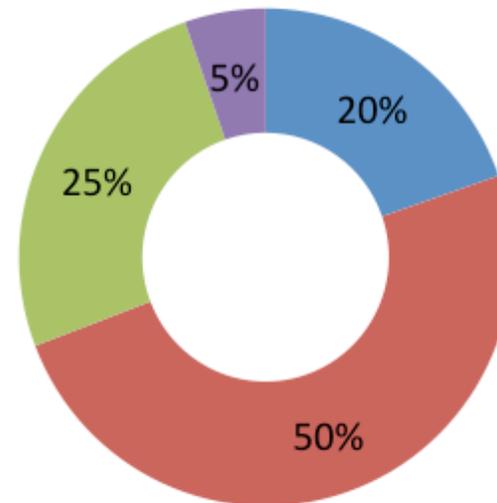
Répartition par expérience

CPU (2011)



■ ALICE ■ ATLAS ■ CMS ■ LHCb

Stockage (2011)

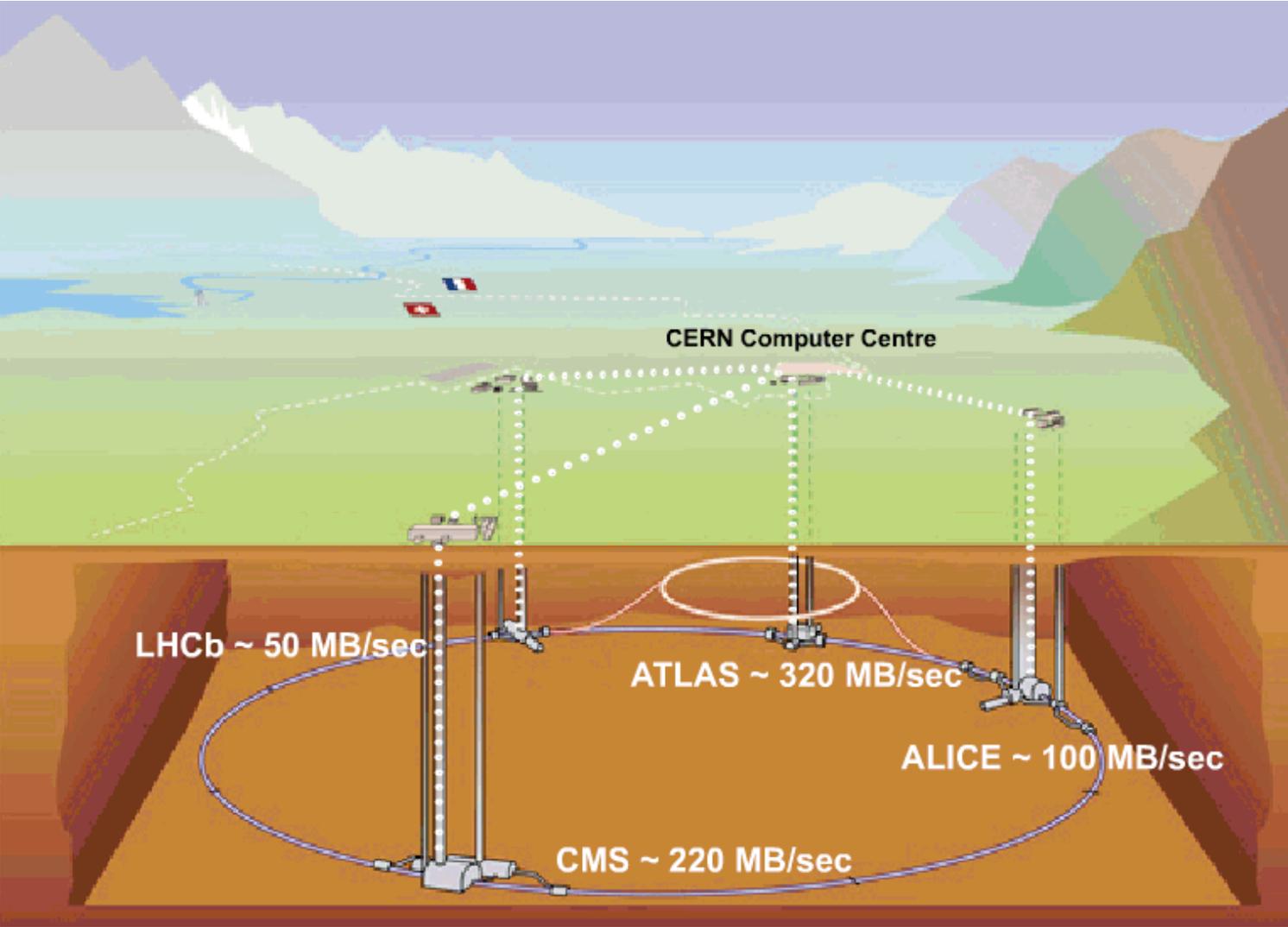


■ ALICE ■ ATLAS ■ CMS ■ LHCb

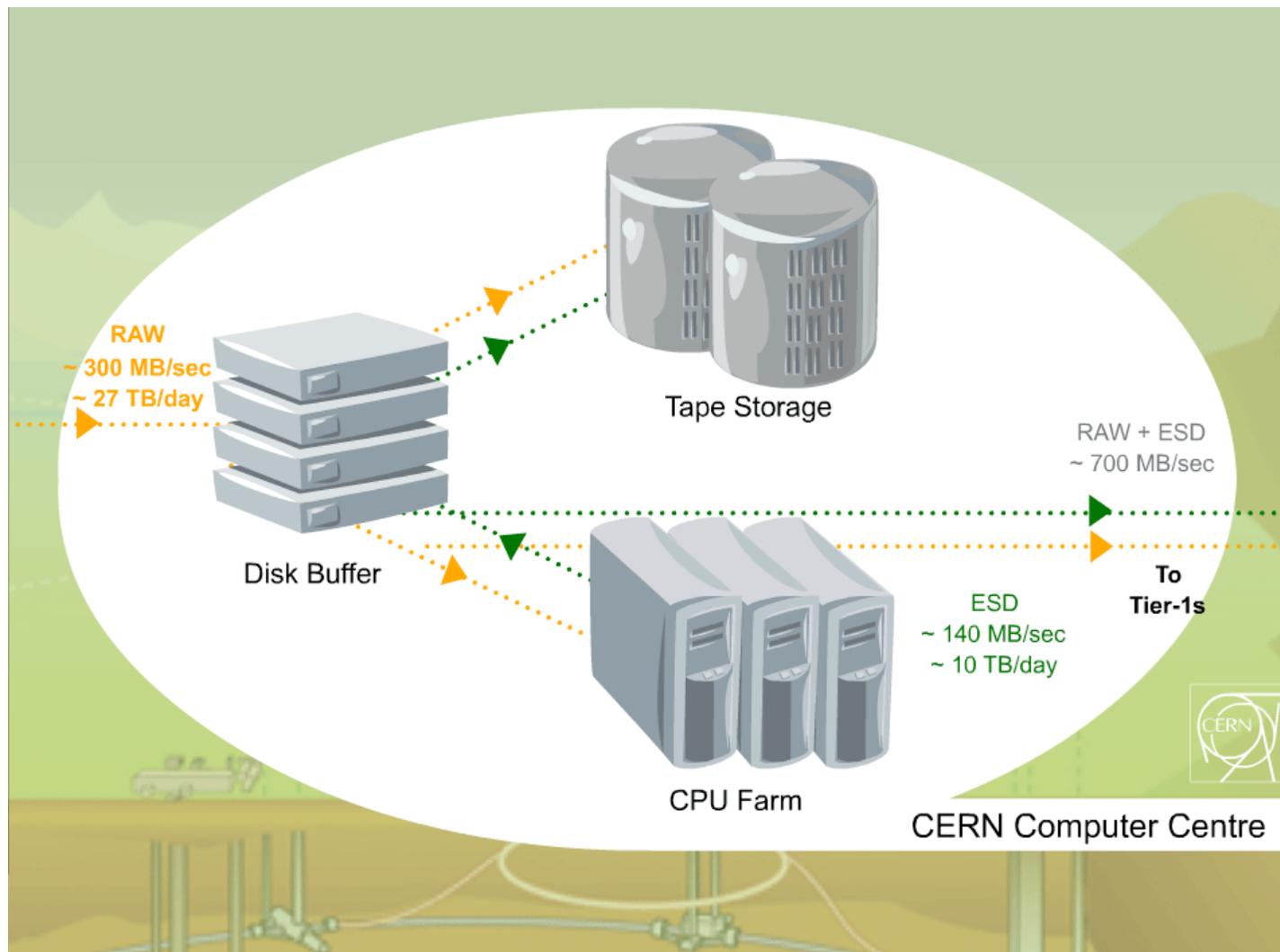
4. LCG : LA GRILLE DU LHC



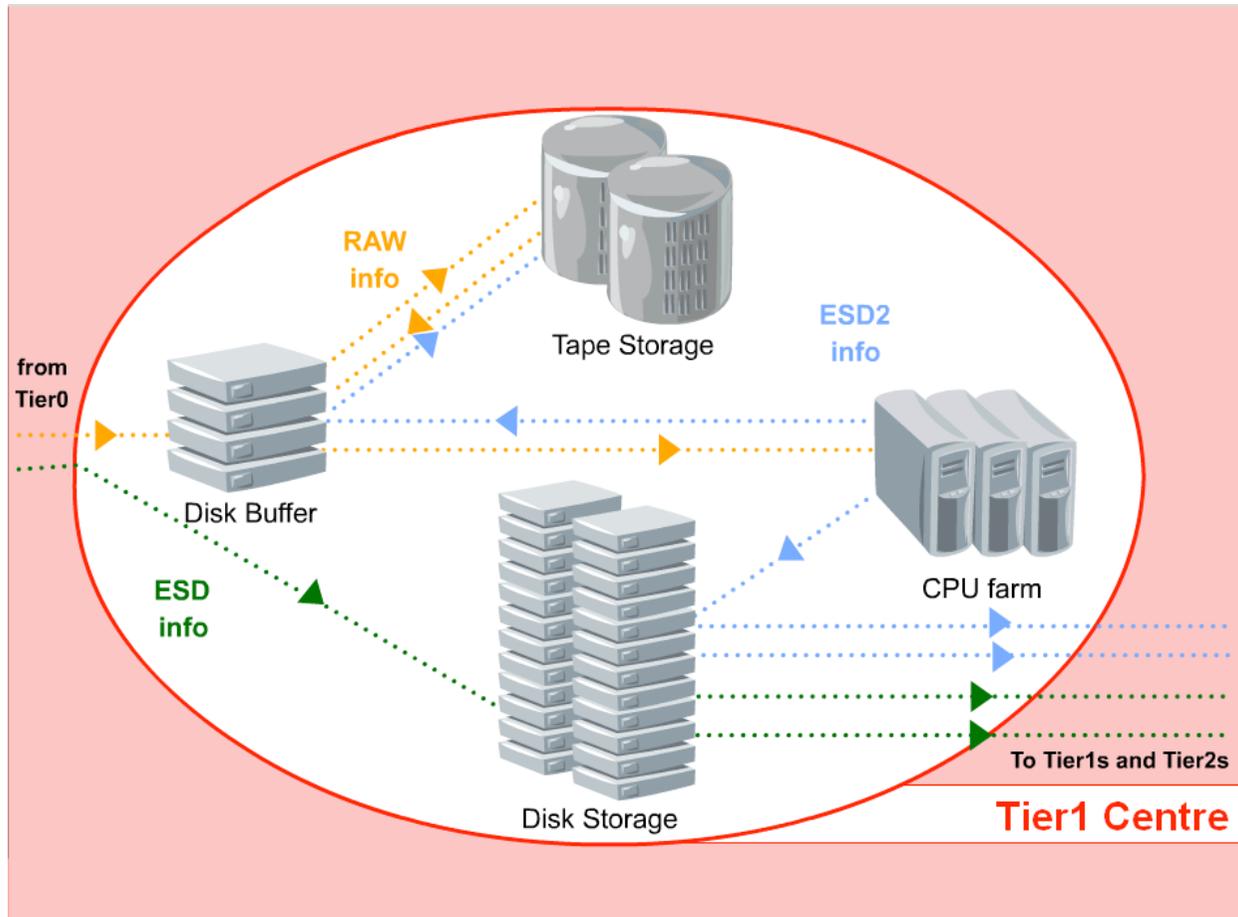
Puits vers Tier 0



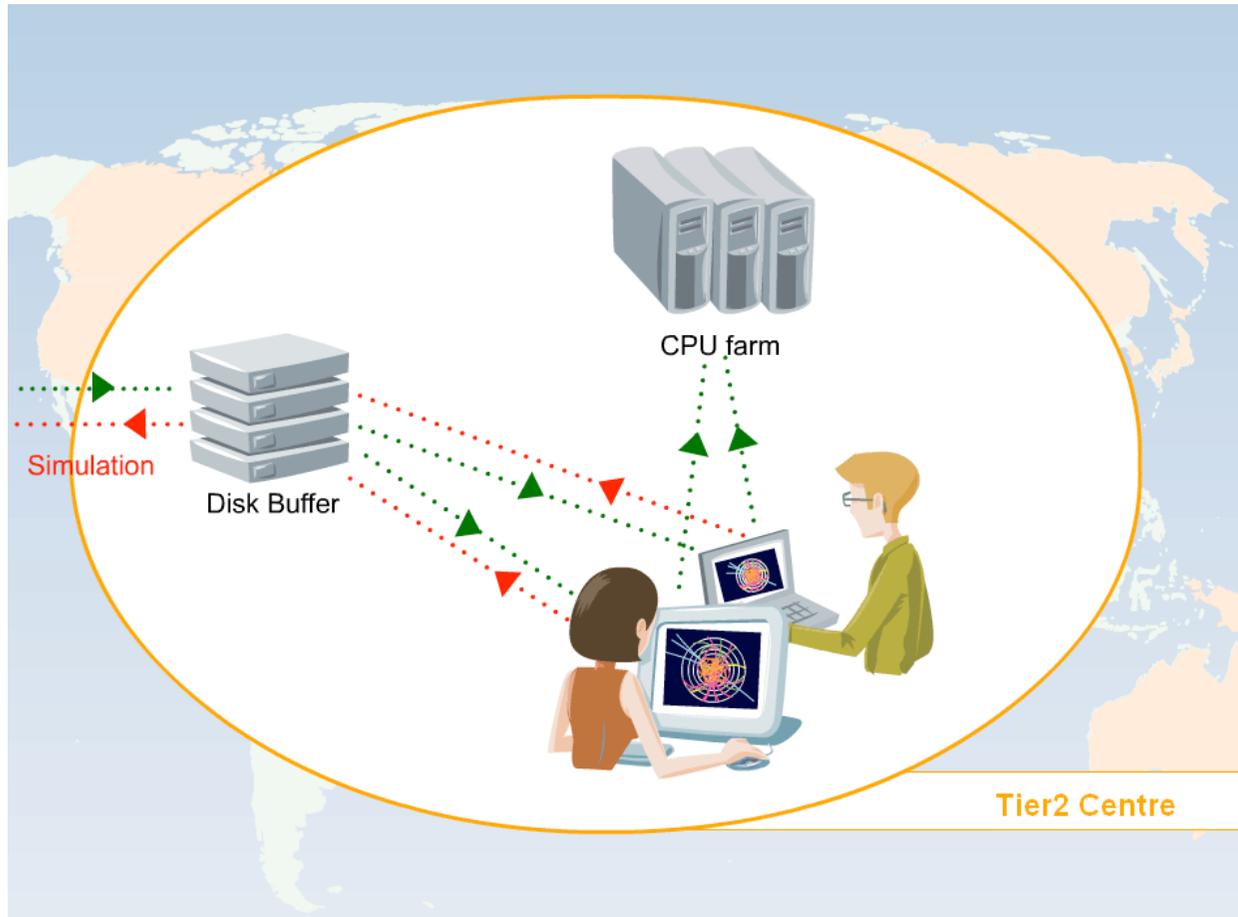
1 CERN Computing Center = T0



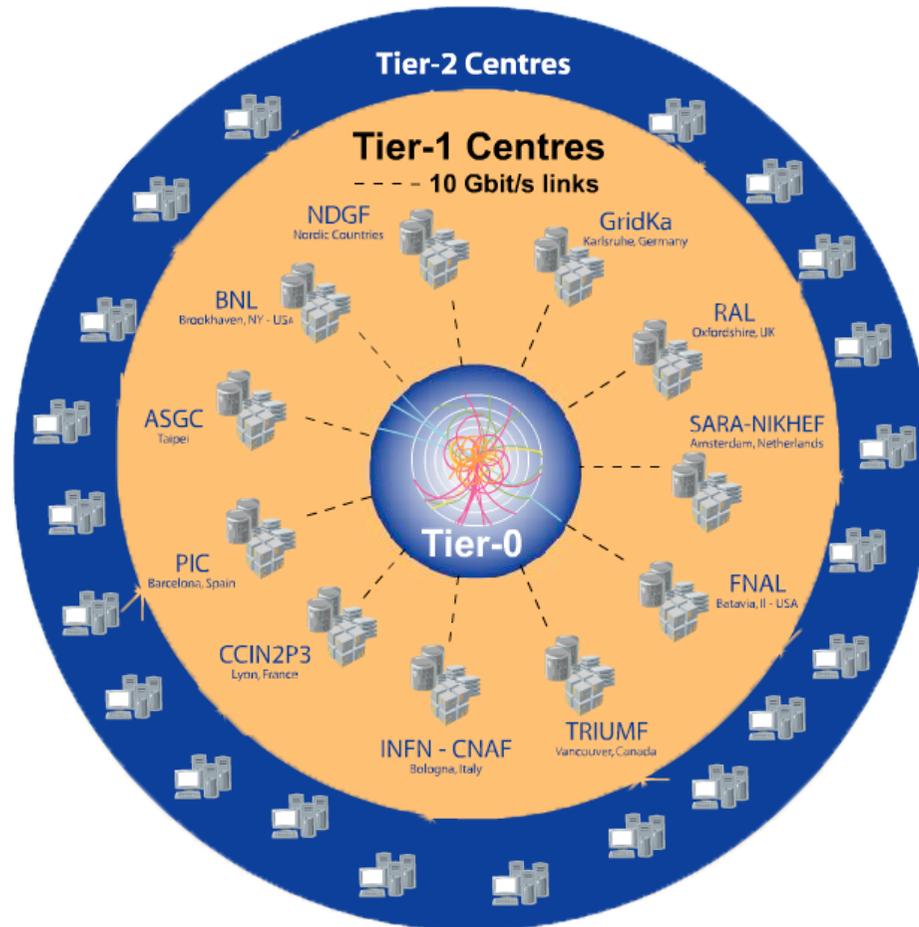
Une dizaine de T1



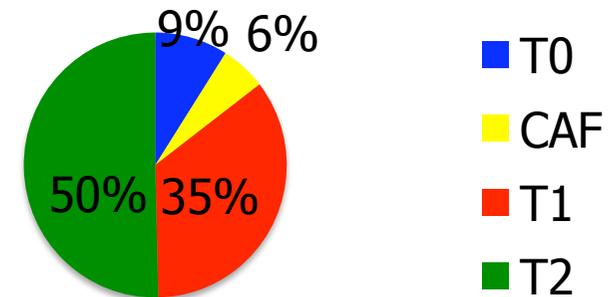
Une centaine de T2



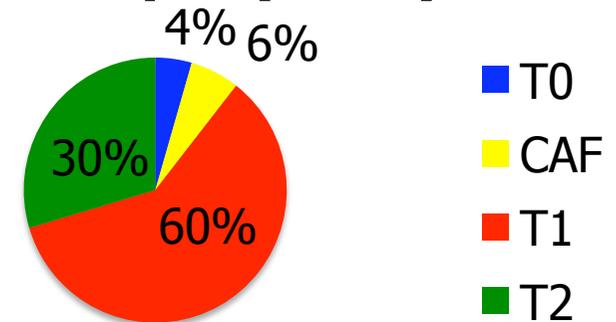
Hierarchie



CPU (2011)



Disque (2011)



5. FONCTIONNEMENT DE LCG



Calcul et stockage

- Element de calcul (CE)
 - Ensemble de serveurs de calcul (Worker Nodes), sur lesquels les jobs sont exécutés
- Element de stockage (SE)
 - Machine de la grille permettant de stocker des fichiers.
 - Service accessible depuis tous les serveurs de calcul (Worker Nodes) et toutes les machines interactives (User Interfaces). Cependant, les fichiers stockés sur cette machine ne sont accessibles que par l'utilisation de commandes spécifiques.
 - Le stockage réel « derrière » peut être disque ou bande (et les protocoles = DPM, xrootd, dCache, CASTOR2, etc...)

Informations

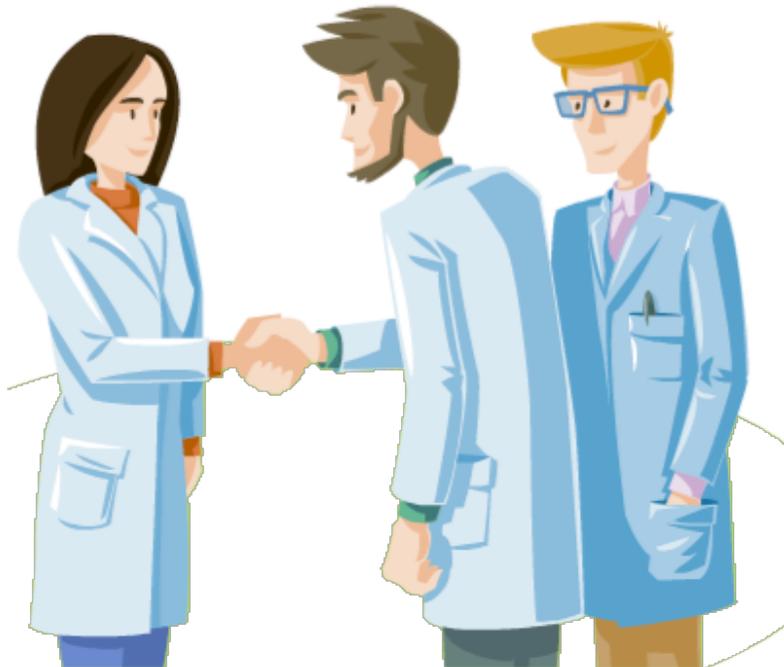
- Service journal de bord (LB)
 - Garde l'historique de la gestion des jobs
- Catalogue des répliques (RC)
 - Les « pages jaunes » de la grille
 - Permet de localiser les données (qui peuvent être présentes à plusieurs endroits)
- Répertoire d'information (IS)
 - L'Agence France Presse de la grille...
 - Permet d'obtenir en temps réel l'état des services de grille de tous les sites

Choix

- Interface utilisateur (UI)
 - Généralement l'ordinateur de l'utilisateur
 - Tout ordinateur possédant le logiciel nécessaire pour accéder à la grille (obtention autorisation temporaire, soumission jobs, etc...)
- Courtier en ressources (RB)
 - L'intermédiaire ultime
 - Sélectionne le CE le mieux adapté à votre job

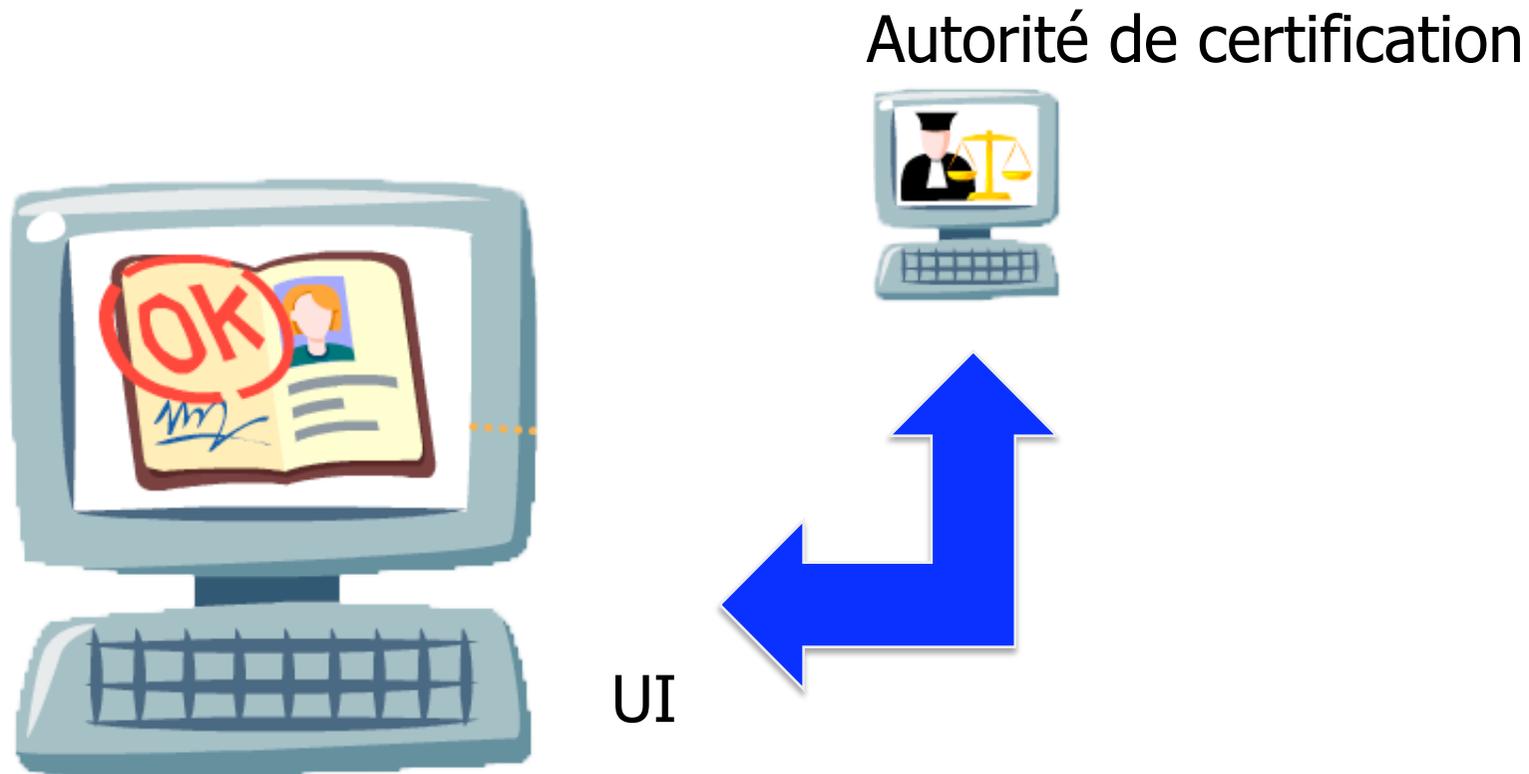
La vie d'un job sur la grille

1. Rejoindre une VO



...avant de pouvoir « jouer »
sur la grille ;-)

2. Obtenir un certificat



3. Installer un certificat

Une fois en possession de votre certificat, vous devez l'**installer dans l'interface utilisateur** - la machine que vous allez utiliser pour accéder à la Grille. Votre certificat a une **validité d'un an**, donc vous n'aurez pas à vous en soucier pendant un certain temps (mais gardez-le en lieu sûr)!



4. Autorisation d'accès temporaire

Néanmoins, chaque fois que vous créez une session sur la Grille, vous aurez besoin d'une **autorisation d'accès temporaire**. Ceci est nécessaire afin de limiter les risques d'exposition de votre certificat sur un réseau non-sécurisé. Les **autorisations d'accès temporaires** ont une durée de vie de 12 heures par défaut.. Durant cette période vous pourrez travailler sur la Grille.



5. Ecrire le job.jdl

Pour faire tourner un job sur la Grille LCG/EGEE, vous devez le décrire dans un **Langage de Description de Job (JDL)**.

Le **JDL** définit les caractéristiques de votre job telles que l'application à utiliser, les données à traiter, les ressources nécessaires etc.



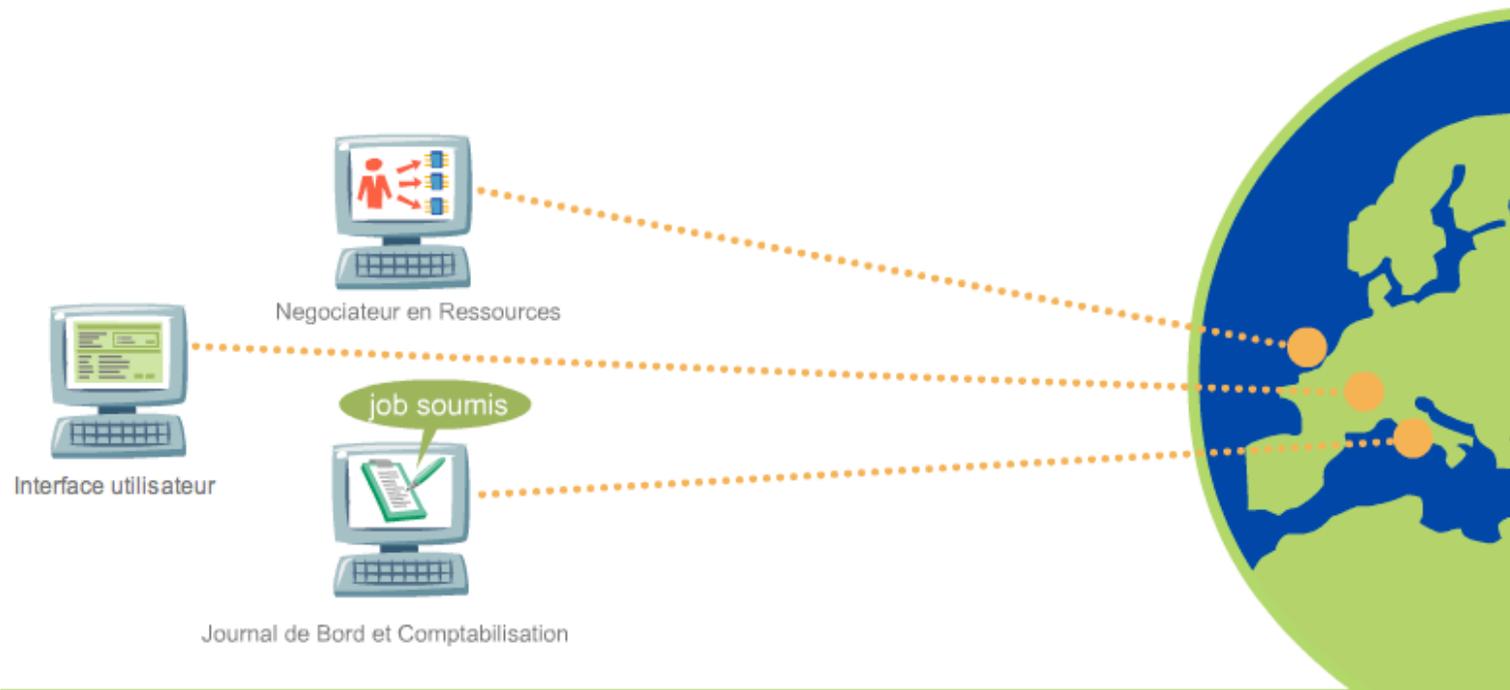
6. Soumission du job

Une fois que vous avez défini le JDL pour votre job, vous pouvez le soumettre au **Negociateur en Ressources**.



6. Soumission du job

Au même moment, le service de **Journal de Bord et Comptabilisation** enregistre le job soumis.



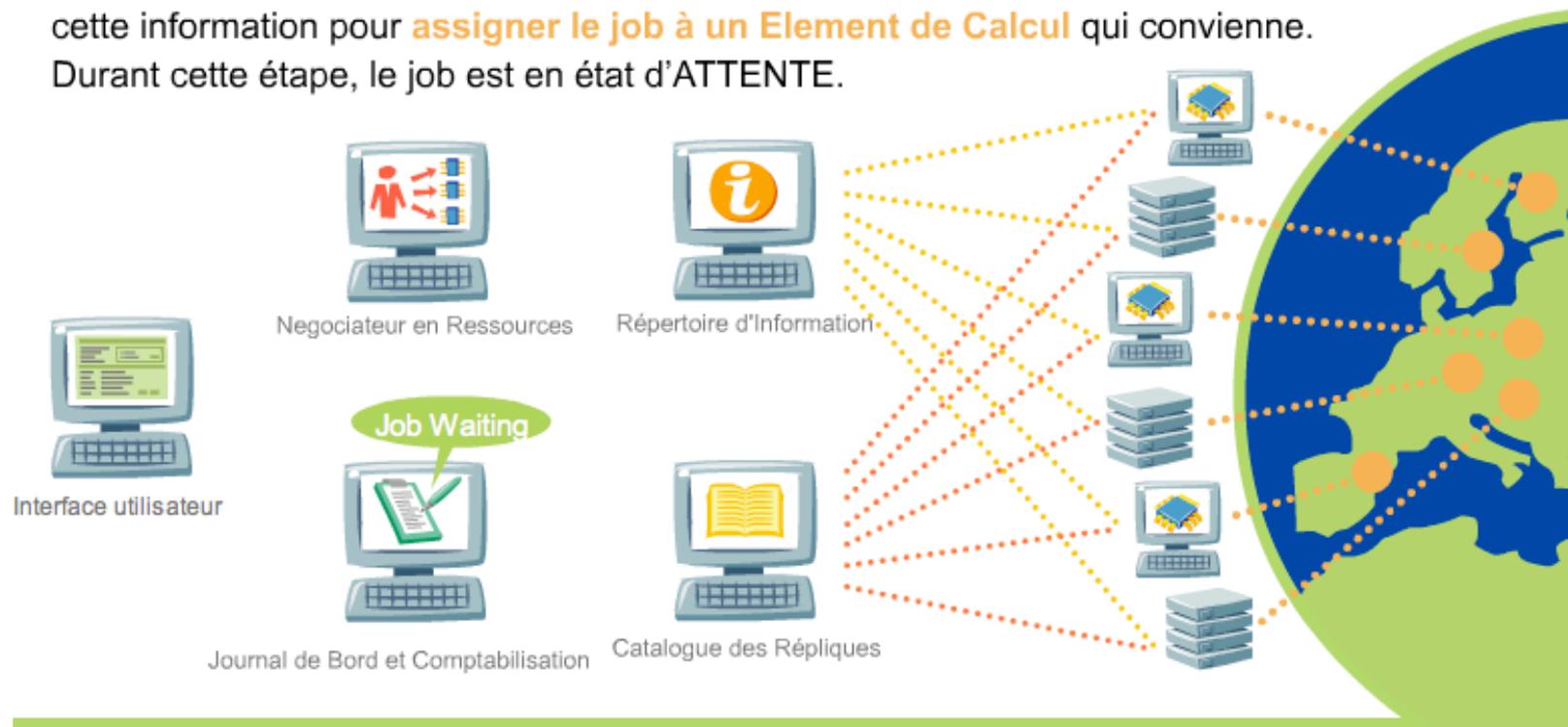
6. Soumission du job

Sur base des informations fournies dans le JDL, le **Negociateur en Ressources** consulte le **Répertoire d'Information** et le **Catalogue des Répliques** pour voir quelles ressources sont disponibles.



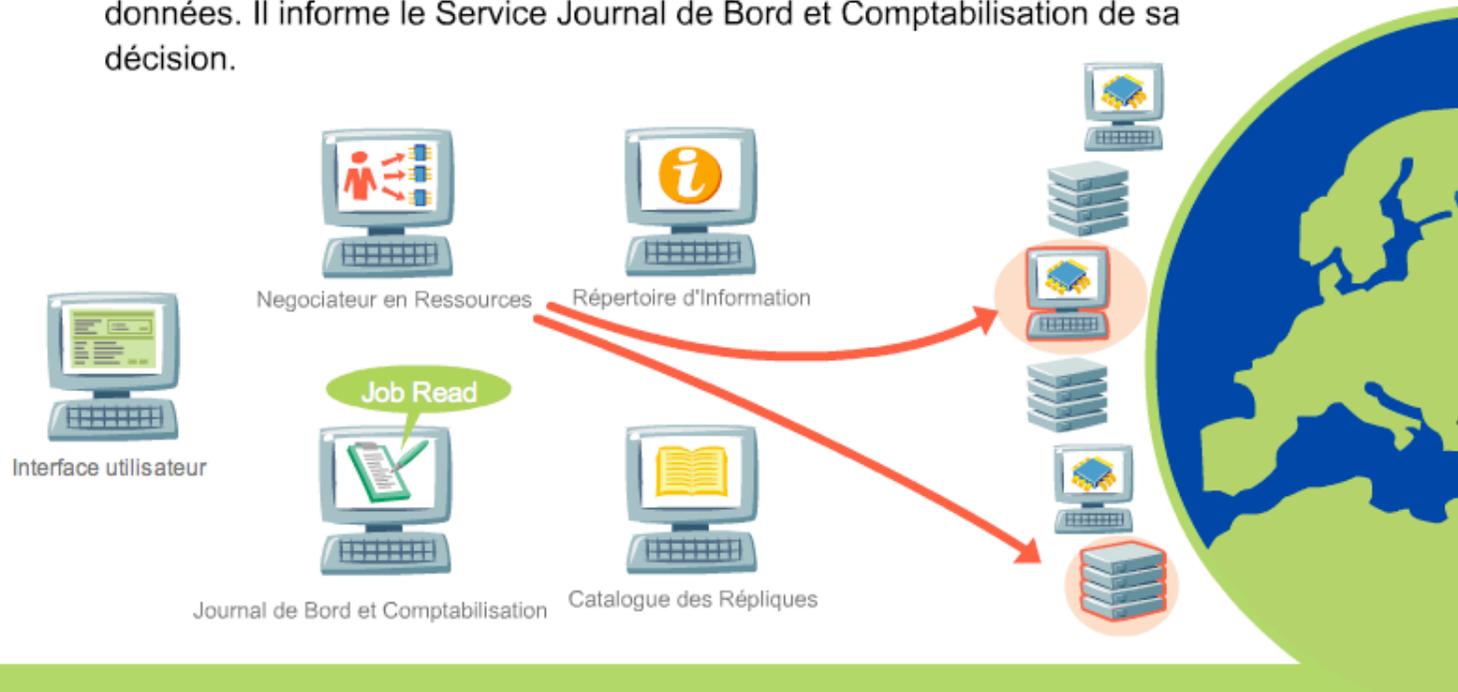
6. Soumission du job

Le **Catalogue des Répliques** et le **Répertoire d'Information** contiennent des informations sur l'état actuel de tous les sites. Le **Negociateur en Ressource** utilise cette information pour **assigner le job à un Element de Calcul** qui convienne. Durant cette étape, le job est en état d'ATTENTE.



6. Soumission du job

Le Négociateur en Ressources **fait son choix**, il a trouvé un **Élément de Calcul** qui convienne et un **Élément de stockage** qui aie les bonnes données. Il informe le Service Journal de Bord et Comptabilisation de sa décision.



6. Soumission du job

Le **Negociateur en Ressources** signale aussi sa décision à l'Élément de Calcul choisi et lui **soumet le job**.



7. Vérifier l'état

Durant tout ce processus, vous pouvez contrôler l'état de votre job en contactant le **Service Journal de Bord et Comptabilisation**.



8. Récupérer le résultat

L'exécution du **job est enfin terminée** sur l'Elément de Calcul.
L'Element de Calcul **transmet le résultat au Negociateur en Ressources.**



8. Récupérer le résultat

Vous pouvez **récupérer le Résultat** sur le Négociateur en Ressources.
Quand tout est terminé, **l'information du Journal de Bord est effacée.**





LIENS UTILES

- <http://gridcafe.org>
- <http://lcg.web.cern.ch/LCG> WLCG
- <http://lcg.in2p3.fr> LCG France

Sites français



La grille servira t'elle à tous ?

*Jules Piccard, a professor at the University of Basel, installed the first telephone in the city, around 1880, between his home and his institute. He showed it proudly to other scientists and got the comment: “**Looks very good, but I doubt it will ever have any practical use**”.*

"The world will only need five computers"
attributed to Thomas J. Watson, IBM

"640 kilobytes is all the memory you will ever need"
attributed to Bill Gates, Microsoft

"There is absolutely no need for a computer in the home"
attributed to Ken Olsen, DEC (once a leading minicomputer manufacturer)