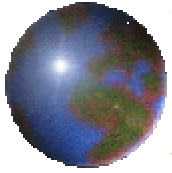


Le modèle de calcul pour ALICE

Yves Schutz (Subatech)

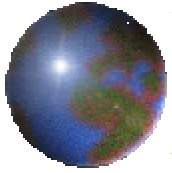




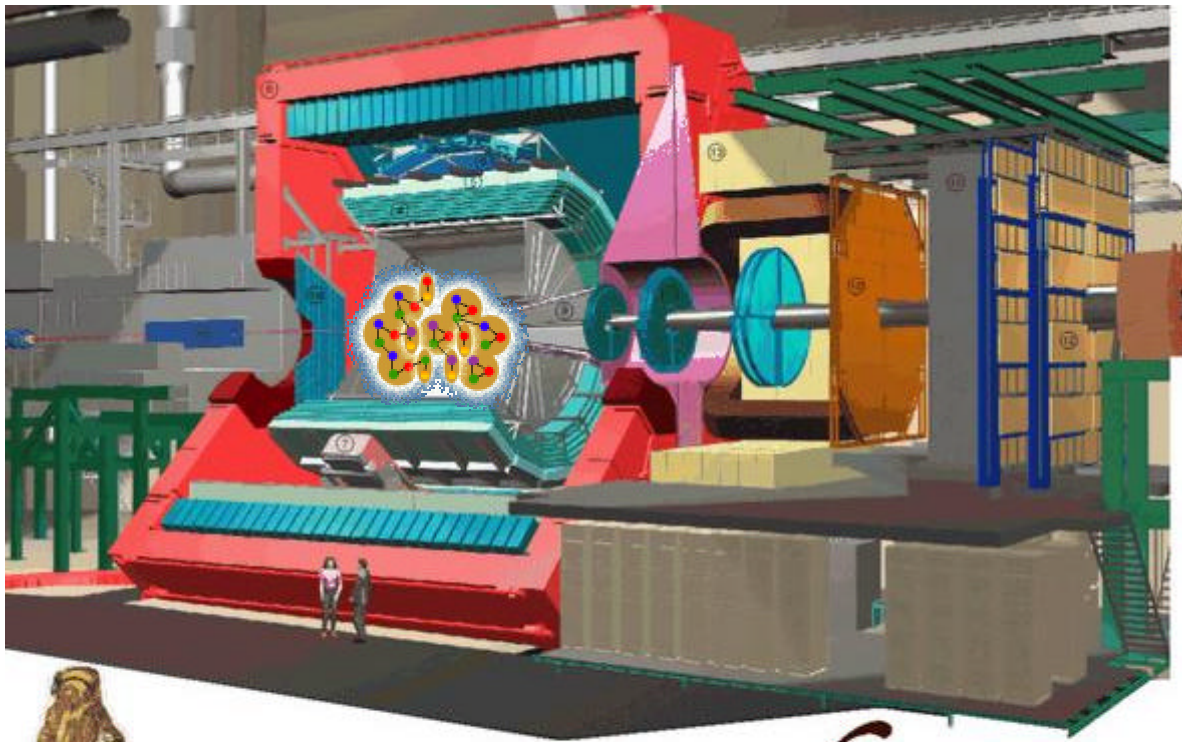
30' avec Alice au pays des ions ...

- Pourquoi ALICE ?
- Caractères propres des données
- Un modèle de calcul distribué
- Éléments précurseurs de la Grille
- Un environnement logiciel OO





Alice : expérience dédiée ions-lourds

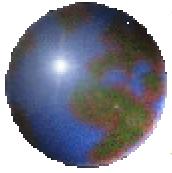


Alice

Juillet 2001

YS-Cargèse

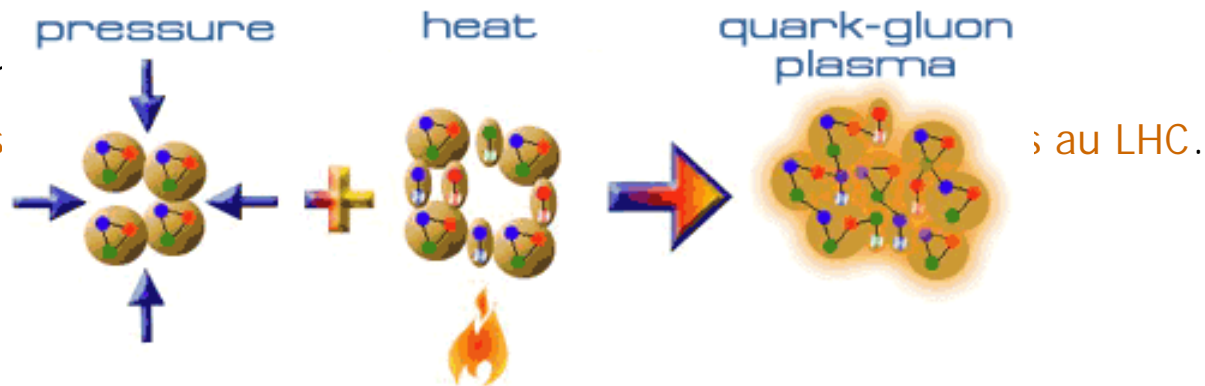
3

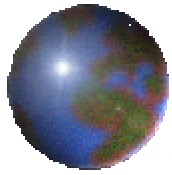


Imaginons...

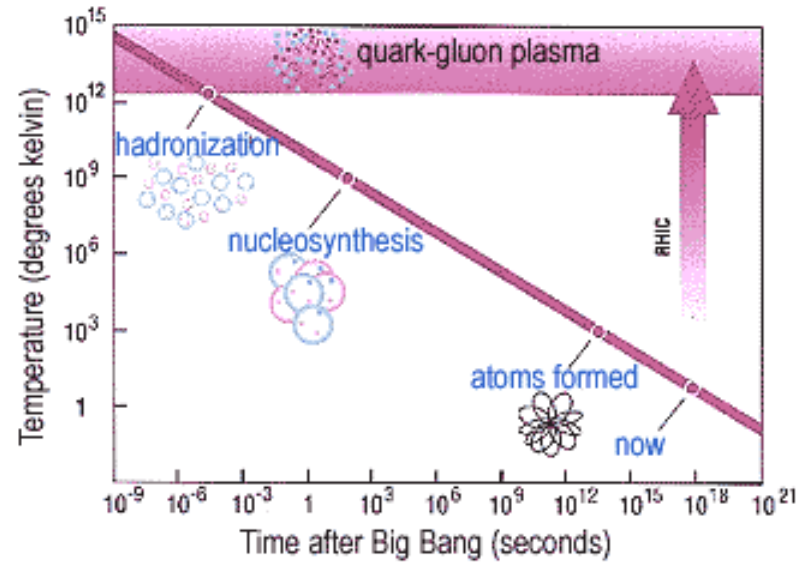
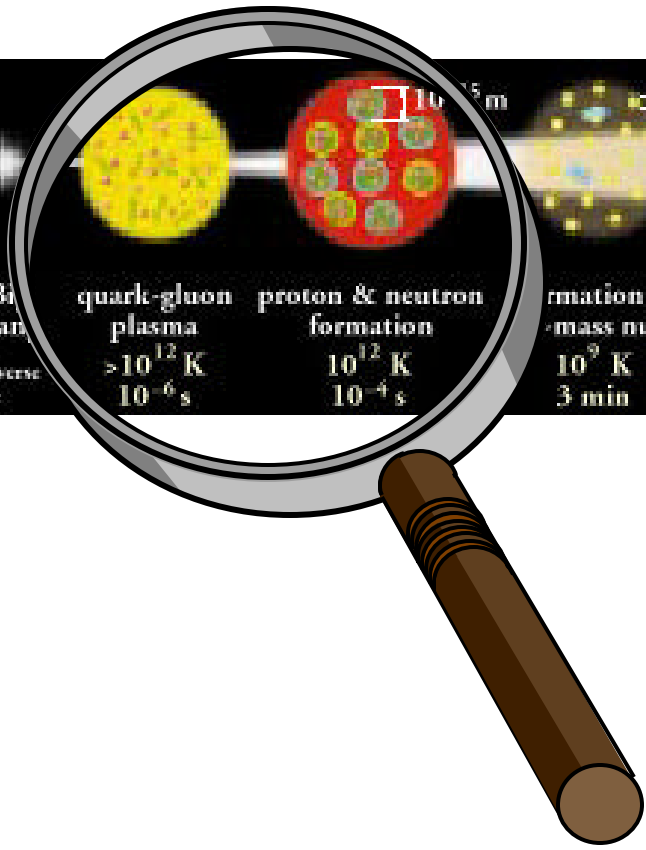
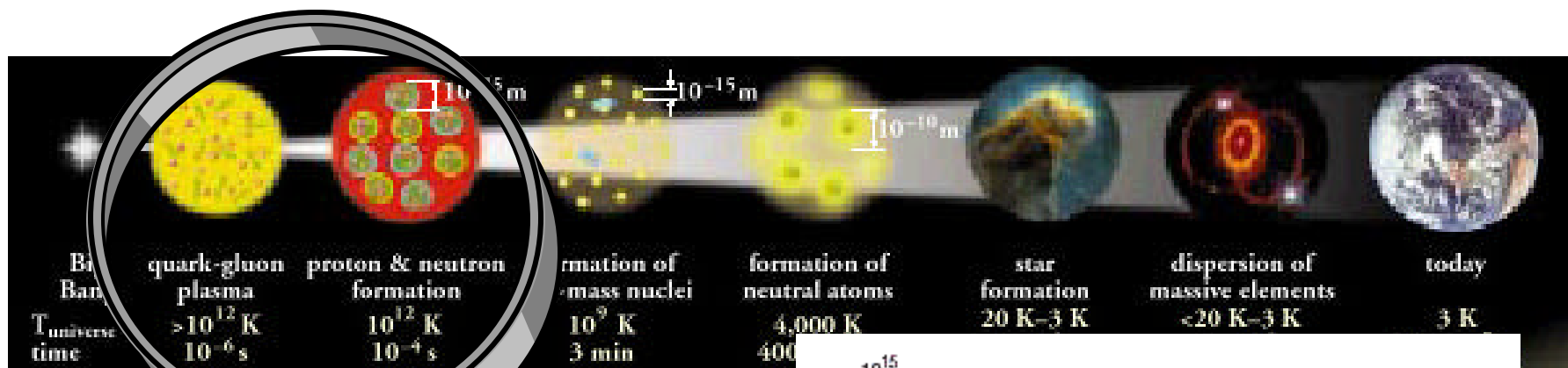
- Un monde gelé où l'eau n'existe qu'à l'état de glace, que
- la glace ne se trouve que sous forme quantifiée : des glaçons,
- des théoriciens soutiennent que la glace doit exister en phase liquide, et que
- la seule façon de chauffer la glace est de frapper un glaçon contre l'autre
- Vous fabriquez donc un gros paquet contenant un million de glaçons
- que vous frappez contre un autre paquet semblable,
- 10 millions de fois chaque seconde.
- Vous produisez ainsi environ 1000 collisions glaçon-glaçon chaque seconde
- et les vous observez du voisinage de Mars ...

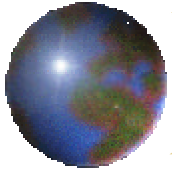
- Changeons n
Ca y est nous



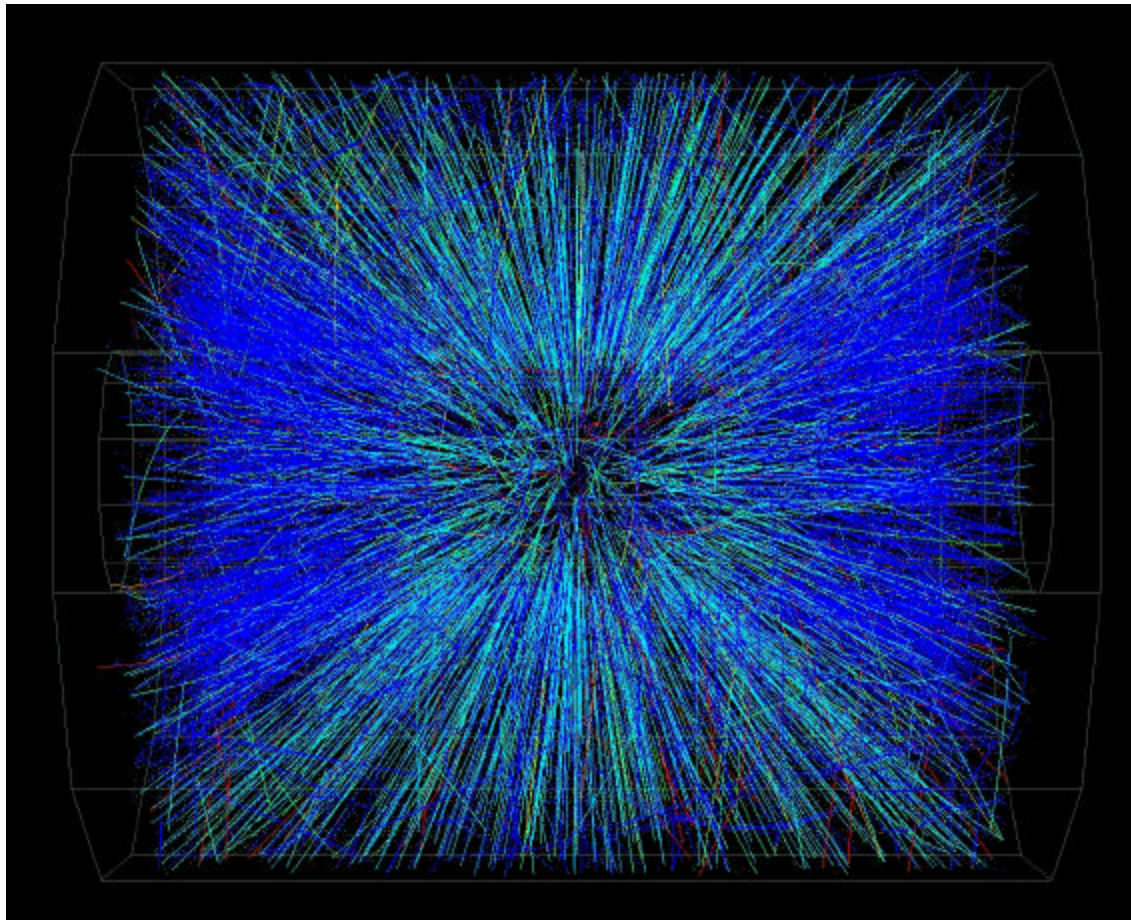


L'expérience a déjà eu lieu ...



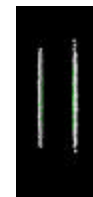
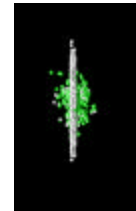
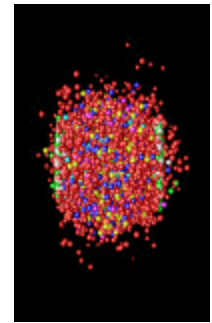
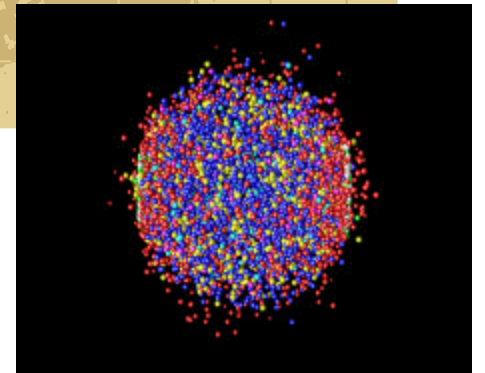


Imageons...



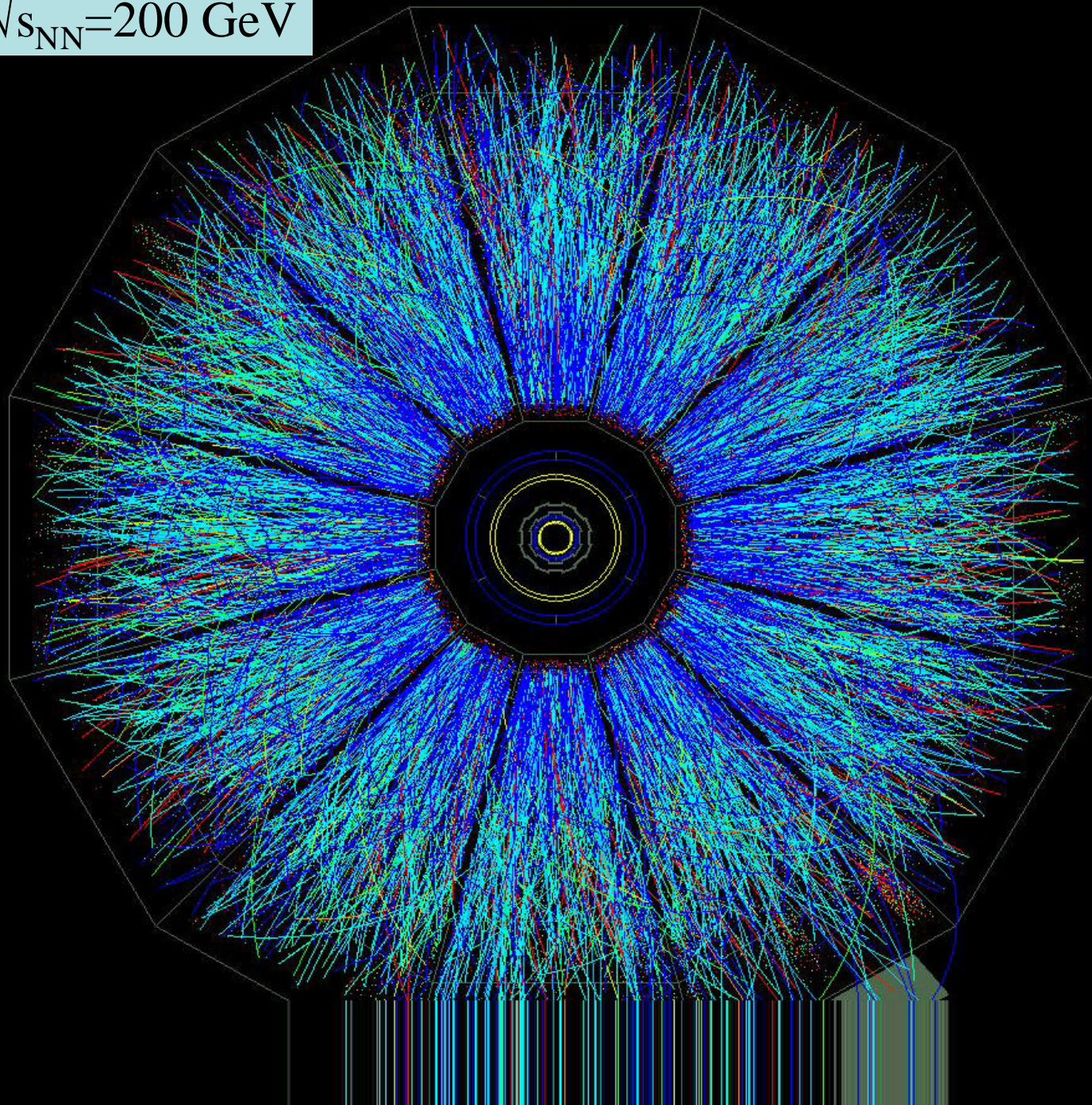
Juillet 2001

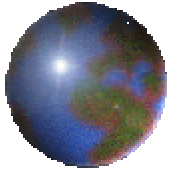
YS-Cargèse



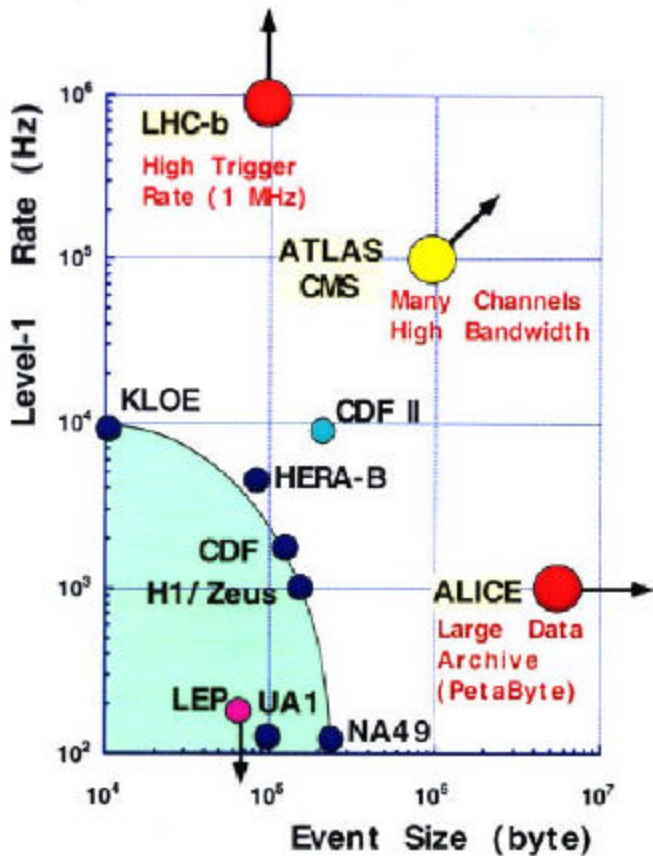
6

Au+Au $\sqrt{s_{NN}}=200$ GeV

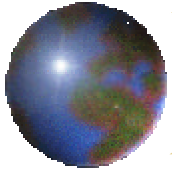




Les données collectées par ALICE



- 16.000 particules dans le détecteur
- échantillonnées à 1 KHz
- 1 événement 80 Mo, <25 Mo>
- enregistrement 1,25 Go/s pendant 10⁶s
- enregistrement 1,00 Mo/s pendant 10⁷s
- reconstruction 5 × 10¹⁰ KSI2000 s
- une année de données 2,5 × 10¹⁵ octets
- simuler 1 événement 25 MSI2000 s
- 500 utilisateurs potentiels répartis dans le monde (26 Pays, 70 Instituts)



Ressources requises pour Alice

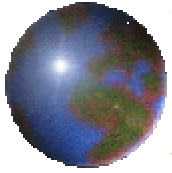
- Stockage de masse

| | Raw | Reconstructed (Tbytes) | | Total | DB | Grand Total |
|--------------|------------------|---------------------------|----------------|------------------|--------------|------------------|
| | | One pass | Many passes | | | |
| Pb+Pb Data | 1 250,000 | 137,600 | 300,400 | 1 688,000 | | |
| p+p Data | 1 000,000 | 34,400 | 75,100 | 1 109,500 | 0,003 | |
| MC Data | 20,000 | 1,376 | 5,504 | 26,880 | | |
| Total | 2 270,000 | 173,376 | 381,004 | 2 824,380 | 0,003 | 2 824,383 |

- CPU

| | First reconstruction | Additional Reconstructions | Simulations & Reconstructions |
|------------|----------------------|----------------------------|-------------------------------|
| | (kSI2000) | | |
| Pb+Pb Data | 9,645E+03 | | 9,645E+03 |
| p+p Data | 3,858E+03 | | 7,716E+03 |
| MC Data | | | 1,447E+01 |

- Réseau : > 1.5 Gbits/s



Traitement distribué des données

- Tier-0 (CERN) :

- ❑ DAQ
- ❑ L3 (niveau de déclenchement logiciel)
- ❑ Stockage (écriture sur bande magnétique)
- ❑ Export (données brutes et reconstruites vers Tier-1)
- ❑ DB (mise à jour et synchronisation)

- Tier-1 (CERN, Allemagne, France, GB, Italie, USA) :

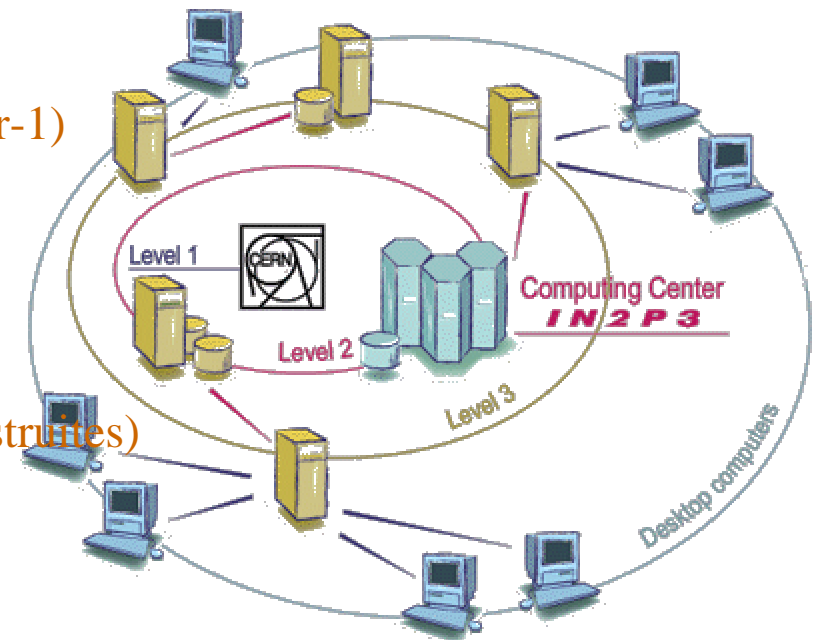
- ❑ Reconstruction (plusieurs par jeu de données)
- ❑ Simulations
- ❑ Stockage (simulations, % données brutes+reconstruites)
- ❑ import/export (de Tier-0/1, vers Tier-0/1/2)
- ❑ Miroir DB (synchronisation)

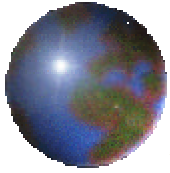
- Tier-2 :

- ❑ Reconstruction (données du Tier-1 local)
- ❑ Import/export (de et vers Tier-1 local)

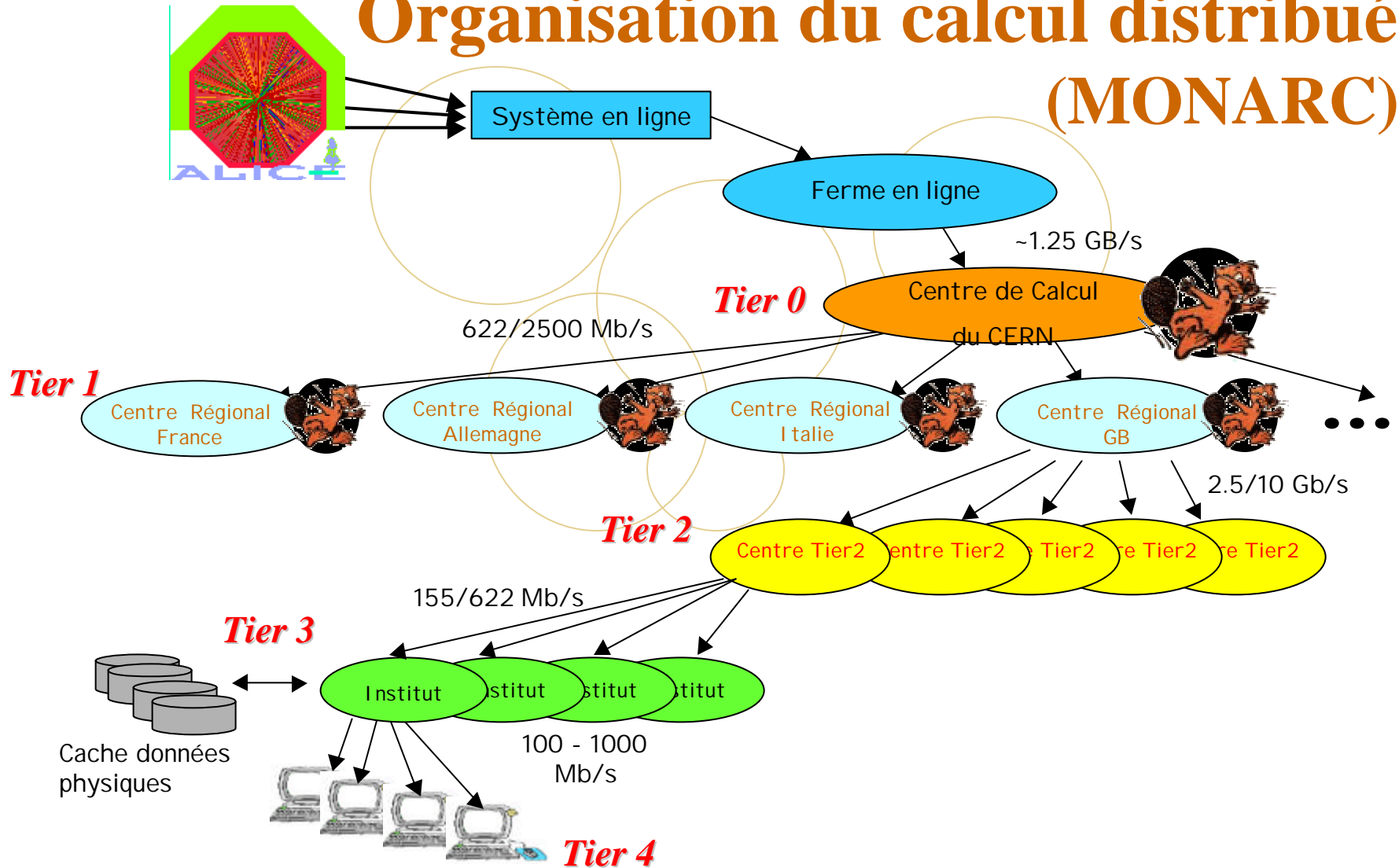
- Tier-3/4 :

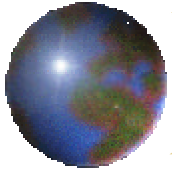
- ❑ Analyse
- ❑ Import/Export





Organisation du calcul distribué (MONARC)

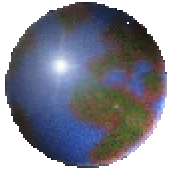




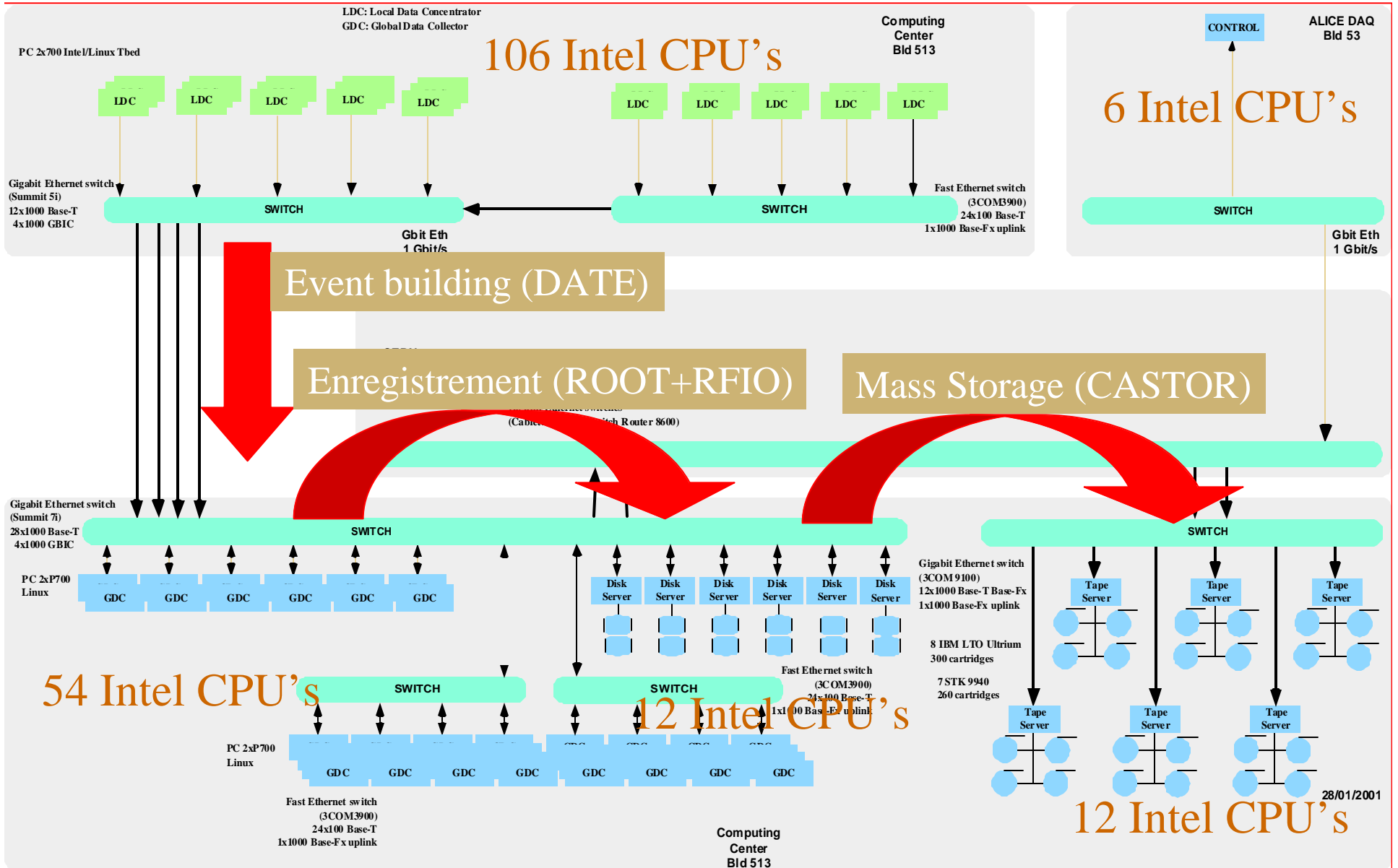
ALICE Data Challenges

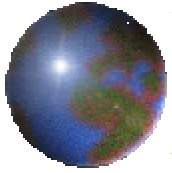
- Objectif 1,25 Goctets/s
- DC réguliers de complexité et taille croissantes
- Aujourd'hui (essai sur 3 mois):
 - 120 Moctets/s pic et 85 Moctets/s moyenne
 - 20M événements, run de 86 heures
 - 100.000 fichiers, 110 Toctets dans CASTOR



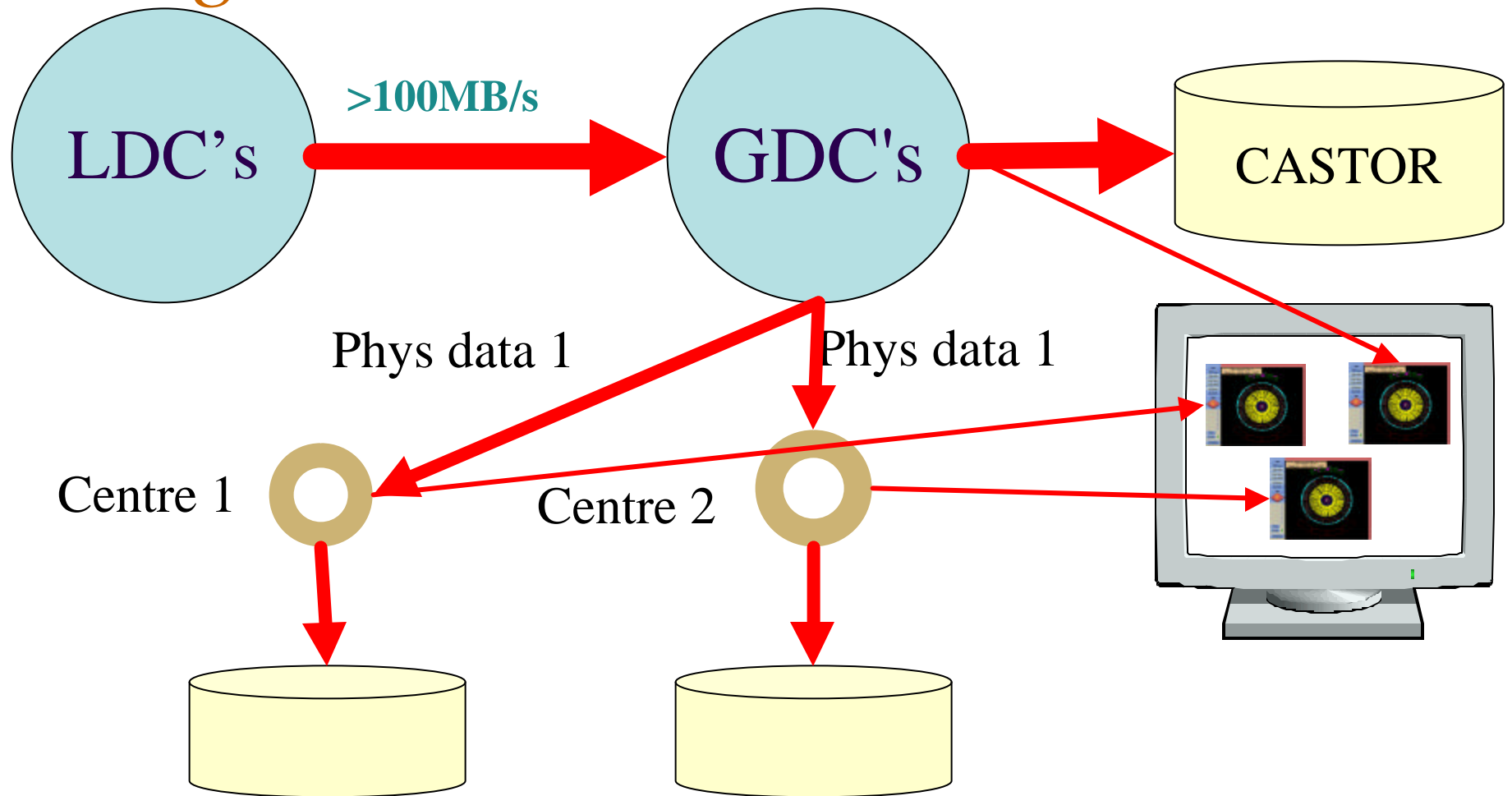


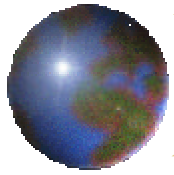
ALICE DC III



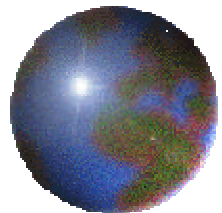
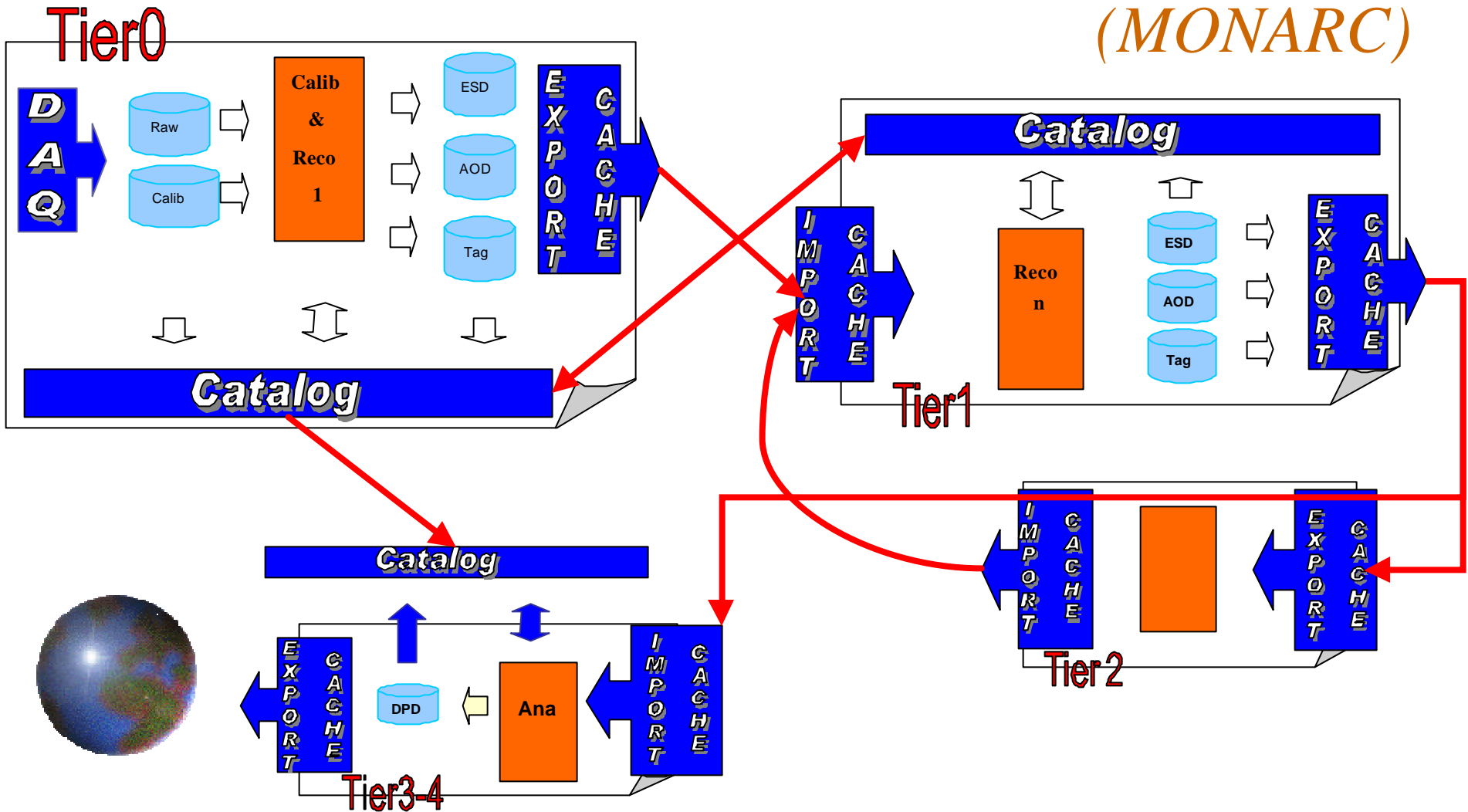


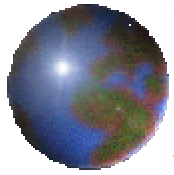
En ligne: ALICE DC III/MONARC



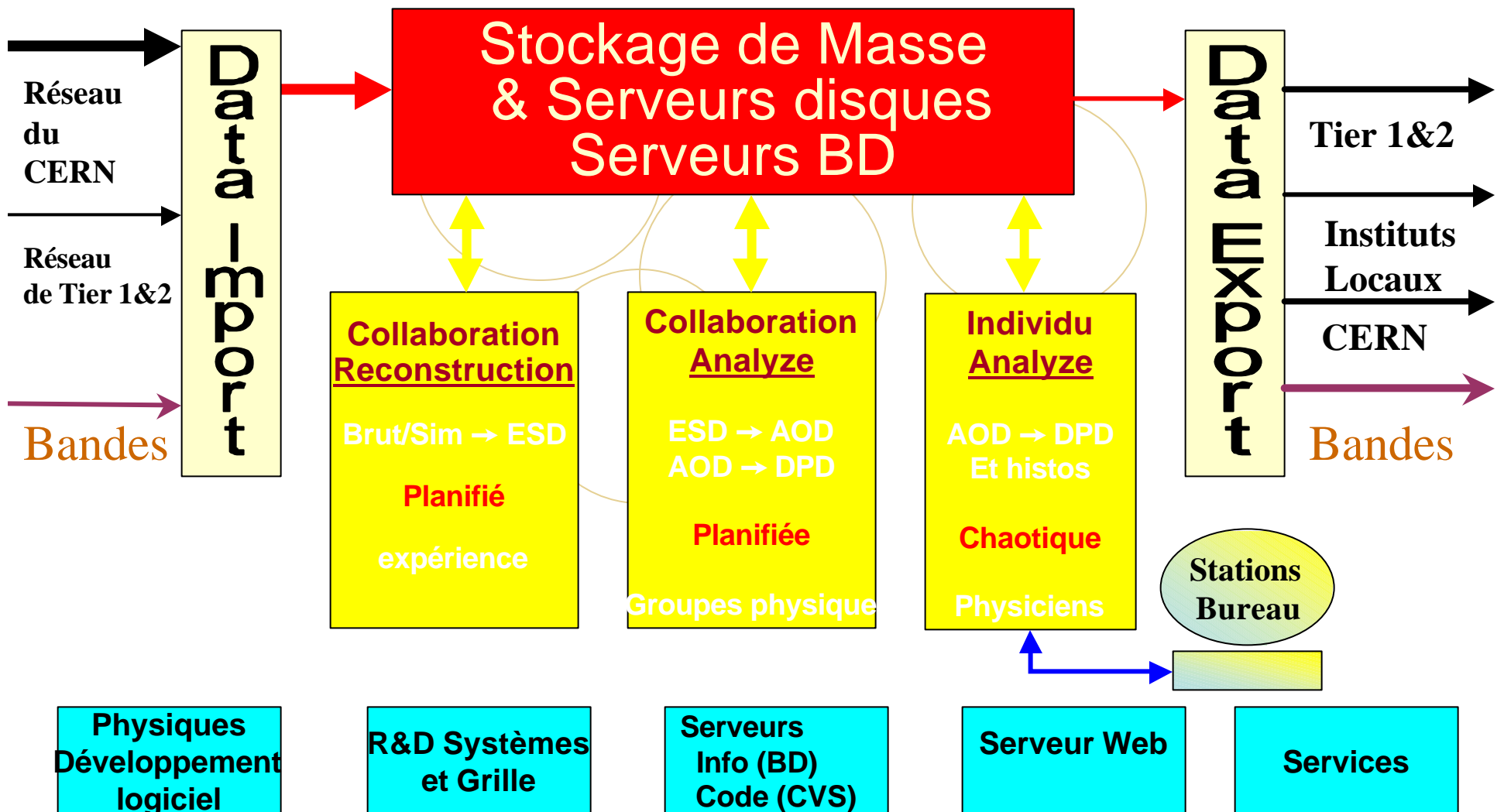


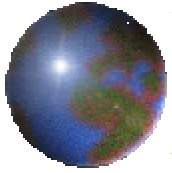
Hors-ligne: Modèle de calcul distribué (MONARC)





Tâches d'un Tier-1





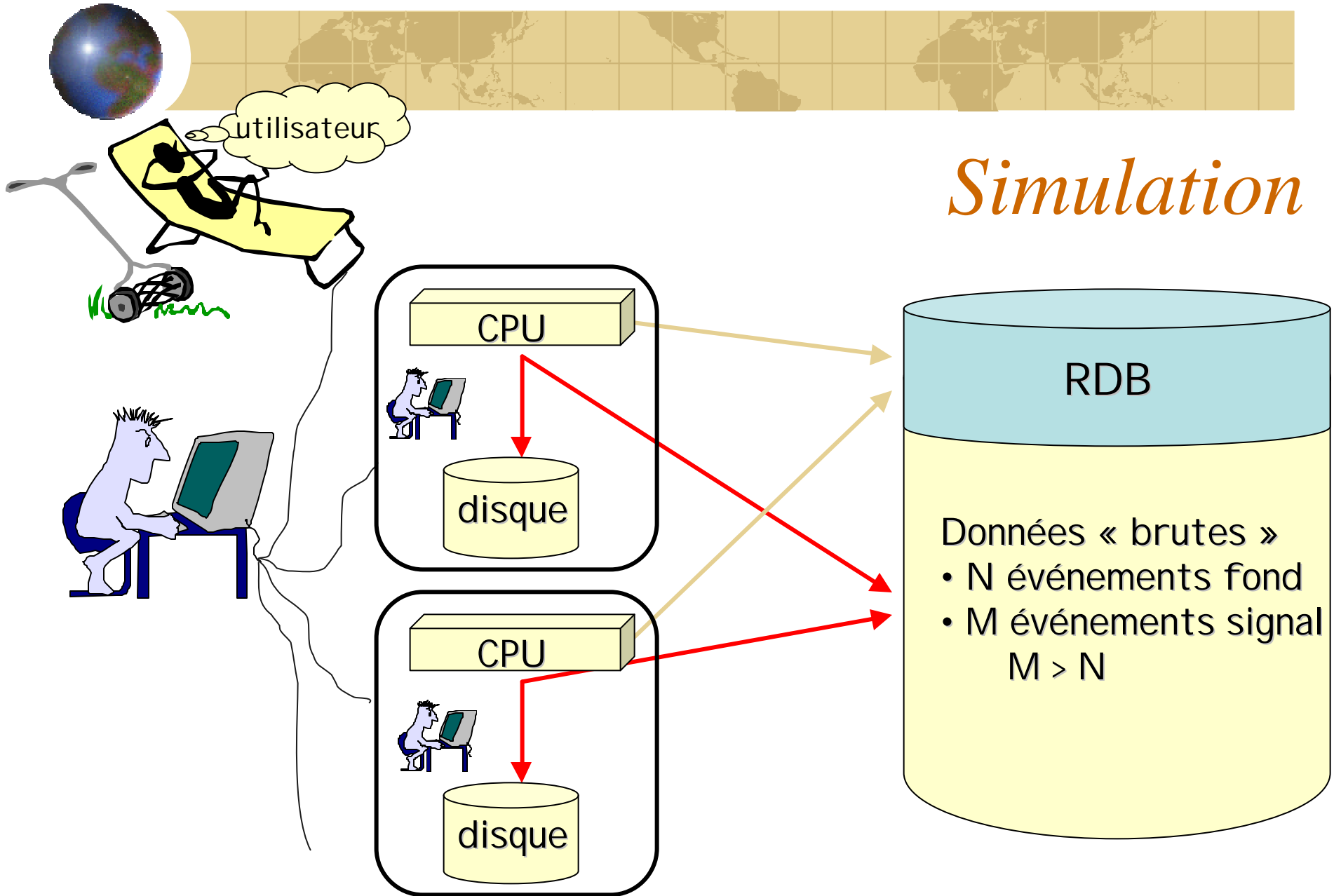
Tier-0/1/2 Alice

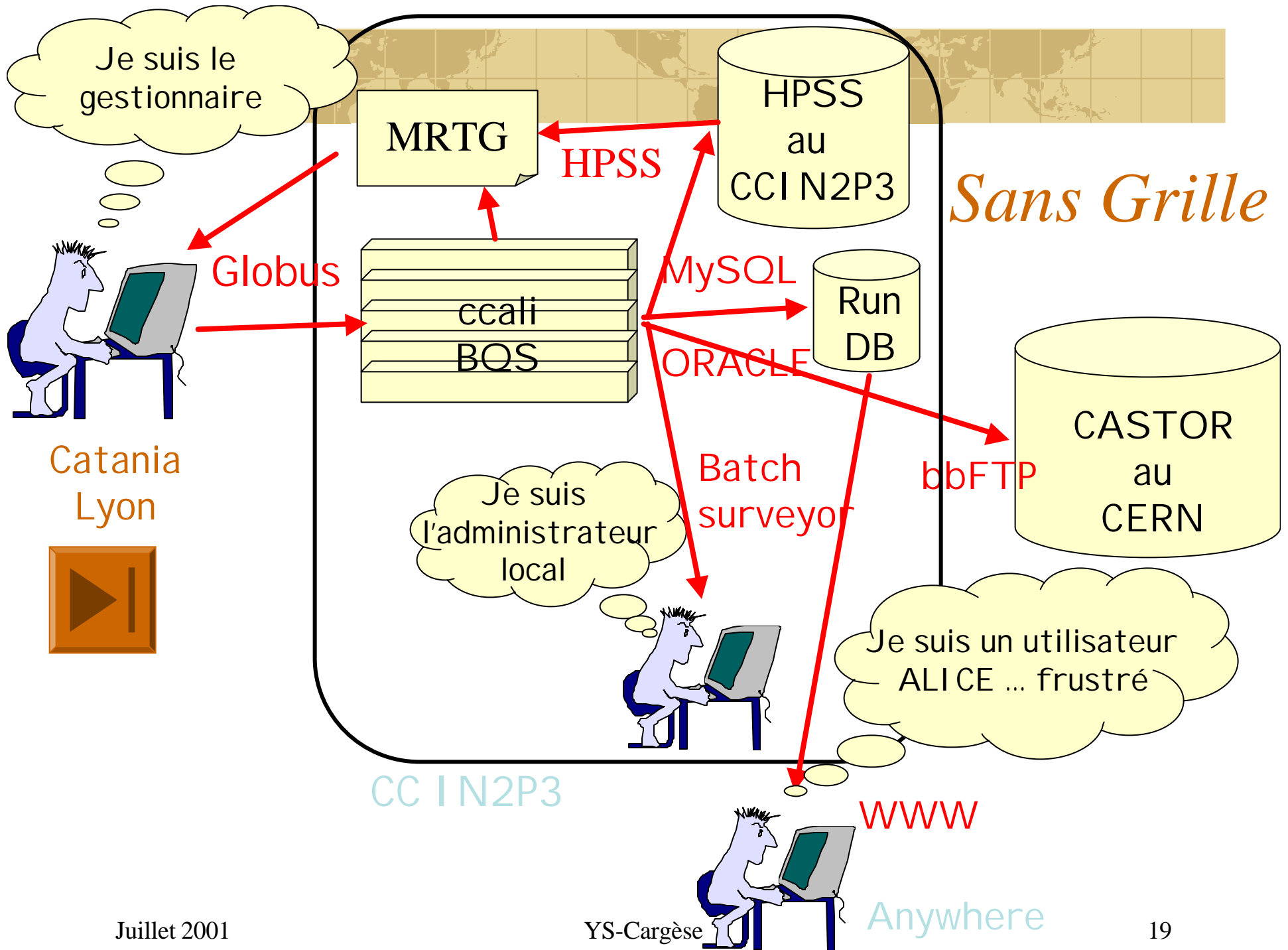
- Stockage de masse

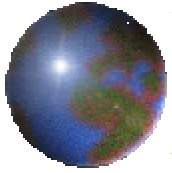
| | Mass Storage (Tbytes) | | | Disk Storage(Tbytes) |
|--------------|-----------------------|---------------|------------------|----------------------|
| | Raw | Reconstructed | Total | |
| Tier0 | 2 250,000 | 172,000 | 2 422,000 | 726,600 |
| Tier1 | 562,500 | 272,595 | 835,095 | 250,529 |
| Total | | | 5 762,380 | 1 728,714 |

- CPU

| | CPU (MSI2000) |
|--------------------|---------------|
| Tier0 | 9,645 |
| Tier1+Tier2 | 2,411 |
| Total | 19,290 |



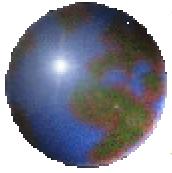




La solution logicielle

- Fin des années 90 :
 - FORTRAN+ZEBRA+PAW
 - Des logiciels qui ont fait leurs preuves
 - ALICE est une « petite » collaboration
 - La route « officielle » était tracée mais ne menait nulle part
- Il fallait :
 - Suivre une seule route
 - Réutiliser du code
 - Que ça marche

Avec des dinosaures, des prophètes et des utilisateurs égarés

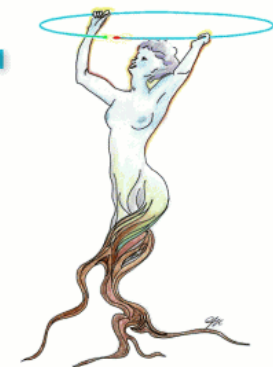


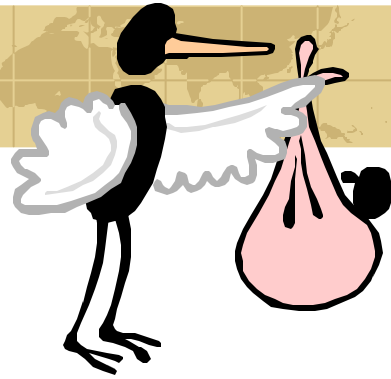
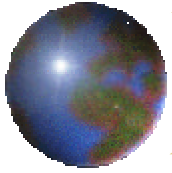
Le choix d'ALICE (1998)

- Approche entièrement OO, implémentation C++
- Des outils qui marchent (G3) sont interfacés
- Tout est bâti sur Root
 - Persistance,
 - I/O,
 - Outils graphiques,
 - Outils de traitement des données
 - GUI,
 - Générateur de documentation, etc. ...

ROOT

An Object-Oriented
Data Analysis Framework

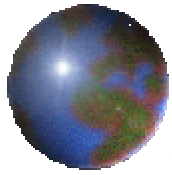




Et naquit AliRoot

- Une couche par dessus ROOT
 - Interface des logiciels externes (G3/4, Fluka, Hijing, Pythia)
 - Implémente tous les détecteurs ALICE
 - Simulation détaillée
 - Reconstruction des données simulées et ... réelles
 - Analyse des données reconstruites
 - 420 classes, 275 Klignes C++

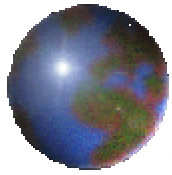




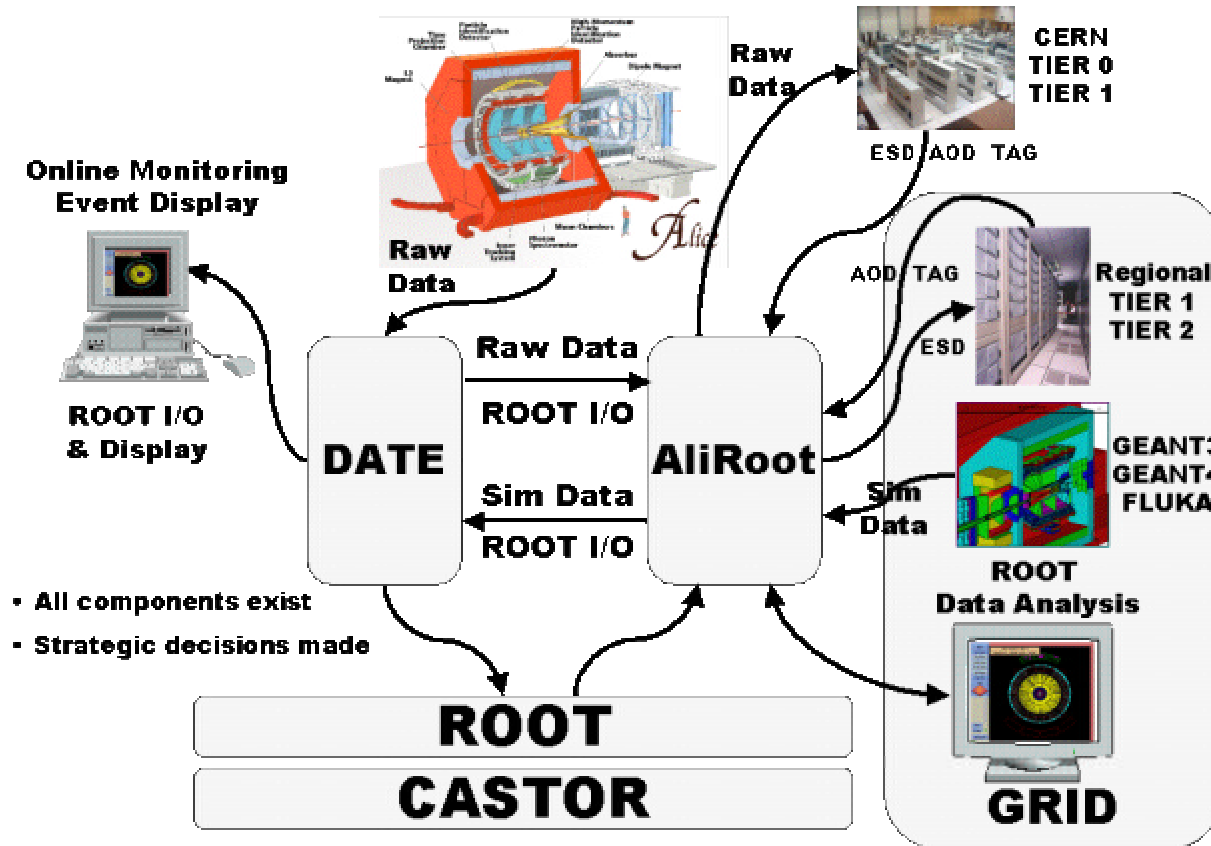
Cathédrale ou bazar

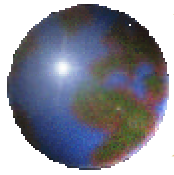


- Prototypage et interaction avec utilisateurs rapides :
 - Les utilisateurs expriment continuellement leurs besoins
- Uniquement ce qui peut être prototypé immédiatement est conçu :
 - Les utilisateurs participent continuellement à la conception
- Tests de qualité et de robustesse :
 - Tests effectués automatiquement chaque nuit
- Interactions utilisateurs/développeurs :
 - Discussion permanente et animée sur forum

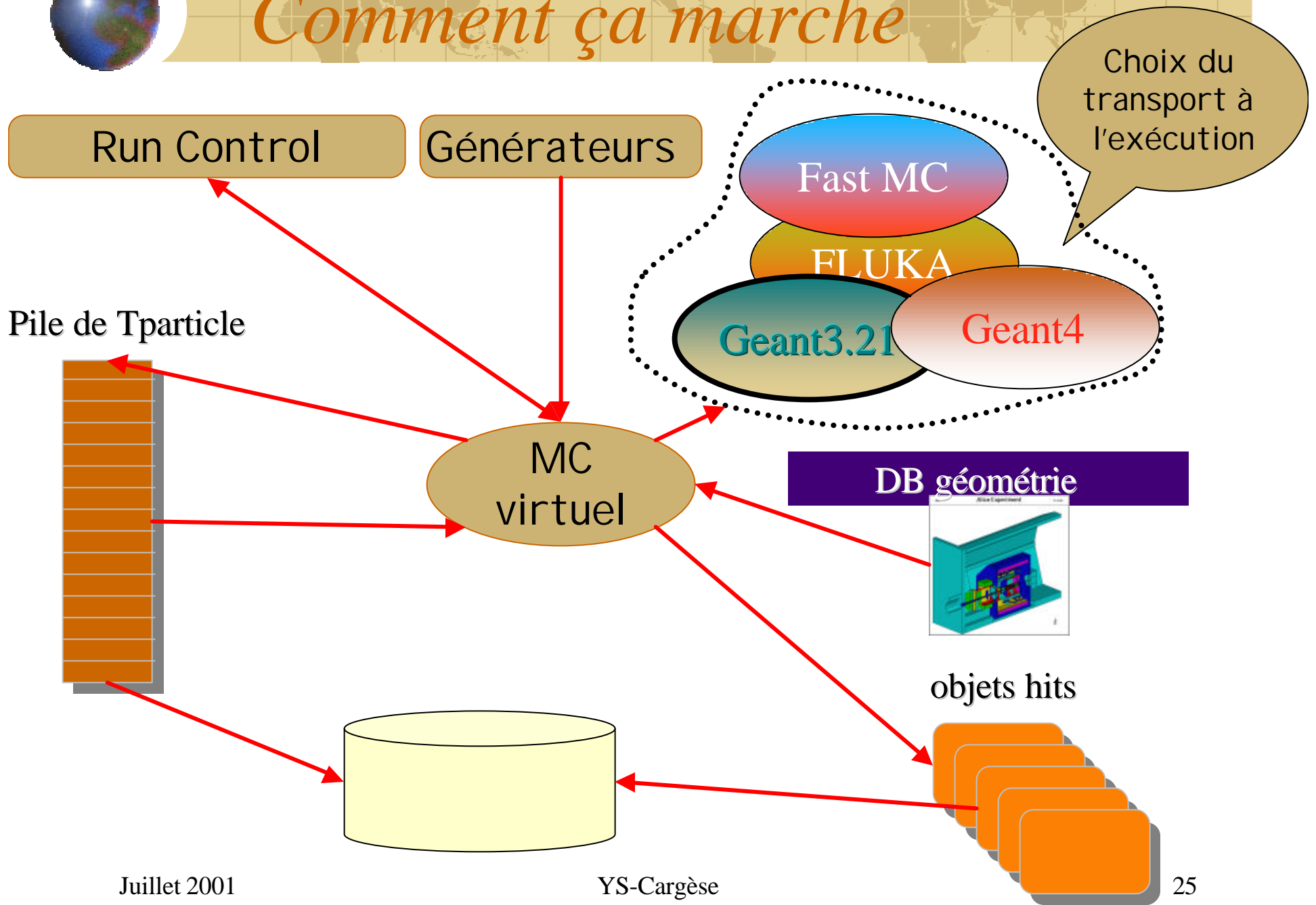


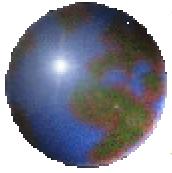
Root et AliRoot dans Alice





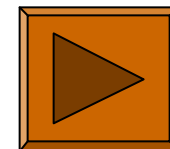
Comment ça marche

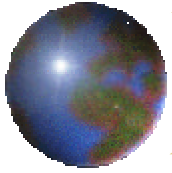




Modèle AliRoot

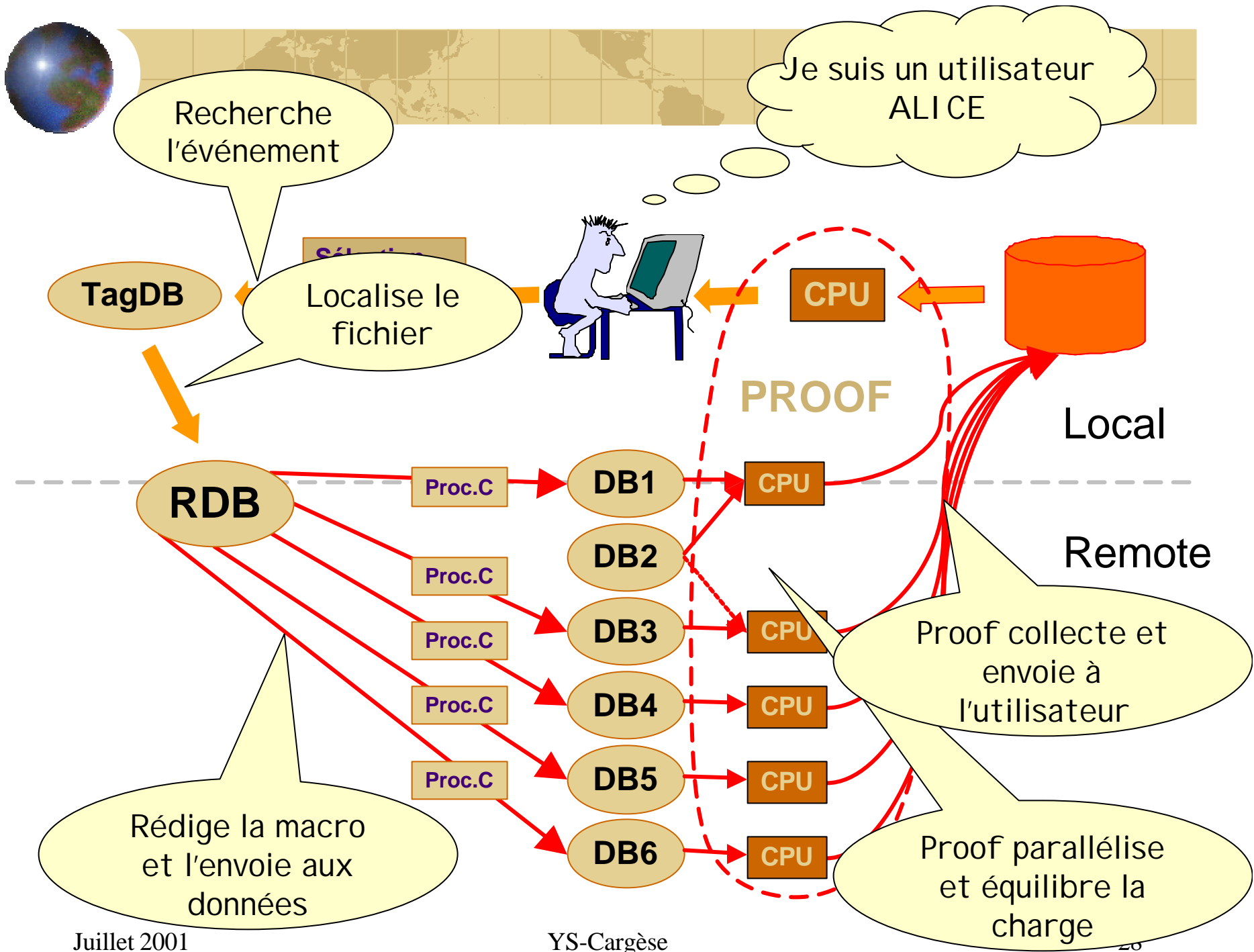
- Modularité : 1 détecteur = 1 *.so
- Séparation des objets OO
 - Objets statiques (géométrie, paramètres, ...)
 - Objets dynamiques
 - Persistant : *tree* -> branche -> conteneur -> objets
 - Transitoire : *folder* à la UNIX -> conteneur * -> objets *
 - Objets algorithmiques
 - Encapsulation des méthodes -> tasks
 - Communication entre objets :
 - *Whiteboard* -> *folders*

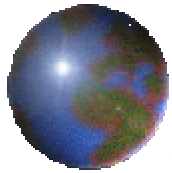




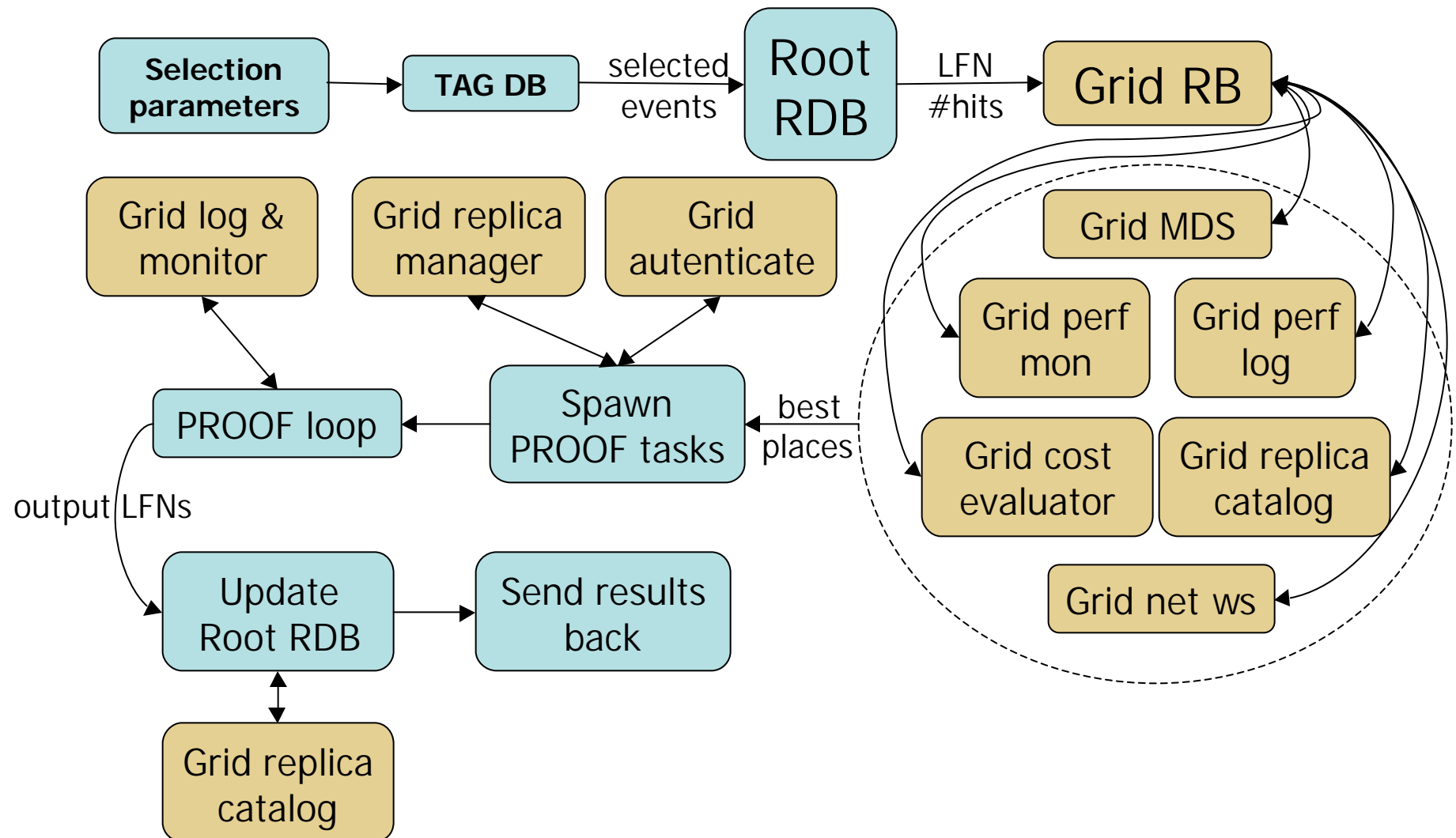
Analyse distribuée ... sans OODB

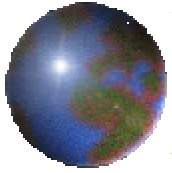
- Amener le code d'analyse aux données
 - Macro C++
 - Interpréteur CINT
 - Compilation au vol
- Exploiter la totalité du CPU disponible
 - Parallélisation du processus
 - Équilibrage de la charge





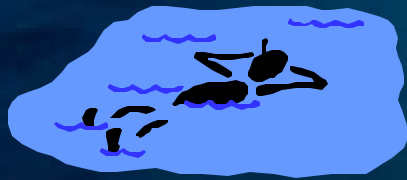
Services DataGrid & ROOT

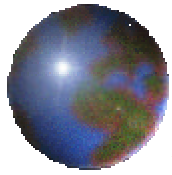




Conclusions

- Le développement des logiciels de Alice est en bonne voie
- Les choix faits par la collaboration sont déjà aujourd'hui payants
- Les essais de simili GRILLE sont lancés





EngineFrame

Alice Site info Host info Alice runs Multi runs ?

Globus Services

- Proxy Services
- Job Services
- LDAP Services
- Alice Services

[Login to EnginFrame at gridct1.ct.infn.it](#)

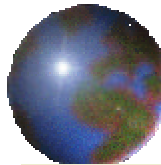
Username:

Password:

© Copyright NICE srl, 1998-2001

NICE **INFN** **Alice**

Web interface for job submission!



ROOT Object Browser

File View Options Help

PHOS

All Folders Contents of ".../WhiteBoard/QAAlarms/PHOS"

root
/afs/in2p3.fr/home/throng/c
ROOT Files
Folders
gAlice
YSAlice
WhiteBoard
 Geometry
 QAAlarms
 + PHOS
 Hits
 + PHOS
 SDigits
 Digits
 RecPoints
 TrackSegments
 RecParticles
tasks
 QA
 + PHOS
 SDigitizer
 PHOS
 Digitizer
 Reconstructioner
AlimCQA

HitsM HitsMB1 HitsMB2 HitsMB3 HitsMB4 HitsMB5 TotEn TotEnB1
TotEnB2 TotEnB3 TotEnB4 TotEnB5

TFolder's

TClonesArray's

TTask's

16 Objects 12 Objects. An array of objects