

Rendez-vous International de la Conception et de la Simulation Numérique Hautes Performances  
*International meeting for High performance Numerical Design and Simulation*

# Forum Ter@tec

27 & 28 juin / june 2012

ECOLE POLYTECHNIQUE PALAISEAU-FRANCE

*Simuler pour concevoir*  
*Design by simulation*

[www.teratec.eu](http://www.teratec.eu)

Sponsors Platinum



Sponsor Gold



Partenaire



Sponsors Silver



DataDirect  
NETWORKS  
INFORMATION IN MOTION



FUJITSU





# Atelier 5

Déluge(s) de données

*Jeudi 28 juin 2012 - 14h00 - 17h30*

*Animateur*

*Jean-Philippe NOMINE*

**CEA**



*[jean-philippe.nomine@cea.fr](mailto:jean-philippe.nomine@cea.fr)*





# Déluge(s) de données



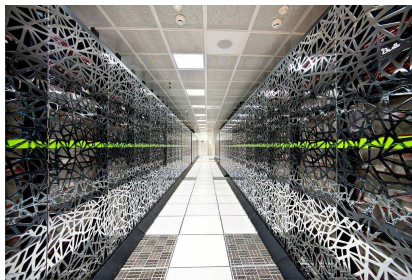
*Chercheurs et ingénieurs sont depuis longtemps producteurs et consommateurs de grandes masses de données : issues de calculs, d'observations, de mesures, ou encore de CAO et de modèles numériques ; de volumes sans cesse croissants, avec une distribution et une circulation de plus en plus poussées sur réseau.*

*Mais d'autres masses de données déferlent désormais autour de nous : commerciales, administratives, financières, multimedia, sur les réseaux sociaux, issues de capteurs, relatives à des objets nomades.... Qu'elles soient privées ou publiques, les modalités de représentation, stockage, traitement, analyse et exploitation de ces données repoussent les limites des approches conventionnelles ou appellent de nouvelles méthodes et techniques pour livrer leur sens et produire de la valeur.*

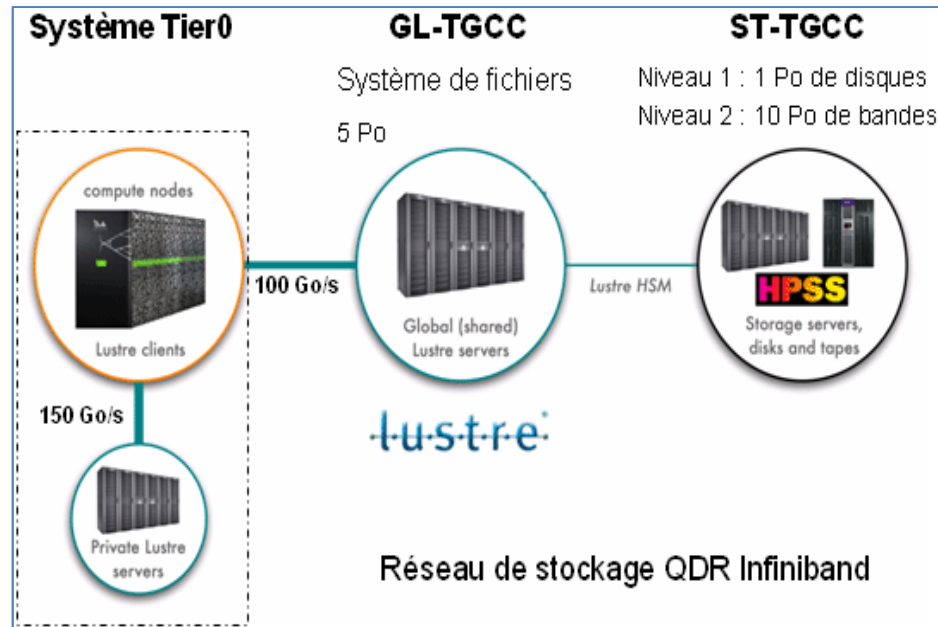
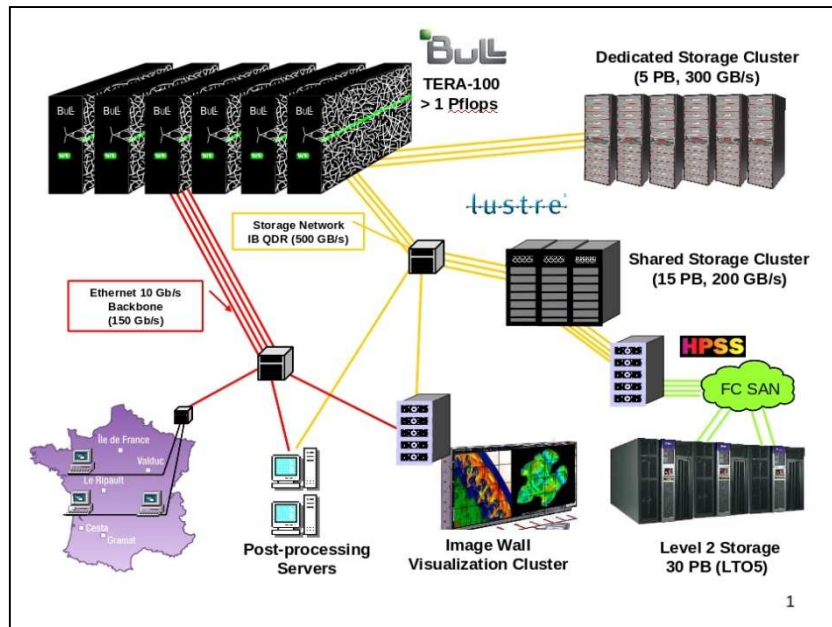
*Cet atelier demande à des observateurs et acteurs éclairés de ces phénomènes de nous livrer leur vision et quelques exemples et éléments pouvant inspirer notre réflexion.*

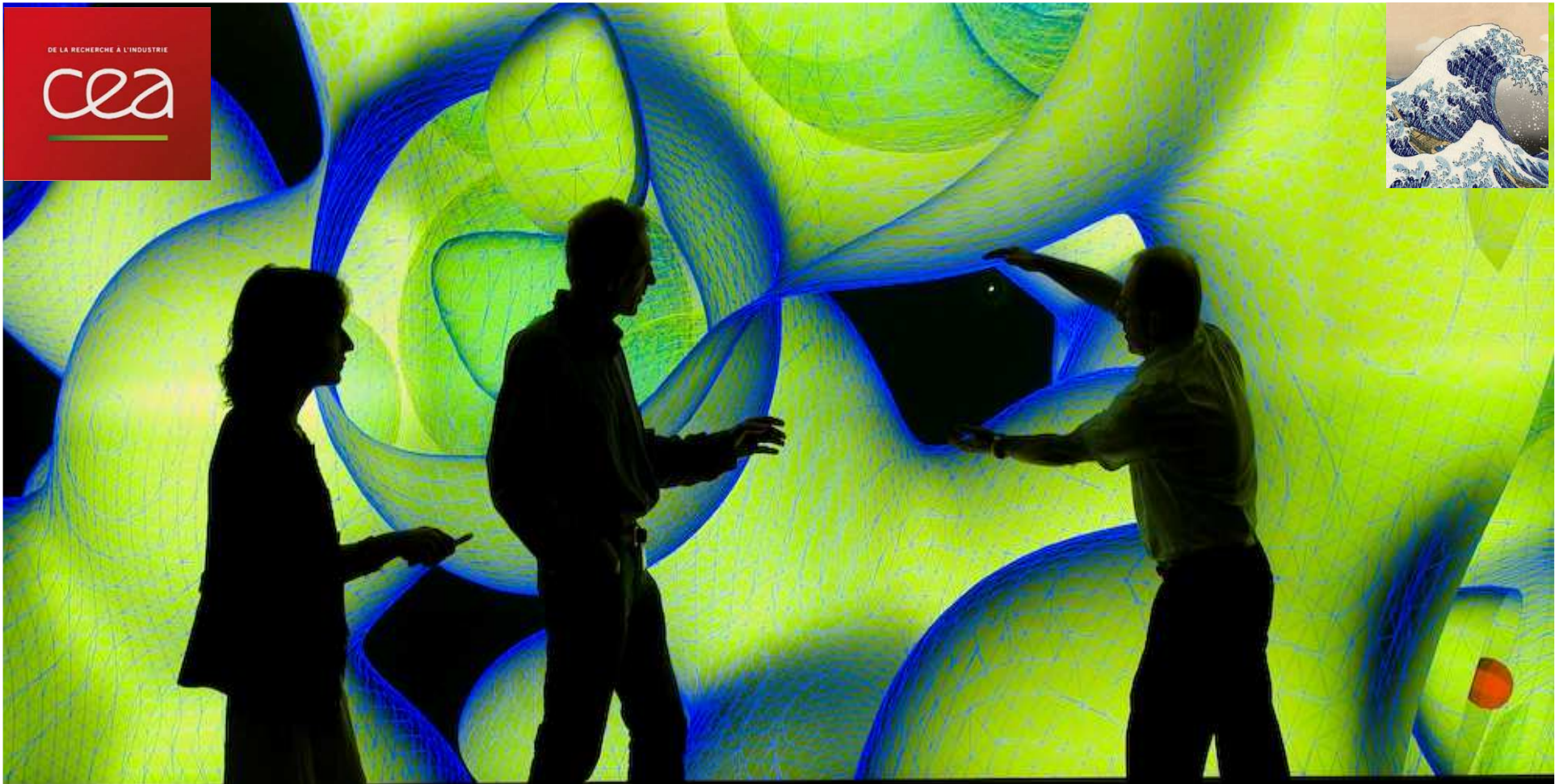


# Centre(s) de calcul...ex: CEA

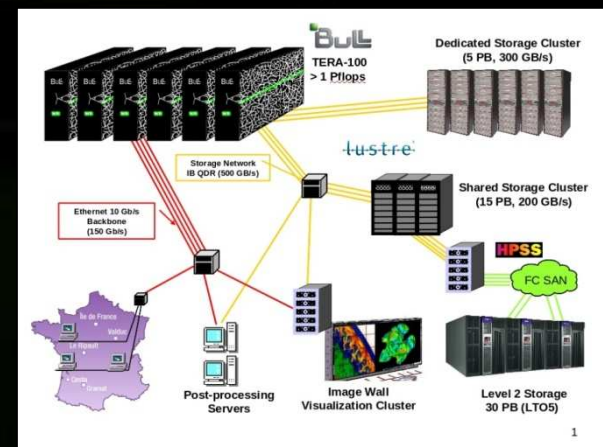


TERA 100 = 1.25 PF / CURIE = 2 PF  
 Lustre, data-centric  
 Quelques PO level 1, jusqu'à 20 PO level 2  
 Quelques 1 à 10 TO produits/jour  
 Centralisé, utilisateurs « distants »  
 Déport de visualisation

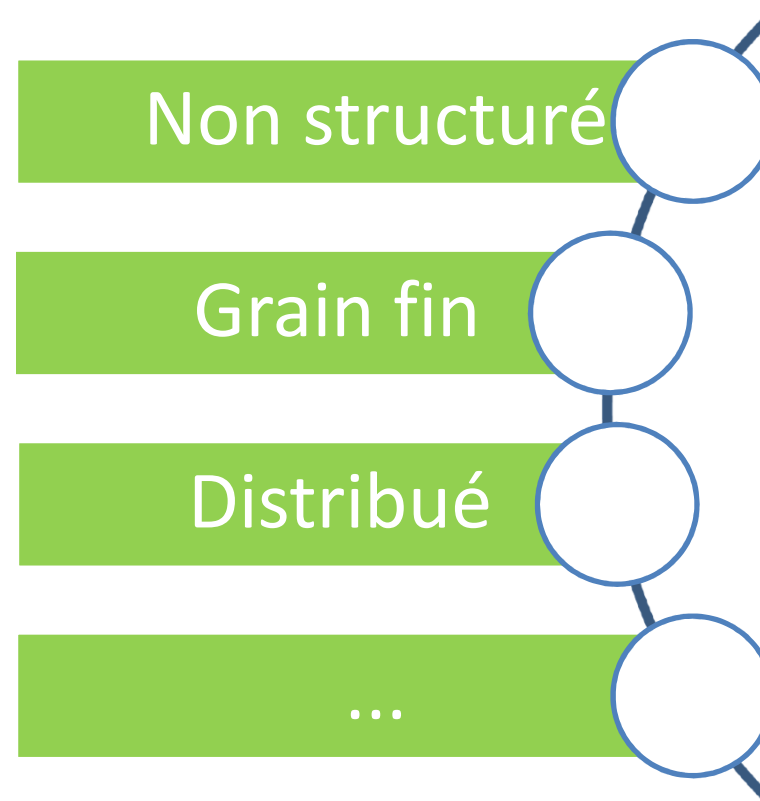
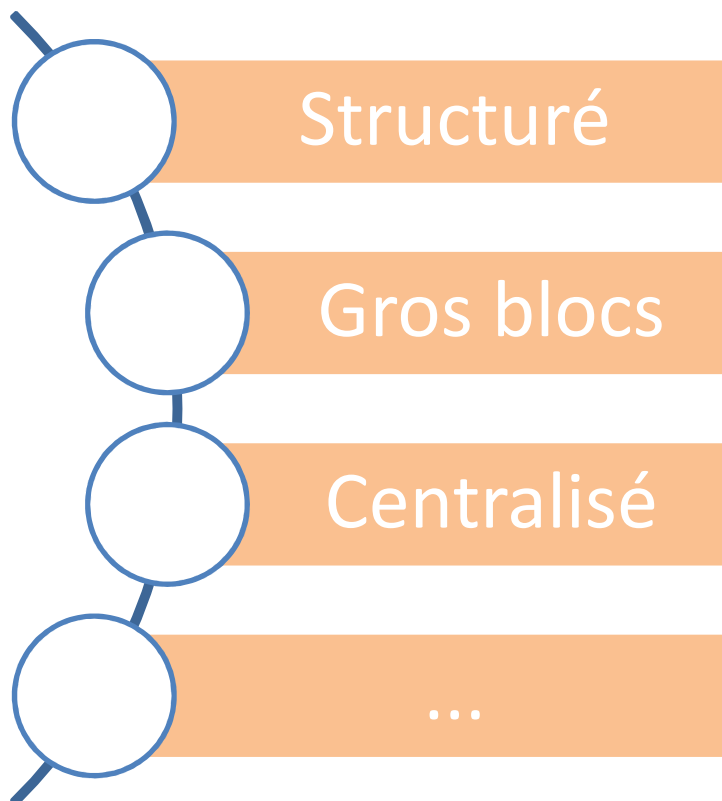




- Maillages(t), particules(t)
- Extractions en temps et en espace
- Opérations sur les extraits temps/espace
- Comparaisons entre cas / paramétrique
- Traitement sur place ou redistribution
- Tendance à plus de données « assimilées » mixées au calcul, à des données moins structurées (génomique...)



# Large Data Sets # Big Data ?





Centre de  
calcul

Vision globale  
Perspectives

Application  
Biologie

Solutions  
Exemples

Solutions  
Exemples

Solutions  
Exemples  
Multimedia  
PME/startup

# Déluge(s) de données



14h00	Introduction Jean-Philippe NOMINE, CEA
14h10	Big Data : More is Different* * Déluge de données: changement d'échelle, changement de paradigme Henri VERDIER, CAP DIGITAL
14h55	Imagerie génétique : des défis computationnels à des implémentations gros grains sur architecture cluster Vincent FROUIN, CEA/DSV
15h25	Pause café - Visite de l'exposition - Networking - Grand Hall
16h00	Voici le déluge: Outils pour l'analyse des Big Data Guy CHESNOT – SGI
16h30	Platform Symphony : une solution industrielle pour relever le défi Big Data en entreprise Philippe BRICARD & Emmanuel LECERF, IBM
17h00	The Visual Factory: Plateforme de reconnaissance visuelle massive Pascal RUBY, XEDIX