

CARRIOCAS

Lustre à 40Gb/s

Pierre-François Honoré

CEA - DSM/Irfu

Plan

i r f u



saclay

Présentation du projet CARRIOCAS 2006-2009

Architecture du réseau à 40 Gb/s

Les applications déployées à l'Irfu :

- Lustre
- proof

Projet CARRIOCAS

irfu
cea
saclay

Projet du pôle de compétitivité System@tic Paris-Région
Calcul Réparti sur Réseau Internet Optique
à Capacité Surmultipliée



Carriocas : le contexte

Forte croissance des besoins de calcul et stockage pour la modélisation et l'analyse de phénomènes complexes et la conception numérique de systèmes complexes



Nécessaire mutualisation des ressources informatiques



Leurs utilisations à distance imposent de fortes exigences sur les performances des réseaux d'interconnexion

Carriocas et GRIF

irfu
cea
saclay

Projet GRIF : fédère les activités des 6 laboratoires parisiens IN2P3 et Irfu impliqués dans LCG/NGI

irfu
cea
saclay



1^{er} démonstrateur de stockage distribué Lustre

Groupe travail Lustre

- Création d'un espace de stockage commun
- Intégration à l'intergiciel de grille

Carriocas : les objectifs

irfu

cea

saclay

Mettre au point des systèmes de transmission à 40 Gb/s compétitifs sur un plan technicoéconomique

Adapter architecture réseau, gestion, protocoles et algorithmes aux besoins des applications distribuées

Tester et valider les approches sur une infrastructure expérimentale avec différentes applications métiers

Implémenter des applications hautes performances

- Visualisation haute résolution interactive et collaborative
- Stockage distribué de volumes massifs de données

Carriocas : organisation

irfu

cea

saclay

LAL : R. Bilhaut, M. Jouvin, G. Philippon

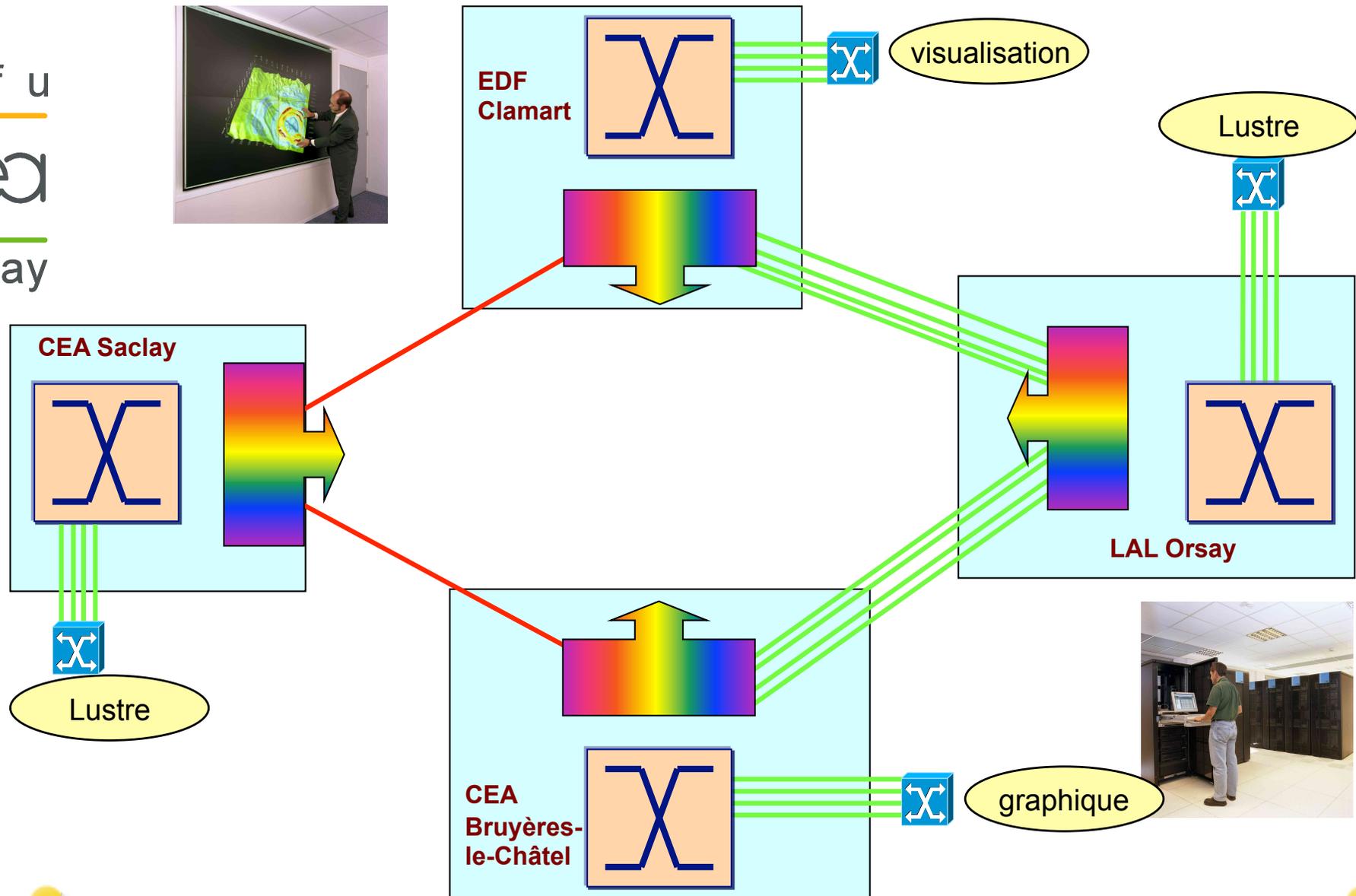
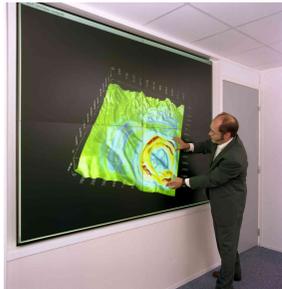
Irfu : M. Airaj, Z. Georgette, A. Gomes, P-F Honoré, G. Herent, J-P. Meyer, M. Winkler

1. Ingénierie des systèmes de transmission à 40 Gbit/s
pilote Alcatel-Lucent
2. R&D de l'architecture et des protocoles de réseaux
pilote Alcatel-Lucent
3. Démonstrateur réseau 40Gbit/s
pilote France Telecom
4. R&D des applications de calcul distribué
pilote CEA



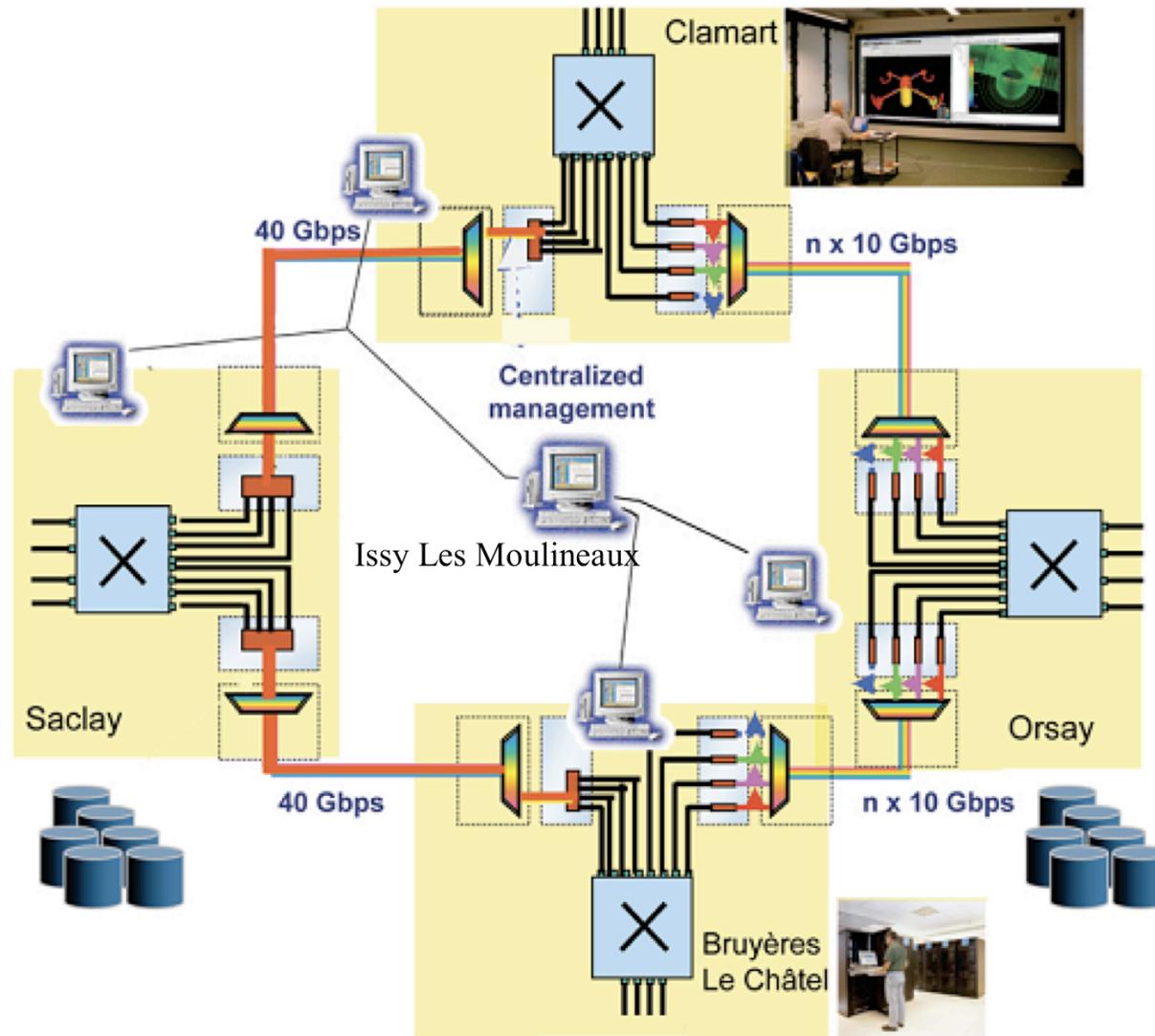
Réseau : schéma de la phase II

irfu
cea
saclay



Réseau : optique

irfu
cea
saclay



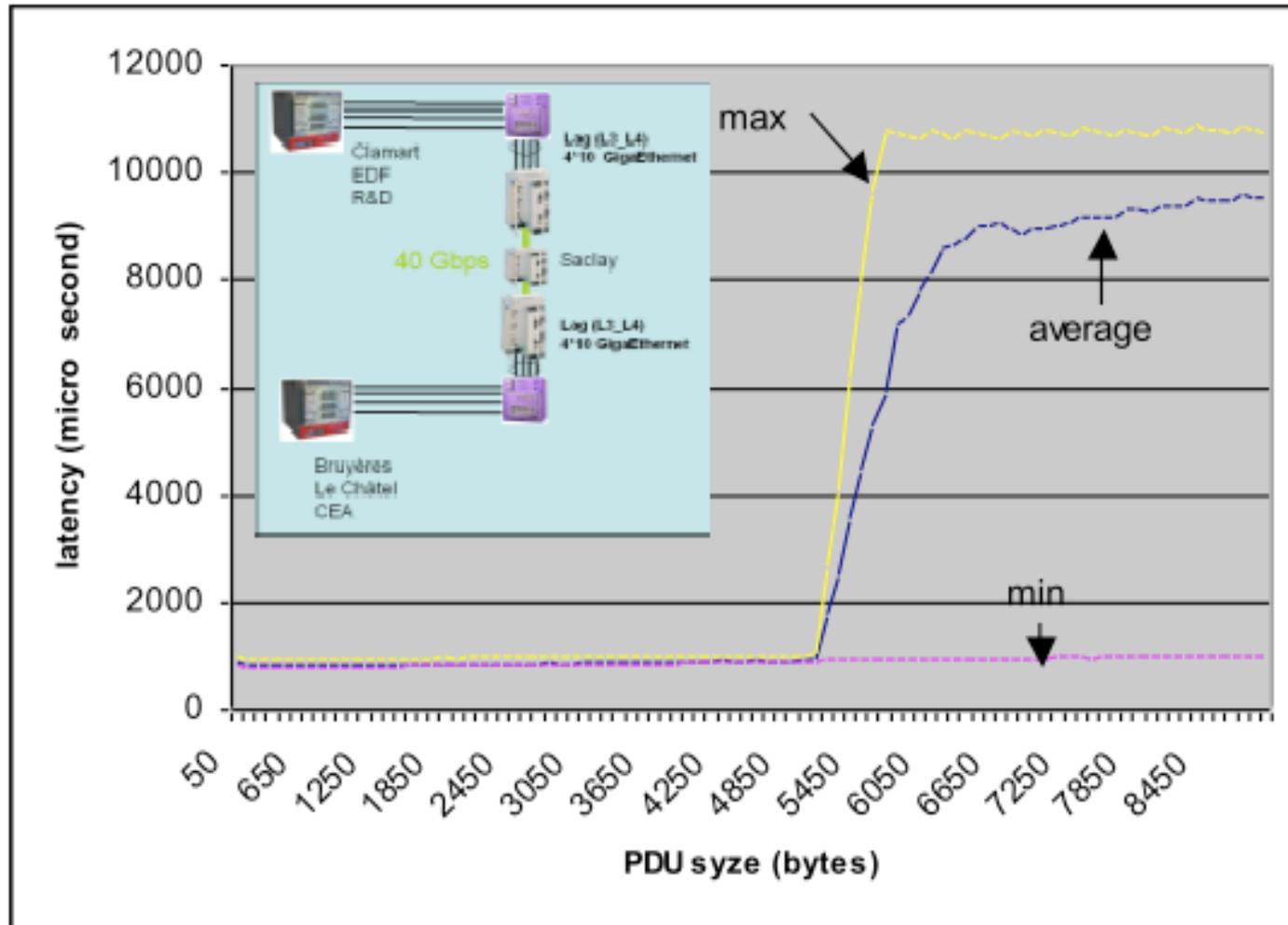
Réseau : latences

Mesure de latence sous forte charge (> 96 %) révèlent des effets dus à l'encapsulation Ethernet - SDH

irfu

cea

saclay

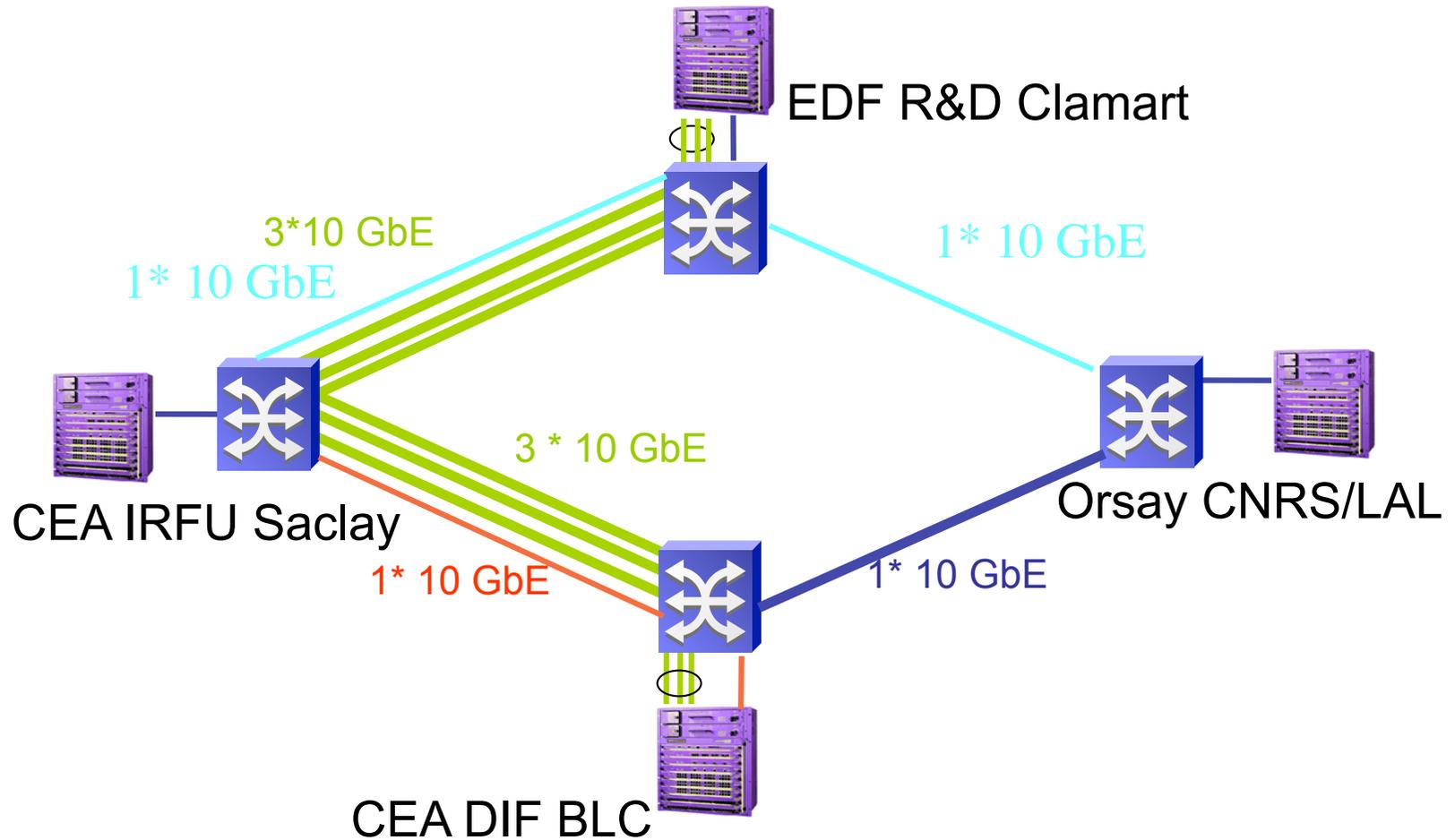


Réseau : EPlines

irfu

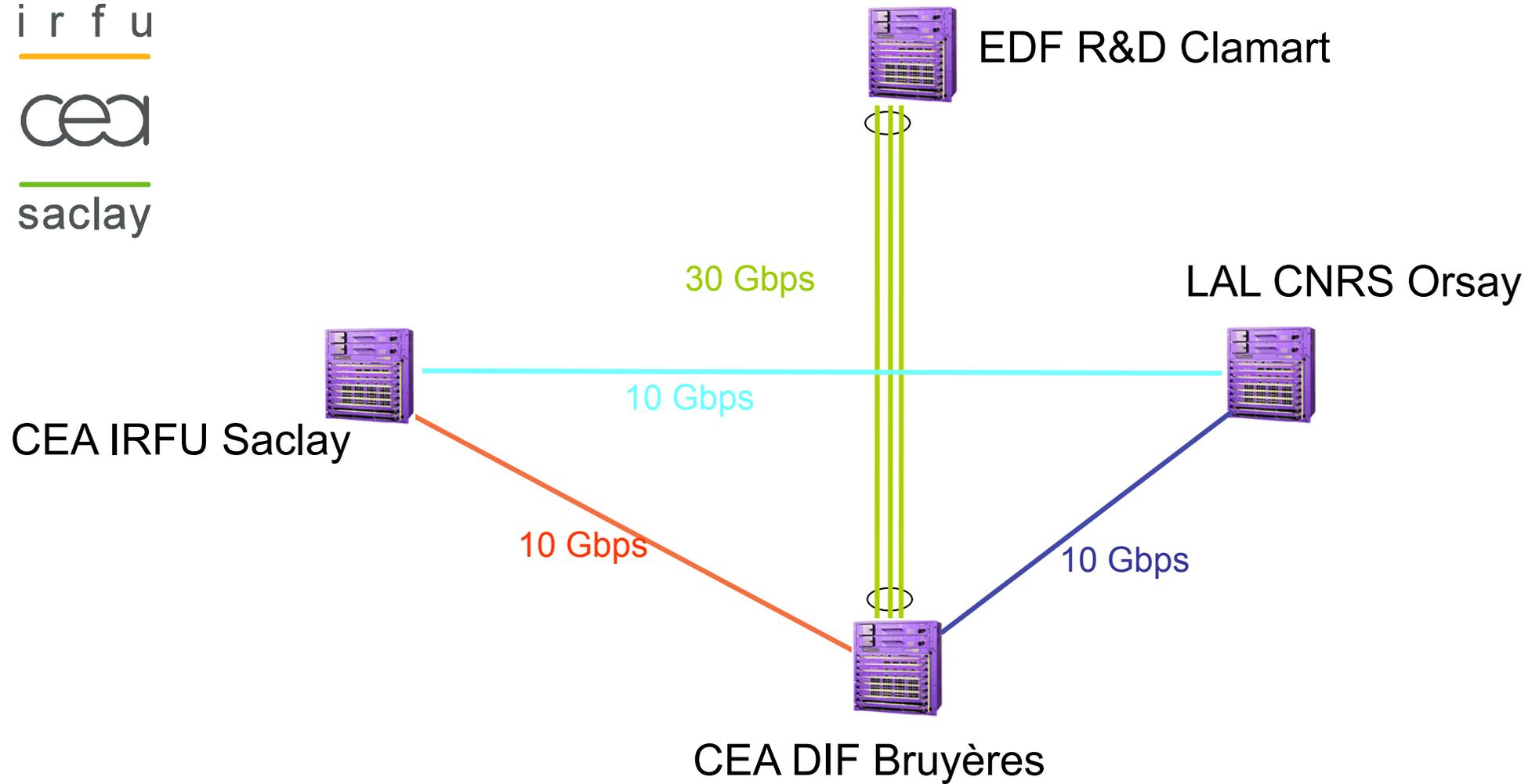
cea

saclay



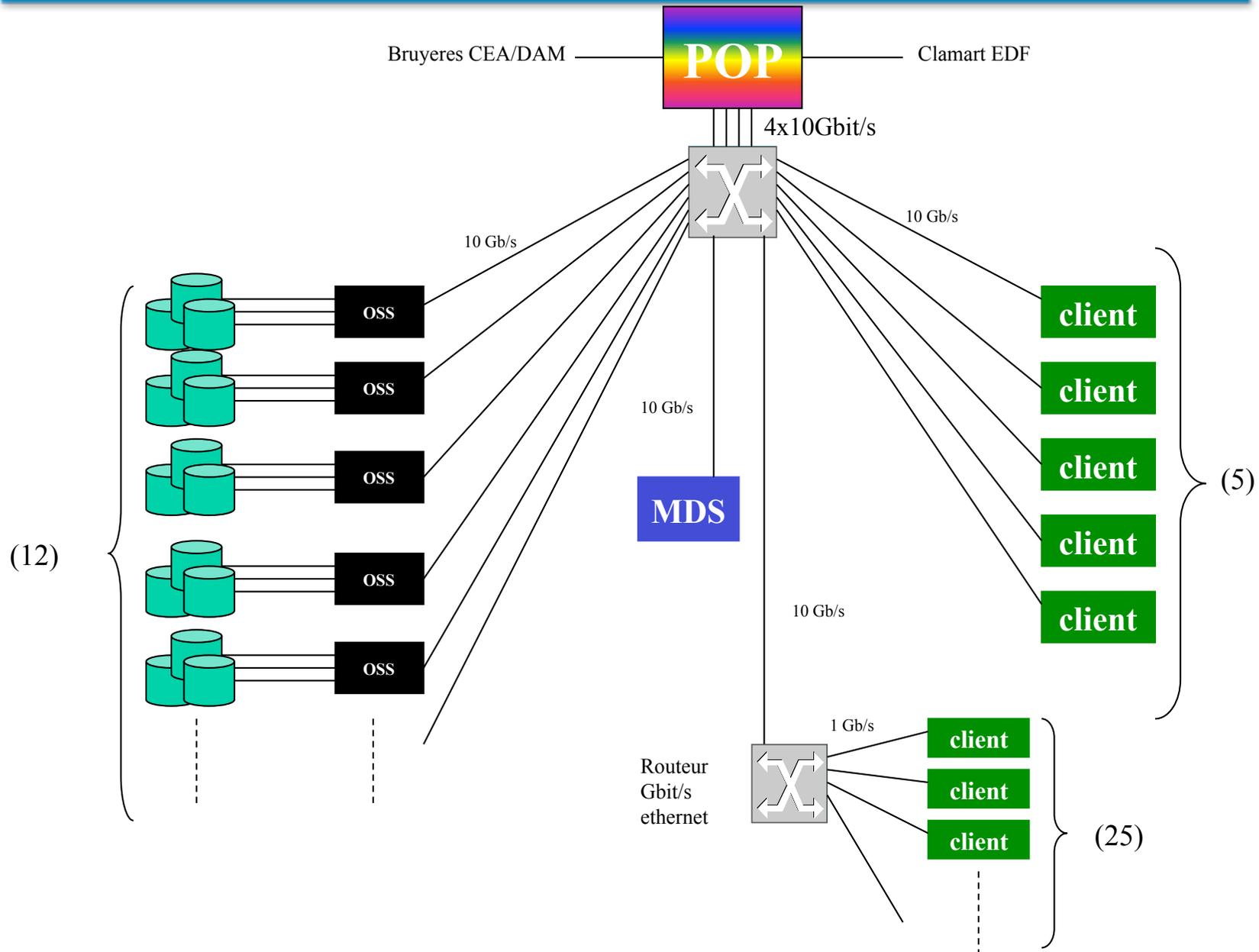
Réseau : vue des applications

irfu
cea
saclay



Lustre – cluster de l'Irfu

irfu
cea
saclay

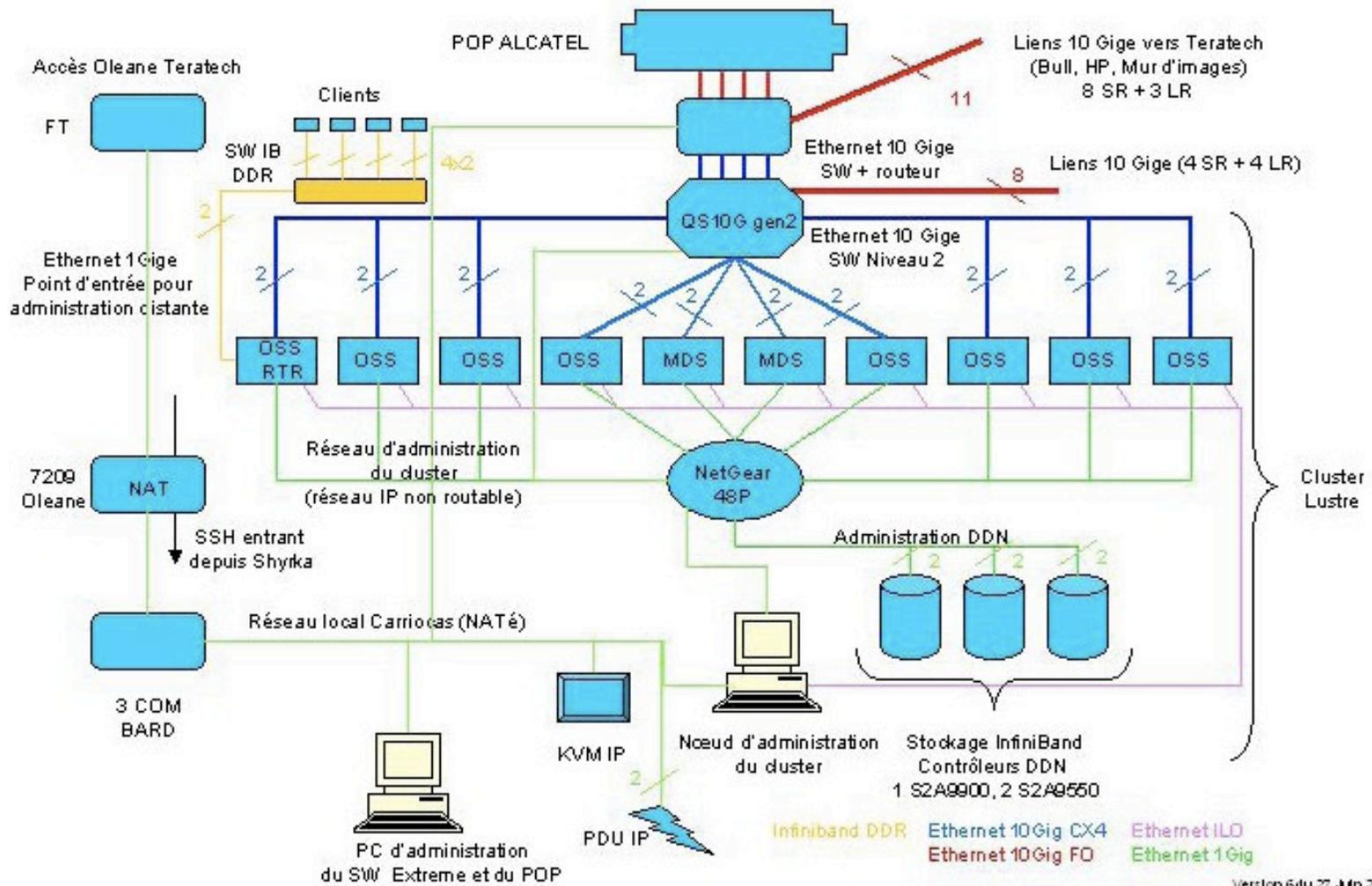


Lustre – cluster de la DIF

irfu

cea

saclay



Irfu – réseau 10Gb/s local

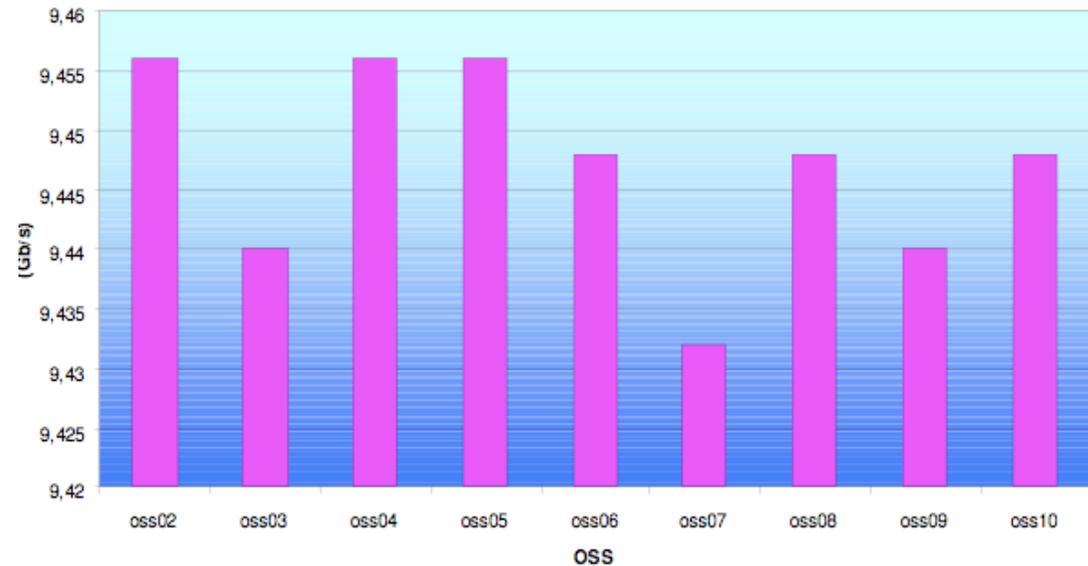
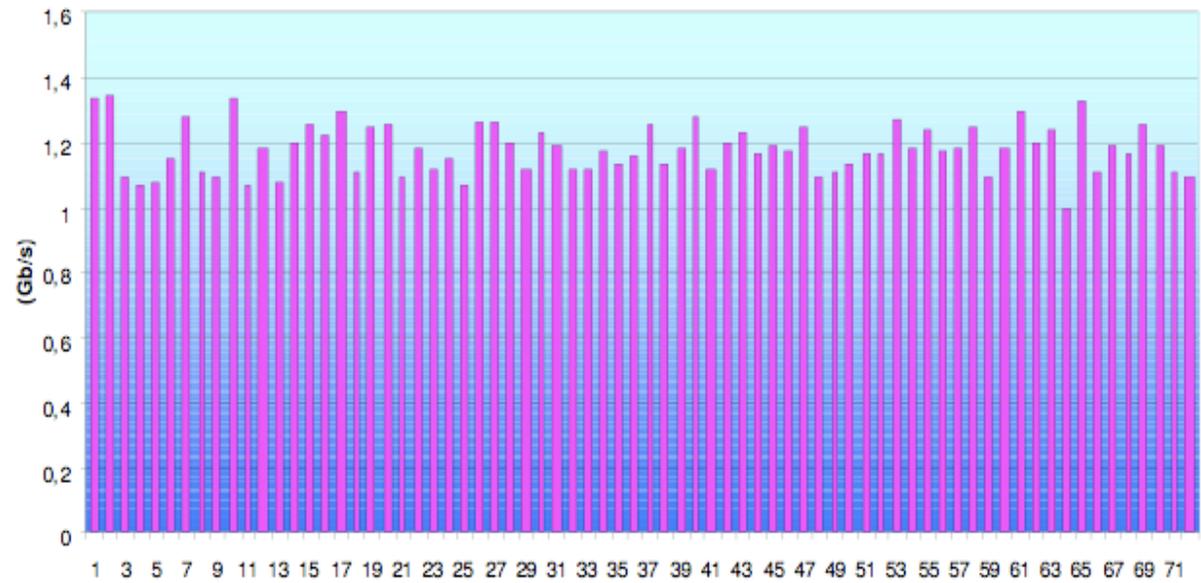
i r f u



saclay

Transfert TCP entre
9 nœuds
simultanément avec
Iperf

Débits par nœud :
94% de la bande passante



Lustre - Performances

J-C. Lafoucrière (CEA/DAM/DIF)

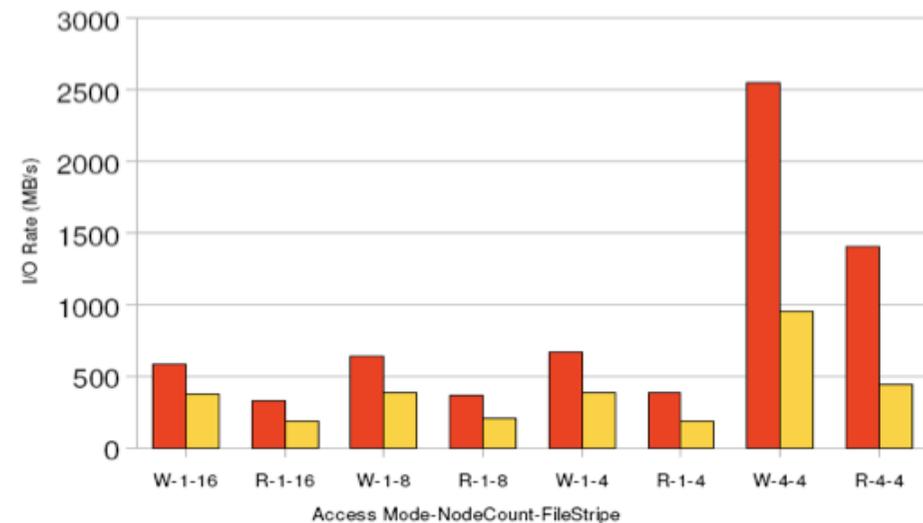
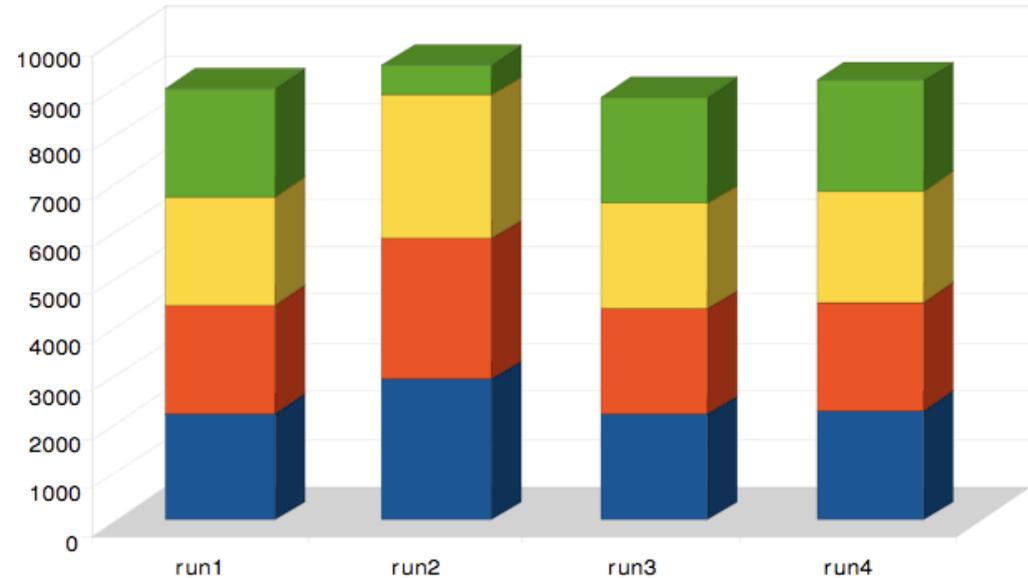
irfu

cea

saclay

Performances réseau
4 noeuds à Bruyères et
4 noeuds à Saclay

Lecture écriture Lustre
WAN
entre Bruyères (distant) et
Saclay (local)



Applications – analyse root

irfu



saclay

Analyse de données simulées de Atlas

	Durée(s)	Charge CPU (%)	Bande passante (MB/s)
Disque local	1222	91,4	3,11
Lustre local Irfu	1209	93,1	3,11
Lustre distant Bruyère	1207	91,8	3,14

Calculs limités par le CPU

Applications - proof

irfu
cea
saclay



Nœuds de calcul



Déploiement d'une
ferme **PROOF**
Parallel ROOT Facility
à Saclay

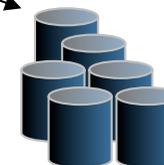
Réseau local
10Gb/s



*Lustre
Saclay*



Réseau WAN CARRIOCAS
10Gbits dédiés LUSTRE



*Lustre
Bruyères le Chatel*

irfu

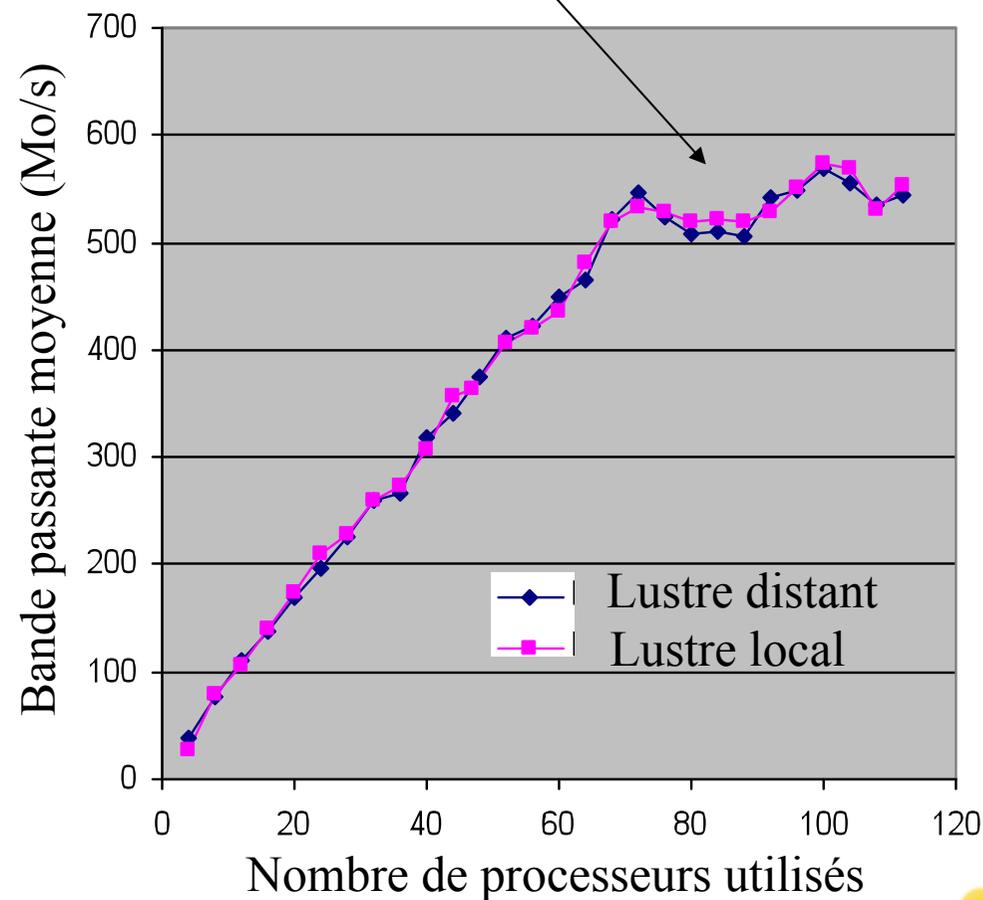


saclay

Des données simulées du LHC sont stockées sur le Lustre local (Irfu) à Saclay et sur le Lustre distant (DIF) à Bruyères

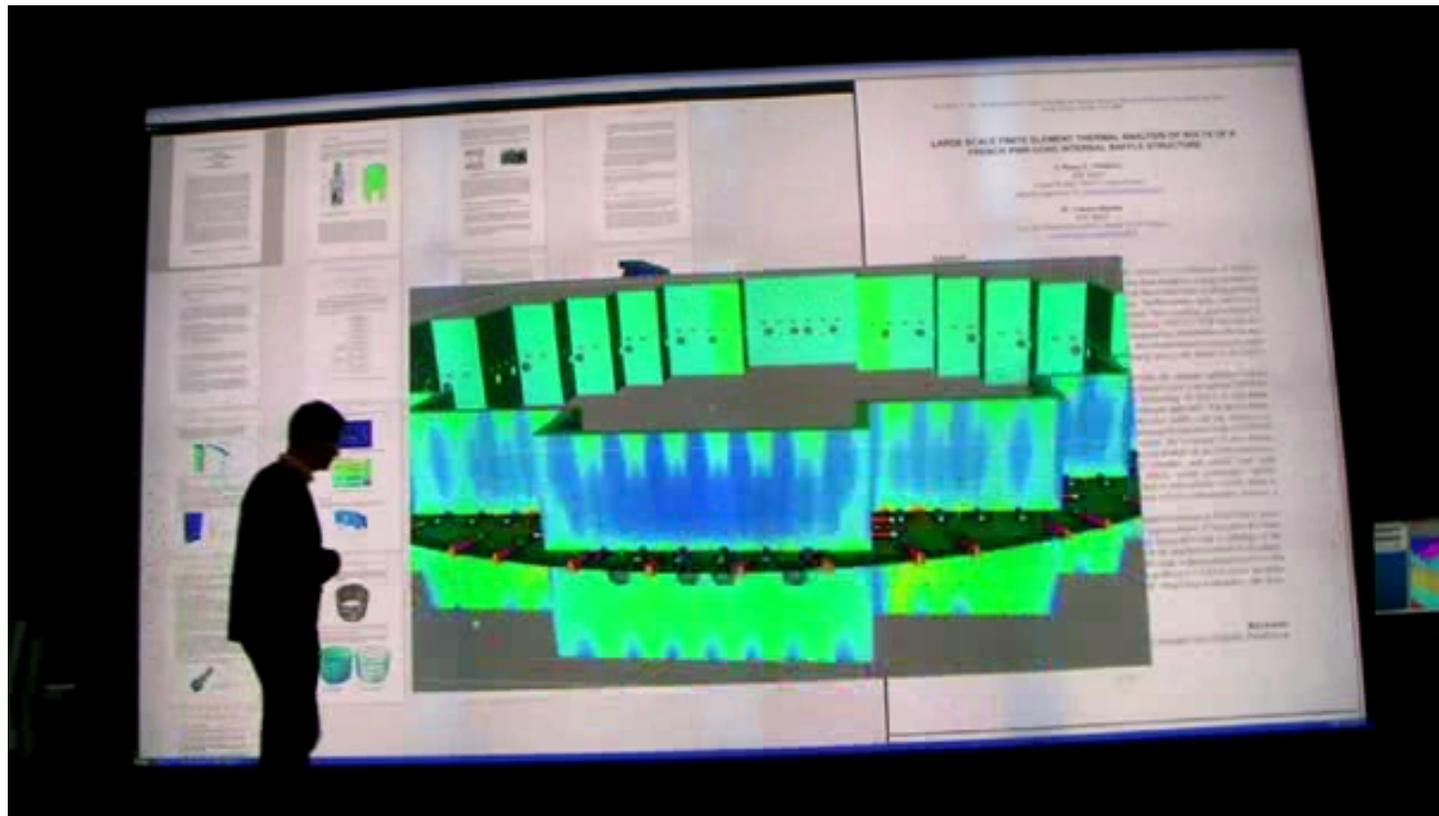
**Validation applicative :
transparence observée du
réseau CARRIOCAS**

Saturation observée due au trafic maître-esclave interne à PROOF



Applications - visualisation

Un mur stéréoscopique, de 6m x 3m et 24M pixels, installé à EDF Clamart permet une visualisation interactive d'images générées à distance ou localement à partir de données du Lustre global



Conclusions

irfu



saclay

Carriocas

Démonstrateur d'un réseau à 40 Gb/s performant, stable, configurable : premières ventes dès 2009

Evolution vers offre Cloud computing ...

Applications

Exploitation déportée d'un cluster graphique : visuportal

Transparence observée de Lustre y compris à haut débit avec des flux multiples