



Grid'5000

Principes et utilisations

Olivier Richard

Labo ID-IMAG / Equipe INRIA MESCAL



Laboratoire
Informatique et
Distribution



CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE



Institut National
Polytechnique
de Grenoble



INSTITUT NATIONAL
DE RECHERCHE EN
INFORMATIQUE ET
EN AUTOMATIQUE



GRENOBLE 1
UNIVERSITÉ
JOSEPH FOURIER
SCIENCES, TECHNOLOGIE, MÉDECINE

Objectif: Etendre le champ du possible

- ▶ Computing As You Want !
 - ▶ Cluster in Cluster
 - ▶ Grid in Cluster
 - ▶ Grid in Grid
 - ▶ Cluster in Grid

Utilisations

- ▶ Traditionnelle
 - ▶ Réservation
 - ▶ Jeu
 - ▶ Libération

- ▶ Avancée
 - ▶ Réservation
 - ▶ Déploiement
 - ▶ *Activation*
 - ▶ Jeu
 - ▶ Libération

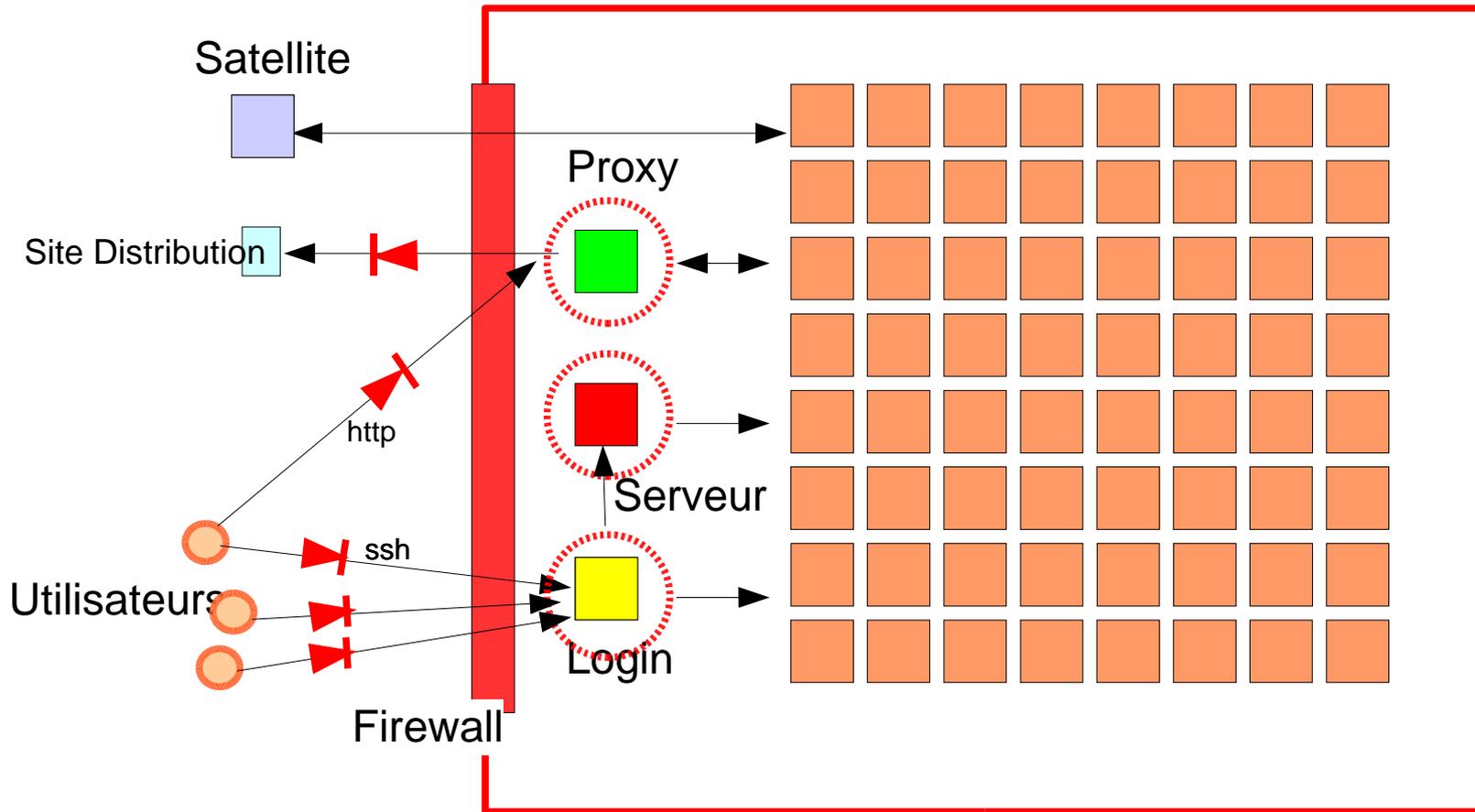
Grid'5000

- ▶ Une grille légère
 - ▶ Homogénéité sur plusieurs services
- ▶ Un réseau confiné (*Renater*)
- ▶ Opteron, Itanium2, G5, Xeon
- ▶ Mécanisme de reboot matérielle + report de console

Services par site

- ▶ 1 machine frontale
- ▶ Authentification **LDAP**
- ▶ Gestionnaire de ressources **OAR** (<http://oar.imag.fr>)
- ▶ Outil de déploiement **Kadeploy2** (<http://ka-tools.imag.fr>)
- ▶ Gestionnaire de ressources **NFS**

1 Site en mode G5K



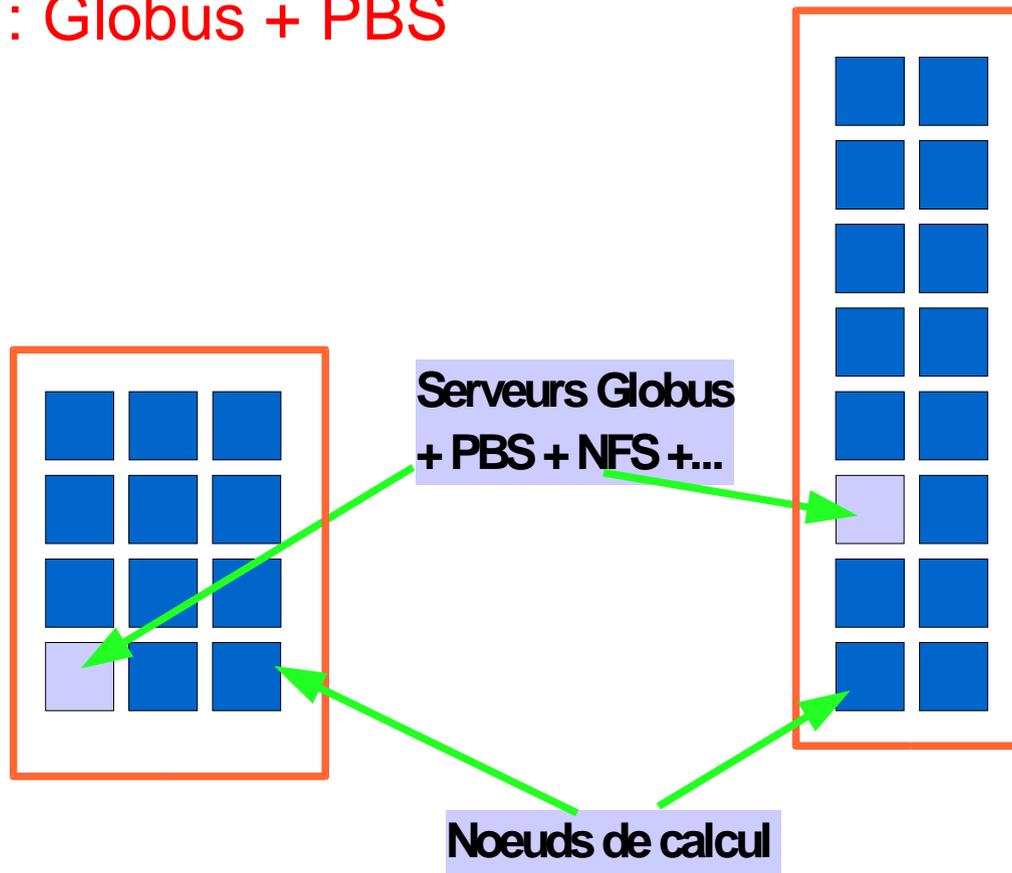
Vers les autres sites G5K

Renater G5K

Serveur : OAR, Kadeploy, NFS, Reboot matériel, Console
Proxy : Http, apt-cacher

Exemple d'une expérience

Environment : Globus + PBS



L'utilisateur dispose d'un **environnements** qui comprend **tous les services** dont il a besoin

Les commandes invoquées

Sur 1 site

```
1 oarsub -I -l nodes=10 -d deploy
2 kadeploy -e debian_g5k -f node_file -p hda6
3 ...utilisation de l'environnement
4 ^D
```

Commande accès_reboot/console matérielle

```
kaconsole -m node38-009
kareboot -h -m node38-009
```

Réservation (coallocation) sur la grille

```
oar_gridsub grenoble:nodes=100,orsay:nodes=200
```

Conclusion

- ▶ Grille pour l'expérimentation
- ▶ Grille légère et confiné
- ▶ Déploiement d'environnement

- ▶ Status
 - ▶ 1300 Proc opérationnels (3200 en 2006)
 - ▶ Mise en place du service de déploiement **En cours (25%)**
 - ▶ **Test global en Avril**
 - ▶ **Stable/opérationnelle 100%** prévu en Juin

Test **gLite** sur Grid'5000 ?

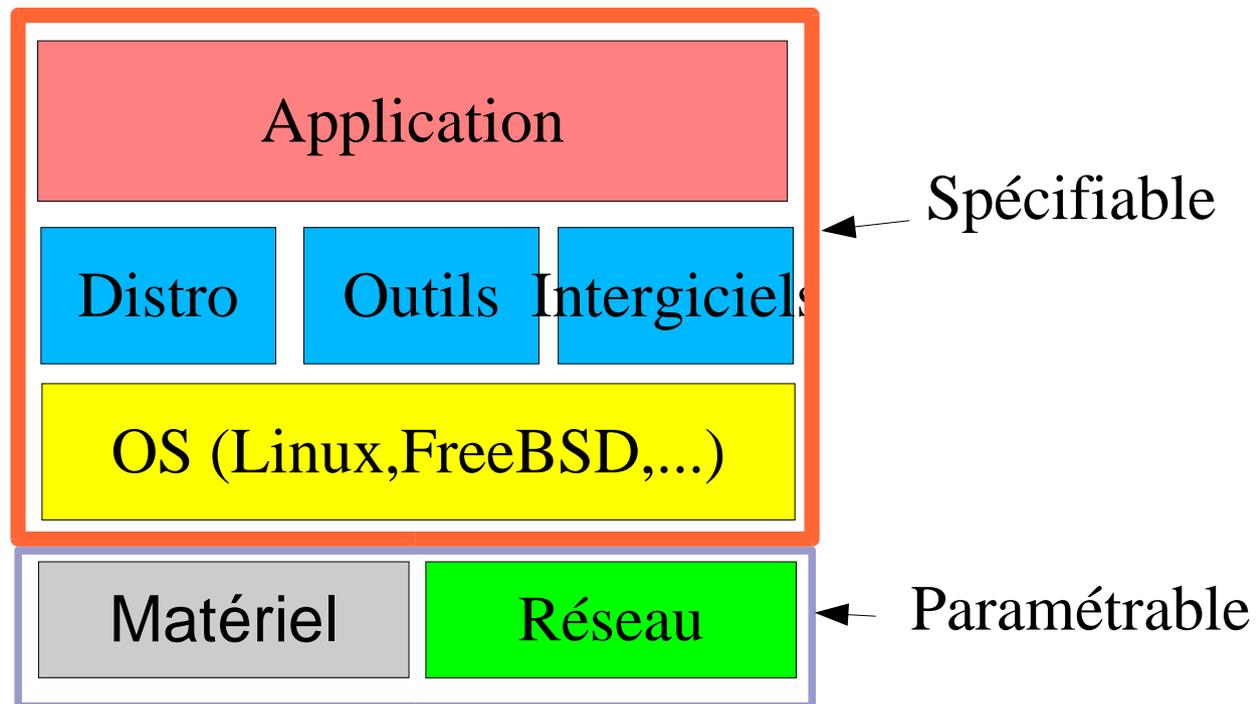
Kadeploy2

◆ Environnement à la demande (EàD)

L'expérimentateur fournit
/choisit son environnement

Exemple :

20 FreeBSD +100 Linux
Debian + softs spécifiques

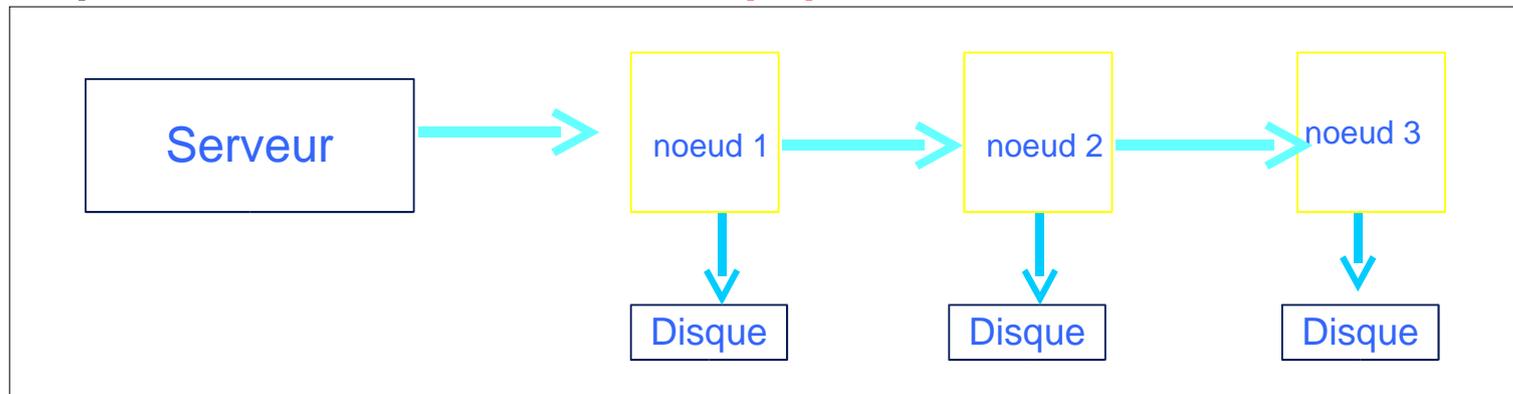


KaDeploy2

- ▶ Permettre le déploiement d'environnement
 - ▶ efficace
 - ▶ robuste
 - ▶ le moins contraignant (environnement/site)
 - ▶ simple
- ▶ Gestion des droits
 - ▶ Capacité de reboot limité
 - ▶ Déploiement autorisé sur les noeuds/partitions alloués
- ▶ Intégration avec les gestionnaires de ressource
 - ▶ OAR ou autre (epilogue/prologue sur serveur kadepoy)

Principes

- ▶ Utilisation : boot réseau/PXE + DHCP+ GRUB (ELILO)
- ▶ Déploiement efficace : **pipeline**



- ▶ **Pas de modification** du MBR (secteur de boot disque)
- ▶ Utilisation du **reboot matériel** (retour au système de référence + reprise après crash)
- ▶ Utilisation du **report de console** pour observer le déroulement des déploiements.
- ▶ **Utilisation d'une BD** (MySQL): état, config, description environ.

Contraintes site / environnement

▶ Environnement

- ▶ Architecture matérielle (processeurs 32bit/64bit)
- ▶ Drivers => auto-detection
- ▶ DHCP, ssh
- ▶ Console sur ttyS0 (pratique :)
- ▶ Pas d'état dépendant de la localisation

▶ Site

- ▶ Serveur DHCP
- ▶ Schéma de partition
- ▶ Clé Root
- ▶ Serveur NFS, LDAP, NTPdate (optionnels)

▶ Dépendances communes : la post-installation

- ▶ Dépend de la distribution (faiblement)
- ▶ En pratique : le site propose des scripts [+ adaptations]

Construction d'environnement

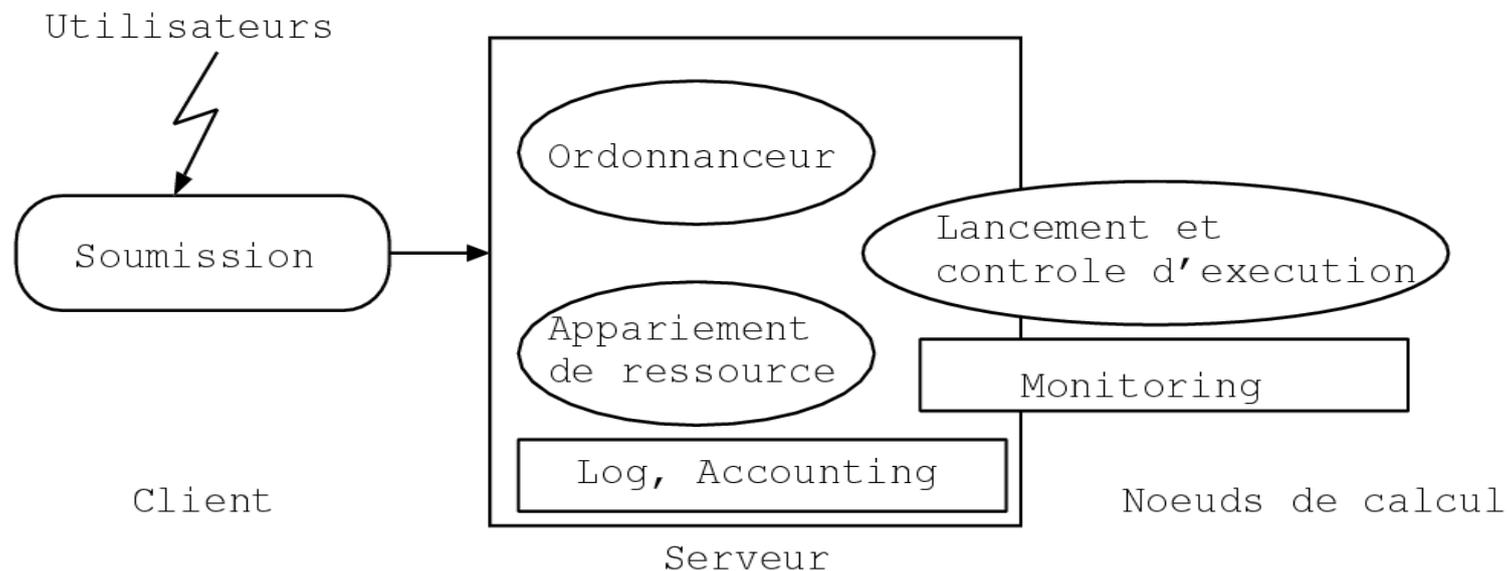
```
1 manip utilisateurs (install, config, patch)
2 tar -czf debian_test.tgz /
3 cp debian_test.tgz /en_lieu_sur/
4 karecordenv -fe envi_test.dsc -n debian_test
```

Contenu de *debian_test.dsc*

```
#Description d'environnement
#name=debian_test
author = auguste@imag.fr
file_env = file:/en_lieu_sur/debian_test.tgz
file_post = file://home/site/postinstall.tgz
kernel = /boot/vmlinuz
initrd = initrd.img
#kernel_param = 3
```

OAR : Gestionnaire de ressource

- ▶ PBS Pro => Payant , PBS (Torque) = pas stable
- ▶ OAR :
 - ▶ Architecture : **medium => BD + Langage Script (Perl) + Taktuk**



Node Status

Monika: icluster2 OAR nodes - Mozilla Firefox

cgi-bin/monika.cgi

Monika: icluster2 OAR nodes | GanttChart | Monika: icluster2 OAR nodes

icluster2 OAR nodes

Summary:

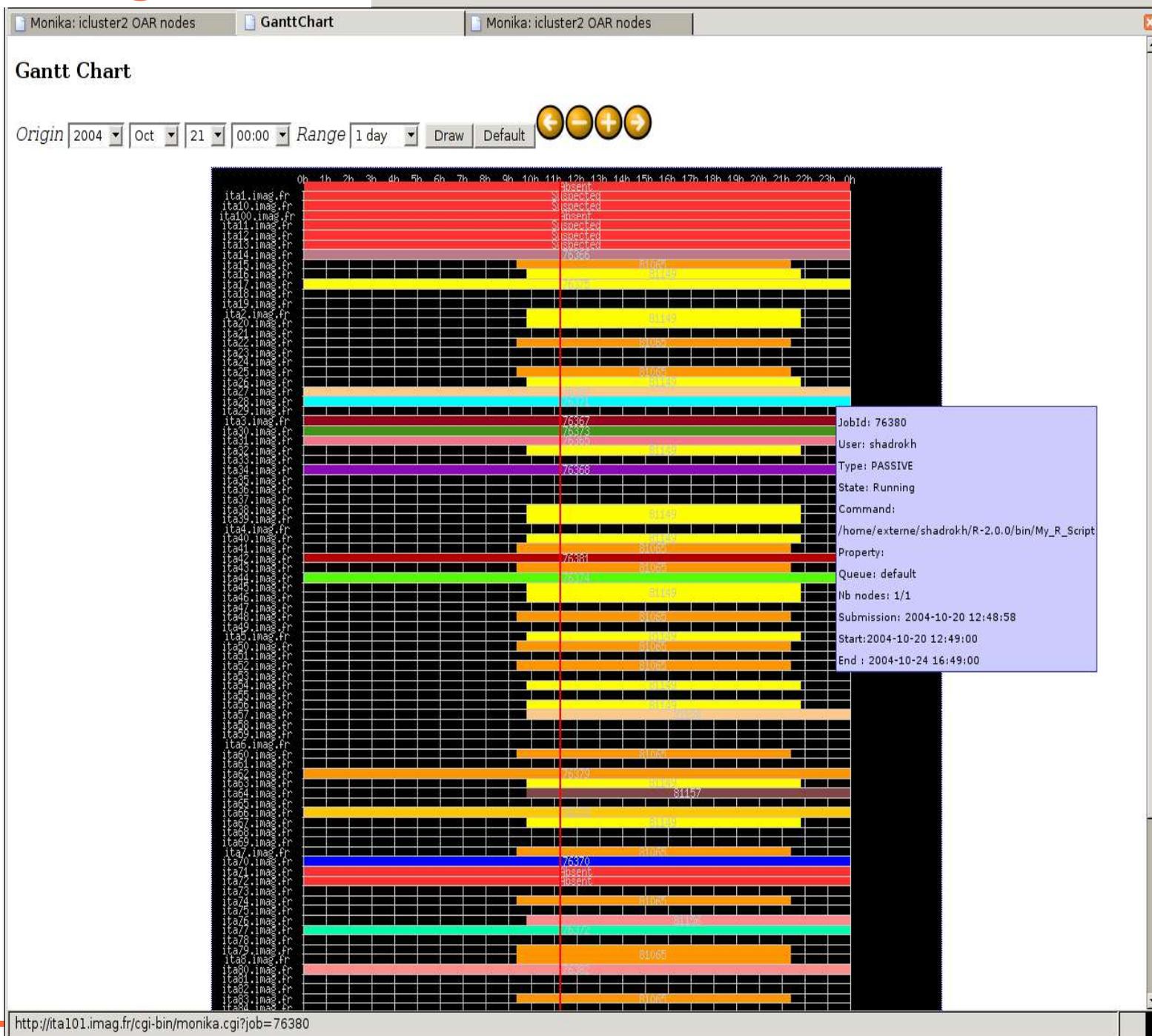
OAR node status	Free	Busy	Total
Nodes	1	87	100

Reservations:

ita1.imaq.fr	Suspected	ita2.imaq.fr	81149	ita3.imaq.fr	76367	ita4.imaq.fr	81298	ita5.imaq.fr	81149
ita6.imaq.fr	81297	ita7.imaq.fr	81065	ita8.imaq.fr	81065	ita9.imaq.fr	81300	ita10.imaq.fr	Suspected
ita11.imaq.fr	Suspected	ita12.imaq.fr	Suspected	ita13.imaq.fr	Suspected	ita14.imaq.fr	76366	ita15.imaq.fr	81065
ita16.imaq.fr	81149	ita17.imaq.fr	76375	ita18.imaq.fr	81293	ita19.imaq.fr	81291	ita20.imaq.fr	81149
ita21.imaq.fr	81299	ita22.imaq.fr	81065	ita23.imaq.fr	81295	ita24.imaq.fr	81289	ita25.imaq.fr	81065
ita26.imaq.fr	81149	ita27.imaq.fr	76380	ita28.imaq.fr	76371	ita29.imaq.fr	81290	ita30.imaq.fr	76373
ita31.imaq.fr	76365	ita32.imaq.fr	81149	ita33.imaq.fr	81303	ita34.imaq.fr	76368	ita35.imaq.fr	81273
ita36.imaq.fr	81280	ita37.imaq.fr	81302	ita38.imaq.fr	81149	ita39.imaq.fr	81149	ita40.imaq.fr	81149
ita41.imaq.fr	81065	ita42.imaq.fr	76381	ita43.imaq.fr	81065	ita44.imaq.fr	76374	ita45.imaq.fr	81149
ita46.imaq.fr	81149	ita47.imaq.fr	81292	ita48.imaq.fr	81065	ita49.imaq.fr	Free	ita50.imaq.fr	81065
ita51.imaq.fr	81305	ita52.imaq.fr	81065	ita53.imaq.fr	81284	ita54.imaq.fr	81149	ita55.imaq.fr	81282
ita56.imaq.fr	81149	ita57.imaq.fr	81154	ita58.imaq.fr	81278	ita59.imaq.fr	81283	ita60.imaq.fr	81065
ita61.imaq.fr	81281	ita62.imaq.fr	76379	ita63.imaq.fr	81149	ita64.imaq.fr	81157	ita65.imaq.fr	81287
ita66.imaq.fr	76378	ita67.imaq.fr	81149	ita68.imaq.fr	81301	ita69.imaq.fr	81288	ita70.imaq.fr	76370
ita71.imaq.fr	Absent	ita72.imaq.fr	Absent	ita73.imaq.fr	81307	ita74.imaq.fr	81065	ita75.imaq.fr	81306
ita76.imaq.fr	81156	ita77.imaq.fr	76372	ita78.imaq.fr	81304	ita79.imaq.fr	81065	ita80.imaq.fr	76382
ita81.imaq.fr	81285	ita82.imaq.fr	81296	ita83.imaq.fr	81065	ita84.imaq.fr	81279	ita85.imaq.fr	81294
ita86.imaq.fr	81155	ita87.imaq.fr	76364	ita88.imaq.fr	81286	ita89.imaq.fr	81065	ita90.imaq.fr	76363
ita91.imaq.fr	76369	ita92.imaq.fr	76377	ita93.imaq.fr	81065	ita94.imaq.fr	81149	ita95.imaq.fr	76376

<http://oar.imaq.fr>

Gantt diagram



<http://oar.imag.fr>