Silicon Graphics, Inc.

Présentation serveur Rackable

Presented by:

Christophe DAAGE
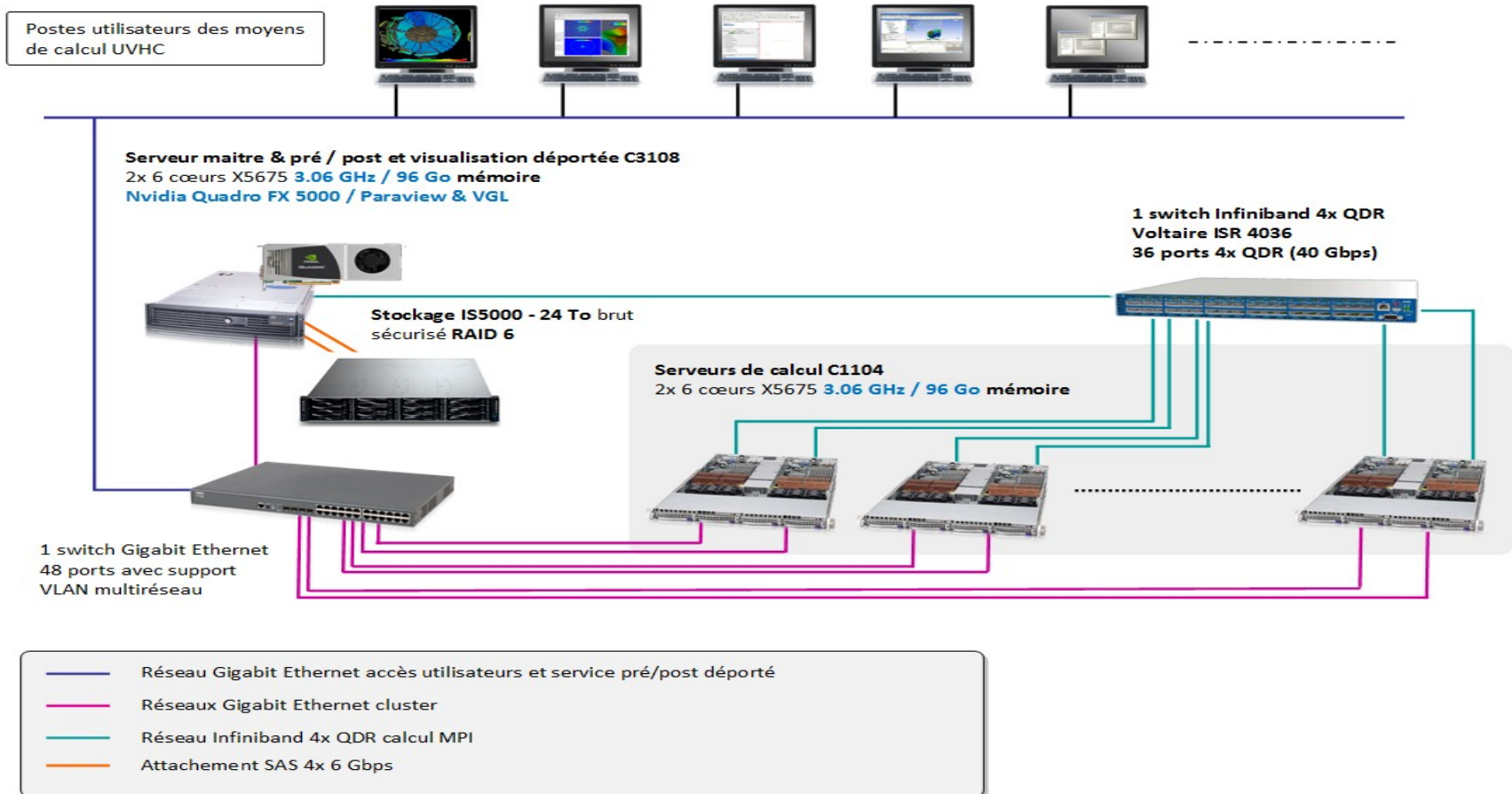
Le 10/11/2011

sgi®

1. la présentation de l'architecture hardware
2. l'utilisation de PBSpro
3. l'utilisation de module
4. l'utilisation des compilateurs
4. les lunchers (array services et mpdboot)
5. l'utilisation de mpt (sgi mpi) et intel mpi
6. les outils contenus dans SGI Foundation
7. les outils contenus dans SGI performance suite
8. les outils de monitoring

SGI
Proprietary

Présentation de l'architecture hardware

Architecture du clusteur de l'UVHC

Postes utilisateurs des moyens de calcul UVHC

Serveur maitre & pré / post et visualisation déportée C3108
2x 6 cœurs X5675 **3.06 GHz / 96 Go** mémoire
Nvidia Quadro FX 5000 / Paraview & VGL

1 switch Infiniband 4x QDR
Voltaire ISR 4036
36 ports 4x QDR (40 Gbps)

Stockage IS5000 - 24 To brut
sécurisé **RAID 6**

Serveurs de calcul C1104
2x 6 cœurs X5675 **3.06 GHz / 96 Go** mémoire

1 switch Gigabit Ethernet
48 ports avec support
VLAN multiréseau

Réseau Gigabit Ethernet accès utilisateurs et service pré/post déporté

Réseaux Gigabit Ethernet cluster

Réseau Infiniband 4x QDR calcul MPI

Attachement SAS 4x 6 Gbps

SGI

Proprietary

sgi

Altix XE Clusteur

Le clusteur repose sur la nouvelle gamme de serveurs Rackable™ C3108 et C1104 à base de processeurs Xeon multi-cœurs nouvelle génération X5675

C pour compute
Taille en nombre de U
Format (0 = half-depth, 1 = std-depth, 2 = ICE Cube optimized)
Nombre de « hotswap » disques
-
Nombre de nodes dans le chassis
Chipset architecture

Le calculateur proposé comporte un nœud maître polyvalent Rackable™ C3108 équipé de deux processeurs hexa cœurs X5675 à 3.06 GHz, 8 Go par cœur, soit 96 Go de mémoire DDR3 ECC et 1 port Infiniband 4x QDR. Ce serveur dispose d'un espace de stockage composé par une baie de disque dernière génération IS 5000 à double contrôleur RAID et attachement rapide SAS 6 Gbps qui comporte 12 disques de SAS 2000 Go de vitesse 7200 tours/min sécurisé en RAID 6 pour une capacité brute de 24 To. [la configuration retenue est de 16 To utile en RAID 6 2x (4+2)].

Le nœud maître polyvalent assure la fonction de nœud frontal et prend en charge les fonctions principales suivantes :
le point d'entrée pour la soumission des travaux de vos utilisateurs au travers du scheduler PBSPro utilisé pour les job.
le service de fichiers sécurisé vers l'ensemble des nœuds de calcul.
la fourniture du service de pré-post traitement Paraview et la fourniture de visualisation déportée
la gestion centralisée du cluster de calcul
l'administration et la distribution des images lors de l'installation de nouveaux nœuds,
le monitoring de la plateforme.

 Features
– 3U dual-socket server
– Ideal HPC management/head node
– Up to 8 x 3.5" hot-swap drives
– 18 DIMM slots for up to 144GB
memory using 8GB DIMMs
– Seven full-height expansion slots
– 1+1 Redundant AC Power Supplies
 Key Model
– C3108-TY11 (formerly Altix XE 500):
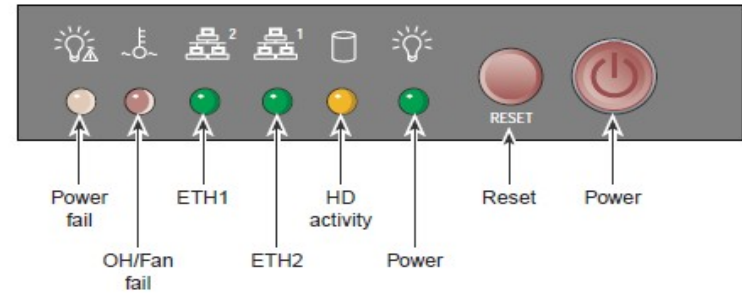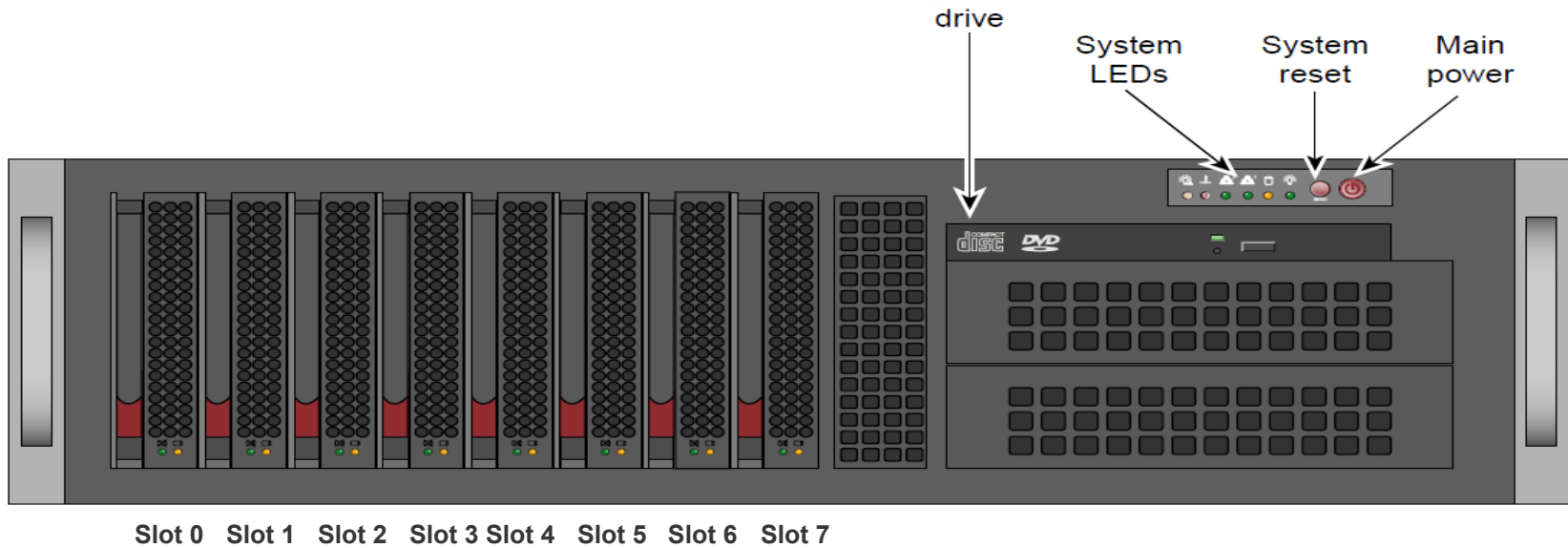Intel Xeon 5600

Altix xe500 (vue avant)

Figure 4-1   Control Panel Interface

Slot 0   Slot 1   Slot 2   Slot 3 Slot 4   Slot 5   Slot 6   Slot 7

Options PCIE sur XE500 (serveur)

GRAPHICS OPTIONS

PCIE-QUADRO-5000 NVIDIA QUADRO 5000 Graphics GPU Card

PCI OPTIONS

PCIE-SAS-9200-8E 6Gb SAS PCIE HBA with two external x4, Gen1
SFF8088 Mini-SAS connectors (low profile)

UIO-RAID-1078E PCIE 3G MEGARAID LSI SAS1078E
Supermicro H8iR UIO card: RAID 0,1,5,6,10

PCIE-CONNX2-1P Mellanox ConnectX-2 QDR, single port, Gen2 IB HCA

Altix xe500 (vue arrière)

- Two full-height PCIe UIO x16 slot (Slots 2 and 6)
- Four full-height PCIe UIO x8 slot (Slots 1, 3, 4, and 7)
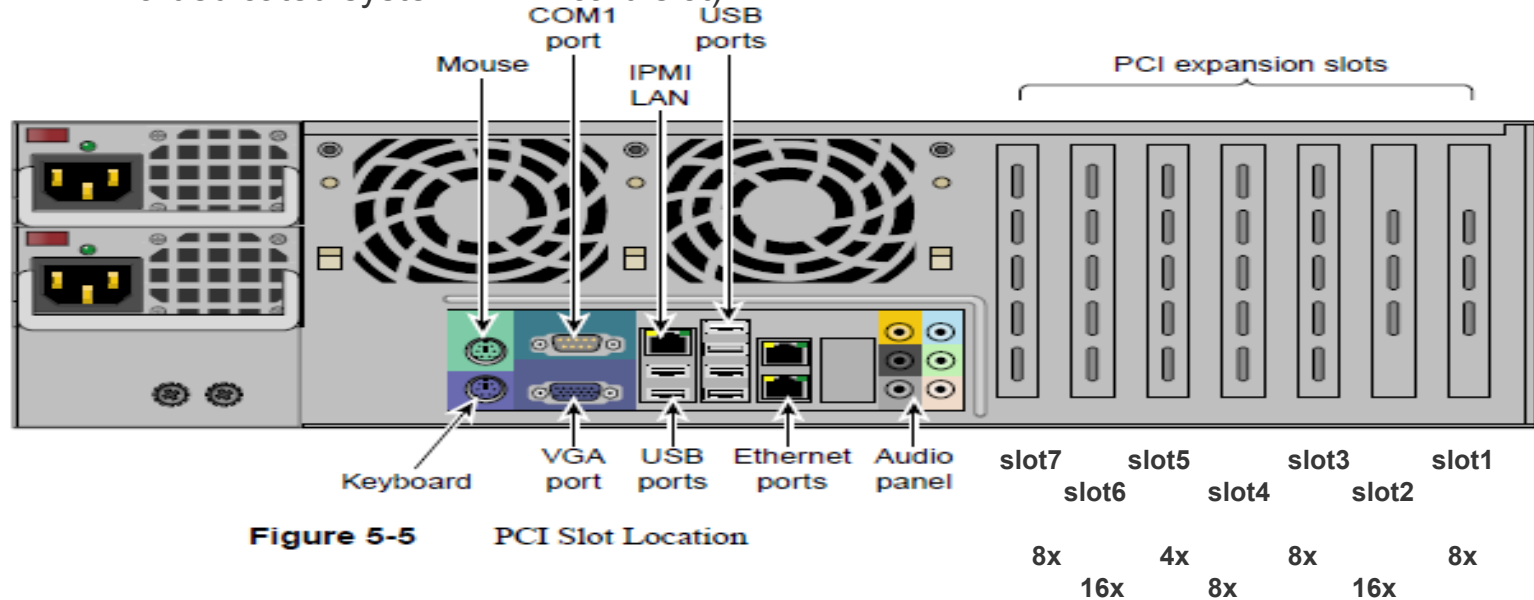- One PCIe UIO x4 slot (Slot 5 is a dedicated system RAID card slot)

Figure 5-5   PCI Slot Location
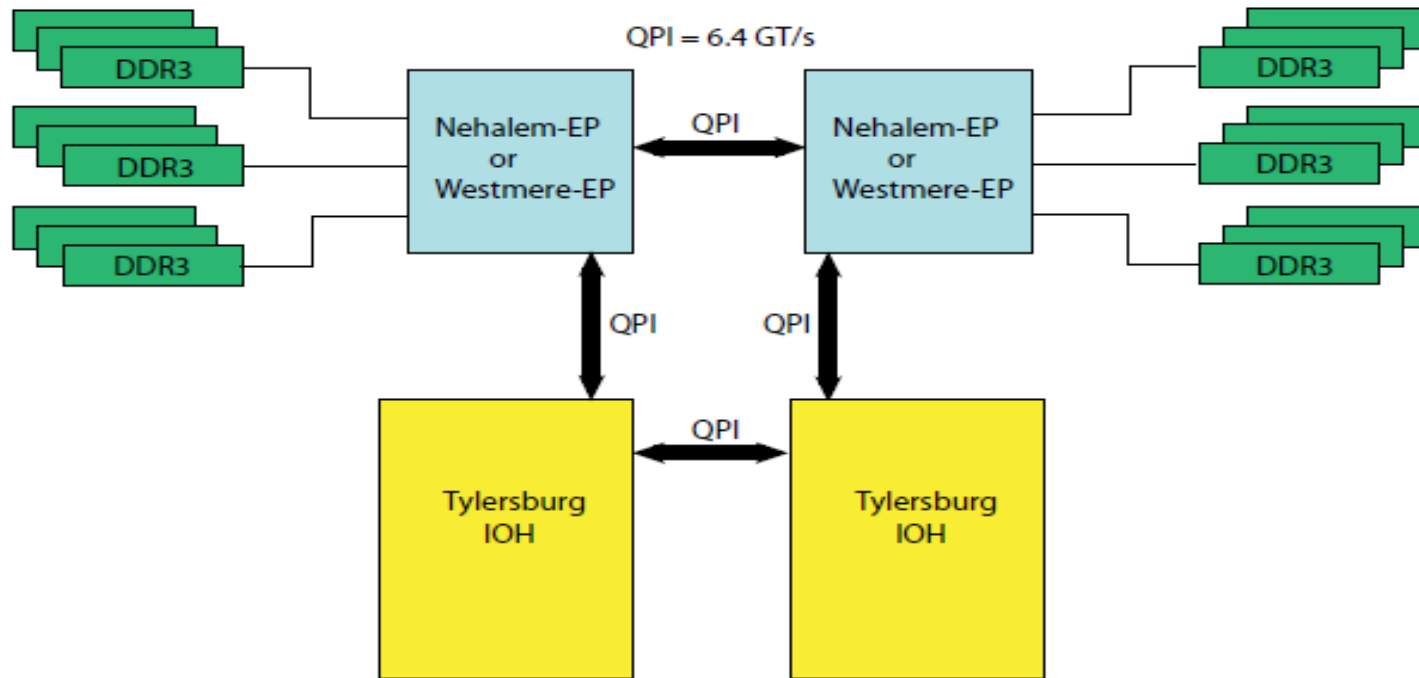
Chipset tylersburg (XE500)

QPI = 6.4 GT/s

DDR 1333
32,6GB
48GB max (par processeur)
pour fonctionner en pleine vitesse
Avec les westmeres
DIMM 8G

2 DIMMS par canal
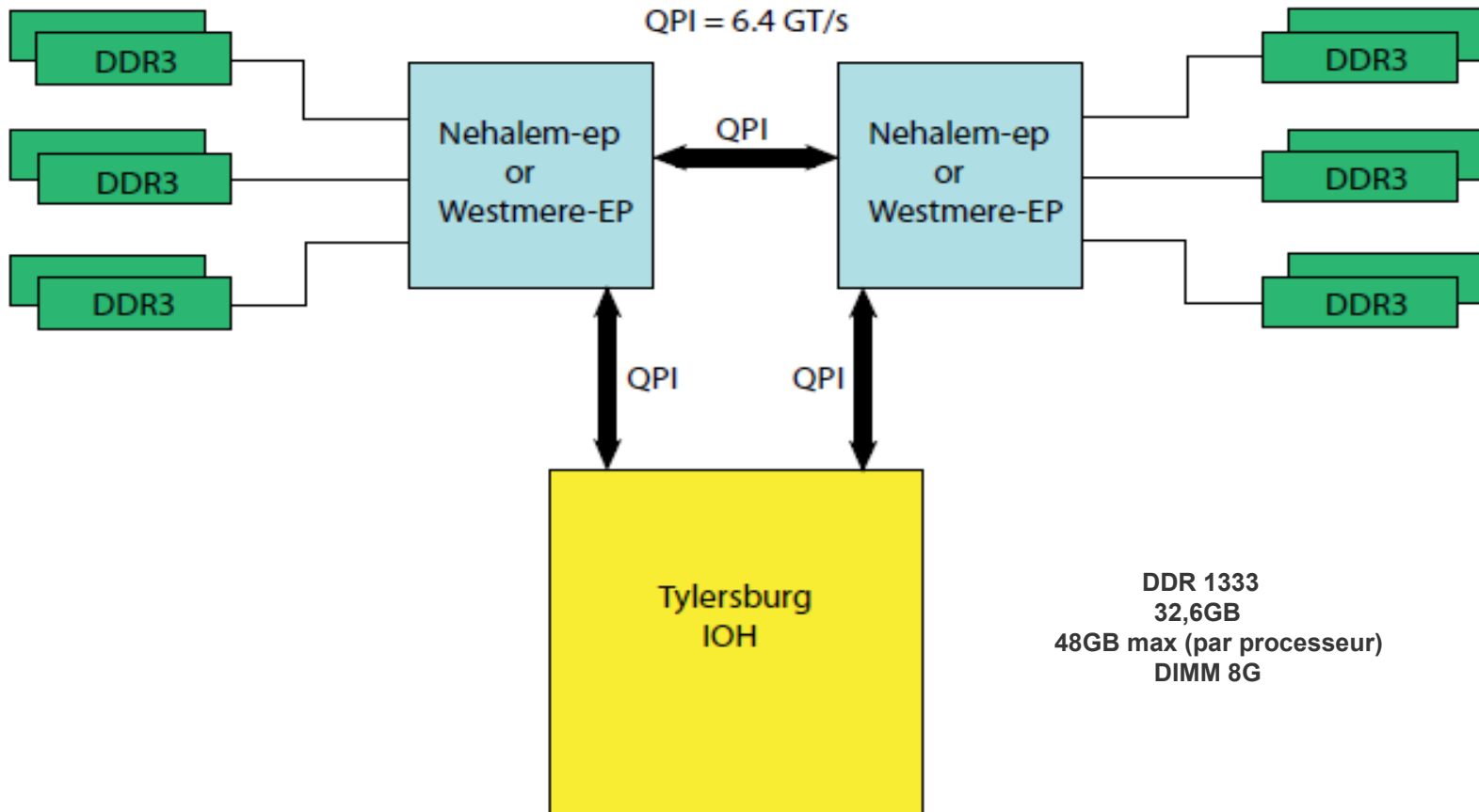
SGI
Proprietary

C1104-2TY9(Gallup3)

Features
– 1U dual-node server
– Ideal for HPC cluster compute nodes
– Up to 4 x 3.5" hot-swap drives (2/node)
– Two low-profile x16 expansion slots
(1/node)
– Optional QSFP DDR or QDR InfiniBand
port/node
 Key Model
– C1104-2TY9 (formerly Altix XE 340):
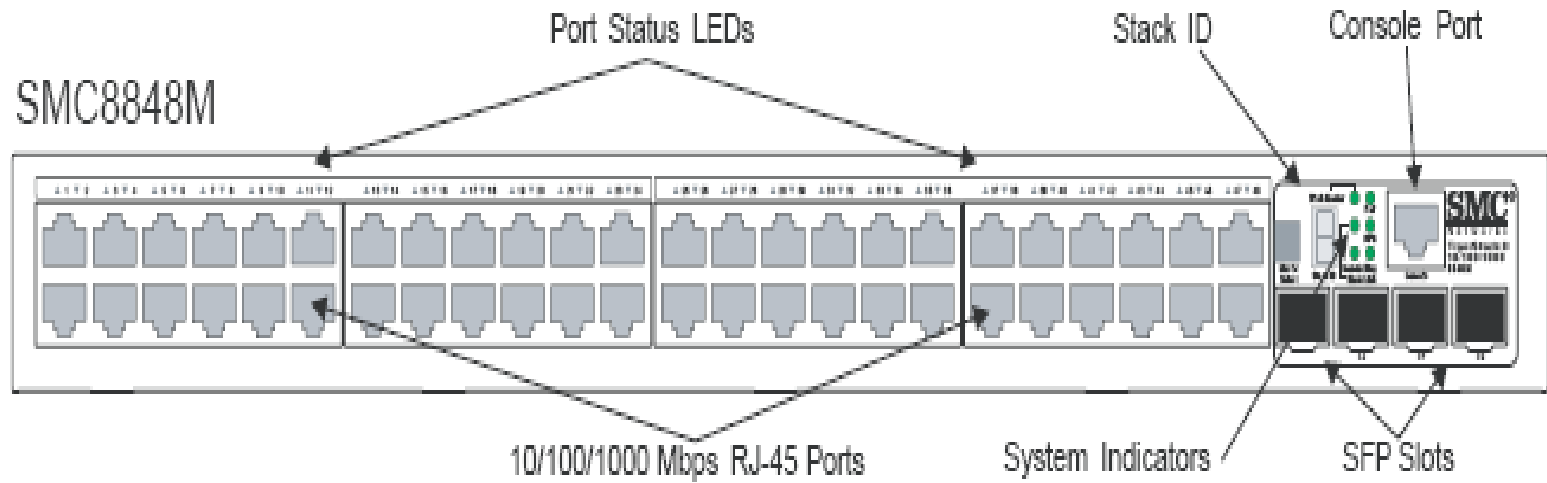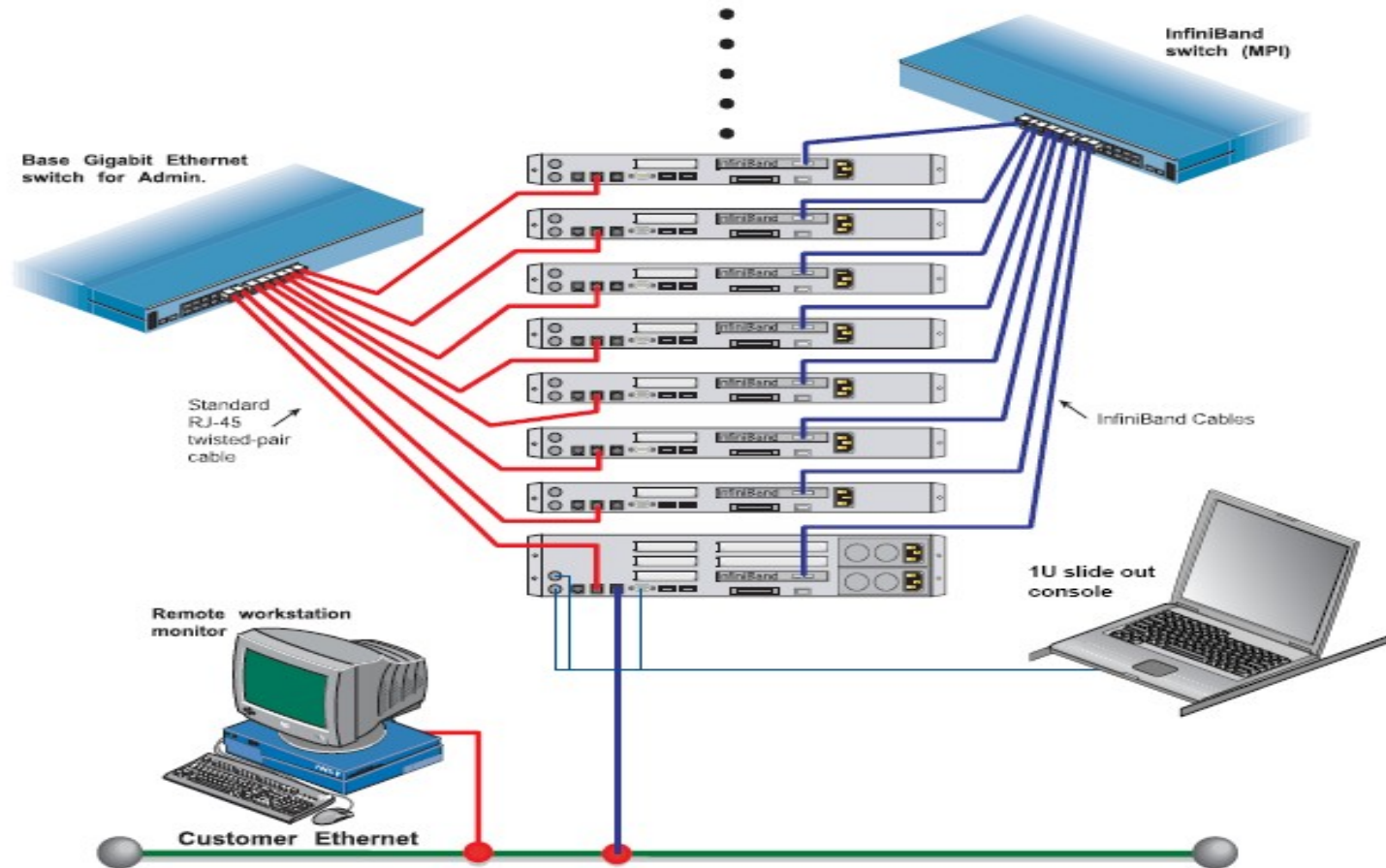Intel Xeon 5600

sgi

Chipset tylersburg (XE340)

SMC gigabit switchs 48 ports

SMC8848M

Port Status LEDs

Stack ID

Console Port

10/100/1000 Mbps RJ-45 Ports

System Indicators

SFP Slots

sgi

Voltaire 36 ports 4 X QDR inf niband

Spécificités tiroir contrôleur et disque

Solution de stockage InfiniteStorage 5000 – 12x 2 To nearline SAS 7200 tours/min	
Interface	4 ports 4x 6 Gbps (2 utilisés sur le serveur maître)
Nombre de tiroirs	1 tiroir double contrôleur + 12 emplacements disques
Niveau de RAID supportés	0, 1, 3, 5, 6, 10 avec support niveaux différents simultanément
Nombre de disques supporté	Maximum 192 disques
Type de disques disponibles	SSD, nearline SAS 7200 tours/min, SAS 10000 tours/min et SAS 15000 tours/min
Backup de cache sécurisé	Oui, sur mémoire flash
Disques fournis	12 disques nearline SAS 2000 To de vitesse de rotation 7200 tours/min
Volumétries sur disques 7200 tours/min	Exemples d'utilisation : 20 To utiles avec un volume en RAID 6 (10+2) 16 To utiles avec découpage en 2 volumes RAID 6 (exemple 4+2 et 4+2)…
Caractéristiques internes	Facteur de forme par tiroir 2 U. Alimentations redondantes, disques, ventilateurs, alimentations hot swap et extractibles à chaud.

sgi

Utilisation de PBSpro

Daemons PBS

PBSpro s'articule autour de trois daemon principaux:
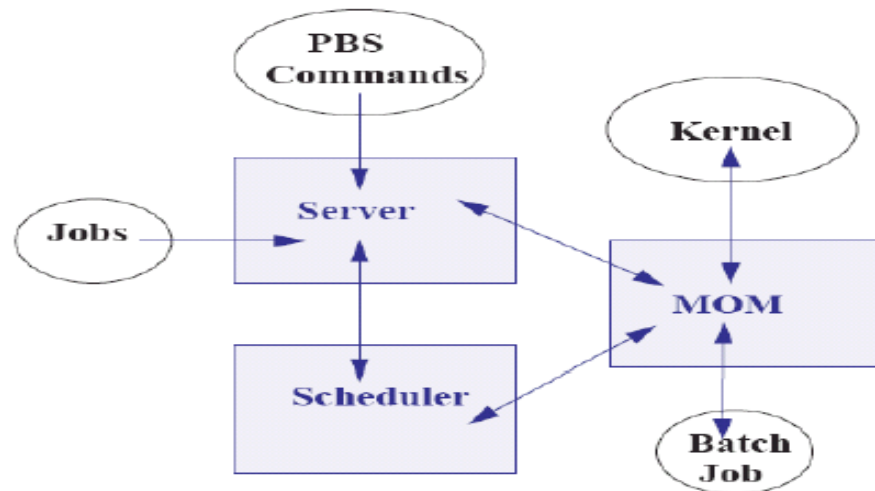
Le Serveur : executable pbs_server

Il a en charge de réceptionner les requêtes des clients, d'organiser les files d'attentes en fonction des ressources demandées et des ressources disponibles

Le MOM : executable pbs_mom

Ce processus est la « mère » de tous les jobs de calcul. Il y aura un MOM par nœud de calcul. Il exécute un job lorsqu'il recevra une copie du serveur. Le MOM crée une session de connexion du soumissionnaire (.login, .cshrc,...) et renvoie les sorties standard et d'erreur vers le serveur en fin d'exécution du job. Il communique au serveur PBS des informations concernant sa charge.

Le Scheduler : executable pbs_sched

Ce processus a en charge l'exécution des jobs en fonction de la politique mise en place sur le site, des jobs présents dans les files d'attente et des ressources disponibles sur les nœuds de calcul.

```
fortune:~ # qmgr -c "p q workq"
#
# Create queues and set their attributes.
#
#
# Create and define queue workq
#
create queue workq
set queue workq queue_type = Execution
set queue workq enabled = True  (autorise les jobs à entrer)
set queue workq started = True (autorise les jobs à s'exécuter)
```

```
fortune:~ # qmgr -c "p q batch"
#
# Create queues and set their attributes.
#
#
# Create and define queue batch
#
create queue batch
set queue batch queue_type = Route
set queue batch route_destinations = small
set queue batch route_destinations += large
set queue batch route_destinations += medium
set queue batch enabled = True
set queue batch started = True
```

SGI

Queue small

queue small

resources_max.pmem = 500mb
resources_min.cput = 00:00:01
resources_max.cput = 00:02:00

queue medium

resources_max.pmem = 1000mb
resources_min.cput = 00:00:01
resources_max.cput = 00:10:00

queue large

resources_min.pmem = 1000mb
resources_max.pmem = 2000mb
resources_min.cput = 00:00:01
resources_max.cput = 10:00:06

SGI

Proprietary

Queue small

```
Qmgr: p q small
#
# Create and define queue small
#
create queue small
set queue small queue_type = Execution
set queue small Priority = 40
set queue small max_running = 1
set queue small from_route_only = True
set queue small resources_max.cput = 00:02:00
set queue small resources_max.pmem = 500mb
set queue small resources_max.pvmem = 500mb
set queue small resources_max.vmem = 500mb
set queue small resources_min.cput = 00:00:01
set queue small resources_min.pmem = 1b
set queue small resources_min.pvmem = 1b
set queue small resources_min.vmem = 1b
set queue small resources_default.cput = 00:01:40
set queue small enabled = True
set queue small started = True
```

SGI
Proprietary

Configuring Anti-express Queues

To configure an anti-express queue, do the following:
• **Create an execution queue called lowprio:**
• **By default, all new queues have a priority of zero. Make sure all queues have a value set for priority, and that lowprio has the lowest priority:**
Qmgr: create queue lowprio
Qmgr: set queue lowprio queue_type=e
Qmgr: set queue lowprio started=true
Qmgr: set queue lowprio enabled=true
Qmgr: set queue lowprio priority=10

Queue priority (express queue)

Express Queues

A queue is an express queue if its priority is greater than or equal
 to the value that defines an express queue.
This value is set in the preempt_queue_prio parameter
in PBS_HOME/sched_priv/sched_config.
The default value for preempt_queue_prio is 150.
You do not need to set by_queue or sort_queues
 to True in order to use express queues.

Configuring express queues
#
 # preempt_queue_prio
#
 #   Defines the priority value of an express queue.
 #   If a queue's priority is this or higher, this queue
 #   becomes an express_queue. All jobs in this queue will
 #   have the "express_queue" preempt priority
#
 #   NOTE: This options works with preempt_prio
#
 #   NO PRIME OPTION
preempt_queue_prio: 150

SGI
Proprietary

sgi

Principales ressources pbs

Resource	Description
arch	Architecture system. L'architecture système est défini au niveau du vnode.
cput	Quantité de temps CPU utilisée par le job pour tous les processus sur tous les vnodes
file	Taille maximum de chaque fichier créé par le job.
host	Nom de noeud d'exécution.
mem	Quantité de mémoire physique alloué au job.
mpiprocs	Nombre de processus MPI pour le chunk. Le nombre de ligne dans PBS_NODEFILE est égal à la somme des valeurs de mpiprocs pour tous les chunks requis pour le job. Pour chaque chunk demandant mpiprocs=P, le nom du noeud sera ecrit P fois dans PBS_NODEFILE
ncpus	Nombre de processeurs requis.
nice	Valeur de Nice avec laquelle sera lancé le job.
pcput	Quantité de temps CPU alloué par processus dans le job.
pmem	Quantité de mémoire physique alloué par processus dans le job.
pvmem	Quantité de mémoire virtuel alloué par processus dans le job.
vmem	Quantité de mémoire virtuel alloué par tous les processus dans le job.
walltime	Temps total (wall-clock) , durant lequel le job peut tourner (défaut: 5ans)

sgi

Resources connues par le scheduler PBS

```
cat /var/spool/PBS/sched_priv/sched_config

#
# resources
#
#     Define resource limits to be honored by PBS Pro.
#     The scheduler will not allow a job to run if the amount of assigned
#     resources exceeds the available amount.
#
#     NOTE: you need to encase the comma separated list of resources in
#           double quotes (")
#     Example:
#
#     resources: "ncpus, mem, arch"
#
#     This is ONLY schedules jobs based on available ncpus, mem, and arch
#     within the cluster. Other resources requested by the job will not
#     evaluated for availability.
#
#     NOTE: Define new resources within
#           $PBS_HOME/server_priv/resourcedef file.
#
#     NO PRIME OPTION

resources: "ncpus, mem, arch, host, vnode, netwins, aoe"
```

SGI

sgi

Resources d'un noeud

```
fortune /home/sgi> qmgr
Max open servers: 49
Qmgr: p n n001
#
# Create nodes and set their properties.
#
#
# Create and define node n001
#
create node n001 Mom=n001.default.domain
set node n001 state = free
set node n001 resources_available.arch = linux
set node n001 resources_available.host = n001
set node n001 resources_available.mem = 99197504kb
set node n001 resources_available.ncpus = 12
set node n001 resources_available.vnode = n001
set node n001 resv_enable = True
set node n001 sharing = default_shared
```

sgi

Queue medium

```
Qmgr: p q medium
#
# Create queues and set their attributes.
#
#
# Create and define queue medium
#
create queue medium
set queue medium queue_type = Execution
set queue medium max_running = 1
set queue medium resources_max.cput = 00:10:00
set queue medium resources_max.pmem = 1000mb
set queue medium resources_min.cput = 00:00:01
set queue medium resources_min.pmem = 1b
set queue medium resources_default.cput = 00:04:00
set queue medium enabled = True
set queue medium started = True
```

SGI

Proprietary

Queue large

```
Qmgr: p q large
#
# Create queues and set their attributes.
#
#
# Create and define queue large
#
create queue large
set queue large queue_type = Execution
set queue large max_running = 1
set queue large resources_max.cput = 10:00:06
set queue large resources_max.pmem = 2000mb
set queue large resources_min.cput = 00:00:01
set queue large resources_min.pmem = 1000mb
set queue large resources_default.cput = 00:04:00
set queue large enabled = True
set queue large started = True
```

Qmgr (print server)

```
#
# Set server attributes
#
set server scheduling = True
set server pnames = router
set server default_queue = workq
set server log_events = 511
set server mail_from = adm
set server query_other_jobs = True
set server resources_default.ncpus = 1
set server default_chunk.ncpus = 1
set server scheduler_iteration = 600
set server resv_enable = True
set server node_fail_requeue = 310
set server max_array_size = 10000
set server pbs_license_info = 6200@fortune
set server pbs_license_min = 1
set server pbs_license_max = 2147483647
set server pbs_license_linger_time = 3600
set server license_count = "Avail_Global:299 Avail_Local:1 Used:0 High_Use:96"
set server eligible_time_enable = False
set server max_concurrent_provision = 5
```

SGI
Proprietary

Principales directives pbs

Directive	description
#PBS –N name	Nom du Job
#PBS –q queue@server	Queue de destination
#PBS –o file_name	Stdout destination
#PBS –e file_name	Stderr destination
#PBS –j	Joindre stdout & stderr
#PBS –m abe	Email notification a: envoie un mail lorsque le job est interrompu b: envoie un mail au début de l'exécution e: envoie un mail en fin d'exécution
#PBS -M userlist	Email recipients
#PBS -h	Hold job
#PBS -a date	Exécution différé
#PBS -l	Exécution interactive
#PBS -k	Laisse les fichiers de sortie sur le noeud d'execution

Job lancé dans la queue de routage « batch »

```
fortune /home/sgi> cat pbs_loop_batch.sh
#!/bin/sh
#PBS -q batch
#PBS -l ncpus=1

/home/sgi/loop

fortune /home/sgi> qsub  pbs_loop_batch.sh
120.Fortune

fortune /home/sgi> qstat -a

fortune:
                                    Req'd  Req'd   Elap
Job ID        Username Queue    Jobname    SessID NDS TSK Memory Time  S Time
-------------- -------- -------- ---------- ------ --- --- ------ ----- - -----
120.fortune    sgi     small    pbs_loop_b  6054  1   1    --  00:01 R 00:00
```

```
fortune /home/sgi> tracejob 120

Job: 120.fortune

11/04/2011 09:58:10  L   Considering job to run
11/04/2011 09:58:10  S   enqueuing into batch, state 1 hop 1
11/04/2011 09:58:10  S   dequeuing from batch, state 1
11/04/2011 09:58:10  S   enqueuing into small, state 1 hop 1
11/04/2011 09:58:10  S   Job Queued at request of sgi@fortune, owner =
             sgi@fortune, job name = pbs_loop_batch., queue =
             small
11/04/2011 09:58:10  S   Job Run at request of Scheduler@fortune on exec_vnode
             (n001:ncpus=1)
11/04/2011 09:58:10  S   Job Modified at request of Scheduler@fortune
11/04/2011 09:58:10  L   Job run
11/04/2011 09:58:10  A   queue=batch
11/04/2011 09:58:10  A   queue=small
11/04/2011 09:58:10  A   user=sgi group=sgi jobname=pbs_loop_batch. queue=small
             ctime=1320418690 qtime=1320418690 etime=1320418690
             start=1320418690 exec_host=n001/0
             exec_vnode=(n001:ncpus=1) Resource_List.cput=00:01:40
             Resource_List.ncpus=1 Resource_List.nodect=1
             Resource_List.place=pack Resource_List.pmem=500mb
             Resource_List.pvmem=500mb
             Resource_List.select=1:ncpus=1:vmem=500mb
             Resource_List.vmem=500mb resource_assigned.ncpus=1
```

SGI

Proprietary

11/04/2011 09:59:50 S   Obit received momhop:1 serverhop:1 state:4 substate:42
11/04/2011 09:59:51 S   Exit_status=137 resources_used.cpupercent=93
              resources_used.cput=00:01:40 resources_used.mem=932kb
              resources_used.ncpus=1 resources_used.vmem=29260kb
              resources_used.walltime=00:01:41
11/04/2011 09:59:51 S   dequeuing from small, state 5
11/04/2011 09:59:51 A   user=sgi group=sgi jobname=pbs_loop_batch. queue=small
              ctime=1320418690 qtime=1320418690 etime=1320418690
              start=1320418690 exec_host=n001/0
              exec_vnode=(n001:ncpus=1) Resource_List.cput=00:01:40
              Resource_List.ncpus=1 Resource_List.nodect=1
              Resource_List.place=pack Resource_List.pmem=500mb
              Resource_List.pvmem=500mb
              Resource_List.select=1:ncpus=1:vmem=500mb
              Resource_List.vmem=500mb session=6054 end=1320418791
              Exit_status=137 resources_used.cpupercent=93
              resources_used.cput=00:01:40 resources_used.mem=932kb
              resources_used.ncpus=1 resources_used.vmem=29260kb
              resources_used.walltime=00:01:41 run_count=1

SGI
Proprietary

sgi

Output du job qui se termine aprés le default cputime de 1m40

fortune /home/sgi> cat pbs_loop_batch..e120
/var/spool/PBS/mom_priv/jobs/120.fortune.SC: line 6: 6091 Killed

Job avec cputime de 5 minutes

```
fortune /home/sgi> cat  pbs_loop_medium.sh
#!/bin/sh
#PBS -q batch
#PBS -l ncpus=1
#PBS -l cput=00:05:00

fortune /home/sgi> qsub pbs_loop_medium.sh
121.Fortune

fortune /home/sgi> qstat -rn

fortune:
```

Job ID	Username	Queue	Jobname	SessID	NDS	TSK	Memory	Req'd Time	S	Elap Time
121.fortune	sgi	large	pbs_loop_m	6530	1	1	--	00:05	R	00:00
n001/0										

Job avec physical mémoire de 500mb

```
fortune /home/sgi> cat pbs_loop_pmem500mb.sh
#!/bin/sh
#PBS -q batch
#PBS -l ncpus=1
#PBS -l pmem=500mb

fortune /home/sgi> qsub pbs_loop_pmem500mb.sh
122.Fortune

fortune /home/sgi> qstat -rn

fortune:
                                    Req'd  Req'd   Elap
Job ID        Username Queue    Jobname   SessID NDS TSK Memory Time  S Time
-------------- -------- -------- ---------- ------ --- --- ------ ----- - -----
122.fortune    sgi     small    pbs_loop_p  6702  1   1    --  00:01 R 00:00
   n001/0
```

sgi

Directives de placement pbs

Les directives de placement doivent être utilisé en conjonction des directives de selection

Cette directive aura la syntaxe suivant

    -l place=[arrangement][: sharing][: grouping]

où

    arrangement pourra prendre la valeur free , pack ou scatter

    sharing pourra prendre la valeur excl ou shared

    grouping sera de la forme group=resource

avec la signification suivante:

    free: Place les chunks sur n'importe quel noeud.

    pack: Tous les chunks seront mis sur un noeud.

    scatter: Les chunks seront repartis sur les noeuds.

    excl: Seul ce job utilisera le noeud choisi.

    Shared: Ce job peut partager lenoeid choisi avec d'autres jobs.

    group=resource: Les Chunks seront regroupé sur les noeuds ayant la meme ressource

Soumission de 2 jobs

```
fortune /home/sgi> qsub  pbs_loop_batch_excl.sh
124.Fortune

fortune /home/sgi> qsub  pbs_loop_batch_excl.sh
125.Fortune

fortune /home/sgi> qstat -rn

fortune:
                                  Req'd  Req'd   Elap
Job ID       Username Queue    Jobname    SessID NDS TSK Memory
Time  S Time
--------------- -------- -------- ---------- ------ --- --- ------ ----- - -----
124.fortune    sgi     small    pbs_loop_b  2141  1   1    --  00:01 R 00:00
  n001/0

fortune /home/sgi> qstat -q

server: fortune

Queue          Memory CPU Time Walltime Node   Run   Que   Lm  State
--------------- ------ -------- -------- ---- ----- ----- ----  -----
workq            --     --      --    --    0    0   --  E R
small           --   00:02:00   --    --    1    1   1  E R
medium           --   00:10:00   --    --    0    0   1  E R
large           --  10:00:06   --    --    0    0   1  E R
batch           --     --      --    --    0    0   --  E R
                                ----- -----
                                  1    1
```

SGI
Proprietary

Place=excl:scatter

```
Qmgr: set queue small max_running = 12

Qmgr:  set queue medium  max_running = 3

fortune /home/sgi> qsub pbs_loop_batch_excl.sh
126.fortune
fortune /home/sgi> qsub pbs_loop_batch_excl.sh
127.Fortune

fortune /home/sgi> qstat -rn

fortune:
                                    Req'd  Req'd   Elap
Job ID        Username Queue    Jobname    SessID NDS TSK Memory Time  S Time
--------------- -------- -------- ---------- ------ --- --- ------ ----- - -----
126.fortune    sgi     small    pbs_loop_b  2342  1  1   --  00:01 R 00:00
  n001/0
127.fortune    sgi     small    pbs_loop_b  9536  1  1   --  00:01 R 00:00
  n002/0

fortune /home/sgi> cat pbs_loop_batch_excl.sh
#!/bin/sh
#PBS -q batch
#PBS -l select=1:ncpus=1 -l place=excl:scatter

/home/sgi/loop
```

Place=pack:shared

```
fortune /home/sgi> qsub  pbs_loop_batch_pack.sh
128.Fortune
fortune /home/sgi> qsub pbs_loop_batch_pack.sh
129.Fortune

fortune /home/sgi> cat  pbs_loop_batch_pack.sh
#!/bin/sh
#PBS -q batch
#PBS -l select=1:ncpus=1 -l place=pack:shared

/home/sgi/loop

fortune /home/sgi> qstat -rn

fortune:
```

| | | | | | | | Req'd | Req'd | Elap | | |
Job ID	Username	Queue	Jobname	SessID	NDS	TSK	Memory	Time	S	Time
128.fortune	sgi	small	pbs_loop_b	3314	1	1	--	00:01	R	00:00
n001/0										
129.fortune	sgi	small	pbs_loop_b	3352	1	1	--	00:01	R	00:00
n001/1										

SGI

sgi

```
top - 12:04:38 up 17 days,  1:14,  2 users,  load average: 1.55, 0.48, 0.17
Tasks: 532 total,   3 running, 529 sleeping,   0 stopped,   0 zombie
Cpu(s):  8.2%us,  0.0%sy,  0.0%ni, 91.8%id,  0.0%wa,  0.0%hi,  0.0%si,  0.0%st
Mem:    96872M total,    4902M used,   91969M free,      1M buffers
Swap:    7999M total,       0M used,    7999M free,    1248M cached

PID USER     PR  NI  VIRT  RES  SHR S %CPU %MEM    TIME+   P COMMAND
5457 sgi      20   0  3700  344  272 R  100  0.0   1:20.83  6 loop
5372 sgi      20   0  3700  344  272 R  100  0.0   1:22.62 12 loop
```

SGI

Proprietary

Place=pack:shared

```
fortune /home/sgi> qsub pbs_loop_batch_pack.sh
132.fortune
fortune /home/sgi> qsub pbs_loop_batch_pack.sh
133.fortune
fortune /home/sgi> qsub pbs_loop_batch_pack.sh
134.fortune
fortune /home/sgi> qsub pbs_loop_batch_pack.sh
135.fortune
fortune /home/sgi> qstat -rn

fortune:
                                     Req'd  Req'd  Elap
Job ID        Username Queue    Jobname   SessID NDS TSK Memory Time  S Time
--------------- -------- -------- ---------- ------ --- --- ------ ----- - -----
132.fortune    sgi     small    pbs_loop_b  7510  1   1    --  00:01 R 00:00  n001/0
133.fortune    sgi     small    pbs_loop_b  7549  1   1    --  00:01 R 00:00  n001/1
134.fortune    sgi     small    pbs_loop_b  7598  1   1    --  00:01 R 00:00  n001/2
135.fortune    sgi     small    pbs_loop_b  7636  1   1    --  00:01 R 00:00  n001/3


Tasks: 535 total,   5 running, 530 sleeping,   0 stopped,   0 zombie
Cpu(s): 16.9%us,  0.0%sy,  0.0%ni, 83.1%id,  0.0%wa,  0.0%hi,  0.0%si,  0.0%st
Mem:    96872M total,    4907M used,    91965M free,       1M buffers
Swap:   7999M total,      0M used,    7999M free,    1249M cached


 PID USER     PR  NI  VIRT  RES  SHR S %CPU %MEM    TIME+   P COMMAND
7673 sgi      20   0  3700  344  272 R  100  0.0  1:27.55  7 loop
7548 sgi      20   0  3700  344  272 R  100  0.0  1:30.95 12 loop
7597 sgi      20   0  3700  344  272 R  100  0.0  1:29.54  9 loop
7635 sgi      20   0  3700  344  272 R  100  0.0  1:28.44  5 loop
```

SGI
Proprietary

Place=pack:shared (avec 6 cores)

Commande cpumap

Sharing of Last Level (3) Caches

```
 Socket    (Logical) Processors
 ------    -------------------------
   0     0   1   2   3   4   5   12   13   14   15   16   17
   1     6   7   8   9   10  11  18   19   20   21   22   23
```

```
top - 13:21:56 up 17 days,  2:31,  2 users,  load average: 2.72, 1.82, 1.00
Tasks: 542 total,   7 running, 535 sleeping,   0 stopped,   0 zombie
Cpu(s): 24.9%us,  0.0%sy,  0.0%ni, 75.1%id,  0.0%wa,  0.0%hi,  0.0%si,  0.0%st
Mem:    96872M total,    4911M used,   91961M free,       1M buffers
Swap:   7999M total,       0M used,    7999M free,    1249M cached
```

```
 PID USER      PR  NI  VIRT  RES  SHR S %CPU %MEM   TIME+   P COMMAND
8523 sgi       20   0  3700  344  272 R  100  0.0   0:33.18 12 loop
8561 sgi       20   0  3700  344  272 R  100  0.0   0:31.80  6 loop
8599 sgi       20   0  3700  344  272 R  100  0.0   0:30.82  7 loop
8637 sgi       20   0  3700  344  272 R  100  0.0   0:29.98  5 loop
8713 sgi       20   0  3700  344  272 R  100  0.0   0:28.41  8 loop
8675 sgi       20   0  3700  344  272 R  100  0.0   0:29.20  4 loop
```

Place=pack:shared (avec 12 cores)

fortune /home/sgi> qstat -rn

fortune:

Job ID	Username	Queue	Jobname	Req'd SessID	Req'd NDS	TSK	Elap Memory Time	S	Time

```
--------------- -------- -------- ---------- ------ --- --- ------ ----- - -----
151.fortune   sgi     small    pbs_loop_b 8924  1   1   --  00:01 R 00:00
                               n001/0
152.fortune   sgi     small    pbs_loop_b 8962  1   1   --  00:01 R 00:00
                               n001/1
153.fortune   sgi     small    pbs_loop_b 9000  1   1   --  00:01 R 00:00
                               n001/2
154.fortune   sgi     small    pbs_loop_b 9039  1   1   --  00:01 R 00:00
                               n001/3
155.fortune   sgi     small    pbs_loop_b 9077  1   1   --  00:01 R 00:00
                               n001/4
156.fortune   sgi     small    pbs_loop_b 9115  1   1   --  00:01 R 00:00
                               n001/5
157.fortune   sgi     small    pbs_loop_b 9153  1   1   --  00:01 R 00:00
                               n001/6
158.fortune   sgi     small    pbs_loop_b 9191  1   1   --  00:01 R 00:00
                               n001/7
159.fortune   sgi     small    pbs_loop_b 9230  1   1   --  00:01 R 00:00
                               n001/8
160.fortune   sgi     small    pbs_loop_b 9268  1   1   --  00:01 R 00:00
                               n001/9
161.fortune   sgi     small    pbs_loop_b 9306  1   1   --  00:01 R 00:00
                               n001/10
162.fortune   sgi     small    pbs_loop_b 9345  1   1   --  00:01 R 00:00
                               n001/11
```

```
Tasks: 560 total,  13 running, 547 sleeping,   0 stopped,   0 zombie
Cpu(s): 50.5%us,  0.0%sy,  0.0%ni, 49.5%id,  0.0%wa,  0.0%hi,  0.0%si,  0.0%s
  Mem:   96872M total,    4916M used,   91955M free,       1M buffers
  Swap:   7999M total,       0M used,    7999M free,    1249M cached

  PID USER      PR NI  VIRT  RES  SHR S %CPU %MEM    TIME+  P COMMAND
 9344 sgi       20  0  3700  344  272 R  100  0.0  0:56.43 1 loop
 9382 sgi       20  0  3700  344  272 R  100  0.0  0:55.72 11 loop
 8999 sgi       20  0  3700  344  272 R  100  0.0  1:05.95 6 loop
 9037 sgi       20  0  3700  344  272 R  100  0.0  1:04.29 7 loop
 9076 sgi       20  0  3700  344  272 R  100  0.0  1:03.23 4 loop
 9114 sgi       20  0  3700  344  272 R  100  0.0  1:02.27 8 loop
 9152 sgi       20  0  3700  344  272 R  100  0.0  1:00.23 3 loop
 9190 sgi       20  0  3700  344  272 R  100  0.0  0:59.31 9 loop
 9228 sgi       20  0  3700  344  272 R  100  0.0  0:58.51 2 loop
 9267 sgi       20  0  3700  344  272 R  100  0.0  0:57.85 10 loop
 9305 sgi       20  0  3700  344  272 R  100  0.0  0:57.08 5 loop
 8961 sgi       20  0  3700  344  272 R   99  0.0  1:07.31 12 loop
```

SGI

Proprietary

Place=excl: (avec 12 cores)

```
fortune sgi/run> qsub pbs_loop_batch_excl.sh
165.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
166.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
167.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
168.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
169.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
170.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
171.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
172.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
173.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
174.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
175.fortune
fortune sgi/run> qsub pbs_loop_batch_excl.sh
176.fortune
```

fortune:

Job ID	Username	Queue	Jobname	SessID	Req'd NDS	Req'd TSK	Elap Memory Time S Time
165.fortune	sgi	small	pbs_loop_b n001/0	10358	1	1	-- 00:01 R 00:00
166.fortune	sgi	small	pbs_loop_b n002/0	15745	1	1	-- 00:01 R 00:00
167.fortune	sgi	small	pbs_loop_b n003/0	30577	1	1	-- 00:01 R 00:00
168.fortune	sgi	small	pbs_loop_b n004/0	30017	1	1	-- 00:01 R 00:00
169.fortune	sgi	small	pbs_loop_b n005/0	30902	1	1	-- 00:01 R 00:00
170.fortune	sgi	small	pbs_loop_b n006/0	4736	1	1	-- 00:01 R 00:00
171.fortune	sgi	small	pbs_loop_b n007/0	4240	1	1	-- 00:01 R 00:00
172.fortune	sgi	small	pbs_loop_b n008/0	16494	1	1	-- 00:01 R 00:00
173.fortune	sgi	small	pbs_loop_b n009/0	12275	1	1	-- 00:01 R 00:00
174.fortune	sgi	small	pbs_loop_b n010/0	13155	1	1	-- 00:01 R 00:00
175.fortune	sgi	small	pbs_loop_b n011/0	12623	1	1	-- 00:01 R 00:00
176.fortune	sgi	small	pbs_loop_b n012/0	12153	1	1	-- 00:01 R 00:00

sgi

Dplace (avec 6 cores)

```
fortune sgi/run> cat pbs_loop_batch_place_pair.sh
#!/bin/sh
#PBS -q batch
#PBS -l select=1:ncpus=6


dplace -c0 /home/sgi/loop &
dplace -c2 /home/sgi/loop &
dplace -c4 /home/sgi/loop &
dplace -c6 /home/sgi/loop &
dplace -c8 /home/sgi/loop &
dplace -c10 /home/sgi/loop
fortune sgi/run> qsub pbs_loop_batch_place_pair.sh
185.fortune
fortune sgi/run> qstat -rn
```

fortune:

Job ID	Username	Queue	Jobname	SessID	NDS	TSK	Req'd Memory	Req'd Time	S	Elap Time
185.fortune	sgi	small	pbs_loop_b	11821	1	6	--	00:01	R	00:00

   n001/0*6

```
top - 14:33:35 up 17 days,  3:43,  2 users,  load average: 2.47, 1.21, 0.55
Tasks: 532 total,   7 running, 525 sleeping,   0 stopped,   0 zombie
Cpu(s): 25.2%us,  0.0%sy,  0.0%ni, 74.7%id,  0.0%wa,  0.0%hi,  0.0%si,
                               0.0%st
Mem:    96872M total,    4907M used,   91964M free,       1M buffers
Swap:    7999M total,       0M used,    7999M free,    1250M cached

PID USER      PR NI VIRT  RES  SHR S %CPU %MEM   TIME+   P
                                COMMAND
11979 sgi     20  0 10064  580  468 R  100  0.0  0:12.88  0 loop
11980 sgi     20  0 10064  580  468 R  100  0.0  0:12.88  2 loop
11981 sgi     20  0 10064  580  468 R  100  0.0  0:12.88  4 loop
11982 sgi     20  0 10064  580  468 R  100  0.0  0:12.88  6 loop
11983 sgi     20  0 10064  580  468 R  100  0.0  0:12.88  8 loop
11984 sgi     20  0 10064  580  468 R  100  0.0  0:12.88 10 loop
```

Dplace (avec 6 cores)

```
fortune sgi/run> qsub pbs_loop_batch_place_impair.sh
187.fortune
fortune sgi/run> qstat -rn

fortune:
                                      Req'd  Req'd   Elap
Job ID        Username Queue   Jobname   SessID NDS TSK Memory Time  S Time
--------------- -------- -------- ---------- ------ --- --- ------ ----- - -----
187.fortune    sgi     small    pbs_loop_b 16539  1   6    --  00:01 R 00:00
   n001/0*6


top - 17:32:20 up 17 days,  6:41,  2 users,  load average: 1.77, 0.70, 0.26
Tasks: 532 total,   7 running, 525 sleeping,   0 stopped,   0 zombie
Cpu(s): 25.5%us,  0.0%sy,  0.0%ni, 74.5%id,  0.0%wa,  0.0%hi,  0.0%si,  0.0%st
Mem:   96872M total,    4907M used,   91964M free,       1M buffers
Swap:   7999M total,       0M used,    7999M free,    1250M cached


 PID USER     PR NI VIRT RES  SHR S %CPU %MEM   TIME+   P COMMAND
16718 sgi     20   0 10064 580  468 R  100  0.0   0:12.16  1 loop
16719 sgi     20   0 10064 580  468 R  100  0.0   0:12.16  3 loop
16720 sgi     20   0 10064 580  468 R  100  0.0   0:12.16  5 loop
16723 sgi     20   0 10064 580  468 R  100  0.0   0:12.16 11 loop
16721 sgi     20   0 10064 580  468 R   99  0.0   0:12.16  7 loop
16722 sgi     20   0 10064 580  468 R   99  0.0   0:12.16  9 loop
```

sgi

Dplace (avec 12 cores et 2 chunks)

```
fortune sgi/run> qsub pbs_loop_batch_place.sh
189.Fortune

fortune sgi/run> qstat -rn
fortune:

                                        Req'd  Req'd   Elap
Job ID        Username Queue    Jobname    SessID NDS TSK Memory
Time  S Time
-------------- -------- -------- ---------- ------ --- --- ------ ----- - -----
189.fortune    sgi    small   pbs_loop_b 17111  2  12   --  00:01 R 00:00
   n001/0*6+n001/1*6

fortune sgi/run> cat pbs_loop_batch_place.sh
#!/bin/sh
#PBS -q batch
#PBS -l select=2:ncpus=6

dplace -c0 /home/sgi/loop &
dplace -c1 /home/sgi/loop &
dplace -c2 /home/sgi/loop &
dplace -c3 /home/sgi/loop &
dplace -c4 /home/sgi/loop &
dplace -c5 /home/sgi/loop &
dplace -c6 /home/sgi/loop &
dplace -c7 /home/sgi/loop &
dplace -c8 /home/sgi/loop &
dplace -c9 /home/sgi/loop &
dplace -c10 /home/sgi/loop &
dplace -c11 /home/sgi/loop
```

```
top - 17:47:50 up 17 days,  6:57,  2 users,  load average: 1.07, 0.45, 0.27
  Tasks: 537 total,  12 running, 525 sleeping,   0 stopped,   0 zombie
Cpu(s): 46.0%us,  0.0%sy,  0.0%ni, 53.9%id,  0.0%wa,  0.0%hi,  0.0%si,
                          0.0%st
  Mem:   96872M total,   4909M used,  91962M free,     1M buffers
  Swap:   7999M total,      0M used,   7999M free,   1250M cached


 PID USER       PR  NI  VIRT  RES  SHR S %CPU %MEM    TIME+   P
                              COMMAND
 17258 sgi      20   0 10064  580  468 R  100  0.0   0:08.31  1 loop
 17259 sgi      20   0 10064  580  468 R  100  0.0   0:08.31  2 loop
 17260 sgi      20   0 10064  580  468 R  100  0.0   0:08.31  3 loop
 17260 sgi      20   0 10064  580  468 R  100  0.0   0:08.31  4 loop
 17262 sgi      20   0 10064  580  468 R  100  0.0   0:08.31  5 loop
 17263 sgi      20   0 10064  580  468 R  100  0.0   0:08.31  6 loop
 17264 sgi      20   0 10064  580  468 R  100  0.0   0:08.31  7 loop
 17265 sgi      20   0 10064  580  468 R  100  0.0   0:08.31  8 loop
 17266 sgi      20   0 10064  580  468 R  100  0.0   0:08.31  9 loop
 17267 sgi      20   0 10064  580  468 R  100  0.0   0:08.31 10 loop
 17268 sgi      20   0 10064  580  468 R  100  0.0   0:08.31 11 loop
 17257 sgi      20   0 10064  580  468 R   99  0.0   0:08.30  0 loop
```

Gui pour la soumission (xpbs)
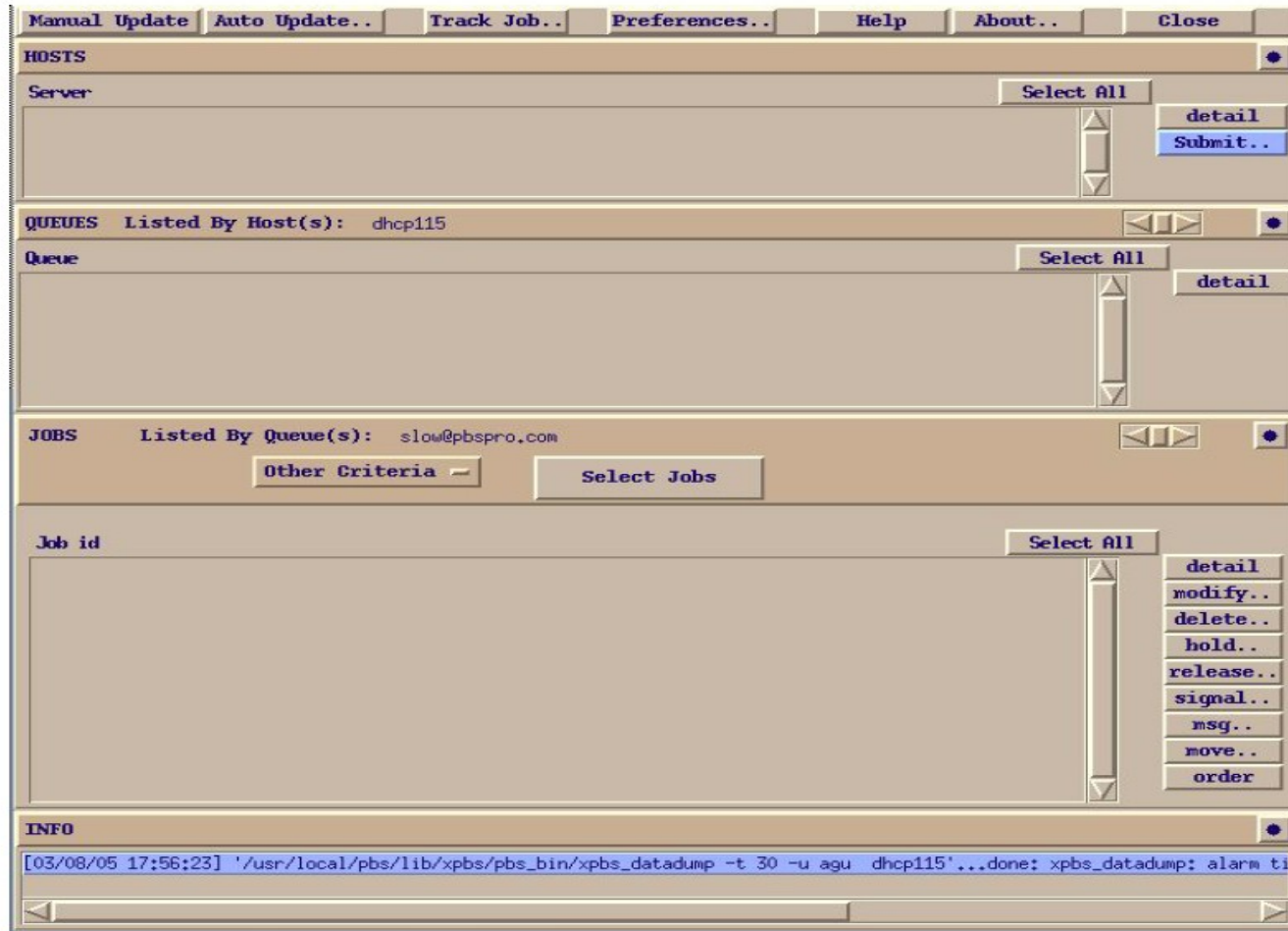
Gui pour le monitoring (xpbsmon)

Utilisation de module

sgi
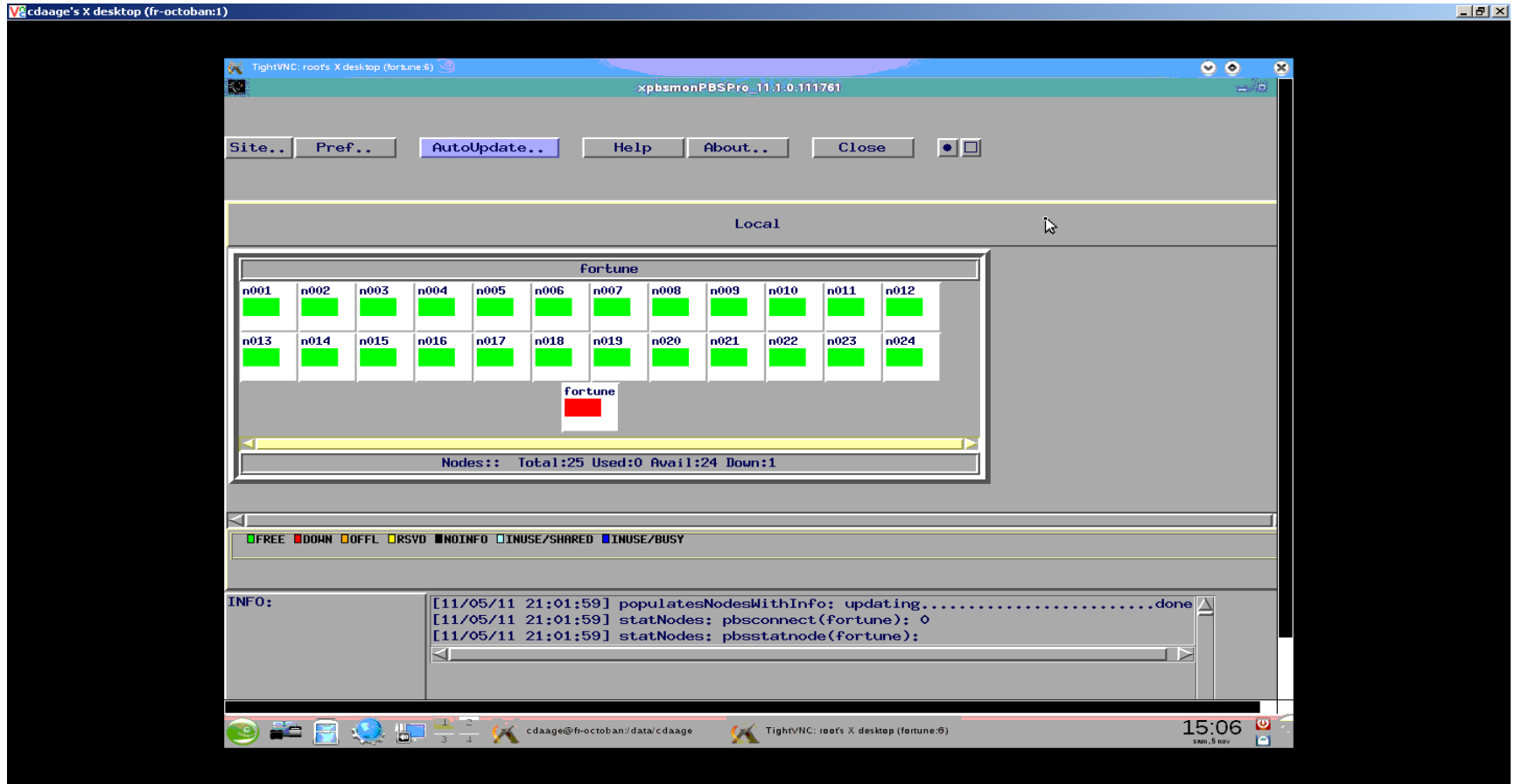
Module (ensemble de scripts qui place des variables)

```
sgi@n001:~> rpm -qa |grep Module
Modules-3.1.6-39.9sgi704rp2.sles11

sgi@n001:~> rpm -ql Modules-3.1.6-39.9sgi704rp2.sles11
/etc/profile.d/modules.csh
/etc/profile.d/modules.sh
/usr/bin/add.modules
/usr/bin/modulecmd
/usr/share/modules
/usr/share/modules/3.1.6
/usr/share/modules/init
/usr/share/modules/init/.modulespath
/usr/share/modules/init/bash
/usr/share/modules/init/csh
/usr/share/modules/init/ksh
/usr/share/modules/init/perl
/usr/share/modules/init/python
/usr/share/modules/init/sh
/usr/share/modules/init/tcsh
/usr/share/modules/init/zsh
/usr/share/modules/modulefiles
/usr/share/modules/modulefiles/dot
/usr/share/modules/modulefiles/module-cvs
/usr/share/modules/modulefiles/module-info
/usr/share/modules/modulefiles/modules
/usr/share/modules/modulefiles/null
/usr/share/modules/modulefiles/use.own
```

```
sgi@n001:~> cat /etc/profile.d/modules.sh
#----------------------------------------------------------------------#
# system-wide profile.modules                                          #
# Initialize modules for all sh-derivative shells                      #
#----------------------------------------------------------------------#
trap "" 1 2 3

case "$0" in
   -bash|bash|*/bash) . /usr/share/modules/init/bash ;;
      -ksh|ksh|*/ksh) . /usr/share/modules/init/ksh ;;
        -sh|sh|*/sh) . /usr/share/modules/init/sh ;;
              *) . /usr/share/modules/init/sh ;;  # default for scripts
esac

trap - 1 2 3
```

SGI
Proprietary

sgi

```
sgi@n001:~> cat /etc/profile.d/modules.csh
#-----------------------------------------------------------------#
# system-wide csh.modules                                         #
# Initialize modules for all csh-derivative shells               #
#-----------------------------------------------------------------#
if ($?tcsh) then
      set modules_shell="tcsh"
else
      set modules_shell="csh"
endif

source /usr/share/modules/init/${modules_shell}

unset modules_shell
```

SGI

Proprietary

Module avail

```
sgi@n001:~> module avail

----------------------------- /usr/share/modules -----------------------------
3.1.6                      modulefiles/modules
modulefiles/MPInside/3.1      modulefiles/mpi/intel-4.0
modulefiles/chkfeature        modulefiles/mpt/2.04
modulefiles/compiler/gcc-4.3.3 modulefiles/null
modulefiles/dot              modulefiles/perfboost
modulefiles/module-cvs        modulefiles/perfcatcher
modulefiles/module-info       modulefiles/use.own

------------------------ /usr/share/modules/modulefiles ------------------------
MPInside/3.1      module-cvs        mpt/2.04          use.own
chkfeature       module-info       null
compiler/gcc-4.3.3 modules          perfboost
dot              mpi/intel-4.0     perfcatcher

------------------------ /opt/intel/Modules/modulefiles ------------------------
intel-cce-12/12.0.0.084  intel-idbe-12/12.0.0.084 intel-mkl-10/10.2.4
intel-cce-12/12.1.0.233  intel-idbe-12/12.1.0.233 intel-mkl-10/10.3.6
intel-fce-12/12.0.0.084  intel-itac-8/8.0.1.9     intel-mpi-4/4.0.1.007
intel-fce-12/12.1.0.233  intel-itac-8/8.0.2.3     intel-mpi-4/4.0.2.003
```

Module load

```
sgi@n001:~> icc
If 'icc' is not a typo you can run the following command to lookup the package that contains the binary:
    command-not-found icc
-bash: icc: command not found

sgi@n001:~> module load intel-cce-12/12.1.0.233

sgi@n001:~> icc -V
Intel(R) C Intel(R) 64 Compiler XE for applications running on Intel(R) 64,
Version 12.1.0.233 Build 20110811
Copyright (C) 1985-2011 Intel Corporation.  All rights reserved.

sgi@n001:~> module list
Currently Loaded Modulefiles:
  1) intel-cce-12/12.1.0.233

sgi@n001:~> ifort -V
Intel(R) Fortran Intel(R) 64 Compiler XE for applications running on Intel(R) 64,
Version 12.1.0.233 Build 20110811
Copyright (C) 1985-2011 Intel Corporation.  All rights reserved.
```

sgi

Module commandes

```
fortune:/ # module -help

 Modules Release 3.1.6 (Copyright GNU GPL v2 1991):
 Available Commands and Usage:
      + add|load            modulefile [modulefile ...]
      + rm|unload           modulefile [modulefile ...]
      + switch|swap         modulefile1 modulefile2
      + display|show        modulefile [modulefile ...]
      + avail              [modulefile [modulefile ...]]
      + use [-a|--append]     dir [dir ...]
      + unuse              dir [dir ...]
      + update
      + purge
      + list
      + clear
      + help               [modulefile [modulefile ...]]
      + whatis             [modulefile [modulefile ...]]
      + apropos|keyword       string
      + initadd            modulefile [modulefile ...]
      + initprepend         modulefile [modulefile ...]
      + initrm             modulefile [modulefile ...]
      + initswitch          modulefile1 modulefile2
      + initlist
      + initclear
```

SGI
Proprietary

Utilisation des compilateurs

C:

icc -openmp omp_hello.c -o hello (compilateurs intel)
pathcc -mp omp_hello.c -o hello (pathscale compilateurs =*www.pathscale.com*)
pgcc -mp omp_hello.c -o hello (portland compilateurs = *www.pgroup.com*)
gcc -fopenmp omp_hello.c -o hello (compilateurs gnu)

Fortran:

ifort -openmp omp_hello.f -o hello
pathf90 -mp omp_hello.f -o hello
pgf90 -mp omp_hello.f -o hello
gfortran -fopenmp omp_hello.f -o hello

Progmp lancer sur 2 chunks (pbs)

```
fortune /home/sgi> cat pbs_progmp.sh
#!/bin/sh
#PBS -q workq
#PBS -l select=2:ncpus=6
export OMP_NUM_THREADS=12
/home/sgi/progmp


fortune /home/sgi> qsub pbs_progmp.sh
45.Fortune

fortune /home/sgi> qstat -rn

fortune:
```

| | | | | | | | Req'd | Req'd | Elap | |
Job ID	Username	Queue	Jobname	SessID	NDS	TSK	Memory	Time	S	Time
45.fortune	sgi	workq	pbs_progmp	3593	2	12	--	--	R	00:00
n001/0*6+n001/1*6										

Visualisation des threads openmp avec top

Top avec f (j) affiche core number + H (voir les threads)

```
PID USER      PR  NI  VIRT  RES  SHR S %CPU %MEM   TIME+   P COMMAND

3745 sgi      20   0  122m 1616  936 R   97  0.0  0:14.40  2 progmp
3744 sgi      20   0  122m 1616  936 R   97  0.0  0:14.37 12 progmp
3747 sgi      20   0  122m 1616  936 R   97  0.0  0:14.42  4 progmp
3748 sgi      20   0  122m 1616  936 R   97  0.0  0:14.41  7 progmp
3749 sgi      20   0  122m 1616  936 R   97  0.0  0:14.40  3 progmp
3751 sgi      20   0  122m 1616  936 R   97  0.0  0:14.39  1 progmp
3754 sgi      20   0  122m 1616  936 R   97  0.0  0:14.41  5 progmp
3742 sgi      20   0  122m 1616  936 R   96  0.0  0:14.36 18 progmp
3746 sgi      20   0  122m 1616  936 R   96  0.0  0:14.34 11 progmp
3750 sgi      20   0  122m 1616  936 R   96  0.0  0:14.32  9 progmp
3752 sgi      20   0  122m 1616  936 R   96  0.0  0:14.33  8 progmp
3753 sgi      20   0  122m 1616  936 R   96  0.0  0:14.37 10 progmp
```

SGI

Proprietary

Comment va etre éxécuté ce binaire

sgi@n001:~> ldd omp_hello (compilé avec gcc / /usr/lib64/libgomp.so.1)
    linux-vdso.so.1 => (0x00007ffff7fdf000)
    libgomp.so.1 => /usr/lib64/libgomp.so.1 (0x00007ffff7dca000)
    libpthread.so.0 => /lib64/libpthread.so.0 (0x00007ffff7bac000)
    libc.so.6 => /lib64/libc.so.6 (0x00007ffff784e000)
    librt.so.1 => /lib64/librt.so.1 (0x00007ffff7645000)
    /lib64/ld-linux-x86-64.so.2 (0x00007ffff7fe0000)

sgi@n001:~> ldd progmp (compilé avec icc / /opt/intel/mkl/10.2.4.032/lib/em64t/libiomp5.so)
    linux-vdso.so.1 => (0x00007ffff7fdf000)
    libm.so.6 => /lib64/libm.so.6 (0x00007ffff7d7d000)
    libiomp5.so => /opt/intel/mkl/10.2.4.032/lib/em64t/libiomp5.so (0x00007ffff7bd7000)
    libgcc_s.so.1 => /lib64/libgcc_s.so.1 (0x00007ffff79c0000)
    libpthread.so.0 => /lib64/libpthread.so.0 (0x00007ffff77a3000)
    libc.so.6 => /lib64/libc.so.6 (0x00007ffff7444000)
    libdl.so.2 => /lib64/libdl.so.2 (0x00007ffff7240000)
    /lib64/ld-linux-x86-64.so.2 (0x00007ffff7fe0000)

SGI

Proprietary

Nm (liste les symboles)

```
fortune /home/sgi> nm omp_hello
        U GOMP_parallel_end@@GOMP_1.0
        U GOMP_parallel_start@@GOMP_1.0
0000000000600e20 d _DYNAMIC
0000000000600fe8 d _GLOBAL_OFFSET_TABLE_
00000000004008a0 R _IO_stdin_used
                 w _Jv_RegisterClasses
0000000000600e00 d __CTOR_END__
0000000000600df8 d __CTOR_LIST__
0000000000600e10 D __DTOR_END__
0000000000600e08 d __DTOR_LIST__
00000000004009e0 r __FRAME_END__
0000000000600e18 d __JCR_END__
0000000000600e18 d __JCR_LIST__
0000000000601040 A __bss_start

            ....
```

Commande qstat –f (sur un job qui a déjà terminé)

```
fortune /home/sgi> qstat -f 45
qstat: Unknown Job Id 45.fortune

 Status information for finished and moved jobs can
be  displayed  using
      the -x and -H options.

fortune /home/sgi> qstat -H
qstat: PBS is not configured to maintain job history
```

Pour avoir l'historique des jobs, il faut lancer sous qmgr:
set server job_history_enable=true

Ou utiliser la commande tracejob

man commande qstat (PBS)

man -M /opt/pbs/default/man qstat

qstat - display status of PBS batch jobs, queues, or servers

SYNOPSIS
    Displaying Job Status
    Default format:
    qstat [-p] [-J] [-t] [-x]
        [[job_identifier | destination] ...]

    Long format:
    qstat -f [-p] [-J] [-t] [-x]
        [[job_identifier | destination] ...]

    Alternate format:
    qstat [-a [-w]| -H | -i | -r] [-G | -M] [-J] [-n [-1][-w]]
        [-s [-1][-w]] [-t] [-T [-w]] [-u user_list]
        [[job_identifier | destination] ...]

sgi

```
fortune /home/sgi> tracejob 45

Job: 45.fortune

10/23/2011 08:00:29  L    Considering job to run
10/23/2011 08:00:29  S    enqueuing into workq, state 1 hop 1
10/23/2011 08:00:29  S    Job Queued at request of sgi@fortune, owner =
                sgi@fortune, job name = pbs_progmp.sh, queue = workq
10/23/2011 08:00:29  S    Job Run at request of Scheduler@fortune on exec_vnode
                (n001:ncpus=6)+(n001:ncpus=6)
10/23/2011 08:00:29  S    Job Modified at request of Scheduler@fortune
10/23/2011 08:00:29  L    Job run
10/23/2011 08:00:29  A    queue=workq
10/23/2011 08:00:29  A    user=sgi group=sgi jobname=pbs_progmp.sh queue=workq
                ctime=1319374829 qtime=1319374829 etime=1319374829
                start=1319374829 exec_host=n001/0*6+n001/1*6
                exec_vnode=(n001:ncpus=6)+(n001:ncpus=6)
                Resource_List.ncpus=12 Resource_List.nodect=2
                Resource_List.place=free
                Resource_List.select=2:ncpus=6
                resource_assigned.ncpus=12
```

```
10/23/2011 08:01:23  S   Obit received momhop:1 serverhop:1 state:4 substate:42
10/23/2011 08:01:32  S   Exit_status=0 resources_used.cpupercent=0
                resources_used.cput=00:10:19 resources_used.mem=2232kb
                resources_used.ncpus=12 resources_used.vmem=216244kb
                resources_used.walltime=00:00:54
10/23/2011 08:01:32  S   dequeuing from workq, state 5
10/23/2011 08:01:32  A   user=sgi group=sgi jobname=pbs_progmp.sh
queue=workq
                ctime=1319374829 qtime=1319374829 etime=1319374829
                start=1319374829 exec_host=n001/0*6+n001/1*6
                exec_vnode=(n001:ncpus=6)+(n001:ncpus=6)
                Resource_List.ncpus=12 Resource_List.nodect=2
                Resource_List.place=free
                Resource_List.select=2:ncpus=6 session=3593
                end=1319374892 Exit_status=0
                resources_used.cpupercent=0
                resources_used.cput=00:10:19 resources_used.mem=2232kb
                resources_used.ncpus=12 resources_used.vmem=216244kb
                resources_used.walltime=00:00:54 run_count=1
```

SGI
Proprietary

sgi

2 sockets (chunk) par nodes (2 processeurs de 6 cores = soit 24 avec hyperthreading)

```
fortune /home/sgi> cpumap
dim. oct. 23 15:42:33 CEST 2011
fortune

model name          : Intel(R) Xeon(R) CPU X5675 @ 3.07GHz
Architecture        : x86_64
cpu MHz             : 1600.000
cache size          : 12288 KB (Last Level)

Total Number of Sockets             : 2
Total Number of Cores               : 12   (6 per socket)
Hyperthreading                      : ON
Total Number of Physical Processors : 12
Total Number of Logical Processors  : 24   (2 per Phys Processor)
==============================================================================

Processor Numbering on Socket(s)

  Socket   (Logical) Processors
  ------   ------------------------
    0      0   1   2   3   4   5   12  13  14  15  16  17
    1      6   7   8   9   10  11  18  19  20  21  22  23
==============================================================================

Sharing of Last Level (3) Caches

  Socket   (Logical) Processors
  ------   ------------------------
    0      0   1   2   3   4   5   12  13  14  15  16  17
    1      6   7   8   9   10  11  18  19  20  21  22  23
==============================================================================

HyperThreading

  Shared Processors
  ----------------
  (   0,  12) (   1,  13) (   2,  14) (   3,  15)
  (   4,  16) (   5,  17) (   6,  18) (   7,  19)
  (   8,  20) (   9,  21) (  10,  22) (  11,  23)
```

sgi

Intel cpuinfo

```
sgi@n001:/opt/intel> find . -name "*cpuinfo*"
./clck/1.6/doc/html/cpuinfo.7.html
./clck/1.6/doc/man/man7/clck-cpuinfo.7.gz
./impi/4.0.0.025/ia32/bin/cpuinfo
./impi/4.0.0.025/intel64/bin/cpuinfo
./impi/4.0.1.007/ia32/bin/cpuinfo
./impi/4.0.1.007/intel64/bin/cpuinfo
./impi/4.0.2.003/ia32/bin/cpuinfo
./impi/4.0.2.003/intel64/bin/cpuinfo
./composerxe-2011.0.084/mpirt/bin/ia32/cpuinfo
./composerxe-2011.0.084/mpirt/bin/intel64/cpuinfo
./composer_xe_2011_sp1.6.233/mpirt/bin/ia32/cpuinfo
./composer_xe_2011_sp1.6.233/mpirt/bin/intel64/cpuinfo
```

SGI

Intel cpuinfo

```
sgi@n001:/opt/intel> cpuinfo
Intel(R) Xeon(R)  Processor (Intel64)
=====  Processor composition  =====
Processors(CPUs)  : 24
Packages(sockets) : 2
Cores per package : 6
Threads per core  : 2
=====  Processor identification  =====
Processor       Thread Id.      Core Id.        Package Id.
0               0               0               0
1               0               1               0
2               0               2               0
3               0               8               0
4               0               9               0
5               0               10              0
6               0               0               1
7               0               1               1
8               0               2               1
9               0               8               1
10              0               9               1
11              0               10              1
12              1               0               0
13              1               1               0
14              1               2               0
15              1               8               0
16              1               9               0
17              1               10              0
18              1               0               1
19              1               1               1
20              1               2               1
21              1               8               1
22              1               9               1
23              1               10              1
=====  Placement on packages  =====
Package Id.     Core Id.        Processors
0               0,1,2,8,9,10       (0,12)(1,13)(2,14)(3,15)(4,16)(5,17)
1               0,1,2,8,9,10       (6,18)(7,19)(8,20)(9,21)(10,22)(11,23)
```

SGI
Proprietary

Intel cpuinfo cache

```
=====  Cache sharing  =====
Cache   Size          Processors
L1      32  KB         (0,12)(1,13)(2,14)(3,15)(4,16)(5,17)(6,18)(7,19)(8,20)(9,21)(10,22)(11,23)
L2      256 KB          (0,12)(1,13)(2,14)(3,15)(4,16)(5,17)(6,18)(7,19)(8,20)(9,21)(10,22)(11,23)
L3      12  MB          (0,1,2,3,4,5,12,13,14,15,16,17)(6,7,8,9,10,11,18,19,20,21,22,23)
```

Cat /proc/cpuinfo

```
sgi@n001:~> cat /proc/cpuinfo | more
processor      : 0
vendor_id      : GenuineIntel
cpu family     : 6
model          : 44
model name     : Intel(R) Xeon(R) CPU        X5675  @ 3.07GHz
stepping       : 2
cpu MHz        : 1600.000
cache size     : 12288 KB
physical id    : 0
siblings       : 12
core id        : 0
cpu cores      : 6
apicid         : 0
initial apicid : 0
fpu            : yes
fpu_exception  : yes
cpuid level    : 11
wp             : yes
flags          : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov
pat pse36 clflush dts acpi mmx fxsr sse sse2 ss ht tm pbe syscall nx pdpe1gb rdt
scp lm constant_tsc arch_perfmon pebs bts rep_good xtopology nonstop_tsc aperfmp
erf pni dtes64 monitor ds_cpl vmx smx est tm2 ssse3 cx16 xtpr pdcm dca sse4_1 sse4_2
popcnt lahf_lm ida arat tpr_shadow vnmi flexpriority ept vpid
bogomips       : 6133.78
clflush size   : 64
cache_alignment : 64
address sizes  : 40 bits physical, 48 bits virtual
power management:
```

La mémoire est commune à 6 cores (1socket = 1 chunk)

```
fortune /home/sgi> numactl --hardware
available: 2 nodes (0-1)
node 0 cpus: 0 1 2 3 4 5 12 13 14 15 16 17
node 0 size: 49142 MB
node 0 free: 29563 MB
node 1 cpus: 6 7 8 9 10 11 18 19 20 21 22 23
node 1 size: 49152 MB
node 1 free: 28567 MB
node distances:
node   0   1
  0:  10  21
  1:  21  10

sgi@n001:~>  numactl --hardware
available: 2 nodes (0-1)
node 0 cpus: 0 1 2 3 4 5 12 13 14 15 16 17
node 0 size: 49142 MB
node 0 free: 45855 MB
node 1 cpus: 6 7 8 9 10 11 18 19 20 21 22 23
node 1 size: 49152 MB
node 1 free: 46610 MB
node distances:
node   0   1
  0:  10  21
  1:  21  10
```

SGI

Cat /proc/meminfo

```
sgi@n001:~> cat /proc/meminfo
MemTotal:       99197504 kB
MemFree:        94633052 kB
Buffers:            1200 kB
Cached:           592912 kB
SwapCached:            0 kB
Active:          2470684 kB
Inactive:         367376 kB
Active(anon):    2244016 kB
Inactive(anon):       52 kB
Active(file):     226668 kB
Inactive(file):   367324 kB
Unevictable:           0 kB
Mlocked:               0 kB
SwapTotal:       8191992 kB
SwapFree:        8191992 kB
Dirty:               184 kB
Writeback:             0 kB
AnonPages:       2244012 kB
Mapped:            66856 kB
Shmem:               120 kB
```

SGI
Proprietary

Cat /proc/meminfo (suite)

```
Slab:                  333868 kB
SReclaimable:          263852 kB
SUnreclaim:             70016 kB
KernelStack:            12080 kB
PageTables:             19584 kB
NFS_Unstable:               0 kB
Bounce:                     0 kB
WritebackTmp:               0 kB
CommitLimit:         57790744 kB
Committed_AS:         5401992 kB
VmallocTotal:     34359738367 kB
VmallocUsed:           483264 kB
VmallocChunk:     34307761356 kB
HardwareCorrupted:          0 kB
HugePages_Total:            0
HugePages_Free:             0
HugePages_Rsvd:             0
HugePages_Surp:             0
Hugepagesize:            2048 kB
DirectMap4k:             5504 kB
DirectMap2M:          2082816 kB
DirectMap1G:         98566144 kB
```

SGI

Utilisation mémoire

MemTotal : la quantité de mémoire vive utilisable mais il y a quelques Mo réquisitionnés par le noyau
MemFree :la quantité de mémoire vive actuellement non allouée.
Buffers : la quantité de mémoire vive actuellement allouée dans des buffers.
Cached : la quantité de mémoire actuellement utilisée en tant que cache
(par exemple, des caches pour les transferts vers et depuis le disque dur).
SwapCached : la quantité de mémoire actuellement sortie du swap,
mais dont une copie est toujours présente dedans.
Active : la quantité de mémoire vive récemment utilisée.
Inactive : la quantité de mémoire qui n'a pas été utilisée depuis un certain temps
(le noyau pioche d'abord dans ces pages s'il a besoin de mémoire).

```
sgi@n001:~> free
              total        used        free      shared    buffers   cached
Mem:       99197504     4756064    94441440           0       1208   1062728
-/+ buffers/cache:    3692128    95505376
Swap:       8191992           0     8191992
```

94441440+1208+1062728= 95505376

sgi

Utilisation mémoire (suite)

MemTotal = **99197504 kB** MemFree + Active + Inactive + Slab + PageTables + VmallocUsed + X **= 94633052+333868+2470684+367376+ 483264**

MemFree = **94633052 kB**

SwapTotal : la taille du swap.

SwapFree : la place disponible dans le swap.

Dirty : la quantité de mémoire en attente d'être écrite sur le disque.

Writeback : la quantité de mémoire qui est actuellement en train d'être écrite sur le disque.

AnonPages : la quantité de mémoire allouée avec la fonction mmap() avec le flag MAP_ANONYMOUS.

Mapped : la quantité de mémoire allouée avec la fonction mmap() à partir de fichiers ou de devices.

Slab : la quantité de mémoire utilisée au sein du noyau, et allouée avec l'allocateur Slab,

pour ses propres objets et structures.

PageTables : la quantité de mémoire utilisée en tant que table de pages pour gérer la pagination.

CommitLimit : une estimation de la quantité maximale de mémoire disponible pour être allouée.

VmallocTotal : la quantité totale de mémoire de l'espace d'adressage virtuel du noyau.

VmallocUsed : la quantité de mémoire utilisée dans l'espace d'adressage virtuel du noyau.

VmallocChunk : le plus grand bloc contigu de mémoire virtuelle (et non physique)

disponible dans l'espace d'adressage virtuel du noyau.

HugePages_Total : le nombre de hugepage allouée. Une hugepage est une page d'une grande taille,

utilisée par des programmes qui ont des besoins précis de mémoire.

HugePages_Free : le nombre de hugepage disponibles.

Hugepagesize : la taille d'une hugepage.

slabtop

```
Active / Total Objects (% used)   : 803312 / 822397 (97.7%)
Active / Total Slabs (% used)     : 78366 / 78373 (100.0%)
Active / Total Caches (% used)    : 104 / 206 (50.5%)
Active / Total Size (% used)      : 314256.37K / 318140.84K (98.8%)
Minimum / Average / Maximum Object : 0.02K / 0.39K / 4096.00K

 OBJS ACTIVE  USE OBJ SIZE  SLABS OBJ/SLAB CACHE SIZE NAME
404800 404162 99%   0.19K  20240     20    80960K dentry
109660 109631 99%   1.00K  27415      4   109660K xfs_inode
 60475  59467 98%   0.06K   1025     59     4100K size-64
 57864  57862 99%   1.00K  14466      4    57864K nfs_inode_cache
 28602  26426 92%   0.18K   1362     21     5448K vm_area_struct
 25715  25628 99%   0.10K    695     37     2780K buffer_head
 19488  17526 89%   0.03K    174    112      696K size-32
 18576  18444 99%   0.08K    387     48     1548K sysfs_dir_cache
 12450  12270 98%   0.12K    415     30     1660K size-128
 10808  10643 98%   0.54K   1544      7     6176K radix_tree_node
  7788   7573 97%   0.62K   1298      6     5192K proc_inode_cache
  7460   5494 73%   0.19K    373     20     1492K filp
  6340   6282 99%   0.19K    317     20     1268K xfs_ili
  4844   4709 97%   1.00K   1211      4     4844K size-1024
  4770   4402 92%   0.07K     90     53      360K Acpi-Operand
  4608   3460 75%   0.02K     32    144      128K anon_vma
```

Sgi commandes spécif ques

```
sgi@n001:~> rpm -qf  /usr/bin/cpumap
numatools-2.0-sgi704rp8.sles11

fortune /home/sgi> rpm -qf /usr/bin/dplace
numatools-2.0-sgi704rp8.sles11

fortune /home/sgi>  rpm -qf /usr/bin/taskset
util-linux-2.16-6.13.1

fortune /home/sgi> rpm -qf /usr/bin/cpuset
cpuset-utils-2.0-sgi704r1.sles11

fortune /home/sgi> rpm -qf /usr/bin/bcfree
numatools-2.0-sgi704rp8.sles11

fortune /home/sgi> rpm -qf /usr/bin/numactl
numactl-2.0.3-0.4.3

fortune /home/sgi> rpm -qf /usr/bin/dlook
numatools-2.0-sgi704rp8.sles11
```

SGI
Proprietary

sgi

les lunchers (array services et mpdboot)

```
n001:/etc/array # cat arrayd.auth
AUTHENTICATION NONE

n001:/etc/array # cat arrayd.conf
# $Id: arrayd.conf.template,v 1.18 2008/06/06 19:19:30 michel Exp $
# Sample array services daemon (arrayd) configuration file
# These are definitions of the array(s) accessible from this system. Each
# array definition consists of an "array" statement which names the array,
# followed by one or more "machine" statements, each of which specifies the
# hostname of a machine in that array.
# If you change this file, you must stop and restart array services
# to make the changes effective.
# The "localhost" entry must always be present when array services is
# standalone, ie. not communicating to any remote hosts. When defining
# an array, include all the hosts, including 'this' host, but not
# localhost.
array default
      machine admin
      machine n001
      machine n002
…
      machine n023
      machine n024
# As an additional security measure, execution of all array commands
# and remote exec requests by root may be suppressed by asserting
# "rootexecution none".  The default is "rootexecution allowed".
#
rootexecution allowed
```

```
#
# These are settings that are local to this system, such as the default
# target array.
#
local
      user  arraysvcs
      group arraysvcs
      port  5434
      options RETRYRESOLV setmachid   # Tell kernel to generate global
ASHs
      destination array default
```

SGI

mpt_hostf le

```
mpirun –f mpt_hostfile

sgi@n001:~> cat mpt_hostfile
n001 12,
n002 12,
n003 12,
n004 12,
n005 12,
n006 12,
n007 12,
n008 12,
n009 12,
n010 12,
n011 12,
n012 12,
n013 12,
n014 12,
n015 12,
n016 12,
n017 12,
n018 12,
n019 12,
n020 12,
n021 12,
n022 12,
n023 12,
n024 12
```

SGI

Proprietary

Intel mpd

```
sgi@n001:~> ls -la  .mpd.conf
-rw------- 1 sgi sgi 15 2010-06-07 20:25 .mpd.conf

sgi@n001:~> cat .mpd.conf
secretword=sgi

sgi@n001:~> cat mpd.hosts
n001
n002
n003
n004
n005
n006
n007
n008
n009
n010
n011
n012
n013
n014
n015
n016
n017
n018
n019
n020
n021
n022
n023
n024
```

SGI
Proprietary

sgi

Intel mpi structure

```
. ${MODULESHOME}/init/sh
module load intel-mpi/3.0.21

mpdboot -n (nb_hosts) -f $PBS_NODEFILE -r ssh
mpiexec -np (nb_process)
mpdallexit
```

Utilisation de mpt (sgi mpi) et intel mpi

```
sgi@n001:~> which mpirun
/opt/intel/impi/4.0.0.025/intel64/bin/mpirun

sgi@n001:~> module load mpt/2.04

sgi@n001:~>  which mpirun
/opt/sgi/mpt/mpt-2.04/bin/mpirun

sgi@n001:~> mpirun -v
MPI: libxmpi.so 'SGI MPT 2.04  05/26/11 10:20:30'

sgi@n001:~> ls /opt/sgi/mpt/mpt-2.04/bin
mpicc       mpif90        mpt_forward       omplace
mpicxx      mpirun        mpt_hugepage_config  rail-config
mpiexec_mpt  mpt_checkpoint  mpt_restart
```

Mpicc (mpt version)

```
sgi@n001:~> cat /opt/sgi/mpt/mpt-2.04/bin/mpicc
#!/bin/bash
#
# Copyright
#   Copyright (c) 2011 SGI, Inc.
#   All rights reserved.
#

defCC=gcc
MPICC_CC=${MPICC_CC:-${defCC:?}}

LDFLAGS+=" -lmpi"

${MPICC_CC} ${CFLAGS} ${LDFLAGS} $@
```

```
sgi@n001:~> mpicc mpi_hello.c -o mpi_hello_mpt

sgi@n001:~> ldd mpi_hello_mpt
        linux-vdso.so.1 =>  (0x00007ffff7fdf000)
        libmpi.so => /opt/sgi/mpt/mpt-2.04/lib/libmpi.so (0x00007ffff7ca5000)
        libc.so.6 => /lib64/libc.so.6 (0x00007ffff793b000)
        /lib64/ld-linux-x86-64.so.2 (0x00007ffff7fe0000)
        libdl.so.2 => /lib64/libdl.so.2 (0x00007ffff7737000)
        librt.so.1 => /lib64/librt.so.1 (0x00007ffff752e000)
        libgcc_s.so.1 => /lib64/libgcc_s.so.1 (0x00007ffff7316000)
        libcpuset.so.1 => /usr/lib64/libcpuset.so.1 (0x00007ffff7109000)
        libpthread.so.0 => /lib64/libpthread.so.0 (0x00007ffff6eec000)
        libbitmask.so.1 => /usr/lib64/libbitmask.so.1 (0x00007ffff6ce7000)
```

La ligne de compilation avec mpt

```
sgi@n001:~>  mpicc mpi_hello.c -o mpi_hello_mpt
sgi@n001:~> ldd  mpi_hello_mpt
    linux-vdso.so.1 =>  (0x00007ffff7fdf000)
    libmpi.so => /opt/sgi/mpt/mpt-2.04/lib/libmpi.so (0x00007ffff7ca5000)
    libc.so.6 => /lib64/libc.so.6 (0x00007ffff793b000)
    /lib64/ld-linux-x86-64.so.2 (0x00007ffff7fe0000)
    libdl.so.2 => /lib64/libdl.so.2 (0x00007ffff7737000)
    librt.so.1 => /lib64/librt.so.1 (0x00007ffff752e000)
    libgcc_s.so.1 => /lib64/libgcc_s.so.1 (0x00007ffff7316000)
    libcpuset.so.1 => /usr/lib64/libcpuset.so.1 (0x00007ffff7109000)
    libpthread.so.0 => /lib64/libpthread.so.0 (0x00007ffff6eec000)
    libbitmask.so.1 => /usr/lib64/libbitmask.so.1 (0x00007ffff6ce7000)

sgi@n001:~> icc mpi_hello.c -o mpi_hello_mpt_intel -lmpi
sgi@n001:~> ldd  mpi_hello_mpt_intel
    linux-vdso.so.1 =>  (0x00007ffff7fdf000)
    libmpi.so => /opt/sgi/mpt/mpt-2.04/lib/libmpi.so (0x00007ffff7ca5000)
    libm.so.6 => /lib64/libm.so.6 (0x00007ffff7a43000)
    libgcc_s.so.1 => /lib64/libgcc_s.so.1 (0x00007ffff782c000)
    libc.so.6 => /lib64/libc.so.6 (0x00007ffff74ce000)
    libdl.so.2 => /lib64/libdl.so.2 (0x00007ffff72c9000)
    /lib64/ld-linux-x86-64.so.2 (0x00007ffff7fe0000)
    librt.so.1 => /lib64/librt.so.1 (0x00007ffff70c0000)
    libcpuset.so.1 => /usr/lib64/libcpuset.so.1 (0x00007ffff6eb3000)
    libpthread.so.0 => /lib64/libpthread.so.0 (0x00007ffff6c95000)
    libbitmask.so.1 => /usr/lib64/libbitmask.so.1 (0x00007ffff6a91000)
```

Job mpi lancer avec pbs par le wrapper mpt mpiexec_mpt

```
fortune /home/sgi> cat qsub_multispeed_mpiexec_mpt.8cores.sh
#!/bin/sh
#PBS -q workq
#PBS -l select=4:ncpus=4:mpiprocs=4
source /etc/profile.d/modules.sh
module load mpt/2.04
module load intel-cce-12/12.1.0.233
export MPI_GROUP_MAX=512
mpiexec_mpt /home/sgi/multispeed_mpt v h 1024

fortune /home/sgi> qstat -rn

fortune:
```

| | | | | | Req'd | Req'd | Elap | | | |
Job ID	Username	Queue	Jobname	SessID	NDS	TSK	Memory	Time	S	Time
49.fortune	sgi	workq	qsub_multi	15528	4	16	--	--	R	00:00

```
  n001/0*4+n001/1*4+n001/2*4+n002/0*4
```

Comment utiliser $PBS_NODEFILE

```
 fortune /home/sgi> cat  pbsnodefile.sh
#!/bin/sh
#PBS -q workq
#PBS -l select=4:ncpus=4:mpiprocs=4
##PBS -l place="pack:shared"
source /etc/profile.d/modules.sh
module load mpt/2.04
module load intel-cce-12/12.1.0.233
export MPI_GROUP_MAX=512
cat $PBS_NODEFILE
mpiexec_mpt /home/sgi/multispeed_mpt v h 1024

cat pbsnodefile.sh.o56
n001
n001
n001
n001
n001
n001
n001
n001
n001
n001
n001
n001
n002
n002
n002
n002
```

Comment utiliser mpiprocs

```
fortune /home/sgi> cat  pbsnodefile.sh
#!/bin/sh
#PBS -q workq
#PBS -l select=4:ncpus=4
##PBS -l place="pack:shared"
source /etc/profile.d/modules.sh
module load mpt/2.04
module load intel-cce-12/12.1.0.233
export MPI_GROUP_MAX=512
cat $PBS_NODEFILE
mpiexec_mpt /home/sgi/multispeed_mpt v h 1024

cat pbsnodefile.sh.o57
n001
n001
n001
n002
```

Compiler en c++

```
sgi@n001:~> cat test.cpp
#include <iostream>
using namespace std;

#include <omp.h>

int main(int argc, char *argv[])
{
  int th_id, nthreads;
  #pragma omp parallel private(th_id) shared(nthreads)
  {
    th_id = omp_get_thread_num();
    #pragma omp critical
    {
      cout << "Hello World from thread " << th_id << '\n';
    }
    #pragma omp barrier

    #pragma omp master
    {
      nthreads = omp_get_num_threads();
      cout << "There are " << nthreads << " threads" << '\n';
    }
  }

  return 0;
}
```

Compiler en c++ avec intel

```
sgi@n001:~> icpc test.cpp -o testexe -openmp

sgi@n001:~> ldd testexe
     linux-vdso.so.1 =>  (0x00007ffff7fdf000)
     libm.so.6 => /lib64/libm.so.6
(0x00007ffff7d7d000)
     libiomp5.so =>
/opt/intel/composer_xe_2011_sp1.6.233/compiler/lib
/intel64/libiomp5.so (0x00007ffff7aae000)
     libstdc++.so.6 => /usr/lib64/libstdc++.so.6
(0x00007ffff77a3000)
     libgcc_s.so.1 => /lib64/libgcc_s.so.1
(0x00007ffff758c000)
     libpthread.so.0 => /lib64/libpthread.so.0
(0x00007ffff736e000)
     libc.so.6 => /lib64/libc.so.6
(0x00007ffff7010000)
     libdl.so.2 => /lib64/libdl.so.2
(0x00007ffff6e0c000)
     /lib64/ld-linux-x86-64.so.2 (0x00007ffff7fe0000)
```

Compiler en c++ avec gcc

```
sgi@n001:~> module purge

sgi@n001:~> module list
No Modulefiles Currently Loaded.

sgi@n001:~> gcc -fopenmp test.cpp -lstdc++ -o test2exe

sgi@n001:~> ldd test2exe
    linux-vdso.so.1 =>  (0x00007ffff7fdf000)
    libstdc++.so.6 => /usr/lib64/libstdc++.so.6 (0x00007ffff7cc8000)
    libgomp.so.1 => /usr/lib64/libgomp.so.1 (0x00007ffff7abe000)
    libpthread.so.0 => /lib64/libpthread.so.0 (0x00007ffff78a1000)
    libc.so.6 => /lib64/libc.so.6 (0x00007ffff7543000)
    libm.so.6 => /lib64/libm.so.6 (0x00007ffff72ec000)
    /lib64/ld-linux-x86-64.so.2 (0x00007ffff7fe0000)
    libgcc_s.so.1 => /lib64/libgcc_s.so.1 (0x00007ffff70d5000)
    librt.so.1 => /lib64/librt.so.1 (0x00007ffff6ecc000)
```

```
sgi@n001:~> cat test.f
      PROGRAM HELLO
      INTEGER ID, NTHRDS
      INTEGER OMP_GET_THREAD_NUM,
OMP_GET_NUM_THREADS
C$OMP PARALLEL PRIVATE(ID)
      ID = OMP_GET_THREAD_NUM()
      PRINT *, 'HELLO WORLD FROM THREAD', ID
C$OMP BARRIER
      IF ( ID .EQ. 0 ) THEN
        NTHRDS = OMP_GET_NUM_THREADS()
        PRINT *, 'THERE ARE', NTHRDS, 'THREADS'
      END IF
C$OMP END PARALLEL
      END
```

SGI
Proprietary

Compiler en fortran77 avec intel

```
sgi@n001:~> module load intel-fce-12/12.1.0.233

sgi@n001:~> ifort test.f -openmp -o test3exe

sgi@n001:~> ldd  test3exe
    linux-vdso.so.1 =>  (0x00007ffff7fdf000)
    libm.so.6 => /lib64/libm.so.6 (0x00007ffff7d7d000)
    libiomp5.so => /opt/intel/composer_xe_2011_sp1.6.233/compiler/lib/intel64/libiomp5.so
(0x00007ffff7aae000)
    libpthread.so.0 => /lib64/libpthread.so.0 (0x00007ffff7891000)
    libc.so.6 => /lib64/libc.so.6 (0x00007ffff7533000)
    libgcc_s.so.1 => /lib64/libgcc_s.so.1 (0x00007ffff731b000)
    libdl.so.2 => /lib64/libdl.so.2 (0x00007ffff7117000)
    /lib64/ld-linux-x86-64.so.2 (0x00007ffff7fe0000)
```

Compiler avec mkl

```
fortune /home/sgi> cat  prog_mkl.f
    CALL ZTRSM
    END

fortune /home/sgi> module load intel-fce-12/12.1.0.233

fortune /home/sgi> ifort prog_mkl.f -lmkl_intel_lp64 -lmkl_sequential -lmkl_core -o prog_mkl
ld: cannot find -lmkl_intel_lp64

fortune /home/sgi> module load intel-mkl-10/10.3.6

fortune /home/sgi> ifort prog_mkl.f -lmkl_intel_lp64 -lmkl_sequential -lmkl_core -o prog_mkl
```

Job intel mpi lancé avec pbs

```sh
#!/bin/sh

# SGI

#PBS -N sgi_mpitest
#PBS -e sgi_mpitest.err
#PBS -o sgi_mpitest.log
#PBS -q workq
#PBS -l place=scatter
#PBS -l select=8:ncpus=12:mpiprocs=12

# shm, sock, ssm, rdma, rdssm
FABRIC=rdma

CORES=$[ `cat $PBS_NODEFILE | wc -l` ]
NODES=$[ `uniq $PBS_NODEFILE | wc -l` ]

cd $PBS_O_WORKDIR

printf "#############################################\n";
printf "Current time is: `date`\n";
printf "Current PBS work directory is: $PBS_O_WORKDIR\n";
printf "Current PBS queue is: $PBS_O_QUEUE\n";
printf "Current PBS job ID is: $PBS_JOBID\n";
printf "Current PBS job name is: $PBS_JOBNAME\n";
printf "PBS stdout log is: $PBS_O_WORKDIR/sgi_mpitest.err\n";
printf "PBS stderr log is: $PBS_O_WORKDIR/sgi_mpitest.log\n";
printf "Fabric interconnect selected is: $FABRIC\n";
printf "This jobs will run on the following ($CORES) processors:\n";
echo `cat $PBS_NODEFILE`
```

Job intel mpi lancé avec pbs

```
. /etc/profile.d/modules.sh

compiler="gcc-$(module avail 2>&1 | grep gcc | awk -F'gcc-' '{print $2}' | awk '{print $1}' | sort -u)"
mpi=intel-4.0


TBEGIN=`echo "print time();" | perl`

module load compiler/$compiler mpi/$mpi && echo successfully loaded modules

printf "MPI compiler/implementation is: $compiler/$mpi\n";
printf "MPI run command location is: `which mpirun`\n";
printf "##############################################\n";

mpdboot --file=${PBS_NODEFILE} -n $NODES -r ssh -v
mpiexec -genv I_MPI_DEVICE $FABRIC -n $CORES /opt/intel/sgi/intel_mpi_hello_world
mpdallexit

TEND=`echo "print time();" | perl`

printf "##############################################\n";
printf "+++ Job finished: `date`\n"
printf "+++ Job walltime: `expr $TEND - $TBEGIN`\n"
printf "##############################################\n";
```

Monitoring du job intel mpi

```
fortune /home/sgi> qstat -rn

fortune:
                                    Req'd  Req'd   Elap
Job ID        Username Queue    Jobname    SessID NDS TSK Memory Time  S Time
--------------- -------- -------- ---------- ------ --- --- ------ ----- - -----
53.fortune     sgi     workq   sgi_mpites 16534  8  96   --   --  R 00:00
  n001/0*12+n002/0*12+n003/0*12+n004/0*12+n005/0*12+n006/0*12+n007/0*12
  +n008/0*12
```

Omplace (jobs hybrides openmp/mpt)

omplace -nt overrides OMP_NUM_THREADS

Ex1. Pour lancer 2 MPI process avec 4 threads par process

> **mpirun -np 2 omplace -nt 4 -vv a.out**

> **The threads would be place as follows:**

>> **rank 0 thread 0 on CPU 0**
>> **rank 0 thread 1 on CPU 1**
>> **rank 0 thread 2 on CPU 2**
>> **rank 0 thread 3 on CPU 3**
>> **rank 1 thread 0 on CPU 4**
>> **rank 1 thread 1 on CPU 5**
>> **rank 1 thread 2 on CPU 6**
>> **rank 1 thread 3 on CPU 7**

Ex2. Pour lancer un job MPI sur les numeros paires CPU sockets
avec 2 CPU cores par socket:

> **mpirun -np 32 omplace -c 0-63:bs=2+st=4 a.out**

Ex3. Pour lancer un job 2 MPI a.out avec 4 threads par process et 3 MPI
b.out process avec 2 threads par process, avec l'option -b
pour bouger le 2 ème groupe de threads sur d'autres cpus

> **mpirun -np 2 omplace -nt 4 a.out : -np 3 omplace -b 8 -nt 2 b.out**

sgi

Omplace (jobs hybrides openmp/mpt)

```
#!/bin/sh
#PBS -q workq
#PBS -l select=1:ncpus=12:mpiprocs=4
source /etc/profile.d/modules.sh
module load mpt/2.04
module load intel-cce-12/12.1.0.233
export MPI_GROUP_MAX=512
cat $PBS_NODEFILE
export OMP_NUM_THREADS=3
time -p mpiexec_mpt omplace -nt 3
/home/sgi/mpiomp
```

SGI

sgi

Omplace (jobs hybrides openmp/mpt)

```
21646 sgi      20  0  126g 3852 2124 R  100  0.0  2:34.64  1 mpiomp
21650 sgi      20  0  126g 3852 2124 R  100  0.0  2:34.65  2 mpiomp
21647 sgi      20  0  126g 3852 2124 R  100  0.0  2:34.67  4 mpiomp
21653 sgi      20  0  126g 3852 2124 R  100  0.0  2:34.64  5 mpiomp
21642 sgi      20  0  126g 3848 2124 R  100  0.0  2:34.67  6 mpiomp
21652 sgi      20  0  126g 3848 2124 R  100  0.0  2:34.68  7 mpiomp
21654 sgi      20  0  126g 3848 2124 R  100  0.0  2:34.67  8 mpiomp
21643 sgi      20  0  126g 3852 2124 R  100  0.0  2:34.68  9 mpiomp
21651 sgi      20  0  126g 3852 2124 R  100  0.0  2:34.68 10 mpiomp
21655 sgi      20  0  126g 3852 2124 R  100  0.0  2:34.68 11 mpiomp
21640 sgi      20  0  126g 3852 2124 R   99  0.0  2:34.53  0 mpiomp
21641 sgi      20  0  126g 3852 2124 R   99  0.0  2:34.35  3 mpiomp
```

les outils contenus dans SGI Foundation

SGI Foundation :

Esp (outil sgi de monitoring d'évenements hardware et software)
esp-4.4-1530sgi704r1.sles11

Lk license management (sgi)
lkSGI-2.2-sgi704r1.sles11.x86_64.rpm

LSI management des cartes FC et SAS
lsiutil-1.0-sgi704r1.sles11

Management du scsi
sgtools-1.2-sgi704r1.sles11

SGI snmp agent
sgi-snmpagent-xe-1.0-sgi704r1.sles11

SGI outils de support (/usr/sbin/cluster_info, netgather, nodetrace, pm_cluster_info_gather, sgi_collect_dump, system_info_gather)
sgi-support-tools-0.2-sgi704r1.sles11

SGI script de configuration système
sgi-base-configuration-1.4-sgi704r1.sles11 (UV)

Outils SGI de tracabilité d'erreurs mémoires
memlog-2.0-sgi704r2.sles11

PCP
pcp-3.5.0-1sgi704r1.sles11

```
fortune:/mnt/iso3/sgi/x86_64 # ls
Modules-3.1.6-39.9sgi704r1.sles11.x86_64.rpm
chkfeature-1.0-sgi704r1.sles11.x86_64.rpm
chkfeature-libs-1.0-sgi704r1.sles11.x86_64.rpm
chkfeature-python-1.0-sgi704r1.sles11.x86_64.rpm
esp-4.4-1530sgi704r1.sles11.x86_64.rpm
lkSGI-2.2-sgi704r1.sles11.x86_64.rpm
lkSGI-java-2.2-sgi704r1.sles11.x86_64.rpm
lkSGI-mgr-2.2-sgi704r1.sles11.x86_64.rpm
lkSGI-python-2.2-sgi704r1.sles11.x86_64.rpm
lkSGI-report-2.2-sgi704r1.sles11.x86_64.rpm
lsiutil-1.0-sgi704r1.sles11.x86_64.rpm
memlog-2.0-sgi704r2.sles11.x86_64.rpm
ml-1.1.2-sgi704r1.sles11.x86_64.rpm
ml_audio-1.1.2-sgi704r1.sles11.x86_64.rpm
ml_nullxcode-1.1.2-sgi704r1.sles11.x86_64.rpm
ml_sgc-1.1.3-sgi704r2.sles11.x86_64.rpm
ml_ustsource-1.1.2-sgi704r1.sles11.x86_64.rpm
ml_v4l-1.1.2-sgi704r1.sles11.x86_64.rpm
msr-tools-1.2-1sgi704r1.sles11.x86_64.rpm
pcp-3.5.0-1sgi704r1.sles11.x86_64.rpm
pcp-gui-1.4.5-1sgi704r1.sles11.x86_64.rpm
pcp-libs-3.5.0-1sgi704r1.sles11.x86_64.rpm
pcp-libs-devel-3.5.0-1sgi704r1.sles11.x86_64.rpm
pcp-pmda-cluster-1.0.4-2sgi704r1.sles11.x86_64.rpm
pcp-pmda-cluster-client-1.0.4-2sgi704r1.sles11.x86_64.rpm
pcp-pmda-infiniband-1.1.1-1sgi704r1.sles11.x86_64.rpm
perl-PCP-MMV-3.5.0-1sgi704r1.sles11.x86_64.rpm
perl-PCP-PMDA-3.5.0-1sgi704r1.sles11.x86_64.rpm
python-networkx-1.0.1-1sgi704r1.sles11.x86_64.rpm
python-pexpect-2.3-23.17sgi704r1.sles11.x86_64.rpm
pyzmq-2.0.8-1sgi704r1.sles11.x86_64.rpm
sgtools-1.2-sgi704r1.sles11.x86_64.rpm
zeromq-2.0.10-1sgi704r1.sles11.x86_64.rpm

sgi-base-configuration-1.4-sgi704r1.sles11.x86_64.rpm
sgi-build-key-1.0-sgi704r1.sles11.x86_64.rpm
sgi-cme_migrate-kmp-default-1.0_2.6.32.12_0.7-sgi704r2.sles11.x86_64.rpm
sgi-cme_migrate-kmp-trace-1.0_2.6.32.12_0.7-sgi704r2.sles11.x86_64.rpm
sgi-hwperf-kmp-default-1.0_2.6.32.12_0.7-sgi704r2.sles11.x86_64.rpm
sgi-hwperf-kmp-trace-1.0_2.6.32.12_0.7-sgi704r2.sles11.x86_64.rpm
sgi-ml_sgc-kmp-default-1.1.3_2.6.32.12_0.7-sgi704r2.sles11.x86_64.rpm
sgi-snmpagent-xe-1.0-sgi704r1.sles11.x86_64.rpm
sgi_irqbalance-1.0-sgi704r1.sles11.x86_64.rpm
```

SGI
Proprietary

```
fortune:/usr/sbin # cluster_info
Checking n02 Done.

The fortune cluster with 24 cnodes has the following
configuration based on Fri Oct 28 23:01:19 CEST 2011 snapshot.

 Product Type: SGI
 Product Name: C1104-2TY9
   Board MFG: Silicon Graphics International        Board Model: X8DTT-IBQ
   Board MFG: Silicon Graphics International        Board Model: X8DTT-IBQF
   BIOS Date: 01/25/2011           HyperThread: Enabled
   Processors: 2 x 6 Cores Intel(R) Xeon(R) CPU X5675 @ 3.07GHz 12288KB Cache
   Total Mem: 99197504 KB    Speed: 1333 MHz (0.8 ns)
   OS Release: SUSE Linux Enterprise Server 11 (x86_64) VERSION = 11 PATCHLEVEL = 1
   Kernel Ver: 2.6.32.45-0.3-default
 SGI Software: SGI Accelerate 1.2, Build 704r5.sles11-1103212004

Following are the file system types detected.
--------------------------------------------------------------------------------
/home on /home type nfs
/opt/intel on /opt/intel type nfs
/scratch on /scratch type nfs
/software on /software type nfs
```

SGI

les outils contenus dans SGI performance suite

sgi

SGI Performance suite

SGI Accelarate :

Cpusets, optimise l'utilisation CPU et mémoire (création de containeurs)

NumaTools outils de placement

Librairie FFIO (runtime I/O performance)

Performance co-pilot (pcp-sgi) monitorer les performances du système

libGRU (UV)

SGISolve(librairies, CRAY code portabilité)

SGI MPT : (mpi de SGI)

Perfcatcher outil de profiling mpi

MPIinside outil de profiling et de mesure de performance

SGI MPI PerfBoost, permet de lancer des applis Platform MPI, Intel MPI, Open MPI et MPICH au travers de mpt.

SGI UPC : (compilateur sgi)

Rpm –qa | grep 704

fortune:/mnt/iso3/sgi/x86_64 # rpm -qa | grep 704
sgi-base-configuration-1.4-sgi704r1.sles11
freeipmi-1.0.3-0.1sgi704rp1.sles11
powerman-2.3.9-0.1sgi704rp3.sles11
Modules-3.1.6-39.9sgi704rp2.sles11
pcp-gui-1.4.5-1sgi704r1.sles11
sgi-mpt-2.04-sgi704rp18.sles11
pcp-pmda-infiniband-1.1.1-1sgi704r1.sles11
sgi-perfsuite-docs-1-sgi704r3.sles11
gltt-2.5.2-704.28
sgi-upc-release-1.2-sgi704r5.sles11
sgimc-1.4.0-sgi704rp74.sles11
conman-0.2.6-0.1sgi704rp3.sles11
sgi-build-key-1.0-sgi704rp2.sles11
discover-common-2.4-sgi704rp2.sles11
discover-server-2.4-sgi704rp2.sles11
memacct-1.2-sgi704r1.sles11
sgimc-server-1.4.0-sgi704rp74.sles11
sgtools-1.2-sgi704r1.sles11
sgi-xpmem-kmp-default-1.6_2.6.32.12_0.7-sgi704rp12.sles11
lkSGI-java-2.2-sgi704rp2.sles11
sgi-release-704-sgi704r5.sles11
sgi-mpi-release-1.2-sgi704r5.sles11
chkfeature-python-1.0-sgi704r1.sles11
libcpuset-1.0sgi-sgi704r1.sles11
lkSGI-python-2.2-sgi704r1.sles11
perfcatcher-25-sgi704r6.sles11
lkSGI-mgr-2.2-sgi704r1.sles11
sgi-common-python-2.4-sgi704rp2.sles11
libFFIO-1.0-sgi704r1.sles11
MPInside-3.1-sgi704r1.sles11
pdsh-mod-machines-2.18-1sgi704rp2.sles11
sgi-mpt-shmem-2.04-sgi704r6.sles11
java-1.6.0-sun-fonts-1.6.0.25-sgi704rp3.sles11
jfreechart-1.0.10-sgi704rp74.sles11
xpmem-1.6-sgi704r2.sles11
lkSGI-2.2-sgi704rp2.sles11
libbitmask-2.0sgi-sgi704r1.sles11

sgi-foundation-release-2.4-sgi704r5.sles11
pdsh-rcmd-exec-2.18-1sgi704rp2.sles11
chkfeature-1.0-sgi704r1.sles11
sgi-support-tools-0.2-sgi704r1.sles11
pcp-sgi-3.1.1-sgi704r2.sles11
mpitests_mpt-2.0-705sgi704r6.sles11
lkSGI-report-2.2-sgi704r1.sles11
jcommon-1.0.16-sgi704rp74.sles11
sgi-foundation-docs-2-sgi704r3.sles11
sgi-procset-1.0-sgi704r1.sles11
sgimc-tftpboot-1.4.0-sgi704rp74.sles11
cpuset-utils-2.0-sgi704r1.sles11
sgi-xvma-kmp-default-1.0_2.6.32.12_0.7-sgi704r2.sles11
sgi-mpt-fs-2.04-sgi704rp18.sles11
pcp-libs-3.5.0-1sgi704r1.sles11
sgi-numatools-kmp-default-2.0_2.6.32.12_0.7-sgi704r2.sles11
sgi-accelerate-release-1.2-sgi704r5.sles11
chkfeature-libs-1.0-sgi704r1.sles11
discover-2.4-sgi704rp2.sles11
perfboost-1.04-sgi704r6.sles11
mkelfImage-2.7-0.1sgi704rp2.sles11
java-1.6.0-sun-alsa-1.6.0.25-sgi704rp3.sles11
MPInside-libs-3.1-sgi704r1.sles11
esp-4.4-1530sgi704r1.sles11
sgi-cm-agnostic-0.5-sgi704rp2.sles11
cattr-2.4-sgi704rp2.sles11
pdsh-rcmd-ssh-2.18-1sgi704rp2.sles11
sgimc-tftp-1.4.0-sgi704rp74.sles11
sgi-upc-1.04-sgi704r6.sles11
sgi-upc-devel-1.04-sgi704r6.sles11
java-1.6.0-sun-plugin-1.6.0.25-sgi704rp3.sles11
sgi-arraysvcs-3.7-sgi704r1.sles11
sgi-management-device-2.4-sgi704rp2.sles11
sgi-numatools-kmp-trace-2.0_2.6.32.12_0.7-sgi704r2.sles11
java-1.6.0-sun-1.6.0.25-sgi704rp3.sles11
pdsh-2.18-1sgi704rp2.sles11
memlog-2.0-sgi704rp9.sles11
numatools-2.0-sgi704rp8.sles11
pcp-3.5.0-1sgi704r1.sles11

SGI

SGI

```
fortune /home/sgi> cat /etc/SuSE-release
SUSE Linux Enterprise Server 11 (x86_64)
VERSION = 11
PATCHLEVEL = 1

fortune /home/sgi> cat /etc/sgi-foundation-release
SGI Foundation Software 2.4, Build 704r5.sles11-1103212004

fortune /home/sgi> cat /etc/sgi-release
SGI Performance Suite 1.2, Build 704r5.sles11-1103212004

fortune /home/sgi> cat /etc/sgi-accelerate-release
SGI Accelerate 1.2, Build 704r5.sles11-1103212004

fortune /home/sgi> cat /etc/sgi-mpi-release
SGI MPI 1.2, Build 704r5.sles11-1103212004

fortune /home/sgi> cat  /etc/sgi-upc-release
SGI UPC 1.2, Build 704r5.sles11-1103212004

Fortune /home/sgi> cat /etc/sgi
SGI: SMC-CMK-MI
Revision: 1.4.0_3
Description: SLES 11.1
Transaction Log: 4
Date: Fri Aug  5 14:25:52 CDT 2011
```

SGI

Proprietary

```
fortune chkfeature/sbin> ./chkfeature
sgi-accelerate     status          on
sgi-array          status          on
sgi-cpuset         status          on
sgi-foundation     status          on
sgi-libbitmask     status          on
sgi-libcpuset      status          on
sgi-lk             status          on
sgi-memacct        status          on
sgi-mpi            status          on
sgi-mpt            status          on
sgi-upc            status          on
```

SGI
Proprietary

sgi

```
fortune:/mnt/iso3/sgi/x86_64 # rpm -ql sgi-upc-1.04-sgi704r6.sles11
/opt/sgi/upc/upc-1.04/lib
/opt/sgi/upc/upc-1.04/lib/libupc.so.0
/usr/share/modules/modulefiles/sgi-upc/1.04
fortune:/mnt/iso3/sgi/x86_64 # rpm -ql sgi-upc-devel-1.04-
sgi704r6.sles11
/opt/sgi/upc/upc-1.04/bin
/opt/sgi/upc/upc-1.04/bin/ir_b2a
/opt/sgi/upc/upc-1.04/bin/sgiupc
/opt/sgi/upc/upc-1.04/include
/opt/sgi/upc/upc-1.04/include/bits/mathinline.h
/opt/sgi/upc/upc-1.04/include/float.h
/opt/sgi/upc/upc-1.04/include/ncc/whirl2c.h
/opt/sgi/upc/upc-1.04/include/ncc/whirl2c_upc.h
/opt/sgi/upc/upc-1.04/include/stdarg.h
/opt/sgi/upc/upc-1.04/include/stdbool.h
/opt/sgi/upc/upc-1.04/include/stddef.h
/opt/sgi/upc/upc-1.04/include/upc.h
/opt/sgi/upc/upc-1.04/include/upc/limits.h
/opt/sgi/upc/upc-1.04/include/upc_collective.h
/opt/sgi/upc/upc-1.04/include/upc_io.h
/opt/sgi/upc/upc-1.04/include/upc_relaxed.h
/opt/sgi/upc/upc-1.04/include/upc_strict.h
/opt/sgi/upc/upc-1.04/include/varargs.h
```

```
/opt/sgi/upc/upc-1.04/lib
/opt/sgi/upc/upc-1.04/lib/libupc.so
/opt/sgi/upc/upc-1.04/lib/upc_init.o
/opt/sgi/upc/upc-1.04/man/man1
/opt/sgi/upc/upc-1.04/man/man1/sgiupc.1
/opt/sgi/upc/upc-1.04/phases
/opt/sgi/upc/upc-1.04/phases/be
/opt/sgi/upc/upc-1.04/phases/be.so
/opt/sgi/upc/upc-1.04/phases/driver
/opt/sgi/upc/upc-1.04/phases/fec
/opt/sgi/upc/upc-1.04/phases/inline
/opt/sgi/upc/upc-1.04/phases/ipl.so
/opt/sgi/upc/upc-1.04/phases/lno.so
/opt/sgi/upc/upc-1.04/phases/whirl2c
/opt/sgi/upc/upc-1.04/phases/whirl2c.so
/opt/sgi/upc/upc-1.04/phases/whirl2c_be
/opt/sgi/upc/upc-1.04/phases/wopt.so
/opt/sgi/upc/upc-1.04/phases/xeon.so
/usr/share/modules/modulefiles/sgi-upc-devel/1.04
```

perfboost

```
fortune sgi/perfboost> ls bin
perfboost

fortune sgi/perfboost> ls lib
libfmpi.so          libmpi_mt.so.3.2        libperfboost_impi.so
libmpichf90.so.1.0   libmpi_mt.so.4
libperfboost_isub.so
libmpichfstub.so.1.0  libmpio.so.1
libperfboost_mpmpi.so
libmpich.so.1.0       libmpi.so.0
libperfboost_mpsub.so
libmpich.so.1.2       libmpi.so.1             libperfboost_null.so
libmpi_f77.so.0       libmpi.so.3.2
libperfboost_ompi.so
libmpi_f90.so.0       libmpi.so.4             libperfboost_osub.so
libmpigf.so.3.2       libopen-pal.so.0
libperfboost_pmpi.so
libmpigf.so.4         libopen-rte.so.0
libperfboost_psub.so
libmpiif.so.3.2       libperfboost_empty_mt.so  libpmpi.so
libmpiif.so.4         libperfboost_empty.so
```

perfboost

```
% module load mpt
% module load perfboost
% mpirun -np 32 perfboost -impi a.out arg1
% mpiexec_mpt perfboost -pmpi b.out arg1
% mpirun host1 32, host2 64 perfboost -impi c.out arg1 arg2


man -M /opt/sgi/man perfboost
```

http://techpubs.sgi.com (007-3773-018) Message Passing Toolkit (MPT) User's Guide Chapter 6

MPInside

setenv LD_PRELOAD /opt/sgi/mpinside/lib/libMPInside_mpt.so

mpirun -np 128 MPInside apps apps_arg

setenv LD_PRELOAD /opt/sgi/mpinside/lib/libMPInside_mpt.so:/usr/lib64/libFFIO.so

mpirun -np 128 MPInside apps apps_arg

man -M /opt/sgi/MPInside/3.1/man MPInside

http://techpubs.sgi.com (007-5780-001) MPInside Reference Guide

perfcatcher (mpt et intel mpi)

http://techpubs.sgi.com (007-3773-018) Message Passing Toolkit (MPT) User's Guide Chapter 9

mpirun -np 64 perfcatch a.out arg1 l

mpirun host1 32, host2 64 perfcatch a.out arg1

Avec intel (-i option)

mpiexec -np 64 perfcatch -i a.out arg1

man -M /opt/sgi/man perfcatch

sgi

Outils de monitoring

Commandes pcp

pmchart

dkvis

mpvis

osvis

nfsvis

clustervis

nodeinfo

diskstat

topdisk

topsys

pmgcluster –t 2

pminfo –T (metrics)

sgi

Pcp pmdas

```
n001:/var/lib/pcp/pmdas # ls
apache   gpsd   ib      lustrecomm mounts     news     process shping    vmware
bonding  kvm       mailq      mysql     pdns      roomtemp summary    weblog
cisco    linux     memcache   named     pmcd      samba     systemtap zimbra
dbping   lmsensors mmv            netfilter postfix sendmail  trace


fortune:/var/lib/pcp/pmdas # rpm -ql pcp-pmda-infiniband-1.1.1-1sgi704r1.sles11
/usr/share/man/man1/pmdaib.1.gz
/var/lib/pcp/pmdas/ib/Install
/var/lib/pcp/pmdas/ib/Remove
/var/lib/pcp/pmdas/ib/domain.h
/var/lib/pcp/pmdas/ib/help
/var/lib/pcp/pmdas/ib/help.dir
/var/lib/pcp/pmdas/ib/help.pag
/var/lib/pcp/pmdas/ib/pmda_ib.so
/var/lib/pcp/pmdas/ib/pmdaib
/var/lib/pcp/pmdas/ib/pmns
/var/lib/pcp/pmdas/ib/root
```

SGI

sgi

Bonus

Altix XE OS

System proxy:

```
fortune:~ # cat /etc/sysconfig/proxy
## Path: Network/Proxy
## Description:
## Type: yesno
## Default: no
## Config: kde,profiles
# Enable a generation of the proxy settings to the profile.
# This setting allows to turn the proxy on and off while # preserving the particular proxy setup.
PROXY_ENABLED="yes"
## Type: string ## Default: ""
# # Some programs (e.g. lynx, arena and wget) support proxies, if set in
# the environment. SuSEconfig can add these environment variables to
# /etc/SuSEconfig/* (sourced by /etc/profile etc.) –
# See http://portal.suse.com/sdb/en/1998/01/lynx_proxy.html for more details.
# Example: HTTP_PROXY="http://proxy.provider.de:3128/"
HTTP_PROXY="http://proxy1-rech.univ-valenciennes.fr:3128/"
## Type: string
## Default: ""
# Some programs (e.g. lynx, arena and wget) support proxies, if set in
# the environment. SuSEconfig can add these environment variables to
# /etc/SuSEconfig/* (sourced by /etc/profile etc.) –
# this setting is for https connections
HTTPS_PROXY="http://proxy1-rech.univ-valenciennes.fr:3128/"
## Type: string
## Default: ""
# Example: FTP_PROXY="http://proxy.provider.de:3128/"
FTP_PROXY=""
## Type: string
## Default: ""
# Example: GOPHER_PROXY="http://proxy.provider.de:3128/"
GOPHER_PROXY=""
## Type: string(localhost)
## Default: localhost
# Example: NO_PROXY="www.me.de, do.main, localhost"
NO_PROXY="localhost, 127.0.0.1"
```

SGI

Proprietary

sgi

Operating System:

```
fortune:~/sgi # zypper ls
# | Alias                                    | Name                                     | Enabled | Refresh | Type
--+------------------------------------------+------------------------------------------+---------+------ -----+---------
1 | nu_novell_com                            | nu_novell_com                            | Yes     | Yes     | ris
2 | SLES11                                   | SLES11                                   | Yes     | No      | yast2
3 | SLES11-Supplimentary                     | SLES11-Supplimentary                     | Yes     | No      | yast2
4 | SUSE Linux Enterprise Server 11 SP1      | SUSE Linux Enterprise Server 11 SP1      | Yes     | No      | yast2
5 | accelerate-1                             | accelerate-1                             | Yes     | No      | yast2
6 | foundation-2                             | foundation-2                             | Yes     | No      | yast2
7 | mpi-1                                    | mpi-1                                    | Yes     | No      | yast2
8 | patch10803                               | patch10803                               | Yes     | No      | plaindir
9 | upc-1                                    | upc-1                                    | Yes     | No      | yast2
```

Operating System:

fortune:~/sgi # zypper ls -d

| https://nu.novell.com/?credentials=NCCcredentials
| dir:///mnt/iso1
| dir:///mnt/iso2
| iso:///?iso=SLES-11-SP1-DVD-x86_64-GM-DVD1.iso&url=dir:///opt/sgi/Factory-Install/ISO/SLES-11-SP1&alias=sles-11-sp-1
| iso:///?iso=accelerate-1.2-cd1-media-sles11-x86_64.iso&url=dir:///opt/sgi/Factory-Install/ISO/accelerate-1&alias=accelerate-1
| iso:///?iso=foundation-2.4-cd1-media-sles11-x86_64.iso&url=dir:///opt/sgi/Factory-Install/ISO/foundation-2&alias=foundation-2 7
| iso:///?iso=mpi-1.2-cd1-media-sles11-x86_64.iso&url=dir:///opt/sgi/Factory-Install/ISO/mpi-1&alias=mpi-1
| dir:///%2Fopt/sgi/Factory-Install/Patches/patch10803/RPMS
| iso:///?iso=upc-1.2-cd1-media-sles11-x86_64.iso&url=dir:///opt/sgi/Factory-Install/ISO/upc-1&alias=upc-1

Chercher un paquage

```
sgi@n001:~> zypper se | grep infiniband
 | ib-bonding                  | infiniband bonding tool                                  | package
 | ib-bonding                  | infiniband bonding tool                                  | srcpackage
 | infiniband-diags            | OpenIB InfiniBand Diagnostic Tools                       | srcpackage
i | infiniband-diags           | OpenIB InfiniBand Diagnostic Tools                       | package
 | infiniband-diags-debuginfo  | Debug information for package infiniband-diags           | package
 | infiniband-diags-debugs     | Debug sources for package infiniband-diags               | package
 | infiniband-diags-devel      | OpenIB InfiniBand Diagnostic Tools SDK                   | package
i | pcp-pmda-infiniband         | PMDA for collecting statistics from Infiniband HCAs and Switches | package
 | pcp-pmda-infiniband         | PMDA for collecting statistics from Infiniband HCAs and Switches | srcpackage
```

SGI

Proprietary

sgi

Job abaqus mono

```
fortune sgi/bench> cat myenv2.job
#/bin/sh
#PBS -l select=1:ncpus=1
#PBS -q workq@fortune
#PBS -N bench
#PBS -j eo
date
env|sort
cd $PBS_O_WORKDIR
input=bench.inp
output=out.$PBS_JOBNAME
export TMPDIR=/tmp/$PBS_JOBID
mkdir -p $TMPDIR
time -p /software/ABAQUS6.11/Commands/abaqus \
j=${output} interactive input=${input} cpus=$(wc -l <$PBS_NODEFILE) scratch=$TMPDIR  >${output}.log 2>&1
ls -l $TMPDIR
grep  -A 7 'M MEMORY' *.dat
grep  -A 7 'JOB TIME ' *.dat
```

Job abaqus parallele

```
fortune sgi/bench> cat bench8.job
#/bin/sh
#PBS -l select=1:ncpus=8:mpiprocs=8  -l place=excl:scatter
#PBS -q workq@fortune
#PBS -N bench8
#PBS -j eo
date
env|sort
cd $PBS_O_WORKDIR
input=3_817_0.4.inp
output=out.$PBS_JOBNAME
export TMPDIR=/tmp/$PBS_JOBID
mkdir -p $TMPDIR
time -p /software/ABAQUS6.11/Commands/abaqus double \
j=${output} interactive input=${input} cpus=$(wc -l <$PBS_NODEFILE) scratch=$TMPDIR  >${output}.log 2>&1
ls -l $TMPDIR
grep  -A 7 'M MEMORY' *.dat
grep  -A 7 'JOB TIME ' *.dat
```

Exemple de job fluent (lancer au travers de pbs)

```ksh
#!/bin/ksh
#PBS -l walltime=0:30:00
#PBS -l select=4:ncpus=24:mpiprocs=12
#PBS -j oe
#PBS -N fluent

. /usr/share/modules/init/ksh
module list
env|sort
input=fl5l1.jou
cd $PBS_O_WORKDIR
ls -l ${PBS_NODEFILE}
nprocs=$(wc -l <${PBS_NODEFILE})
time -p /store/mdk/FLUENT/fluent12.1.9/ansys_inc/v121/fluent/bin/fluent 3ddp \
-ssh  -pib -cnf=${PBS_NODEFILE}  -t$nprocs  -i $input -g
#> log.$nprocs 2>&1
```

FIN