

Le système de batch

Le système de batch

Les groupes de serveurs

Les groupes de serveurs avec proc Intel

Groupe de serveur	Type	Fréquence	Processeur	#coeurs * x2 logiques	Mémoire	Disque /scratch	Réseau
c6100lin24ib	Bi-socket	2,66 Ghz	Intel Westmere	6 physiques* par socket 12 au total*	24 GO 2 GO/coeur	250 GO	IB QDR
r410lin24ibA	Bi-socket	2,66 Ghz	Intel Nehalem	4 physiques* par socket 8 au total*	24 GO 3 GO/coeur	140 GO	IB QDR
r410lin24ibB	Bi-socket	2,66 Ghz	Intel Nehalem	4 physiques* par socket 8 au total*1	24 GO 3 GO/coeur	140 GO	IB QDR
r410lin24gbC	Bi-socket	2,66 Ghz	Intel Nehalem	4 physiques* par socket 8 au total*	24 GO 3 GO/coeur	140 GO	GB
x41zlin16ib	Bi-socket	2,66 Ghz	Intel Hppertown	4 physiques par socket 8 au total	16 GO 2 GO/coeur	400 GO	IB DDR
x41zlin32gb	Bi-socket	2,66 Ghz	Intel Hppertown	4 physiques par socket 8 au total	32 GO 4 GO/coeur	900 GO	GB

Les groupes de serveurs avec proc AMD

Groupe de serveur	Type	Fréquence	Processeur	#coeurs	Mémoire	Disque /scratch	Réseau
v40zlin16ib	Quadri-socket	2,4 Ghz	Amd Opteron	1 physique par socket 4 au total	16G0 4G0/coeur	73 GO	IB SDR
v22zlin4ib	Bi-socket	2,2 Ghz	Amd Opteron	2 physiques par socket 4 au total	4G0 1G0/coeur	73 GO	IB SDR
dl165lin24ib	Bi-socket	2,66 Ghz	Amd Shangäi	4 physiques par socket 8 au total	24 GO 2,5 GO/ coeur	250 GO	IB DDR
dl175lin32gb	Bi-socket	2,66 Ghz	Amd Istanbul	6 physiques par socket 12 au total	32 GO 2,5 GO/ coeur	250 GO	GB
r815lin128ib	Quadri-socket	2,3 Ghz	Amd Magny-Cours	8 physiques par socket 32 au total	128 GO 4 GO/coeur	1 TO	IB QDR

Convention du nom des queues batch

- Exemples
 - r410lin24ibA
 - dl175lin32gb

r410	lin	24	ib	A
dl175	lin	32	gb	
Type de serveur	OS	Mémoire	Réseau dédié aux calculs	Numéro si nécessaire

Le système de batch

Utilisation des ressources

Utilisation optimale des ressources

Pour utiliser au mieux les ressources, il convient de bien remplir les serveurs.

Deux moyens:

- remplir au mieux
- remplir par multiples de n cœurs ou n est le nombre de cœurs physiques par serveur.

Utilisation optimale des ressources

Remplir au mieux mène rapidement à un morcellement entre les serveurs des applications parallélisées via mpi.

Donc le remplissage au mieux n'est pas mis en place, seul le remplissage par multiple d'un serveur entier est mis en place.

Le système de batch

Les queues et environnements
parallèles ou sous-queues

Les queues batch et leurs environnements de programmation (ou sous-queues)

	mpi32 namd32	mpi12 namd12	mpi8 namd8	openmp12	openmp8	gaussian12	gaussian8	gaussian4
c6100lin24ib		*		*		*		
r410lin24ibA	*		*					
r410lin24ibB			*					
r410lin24ibC			*		*			
dl175lin32gb		*		*		*		
dl165lin24ib			*		*		*	
dl165adf			*					
x41zlin24ib			*		*		*	
x41zlin32gb			*		*		*	*
r815lin128ib	mpi64			openmp32			gaussian8	

Les queues batch et leurs environnements

	mpi12 namd12	mpi8 namd8	openmp12	openmp8	gaussian12	gaussian8	gaussian4
c6100lin24sl			*		*		
dl165monobig				*		*	*
dl165mono							
x41zmono							
molcas1,2,3,4							
v40zmono							
v22zmono							
matlab							

Les queues batch mpi tests et leurs environnements

	test
R128_ib_test	*
c6100_ib_test	*
r410A_ib_test	*
r410B_ib_test	*
dl175_gb_test	*
dl165_ib_test	*
x41z_ib_test	*

Le système de batch

La commande qstat

La commande qstat: 3 exemples utiles

- `qstat -u "*" -s r`

- `qstat -g c`

- `qstat -u "*" -s p -r`

qstat -u "*" -s r

job-ID	prior	name	user	state	submit/start at	queue	slots	ja-task-ID
15200	0.31000	EINT_MP2_0	pfleura2	r	08/25/2010 12:31:28	c6100lin24sl@c6100lin38.ens-ly	12	
21201	0.30303	batchrun_m	anaso	r	09/13/2010 12:33:06	matlab@r610lin3.ens-lyon.fr	1	
21534	0.12756	aoutzut06	delbiond	r	09/15/2010 17:10:15	r410lin24gbC@r410lin83.ens-lyo	8	
21535	0.12756	aoutzut07	delbiond	r	09/15/2010 17:10:45	r410lin24gbC@r410lin75.ens-lyo	8	
21536	0.12756	aoutzut08	delbiond	r	09/15/2010 17:11:15	r410lin24gbC@r410lin74.ens-lyo	8	
21558	0.12752	aout00zut	delbiond	r	09/16/2010 07:57:15	r410lin24gbC@r410lin80.ens-lyo	8	
21614	0.11375	H2O_SLAB_6	fcinquin	r	09/16/2010 10:11:45	x41zlin16ib@x41zlin2.ens-lyon.	16	
21688	0.11498	Pt96dynL5	fcinquin	r	09/16/2010 13:21:15	c6100lin24ib@c6100lin8.ens-lyo	48	
21689	0.11498	Pt98dynL5	fcinquin	r	09/16/2010 13:21:45	c6100lin24ib@c6100lin28.ens-ly	48	
21690	0.11498	Pt101dynL5	fcinquin	r	09/16/2010 13:21:45	c6100lin24ib@c6100lin20.ens-ly	48	
21692	0.11498	Pt96dynL3	fcinquin	r	09/16/2010 16:47:16	dl175lin32@dl175lin3.ens-lyon.	48	
21693	0.11498	Pt98dynL3	fcinquin	r	09/16/2010 16:49:15	dl165lin24ib@dl165lin13.ens-ly	48	

qstat -g c

```
qstat -g c
CLUSTER QUEUE    CQLOAD  USED  RES  AVAIL  TOTAL  aoACDS  cdsuE
c6100lin24ib    0.17    156   0    276   432    0       0
c6100lin24sl    0.25     12   0     0    12     0       0
dl165adf        0.00     0    0    32    32     0       0
dl165lin24ib    0.01     0    0   160   160    0       0
dl165mono       0.00     0    0    48    56     0       8
dl175lin32gb    0.25     0    0     0   144    0      144
molcas1         0.09     1    0     3     4     0       0
molcas2         0.00     0    0     4     4     0       0
r410lin24gbC    0.29     48   0    48    96     0       0
r410lin24ibA    0.19    112   0   176   288    0       0
r410lin24ibB    0.22    128   0   160   288    0       0
r815lin         0.00     0    0    48    48     0       0
```


qstat -u "*" -s p -r

```
qstat -u "*" -s p -r
```

job-ID	prior	name	user	state	submit/start at	queue	slots	ja-task-ID
--------	-------	------	------	-------	-----------------	-------	-------	------------

22527	0.15756	TestVR3	htouil	qw	09/21/2010 16:19:46		240	
-------	---------	---------	--------	----	---------------------	--	-----	--

Full jobname: TestVR3

Requested PE: c6100_ib 240

Hard Resources:

Soft Resources:

Hard requested queues: c6100lin24ib

22614	0.10488	gv136s	xrozansk	qw	09/22/2010 09:14:47		4	
-------	---------	--------	----------	----	---------------------	--	---	--

Full jobname: gv136s

Requested PE: gaussian12 4

Hard Resources:

Soft Resources:

Hard requested queues: dl175lin32

22615	0.10297	armchair.d	fcinquin	qw	09/22/2010 09:15:59		16	
-------	---------	------------	----------	----	---------------------	--	----	--

Full jobname: armchair.dyn

Requested PE: vasp 16

Hard Resources:

Soft Resources:

Hard requested queues: r410lin24ibA

Le système de batch

Scripts de soumission d'un job
séquentiel

Les fichiers de soumission au système de batch (version séquentielle)

```
#!/bin/bash
```

```
#
```

```
QUEUE= "dl165mono,x41zmono,v40zmono,v22zmono"
```

```
#
```

```
# À changer
```

```
HOMEDIR="/home/..."
```

```
#
```

```
# man qsub pour obtenir une explication des options
```

```
#
```

```
/usr/local/sgc/bin/lx24-amd64/qsub -m b -m e -e ${HOMEDIR} -o ${HOMEDIR} -q $QUEUE ${HOMEDIR}/batchrun
```

Les fichiers soumis au système de batch (version séquentielle)

```
#!/bin/bash
#
# À modifier
WORKDIR="/home/..."
EXECDIR="/home/... "
#
# Important sinon l'exécutable monprogramme est lancé depuis ~
cd ${WORKDIR}
#
# À ne pas modifier
#$ -cwd
#$ -V
#
${EXECDIR}/monprogramme <Monfichierdedonnées > Monfichierderésultats
```

Le système de batch

Scripts de soumission d'un job
parallélisé via openmpi

Les fichiers de soumission au système de batch (version parallèle via openmpi)

```
#!/bin/bash
#
QUEUE="r410lin24ibA"
SOUSQUEUE="mpi8"
NSLOTS= "24 "
#
# À changer
HOMEDIR="/home/..."
#
# man qsub pour obtenir une explication des options
#
/usr/local/sgе/bin/lx24-amd64/qsub -m b -m e -e ${HOMEDIR} -o ${HOMEDIR}
-q $QUEUE -pe $SOUSQUEUE $NSLOTS ${HOMEDIR}/batchrun
```

Les fichiers soumis au système de batch (version parallèle via openmpi)

```
#!/bin/bash
#
# À modifier
WORKDIR="/home/... »
EXECDIR="/home/... »
#
# Donné par le système de batch($NSLOTS aussi)
HOSTFILE=$TMPDIR/machines
#
# Pour selectionner la commande mpirun associée au compilateur utilisé (à modifier)
PREFIX=/softs/openmpi-1.4.1-intel-11.1.069
MPIRUN=$PREFIX/bin/mpirun
#
# Important sinon la commande $MPIRUN est lancée depuis ~
cd ${WORKDIR}
#
```

Les fichiers soumis au système de batch (version parallèle via openmpi)

```
#  
#$ -cwd  
#$ -V  
#  
$MPIRUN -prefix $PREFIX -mca btl openib,tcp,self -mca btl_tcp_if_include eth1,eth0  
-np $NSLOTS ${EXECDIR}/Monprogrammeopenmpi
```


Le système de batch

Scripts de soumission d'un job parallélisé
via openmpi (cas particuliers)

Les fichiers de soumission au système de batch (version parallèle via openmpi) cas particuliers

```
#!/bin/bash
```

```
#
```

```
QUEUE= "x41zlin16ib,dl165lin24ib, r410lin24gbC,x41zlin32gb"
```

```
SOUSQUEUE="mpi8"
```

```
NSLOTS="8 "
```

```
#
```

Le système de batch

Scripts de soumission d'un job
via openmp ou équivalent

Les fichiers de soumission au système de batch de jobs gaussian

```
#!/bin/bash
```

```
#
```

```
QUEUE= "x41zlin32gb"
```

```
SOUSQUEUE= "gaussian8"
```

```
NSLOTS= "8"
```

```
#
```

```
ou
```

```
#!/bin/bash
```

```
#
```

```
QUEUE="x41zlin32gb"
```

```
SOUSQUEUE= "gaussian4"
```

```
NSLOTS="4"
```

```
#
```

Les fichiers de soumission au système de batch de jobs gaussian (suite)

```
#!/bin/bash  
#  
QUEUE= "dl175lin32gb"  
SOUSQUEUE= "gaussian12"  
NSLOTS= "12"  
#
```

Les fichiers soumis au système de batch de jobs gaussian

```
#!/bin/bash
#
# À modifier
WORKDIR="/home/..."
EXECDIR="/home/... "
#
# Important sinon l'exécutable monprogramme est lancé depuis ~
cd ${WORKDIR}
#
# À ne pas modifier
#$ -cwd
#$ -V
#
${EXECDIR}/monprogramme <Monfichierdedonnées > Monfichierderésultats
```

Les fichiers de soumission au système de batch de jobs openmp

```
#!/bin/bash
#
QUEUE= "x41zlin32gb, r410lin24gbC,x41zlin16ib,dl165lin24ib"
SOUSQUEUE= "openmp8"
NSLOTS="8 "
#
ou
#!/bin/bash
#
QUEUE= "dl175lin32gb"
SOUSQUEUE= "openmp12"
NSLOTS= "12"
#
```

Les fichiers soumis au système de batch de jobs openmp

```
#!/bin/bash
#
# À modifier
WORKDIR="/home/..."
EXECDIR="/home/... "
#
# Important sinon l'exécutable monprogramme est lancé depuis ~
cd ${WORKDIR}
#
# À ne pas modifier
#$ -cwd
#$ -V
#
${EXECDIR}/monprogramme <Monfichierdedonnées > Monfichierderésultats
```


Le système de batch

Scripts de soumission d'un job parallélisé
via openmpi (directives particulières)

Les fichiers de soumission au système de batch version parallèle via openmpi (directives particulières)

```
#!/bin/bash
#
QUEUE="r410lin24ibA"
SOUSQUEUE="mpi8"
NSLOTS= "128 "
#
# À changer
HOMEDIR="/home/..."
#
# man qsub pour obtenir une explication des options
#
/usr/local/sge/bin/lx24-amd64/qsub -m b -m e -e ${HOMEDIR} -o ${HOMEDIR}
-q $QUEUE -pe $SOUSQUEUE $NSLOTS ${HOMEDIR}/batchrun
```

Les fichiers soumis au système de batch batch version parallèle via openmpi (directives particulières)

```
#!/bin/bash
#
# À modifier
WORKDIR="/home/..."
EXECDIR="/home/..."
#
# Donné par le système de batch($NSLOTS aussi)
HOSTFILE=$TMPDIR/machines
#
# Pour selectionner la commande mpirun associée au compilateur utilisé (à modifier)
PREFIX=/softs/openmpi-1.4.1-intel-11.1.069
MPIRUN=$PREFIX/bin/mpirun
#
# Important sinon la commande $MPIRUN est lancée depuis ~
cd ${WORKDIR}
#
```

Les fichiers soumis au système de batch version parallèle via openmpi (directives particulières)

```
#  
#$ -cwd  
#$ -V  
#  
NPERCKET="4"  
#  
$MPIRUN -prefix $PREFIX -mca btl openib,tcp,self -mca btl_tcp_if_include eth1,eth0  
-npsocket ${NPERCKET} -bycore -bind-to-core -report-bindings  
-np $NSLOTS ${EXECDIR}/Monprogrammeopenmpi
```