

Εκτέλεση

Έχουμε 4 βασικές ουρές:

- clones (1-26,29)
- xenons (1-8)
- termis (2-6)
- serial (nehalem, dunnington, rouxlas)

και 1 γενική:

- parallel (clones, xenons, termis)

1. για εκτέλεση σε κάποιο clone:

```
qsub -q clones script.sh
```

ή για interactive mode:

```
qsub -q clones -I
```

2. για εκτέλεση σε κάποιο serial με 12 cores και 10G μνήμη:

```
qsub -q serial -l nodes=1:ppn=12,vmem=10000m script.sh
```

3. για εκτέλεση σε 4 xenons με 4 cores και 2G μνήμη ανά node:

```
qsub -q xenons -l nodes=4:ppn=4,vmem=1900m script.sh
```

4. για εκτέλεση σε 1 node (parallel), με 8 cores, 10G μνήμη και local storage:

```
qsub -q parallel -l nodes=1:ppn=8:local,vmem=10000m script.sh
```

5. αν είστε συναισθηματικοί με το clone που πάντα είχατε (έστω clone4):

```
qsub -q clones -l nodes=clone4:ppn=8,vmem=1800m -I
```

ή με τα clones που πάντα είχατε (έστω clone4-clone8):

```
qsub -q clones -l nodes=clone4:ppn=8+clone5:ppn=8+clone6:ppn=8+clone7:
```

Παράδειγμα 1: Interactive, 4 nodes, 8 cores each, unlimited memory usage (CAREFUL)

```
ananos@scirouter:~$ qsub -q clones -l nodes=4:ppn=8 -I
qsub: waiting for job 5435.localhost to start
qsub: job 5435.localhost ready

ananos@clone1:~$ cat /var/spool/pbs/aux/5435.localhost |uniq
clone1
clone2
clone3
clone4
ananos@clone1:~$ ulimit -v
unlimited
ananos@clone1:~$
```

Παράδειγμα 2: Interactive, 4 nodes, 8 cores each, limited memory usage

```
ananos@scirouter:~$ qsub -q clones -l nodes=4:ppn=8,vmem=1500m -I
qsub: waiting for job 5436.localhost to start
qsub: job 5436.localhost ready

ananos@clone1:~$ cat /var/spool/pbs/aux/5436.localhost |uniq
clone1
clone2
```

```
clone3
clone4
ananos@clone1:~$ ulimit -v
1536000
ananos@clone1:~$
```

Παράδειγμα 3: Interactive 3 nodes, 16 cores each, serial, limited memory usage

```
ananos@scirouter:~$ qsub -q serial -l nodes=3:ppn=16,vmem=10000m -I
qsub: waiting for job 5437.localhost to start
qstqsub: job 5437.localhost ready

ananos@nehalem:~$ cat /var/spool/pbs/aux/5437.localhost |uniq
nehalem
dunnington
rouxlas
ananos@nehalem:~$ ulimit -v
10240000
ananos@nehalem:~$
```

Παράδειγμα 4: NON-Interactive 8 nodes, 4 cores each, parallel, limited memory usage per core

Το αρχείο που θα κάνουμε qsub:

```
ananos@scirouter:~$ cat advection.pbs_small
#!/bin/bash -l

#PBS -l nodes=8:ppn=4,pvmem=400m
#PBS -N advective_small

cd advective3D
for itype in mx tcp
do
    for x in 32
    do

        MX_RCACHE=2 \
            /usr/local/open-mpi/bin/mpirun \
            -x MX_RCACHE=2 \
            -np $x --mca btl $itype,sm,self -bynode \
            adv3D.comm -T 100 -X 256 -Y 256 -Z 256 -k 2 -x 1 -y 1 -z $x

        MX_RCACHE=2 \
            /usr/local/open-mpi/bin/mpirun \
            -x MX_RCACHE=2 \
            -np $x --mca btl $itype,sm,self -bynode \
            adv3D.comm -T 100 -X 256 -Y 256 -Z 256 -k 2 -x 1 -y $x -z 1

        MX_RCACHE=2 \
            /usr/local/open-mpi/bin/mpirun \
            -x MX_RCACHE=2 \
            -np $x --mca btl $itype,sm,self -bynode \
            adv3D.comm -T 100 -X 256 -Y 256 -Z 256 -k 2 -x $x -y 1 -z 1

    done
done
```

η εντολή:

```
ananos@scirouter:~$ qsub -q clones advection.pbs_small
5438.localhost
ananos@scirouter:~$
```

παρακολούθηση της εκτέλεσης:

```
ananos@scirouter:~$ qstat
Job id              Name                User              Time Use S Queue
-----
5438.localhost      advective_small     annanos           0 R clones
ananos@scirouter:~$
```

και το output:

```

ananos@scirouter:~$ cat advective_small.o5438
1x1x32 1.171 0.906 0.427 0.479 0.266
1x32x1 0.533 0.349 0.032 0.317 0.184
32x1x1 0.500 0.319 0.041 0.278 0.182
1x1x32 2.370 2.104 0.328 1.776 0.266
1x32x1 2.051 1.840 0.021 1.818 0.211
32x1x1 2.070 1.877 0.026 1.850 0.193

```

Αντίστοιχα για 2 ουρές:

```

ananos@scirouter:~$ qsub -q clones advection.pbs_small
5441.localhost
ananos@scirouter:~$ qsub -q xenons advection.pbs_small
5442.localhost
ananos@scirouter:~$ qstat
Job id                Name                User                Time Use S Queue
-----
5441.localhost        advective_small      ananos              0 R clones
5442.localhost        advective_small      ananos              0 R xenons

```

και το output:

```

ananos@scirouter:~$ cat advective_small.o5441
1x1x32 1.131 0.887 0.469 0.418 0.244
1x32x1 0.801 0.478 0.031 0.448 0.323
32x1x1 0.507 0.330 0.044 0.286 0.177
1x1x32 2.181 1.909 0.323 1.585 0.272
1x32x1 2.065 1.835 0.020 1.815 0.230
32x1x1 2.055 1.864 0.026 1.838 0.191
ananos@scirouter:~$ cat advective_small.o5442
1x1x32 3.470 2.702 1.016 1.686 0.768
1x32x1 2.170 1.335 0.087 1.248 0.835
32x1x1 2.064 1.504 0.089 1.415 0.560
1x1x32 4.181 3.404 0.757 2.646 0.778
1x32x1 3.365 2.447 0.076 2.371 0.918
32x1x1 3.253 2.603 0.105 2.497 0.650

```

Tips

- 1. Αν θέλουμε να τρέχουμε interactive jobs, για να μπορούμε να τα αφήσουμε να τρέχουν και να τα ανακτήσουμε αργότερα τρέχουμε

```
screen
```

στον scirouter και μετά qsub κλπ. για να βγούμε από το screen, αρκεί το ctrl-A-d (detach) και για να ξαναμπούμε κάποια στιγμή το

```
screen -rx
```

- 2. Προσοχή στη χρήση της μνήμης: Η εντολή

```
pbsnodelist
```

μας δίνει την κατάσταση των μηχανημάτων. Το 5ο πεδίο είναι η διαθέσιμη μνήμη του μηχανήματος σε MB. Φροντίζουμε να ορίζουμε πάντα το vmem=XXXXm για να είμαστε σίγουροι ότι θα τρέξουμε σε μηχανήματα που σγκώνει την εφαρμογή μας.

- 3. Προς το παρόν ακολουθούμε την πεπατημένη χρησιμοποιώντας ολόκληρα μηχανήματα για όσο διαρκεί η δουλειά που θέλουμε να τρέξουμε.
- 4. Οι κόμβοι, σύμφωνα με το pbsnodelist έχουν συγκεκριμένα χαρακτηριστικά (πχ. mx10g -- myrinet 10gbps, mx2g -- myrinet 2+2gbps, local -- local storage, clover -- intel core αρχιτεκτονική code name: clovertown κλπ.) επομένως αν θέλουμε να τρέξουμε σε myri10g με local storage γράφουμε:

```
qsub -q clones -l nodes=4:ppn8:mx10g:local -I
```