Queues d’ordonnancement fonctionnement actuel & évolutions possibles

slurm

Rué François – SED Bordeaux - PlaFRIM

Queues actuelles

Différentes ressources

- Miriel
- Sirocco0[1-5]
- sirocco06
- Mistral
- Power8
- Knl
- Souris
- Mirage

Différentes queues

Queues actuelles

Différentes ressources

- Miriel
- Sirocco0[1-5]
- sirocco06
- Mistral
- Power8
- Knl
- Souris
- Mirage

Différentes queues

- ibm_pw8
- defq
- longq
- court
- multipart
- special
- souris
- court_souris
- long_souris
- court_mirage
- long_mirage
- special_mirage
- court_mistral
- long_mistral
- court_sirocco
- long_sirocco
- intel_knl

Queues actuelles

Différentes ressources

- Miriel
- Sirocco0[1-5]
- sirocco06
- Mistral
- Power8
- Knl
- Souris
- Mirage

Différentes queues

- ibm_pw8
- defq
- longq
- court
- multipart
- special
- souris
- court_souris
- long_souris
- court_mirage
- long_mirage
- special_mirage
- court_mistral
- long_mistral
- court_sirocco
- long_sirocco
- intel_knl

Queues actuelles

Différentes ressources

- Miriel
- Sirocco0[1-5]
- sirocco06
- Mistral
- Power8
- Knl
- Souris
- Mirage

Différentes queues

- ibm_pw8
- defq
- longq
- court
- multipart
- special
- souris
- court_souris
- long_souris
- court_mirage
- long_mirage
- special_mirage
- court_mistral
- long_mistral
- court_sirocco
- long_sirocco
- intel_knl

Queues actuelles

Différentes ressources

- Miriel
- Sirocco0[1-5]
- sirocco06
- Mistral
- Power8
- Knl
- Souris
- Mirage

Différentes queues

- ibm_pw8
- defq
- longq
- court
- multipart
- special
- souris
- court_souris
- long_souris
- court_mirage
- long_mirage
- special_mirage
- court_mistral
- long_mistral
- court_sirocco
- long_sirocco
- intel_knl

Queues actuelles

Différentes ressources

- Miriel
- Sirocco0[1-5]
- sirocco06
- Mistral
- Power8
- Knl
- Souris
- Mirage

pas sur le même switch …

Différentes queues

- ibm_pw8
- defq
- longq
- court
- multipart
- special
- souris
- court_souris
- long_souris
- court_mirage
- long_mirage
- special_mirage
- court_mistral
- long_mistral
- court_sirocco
- long_sirocco
- intel_knl

Queues actuelles

Différentes ressources

- Miriel
- Sirocco0[1-5]
- sirocco06
- Mistral
- Power8
- Knl
- Souris
- Mirage

Différentes queues

- ibm_pw8
- defq
- longq
- court
- multipart
- special
- souris
- court_souris
- long_souris
- court_mirage
- long_mirage
- special_mirage
- court_mistral
- long_mistral
- court_sirocco
- long_sirocco
- intel_knl

Queues actuelles

Différentes ressources

- Miriel
- Sirocco0[1-5]
- sirocco06
- Mistral
- Power8
- Knl
- Souris
- Mirage

Différentes queues

- ibm_pw8
- defq
- longq
- court
- multipart
- special
- souris
- court_souris
- long_souris
- court_mirage
- long_mirage
- special_mirage
- court_mistral
- long_mistral
- court_sirocco
- long_sirocco
- intel_knl

Limitées à quelques équipes ...

Queues actuelles - qos

qos : qualité de service pour chaque travail soumis sur le cluster à travers slurm

```
[rue@devel13 ~]$ sacctmgr show qos format=Name%20,GrpCPUS,GrpJobs,GrpNodes,GrpSubmit,MaxNodes,MaxWall,MaxJob,MaxSubmit
                Name  GrpCPUs GrpJobs GrpNodes GrpSubmit MaxNodes     MaxWall MaxJobs MaxSubmit
-------------------- -------- ------- -------- --------- -------- ----------- ------- ---------
              normal
               multi                3                4                           1         2
               longq     1080      16       45        64       16  3-00:00:00       2        10
               court     1080               45        64             04:00:00       2        10
             special                40       77        50                           10        20
            verylong      192               8
             sirocco
             mistral
          qoscompute                                   1
      qoscrt_mistral                20                20                           2        10
      qoscrt_sirocco                 5                10                           2         4
       qoslg_mistral                10                10                           2         5
       qoslg_sirocco                 2                 4                           1         2
       qoscrt_mirage                            8       32                          16        16
        qoslg_mirage                            8       32                          16        16
           qos_souris
```

Queues actuelles - qos

qos : qualité de service pour chaque travail soumis sur le cluster à travers slurm

```
[rue@devel13 ~]$ sacctmgr show qos format=Name%20,GrpCPUS,GrpJobs,GrpNodes,GrpSubmit,MaxNodes,MaxWall,MaxJob,MaxSubmit
                Name  GrpCPUs GrpJobs GrpNodes GrpSubmit MaxNodes     MaxWall MaxJobs MaxSubmit
-------------------- -------- ------- -------- --------- -------- ----------- ------- ---------
              normal
```

```
# Partitions
PartitionName=ibm_pw8 Default=NO MinNodes=1 MaxNodes=1 DefaultTime=00:10:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=YES:4 GraceTime=0 Pre
emptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP
PartitionName=defq Default=YES MinNodes=1 MaxNodes=4 DefaultTime=01:00:00 MaxTime=02:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=YES:4 GraceTime=0 Preem
ptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=miriel[001-088]
PartitionName=sirocco Default=NO MinNodes=1 MaxNodes=3 DefaultTime=03:00:00 MaxTime=10:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 Preem
ptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=DRAIN
PartitionName=mistral Default=NO MinNodes=1 MaxNodes=10 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 Pree
mptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=DRAIN
PartitionName=longq Default=NO MinNodes=1 MaxNodes=16 DefaultTime=01:00:00 MaxTime=3-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 Preem
ptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=longq LLN=NO State=UP Nodes=miriel[001-088]
PartitionName=court Default=NO MinNodes=1 MaxNodes=42 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 Preempt
Mode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=court LLN=NO State=UP Nodes=miriel[001-088]
PartitionName=multiPart Default=NO MinNodes=1 MaxNodes=40 DefaultTime=00:10:00 MaxTime=01:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 Pre
emptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=multi LLN=NO State=UP Nodes=miriel[001-088],mistral[11-18],sirocco[01-05]
PartitionName=special Default=NO MinNodes=1 MaxNodes=77 DefaultTime=00:10:00 MaxTime=00:30:00 AllowGroups=ALL Priority=100 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 Pree
mptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=special LLN=NO State=UP Nodes=miriel[001-088]
PartitionName=souris Default=NO MinNodes=1 MaxNodes=1 DefaultTime=00:10:00 MaxTime=7-00:00:00 AllowGroups=lfant,geostat,sed-bdx,cellule-imb Priority=1000 DisableRootJobs=NO RootOnly=NO Hidde
n=NO Shared=YES:4 GraceTime=0 PreemptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=souris
PartitionName=computeq Default=NO MinNodes=1 MaxNodes=16 DefaultTime=01:00:00 MaxTime=7-00:00:00 AllowGroups=ALL Priority=100 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0
PreemptMode=OFF ReqResv=YES AllowAccounts=ALL AllowQos=qoscompute LLN=NO State=UP Nodes=mistral[11-18]
PartitionName=mirabelle Default=NO MinNodes=1 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 PreemptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQo
s=ALL LLN=NO State=DRAIN
PartitionName=mirage Default=NO MinNodes=1 MaxNodes=3 DefaultTime=03:00:00 MaxTime=10:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 Preemp
tMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=DRAIN
PartitionName=court_mirage Default=NO MinNodes=1 MaxNodes=8 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 P
reemptMode=OFF ReqResv=NO DefMemPerCPU=3016 MaxMemPerNode=36195 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=mirage[01-09]
PartitionName=long_mirage Default=NO MinNodes=1 MaxNodes=8 DefaultTime=01:00:00 MaxTime=3-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0
PreemptMode=OFF ReqResv=NO DefMemPerCPU=3016 MaxMemPerNode=36195 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=mirage[01-09]
PartitionName=special_mirage Default=NO MinNodes=1 MaxNodes=8 DefaultTime=01:00:00 MaxTime=1-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime
=0 PreemptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP
PartitionName=long_souris Default=NO MinNodes=1 DefaultTime=01:00:00 MaxTime=3-00:00:00 AllowGroups=lfant,geostat,sic,sed-bdx,runtime,hiepacs,cellule-imb,iparla Priority=10 DisableRootJobs=N
O RootOnly=NO Hidden=NO Shared=NO GraceTime=0 PreemptMode=OFF ReqResv=NO MaxMemPerNode=2997008 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=souris
PartitionName=court_souris Default=NO MinNodes=1 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=YES:4 GraceTime=0 PreemptMo
de=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=souris
PartitionName=court_mistral Default=NO MinNodes=1 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 PreemptMode=OFF ReqResv=NO DefMe
mPerCPU=6441 MaxMemPerNode=128832 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=mistral[11-18]
PartitionName=court_sirocco Default=NO MinNodes=1 MaxNodes=5 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0
PreemptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=qoscrt_sirocco LLN=NO State=UP Nodes=sirocco[01-05]
PartitionName=long_sirocco Default=NO MinNodes=1 MaxNodes=2 DefaultTime=01:00:00 MaxTime=3-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0
 PreemptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=qoslg_sirocco LLN=NO State=UP Nodes=sirocco[01,02]
PartitionName=long_mistral Default=NO MinNodes=1 MaxNodes=16 DefaultTime=01:00:00 MaxTime=7-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=
0 PreemptMode=OFF ReqResv=NO DefMemPerCPU=6441 MaxMemPerNode=128832 AllowAccounts=ALL AllowQos=qoslg_mistral LLN=NO State=UP Nodes=mistral[11-18]
PartitionName=qholiday Default=NO MinNodes=1 MaxNodes=36 DefaultTime=00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 PreemptMode=OFF ReqR
esv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP
PartitionName=testpreempt Default=NO MinNodes=1 MaxNodes=80 DefaultTime=01:00:00 MaxTime=08:00:00 AllowGroups=ALL Priority=1 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=17
PreemptMode=REQUEUE ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=mirage[01-09],miriel[001-088],mistral[11-18],sirocco[01-05],souris
PartitionName=intel_knl Default=NO MinNodes=1 MaxNodes=1 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=YES:4 GraceTime=0 P
reemptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP
```

Queues actuelles - qos

qos : qualité de service pour chaque travail soumis sur le cluster à travers slurm

```
[rue@devel13 ~]$ sacctmgr show qos format=Name%20,GrpCPUS,GrpJobs,GrpNodes,GrpSubmit,MaxNodes,MaxWall,MaxJob,MaxSubmit
                Name  GrpCPUs GrpJobs GrpNodes GrpSubmit MaxNodes     MaxWall MaxJobs MaxSubmit
-------------------- -------- ------- -------- --------- -------- ----------- ------- ---------
              normal
```

```
# Partitions
PartitionName=ibm_pw8 Default=NO MinNodes=1 MaxNodes=1 DefaultTime=00:10:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=YES:4 GraceTime=0 Pre
emptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP
PartitionName=defq Default=YES MinNodes=1 MaxNodes=4 DefaultTime=01:00:00 MaxTime=02:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=YES:4 GraceTime=0 Preem
ptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=miriel[001-088]
PartitionName=sirocco Default=NO MinNodes=1 MaxNodes=3 DefaultTime=03:00:00 MaxTime=10:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 Preem
ptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=DRAIN
PartitionName=mistral Default=NO MinNodes=1 MaxNodes=10 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 Pree
mptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=DRAIN
PartitionName=longq Default=NO MinNodes=1 MaxNodes=16 DefaultTime=01:00:00 MaxTime=3-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 Preem
ptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=longq LLN=NO State=UP Nodes=miriel[001-088]
PartitionName=court Default=NO MinNodes=1 MaxNodes=42 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 Preempt
Mode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=court LLN=NO State=UP Nodes=miriel[001-088]
PartitionName=multiPart Default=NO MinNodes=1 MaxNodes=40 DefaultTime=00:30:00 MaxTime=01:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 Pre
emptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=multi LLN=NO State=UP Nodes=miriel[001-088],mistral[11-18],sirocco[01-05]
PartitionName=special Default=NO MinNodes=1 MaxNodes=77 DefaultTime=00:10:00 MaxTime=00:30:00 AllowGroups=ALL Priority=100 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 Pree
mptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=special LLN=NO State=UP Nodes=miriel[001-088]
PartitionName=souris Default=NO MinNodes=1 MaxNodes=1 DefaultTime=00:10:00 MaxTime=7-00:00:00 AllowGroups=lfant,geostat,sed-bdx,cellule-imb Priority=1000 DisableRootJobs=NO RootOnly=NO Hidde
n=NO Shared=YES:4 GraceTime=0 PreemptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=souris
PartitionName=computeq Default=NO MinNodes=1 MaxNodes=16 DefaultTime=01:00:00 MaxTime=7-00:00:00 AllowGroups=ALL Priority=100 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0
PreemptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=qoscomputeq LLN=NO State=UP Nodes=mistral[11-18]
PartitionName=mirabelle Default=NO MinNodes=1 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 PreemptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQo
s=ALL LLN=NO State=DRAIN
PartitionName=mirage Default=NO MinNodes=1 MaxNodes=3 DefaultTime=03:00:00 MaxTime=10:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 Preemp
tMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=DRAIN
PartitionName=court_mirage Default=NO MinNodes=1 MaxNodes=8 DefaultTime=01:00:00 MaxTime=4:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 P
reemptMode=OFF ReqResv=NO DefMemPerCPU=3016 MaxMemPerNode=36195 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=mirage[01-09]
PartitionName=long_mirage Default=NO MinNodes=1 MaxNodes=8 DefaultTime=01:00:00 MaxTime=3-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0
PreemptMode=OFF ReqResv=NO DefMemPerCPU=3016 MaxMemPerNode=36195 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=mirage[01-09]
PartitionName=special_mirage Default=NO MinNodes=1 MaxNodes=8 DefaultTime=01:00:00 MaxTime=1-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime
=0 PreemptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP
PartitionName=long_souris Default=NO MinNodes=1 DefaultTime=01:00:00 MaxTime=3-00:00:00 AllowGroups=lfant,geostat,sic,sed-bdx,runtime,hiepacs,cellule-imb,iparla Priority=10 DisableRootJobs=N
O RootOnly=NO Hidden=NO Shared=NO GraceTime=0 PreemptMode=OFF ReqResv=NO MaxMemPerNode=2997008 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=souris
PartitionName=court_souris Default=NO MinNodes=1 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=YES:4 GraceTime=0 PreemptMo
de=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=souris
PartitionName=court_mistral Default=NO MinNodes=1 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0 PreemptMode=OFF ReqResv=NO DefMe
mPerCPU=6441 MaxMemPerNode=128832 AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=mistral[11-18]
PartitionName=court_sirocco Default=NO MinNodes=1 MaxNodes=5 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0
PreemptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=qoscrt_sirocco LLN=NO State=UP Nodes=sirocco[01-05]
PartitionName=long_sirocco Default=NO MinNodes=1 MaxNodes=2 DefaultTime=01:00:00 MaxTime=3-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=0
 PreemptMode=OFF ReqResv=NO DefMemPerCPU=5330 MaxMemPerNode=128400 AllowAccounts=ALL AllowQos=qoslg_sirocco LLN=NO State=UP Nodes=sirocco[01,02]
PartitionName=long_mistral Default=NO MinNodes=1 MaxNodes=16 DefaultTime=01:00:00 MaxTime=7-00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO GraceTime=
0 PreemptMode=OFF ReqResv=NO DefMemPerCPU=6441 MaxMemPerNode=128832 AllowAccounts=ALL AllowQos=qoslg_mistral LLN=NO State=UP Nodes=mistral[11-18]
PartitionName=qholiday Default=NO MinNodes=1 MaxNodes=36 DefaultTime=00:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=0 PreemptMode=OFF ReqR
esv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP
PartitionName=testpreempt Default=NO MinNodes=1 MaxNodes=80 DefaultTime=01:00:00 MaxTime=08:00:00 AllowGroups=ALL Priority=1 DisableRootJobs=NO RootOnly=NO Hidden=YES Shared=NO GraceTime=17
PreemptMode=REQUEUE ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP Nodes=mirage[01-09],miriel[001-088],mistral[11-18],sirocco[01-05],souris
PartitionName=intel_knl Default=NO MinNodes=1 MaxNodes=1 DefaultTime=01:00:00 MaxTime=04:00:00 AllowGroups=ALL Priority=10 DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=YES:4 GraceTime=0 P
reemptMode=OFF ReqResv=NO AllowAccounts=ALL AllowQos=ALL LLN=NO State=UP
```

Temps moyen d'attente : 5757.16 secondes soit environ 9 minutes

Temps moyen de calcul : 2076.91 secondes soit environ 34 minutes

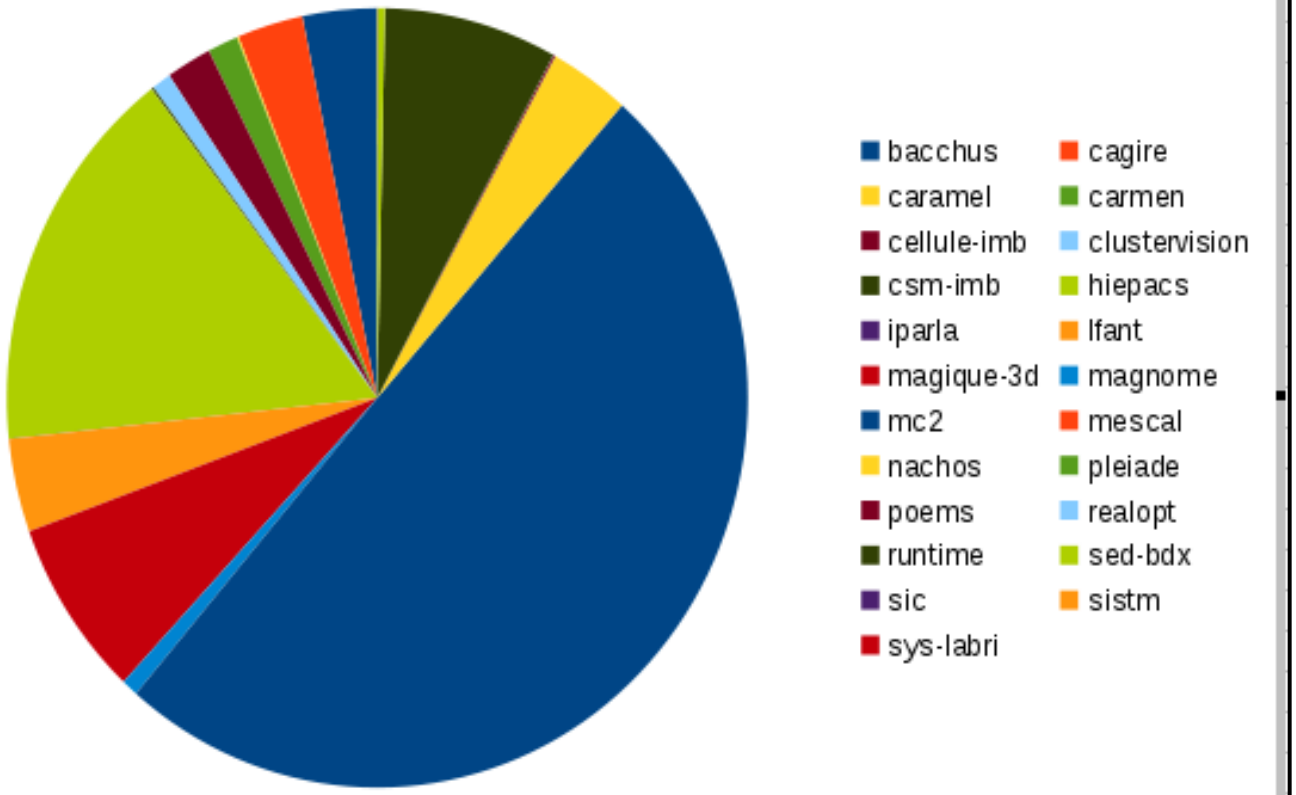
Queues actuelles - qos

```
[root@bright ~]# sreport user top Start="2016-01-01" End="2017-01-23"
--------------------------------------------------------------------------------
Top 10 Users 2016-01-01T00:00:00 - 2017-01-22T23:59:59 (33523200 secs)
Time reported in CPU Minutes
--------------------------------------------------------------------------------
  Cluster     Login     Proper Name         Account       Used     Energy
--------- --------- --------------- --------------- ---------- ----------
slurm_cl+   bruneau Charles-Henri +             mc2   92276099          0
slurm_cl+  jguillot  Jeremy Guillot         realopt   32095204          0
slurm_cl+  jguillot  Jeremy Guillot         realopt   30081335          0
slurm_cl+  fbernard Florian Bernard             mc2   27034303          0
slurm_cl+    araeli     Alice Raeli             mc2   24758892          0
slurm_cl+  bmatschk Benjamin Matsc+         a2x-imb   23996769          0
slurm_cl+  allomber  Bill Allombert     cellule-imb   19372307          0
slurm_cl+  adebraue Alexia deBrauer             mc2   18710167          0
slurm_cl+    qviaud   Quentin Viaud         realopt   18574108          0
slurm_cl+     milio      Enea Milio           lfant   18464130          0
[root@bright ~]#
```

Queues actuelles - qos

```
[root@br
--------
Top 10 U
Time rep
--------
  Cluste
--------
slurm_cl
slurm_cl
slurm_cl
slurm_cl
slurm_cl
slurm_cl
slurm_cl
slurm_cl
slurm_cl
[root@br
```

	A	B
1	bacchus	3,25
2	cagire	2,87
3	caramel	0,08
4	carmen	1,32
5	cellule-imb	1,98
6	clustervision	0,84
7	csm-imb	0,1
8	hiepacs	16,26
9	iparla	0,01
10	lfant	3,88
11	magique-3d	7,43
12	magnome	0,72
13	mc2	49,84
14	mescal	0,01
15	nachos	3,49
16	pleiade	0,01
17	poems	0,09
18	realopt	0,01
19	runtime	7,55
20	sed-bdx	0,31
21	sic	0,01
22	sistm	0,01
23	sys-labri	0



Priority

Job Priority Factors In General

The job's priority at any given time will be a weighted sum of all the factors that have been enabled in the slurm.conf file. Job priority can be expressed as:

```
Job_priority =
        (PriorityWeightAge) * (age_factor) +
        (PriorityWeightFairshare) * (fair-share_factor) +
        (PriorityWeightJobSize) * (job_size_factor) +
        (PriorityWeightPartition) * (partition_factor) +
        (PriorityWeightQOS) * (QOS_factor) +
        SUM(TRES_weight_cpu * TRES_factor_cpu,
            TRES_weight_<type> * TRES_factor_<type>,
            ...)
```

Priority

Job Priority Factors In General

The job's priority at any given time will be a weighted sum of all the factors that have been enabled in the slurm.conf file. Job priority can be expressed as:

```
Job_priority =
        (PriorityWeightAge) * (age_factor) +
        (PriorityWeightFairshare) * (fair-share_factor) +
        (PriorityWeightJobSize) * (job_size_factor) +
        (PriorityWeightPartition) * (partition_factor) +
        (PriorityWeightQOS) * (QOS_factor) +
        SUM(TRES_weight_cpu * TRES_factor_cpu,
            TRES_weight_<type> * TRES_factor_<type>,
            ...)
```

Age factor :

The age factor represents the length of time a job has been sitting in the queue and elligible to run

Priority

Job Priority Factors In General

The job's priority at any given time will be a weighted sum of all the factors that have been enabled in the slurm.conf file. Job priority can be expressed as:

```
Job_priority =
        (PriorityWeightAge) * (age_factor) +
        (PriorityWeightFairshare) * (fair-share_factor) +
        (PriorityWeightJobSize) * (job_size_factor) +
        (PriorityWeightPartition) * (partition_factor) +
        (PriorityWeightQOS) * (QOS_factor) +
        SUM(TRES_weight_cpu * TRES_factor_cpu,
            TRES_weight_<type> * TRES_factor_<type>,
            ...)
```

Job size factor :

The job size factor correlates to the number of nodes or CPUs the job has requested.

Priority

Job Priority Factors In General

The job's priority at any given time will be a weighted sum of all the factors that have been enabled in the slurm.conf file. Job priority can be expressed as:

```
Job_priority =
        (PriorityWeightAge) * (age_factor) +
        (PriorityWeightFairshare) * (fair-share_factor) +
        (PriorityWeightJobSize) * (job_size_factor) +
        (PriorityWeightPartition) * (partition_factor) +
        (PriorityWeightQOS) * (QOS_factor) +
        SUM(TRES_weight_cpu * TRES_factor_cpu,
            TRES_weight_<type> * TRES_factor_<type>,
            ...)
```

Partition factor :

Each node partition can be assigned an integer priority. The larger the number, the greater the job priority will be for jobs that request to run in this partition.

Priority

Job Priority Factors In General

The job's priority at any given time will be a weighted sum of all the factors that have been enabled in the slurm.conf file. Job priority can be expressed as:

```
Job_priority =
        (PriorityWeightAge) * (age_factor) +
        (PriorityWeightFairshare) * (fair-share_factor) +
        (PriorityWeightJobSize) * (job_size_factor) +
        (PriorityWeightPartition) * (partition_factor) +
        (PriorityWeightQOS) * (QOS_factor) +
        SUM(TRES_weight_cpu * TRES_factor_cpu,
            TRES_weight_<type> * TRES_factor_<type>,
            ...)
```

QOS factor :

Each qos can be assigned an integer priority. The larger the number, the greater the job priority will be for jobs that request to run in this qos.

Priority

Job Priority Factors In General

The job's priority at any given time will be a weighted sum of all the factors that have been enabled in the slurm.conf file. Job priority can be expressed as:

```
Job_priority =
        (PriorityWeightAge) * (age_factor) +
        (PriorityWeightFairshare) * (fair-share_factor) +
        (PriorityWeightJobSize) * (job_size_factor) +
        (PriorityWeightPartition) * (partition_factor) +
        (PriorityWeightQOS) * (QOS_factor) +
        SUM(TRES_weight_cpu * TRES_factor_cpu,
            TRES_weight_<type> * TRES_factor_<type>,
            ...)
```

Fair share factor :

The fair-share component to a job's priority influences the order in which a user's queued jobs are scheduled to run based on the portion of the computing resources they have been allocated and the resources their jobs have already consumed. The fair-share factor does not involve a fixed allotment, whereby a user's access to a machine is cut off once that allotment is reached.

Priority

Job Priority Factors In General

The job's priority at any given time will be a weighted sum of all the factors that have been enabled in the slurm.conf file. Job priority can be expressed as:

```
Job_priority =
        (PriorityWeightAge) * (age_factor) +
        (PriorityWeightFairshare) * (fair-share_factor) +
        (PriorityWeightJobSize) * (job_size_factor) +
        (PriorityWeightPartition) * (partition_factor) +
        (PriorityWeightQOS) * (QOS_factor) +
        SUM(TRES_weight_cpu * TRES_factor_cpu,
            TRES_weight_<type> * TRES_factor_<type>,
            ...)
```

Fair share factor :

The fair-share component to a job's priority influences the order in which a user's queued jobs are scheduled to run based on the portion of the computing resources they have been allocated and the resources their jobs have already consumed. The fair-share factor does not involve a fixed allotment, whereby a user's access to a machine is cut off once that allotment is reached.

Queues actuelles

- avantages
- inconvénients

Queues actuelles

- avantages
  - Partage de la machine en fonction des usages
  - Définition fine des partitions
  - Spécialisation de certaines partitions
- inconvénients

Queues actuelles

- avantages
  - Partage de la machine en fonction des usages
  - Définition fine des partitions
  - Spécialisation de certaines partitions
- Inconvénients
  - Gestion par partition
  - Démultiplication des partitions
  - Pas de gestion globale
  - Pas de gestion des multiples interfaces

Partitions – propositions

- 2 partitions : court et long
- Des ressources spécifiques

Interfaces réseau :

Opa
Ib
...

Accélérateurs :

GPU
MIC

Machines spécialisées :

Power 8
KNL
Souris ?

Partitions - proposition

- Partition

  court/long --gres = none
  - gpu
  - mic
  - opa
  - …

Partitions - proposition

- Partition

  court/long --gres =
  - none
  - gpu
  - mic
  - opa
  - …

  - Simplifier l'usage des différentes partitions et des ressources matérielles associées
    - Par ex : au lieu de -p court_sirocco –gres=GPU on écrirait juste –gres=GPU, de même pour IB et OPA sur miriel

Partitions - proposition

- Partition

  court/long --gres = none
  - gpu
  - mic
  - opa
  - …

  - Simplifier l'usage des différentes partitions et des ressources matérielles associées
    - Par ex : au lieu de -p court_sirocco –gres=GPU on écrirait juste –gres=GPU, de même pour IB et OPA sur miriel
  - Court tous les nœuds avec une limite de 4h, long la moitié des nœuds avec 3 jours
  - Gerer les valeurs des parametres dans job priority

Partitions - proposition

- A côté de tout cela ?
  - **Les modules**
  - Les modules sont pensés et configurés en fonction du matériel
  - Les modules seront montés en fonction des ressources auquelles l'utilisateur souhaite accéder
  - **Les groupes**
  - Les utilisateurs ne seront plus attachés à un groupe
  - L'usage que les utilisateurs feront de l'ordonnanceur ne sera pas impacté par les autres membres du groupe