

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

MAUI or: How I stop Worrying and Love the Cluster

Andrea Cucca

February 14, 2009

Table of contents

- 1 What MAUI is
- 2 MAUI vs PBS
 - Basic MAUI Concepts
- 3 QOS
- 4 The new system
- 5 Everyday use
 - showbf
 - checkjob
 - showq
 - checknode
 - showres
 - Other Commands
 - Standard procedures
- 6 Improvements

A satellite view

The island of Maui is the second-largest of the Hawaiian Islands at 727.2 square miles (1883.5 km^2)

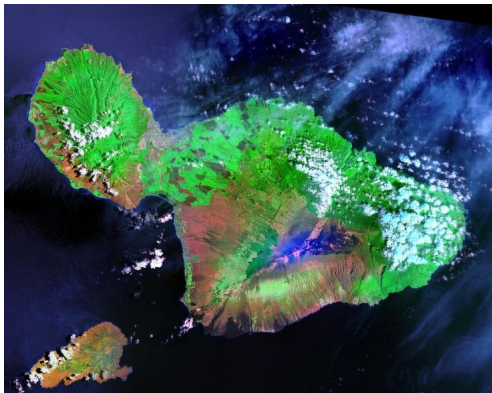


Figure: MAUI from satellite

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

A short description

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- Maui is a job scheduler that provides a mechanism for submitting, launching, and tracking jobs on a shared resource.
- MAUI determines, when, where, and how jobs are run; it maximizes the cluster performances, and takes care of traffic control, preventing jobs from interfering with each other (for example using the same resources).
- MAUI IS NOT a resource manager (like torque, PBS, Loadlever are) but needs a resource manager to work.

Differences between MAUI and PBS

- A resource manager (for example PBS) manages the job queue and manages the compute nodes.
- A scheduler tells the resource manager what to do, when to run jobs, and where.
- Users can submit jobs and query the state of the machine and jobs through the resource manager or, when Maui is running, they have other commands which provide additional information and capabilities.
- Example: PBS commands like `qstat` or `pbsnodes` display the actual state of jobs in a queue but don't give any further information about when or in which nodes the job will run.
On the other side, MAUI commands like `checkjob` or simply `showq` show you this information (and much more).

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

`showbf`

`checkjob`

`showq`

`checknode`

`showres`

Other Commands

Standard procedures

Improvements

What MAUI adds to PBS?

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- The possibility to create groups of resources dedicated to a selected type of jobs. (PBS do it too but... doesn't work!).
- The possibility to create resource reservations which guarantee resource availability at particular times
- The possibility to better manage cluster resources

Main concepts in MAUI

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

To better understand MAUI is worth to make clear some terminology

- 1 Resources and Nodes
- 2 Jobs and Tasks
- 3 Queues and Reservations
- 4 Backfill

Resources and Nodes

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- In the MAUI jargon a **resource** is every item that interacts with the scheduler (memory, CPU, disk space, software...) and a **node** is just a collection of resources.
- Information about nodes is provided to the scheduler chiefly by the resource manager.

Jobs and Tasks - Part I

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- Every **job** is defined by a certain number of “requirements” each requesting a number of resources, according to the information sent by the resource manager.
- In the simplest case a job is composed by only one requirement, for example 4 nodes of 4GB RAM.
- In a more complicated way of thinking a job may be composed by two requirements, one asking for 1 node of at least 512 MB of RAM and the second asking for 4 nodes of at least 1GB RAM.
- Each requirement consists of **tasks**. Tasks are the minimal independent unit of resources and are displayed by the various MAUI commands.

Jobs and Tasks - Part II

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- A key aspect of a task is that the resources associated with the task must be allocated as an atomic unit, without spanning node boundaries. A task requesting 1 proc and 1GB of memory cannot be satisfied by allocating 1 proc on one node and memory on another.
- This means that when you ask for 2 procs in 2 nodes you are actually asking to the scheduler to manage 4 tasks (`checkjob` reports this info).

Queues and Reservations

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- MAUI consider a queue as an abstract object to implement some policies to jobs and calls it a **class**
- Queues must be defined within the resource manager
- MAUI can be used to specify additional constraints to the queues defined by the resource manager (nodes subsets, priority by user or group etc.)
- Each time you submit a job to a queue you reserve a block of specific resources for a particular use, in a given time
- For example, next week, users Feffe and Julien may want to reserve 20 processors and 128 GB of memory from Monday 00:00 AM to Sunday 23:59 PM. If they correctly instruct MAUI, they can do it.

Backfill

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- Backfill is a scheduling algorithm that optimizes performances by distributing workload on resources.
- In the scheduler approach each job must provide an estimation of how long it will need to run (this is referred as wallclock limit so the walltime is simply the job duration).
- The more accurate the wallclock limit, the more 'holes' Maui can find to start your job early.
- Because of this backfill tends to favor smaller and shorter running jobs more than larger and longer running ones.
- The `showbf` command displays exactly what resources are available for immediate use.

Queues versus Quality of Service

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QoS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- Queues definition are independent from MAUI but rely only on the resource manager (PBS)
- When an idle job becomes eligible to run, it is assigned a priority. This priority is used to sort the jobs.
- Typically queues are used to divide and classify the workload. Each queue may have a fixed priority and each job can have a second priority parameter to sort it from the queue.
- Taking into account all parameters that set a batch job policy, you may end up with more queues than jobs.
- MAUI commands does not show you the queue in wich your job is running but they display the class and QoS of your job
- QoS (**Quality of Service**) is not a hierarchical scheme. It is merely a method of setting the parameters of a job when it enters the scheduler. All jobs eligible to run remain in one common idle-queue and their priorities are compared with all others

New implementation on nero

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QoS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- A new queue system based on amount of memory per CPU, so separating the nodes in 4 groups (mem_2; mem_4; mem_8; mem_16; mem_32) inside PBS
- A class system (with same names as queues) inside MAUI with restriction on node types, memory, CPU number, number jobs per user etc.
- A QoS system (with same names as queues) inside MAUI which a few parameters defining the priority
- There is no difference between queues at the moment concerning cputime or walltime

Restrictions on PBS queues

```
create queue mem.8
set queue mem.8 queue_type = Execution
set queue mem.8 max_queuable = 30
set queue mem.8 max_running = 20
set queue mem.8 acl_host_enable = False
set queue mem.8 acl_hosts = compute-0-9.local
set queue mem.8 acl_hosts += compute-0-8.local
set queue mem.8 acl_hosts += compute-0-7.local
set queue mem.8 acl_hosts += compute-0-6.local
set queue mem.8 acl_hosts += compute-0-5.local
set queue mem.8 acl_hosts += compute-0-4.local
set queue mem.8 acl_hosts += compute-0-3.local
set queue mem.8 acl_hosts += compute-0-10.local
set queue mem.8 resources_max.cput = 840:00:00
set queue mem.8 resources_max.mem = 8192mb
set queue mem.8 resources_max.ncpus = 8
set queue mem.8 resources_max.nodect = 4
set queue mem.8 resources_max.walltime = 999:00:00
set queue mem.8 resources_min.ncpus = 1
set queue mem.8 resources_min.nodect = 1
set queue mem.8 resources_default.cput = 504:00:00
set queue mem.8 resources_default.mem = 8000mb
set queue mem.8 resources_default.ncpus = 1
set queue mem.8 resources_default.nodect = 1
set queue mem.8 max_user_run = 6
set queue mem.8 enabled = True
set queue mem.8 started = True
```

Restrictions on classes

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- CLASSCFG[mem_2] QDEF=mem_2 MAXJOB=4 MAXPROCS=4 MAXJOBQUEUED=10 PRIORITY=1000 MAXMEM=2000 FRAME=5 FLAGS=PREEMPTEE
- CLASSCFG[mem_4] QDEF=mem_4 MAXJOB=16 MAXPROCS=16 MAXJOBQUEUED=30 PRIORITY=1000 MAXMEM=4000 FRAME=4 FLAGS=PREEMPTEE
- CLASSCFG[mem_8] QDEF=mem_8 MAXJOB=20 MAXPROCS=20 MAXJOBQUEUED=30 PRIORITY=1000 MAXMEM=8000 FRAME=3 FLAGS=PREEMPTEE
- CLASSCFG[mem_16] QDEF=mem_16 MAXJOB=6 MAXPROCS=6 MAXJOBQUEUED=10 PRIORITY=1000 MAXMEM=16000 FRAME=2 FLAGS=PREEMPTEE
- CLASSCFG[mem_32] QDEF=mem_32 MAXJOB=3 MAXPROCS=3 MAXNODE=1 MAXJOBQUEUED=6 PRIORITY=5000 MAXMEM=32000 FRAME=1 FLAGS=PREEMPTEE

Restrictions on QoS

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- QOSCFG[mem_2] MAXJOB=4 MAXPROCS=4 MAXJOBQUEUED=10 PRIORITY=1000
MAXMEM=2000 FRAME=5 FLAGS=PREEMPTTEE
- QOSCFG[mem_4] MAXJOB=16 MAXPROCS=16 MAXJOBQUEUED=30 PRIORITY=1000
MAXMEM=4000 FRAME=4 FLAGS=PREEMPTTEE
- QOSCFG[mem_8] MAXJOB=20 MAXPROCS=20 MAXJOBQUEUED=30 PRIORITY=1000
MAXMEM=8000 FRAME=3 FLAGS=PREEMPTTEE
- QOSCFG[mem_16] MAXJOB=6 MAXPROCS=6 MAXJOBQUEUED=10 PRIORITY=1000
MAXMEM=16000 FRAME=2 FLAGS=PREEMPTTEE
- QOSCFG[mem_32] MAXJOB=3 MAXPROCS=3 MAXNODE=1 MAXJOBQUEUED=6
PRIORITY=5000 MAXMEM=32000 FRAME=1 FLAGS=PREEMPTTEE

Same flags as above.

Lots of customizations possible: MAXWC, FSTARGET, QTWEIGHT, QTTARGET, XFWEIGHT, XFTARGET, PLIST, PDEF, QFLAGS,

Stop this bla bla and show us something useful!

Job submission is unchanged. You still use the same submission script and you submit the job with `qsub`
Then you have 2 options to manage your job:

- 1 Forget that you have MAUI installed and work as usual, with `qsub`, `qstat`, etc.
- 2 Use the MAUI-integrated commands.

showbf

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

`showbf`

`checkjob`

`showq`

`checknode`

`showres`

Other Commands

Standard procedures

Improvements

Use `showbf` when you want to know which resources are available for use.

Useful flags are:

- `showbf -S` (show the available resources per node)
- `showbf -v` (show the status of every node)

showbf

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

```
$ showbf -S
```

HostName	Procs	Memory	Disk	Swap	Time Available
compute-0-0.local	1	136	1	17437	41:03:56:47
compute-0-2.local	1	136	1	2054	41:03:55:28
compute-0-11.local	4	8040	1	8922	INFINITY
compute-0-12.local	8	32143	1	33051	INFINITY
compute-0-13.local	2	20143	1	11542	34:19:19:55

showbf

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showes

Other Commands

Standard procedures

Improvements

```
$ showbf -v
```

```
backfill window (user: 'cucca' group: 'users' partition: ALL) Tue Feb 3 11:37:06
```

```
16 procs available for 38:16:18:26
```

```
14 procs available with no timelimit
```

```
node compute-0-0.local is blocked by reservation 11346 in INFINITY
node compute-0-1.local is unavailable (state 'Down')
node compute-0-2.local is blocked by reservation 11345 in INFINITY
node compute-0-3.local is unavailable (state 'Busy')
node compute-0-4.local is unavailable (state 'Busy')
node compute-0-5.local is unavailable (state 'Busy')
node compute-0-6.local is unavailable (state 'Busy')
node compute-0-7.local is unavailable (state 'Busy')
node compute-0-8.local is unavailable (state 'Busy')
node compute-0-9.local is unavailable (state 'Busy')
node compute-0-10.local is unavailable (state 'Busy')
node compute-0-11.local is blocked by reservation NONE in INFINITY
node compute-0-12.local is blocked by reservation NONE in INFINITY
node compute-0-13.local is blocked by reservation 11325 in 38:16:18:26
```

checkjob

MAUI or: How I
stop Worrying and
Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

checkjob means more or less “Everything You Always Wanted to Know About Your Job* (*But Were Afraid to Ask) ”

Use **checkjob** when you want to know job state, resource requirements, environment, constraints, credentials, history, allocated resources, and resource utilization

checkjob

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

\$ checkjob

checking job 11326

State: Running

Creds: user:fiori group:users class:mem_8 qos:mem_8

WallTime: 2:12:25:12 of 41:15:00:00

SubmitTime: Sat Jan 31 12:58:06

(Time Queued Total: 10:39:17 Eligible: 10:38:49)

StartTime: Sat Jan 31 23:37:23

Total Tasks: 1

Req[0] TaskCount: 1 Partition: DEFAULT

Network: [NONE] Memory \geq 0 Disk \geq 0 Swap \geq 0

Opsys: [NONE] Arch: [NONE] Features: [NONE]

Dedicated Resources Per Task: PROCS: 1 MEM: 8000M

Allocated Nodes: [compute-0-6.local:1]

IWD: [NONE] Executable: [NONE]

Bypass: 0 StartCount: 1

PartitionMask: [ALL]

Reservation '11326' (-2:12:24:33 \rightarrow 39:02:35:27 Duration: 41:15:00:00)

PE: 1.12 StartPriority: 6388

showq

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

`showq` shows information about running, idle, and non-queued jobs.

Compared to `qstat` that shows only jobs owned by the user, this command displays all jobs in active, idle, and non-queued states for every user.

It doesn't show queue names because queue names appear only at the resource manager level, not at the scheduler level.

showq output

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

\$ showq

ACTIVE JOBS

JOBNAME	USERNAME	STATE	PROC	REMAINING	STARTTIME
11196	guzzo	Running	4	32:23:35:29	Sun Jan 25 22:35:43
11227	gatti	Running	4	34:16:16:48	Tue Jan 27 15:17:02
11291	vidal	Running	1	36:11:22:34	Thu Jan 29 10:22:48
11292	fiori	Running	1	36:15:29:25	Thu Jan 29 14:29:39
11307	guzzo	Running	6	38:13:51:38	Sat Jan 31 12:51:52
11322	fiori	Running	1	38:13:54:44	Sat Jan 31 12:54:58
11324	fiori	Running	1	38:13:56:05	Sat Jan 31 12:56:19
11325	lopez	Running	1	38:13:57:19	Sat Jan 31 12:57:33
11326	fiori	Running	1	39:00:37:09	Sat Jan 31 23:37:23
11327	fiori	Running	1	40:07:10:46	Mon Feb 2 06:11:00
11328	fiori	Running	1	40:07:10:46	Mon Feb 2 06:11:00
11336	guzzo	Running	4	40:08:58:10	Mon Feb 2 07:58:24
11345	lopez	Running	1	41:00:52:21	Mon Feb 2 23:52:35
11346	lopez	Running	1	41:00:53:40	Mon Feb 2 23:53:54
14 Active Jobs	28 of 44	Processors Active		(63.64%)	
	11 of 13	Nodes Active		(84.62%)	

IDLE JOBS

JOBNAME	USERNAME	STATE	PROC	WCLIMIT	QUEUETIME
11333	lopez	Idle	4	41:15:00:00	Sat Jan 31 15:20:21

1 Idle Job

BLOCKED JOBS

JOBNAME	USERNAME	STATE	PROC	WCLIMIT	QUEUETIME
---------	----------	-------	------	---------	-----------

checknode

```
$ checknode compute-0-10
```

```
checking node compute-0-10.local
```

```
State: Busy (in current state for 1:01:45:22)
```

```
Configured Resources: PROCS: 4 MEM: 31G SWAP: 12G DISK: 1M
```

```
Utilized Resources: PROCS: 4
```

```
Dedicated Resources: PROCS: 4 MEM: 6000M
```

```
Opsys: linux Arch: [NONE]
```

```
Speed: 1.00 Load: 3.530
```

```
Network: [DEFAULT]
```

```
Features: [NONE]
```

```
Attributes: [Batch]
```

```
Classes: [mem_8 2:4][route 4:4]
```

```
Total Time: INFINITY Up: INFINITY (92.63%) Active: INFINITY (84.96%)
```

```
Reservations:
```

```
Job '11196'(x2) -8:18:10:57 ⇒ 32:20:49:03 (41:15:00:00)
```

```
Job '11336'(x2) -1:08:48:16 ⇒ 40:06:11:44 (41:15:00:00)
```

```
JobList: 11196,11336
```

showres

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- `showres` displays all reservations currently in place within Maui. The default behavior is to display reservations on a reservation-by-reservation basis.
- `showres -n` display information regarding all nodes reserved

showres Output

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

\$ showres

Reservations

ResID	Type	S	Start	End	Duration	N/P	StartTime
11196	Job	R	-9:00:38:36	32:14:21:24	41:15:00:00	2/4	Sun Jan 25 22:35:56
11227	Job	R	-7:07:57:17	34:07:02:43	41:15:00:00	1/4	Tue Jan 27 15:17:15
11291	Job	R	-5:04:31:16	36:10:28:44	41:15:00:00	1/1	Thu Jan 29 18:43:16
11292	Job	R	-5:08:44:41	36:06:15:19	41:15:00:00	1/1	Thu Jan 29 14:29:51
11307	Job	R	-3:10:22:27	38:04:37:33	41:15:00:00	3/6	Sat Jan 31 12:52:05
11322	Job	R	-3:10:19:22	38:04:40:38	41:15:00:00	1/1	Sat Jan 31 12:55:10
11324	Job	R	-3:10:18:01	38:04:41:59	41:15:00:00	1/1	Sat Jan 31 12:56:31
11325	Job	R	-3:10:16:46	38:04:43:14	41:15:00:00	1/1	Sat Jan 31 12:57:46
11326	Job	R	-2:23:36:57	38:15:23:03	41:15:00:00	1/1	Sat Jan 31 23:37:35
11327	Job	R	-1:17:03:19	39:21:56:41	41:15:00:00	1/1	Mon Feb 2 06:11:13
11328	Job	R	-1:17:03:19	39:21:56:41	41:15:00:00	1/1	Mon Feb 2 06:11:13
11336	Job	R	-1:15:17:25	39:23:42:35	41:15:00:00	2/4	Mon Feb 2 07:57:07
11345	Job	R	-23:21:57	40:15:38:03	41:15:00:00	1/1	Mon Feb 2 23:52:35
11346	Job	R	-23:20:38	40:15:39:22	41:15:00:00	1/1	Mon Feb 2 23:53:54
11348	Job	R	-6:53:46	41:08:06:14	41:15:00:00	2/4	Tue Feb 3 16:20:46
11352	Job	R	-4:36:32	41:10:23:28	41:15:00:00	1/4	Tue Feb 3 18:38:00

16 reservations located

showres Output

```
$ showres -n
```

```
reservations on Tue Feb 3 22:51:26
```

Node	Type	ResID	JobState	Task	Start	Duration	StartTime
c0-0	Job	11346	Running	1	-22:57:32	41:15:00:00	Mon Feb 2 23:53:54
c0-1	Job	11291	Running	1	-5:04:08:10	41:15:00:00	Thu Jan 29 18:43:16
c0-2	Job	11345	Running	1	-22:58:51	41:15:00:00	Mon Feb 2 23:52:35
c0-3	Job	11292	Running	1	-5:08:21:35	41:15:00:00	Thu Jan 29 14:29:51
c0-3	Job	11324	Running	1	-3:09:54:55	41:15:00:00	Sat Jan 31 12:56:31
c0-4	Job	11307	Running	2	-3:09:59:21	41:15:00:00	Sat Jan 31 12:52:05
c0-5	Job	11307	Running	2	-3:09:59:21	41:15:00:00	Sat Jan 31 12:52:05
c0-6	Job	11322	Running	1	-3:09:56:16	41:15:00:00	Sat Jan 31 12:55:10
c0-6	Job	11326	Running	1	-2:23:13:51	41:15:00:00	Sat Jan 31 23:37:35
c0-7	Job	11196	Running	2	-9:00:15:30	41:15:00:00	Sun Jan 25 22:35:56
c0-8	Job	11336	Running	2	-1:14:54:19	41:15:00:00	Mon Feb 2 07:57:07
c0-9	Job	11307	Running	2	-3:09:59:21	41:15:00:00	Sat Jan 31 12:52:05
c0-9	Job	11327	Running	1	-1:16:40:13	41:15:00:00	Mon Feb 2 06:11:13
c0-9	Job	11328	Running	1	-1:16:40:13	41:15:00:00	Mon Feb 2 06:11:13
c0-10	Job	11196	Running	2	-9:00:15:30	41:15:00:00	Sun Jan 25 22:35:56
c0-10	Job	11336	Running	2	-1:14:54:19	41:15:00:00	Mon Feb 2 07:57:07
c0-11	Job	11352	Running	4	-4:13:26	41:15:00:00	Tue Feb 3 18:38:00
c0-12	Job	11348	Running	2	-6:30:40	41:15:00:00	Tue Feb 3 16:20:46
c0-13	Job	11227	Running	4	-7:07:34:11	41:15:00:00	Tue Jan 27 15:17:15
c0-13	Job	11325	Running	1	-3:09:53:40	41:15:00:00	Sat Jan 31 12:57:46
c0-13	Job	11348	Running	2	-6:30:40	41:15:00:00	Tue Feb 3 16:20:46

```
21 nodes reserved
```

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

Other Useful Commands

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

- **canceljob** <JOBID> selectively cancels the specified job(s) (active, idle, or non-queued) from the queue
Example: `$ canceljob 12345`
- **setqos** <QOS> <JOBID> set Quality Of Service for the specified job.
QOS ranges from 0 (lowest) to 8 (highest).
Example: `$ setqos 5 12345`
- **releaseres** <RESERVATION_ID> allows users to release reservation on their own jobs.
Note that releasing a reservation on an active job has no effect since the reservation will be automatically recreated.
Example: `$ releaseres 12345`
- **showstart** <JOBID>
This command displays the earliest possible start time of a job. If the job already possesses a reservation, the start time of this reservation will be reported. If no such reservation exists, this command will determine the earliest time a reservation would be created assuming this job was highest priority. If this job does not have a reservation and it is not highest priority, the value of returned information may be limited.
Example: `$ showstart 12345`

Basic workflow

- 1 Check resources availability with `showbf -S` command
- 2 Edit your PBS script and customize your job parameters (namely the queue in which you want your job will run)
- 3 Run as usual with `qsub scriptname.pbs`
- 4 Eventually check your job status (`checkjob -v <JOBID>`)
- 5 Cross your fingers and wait for the end of the job.

Use the internal node disk

- 1 Each disk node has an internal partition of nearly 60 GB (/state/partition1) with a directory for each user on this space
- 2 To write on this space you need to modify your PBS submission script in the following way:

```
#!/bin/bash
. ~/.bashrc
export SCRATCHDIR=/state/partition1/$PBS_O_LOGNAME
cd $SCRATCHDIR
echo Starting execution at `date`
$PBS_O_WORKDIR/abinis < $PBS_O_WORKDIR/input.file > $SCRATCHDIR/my.output
mv $SCRATCHDIR/my.output /anywhere/you/want
echo Finished at `date`
```

Where:

- \$PBS_O_WORKDIR is the directory from which the job was submitted
- \$SCRATCHDIR is the directory in the node
- \$PBS_O_LOGNAME is your login name

Mandatory flags:

- `#PBS -S /bin/sh` ⇒ defines a shell used by OpenPBS
- `#PBS -N your_job_name` ⇒ selects a name for the job
- `#PBS -q mem_32` ⇒ selects the queue
- `#PBS -l ncpus=1` ⇒ selects the number of cpu you want to reserve

Useful flags :

- `#PBS -l pmem=1000mb` ⇒ requests 1000 mb of memory per processor
- `#PBS -l nodes=2:ppn=2` ⇒ requests two nodes and two processors per node (ppn)
- `#PBS -l cput=4:00:00` ⇒ requests 4 hours 00 minutes of CPU time
- `#PBS -l walltime=4:30:00` ⇒ requests 4 hour 30 minutes of wallclock time (not CPU time!)
- `#PBS -M your.adress@polytechnique.edu` ⇒ sends you an email
- `#PBS -r y (n)` ⇒ sets the job as (not)re-runnable
- `#PBS -m aeb` ⇒ Mail to user on a=abort, b=begin, e=end
- `#PBS -o my_job.pbslog` ⇒ sets the job logname
- `#PBS -j oe` ⇒ Joins stderr to stdout
- `#PBS -e error` ⇒ Output error file separated

MAUI or: How I stop Worrying and Love the Cluster

Andrea Cucca

What MAUI is

MAUI vs PBS

Basic MAUI Concepts

QOS

The new system

Everyday use

showbf

checkjob

showq

checknode

showres

Other Commands

Standard procedures

Improvements

Thanks for your attention

(The End)