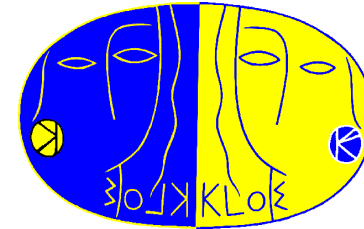

User Batch Queues



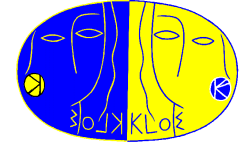
M. Moulson, 25 November 2003

Offline Discussion

Outline:

- *User computer resources*
 - *Current batch queues*
 - *Problems*
 - *Proposals*
-

KLOE user/analysis CPU resources



Currently installed:

fibm11-fibm34 (except fibm16)

23 IBM B80, 4×375 MHz Power3 (92 CPU's)

Typically need 76 CPU's for offline group

fsun01-fsun10

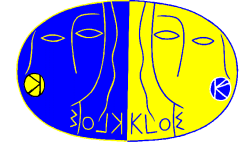
10 Sun E450, 4×400 MHz UltraSPARC-II (~20 B80 CPU's)

Not really used at present

Ordered, available in January?

10 IBM p630, 4×1.45 GHz Power4+ (~100 B80 CPU's)

KLOE user/analysis CPU resources



For 2004 data taking, need ~150 B80 to follow acquisition

Reconstruction, DST production, and MC

Assumes $\mathcal{L} = 1 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ and background same as in 2002

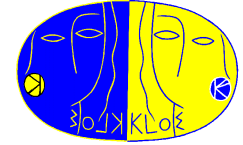
192 B80 available for offline (assuming Sun not used)

~40 B80 on IBM machines can be allocated to users

Sun machines cannot be managed with LoadLeveler (?)

Could be made available for interactive use

Resources currently allocated to users



fibm11-fibm14 (16 CPU)

Interactive use, batch submission

fibm15, fibm17-19 (16 CPU)

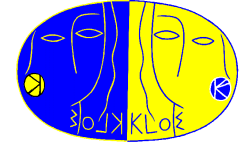
Batch execution

fsun01-fsun02 (~4 CPU equiv.)

Interactive use

Could be possible to allocate two more servers to users

LoadLeveler in a nutshell



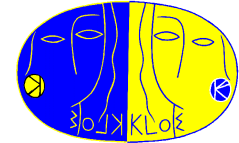
Write .cmd file with job description:

```
# @ class = ktest ! KLOE user queue
# @ job_name = my_job ! Optional simple name
# @ executable = my_script.csh ! Executable script or image
# @ arguments = arg1 arg2 arg3 ! List of args, if any
# @ output = my_logfile ! Standard output redirected here
# @ error = my_errfile ! Standard error redirected here
# @ notification = never ! Default is to e-mail the logfile
# @ queue
```

Use llsbmit to launch job:

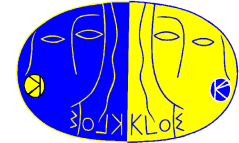
```
llsbmit my_job.cmd
```

Queues currently defined



Name	Slots	Limits	Description
krec	15	none	Offline (bgg)
kcale	1	none	Energy calibration
kcalt	1	none	Timing calibration
kmc	60	none	MC production
ktest	18	none	User analysis
kfast	1	5 min	User analysis

The ktest queue



18 slots total:

5 each on fibm17-fibm19

3 on fibm15

Users can enqueue unlimited number of jobs

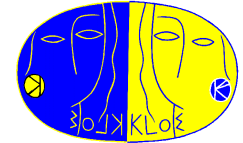
User analysis with 1 job/run in data set = 5000 jobs

Users can occupy up to (only) 6 slots at a time

Job timeout determined by afs token life at job submission

7 days maximum

The ktest queue: problems



Jobs execute on first come/first served basis

If 3 users obtain 6 slots each for 5 days of running, no slots available for 5 days

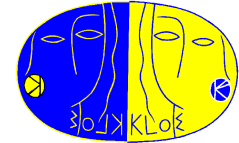
Problem is neither better nor worse if each user submits:

5000 1-minute jobs or

6 833-minute jobs

Either way, the user has an effective lock on slots until all jobs are complete

The ktest queue: solutions?



Trivial provisions:

Expand total number of slots available

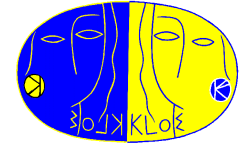
Decrease number of slots available to each user

Easy to decrease number of slots available to user, but:

Is it right to make users wait when queues are empty?

Management of user slot availability needs to be dynamic

The ktest queue: solutions?

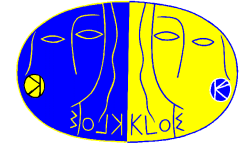


Fabio Fortugno is testing system for dynamic load management

- Implemented at script level
- If a single user has more than 100 jobs in queue, temporarily takes remaining jobs out of queue
- Allows new users to enter
- If no other users enter new jobs, puts jobs back on the queue in batches of 100

Not effective if users submit few, very long jobs

Other solutions



Open up kfast queue (or create **kmed** queue)

- Add many slots
- Allow users a small number of slots (1-2)
- Increase time limit (1-2 hours, maybe more)

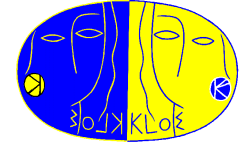
Users who need fast turnaround submit to kfast

Users who need to process large amounts of data can wait by definition and submit to ktest

No point in submitting abusive jobs to kfast

They will just go slowly and/or die

Conclusions proposal



- Open up several new Sun machines to interactive use
 - Close fibm14 (and maybe fibm13) to interactive use and add to ktest
 - Add fibm20 and fibm21 to batch system
 - Create kmed queue, 10-12 slots, 1-2 jobs/user, few-hour time limit?
 - Stay with Fabio's management system for ktest
-