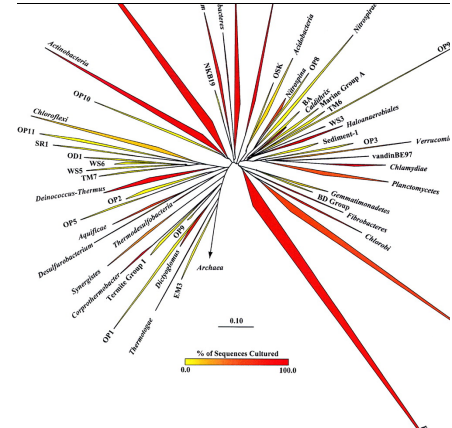
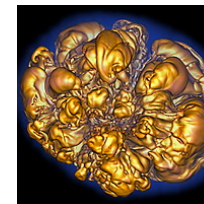
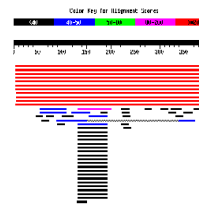
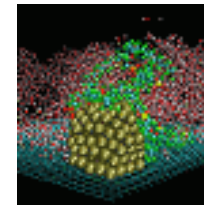
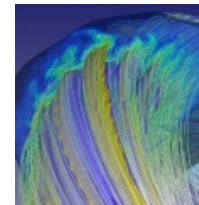
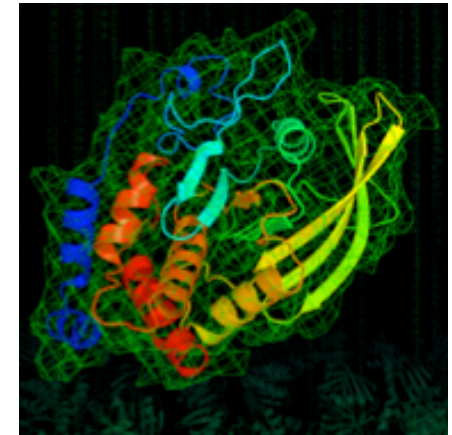


Monitoring and Debugging Jobs on Genepool



Douglas Jacobsen
NERSC Bioinformatics Computing
Consultant

February 12, 2013

Checking on your jobs

- **You've mastered qsub and submitted your jobs**
 - now what?
- **Monitoring jobs in the queue/running**
 - qstat [Univa GridEngine]
 - qs [NERSC]
 - isjobcomplete [NERSC]
 - NERSC genepool website
- **Investigating completed jobs**
 - qacct [Univa GridEngine]
 - qqacct [NERSC]
 - qqacct with qqplot.py [NERSC]
 - NERSC genepool website

```
dmj@phoebe:~$ qstat
job-ID prior name user state submit/start at queue jclass slots ja-task-ID
-----
336024 0.44577 testJob_1 dmj r 02/11/2013 19:30:03 normal.q@sgj07a26.nersc.gov 1
336025 0.39718 testJob_2 dmj r 02/11/2013 19:30:03 normal.q@sgj07b08.nersc.gov 1
336026 0.37289 testJob_3 dmj r 02/11/2013 19:30:03 normal.q@sgj07b13.nersc.gov 1
336027 0.00000 env dmj qw 02/11/2013 19:30:08 1
dmj@phoebe:~$
```

- By default, qstat only shows *your* jobs
- To see others, qstat -u <username> or qstat -u *
- State:
 - r: “running”
 - qw: “queue-wait”
 - R<state>: “rescheduled <basic state>”
 - E<state>: “error <basic state>”
 - h<state>: “hold <basic state>”

qstat – job detail

- **qstat -j <job id>**

```
dmj@phoebe:~$ qstat -j 336025
```

```
=====
job_number:                336025
jclass:                    NONE
exec_file:                 job_scripts/336025
submission_time:          Mon Feb 11 19:29:54 2013
owner:                    dmj
uid:                      56094
group:                    dmj
gid:                      56094
sge_o_home:               /global/homes/d/dmj
sge_o_log_name:           dmj
sge_o_path:               /usr/common/usg/languages/gcc/4.6.3/bin:/usr/common/usg/bin:/usr/common/mss/
common/nsg/bin:/jgi/tools/bin:/usr/syscom/nsg/bin:/opt/uge/phoebe/uge/bin/lx-amd64:/usr/syscom/nsg/opt/M
3.2.9/bin:/usr/local/bin:/usr/bin:/bin:/usr/bin/X11:/usr/games
sge_o_shell:              /bin/bash
sge_o_workdir:            /global/u2/d/dmj
sge_o_host:               phoebe
account:                  sge
hard_resource_list:       ram.c=4G,h_vmem=4G,h_rt=43200
notify:                   FALSE
job_name:                 testJob_2
jobshare:                 0
hard_queue_list:         normal.q
env_list:                 SGE_PPN=1
job_args:                 500
script_file:              sleep
project:                  system.p
binding:                  NONE
binding:                  NONE
1:                         cpu=00:07:44, mem=0.00186 GBs, io=0.00142, vmem=4.000G, maxvmem=4.000G
                           (Collecting of scheduler job information is turned off)
```

qs – “qstat-safe”

- qs is supplied from a cached copy of the job database – the cache is polled every 2 minutes
- qs -S: queue summary

```
dmj@genepool01:~$ qs -S
```

Queue/Resource	r	qw	Eqw	ERq	hRq	hqw
high.q	2:2	0:0	0:0	0:0	0:0	0:0
48-120.c	2:2	0:0	0:0	0:0	0:0	0:0
highmemsys.	2:2	0:0	0:0	0:0	0:0	0:0
long.q	81:135	0:0	7:8	0:0	1:1	691:6279
long_excl.q	6:6	26:26	0:0	0:0	0:0	0:0
120-250.c	2:2	0:0	0:0	0:0	0:0	0:0
250-500.c	3:3	0:0	0:0	0:0	0:0	0:0
1000-2000.c	1:1	0:0	0:0	0:0	0:0	0:0
exclusive.c	6:6	26:26	0:0	0:0	0:0	0:0
highmemsys.	6:6	0:0	0:0	0:0	0:0	0:0
normal.q	46:1400	741:13975	30:556	27:27	1:1	3570:3570
highmemsys.	0:0	0:0	0:0	0:0	0:0	0:0
plantdb.c	1:1	0:0	0:0	0:0	0:0	0:0
normal_excl.q	1:6	18:153	0:0	0:0	0:0	4:4
120-250.c	1:6	1:136	0:0	0:0	0:0	0:0
250-500.c	0:0	1:1	0:0	0:0	0:0	0:0
500-1000.c	0:0	12:12	0:0	0:0	0:0	0:0
exclusive.c	1:6	18:153	0:0	0:0	0:0	4:4
highmemsys.	1:6	14:149	0:0	0:0	0:0	0:0
timelogic.q	0:0	0:0	0:0	0:0	36:36	84:84

Cluster summary

```
Node fractions reserved: 278.05 / 776 (35.83%)
Nodes scheduled: 377 / 776 (48.58%)
Node utilization: 235.67 / 776 (30.37%)
2013-02-11 19:42:36.250129
```

Cache Time

Jobs

Array Tasks

qs – Detail View

- qs detail view shows all jobs by default; the flags AND the jobs and display
- qs -s qw | more
 - View all jobs in state “qw” ordered by priority
- qs -q normal_excl.q -s qw | more
 - View jobs in state “qw” ordered by priority

qs Allowed Flags

```
-s <state>[,<state>]
-q <queue>[,...]
-p <project>[,...]
-u <user>[,...]
-j <job>[,...]
-l <resource>[,...]
```

```
dmj@genepool01:~$ qs -q normal_excl.q -s qw
```

JOBID	ST	PRIOR	USER	PROJECT	QUEUE	NAME	R_N:s	TS	R_RAM/N	R_RAM/s	R_TIME	U_TIME	START/SUB_TIME	TASK
4795953	qw	2.3734	regan	gentech-rnd.	normal	jgi_job_	32		15.6G	15.6G	12:00:00	--:--:--	2013-02-11 14:37:46	
4793371	qw	0.1478	bushnell	prok-assembl	normal	IBYX_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:56:57	
4793373	qw	0.1470	bushnell	prok-assembl	normal	IBYY_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:56:58	
4793374	qw	0.1463	bushnell	prok-assembl	normal	IBYZ_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:00	
4793375	qw	0.1452	bushnell	prok-assembl	normal	IBZA_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:01	
4793376	qw	0.1445	bushnell	prok-assembl	normal	IBZC-1_m	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:02	
4793359	qw	0.1442	bfoster	prok-assembl	normal	command.	1:16	16	120.0G	7.5G	11:00:00	--:--:--	2013-02-11 10:56:22	
4793377	qw	0.1438	bushnell	prok-assembl	normal	IBZC-2_m	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:04	
4793378	qw	0.1428	bushnell	prok-assembl	normal	IBZF_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:05	
4793379	qw	0.1424	bushnell	prok-assembl	normal	ICTO_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:06	
4793380	qw	0.1421	bushnell	prok-assembl	normal	ICTP_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:08	
4793381	qw	0.1418	bushnell	prok-assembl	normal	IGAA_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:09	
4793382	qw	0.1414	bushnell	prok-assembl	normal	IIFY_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:10	
4793383	qw	0.1411	bushnell	prok-assembl	normal	IIFYG_map	1:32	32	1000.0G	31.2G	11:00:00	--:--:--	2013-02-11 10:57:12	
4784554	qw	0.1238	kmfagnan	funga1-annot	normal	phyl_lar	50:16	80	120.0G	7.5G	12:00:00	--:--:--	2013-02-10 15:48:20	
4795161	qw	0.0709	ssunkara	plant-assemb	normal	submit_s	1:16	16	120.0G	7.5G	12:00:00	--:--:--	2013-02-11 12:18:38	
4795169	qw	0.0709	ssunkara	plant-assemb	normal	submit_s	1:16	16	120.0G	7.5G	12:00:00	--:~:~:~	2013-02-11 12:20:27	
4751025	qw	0.0708	eugeneg	plant-assemb	normal	linkedSc	1:24	24	252.0G	10.5G	12:00:00	--:~:~:~	2013-02-06 17:06:26	58-193:~

Requested Nodes : slotsPerNode | Total Slots

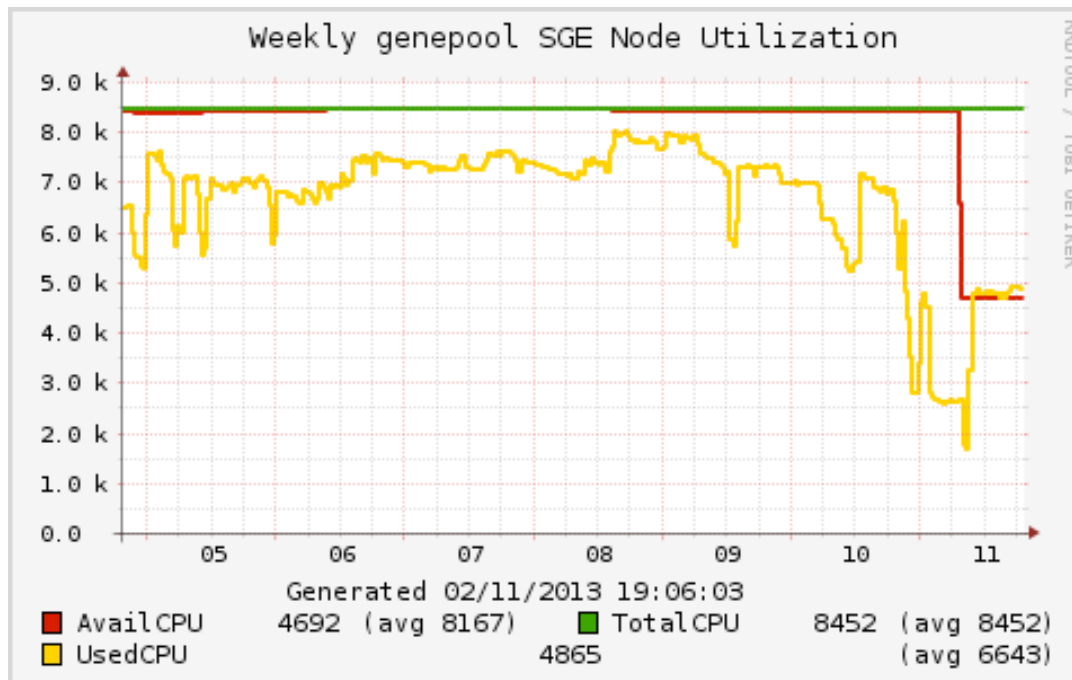
Workflow Control - isjobcomplete

- **isjobcomplete** – calls `qs` then `qstat` to determine if a job number is still in the queue
 - Exit status 0 if the job is not in the queue
 - Exit status 1 if the job is still in the queue – regardless of state (`r,qw,Eqw,...`)

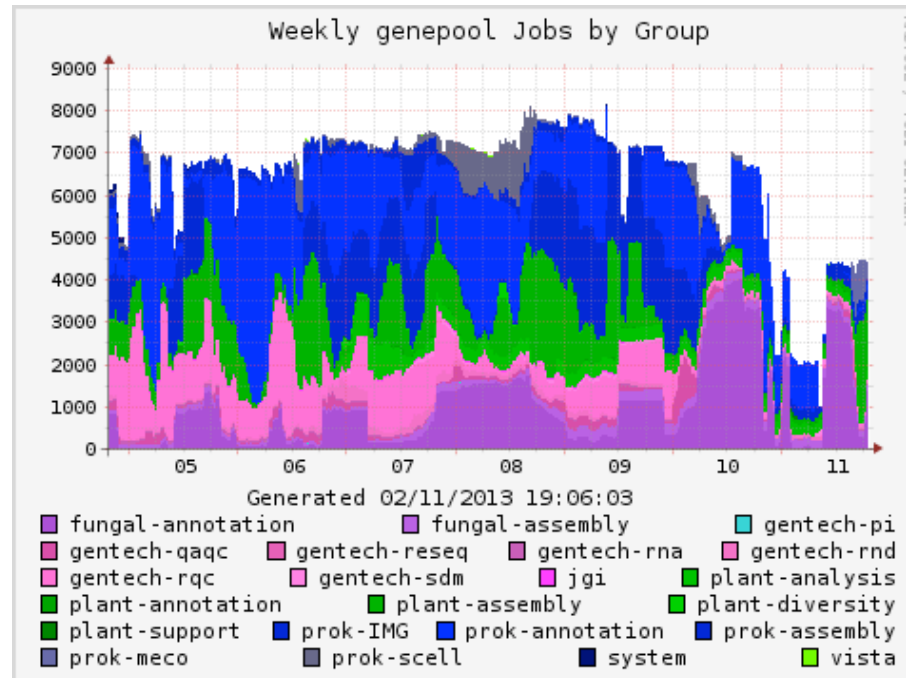
```
dmj@genepool01:~$ isjobcomplete 4792409
4792409 queued/running
dmj@genepool01:~$ echo $?
1
dmj@genepool01:~$
```

- **Benefit: simple, robust call to determine if job is completed**

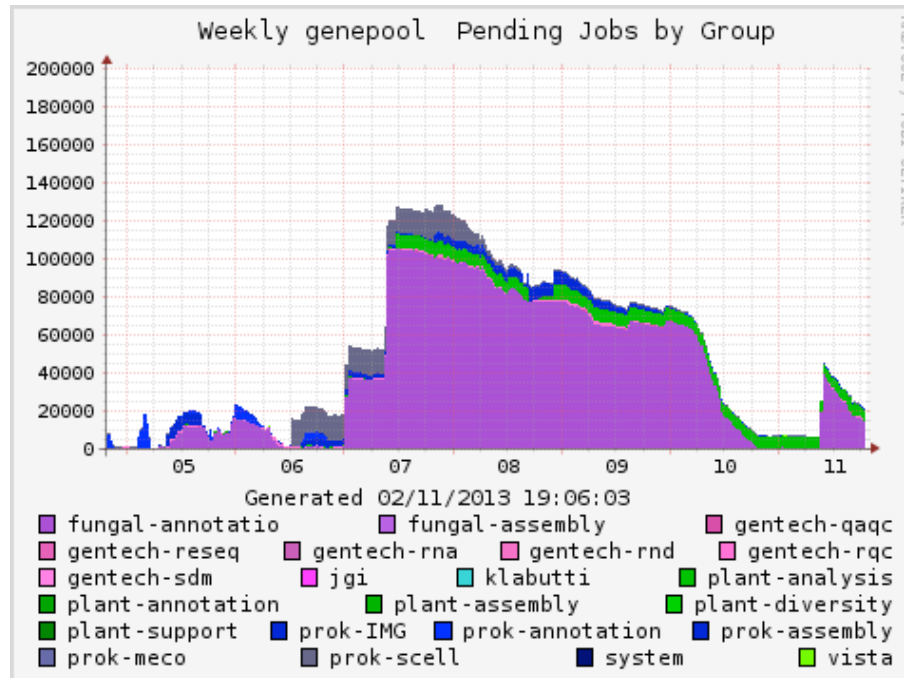
- Search for “genepool cluster statistics”
 - <http://www.nersc.gov/users/computational-systems/genepool/cluster-statistics/>



- Search for “genepool cluster statistics”
 - <http://www.nersc.gov/users/computational-systems/genepool/cluster-statistics/>



- Search for “genepool cluster statistics”
 - <http://www.nersc.gov/users/computational-systems/genepool/cluster-statistics/>



Genepool Completed Jobs

- Search for “genepool completed jobs” on nersc.gov

The screenshot shows a web browser window with the URL <https://www.nersc.gov/users/computational-systems/genepool/genepc>. The page title is "Genepool Queue". The NERSC logo is prominent, with the tagline "Powering Scientific Discovery Since 1974". A navigation menu includes links for HOME, ABOUT, SCIENCE AT NERSC, SYSTEMS, FOR USERS (highlighted), NEWS & PUBLICATIONS, R & D, EVENTS, LIVE STATUS, and STAFF ONLY. The main heading is "GENEPOOL COMPLETED JOBS". Below this, a message states: "This is a Beta site, please send questions and comments to consult@nersc.gov." The search interface includes several sections: "Select a time period" with dropdowns for "Show jobs that completed after" (Feb 4, 2013 @ 0:00 Pacific Time) and "and completed on or before" (Feb 11, 2013 @ 19:18:59 Pacific Time); "Limit the number of jobs to display" with radio buttons for 50, 100, 500 (selected), 1,000, 5,000, and 10,000; "Refine search" with input fields for User, Project, Exec Queue (All), JobID, and Node Name; and "Select columns to display" with checkboxes for Hostname, Job ID, Job Name, User, Exec Queue, Project, Status, Core, Submitted, Start, and Complete. The footer shows "View Page in: CMS Draft Site Published Site" and "Logged in as Douglas M Jacobsen - Log out".

Investigating Completed Jobs

- **GridEngine saves accounting information for all completed and errored-out jobs**
- **These records reflect what your project has been billed for fair-share calculations**
- **Also show the total resource utilization figures**
 - Can be useful (but not perfect) when trying to understand why a job crashed

Investigating Completed Jobs

- **qacct – provided by Univa**
 - Shows detailed and aggregate info about completed jobs
 - NERSC rotates accounting logs daily severely limiting utility of qacct to current days' completed jobs
 - Interface options are essentially: qacct -j <jobid> ; qaact -o <user>
- **qqacct – provided by NERSC**
 - Allows arbitrary query of any field (65 fields), or mathematical transformations of any field in the accounting data
 - Can search any date range
 - Simple queries can be a little verbose to enter
 - Output in csv delimited format for easy post-processing

Investigating Completed Jobs

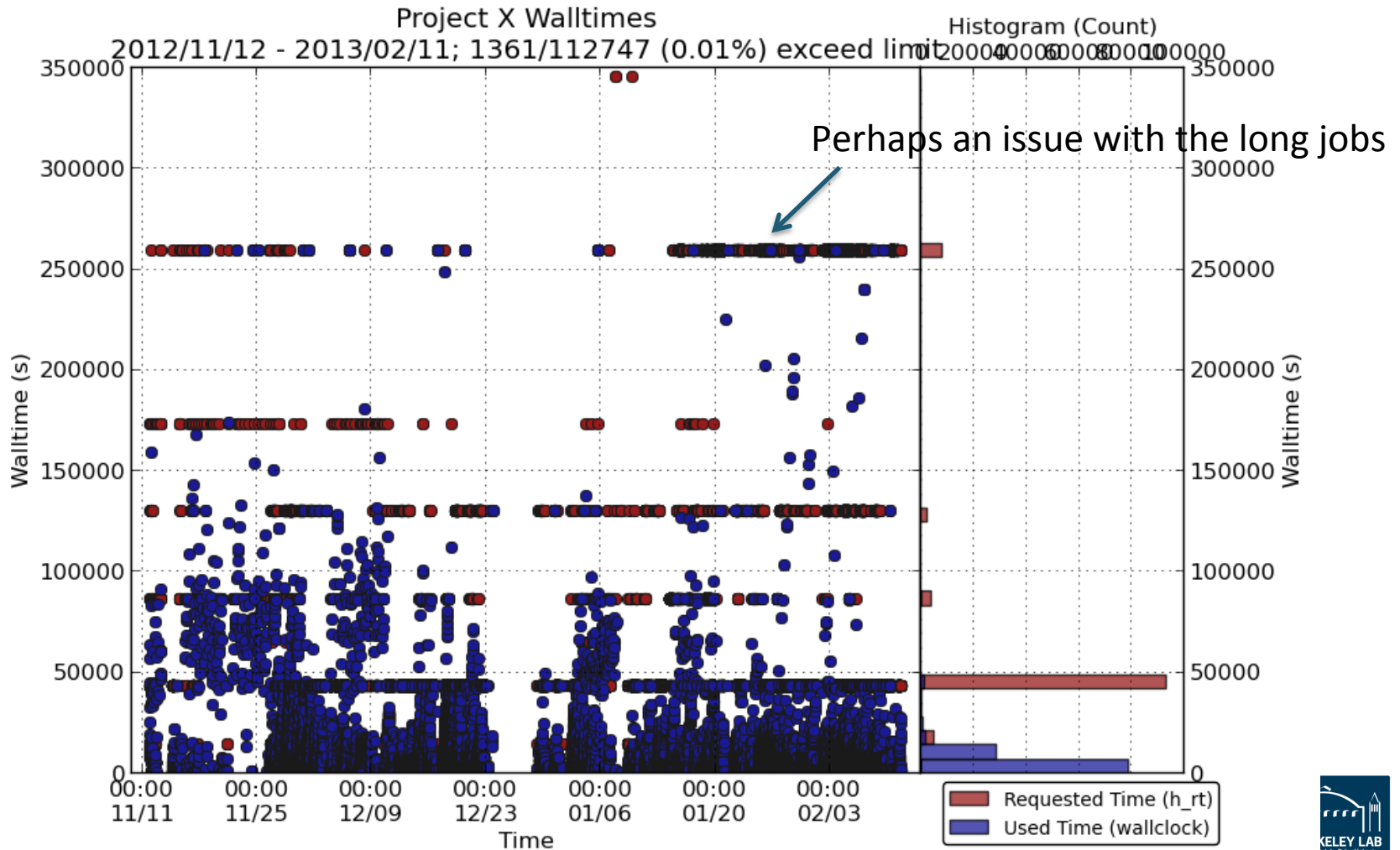
- **Check your jobs for the past 90 days:**
 - `qqacct -D 90 -q 'user=="dmj"'`
- **Just the jobs UGE thinks failed over past 3 days (default)**
 - `qqacct -q 'user=="dmj" && failed != 0'`
- **Just the jobs UGE thinks failed with time/memory info**
 - `qqacct -q 'user=="dmj" && failed !=0' -c
'job_number,failed,memory(ppn*h_vmem),memory(maxvmem),
h_rt,wall'`
- **Always put query in single quotes – the shell is likely to try to parse many of the characters in the query**
- **“-c” overrides default output columns**

Investigating Completed Jobs

Make sure to load
the python module
before using qqplot.py



```
dmj@genepool01:~$ qqacct -D 90 -q 'project=="projectx.p" && end != 0' | qqplot.py  
--wall -t "Project X Walltimes" -o px_wall.png
```

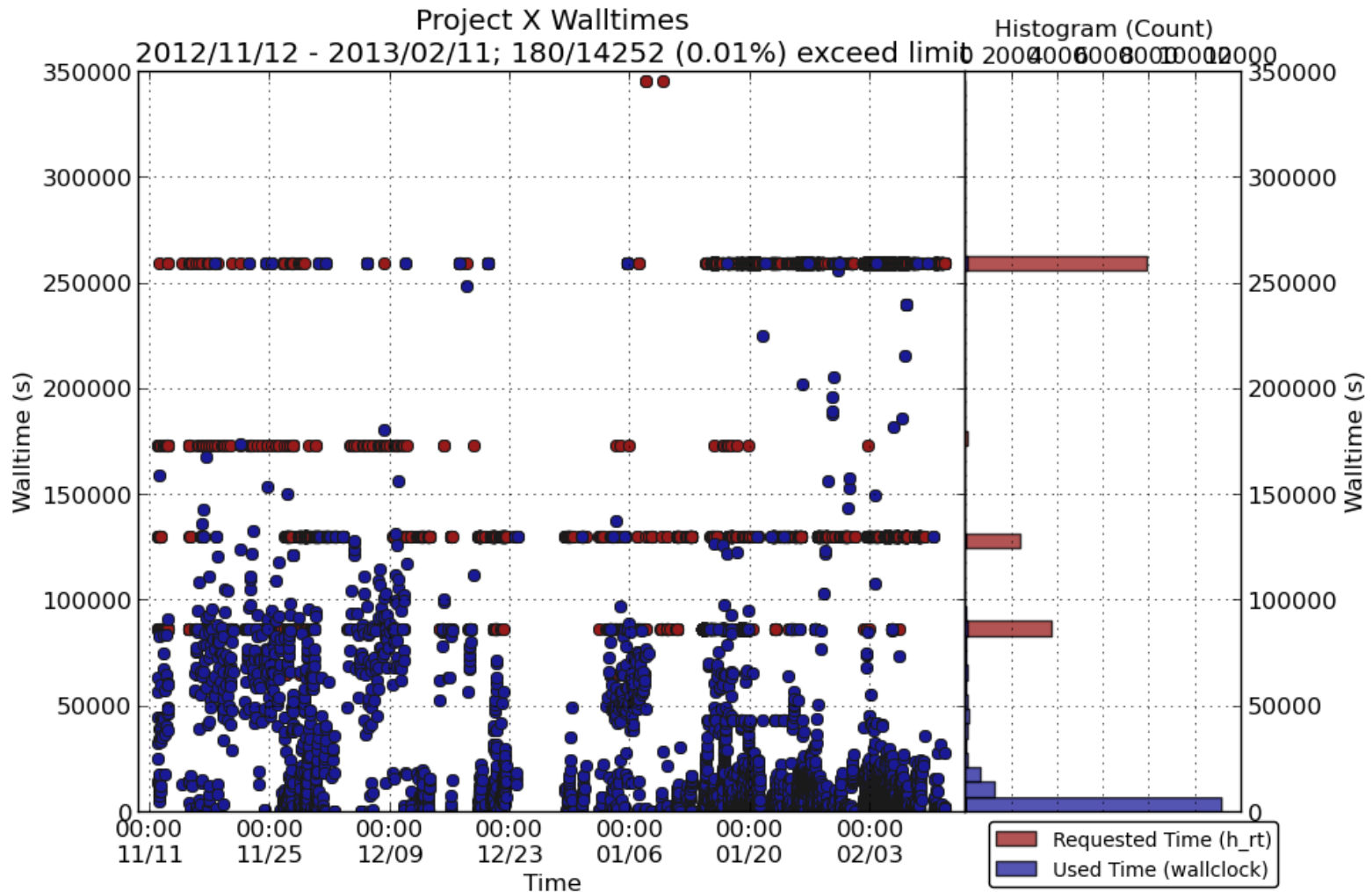


Investigating Completed Jobs

Make sure to load
the python module
before using qqplot.py



```
dmj@genepool01:~$ qqacct -D 90 -q 'project=="projectx.p" && end != 0 && h_rt > 12:00:00' | qqplot.py --wall -t "Project X Walltimes" -o px_wall_long.png
```



Investigating Completed Jobs

Make sure to load
the python module
before using qqplot.py



```
dmj@genepool01:~$ qqacct -D 90 -q 'project=="projectx.p" && end != 0 && h_rt > 12:00:00 && wall > 12:00:00' | qqplot.py --wall -t "Project X Walltimes" -o file.png
```

