

Primer for Site Debugging

This talk introduces key concepts and tools used in the following talk on site debugging

By Jeff Dost (UCSD)

Overview

- Monitoring and Reports
- Logs
- Tools

Overview

- **Monitoring and Reports**
- Logs
- Tools

factoryStatusNow

Select a view: Active only Troubleshoot Period

Total (Default) ▾

Update Table

Reset All Selections

XML last update: Wed Jun 25 03:19:07 2014

Entry Name		Status:								Requested:		Claimed	User run here
		Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle		
CMS_T2_US_Caltech_cit	↑	1254	319	0	319	0	0	0	1	2588	313	980	978
CMS_T1_US_FNAL_ce	↑	869	153	0	153	0	0	0	0	2568	150	847	845
CMS_T2_US_Florida_logw1	↑	761	118	0	118	0	2	0	668	5293	151	824	824

factoryStatusNow

XML last update: Wed Jun 25 03:19:07 2014

Entry Name		Status:								Requested:		Claimed	User run here
		Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle		
CMS_T2_US_Caltech_cit	↑	1254	319	0	319	0	0	0	1	2588	313	980	978
CMS_T1_US_FNAL_ce	↑	869	153	0	153	0	0	0	0	2568	150	847	845
CMS_T2_US_Florida_iogw1	↑	761	118	0	118	0	2	0	668	5293	151	824	824

- **Waiting** and **Pending** are 2 categories of **Idle**
Idle = Waiting + Pending

factoryStatusNow

XML last update: Wed Jun 25 03:19:07 2014

Entry Name		Status:								Requested:		Claimed	User run here
		Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle		
CMS_T2_US_Caltech_cit	↑	1254	319	0	319	0	0	0	1	2588	313	980	978
CMS_T1_US_FNAL_ce	↑	869	153	0	153	0	0	0	0	2568	150	847	845
CMS_T2_US_Florida_iogw1	↑	761	118	0	118	0	2	0	668	5293	151	824	824

- **Waiting** – the glidein never left the factory (only on our local queue)

factoryStatusNow

XML last update: Wed Jun 25 03:19:07 2014

Entry Name		Status:								Requested:		Claimed	User run here
		Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle		
CMS_T2_US_Caltech_cit	↑	1254	319	0	319	0	0	0	1	2588	313	980	978
CMS_T1_US_FNAL_ce	↑	869	153	0	153	0	0	0	0	2568	150	847	845
CMS_T2_US_Florida_iogw1	↑	761	118	0	118	0	2	0	668	5293	151	824	824

- **Pending** – the glidein made it to the site batch system, but it has not been assigned to a worker node yet
(made it to the site queue)

factoryStatusNow

XML last update: Wed Jun 25 03:19:07 2014

Entry Name		Status:								Requested:		Claimed	User run here
		Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle		
CMS_T2_US_Caltech_cit	↑	1254	319	0	319	0	0	0	1	2588	313	980	978
CMS_T1_US_FNAL_ce	↑	869	153	0	153	0	0	0	0	2568	150	847	845
CMS_T2_US_Florida_iogw1	↑	761	118	0	118	0	2	0	668	5293	151	824	824

- If either Waiting or Pending* are high but Running ≈ 0 we should investigate

* High Pending and 0 Running is not necessarily a problem unless no jobs start for a significant period of time (~24 hrs or more)

factoryStatusNow

Status:								Requested:		Client Monitor:						
Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle	Claimed	User run here	User running	Unmatched	User idle	Registered	Info age
1254	319	0	319	0	0	0	1	2588	313	980	978	129526	16	128883	997	265
869	153	0	153	0	0	0	0	2568	150	847	845	78081	31	120723	879	534
761	118	0	118	0	2	0	668	5293	151	824	824	65730	2	109394	826	431

- **Requested Idle** – number the frontend is requesting (pressure value)

factoryStatusNow

Status:								Requested:		Client Monitor:						
Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle	Claimed	User run here	User running	Unmatched	User idle	Registered	Info age
1254	319	0	319	0	0	0	1	2588	313	980	978	129526	16	128883	997	265
869	153	0	153	0	0	0	0	2568	150	847	845	78081	31	120723	879	534
761	118	0	118	0	2	0	668	5293	151	824	824	65730	2	109394	826	431

- A well behaved entry should have*
Req Idle \approx Idle

* An exception is when we significantly limit max idle in the factory config

factoryStatusNow

7 2014

	Status:								Requested:		Client Monitor:						
	Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle	Claimed	User run here	User running	Unmatched	User idle	Registered	Info age
↑	1254	319	0	319	0	0	0	1	2588	313	980	978	129526	16	128883	997	265
↑	869	153	0	153	0	0	0	0	2568	150	847	845	78081	31	120723	879	534
↑	761	118	0	118	0	2	0	668	5293	151	824	824	65730	2	109394	826	431

- The Frontend reports back stats about its User Collector
- **Registered** is the number of glideins actually connected to the Collector
- **Registered** should roughly be equal to **Running**

factoryStatusNow

7 2014

	Status:								Requested:		Client Monitor:						
	Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle	Claimed	User run here	User running	Unmatched	User idle	Registered	Info age
↑	1254	319	0	319	0	0	0	1	2588	313	980	978	129526	16	128883	997	265
↑	869	153	0	153	0	0	0	0	2568	150	847	845	78081	31	120723	879	534
↑	761	118	0	118	0	2	0	668	5293	151	824	824	65730	2	109394	826	431

- We define **Rundiff** as:
Running – Registered
- Rundiff >> 0 should be investigated

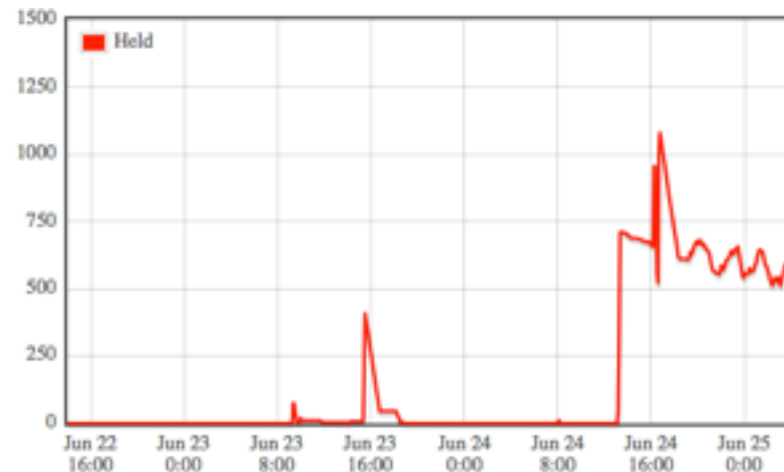
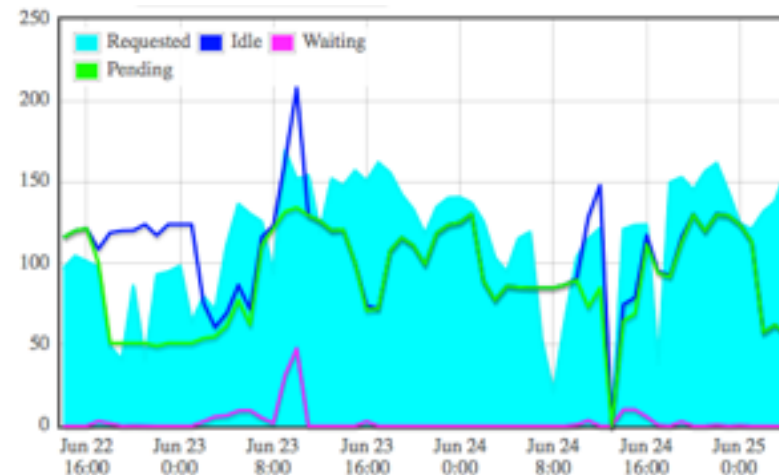
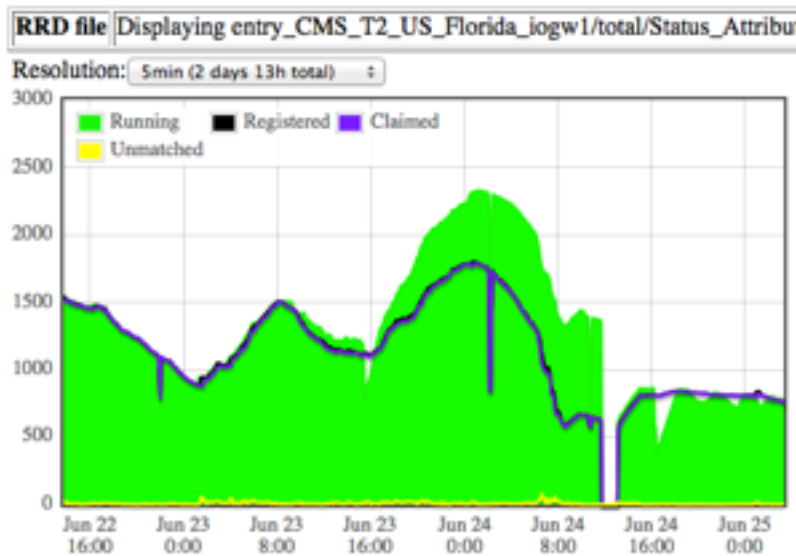
factoryStatusNow

Status:								Requested:		Client Monitor:						
Running	Idle	Waiting	Pending	Staging in	Staging out	Unknown	Held	Max glideins	Idle	Claimed	User run here	User running	Unmatched	User idle	Registered	Info age
254	319	0	319	0	0	0	1	2588	313	980	978	129526	16	128883	997	265
69	153	0	153	0	0	0	0	2568	150	847	845	78081	31	120723	879	534
61	118	0	118	0	2	0	668	5293	151	824	824	65730	2	109394	826	431

- Frontend subcategories of **Registered**:
 - **Claimed** – glideins running user jobs
 - **Unmatched** – glideins available, but 0 jobs match their requirements
 - Registered = Claimed + Unmatched
 - Unmatched >> Claimed should also be investigated

factoryStatus

- Same stats as factoryStatusNow but plotted over time:



analyze_entries

- Report excerpt:

```
frontend_CMSG-v1_0_cmnpilot:
```

	strt	fval	0job		val	idle	wst	badp		waste	time	total
Total/Average	8%	0%	16%		0%	1%	2%	37%		6930	280120	23370
-----	---	---	---		---	---	---	---		---	---	---
CMS_T3_US_PuertoRico_grid0	4%	0%	45%		16%	22%	25%	52%		210	811	165
CMS_T2_US_Purdue_hadoop	64%	23%	72%		23%	10%	34%	70%		199	575	672
CMS_T3_UK_SGrid_Oxford_ce06_medium	1%	1%	86%		2%	57%	60%	69%		30	50	73

- Legend:

```
strt - % of jobs where condor failed to start
fval - % of glideins that failed to validate (hit 1000s limit)
0job - % 0 jobs/glidein
-----
val - % of time used for validation
idle - % of time spend idle
wst - % of time wasted (Lasted - JobsLasted)
badp - % of badput (Lasted - JobsGoodput)
-----
waste - wallclock time wasted (hours) (Lasted - JobsLasted)
time - total wallclock time (hours) (Lasted)
total - total number of glideins
-----
Lasted      - total wallclock time
JobsLasted  - wallclock time used to run jobs
JobsGoodput - wallclock time used by jobs terminatig with exit code 0
```

analyze_entries

- All of the following are counted as **Waste**:
 - Condor failing startup
 - Failing validation
 - 0job
 - Idle

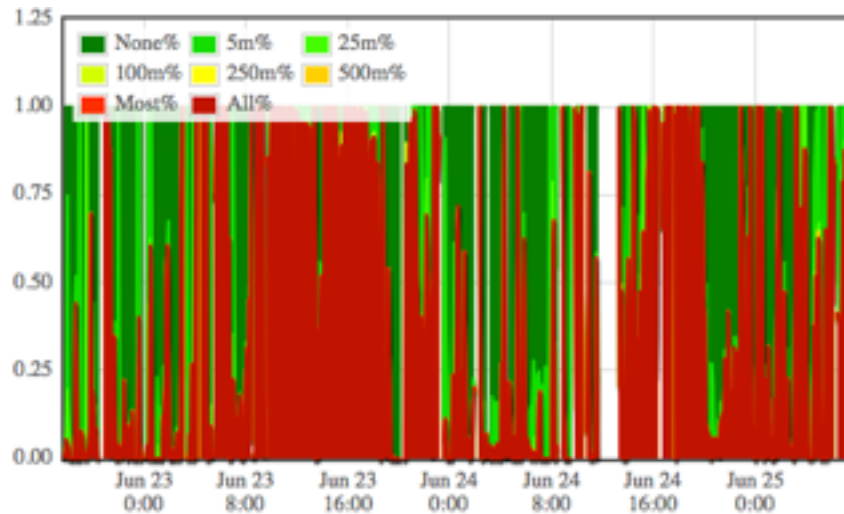
NOTE in this report, idle refers to time glidein spent running but not running user jobs, e.g. Unmatched

- We want to investigate whenever waste is high for an entry

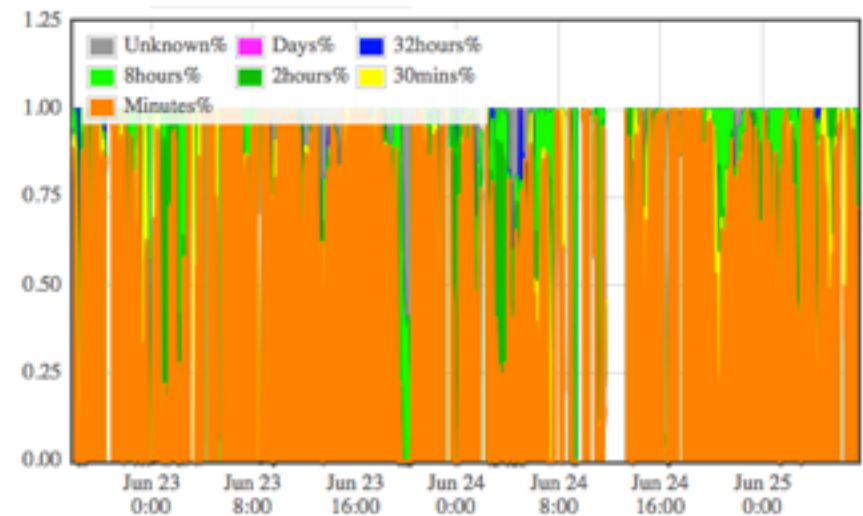
factoryCompletedStats

- Useful to see validation over time and short running glideins:
CMS_T2_US_Purdue_hadoop (has problems!)

Fraction wasted due to node validation



Wallclock time per glidein



Overview

- Monitoring and Reports
- **Logs**
- Tools

Logs

- Glidiens have three logs associated with them, job.*.out, job.*.err, condor_activity
- *.err logs contain compressed condor daemon logs, as well as an XML report containing statistics
- Tools provided to extract the compressed logs:
 - cat_MasterLog
 - cat_StartdLog
 - cat_StartdHistoryLog
 - cat_StarterLog
 - cat_XMLResult

Logs

- `job.*.out` and `job.*.err` logs contain lots of diagnostic info, and also include any stdout or stderr written by validation scripts
- If a validation script provides an XML report, it is often enough to read the summary to discover validation errors

Logs

```
$ cat_XMLResult job.2106335.3.out
<?xml version="1.0"?>
<OSGTestResult logname="job.2106335.3.out" id="glidein_startup.sh" version="4.3.1">
  <operatingenvironment>
    <env name="client_name">CMSG-v1_0.main</env>
    <env name="client_group">main</env>
    <env name="user">cuser13</env>
    <env name="arch">x86_64</env>
    <env name="os">CentOS release 6.4 (Final)</env>
    <env name="hostname">cabinet-8-8-11.t2.ucsd.edu</env>
    <env name="cwd">/data1/condor_local/execute/dir_16146</env>
  </operatingenvironment>
  <test>
    <tStart>2014-06-19T23:48:08-07:00</tStart>
    <tEnd>2014-06-19T23:49:22-07:00</tEnd>
  </test>
  <result>
    <status>ERROR</status>
    <metric name="TestID" ts="2014-06-19T23:49:21-07:00" uri="local">main/validate_node.sh</metric>
    <metric name="failure" ts="2014-06-19T23:49:21-07:00" uri="local">WN_Resource</metric>
    <metric name="CwdFreeKb" ts="2014-06-19T23:49:21-07:00" uri="local">751952</metric>
    <metric name="CwdMinKb" ts="2014-06-19T23:49:21-07:00" uri="local">1048576</metric>
  </result>
  <detail>
    Validation failed in main/validate_node.sh.

    Space on '.' not enough.
    At least 1024 MBs required, found 751952 KBs
  </detail>
</OSGTestResult>
```

Logs

- The `condor_activity` log contains state transitions for each glidein during its lifetime:

```
000 (2118175.001.000) 06/24 13:16:44 Job submitted from host: <169.228.38.36:46438>
...
027 (2118175.001.000) 06/24 13:16:57 Job submitted to grid resource
  GridResource: condor osg-gw-7.t2.ucsd.edu osg-gw-7.t2.ucsd.edu:9619
  GridJobId: condor osg-gw-7.t2.ucsd.edu osg-gw-7.t2.ucsd.edu:9619 533839.0
...
001 (2118175.001.000) 06/24 13:26:57 Job executing on host: condor osg-gw-7.t2.ucsd.edu osg-gw-7.t2.ucsd.edu:9619
...
005 (2118175.001.000) 06/25 01:32:12 Job terminated.
  (1) Normal termination (return value 0)
      Usr 0 00:00:00, Sys 0 00:00:00 - Run Remote Usage
      Usr 0 00:00:00, Sys 0 00:00:00 - Run Local Usage
      Usr 0 00:00:00, Sys 0 00:00:00 - Total Remote Usage
      Usr 0 00:00:00, Sys 0 00:00:00 - Total Local Usage
  0 - Run Bytes Sent By Job
  0 - Run Bytes Received By Job
  0 - Total Bytes Sent By Job
  0 - Total Bytes Received By Job
...
```

Logs

- `condor_activity` for a site with problems:

```
000 (4461966.006.000) 01/17 01:54:10 Job submitted from host: <169.228.38.36:8938?addr=169.228.38.36-8938>
...
012 (4461966.006.000) 01/17 01:55:18 Job was held.
    CREAM_Delegate Error: Received NULL fault; the error is due to another cause: FaultString=[connection
error] - FaultCode=[SOAP-ENV:Client] - FaultSubCode=[SOAP-ENV:Client] - FaultDetail=
[Connection timed out]
    Code 0 Subcode 0
...
013 (4461966.006.000) 01/17 02:00:34 Job was released.
    via condor_release (by user gfactory)
...
012 (4461966.006.000) 01/17 02:01:43 Job was held.
    CREAM_Delegate Error: Received NULL fault; the error is due to another cause: FaultString=[connection
error] - FaultCode=[SOAP-ENV:Client] - FaultSubCode=[SOAP-ENV:Client] - FaultDetail=[Connection timed out]
    Code 0 Subcode 0
...
009 (4461966.006.000) 01/17 07:02:47 Job was aborted by the user.
    via condor_rm (by user gfactory)
...
```

Overview

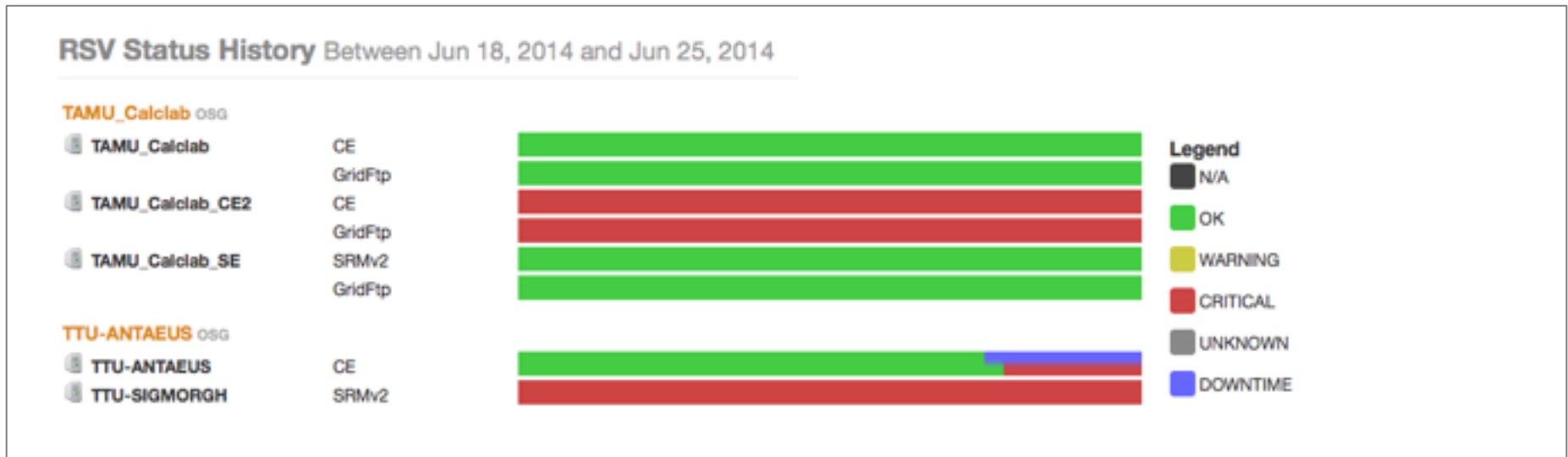
- Monitoring and Reports
- Logs
- **Tools**

Tools

- Summary of tools cited in next talk:
 - **entry_q** – convenience wrapper for condor_q to filter by entry name
 - **entry_ls** – list all .err or .out logs for a particular entry, FE, date combination
 - **get_wns** – extracts workernode hostnames from glidein XML reports

OSG Status Website

- All OSG Sites report to RSV:



- It is useful to check here when a CE is unreachable, to see if the Site is down for maintenance
 - <http://myosg.grid.iu.edu/about>