



# Troubleshooting Batch Job Submission

Shreyas Bhat  
FIFE Workshop  
22 June 2017

- What kind of things do we see the most?
- Why can't I submit my jobs?
- Why isn't my job starting?
- Job is running. Why did it go held?
- Why is the job taking way longer than I expected?
- Fetching logs
- Final notes
- What we want to know in your SNOW ticket

# But first....a quick plug

- Fifemon! <https://fifemon.fnal.gov/monitor/dashboard/db/fifemon-home>
- Multiple experiment/user-specific, as well as general monitoring dashboards
- EXTREMELY USEFUL!
- “There’s a dashboard for that.”™ - K. Retzke

**Why can't I submit jobs?  
Why don't they start?**

- Are you registered properly in VOMS (Virtual Organization Management System)?
  - If not, you might get message like this when jobsub\_submit tries to run voms-proxy-init:  
Error: des: User unknown to this VO.
  - Is your cert registered in VOMS for your VO?
  - Are you using the correct role? (Are you registered with the role you're passing into jobsub\_submit or voms-proxy-init?)
- If you think you have a proxy, check it! What's wrong with this?

```
bash-4.1$ voms-proxy-info -all
```

```
subject : /DC=org/DC=cilogon/C=US/O=Fermi National Accelerator Laboratory/OU=People/CN=Shreyas Bhat/CN=UID:sbhat
```

```
...
```

```
timeleft : 167:59:26
```

```
key usage : Digital Signature, Key Encipherment, Data Encipherment
```

# Authentication, continued

- The problem with previous output - no VOMS extension!
- This is what the output of voms-proxy-info -all should look like:

```
subject : /DC=org/DC=cilogon/C=US/O=Fermi National Accelerator
```

```
...
```

```
type : RFC compliant proxy
```

```
...
```

```
timeleft : 11:59:49
```

```
...
```

```
=== VO fermilab extension information ===
```

```
VO : fermilab
```

```
subject : /DC=org/DC=cilogon/C=US/O=Fermi National Accelerator
```

```
Laboratory/OU=People/CN=Shreyas Bhat/CN=UID:sbhat
```

```
issuer : /DC=org/DC=opensciencegrid/O=Open Science Grid/OU=Services/CN=voms2.fnal.gov
```

```
attribute : /fermilab/minos/Role=Analysis/Capability=NULL
```

```
attribute : /fermilab/annie/Role=NULL/Capability=NULL
```

```
...
```

Great resource for Security Basics (J. Teheran):

<http://cd-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=5884>

# I submitted my jobs; it's been ten minutes.....

- For job to start, resources requested in jobs need to be available
  - Memory (default 2000 MiB)
  - # CPUs (default 1)
  - Time left for slot (default 8 hours)
  - Disk (Default 35,000,000 KiB disk)
  - Allocation
- Remember that changing the defaults can help you get jobs going faster (see Ken's talk)
- Fifemon "Why isn't my job starting?" dashboard (<https://fifemon.fnal.gov/monitor/dashboard/db/why-isnt-my-job-running>)
- Check availability of slots by requirements: <https://fifemon.fnal.gov/monitor/dashboard/db/fifebatch-slots-unclaimed>

# Available slots and allocations (Onsite)

- If experiment is under their allocation for number of slots claimed, a few other things to consider
  - Can available slots run your jobs (job requirements)?
  - User priority
  - Subgroup usage?
  - Fifemon “Experiment Batch Details” dashboard:  
<https://fifemon.fnal.gov/monitor/dashboard/db/experiment-batch-details>
- Condor negotiator will try to negotiate per group, will keep negotiating until quota is filled or no more “available” slots
- Note: Allocations scale with unclaimed slots, groups requesting slots
- Note: Fermilab doesn’t preempt jobs. If available slots can’t run the jobs you want, we won’t stop someone else’s job even if they’re over their allocation.



# My job started to run - what went wrong?

- Jobs that go over requested memory, disk, or time, are automatically held - this is what we mostly see
- “Held”
  - job terminated immediately
  - stdout, stderr are vacated
  - Job goes back in queue - will stay there until released
- “Why are my jobs held” dashboard:  
(<https://fifemon.fnal.gov/monitor/dashboard/db/why-are-my-jobs-held>)
- HoldReason something like “SYSTEM\_PERIODIC\_HOLD Memory/limit 2682/1.9500000000000000E+03”
- Options:
  - Increase requested resources (SNOW Ticket currently)
  - Remove, resubmit jobs with increased requests
  - Remove, reformulate jobs to use less resources, resubmit
- Can run job manually on [gpgtest.fnal.gov](http://gpgtest.fnal.gov) to test - simulate worker node
- Other reasons for held jobs (like glexec errors) exist as well

# Why is my job taking a very long time?

- File transfer issues
- Jobs waiting for files is one of the most common reasons we find. See M. Mengel's talk on Data Handling Best Practices
- Does the file you're trying to transfer exist (e.g. did it disappear from scratch dCache)?
- Overwrite files on dCache failed
- General traffic on dCache or Bluearc CPN locks
- IFDH\_DEBUG=1
- Fifemon dashboards:
  - Experiment overview dashboard  
(<https://fifemon.fnal.gov/monitor/dashboard/db/experiment-overview>)
  - Specific dCache dashboards such as  
(<https://fifemon.fnal.gov/monitor/dashboard/db/dcache-transfer-overview>)

# Cluster Emails (FIFE Job Summary)

- Have been sent since May 30 for each cluster
- Capture resource utilization of jobs - inform users
- Use these to tune resource requests for future jobs
- Can get per cluster or in digest form (daily)
- More info:

[https://cdcvs.fnal.gov/redmine/projects/fife/wiki/Email\\_Reports](https://cdcvs.fnal.gov/redmine/projects/fife/wiki/Email_Reports)

<b>Cluster</b>	
<b>Number of Jobs</b>	4976
<b>Submitted</b>	2017-06-17 17:08:35 -0500 CDT
<b>Owner/Group</b>	
<b>Command</b>	sh
<b>Requested Memory</b>	2400 MiB
<b>Requested Disk</b>	9.8 GiB
<b>Expected Wall Time</b>	0s

[View this cluster on Fifemon](#)

Average time waiting in queue: 1h12m22s

NOTE: Job statistics are collected every 10 minutes.

1886 jobs were excluded from these statistics due to walltime under 10 min.

<b>Used</b>	<b>Min</b>	<b>Max</b>	<b>Avg</b>
<b>Memory</b>	3.4 MiB	2318.2 MiB	1585.7 MiB
<b>Disk</b>	0.0 GiB	1.4 GiB	1.3 GiB
<b>Wall Time</b>	13m39s	78h32m46s	8h35m24s
<b>CPU Time</b>	0s	11h39m49s	7h19m20s

<b>Efficiency</b>	<b>Min</b>	<b>Max</b>	<b>Avg</b>
<b>Memory</b>	0.0%	98.9%	67.5%
<b>Disk</b>	0.0%	14.4%	13.4%
<b>CPU</b>	0.0%	282.9%	82.7%
<b>Time</b>	100.0%	100.0%	100.0%

1 Job(s) are held for the reasons below. For more information see [Why Are My Jobs Held?](#) on Fifemon.

<b>#</b>	<b>Reason</b>	<b>Code</b>
1	SYSTEM_PERIODIC_HOLD: Job exceeded requested resources.	26

[Opt-out or change your subscription preferences.](#)  
[About these reports.](#)

<b>Owner</b>	novapro
<b>Number of Jobs</b>	1871

Average time waiting in queue: 3h46m43s

NOTE: Job statistics are collected every 10 minutes.

708 jobs were excluded from these statistics due to walltime under 10 min.

Requested	Min	Max	Avg
Memory	1900.0 MiB	2400.0 MiB	1949.5 MiB
Disk	9.8 GiB	10.0 GiB	9.8 GiB
Time	0s	0s	0s

Used	Min	Max	Avg
Memory	3.4 MiB	1790.1 MiB	1300.7 MiB
Disk	0.0 GiB	3.4 GiB	1.4 GiB
Wall Time	10m47s	82h1m8s	22h15m7s
CPU Time	0s	27h14m12s	7h36m52s

Efficiency	Min	Max	Avg
Memory	0.0%	83.9%	66.6%
Disk	0.0%	34.9%	14.0%
CPU	0.0%	99.1%	63.0%
Time	100.0%	100.0%	100.0%

1 Job(s) are held for the reasons below. For more information see [Why Are My Jobs Held?](#) on Fifemon.

#	Reason	Code
1	SYSTEM_PERIODIC_HOLD: Job exceeded requested resources.	26

Clusters included in this digest:

Cluster ID	Group	Command
<a href="https://18175292@fifebatch1.fnal.gov">18175292@fifebatch1.fnal.gov</a>	nova	rijeeshk-prod_full_chain_fd_cry_fhc_all-20170617_1707.sh
<a href="https://18176705@fifebatch1.fnal.gov">18176705@fifebatch1.fnal.gov</a>	nova	rijeeshk-fd_data_period2-20170617_1756.sh
<a href="https://18179844@fifebatch1.fnal.gov">18179844@fifebatch1.fnal.gov</a>	nova	rijeeshk-fd_data_period3_draining-20170617_1937.sh
<a href="https://18180214@fifebatch1.fnal.gov">18180214@fifebatch1.fnal.gov</a>	nova	rijeeshk-prod_full_chain_fd_cry_fhc_all-20170617_1704_3.sh
<a href="https://18207044@fifebatch1.fnal.gov">18207044@fifebatch1.fnal.gov</a>	nova	rijeeshk-fd_data_period1-20170618_1124.sh
<a href="https://18280883@fifebatch1.fnal.gov">18280883@fifebatch1.fnal.gov</a>	nova	jwolcott-jwolcott-reco-keepup-Offsite-S16-12-07-neardet-numi-7_days_ago-20170620_0210.sh
<a href="https://18324129@fifebatch1.fnal.gov">18324129@fifebatch1.fnal.gov</a>	nova	rijeeshk-prod_full_chain_nd_genie_fhc_slicertest-20170620_1624-testjobs.sh
<a href="https://18328941@fifebatch1.fnal.gov">18328941@fifebatch1.fnal.gov</a>	nova	rijeeshk-prod_full_chain_nd_numi_fhc_slicertest-20170620_1721.sh

# Error Codes and Logs

- We see many job executables trap signals (e.g. if you get pre-empted somewhere) and with various exit codes
  - This can point you to what kinds of problems job ran into
- Can't underestimate how much good logging helps
  - “If you had to debug this at 3 a.m., what would you want to see?”
  - Good idea to redirect physics executable .err, .out files, ifdh cp back
- jobsub\_fetchlog OR Fifemon “Job Cluster Summary” page (<https://fifemon.fnal.gov/monitor/dashboard/db/job-cluster-summary>)
  - .cmd file = condor submit file for cluster
  - .sh file = executable for cluster
  - \*wrap.sh = jobsub-wrapped executable for cluster
  - .log = HTCondor log (Job submitted, started, got held, finished, etc.) for cluster
  - .err, .out files = stderr and stdout from the individual jobs

# Offsite Issues

- It's tougher to notice and diagnose offsite problems
- If you see a large number of failures at a certain site, please open ServiceNow ticket
- Aside from Fifemon dashboards, daily production email reports indicate this clearly (many experiments get these - let us know if you want them as well)

## Summary

SITE	TOTAL JOBS	FAILED JOBS	SUCCESS RATE
Tusker	933	0	100.0
GPGrid	1949	2	99.9
SU-OG-CE	2648	4	99.8
AGLT2	413	0	100.0
Nebraska	1971	1971	0.0
USCMS-FNAL-WC1	5053	1	100.0
BNL-ATLAS	460	0	100.0



Mu2e Production Jobs Success Rate on the OSG Sites (2017-06-16 06:30:01 - 2017-06-17 06:30:01)

Summary

SITE	TOTAL JOBS	FAILED JOBS	SUCCESS RATE
Tusker	43	14	67.4
UCSDT2	32	0	100.0
AGLT2	227	0	100.0
GPGrid	3523	1	100.0

Failed Job Details (100 clusters shown here)

Click on each job for more details.

CLUSTER	USERID	TOTAL JOBS	FAILED JOBS	JOBID	START TIME	END TIME	SITE	HOST	EXITCODE
21164701	mu2epro	4491	274						
				<a href="mailto:21164701.34@fifebatch2.fnal.gov">21164701.34@fifebatch2.fnal.gov</a>	2017-06-16 16:19:37	2017-06-16 16:49:48	BNL-ATLAS	acas1564.usatlas.bnl.gov	18
				<a href="mailto:21164701.6227@fifebatch2.fnal.gov">21164701.6227@fifebatch2.fnal.gov</a>	2017-06-17 00:17:01	2017-06-17 02:48:33	BNL-ATLAS	acas1044.usatlas.bnl.gov	18
				<a href="mailto:21164701.5359@fifebatch2.fnal.gov">21164701.5359@fifebatch2.fnal.gov</a>	2017-06-16 23:42:41	2017-06-17 02:42:39	BNL-ATLAS	acas1084.usatlas.bnl.gov	18
				<a href="mailto:21164701.720@fifebatch2.fnal.gov">21164701.720@fifebatch2.fnal.gov</a>	2017-06-16 16:31:06	2017-06-16 17:01:10	BNL-ATLAS	acas1073.usatlas.bnl.gov	18

Site Details

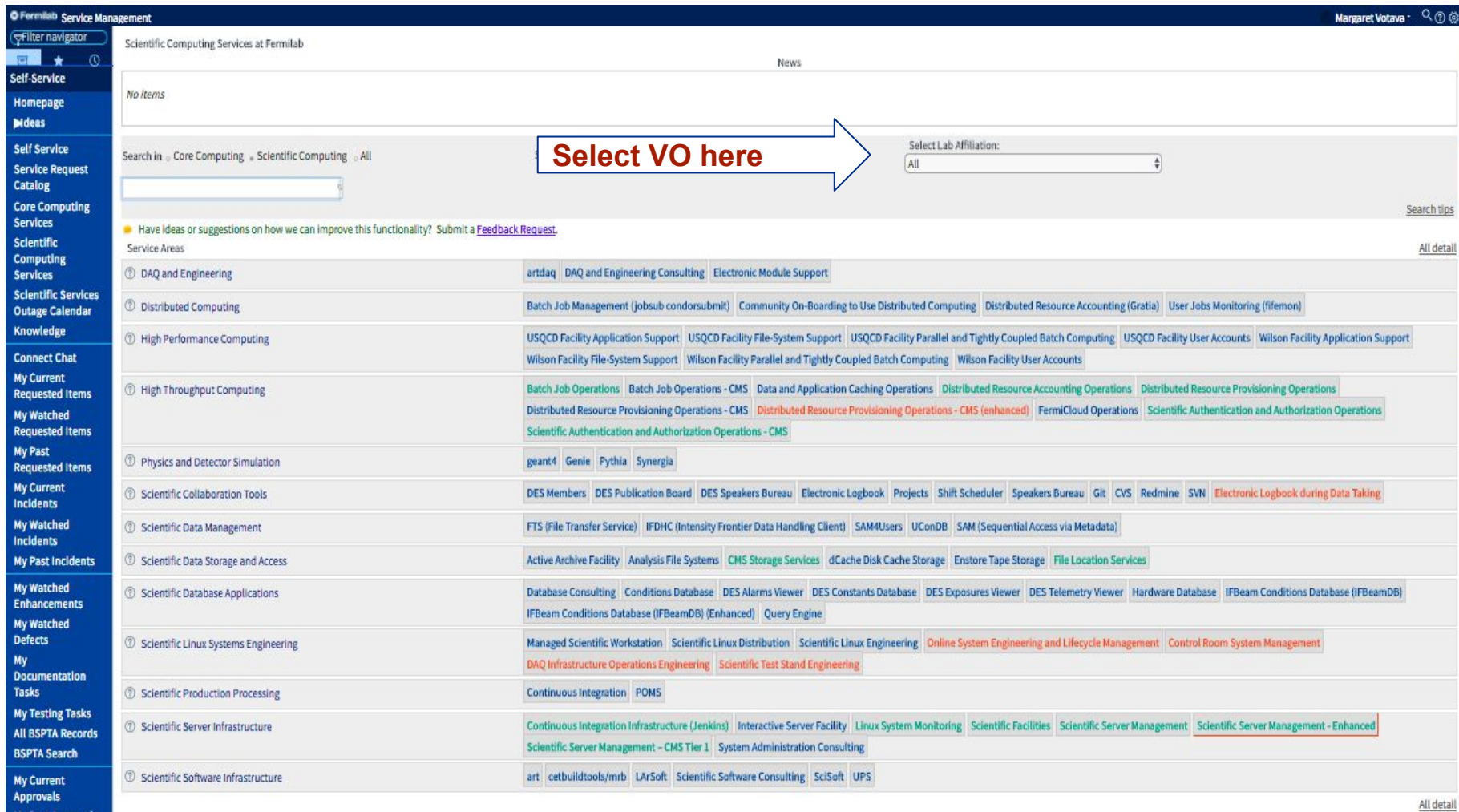
SITE	TOTAL JOBS	FAILED JOBS	SUCCESS RATE	BAD HOST	EXIT CODE	FAILED JOBS ON HOST
BNL-ATLAS	798	588	26.3			
				acas1062.usatlas.bnl.gov	18	67
				acas1086.usatlas.bnl.gov	18	48
				acas1015.usatlas.bnl.gov	18	42
				acas1084.usatlas.bnl.gov	18	36

# What if I couldn't find the issue?

- Consult with experiment computing teams
- File a ServiceNow ticket
  - [www.servicedesk.fnal.gov](http://www.servicedesk.fnal.gov) (may require an 8x5 human)
  - phone x2345 (requires 8x5 human to answer and route)
  - email [servicedesk@fnal.gov](mailto:servicedesk@fnal.gov) (requires 8x5 human to read and route)
- New from last year
  - [Almost] all services are in the SCD service catalog. Please start here when initiating a ticket. (12/1/2016)
  - If you select your affiliation, aka experiment, aka VO, the list of services you will see, are limited to only what your VO uses (6/1/2017)
  - Outage calendar/records. See what is currently down or planned to be down (5/1/2017)

# Where to find things on the service catalog

# SCD catalog before VO filtering



Scientific Computing Services at Fermilab

Search in Core Computing Scientific Computing All

Select Lab Affiliation: All

Have Ideas or suggestions on how we can improve this functionality? Submit a [Feedback Request](#).

Service Areas

DAQ and Engineering	artdaq	DAQ and Engineering Consulting	Electronic Module Support									
Distributed Computing	Batch Job Management (jobsub condorsubmit)	Community On-Boarding to Use Distributed Computing	Distributed Resource Accounting (Gratia)	User Jobs Monitoring (fifemon)								
High Performance Computing	USQCD Facility Application Support	USQCD Facility File-System Support	USQCD Facility Parallel and Tightly Coupled Batch Computing	USQCD Facility User Accounts	Wilson Facility Application Support	Wilson Facility File-System Support	Wilson Facility Parallel and Tightly Coupled Batch Computing	Wilson Facility User Accounts				
High Throughput Computing	Batch Job Operations	Batch Job Operations - CMS	Data and Application Caching Operations	Distributed Resource Accounting Operations	Distributed Resource Provisioning Operations	Distributed Resource Provisioning Operations - CMS	Distributed Resource Provisioning Operations - CMS (enhanced)	FermiCloud Operations	Scientific Authentication and Authorization Operations	Scientific Authentication and Authorization Operations - CMS		
Physics and Detector Simulation	geant4	Genie	Pythia	Synergia								
Scientific Collaboration Tools	DES Members	DES Publication Board	DES Speakers Bureau	Electronic Logbook	Projects	Shift Scheduler	Speakers Bureau	Git	CVS	Redmine	SVN	Electronic Logbook during Data Taking
Scientific Data Management	FTS (File Transfer Service)	IFDHC (Intensity Frontier Data Handling Client)	SAMUsers	UConDB	SAM (Sequential Access via Metadata)							
Scientific Data Storage and Access	Active Archive Facility	Analysis File Systems	CMS Storage Services	dCache Disk Cache Storage	Enstore Tape Storage	File Location Services						
Scientific Database Applications	Database Consulting	Conditions Database	DES Alarms Viewer	DES Constants Database	DES Exposures Viewer	DES Telemetry Viewer	Hardware Database	IFBeam Conditions Database (IFBeamDB)	IFBeam Conditions Database (IFBeamDB) (Enhanced)	Query Engine		
Scientific Linux Systems Engineering	Managed Scientific Workstation	Scientific Linux Distribution	Scientific Linux Engineering	Online System Engineering and Lifecycle Management	Control Room System Management	DAQ Infrastructure Operations Engineering	Scientific Test Stand Engineering					
Scientific Production Processing	Continuous Integration	POMS										
Scientific Server Infrastructure	Continuous Integration Infrastructure (Jenkins)	Interactive Server Facility	Linux System Monitoring	Scientific Facilities	Scientific Server Management	Scientific Server Management - Enhanced	Scientific Server Management - CMS Tier 1	System Administration Consulting				
Scientific Software Infrastructure	art	cetbuildtools/mrb	LarSoft	Scientific Software Consulting	SciSoft	UPS						

# SCD catalog VO filtering (mu2e)

Fermilab Service Management Margaret Votava

Filter navigator

Scientific Computing Services at Fermilab

News

Self-Service

- Homepage
- Ideas
- Self Service
- Service Request
- Catalog
- Core Computing Services
- Scientific Computing Services
- Scientific Services
- Outage Calendar
- Knowledge
- Connect Chat
- My Current Requested Items
- My Watched Requested Items
- My Past Requested Items
- My Current Incidents
- My Watched Incidents
- My Past Incidents
- My Watched Enhancements
- My Watched Defects
- My Documentation

Search in [Core Computing](#) [Scientific Computing](#) [All](#)      Select Visibility Type: [All](#)      Select Lab Affiliation: [E-973 \(Mu2e Muon to ElectronConversion\)](#)

[Search tips](#)

Have ideas or suggestions on how we can improve this functionality? Submit a [Feedback Request](#).

Service Areas [All detail](#)

DAQ and Engineering	<a href="#">artdaq</a> <a href="#">DAQ and Engineering Consulting</a>
Distributed Computing	<a href="#">Batch Job Management (jobsub condorsubmit)</a>
High Throughput Computing	<a href="#">Batch Job Operations</a> <a href="#">FermiCloud Operations</a>
Physics and Detector Simulation	<a href="#">geant4</a>
Scientific Collaboration Tools	<a href="#">Electronic Logbook</a> <a href="#">Git</a> <a href="#">CVS</a>
Scientific Data Management	<a href="#">FTS (File Transfer Service)</a> <a href="#">IFDHC (Intensity Frontier Data Handling Client)</a>
Scientific Data Storage and Access	<a href="#">dCache Disk Cache Storage</a> <a href="#">Ernstape Tape Storage</a>
Scientific Database Applications	<a href="#">Conditions Database</a> <a href="#">Hardware Database</a>
Scientific Linux Systems Engineering	<a href="#">Managed Scientific Workstation</a>
Scientific Production Processing	<a href="#">Continuous Integration</a>
Scientific Server Infrastructure	<a href="#">Continuous Integration Infrastructure (Jenkins)</a>
Scientific Software Infrastructure	<a href="#">art</a>

[All detail](#)

# Outage calendar

Scientific Services Outage Calendar

Filter navigator

Self-Service

- Homepage
- Ideas
- Self Service
- Service Request Catalog
- Core Computing Services
- Scientific Computing Services
- Scientific Services Outage Calendar
- Knowledge
- Connect Chat
- My Current Requested Items
- My Watched Requested Items
- My Past Requested Items
- My Current Incidents
- My Watched Incidents
- My Past Incidents
- My Watched Enhancements
- My Watched Defects

Scientific Services Outage Calendar

MV Margaret Votava

Type of Service Outage

- Unplanned Outage
- Degraded
- Planned Outage
- Maintenance

SCD Service Outage Report

Day Week Month Year June 2017 Today << < > >>

W	Mon	Tue	Wed	Thu	Fri	Sat	Sun
22	29	30	31	1	2	3	4
		8:53 CMS Disk phedex serv		16:16 CMS phedex tape noi 18:32 CMS EOS down. Rest			
23	5	6	7	8	9	10	11
	6:30 SL8500#5 bot 3,3,0,1,C						
24	12	13	14	15	16	17	18
	10:40 About 1100 jobs were		10:40 The NOVA SAM datab	9:00 Replace chassis in Nex 9:07 Replace chassis in Nex 10:20 Replace chassis in Ne		9:30 The fifebatch2 server s	
25	19	20	21	22	23	24	25
26	26	27	28	29	30	1	2
27	3	4	5	6	7	8	9

Highlight based on: Type

# What will help us?

- VO (required)
- Your FNAL username/username of job owner (for example, \*pro) and Role
- Cluster/job ID of jobs
- When jobs were submitted
- Any errors you see
- Where applicable, expected behavior
- Some others might disagree - I personally find it helpful to know what user thinks is the issue, if he/she has an idea

# Other tips/suggestions?