



Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

The Mu2e Experiment at Fermilab : Experience with OSG Opportunistic

Ray Culbertson

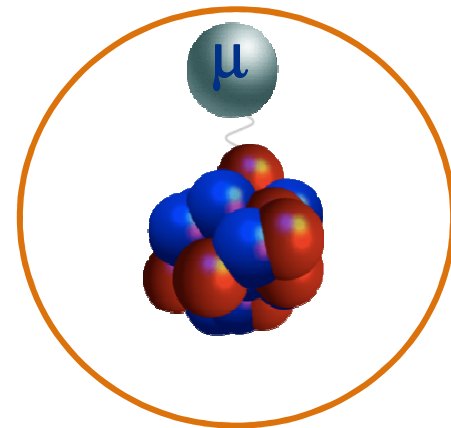
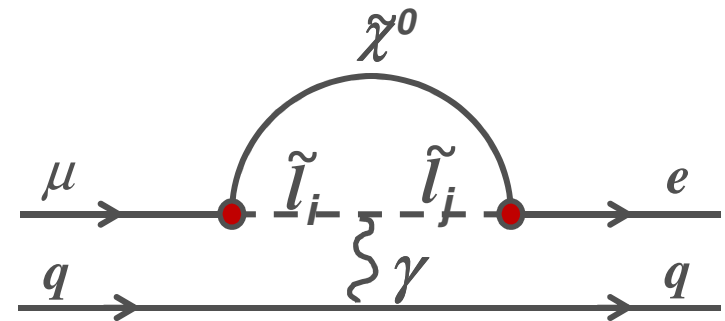
OSG All Hands Meeting 2016

16 Mar 2016

The Mu2e Experiment

The Mu2e Idea

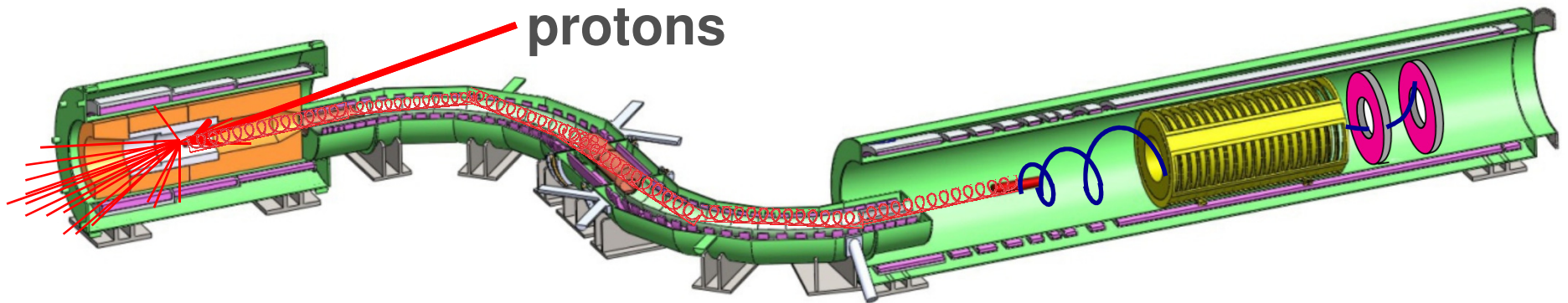
- FNAL, commissioning in 2021
- Search for
Charged-Lepton Flavor Violation
- $\mu^- + N \rightarrow e^- + N$
- Conversion of μ to e (no ν 's) is very(!) small in SM w/neutrino mass
- It does happen in many New Physics scenarios
- Looking for a few signal events in three years of running
- Must Control backgrounds!
- Factor 10^4 improvement over the current limit! (S.E.S of 3×10^{-17} !)



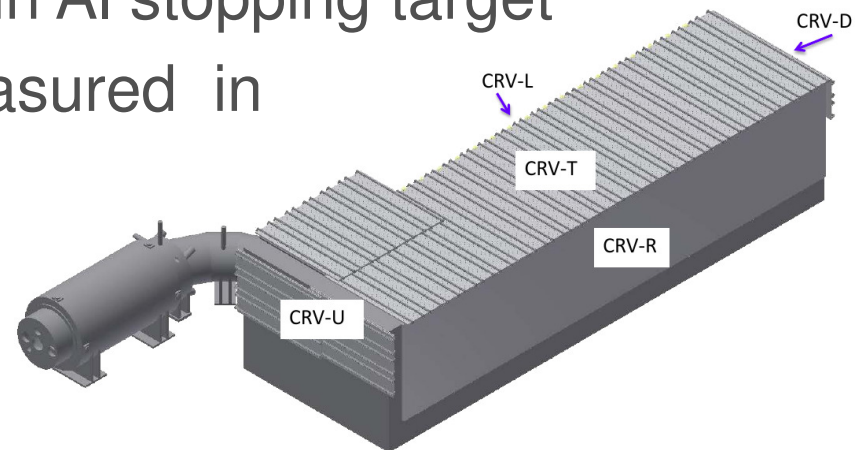
mu2e.fnal.gov
for more info

The Experimental Concept

- 8 GeV proton beam strikes tungsten target
- Graded field pushes pions and muons into solenoid channel



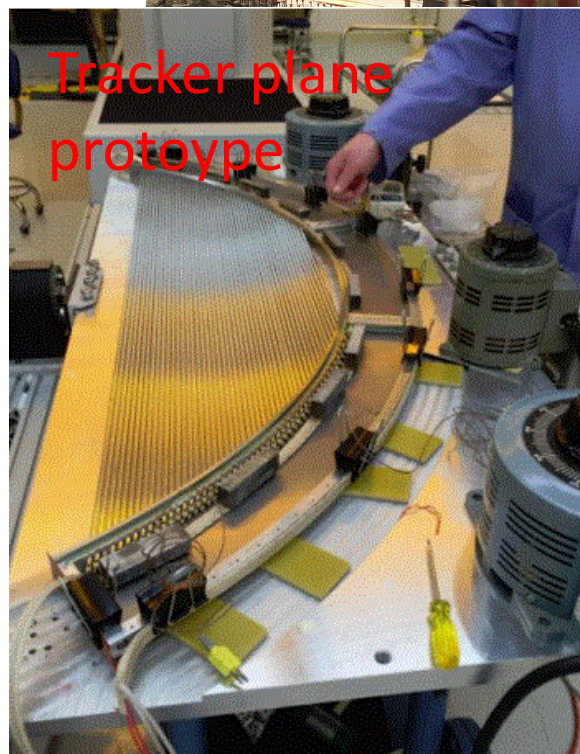
- Pions decay, muons get stopped in Al stopping target
- Any conversion electrons are measured in
 - tracker
 - calorimeter
- Protected by cosmic-ray veto:



Some Nice Pictures



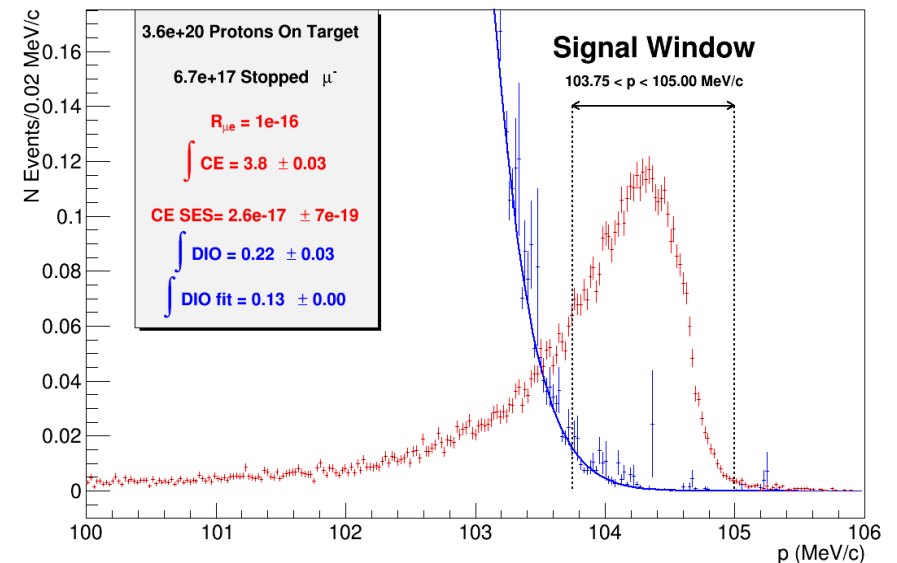
Civil construction (not prototype!)



Simulation Challenges

- We are seeking Dept. of Energy CD-3 approval
- Design has evolved, need to provide updated information for reviews in summer 2016
 - new physics models
 - evolved tracker
 - evolved calorimeter
 - new material
 - new shielding
- What is the updated background estimate?

Reconstructed electron momentum shape for background and conversion signal



- New signal efficiencies?
- Better statistics in several background estimates
- What are the effects of radiation from beam flash?

The CPU Problem and How It Was Solved

OSG Solves the Mu2e CD-3 CPU Problem

- The view from about one year ago...
 - Need 14 Mh CPU in less than 5 months
 - Have 3.5Mh reserved to the experiment on FNAL grid
 - Might get 7Mh, using opportunistic on FNAL grid
- Therefore we need to go to OSG!
- Commissioning
 - Dec 2014 – started commissioning cvmfs
 - Feb 2015 started commissioning OSG
 - Apr 2015 running tests of 5K slots regularly, adding sites
 - Jun 2015 start production
 - Sep 2015 succeed making the baseline checkpoint goal
 - Production almost continuous, and only winding down now
- Met all basic goals, and went on to meet stretch goals!

Thanks (and disclaimer)

- I'm a physicist running simulation jobs, and the following is only a “user perspective” of what it was like to commission and run a project on opportunistic OSG for a year
 - I'm not an expert in OSG software, condor, etc, and don't even have access to those layers
- All technical kudos belong to the FNAL support teams
 - Ken Herner was our primary contact – countless hours..
 - Also critical: Robert Illingworth, Marc Mengel, Bo Jayatilaka, Mike Kirby, Joe Boyd, Tanya Levshina, Mike Ghieth, Dave Dykstra, Chris Green, Yujun Wu, Dmitri Litvinsev, Mike Diesburg ...
 - Also Mu2e physicists Rob Kutschke and Andrei Gaponenko
- All technical misstatements are mine...

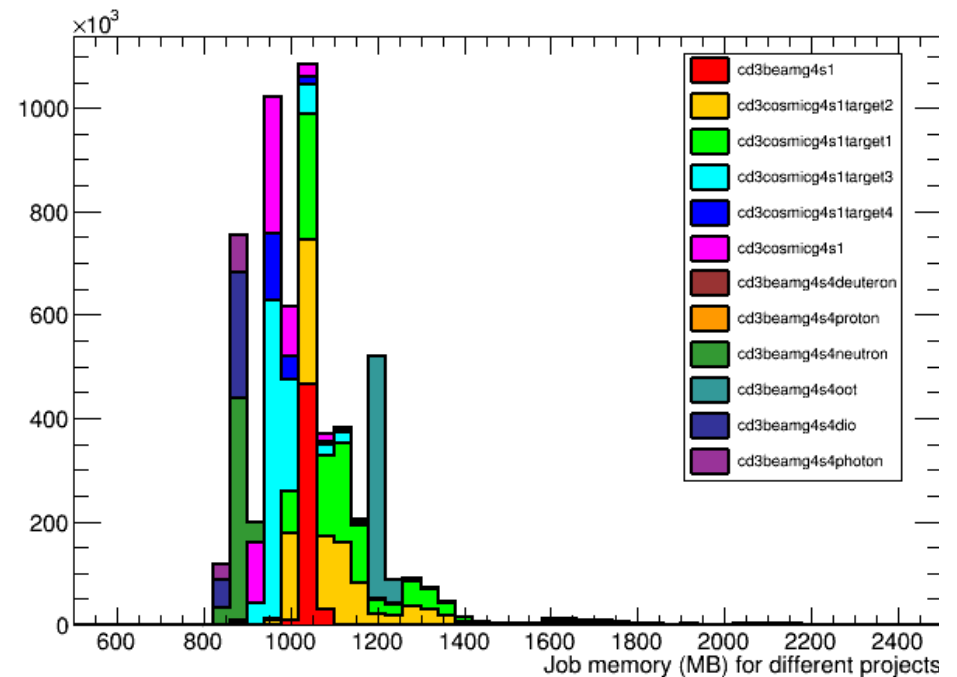
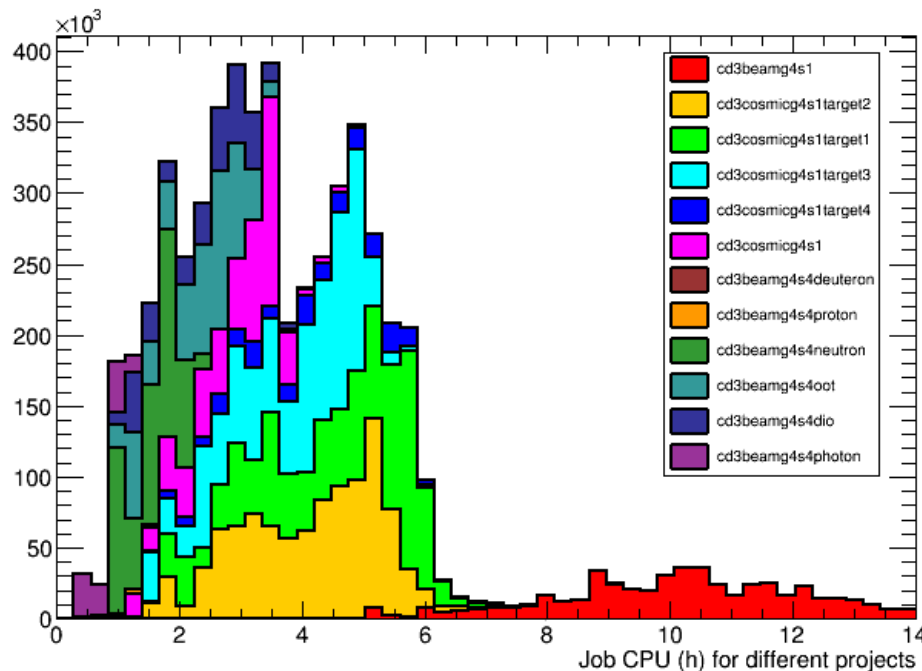
Of course, copious thanks also to all of OSG and sites!!

The Mu2e Simulation Job

- Submit via **jobsub** (FNAL interface to condor)
- Command file moved to worker node by condor
- FNAL infrastructure packages and Mu2e code are on **cvmfs**
- Copy in a small configuration file
 - via **ifdh** (FNAL interface to gridftp and other transport)
- Run Mu2e executable (2-16 h)
- Copy back 5-10 small files (20 MB or less)
 - Also via **ifdh**
- Input and output to **dCache** (distributed disk at FNAL)
- Hand-run run scripts to validate job output
- Move to tape and final tape-backed dCache location
 - **FTS** (Fermilab process to upload files)

What Ran on Sites

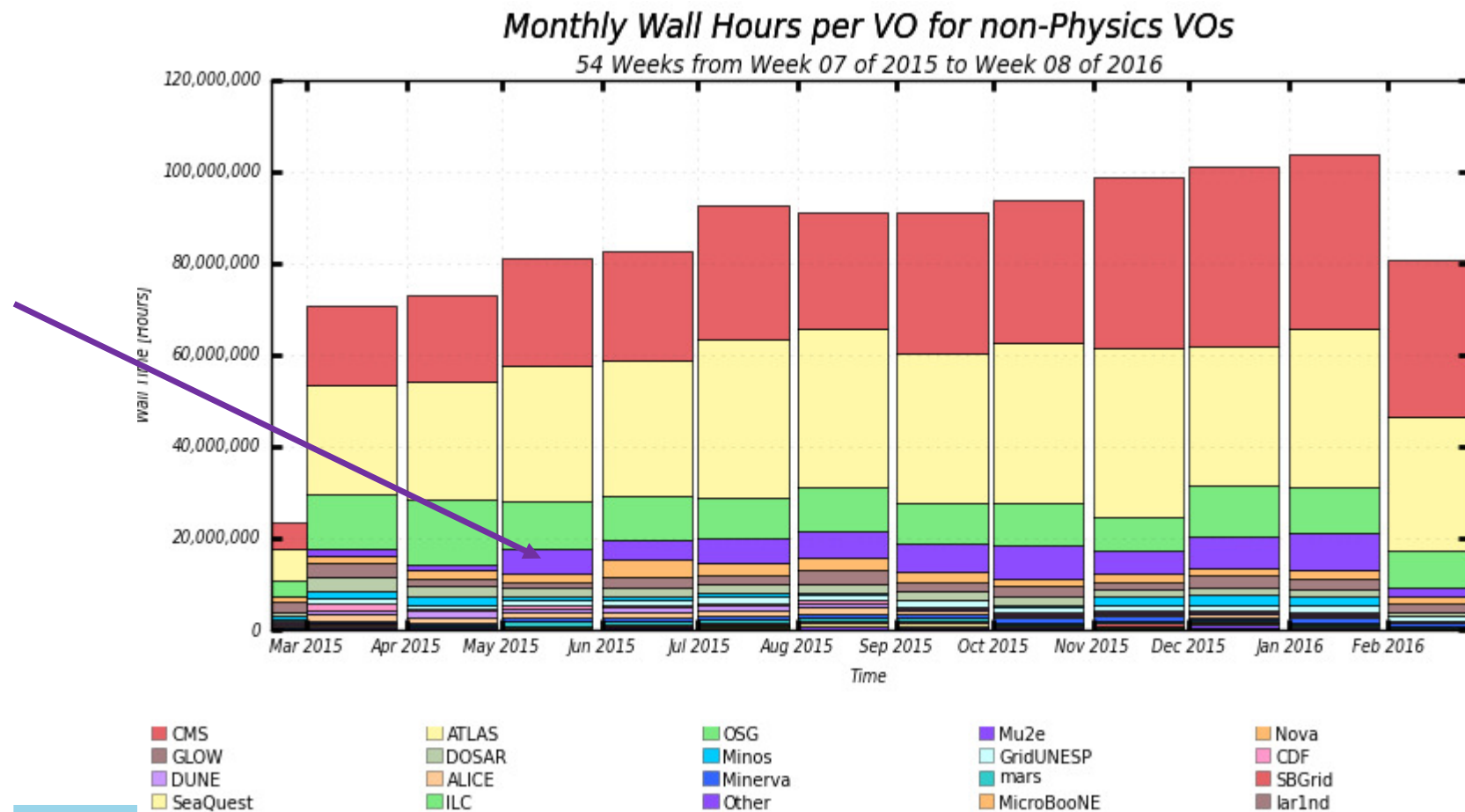
- CPU and memory separated by project



- CPU mostly 2-6h, one job to 16h
- memory 0.8-1.8 GB

Totals – Gratia Summaries

- 5% of all OSG wall time (4th largest),
- 20% of all opportunistic OSG wall time (2nd largest after “OSG”)

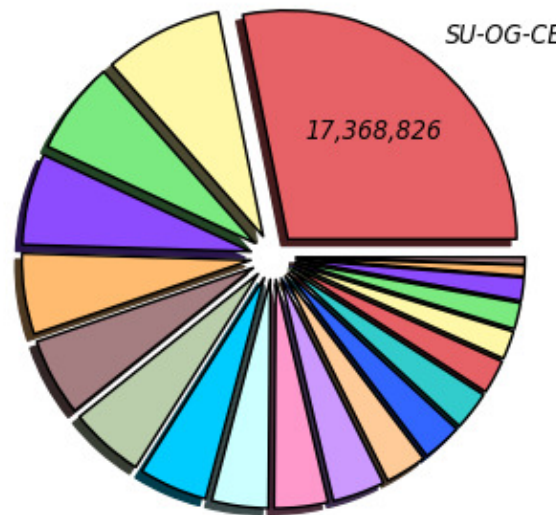


Major Sites

- Syracuse, Caltech, FNAL CMS tier 1, Nebraska – 17+ in all!
- 60 Mh wall time and counting (more than 14 Mh due to stretch goals, new projects, tests, false starts, restarts, CPU variations)

Wall Hours by Facility (Sum: 61,699,270 Hours)

62 Weeks from Week 00 of 2015 to Week 10 of 2016



SU-OG-CE (17,368,826)
FNAL_GPGRID_3 (3,985,037)
FNAL_GPGRID_QUOTA_3 (3,258,628)
FNAL_GPGRID_2 (2,342,681)
MIT_CMS (1,726,765)
Other (1,234,339)
gpce01.fnal.gov (468,755)

CIT_CMS_T2B (5,153,139)
Nebraska-HCC (3,524,217)
CIT_CMS_T2 (2,991,417)
red-gateway1 (2,289,879)
MWT2_CE_UIUC (1,699,213)
NWICG_NDCMS (1,142,957)
BNL_ATLAS_1 (306,665)

USCMS-FNAL-WC1-CE3 (4,047,095)
FNAL_GPGRID_OPP_3 (3,400,501)
GLOW-OSG (2,469,283)
red-gateway2 (1,816,184)
MIT_CMS_2 (1,481,513)
Nebraska (992,176)

Some Details of the Story

Commissioning Starts with a Probe/Debug Job

- Bash script, born of immediately necessity
- printenv, cat /proc/cpu
- Print authentication
- Check OS, glibc, ifdh (gridftp), uberftp, FNAL UPS
- Run cvmfs-uptodate, cvmfs_config stat -v
- Test cvmfs
- Test software setup procedures
- Copy in 1-1000 MB (tunable)
- Run typical Mu2e simulation job (tunable)
- Make validation plots
- Copy back 1-1000 MB (tunable)
- Report all results and job timing in a easily-parsed format

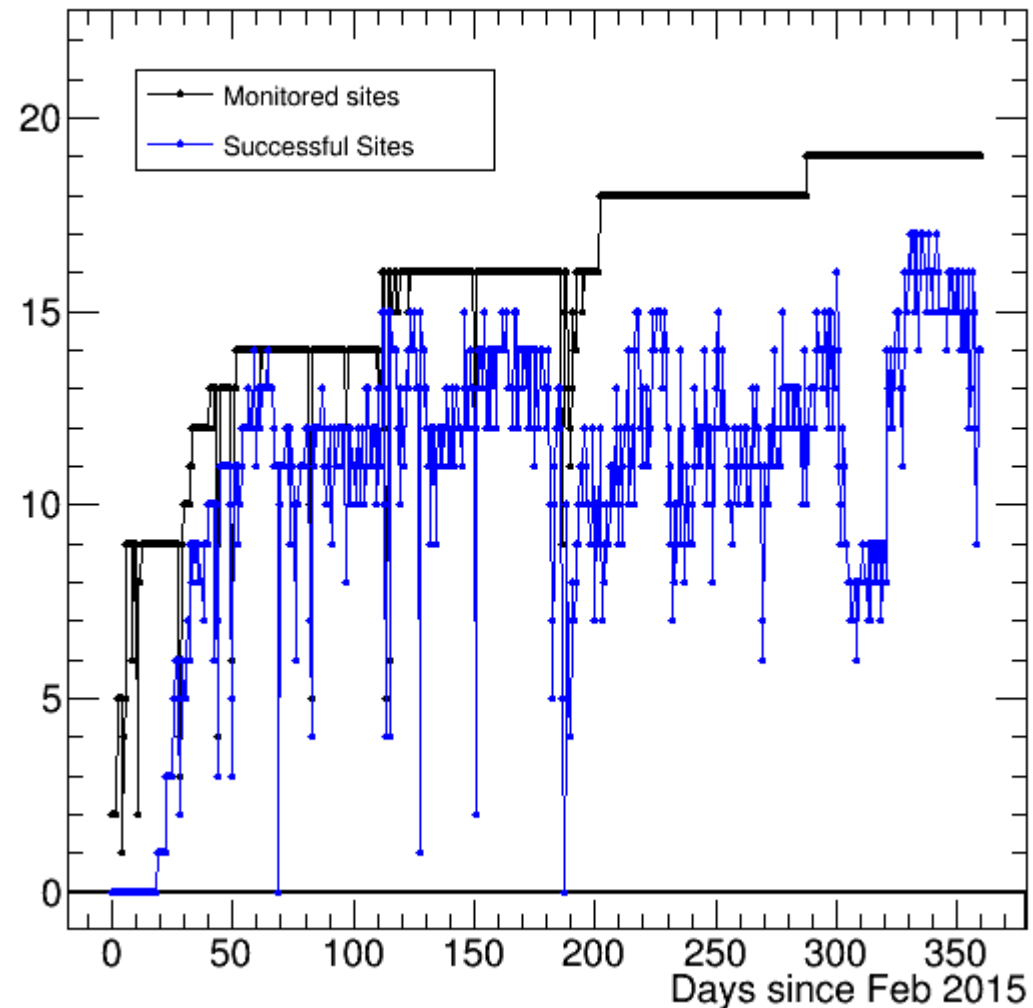
Meeting needs along the way – Site Monitor

- Now submit that probe job on all sites
 - Test the whole infrastructure chain *for Mu2e needs*
- Submit twice a day, kill it if it doesn't run in 12 hours
- Summarize all results
 - A 5 is complete success, lower means some step failed

Time	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site
2016-02-22 06:00	5	5	5	5	0	5	5	5	5	5	5	5	5	0	5	4	5	0	5
2016-02-21 18:00	5	5	5	5	0	5	5	5	5	5	0	5	5	0	5	0	5	0	5
2016-02-21 06:00	5	5	5	5	0	5	5	5	5	5	5	5	5	0	5	0	5	0	5
2016-02-20 18:00	5	5	5	5	0	5	0	5	5	5	5	5	5	0	5	0	5	0	5
2016-02-20 06:00	5	5	5	5	0	5	0	0	5	5	5	5	5	0	5	0	5	0	5
2016-02-19 18:00	5	5	5	5	0	5	5	0	5	5	5	5	5	0	5	0	5	0	5
2016-02-19 06:00	5	5	5	5	5	5	5	0	5	5	5	5	5	0	5	4	5	0	5
2016-02-18 18:00	5	5	5	5	5	5	5	4	5	0	5	5	5	0	5	4	5	0	4
2016-02-18 06:00	5	5	5	5	0	5	0	0	5	5	5	5	5	0	5	0	5	0	5
2016-02-17 18:00	5	5	5	5	0	5	5	0	5	5	5	5	5	0	5	0	5	0	5
2016-02-17 06:00	5	5	5	5	0	5	5	0	5	5	5	5	5	0	5	0	5	0	0
2016-02-16 18:00	5	5	5	5	0	5	5	0	5	5	5	5	5	0	5	0	5	0	5
2016-02-16 06:00	5	5	5	5	0	5	0	0	5	5	0	5	5	0	5	4	5	0	5
2016-02-15 18:00	5	5	5	5	0	5	5	5	5	5	0	5	5	0	5	0	5	0	5

The Site Probe Monitor Tells the Long-term Story

- Summary of monitored sites over the year
- Include trace of sites with complete, level 5, success
- Maintaining 12-15 useable sites
- Dips around Aug and Holidays



One Year of Servicedesk Tickets

A very rough sorting/analysis of 179 tickets related to Mu2e production in the last year...

- 51 submission infrastructure
 - Jobsub, fifebatch, condor, glideins, monitoring, condor logs
- 37 dCache
- 35 file handling at FNAL (SAM, FTS and other)
- 26 ifdh and gridftp
- 18 CMVFS
 - Many of these are problems at specific OSG sites
- 12 OSG site issues

Scale is 2/week for lab infrastructure, 1/week for OSG sites, plus numerous email threads, conversations

Some Typical Issues local to FNAL

- Fifebatch (condor servers) overloaded or crashed
- “Sandbox”
 - No disk space, Cant change ownership
 - Glide-in problems
- Monitoring down or incorrect
- Gridftp servers
 - overloaded
 - rejecting authentication
- dCache
 - Overloaded, not responding
 - Components crashed
 - Missing directory entries
 - Rate issues

Site Issues I

- Local software
 - uberftp not at latest version, triggers known bugs
 - /usr/bin/time command not installed (reports memory)
 - eventually we customized it and put on cvmfs
 - pass all signals, memory incorrect by a factor of 4
- Kernels
 - Request SL6, see 99% 2.6 1% 3.x 0.1% 4.x
 - Also matched SUSE, CentOS with no problems
 - This variation mostly manifested in FNAL “setup” infrastructure
- Optional libraries
 - Some sites do not install X11 display libraries
 - We now provide them on cvmfs and include them in library path
 - Developing graphics-free builds, but would rather it just works

Site Issues II

- Authentication
 - By VO or user
 - Hard to differentiate from no slots available and ad mismatch
- CMVFS
 - Not mounted, wrong version
 - Cache not up to date
 - Corrupt, causes seg faults and missing files
- Single-node black holes
 - Often CVMFS errors
 - Hardware errors: seg fault, bus error, input/output error, disk full
- Job restarts (see next)

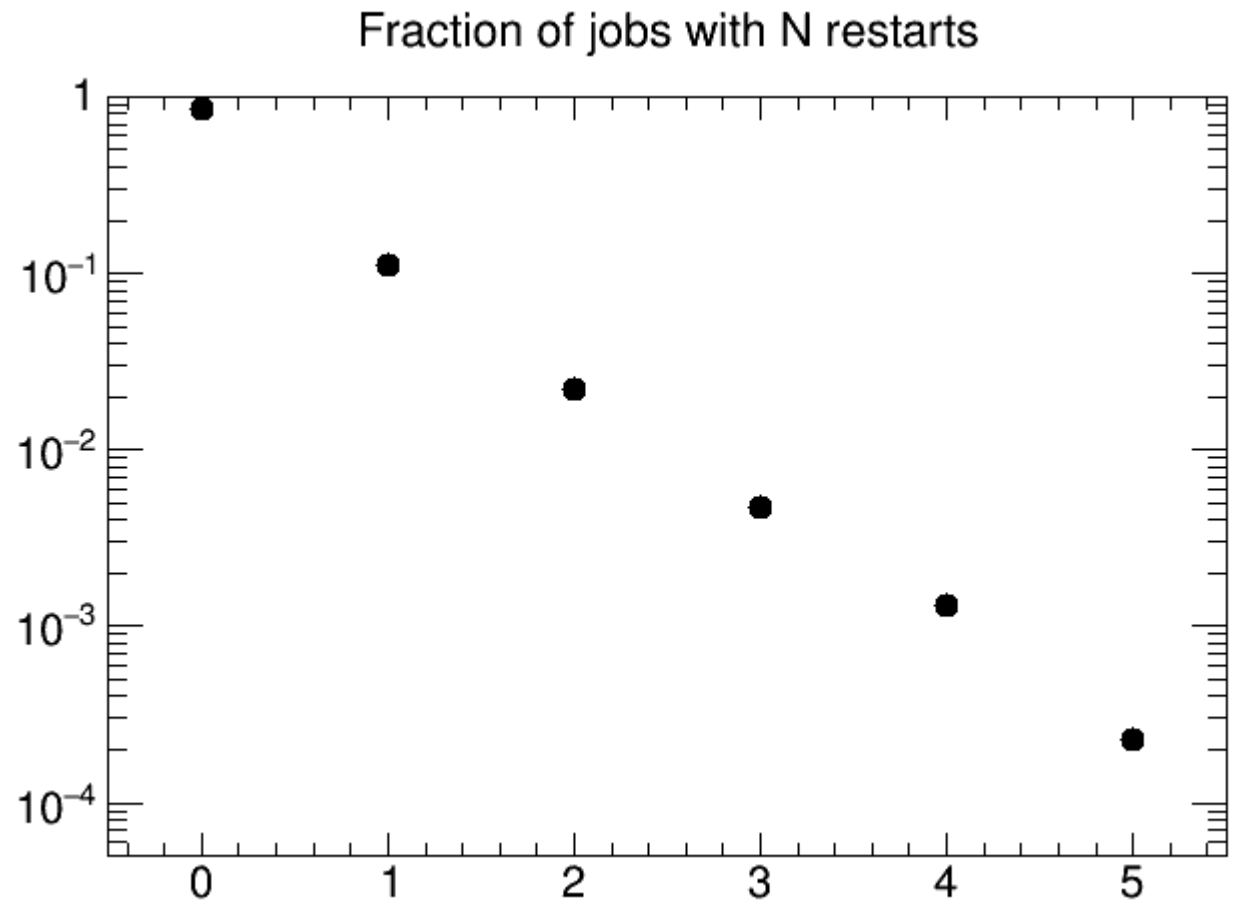
Job Restarts are Disappointing

- By condor log status
- Probability for a job to get disconnected or evicted by site
- Overall 17% of jobs get disconnected or evicted
- Some evictions are reported as disconnects

	Disconnect Prob (%)	Eviction Prob(%)
Site	0	0
	3	0
Site	0	3
	16	3
Site	1	0
	0	0
Site	54	5
	73	0
Site	1	4
	0	1
Site	44	0
	16	0
Site	34	0
	10	0
Site	2	0
	36	0

Job Restarts

- The number of times a single job got restarted
- Submissions of 10K jobs usually see a long tail of restarts, up to 4 or 5 times
- Disconnect confirmation timeout adds ~6 h each time
- We usually have to wait for a submission to finish to proceed
= several days delay, then resubmit failures, repeat



Meeting needs along the way – Gratia Report

- Daily email (gratia via Tanya Levshina)
- 24-hour response to site and single-node black holes



Mu2e Production Jobs Success Rate on the OSG Resources (2016/02/24 07:00:01 - 2016/02/25 07:00:01).

Summary

SITE		TOTAL JOBS		FAILED JOBS	SUCCESS RATE
Tusker		926		0	100.0
UCSDT2		1307		0	100.0
GPGri	Sandhills	625	5	99.2	
AGLT				c6003.sandhills.hcc.unl.edu	126 1
USCN				c6003.sandhills.hcc.unl.edu	134 1

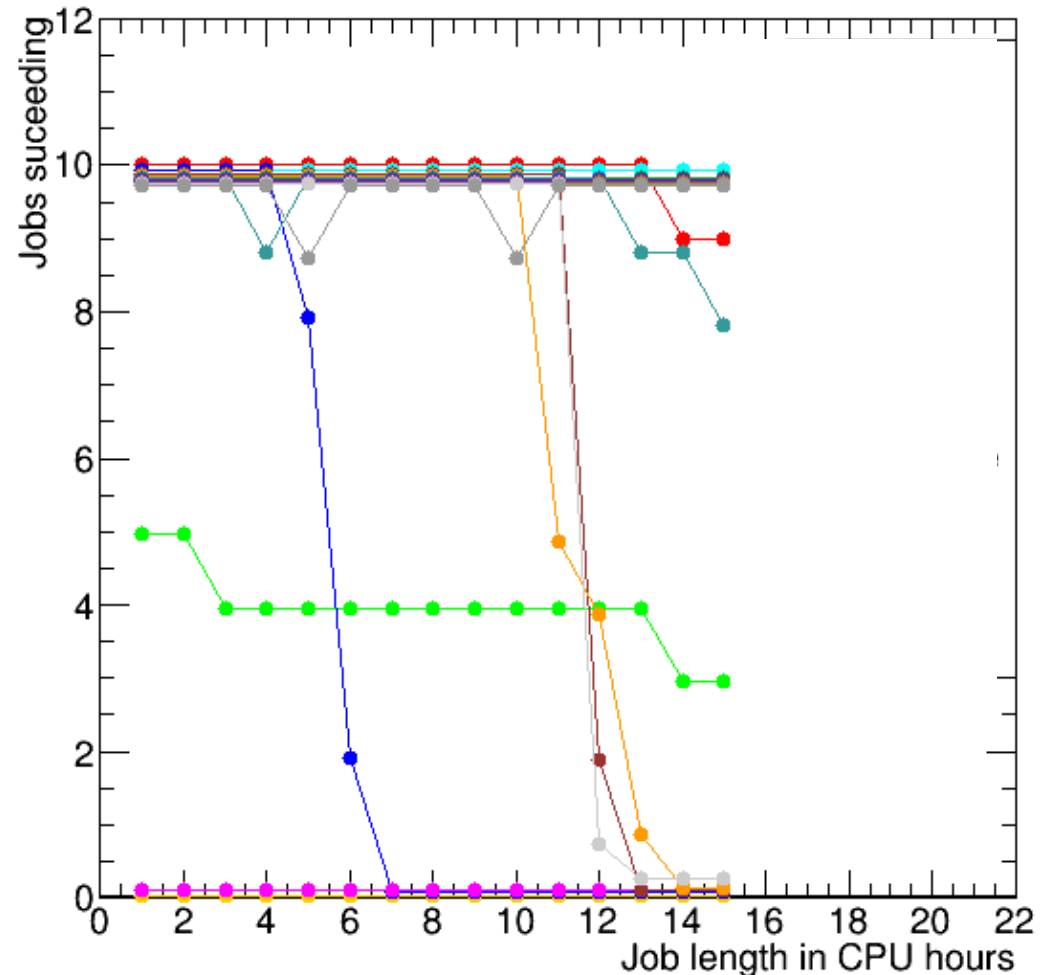
Meeting needs along the way – Monitor Condor logs

- Every hour, unpack all condor logs of recent jobs
- Primary purpose is to “immediately” see if there are widespread failures
- Also by site
- More work required to find a single problematic node
- FNAL starting up similar monitors!

Date	hr	run	start	end	fail	disc(%%)	evict	shadow
02/25	12	6876	972	941	1	372(5)	0	1
02/25	11	7239	2409	2684	1	242(3)	0	3
02/25	10	7870	682	404	28	227(2)	0	9
02/25	09	8008	4800	1177	209	378(4)	28	194
02/25	08	6399	227	97	97	323(5)	25	183
02/25	07	6496	113	10	10	208(3)	2	14
02/25	06	6506	273	0	0	571(8)	0	2
02/25	05	6393	558	0	0	352(5)	0	2
02/25	04	6239	1135	711	0	415(6)	0	18
02/25	03	6950	1509	1172	1	281(4)	0	2
02/25	02	8120	1909	1216	1	615(7)	0	8
02/25	01	9297	1789	1470	2	556(5)	0	1
02/25	00	10568	1529	1068	1	354(3)	0	1
02/24	23	11523	1829	942	1	270(2)	0	0
02/24	22	12225	1394	885	0	586(4)	0	2
02/24	21	12814	980	788	1	294(2)	0	1
02/24	20	13061	1221	753	0	333(2)	0	2
02/24	19	13224	1008	851	0	433(3)	0	4
02/24	18	13279	1018	926	0	518(3)	0	0
02/24	17	13230	1561	962	15	445(3)	0	1
02/24	16	12632	1720	672	116	675(5)	0	17
02/24	15	11584	1348	830	2	291(2)	1	1
02/24	14	11066	965	701	0	295(2)	0	0
02/24	13	10802	1038	824	0	597(5)	0	1
02/24	12	10588	1358	1273	1	82(0)	0	4
02/24	11	10503	1269	1104	0	442(4)	0	2

Meeting needs along the way – Site Limits Study

- Submit probe jobs to each site
- Vary the length of the CPU
- Analyze the success rate as a function of time
- Reverse-engineer the site time policy
- Recently FNAL support started compiling this information for us!



Some Thoughts on the Process

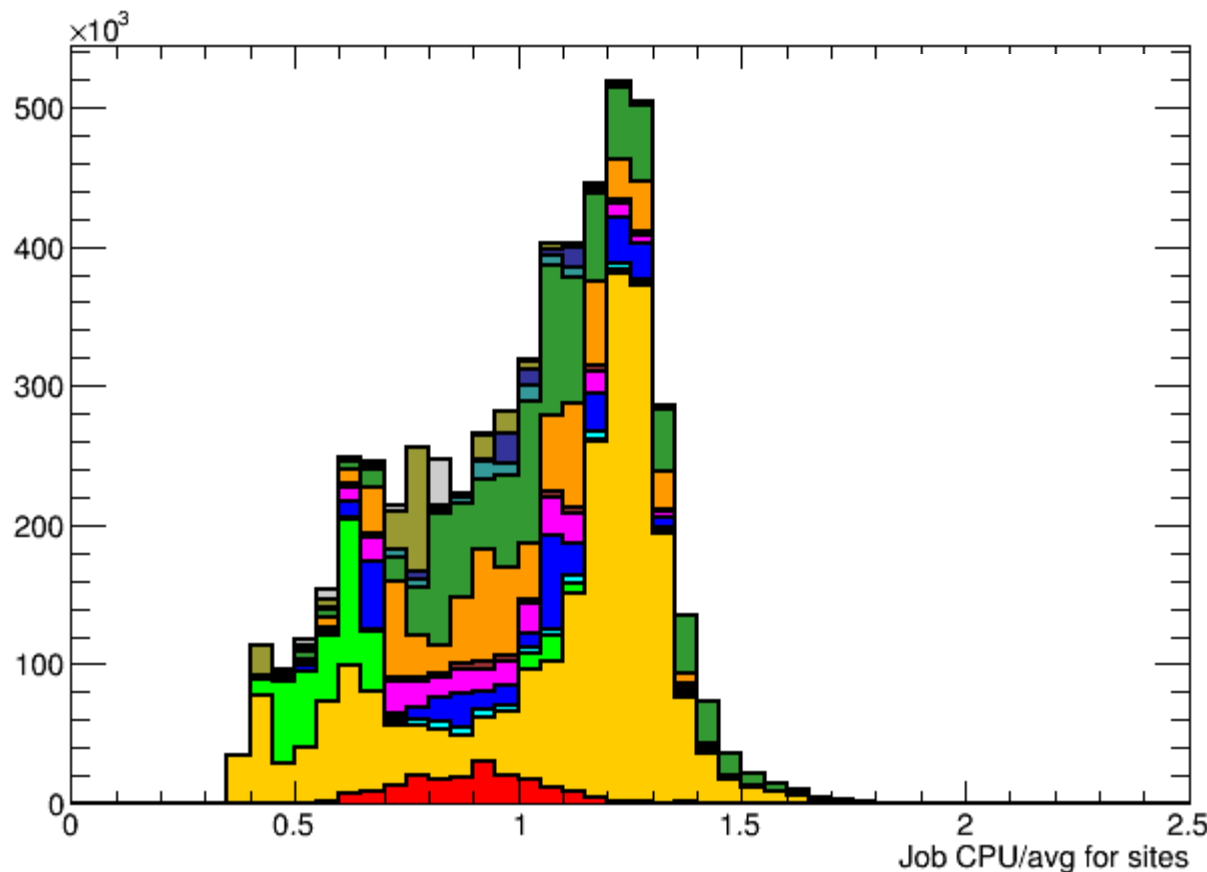
Possible Improvements

After a year of using OSG Opportunistic, this is what I wish for, with no restriction by history, practicality, ongoing work, etc.

- Guarantee 8 h running time before an eviction ←
- Announce capabilities in detail (OS, memory, disk, counts, etc.)
- Announce eviction policy in detail
- Label evictions as evictions, reduce disconnect timeout
- Announce opportunistic availability continuously
- Provide access to example nodes, and node debugging
- Mitigate black holes – detect, throttle/blacklist, and report nodes which show anomalous job lifetime or exit codes?
- Uniformity – software, all sites offer opportunistic to all VO's
- Optimal time command available everywhere
- Lifetimes by normalized CPU (see next)

CPU power is variable

- CPU variation between jobs in a project is order 2%
- CPU/(avg CPU for project) shows factor 3 variation
- Without normalization, makes planning harder..



Thank you for these resources!

- The OSG resources were critical to the Mu2e success for CD3 and will be critical for our future efforts
- Effort was quite manageable
 - Order 1 FTE-month to get going on 10-15 sites
 - Maintenance is order 5-10% FTE
- Seeing continuous improvements...
 - Reliability, monitoring and information
 - Startup getting easier for the next experiment
- See wishlist for my takeaway

The Mu2e experiment would like to thank all those involved from the OSG, the sites, and the lab infrastructure support. You made this success possible!