Large-scale Simulation and Data Processing in the NOvA Experiment
DPF 2017

Adam Moren

DPF 2017

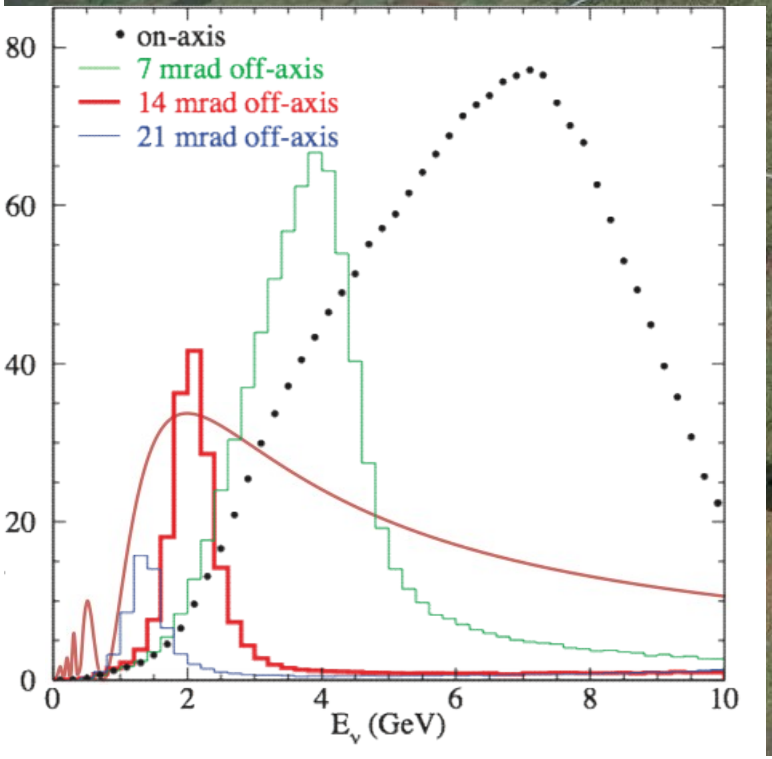
Thursday August 3rd, 2017

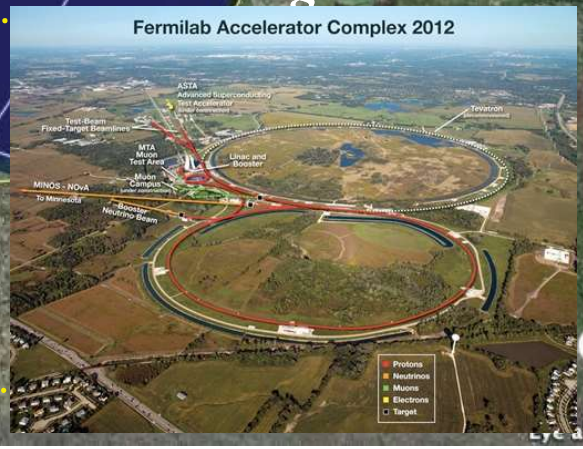
NOvA Experiment

Ash River, MN
810 km from Fermilab

Far detector on the surface

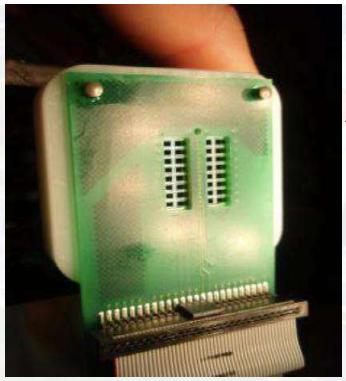
NuMI beam at 700 kW and
Near detector underground

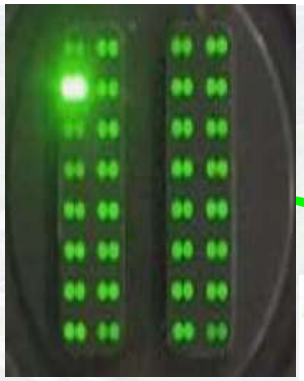




NOνA Detectors:

- Fine-grained, low-*Z*, highly-active tracking calorimeters
- 11 M liters of scintillator
- λ-shifting fiber and APDs





Fiber pairs from 32 cells

~60 m

15.6 m

Far Detector
(*completed Nov. 2014*)
14 kton, 896 layers,
344,000 Channels

Alternating planes
(x view and y view)

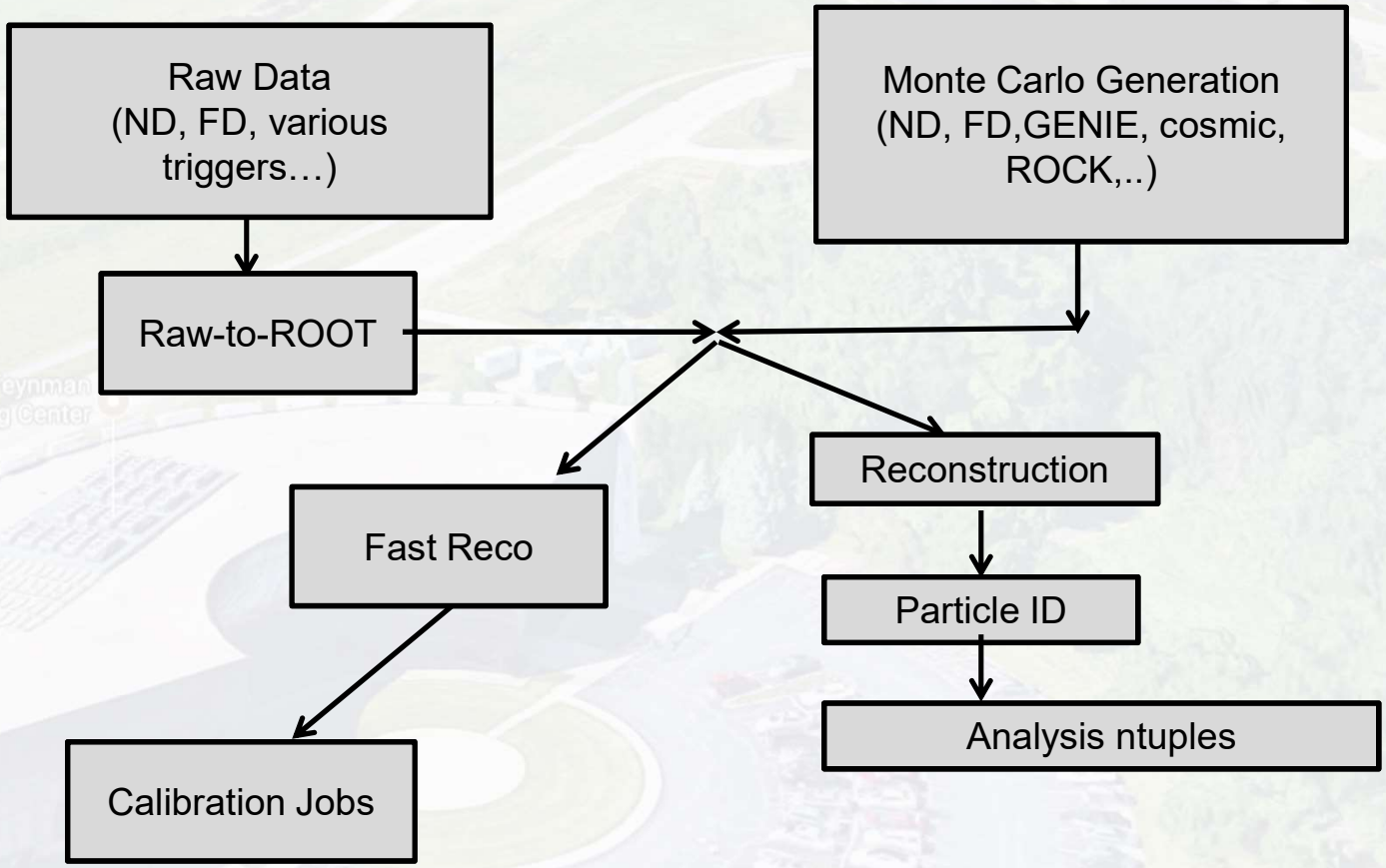
4.1 m

Near Detector
(*completed Aug. 2014*)
0.3 kton, 206 layers,
18,000 Channels

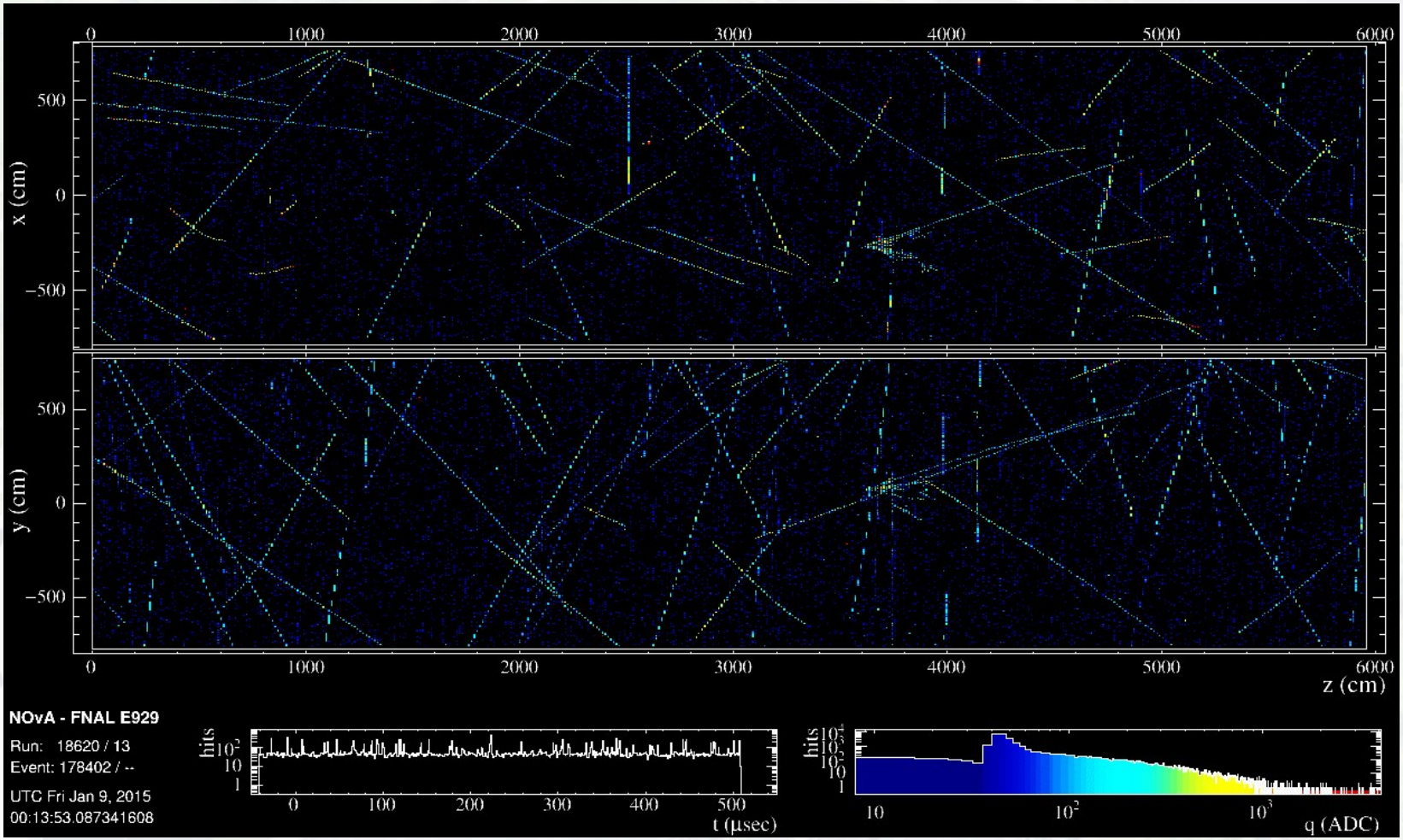
NOvA Data Rates

- The NOvA detectors have produced over 5PB of raw data since 2014
  - Beam time is 0.0005 s every 1.3s. That means for 99.99% of the time we are not recording NuMI data with our detector.
  - 10k RAW files/day (bursting to 12k!)
  - 150kHz cosmic ray rate
  - 5.5 PB of RAW Data Files
- Official NOvA Analysis (2015-2017)
  - 20 Million Files
  - 17 Billion Events
  - 5.6 PB of processed events

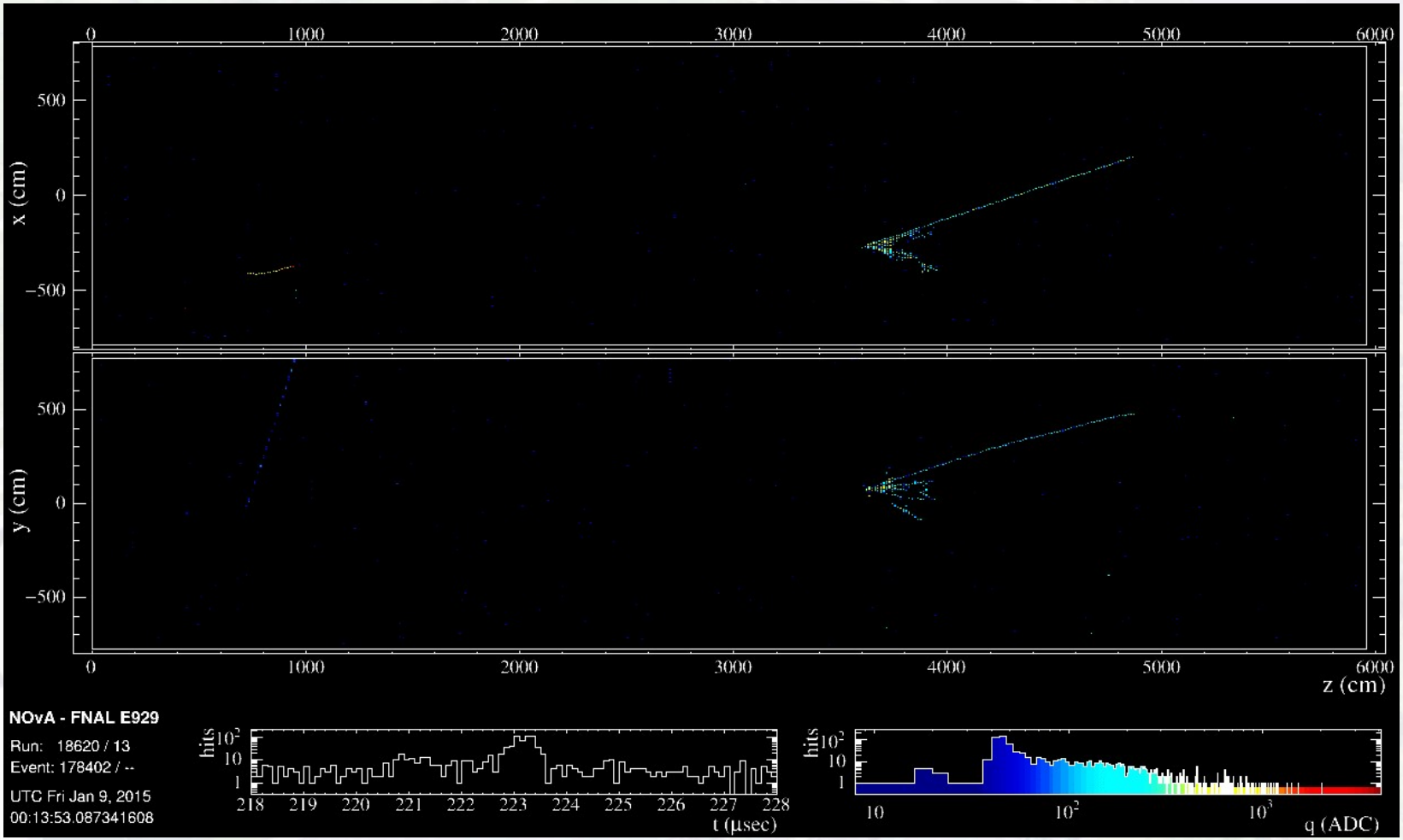
Production Workflow



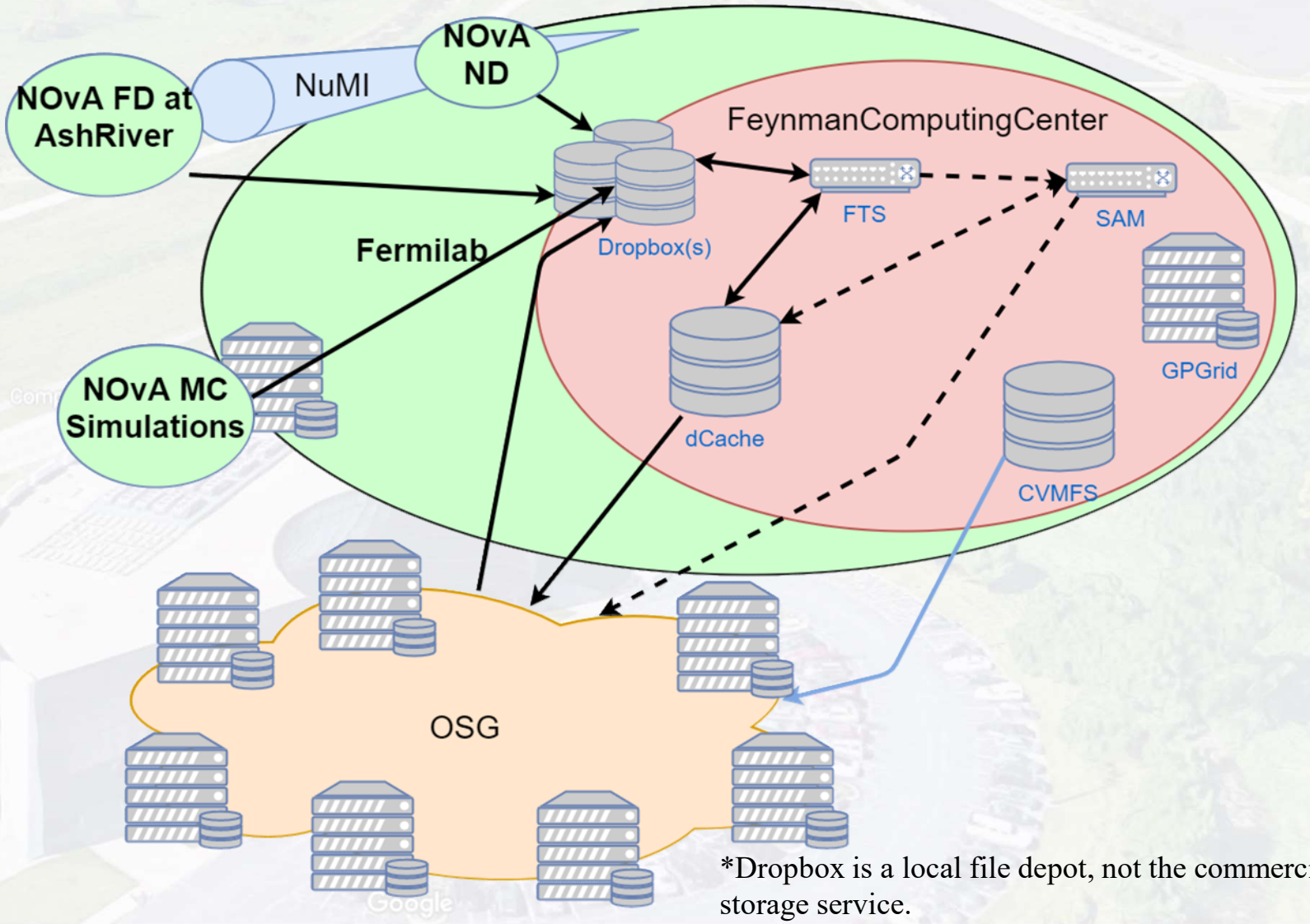
RAW


0
1000
2000
3000
4000
5000
6000
500
0
−500
x (cm)
y (cm)
z (cm)
NOvA - FNAL E929
Run: 18620 / 13
Event: 178402 / --
UTC Fri Jan 9, 2015
00:13:53.087341608
hits
t (μsec)
q (ADC)

reco


0
1000
2000
3000
4000
5000
6000
500
0
−500
x (cm)
y (cm)
z (cm)
NOvA - FNAL E929
Run: 18620 / 13
Event: 178402 / --
UTC Fri Jan 9, 2015
00:13:53.087341608
hits
t (μsec)
q (ADC)

How NOvA Gets it Done



*Dropbox is a local file depot, not the commercial cloud storage service.

FTS/Dropbox

- Dropbox(s)
  - Receives files from grid nodes and detectors
  - Write directories hash to evenly distribute files into tiered directory structure from offline/grid jobs
- FTS
  - Monitors Dropbox folders for new files
  - Declares new files to SAM
    - Metadata Associated with the file
    - Copies files into dCache persistent or tape space

FTS/Dropbox

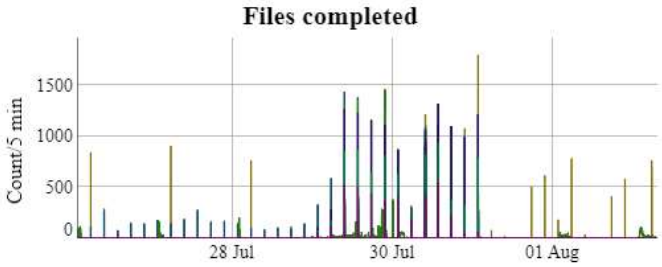
FTS status for nova-fts-novasamgpvm01

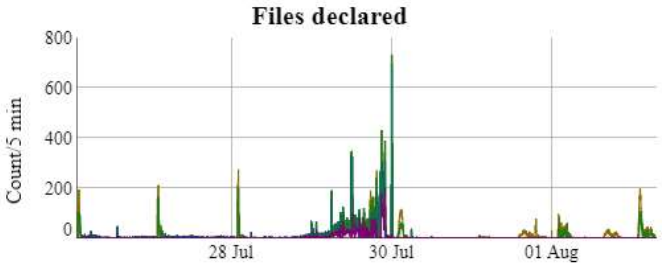
Generated at 2017-08-02 07:33:37 CDT (update)

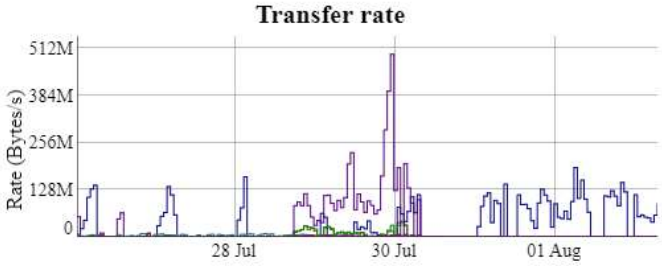
Summary

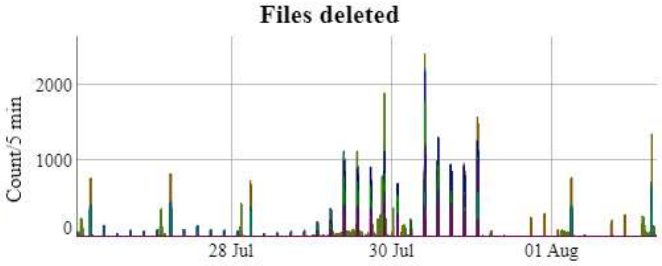
FTS: OK	SAM: OK

Completed files:	42991
Failed transfers:	0
All error files:	3238
Waiting on tape:	98
Other pending files:	0
New files:	2

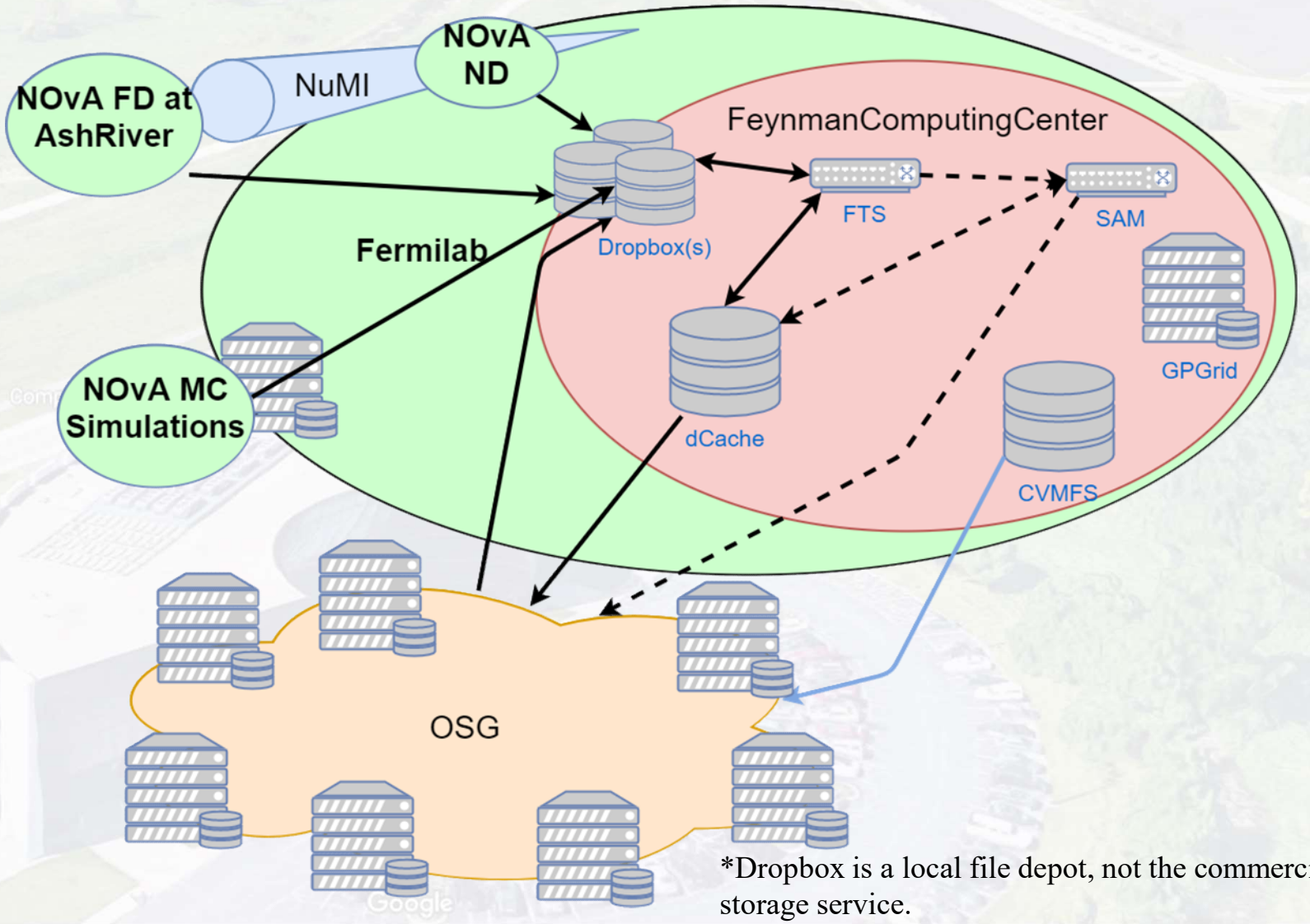








How NOvA Gets it Done



*Dropbox is a local file depot, not the commercial cloud storage service.

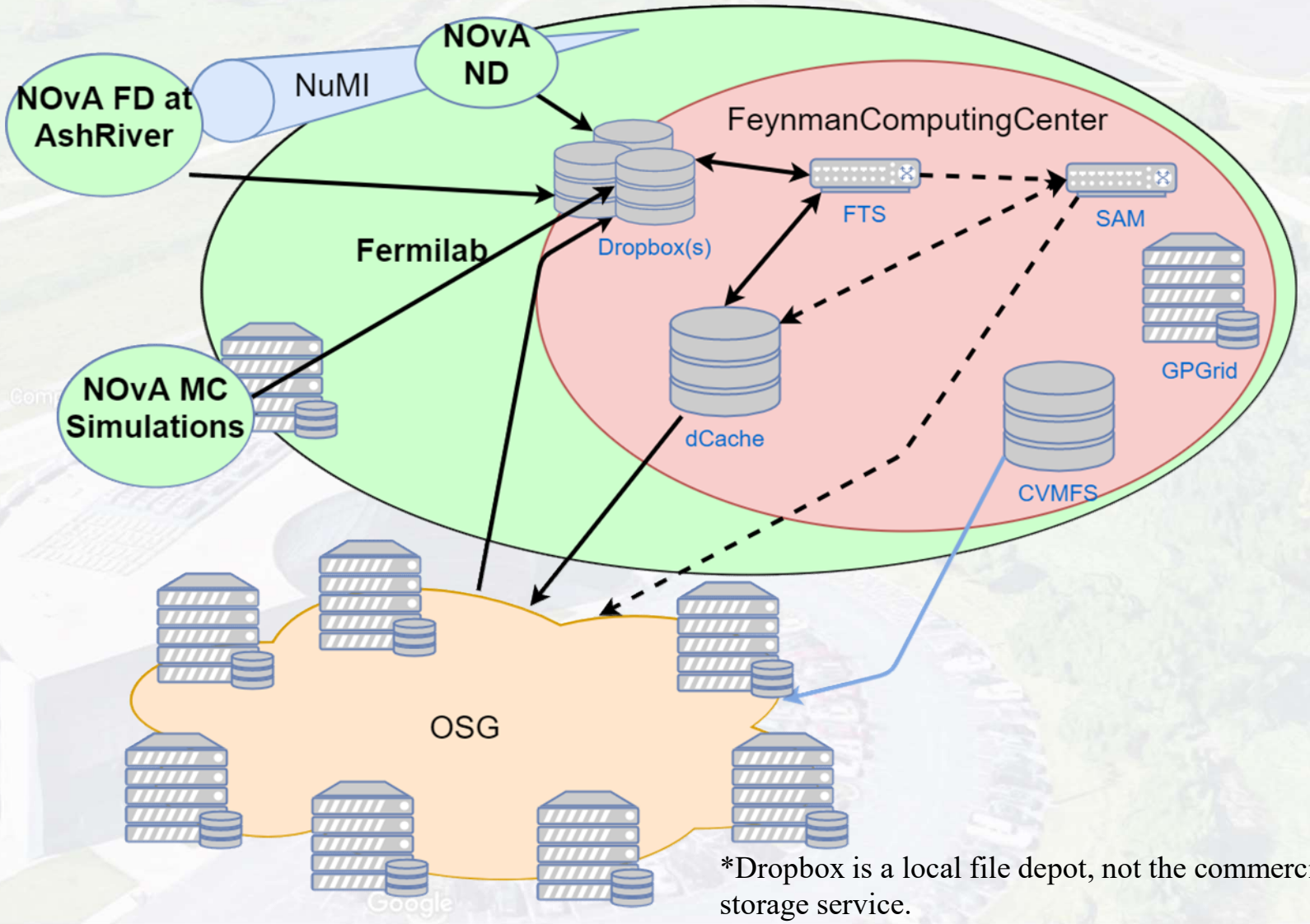
dCache/enstore

- Distributed, multi-petabyte, scalable disk storage system with a single rooted file system providing location independent file access
  - 1-5GBps ingress
  - 10+GBps egress
  - areas
    - Scratch disk (serves out grid jobs)
    - Persistent disk (325TB)
    - Tape-backed (enstore)
  - Files are immutable
  - Pseudo-infinite in size

dCache


Dcache IO R...
Zoom Out
Last 30 d...
vo
nova
PoolGroup
All
Inbound to dcache
6.0 GBs
5.0 GBs
4.0 GBs
3.0 GBs
2.0 GBs
1.0 GBs
0 kBs
7/4
7/7
7/10
7/13
7/16
7/19
7/22
7/25
7/28
7/31
nova
Outbound from dcache
20 GBs
15 GBs
10 GBs
5 GBs
0 kBs
7/4
7/7
7/10
7/13
7/16
7/19
7/22
7/25
7/28
7/31
nova

How NOvA Gets it Done



*Dropbox is a local file depot, not the commercial cloud storage service.

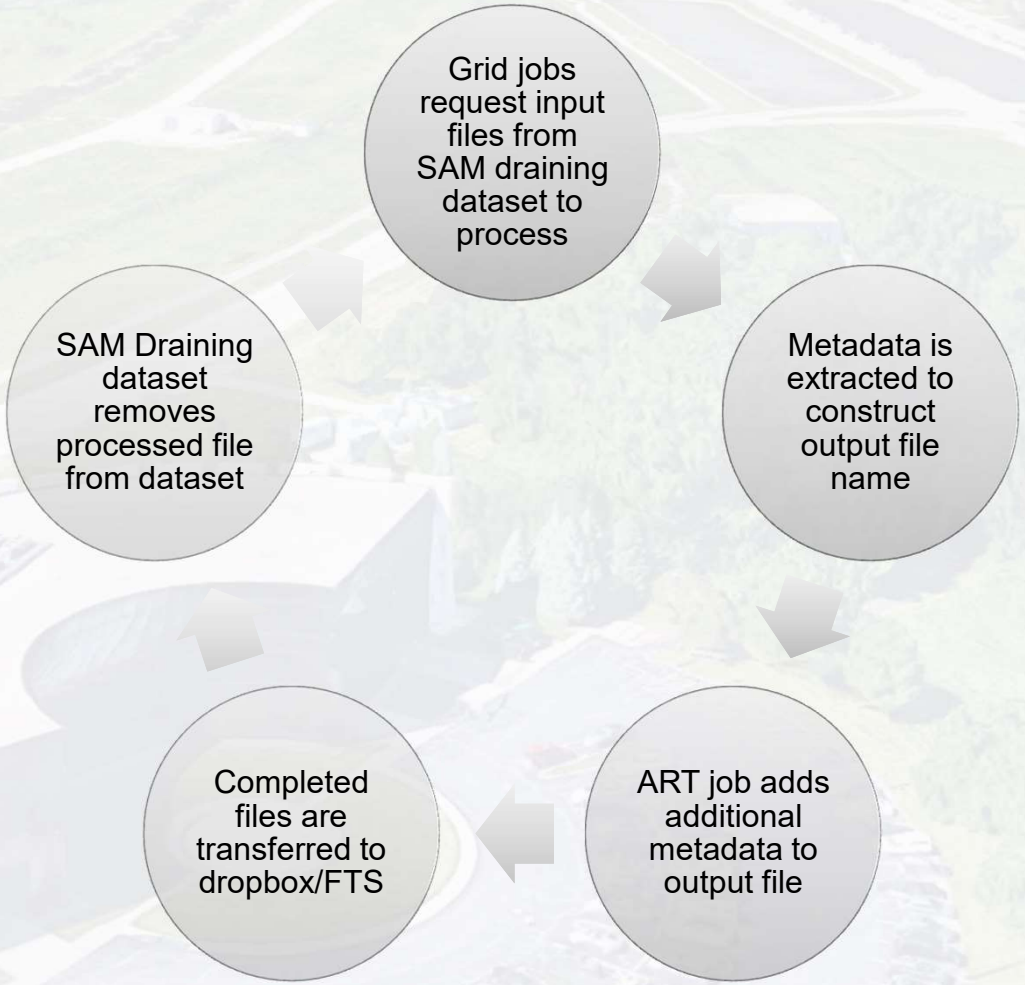
SAM

- Sequential Access via Metadata (SAM)
  - File indexing according to metadata
  - Cataloging the physical location
  - Dataset management
    - Users can create their own datasets based on metadata attributes they define
  - Facilitating data transfer to off-site grids
  - Created here at Fermilab by the Computing Division

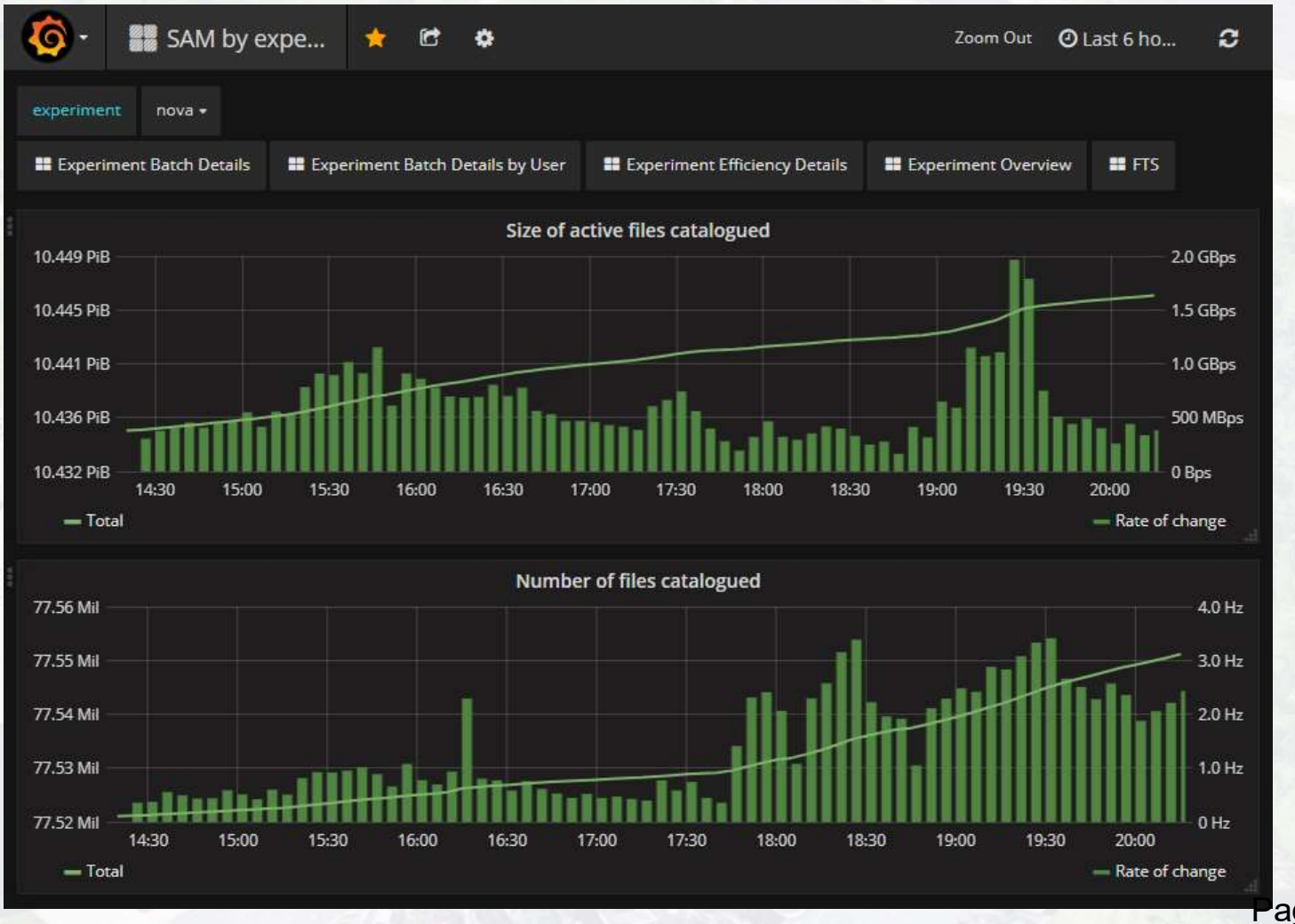
SAM

File Name	fardet_genie_fhc_swap_none_1000_r00014566_s59_c000_v01.72_v1_20170330_145648.sim.daq.root
File Id	310534890
Create Date	2017-04-01T03:03:59+00:00
User	novapro
Update Date	2017-04-03T06:48:47+00:00
Update User	novapro
File Size	505573395
Checksum	
Content Status	good
File Type	importedSimulated
File Format	artroot
Group	nova
Data Tier	artdaq
Application	nova eventmixer r17-03-09-prod3genie.c
Event Count	1000
First Event	1
Last Event	1000
Start Time	2017-04-01T01:39:05+00:00
End Time	2017-04-01T02:51:46+00:00
Data Stream	out1
FCL.Version	v01.72
NOVA.DetectorID	fd
NOVA.HornConfig	mn000z200i
NOVA.HornPolarity	fhc
NOVA.Label	beta
NOVA.Release	R17-03-09-prod3genie.c

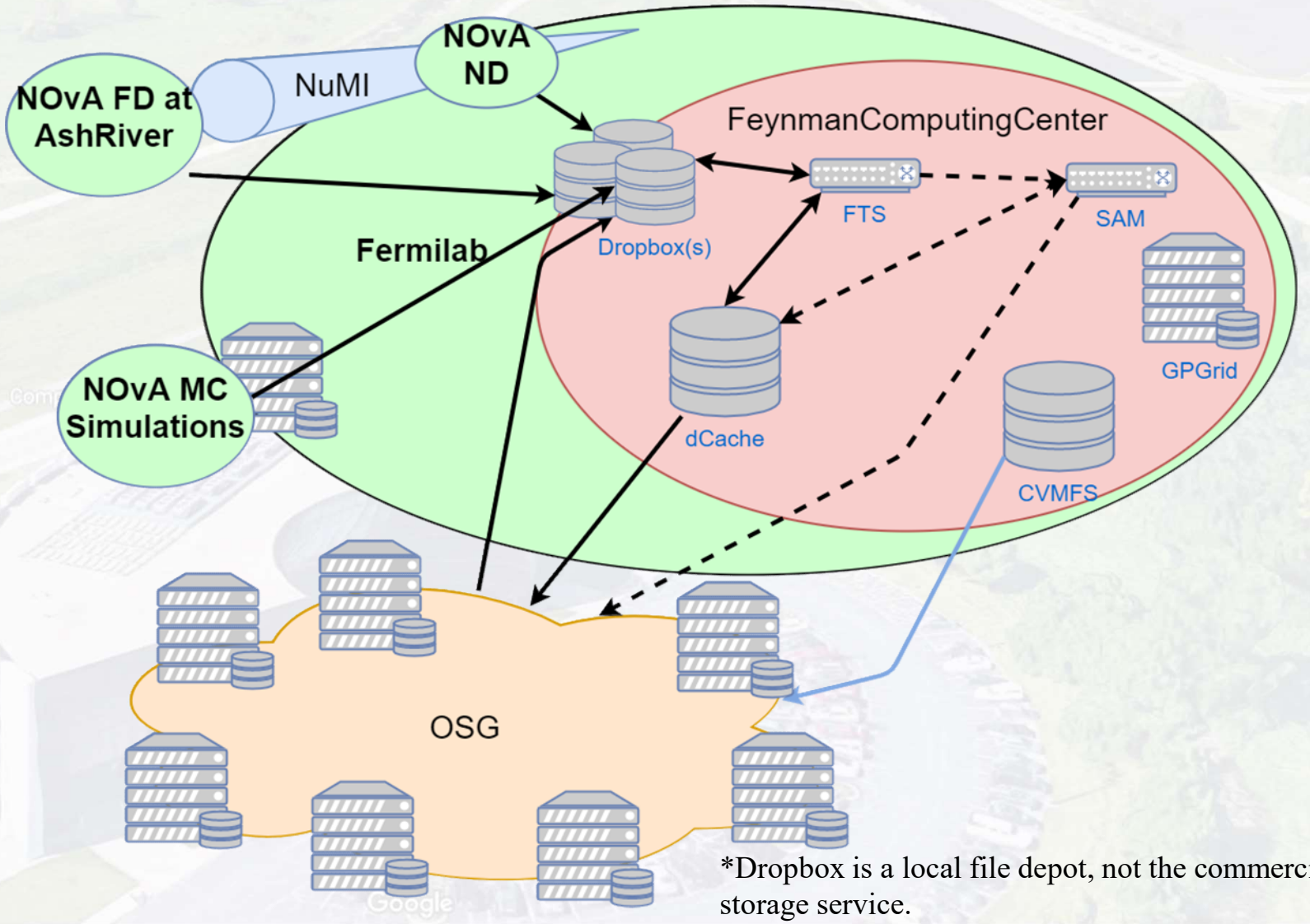
SAM Draining Datasets



SAM


SAM by expe...
Zoom Out
Last 6 ho...
experiment
nova
Experiment Batch Details
Experiment Batch Details by User
Experiment Efficiency Details
Experiment Overview
FTS
Size of active files catalogued
10.449 PiB
10.445 PiB
10.441 PiB
10.436 PiB
10.432 PiB
2.0 GBps
1.5 GBps
1.0 GBps
500 MBps
0 Bps
14:30
15:00
15:30
16:00
16:30
17:00
17:30
18:00
18:30
19:00
19:30
20:00
Total
Rate of change
Number of files catalogued
77.56 Mil
77.55 Mil
77.54 Mil
77.53 Mil
77.52 Mil
4.0 Hz
3.0 Hz
2.0 Hz
1.0 Hz
0 Hz
14:30
15:00
15:30
16:00
16:30
17:00
17:30
18:00
18:30
19:00
19:30
20:00
Total
Rate of change

How NOvA Gets it Done

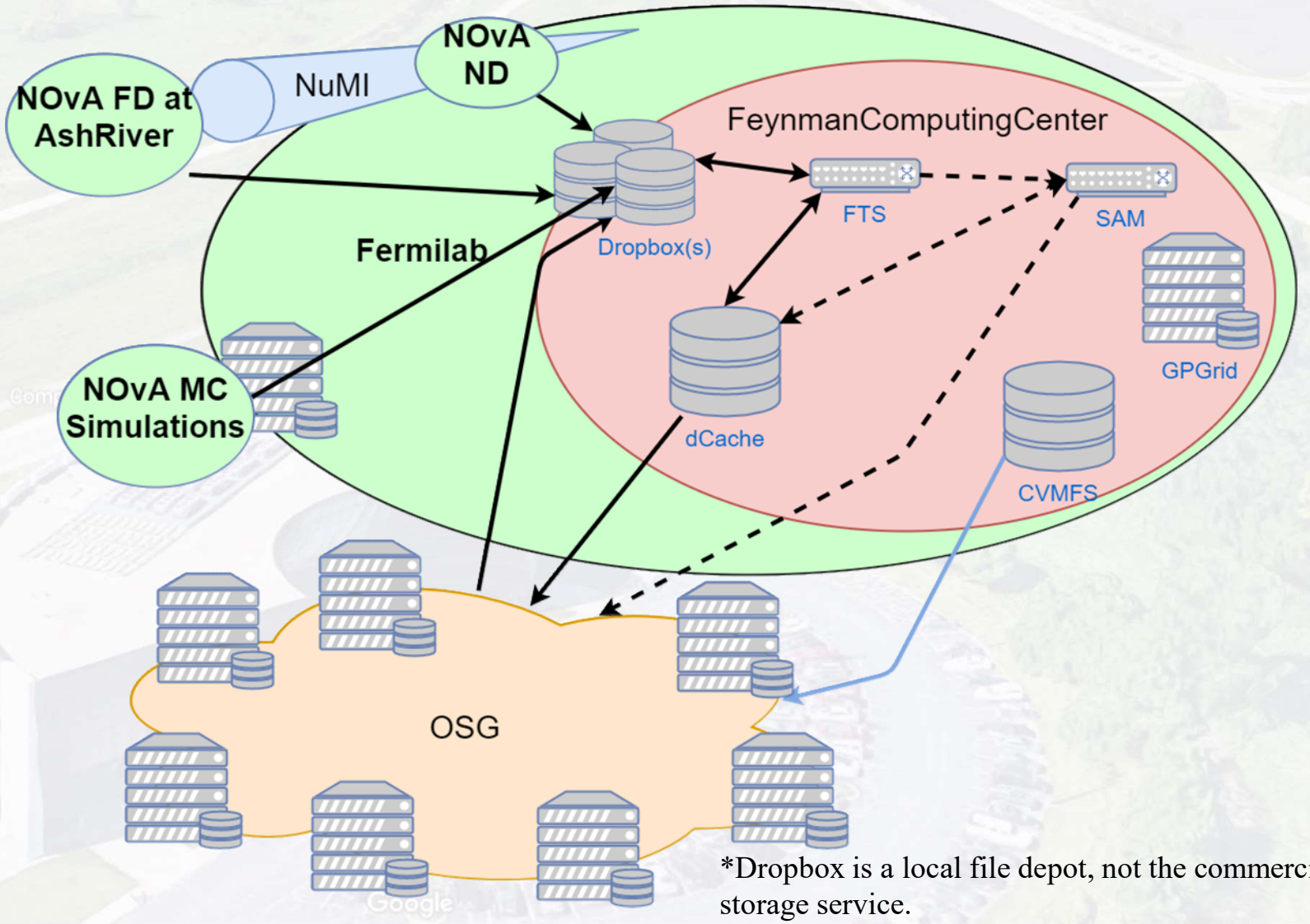


*Dropbox is a local file depot, not the commercial cloud storage service.

CVMFS

- CERN VM File System
  - Software Distribution Service
    - allows NOvA jobs to run on GPGrid and OSG using the same executables everywhere
  - HTTP distribution for easy firewall traversal
  - Mounted as Read Only File System
  - Local caching

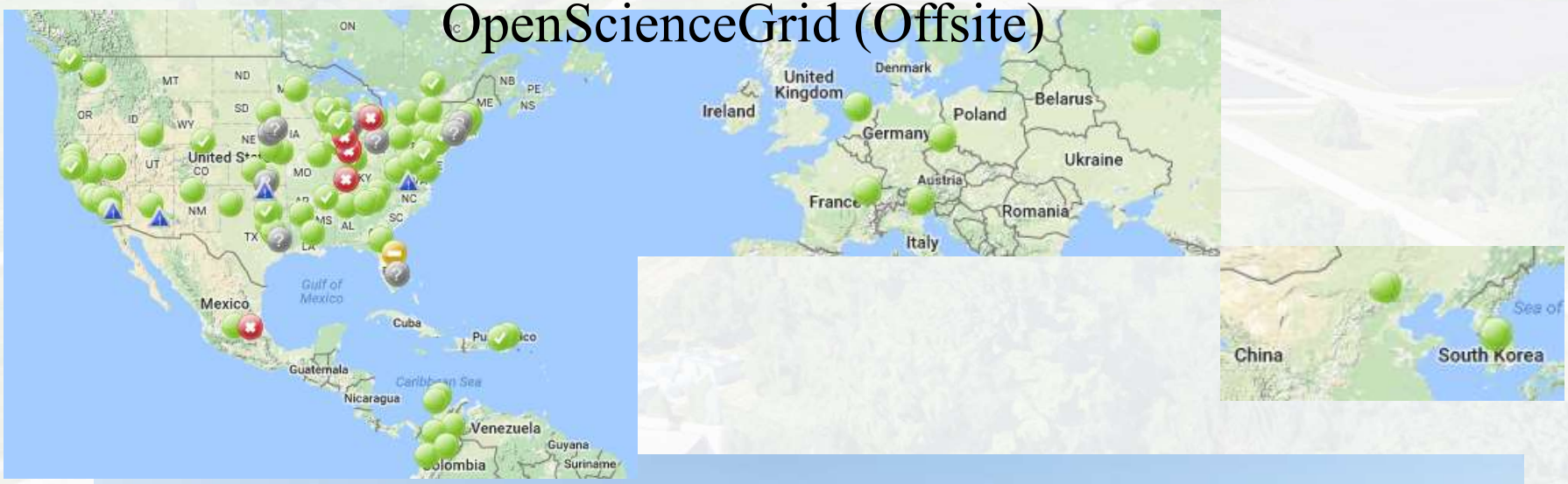
How NOvA Gets it Done



*Dropbox is a local file depot, not the commercial cloud storage service.

Compute

OpenScienceGrid (Offsite)



GPGrid (Onsite) ~17k cores

OSG


Bin size
auto
VO
nova
Job Type
Payload
Filter
Total Wall Hours
22.2 Mil
Total CPU Hours
12.66 Mil
Wall Hours by Usage Model by 1d
600 K
500 K
400 K
300 K
200 K
100 K
0
3/1
4/1
5/1
6/1
7/1
8/1
Unknown Total: 15.974 Mil
OPPORTUNISTIC Total: 3.548 Mil
DEDICATED Total: 2.681 Mil
Wall Hours by Site per 1d
FNAL USCMS Tier1
FZU
SU OrangeGrid
GLOW
Caltech CMS Tier2
JINR_Cloud
Brookhaven ATLAS Tier1
OSC
Nebraska-Omaha
AGLT2
Nebraska-CMS
Nebraska-Lincoln
MWT2 ATLAS UC
STANFORD_HOSTED
MIT CMS
UCSD CMS Tier2
Clemson IT
SMU_HPC
UW-IT
FermiGrid
NWICG_NDCMS
GPGrid
JINR_NOVA_CE1
JINR_NOVA_CE2
N/A

Grids


Job Summary
20000
15000
10000
5000
0
6.0 Bil
5.0 Bil
4.0 Bil
3.0 Bil
2.0 Bil
1.0 Bil
0
10:00
11:00
12:00
13:00
14:00
15:00
16:00
17:00
18:00
19:00
20:00
21:00
Running Current: 8629
Idle Current: 5769
Held Current: 137
Priority Current: 0
Onsite Jobs
8000
6000
4000
2000
0
8500
8000
7500
7000
6500
6000
5500
10:00
11:00
12:00
13:00
14:00
15:00
16:00
17:00
18:00
19:00
20:00
21:00
FNAL_GPGrid
Idle

Management

- Shifts
- ECL
- Wiki
- Slack Channel
- Trello/Scrum board

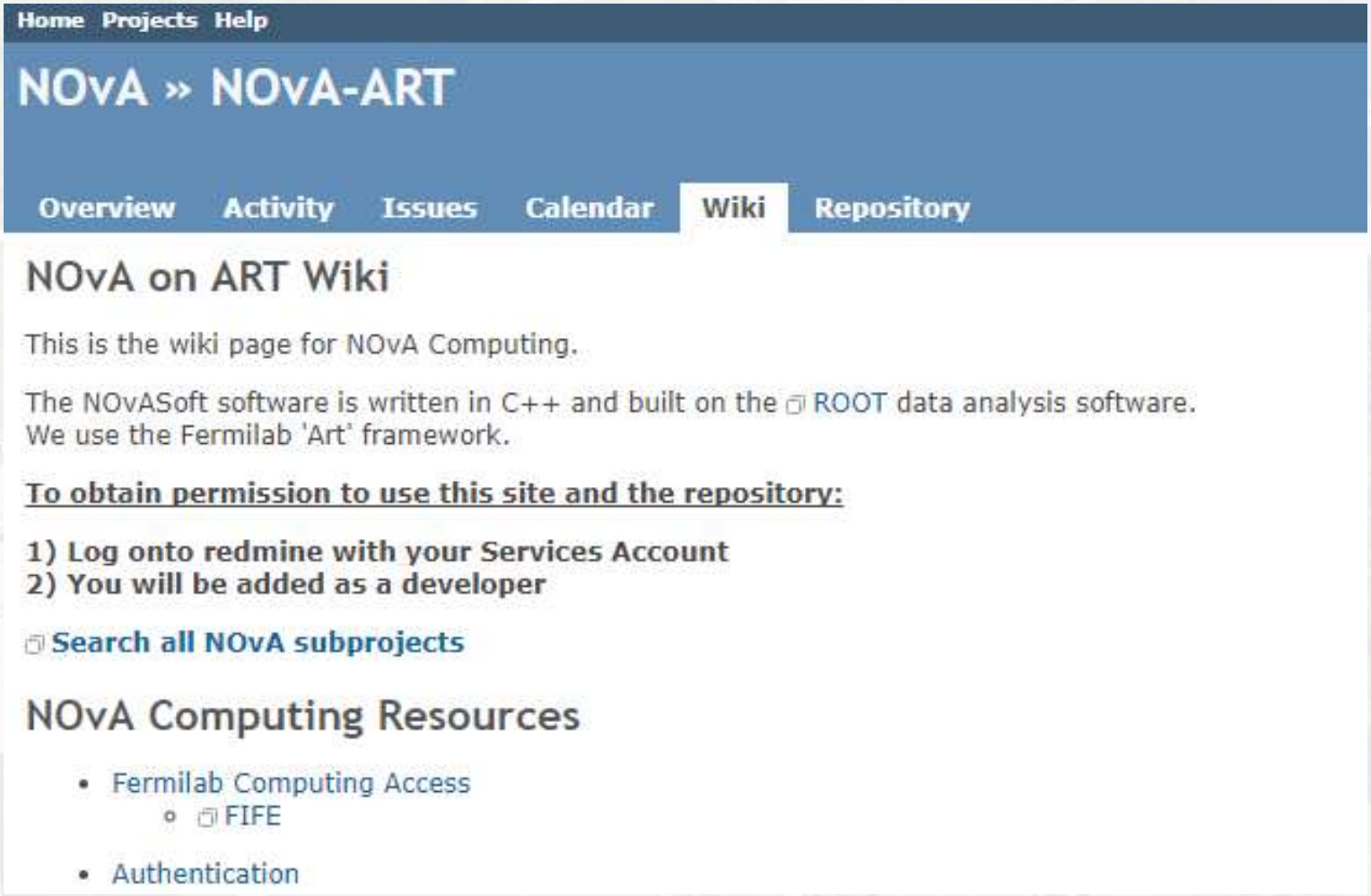
Shifts

- Improved efficiency
  - Shifters will check all the systems during their week.
  - Day-to-day optimization load is centralized.
  - Clearly defines the amount of time individuals are expected to spend on production
  - Conveners distribute time in a predictable way to even out the workload

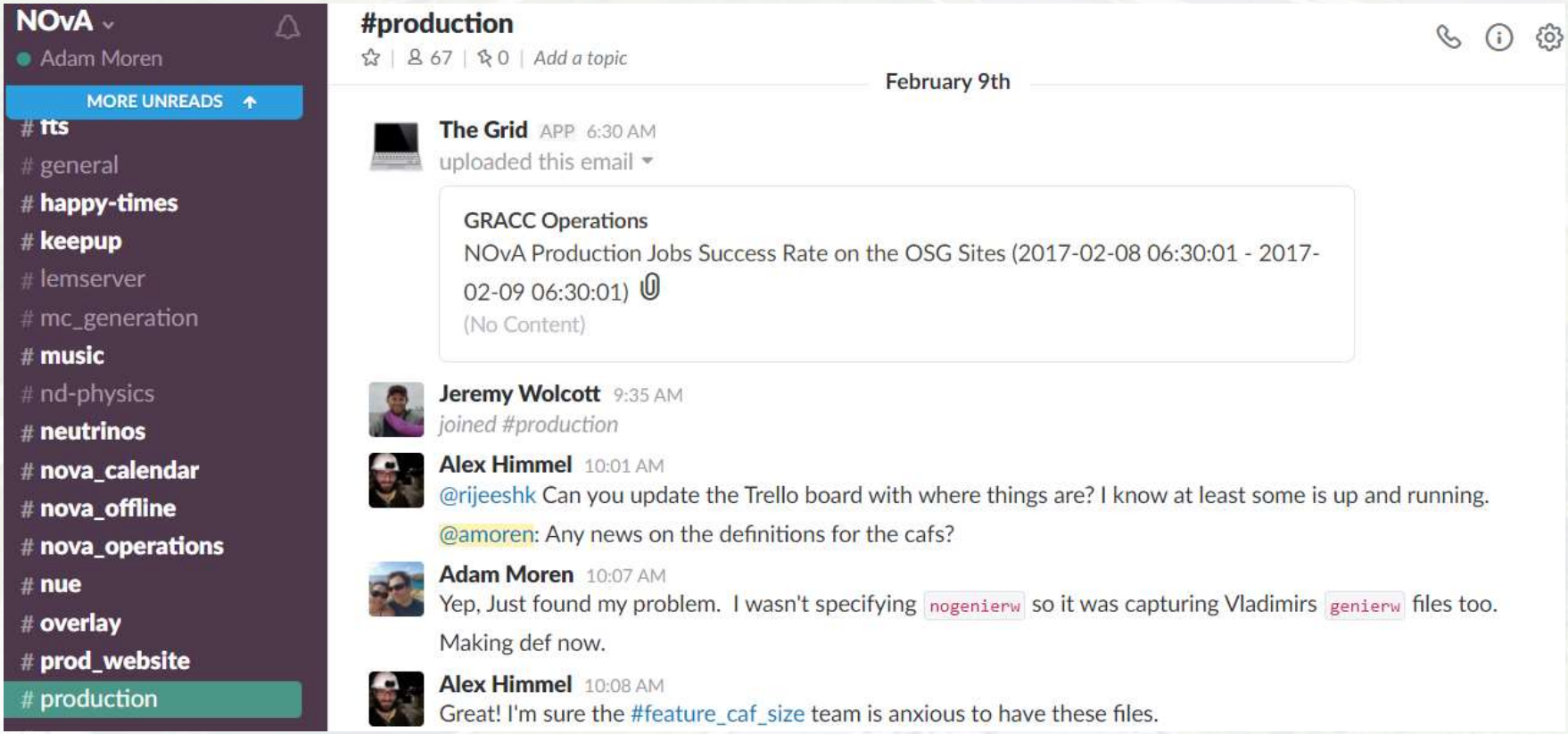
Electronic Collaboration Logbook (ECL)

02/06/2017	**ProductionJobs**	**[amoren]**	Generation	MC	ND	ProductionJobChecklist	*ND MC FHC Rock with R16-11-22-prod3geniepreview.d*
02/06/2017	**FCLs and Definitions**	**[amoren]**	Generation	MC	ND	ProductionJobChecklist	*ND MC FHC Rock with R16-11-22-prod3geniepreview.d*
02/05/2017	**ProductionJobs**	**[amoren]**	MC	ND		ProductionJobChecklist	*ND MC FHC CAF Epoch1-3c (feature_caf_size.b)*
02/04/2017	**ProductionJobs**	**[amoren]**	FD	MC		ProductionJobChecklist	*FD MC FHC Tauswap CAF Epoch3c*
02/04/2017	**ProductionJobs**	**[amoren]**	FD	MC		ProductionJobChecklist	*FD MC FHC Tauswap CAF Epoch3b*
02/04/2017	**ProductionJobs**	**[amoren]**	FD	MC		ProductionJobChecklist	*FD MC FHC Tauswap CAF Period2*
02/04/2017	**ProductionJobs**	**[amoren]**	FD	MC		ProductionJobChecklist	*FD MC FHC Tauswap CAF Period1*
02/03/2017	**ProductionJobs**	**[amoren]**	(de)CAF Respin	FD		ProductionJobChecklist	*FD MC FHC Fluxswap Cosmics Overlay*
02/03/2017	**ProductionJobs**	**[amoren]**	(de)CAF Respin	FD		ProductionJobChecklist	*FD MC FHC NonSwap Cosmic Overlay*
02/03/2017	**ProductionJobs**	**[amoren]**	(de)CAF Respin	FD		ProductionJobChecklist	*FD MC FHC NonSwap Cosmic Overlay*
02/02/2017	**ProductionJobs**	**[amoren]**	(de)CAF Respin	FD		ProductionJobChecklist	*FD MD FHC Fluxswap New CAF*
02/02/2017	**ProductionJobs**	**[amoren]**	(de)CAF Respin	FD		ProductionJobChecklist	*FD MC FHC*
01/17/2017	**ProductionJobs**	**[amoren]**	Calib	FD		ProductionJobChecklist	*FD epoch5a cosmics with R16-11-02-prod3calibpreview.a*
01/17/2017	**ProductionJobs**	**[amoren]**	Calib	ND		ProductionJobChecklist	*ND epoch5a ddactivity with R16-11-02-prod3calibpreview.a*

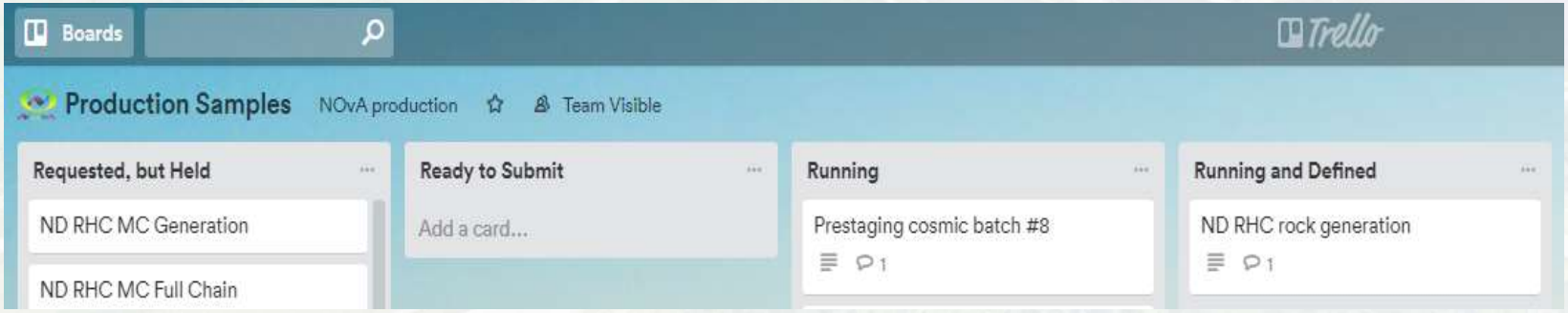
Wiki's

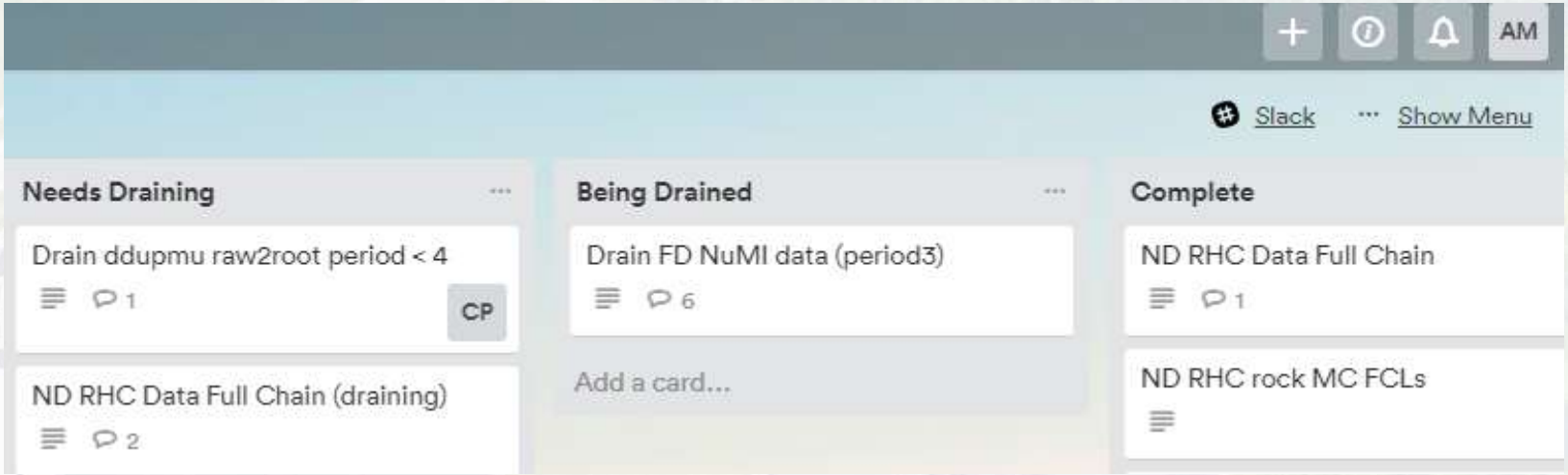

Home Projects Help
NOvA » NOvA-ART
Overview Activity Issues Calendar Wiki Repository
NOvA on ART Wiki
This is the wiki page for NOvA Computing.
The NOvASoft software is written in C++ and built on the ROOT data analysis software.
We use the Fermilab 'Art' framework.
To obtain permission to use this site and the repository:
1) Log onto redmine with your Services Account
2) You will be added as a developer
Search all NOvA subprojects
NOvA Computing Resources
Fermilab Computing Access
FIFE
Authentication

Slack


NOvA
Adam Moren
MORE UNREADS
tts
general
happy-times
keepup
lemserver
mc_generation
music
nd-physics
neutrinos
nova_calendar
nova_offline
nova_operations
nue
overlay
prod_website
production
#production
67 | 0 | Add a topic
February 9th
The Grid APP 6:30 AM
uploaded this email
GRACC Operations
NOvA Production Jobs Success Rate on the OSG Sites (2017-02-08 06:30:01 - 2017-02-09 06:30:01)
(No Content)
Jeremy Wolcott 9:35 AM
joined #production
Alex Himmel 10:01 AM
@rijeeshk Can you update the Trello board with where things are? I know at least some is up and running.
@amoren: Any news on the definitions for the cafs?
Adam Moren 10:07 AM
Yep, Just found my problem. I wasn't specifying nogenierw so it was capturing Vladimirs genierw files too.
Making def now.
Alex Himmel 10:08 AM
Great! I'm sure the #feature_caf_size team is anxious to have these files.

Trello


Boards
Trello
Production Samples
NOvA production
Team Visible
Requested, but Held
ND RHC MC Generation
ND RHC MC Full Chain
Ready to Submit
Add a card...
Running
Prestaging cosmic batch #8
1
Running and Defined
ND RHC rock generation
1


AM
Slack
Show Menu
Needs Draining
Drain ddupmu raw2root period < 4
1
CP
ND RHC Data Full Chain (draining)
2
Being Drained
Drain FD NuMI data (period3)
6
Add a card...
Complete
ND RHC Data Full Chain
1
ND RHC rock MC FCLs

Conclusions

- NOvA's detectors produce large pipelines of data
- NOvA negotiates that datastream using numerous large-infrastructure tools supplied & supported by Fermilab and the wider HEP community
- Detailed resource and personnel management strategies have proven essential in smooth end-to-end simulation & reconstruction processing

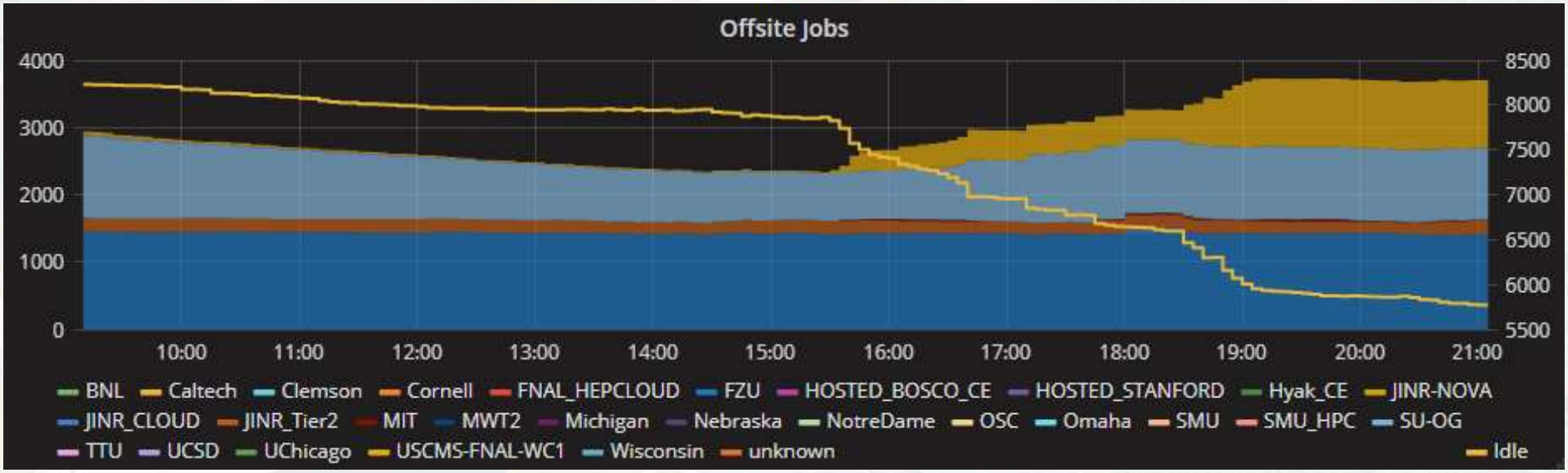
Acknowledgements

- Data Handling with SAM and ART at the NOvA Experiment, Adam A. et al., CHEP 2015
- INTRODUCTION TO NOVA AND DATA TAKING, Louise Suter Argonne National Lab Workshop March Collaboration meeting, 2016
- Data handling for NOvA analyzers - A tutorial on the dCache system and “SAM for users” tools Craig Group & Pengfei Ding, 2015
- Recent Evolution of the Offline Computing Model of the NOvA Experiment Talk, Craig Group & Alec Habig, CHEP 2015
- NOvA Data Acquisition System, Xinchun Tian, Department of Physics and Astronomy Brown, 2011
- CVMFS - http://cernvm.cern.ch/portal/filesystem
- NOvA Production Group!!

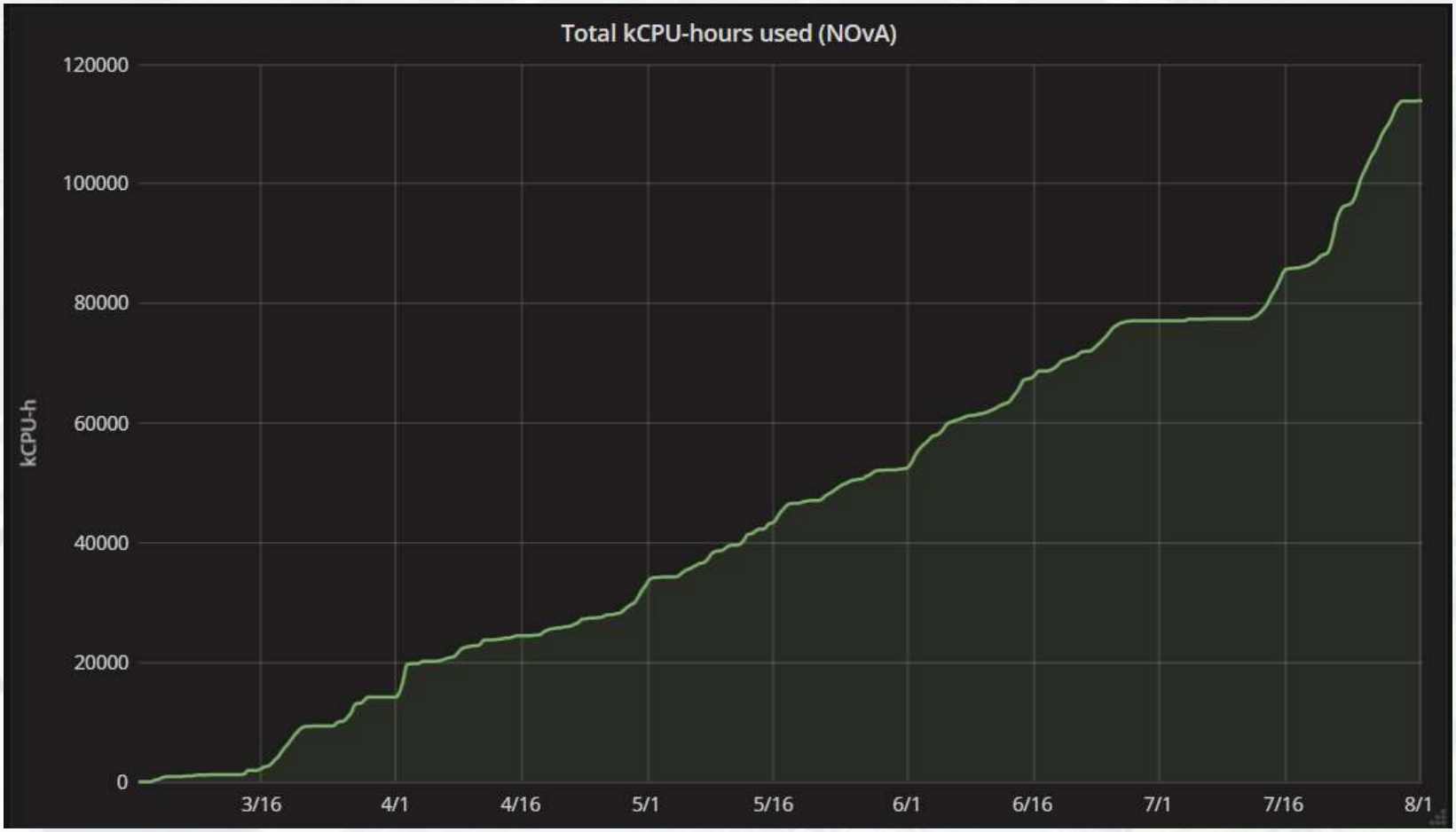
The speaker is supported by
NSF RUI grant #1607381

Backup Slides

OSG/Offsite


Offsite Jobs
4000
3000
2000
1000
0
8500
8000
7500
7000
6500
6000
5500
10:00
11:00
12:00
13:00
14:00
15:00
16:00
17:00
18:00
19:00
20:00
21:00
BNL
Caltech
Clemson
Cornell
FNAL_HEPCLOUD
FZU
HOSTED_BOSCO_CE
HOSTED_STANFORD
Hyak_CE
JINR-NOVA
JINR_CLOUD
JINR_Tier2
MIT
MWT2
Michigan
Nebraska
NotreDame
OSC
Omaha
SMU
SMU_HPC
SU-OG
TTU
UCSD
UChicago
USCMS-FNAL-WC1
Wisconsin
unknown
Idle

Grids


Total kCPU-hours used (NOvA)
kCPU-h
120000
100000
80000
60000
40000
20000
0
3/16
4/1
4/16
5/1
5/16
6/1
6/16
7/1
7/16
8/1