

FIFE Workshop

Welcome

Thank you for coming

SCPMT was valuable

- Report is [here](#).
- Several of the recommendations/comments are out of SCD control - ie up to the projects/collaborations.
- Working through your use cases on storage. Details to follow in this presentation.

Do you want to be here?



The long term roadmap



- Directorate level theme: Fermilab the international home for DUNE. We will feel it on many fronts from updating the user onboarding process to authentication/authorization across scientific computing services.
- OCIO level theme (strong DOE push for this): common solutions. FIFE is a poster child of success in that regard, but looking for a larger scale - IF and CMS. Potential areas are:
 - Frameworks
 - Data management
 - Workflow management



Communication channels

SCD Initiated	Experiment Initiated
SCPMT	SPPM
FIFE Workshop	SNOW Tickets
Redmine Roadmaps	AEM
E-Mail updates	

What is the right set (ie, most effective) communication channels for you?
How can both sides improve the listening aspects. Is SCD responsive to your priorities? Are you responsive to SCDs?

Homework

What is the right set (ie, most effective) communication channels for you? How can both sides improve the listening aspects. Is SCD responsive to your priorities? Are you responsive to SCDs?



FIFE Roadmap Workshop

FIFE Roadmap Workshop

June 14, 2018



FABRIC FOR FRONTIER EXPERIMENTS

FIFE Roadmap Workshop

- The goal of the roadmap discussion is to both inform experiments and gather feedback about strategic infrastructure changes and computing service modifications.
- SCD has budgetary, security, and effort constraints that must be met.
- Must also keep track of industry best practices and available tools.
- Constant large meetings between experiments and service providers are not productive to developing strategy.
- Feedback concerning specific enhancements, changes, and bugs are important on day-to-day basis.



Does it all seem confusing and directionless?

Hopefully, at the end of the day at least a few of the signs will be more understood

Expectations from the workshop

- Experiments and providers communicate about current services in a productive conversation
- Service providers present plans for the roadmap of their services and changes for the next 6-12 months
- Experiments help to guide the process of developing requirements for changes and new services (should not develop requirements during this workshop)
- Everyone has a better understanding of the direction of SCD services and experiment computing models
- Please feel free to contact FIFE to schedule working meeting with service providers - especially if coordination is required

Agenda for the Workshop

- Communications with SCD
 - new SNOW interface
- HEPCloud
- Container Building
- GPU update
- jobsub
- CVMFS
- dCache
- DM and Rucio
- Security
- FERRY
- Landscape
- POMS
- CI Validation
- CSV API
- art framework
- Online support



SCD Communication

Mike Kirby

FIFE Roadmap Workshop

June 14, 2018



New Service Now Interface

The screenshot shows the new Service Now interface for Fermilab. The header is blue and contains the Fermilab logo, a search bar with the text "Search for answers or items", and a "SEARCH" button. Below the header is a navigation grid of six service tiles:

- Get Help** (Blue tile with a person icon)
- Request Something** (Grey tile with a shopping cart icon)
- Browse Services** (Blue tile with a grid icon)
- My Tickets** (Grey tile with a document icon)
- Service Status** (Blue tile with a warning triangle icon)
- Get Information** (Grey tile with an information icon)

A right-hand sidebar displays user information for Michael Kirby, announcements, and a list of tickets:

- Announcements**
 - Common web portal questions and workarounds - updated
 - OSG certificate authority shutting down after May 31
 - Web portal user instructions
- My Tickets**
 - SQLite error returned on jobsub logs
 - "Geant4" and "elegant" on APC cluster
 - procurement status of new EVB and RAID array
 - E-974 (MicroBooNE) Account Request
- My Approvals**
 - You have no pending approvals
- Service Status**

Suggestions for Service Now

The impression we have from experimenters is that users are not very familiar with the SNOW interface.

- Do we need tutorials?
- How do we make understanding the usefulness of SNOW easier?

Requested improvements to Service Now:

- the ability to search for and see tickets you are not already a watcher on
- the ability to add people to the watch list of a ticket
- have SNOW interact better with email, specifically
 - Shortening subject lines
 - Keeping the subject consistent across all ticket stages
 - Not sending duplicate emails
 - Not frequently re-duplicating the ticket history
- Continue to use SNOW to open tickets for feature requests and downtimes



HEP Cloud

Parag Mhashilkar

FIFE Roadmap Workshop

June 14, 2018



HEP Cloud Roadmap

- HEPCloud facility will be in production by end of 2018.
- Why should you care?
 - Nova used HEP Cloud to run 1M cores using HEP Cloud on NERSC in May
 - Do you have an urgent need for computing power? Your experiment could be next one!
 - HEP Cloud extends the possibility of using several resource providers Google, Amazon, NERSC, OSG, Local cluster when submitted to HEP Cloud - greatly expanded resources
 - No free lunch - requires changes on how job submission happens and we want to make this as friendly as possible.



HEP Cloud Roadmap

- What should you change to use new features?
 - Jobs should provide additional information about itself
 - Jobs should focus on resource needs and not on specific external sites or resource types - users *will* have less control over site location
- What is next?
 - There will be an experiment onboarding/offboarding process to run on Cloud and HPC centers
 - Details are being worked out and will be available by end of the summer 2018

HEP Cloud Scope for 2018

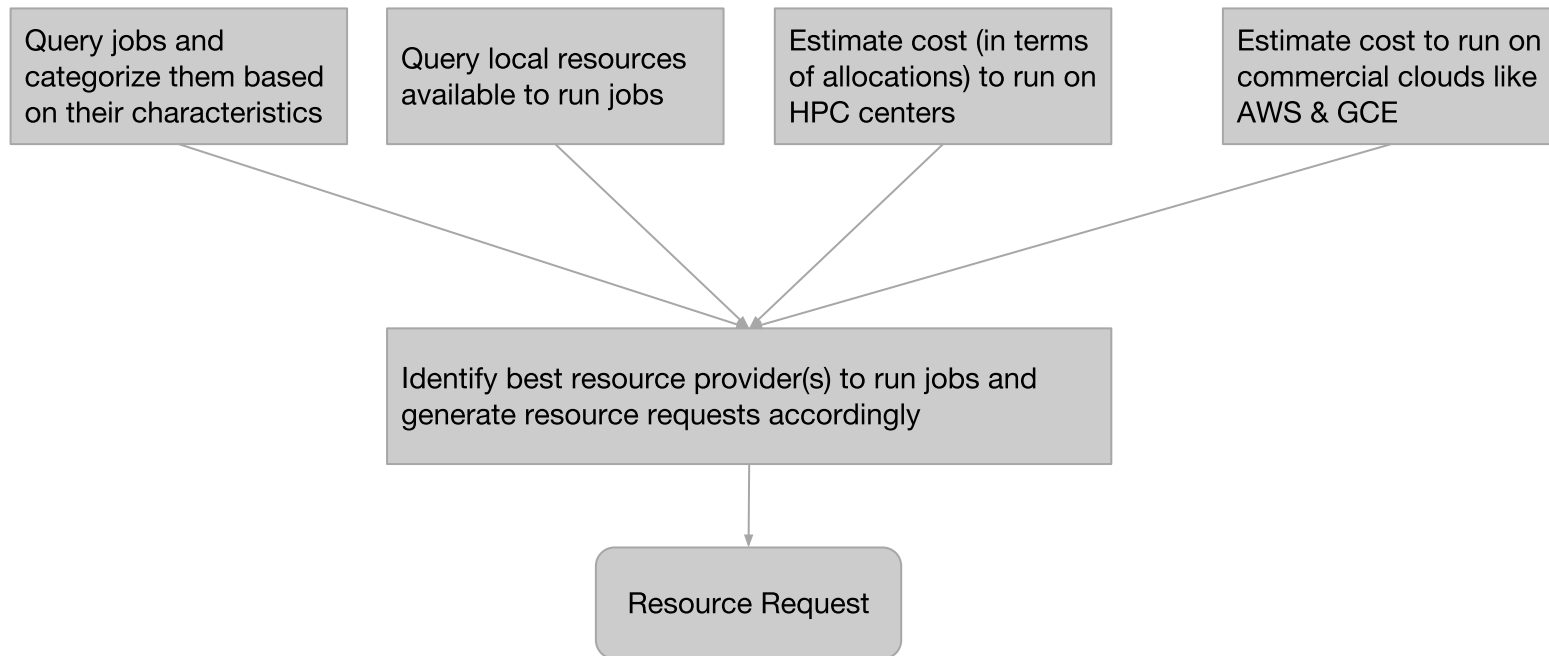
HEP Cloud Facility to date:

- Extend the Fermi Grid facility to integrate external resources
 - Opportunistic, HPC, Commercial Cloud Resources
- Execute scientific workflows identified in conjunction with the experiments' leadership
- Manage the access interfaces to the facility
- Automate workflow placement decisions based on multiple metrics
- Monitor operations of the facility
- Financially account and plan for use of the facility
- Manage contractual and similar agreements related to the facility
 - Example: Do you have an allocation at NERSC or will you run using FIFE allocation at NERSC?

HEP Cloud Path to Production

- **HEPCloud Facility**
 - Support running jobs at several resource providers: Google, Amazon, NERSC, OSG, Local cluster
 - Develop Facility Governance Model & Economic Model
 - Document: Facility limits and assumptions, Experiment onboarding/offboarding
- **Decision Engine**
 - Designing and implementing resource provisioning algorithms
- **Facility Monitoring**
 - Full suite of monitoring of each resource provider and available resources for accounting, auditing, budgeting, etc.
- **Comply with Fermi Security requirements**
 - Risk Analysis, Security Plan, Security Controls, etc
- **Obtain Authority to Operate (ATO) from the DOE**

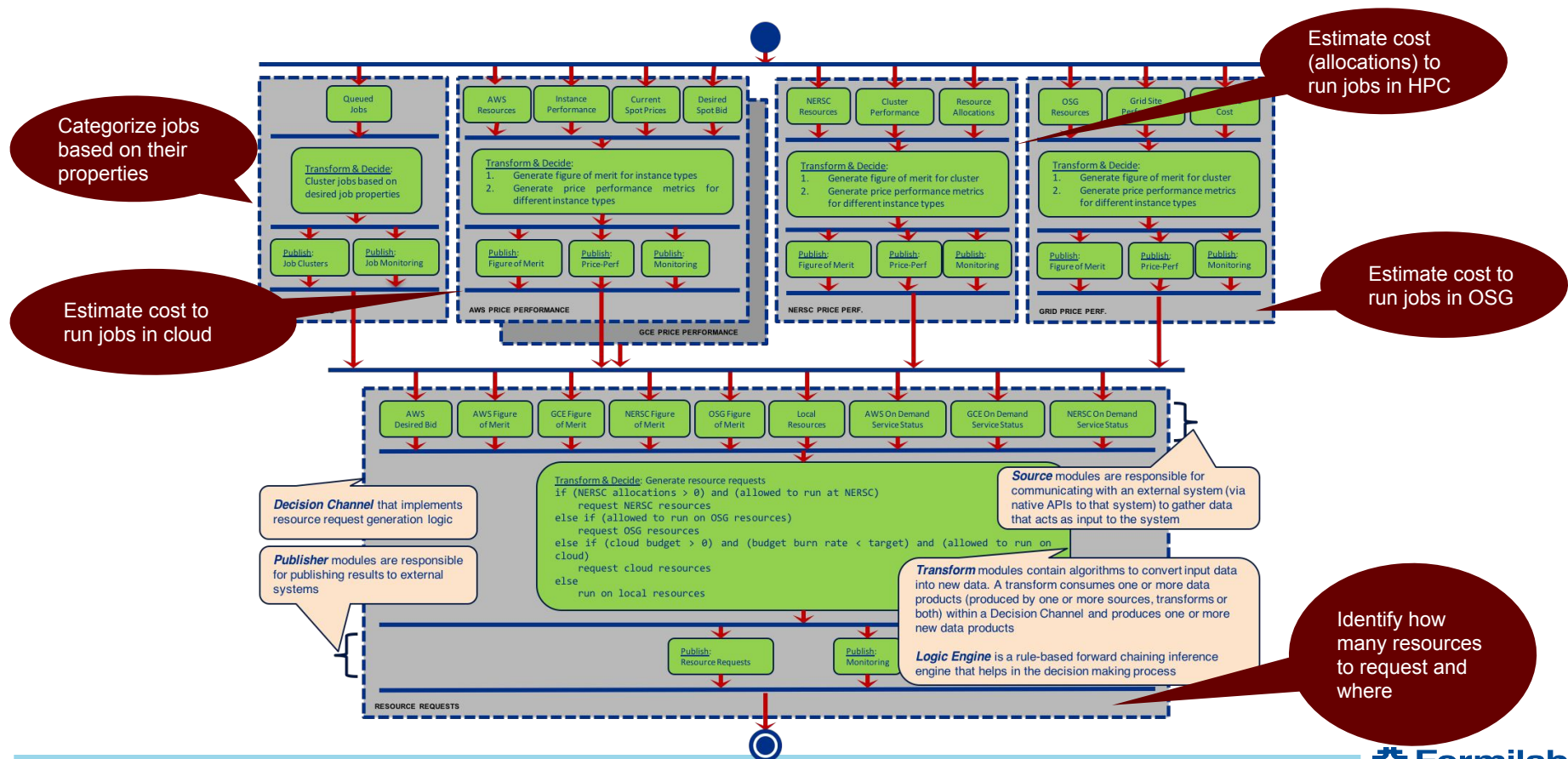
HEP Cloud Decision Engine Provisioning Workflow



HEP Cloud Production FIFE Impact

- Will still use Jobsub to submit jobs
 - Expect some new job attributes to be introduced, details to follow by end of July
 - IO rates, memory, timing constraints, special needs like GPUs, etc
- Job submission should ***focus on resource needs*** and ***not on specific external sites or resource types***
 - Provides flexibility on acquiring resources to run jobs
- Experiments will have an on-boarding process to use HPC and Cloud
 - Understanding the classification of jobs suitable for resource types:
 - New job attributes (if not already present)
 - Cloud requires accurate input and output data size as it directly impacts cost of running in Cloud
 - Job idling while waiting for large amount of input data results in CPU wastage and low job efficiency
 - Shipping data out of cloud is expensive
 - Creating experiment specific containers for HPC resources
 - Only PRODUCTION jobs can be forwarded to Cloud and HPC resources

Decision Engine Provisioning Workflow Details





Container Building Interim Roadmap

Tony Tiradani, Ed Simmonds

FIFE Roadmap Workshop

June 14, 2018

Containers Are The Future (and the Present)



- We recognize that containers are becoming more important.
- A solution is needed for various operations:
 - Developing containers
 - Testing containers
 - Storing containers
 - Distributing containers
- We are working on longer term solutions, but requests are being made right now for places to develop containers.
- Docker is inherently insecure, so this must be provided in a controlled manner.

Container Build Service (pilot version)



- **What's new?**
 - A centralized way to build containers for your experiment
- **Why should you care?**
 - will provide a build node to construct customized containers
 - currently these containers cannot be distributed to worker nodes
- **What should you change to use new features?**
 - No changes to other procedures required in this initial pilot phase. If you need a container, this will let you build it.
- **What issues will be addressed next, and why should you care about them?**
 - If the pilot is successful, we plan to add: continuous integration workflow for container building
 - linkage to jobs running on grid or other facilities may come later

Interim Solution for Building containers:

- Very soon, each experiment may request a container build machine.
 - Will look like an experiment interactive machine (SL7 only)
 - NFS mounts very limited; CVMFS available if desired
 - Specific users will be added to docker group
 - Will have Docker and Singularity installed
 - Yes, we're intentionally being somewhat vague on the timeline
- Private Docker container registry will be available as well.
 - One project per experiment
 - Experiments may set permissions on who can read and post
 - (We can't provide a private Singularity registry at this time)

Two General Use Cases for Containers

- Containers that run jobs “somewhere”
- Containers that run arbitrary services

Disclaimers

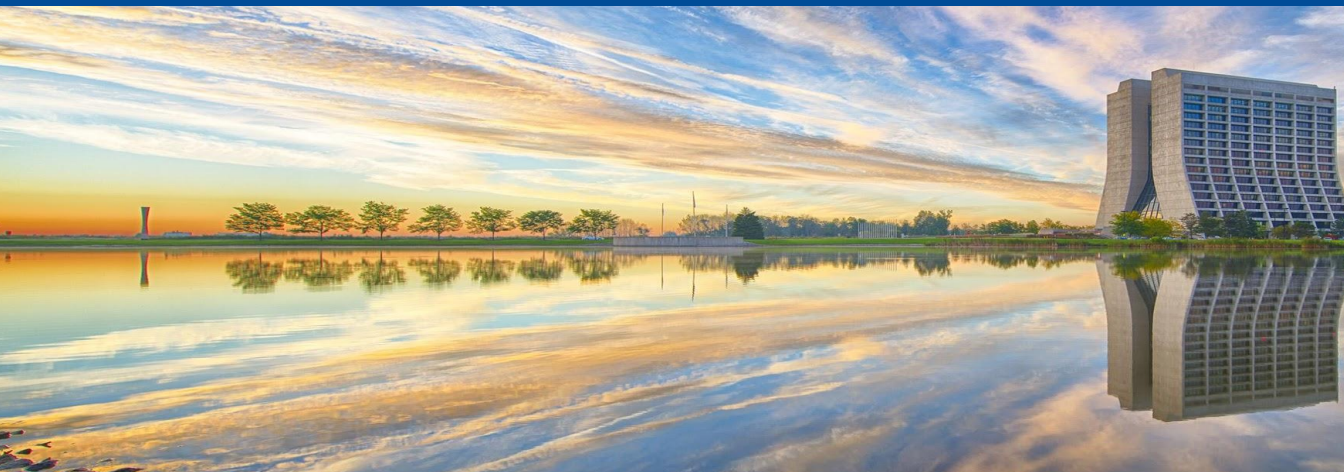
- We are not providing a production environment to deploy containers (yet).
- The expectation is that experiments will select a small number of developers to use these container build servers.
- We do not yet have a good understanding of resource requirements (disk space in particular).
- This is an interim solution, really a pilot at this point.
- There are no backups of these systems, or the on-site Docker registry. The on-site Docker registry will have replication to a standby server.

Disclaimers continued (really?)

- Experiments will have quotas, and will need to police their own users.
- These build servers are not intended for running ‘production’ containers, only for development and testing of containers, so you cannot run service containers on the build node.
- No hard limits initially on container sizes, other than a 100GB total disk space limit per experiment. This may change.

Future?

- Integration testing of containers with a Jenkins service
- Workflow for integration with OSG and CMS batch job operations
- An environment that can support arbitrary containers—intended for persistent services.



GPU resources and availability

Ken Herner
FIFE Roadmap Workshop
Jun 14, 2018

14 Jun 1822:
Charles Babbage presents a “Note on the application of machinery to the computation of astronomical and mathematical tables” to the Royal Astronomical Society

211

NOTE ON THE APPLICATION OF MACHINERY TO THE
COMPUTATION OF ASTRONOMICAL AND MATHEMATICAL TABLES.

By CHARLES BABBAGE, Esq., F.R.S.

From the Memoirs of the Astronomical Society of London.

Read 14th June, 1822.

It is known to several of the members of this society that I have been engaged during the last few months in the contrivance of machinery which, by the application of a moving force, may calculate any tables that may be required. I am now able to acquaint the society with the successful results at which I have arrived; and although it might at the first view appear a bold undertaking to attempt the construction of an engine which should execute operations so various as those which contribute to the formation of the numerous tables that are constantly required for astronomical purposes, yet to those who are acquainted with the method of differences the difficulty will be in a considerable degree removed.

I have taken the method of differences as the principle on which my machinery is founded; and in the engine which is just finished I have limited myself to two orders of differences. With this machine I have repeatedly constructed tables of square and triangular numbers, as well as a table from the singular formula $x^2 + x + 41$, which comprises amongst its terms so many prime numbers.

These, as well as any others which the engine is competent to form, are produced almost as rapidly as an assistant can write them down. The machinery by which these calculations are effected is extremely simple in its kind, consisting of a small number of different parts frequently repeated.

In the prosecution of this plan, I have contrived methods by which types shall be set up by the machine in the order determined by the calculation; and the arrangements are of such a nature that, if executed, there shall not exist the possibility of error in any printed copy of tables computed by this engine. Of several of these latter contrivances I have made models; and, from the experiments I have already made, I feel great confidence in the complete success of the plans I have proposed.

C. BABBAGE.

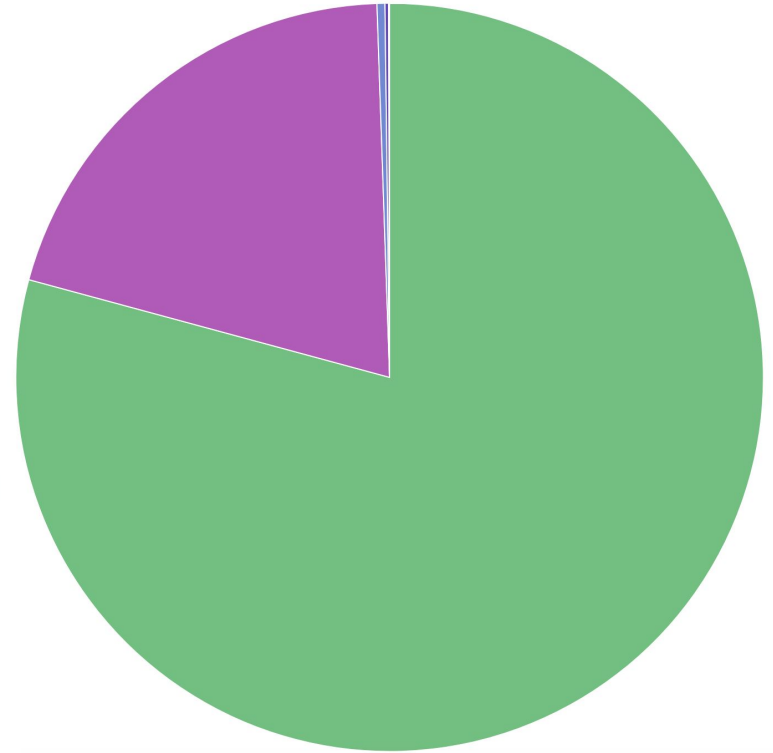
DEVONSHIRE STREET, PORTLAND PLACE,
JUNE 2, 1822.

GPU Access

- Some of you may be familiar with the [Wilson Cluster at Fermilab](#); this requires a separate account
- Other GPU clusters available via OSG and reachable via jobsub, with some extra options
 - Nebraska, Syracuse, and UCSD have slots now. Best to NOT request a specific OS during submission to get the most slots
- General GPU advice
 - Use the following slides to get information about the GPUs in the job
 - Don't confine yourself to only the latest and greatest models unless absolutely necessary. They have the longest wait.

GPUs in action in the past year

- 18,693 jobs
- 5 users, 3 of whom were testing
- **Feedback requested:**
 - No demand?
 - Submission too hard?
 - Resources inadequate?



A GPU test job

- In general you just have to add **--lines='+RequestGPUs=1'** to your `jobsub_submit` command
 - The FNAL frontend will steer you to an appropriate site
 - Can request >1 but no promises made about availability or fast start time. Could also add options for specific site if needed (not recommended)
- So try the following:

```
$ jobsub_submit -G your_expt -M --memory=1000MB --disk=1GB \  
--expected-lifetime=1h -N 8 \  
--resource-provides=usage_model=OFFSITE --lines='+RequestGPUs=1' \  
file:///nashome/k/kherner/basicscript_GPU.sh
```

Getting GPU capabilities

- GPU information is now advertised in machine classad
- You can look for CUDA or OpenCL support
- NVIDIA driver info
- Attributes are advertised in the form of:
<LANG><KEY>=<ATTR>
- To get all info about the machine, you can do "`cat \${_CONDOR_JOB_IWD}/.machine.ad`" somewhere in your executable. Then you can look for specific attributes

GPU attribute examples

- `CUDAGlobalMemoryMb` = 12194
- `CUDARuntimeVersion` = 6.0
- `CUDACapability` = 6.0
- `CUDAComputeUnits` = 56
- `CUDAClockMhz` = 1328.5
- `CUDACoresPerCU` = 192
- `CUDADriverVersion` = 8.0
- `CUDAECCEEnabled` = false
- `CUDADeviceName` = "Tesla P100-PCIE-12GB"
- `NV_DRIVER` = 375.66

Example of machine.ad dump on a
CUDA-installed node

- `OCLOpenCLVersion` = 1.2
- `OCLGlobalMemoryMb` = 2001
- `OCLComputeUnits` = 5
- `OCLECCEEnabled` = false
- `OCLDeviceName` = "GeForce GTX 750 Ti"
- `OCLClockMhz` = 1084
- `NV_DRIVER` = 375.66

Example of machine.ad dump on a
OpenCL-installed node

Using these variables in job requirements

- Using jobsub
- Example where we want to get a Tesla P100 GPU:
 - `jobsub_submit ... --lines='+RequestGPUs=1'`
`--append_condor_requirements='(regexp(\"^Tesla\\P100.+\\$\",TARGET.CUDADeviceName,\"i\")||regexp(\"^Tesla\\P100.+\\$\",TARGET.OCLDeviceName,\"i\"))'`
 - Be careful about cutting/pasting this -- unicode/ascii issues with python 2!
- Find OPENCL support:
 - `--lines='+RequestGPUs=1' \\`
`--append_condor_requirements='isUndefined(TARGET.OCLDeviceName)==FALSE'`
- Finding CUDA support is similar to above (OCLDeviceName → CUDADeviceName)
- Any other combination of vars on previous slide possible. Need to be careful about escaping quotes and spaces with the backslash (\\) though.



Jobsub Status and Plans

Dennis Box

FIFE Roadmap Workshop

June 14, 2018

June 14 2018



JobSub Project Status

- Project was relegated to maintenance mode FY 2017
 - 25% of 1 FTE allocated
 - This turned out to be somewhat unrealistic.
- With this effort level the following was planned:
 - GPGGrid refactoring support, Condor changes
 - High priority bug fixes (see redmine issue list)
- Unplanned changes that arose:
 - GUMS replacement by Ferry
 - Bluearc user area dismounting, replacement with dCache

Increase in effort and future development

- A developer was added to the project, bringing planned effort up to 50% of 1 FTE
 - Welcome Shreyas Bhat!
 - With this surfeit of manpower, it is our obligation, our sacred duty, to wage a campaign of world domination.
 - Jobsub will become the planets pre-eminent grid job submission system!
- Realistically:
 - Jobsub will continue to respond to changes in user requirements and external operating environment.
 - Security takes priority as always
 - New features on a ‘best effort’ basis

Jobsub upgrades coming soon

- SL6 to SL7 migration
 - SL6 end of life 2024, BUT..
 - SL6 uses unsupported Python 2.6
 - No bug fixes/security updates
 - Upgrading SL6 to supported python problematic -> SL7 preferred
- Python 2.7 to 3.x migration
 - 2.7 end of life Jan 1 2020 (18 months)
 - Won't start until SL7 migration happens
- HEPCloud migration
 - We think config file changes will do most heavy lifting
- Highest priority feature requests
 - Tarball distribution via dCache/CVMFS (ongoing)
 - Job submission via config/text file (soon)
- jobsub 1.2.7 is being tested by FIFE, experiments should expect requests to test the new version in the next few days



CVMFS path forward

Dave Dykstra

FIFE Roadmap Workshop

June 14, 2018

June 14 2018



FIFE user code distribution in CVMFS



- Currently user tarballs submitted through jobsub are distributed to worker nodes via dCache resilient pools
 - Not a good use of resources especially with large numbers of jobs
 - User jobs see delays while downloading the code with the current implementation
- We are designing a replacement using automated publishing in CVMFS
 - Should be transparent to the users
 - Includes overcoming a number of challenges with speed of publishing and distribution
 - Upcoming CVMFS features will help with the speed
 - Will use repositories shared across experiments
 - The timeline is under discussion, probably 6 to 12 months

Reminder - StashCache is available for handling partially reused data files

https://cdcv.s.fnal.gov/redmine/projects/fife/wiki/Introduction_to_FIFE_and_Component_Services#OASISCVMFS-process-for-handling-partially-reused-data-files-StashCache

Improvements coming to CVMFS publishing



These improvements are relevant to user code distribution and may be useful for experiment code distribution.

- **Faster propagation of updates**
 - Time for stratum 1 updates and worker node updates is already reduced; should now be on the order of 15 minutes between publish and availability on worker nodes
 - Longer delays can happen on stratum 1s under load
 - These times will be further reduced
- **Parallel publishing in one repository**
 - Current release is not production ready and hard to deploy; requires separate servers or docker containers
- **Direct publishing of tarballs**
 - Coming in next major cvmfs release
 - Speeds up publishing and allows parallel publishes in one server without containers

- A couple of U.S. HPC sites (PSC and TACC) have installed cvmfs on worker nodes, and others are providing the files through other means
 - Some require a single image file with all code, which can be quite a challenge to generate
- Singularity, which is popular on HPC systems, has future plans to support a plugin to start cvmfs using its setuid privileges
- The latest Linux kernel is about to get a patch to allow mount of fuse filesystems in unprivileged namespaces
 - could be used for cvmfs after the HPC kernels eventually get upgraded



Storage

Dmitry Litvintsev

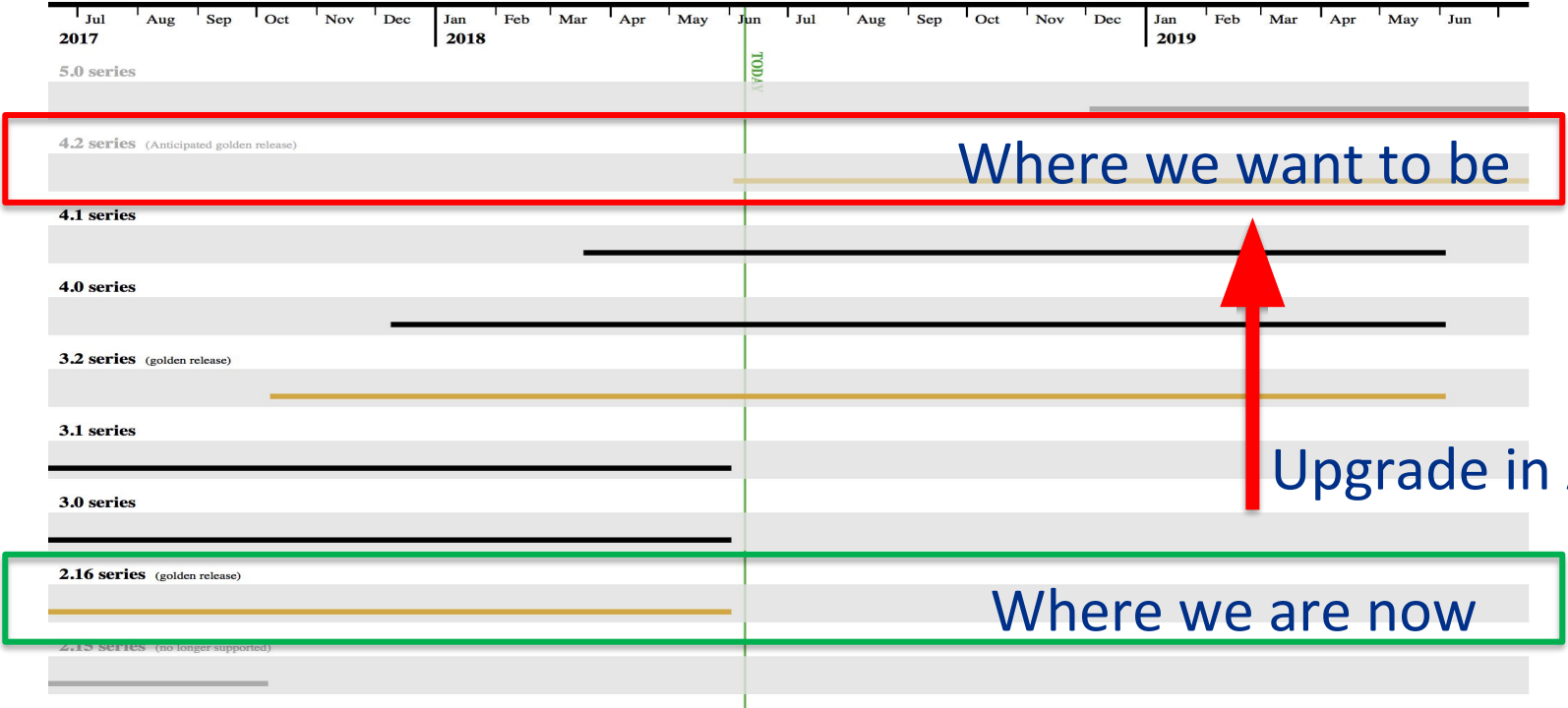
FIFE Roadmap Workshop

June 14, 2018

dCache, what is new

dCache server releases

... along with the series support durations.



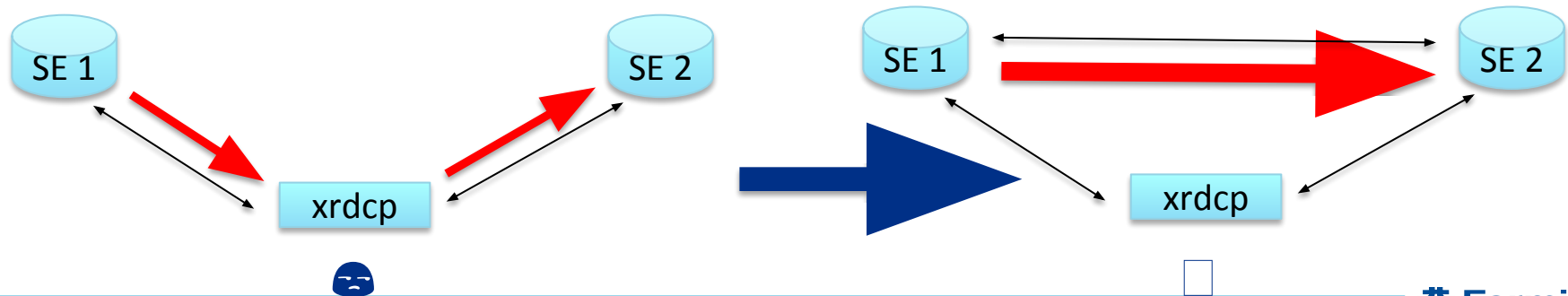
Where we want to be

Upgrade in Aug/Sep

Where we are now

dCache 4.2.0 what is new

- Interface to Frontier Experiments Registry (FERRY) a RESTful service that provides DN/FQAN maps (like gridmap and vorole-map) that can be periodically pulled by dCache authorization service:
 - removes dependency on GUMS (which is a SPOF and unsupported s/w component)
- Many improvements in dCache NFS server targeted at better handling misbehaving clients, overloaded pools and network glitches.
- Experimental support for mutable files. It will be possible to modify files in dCache (with some restrictions).
- “Real” 3rd party xrootd transfers:



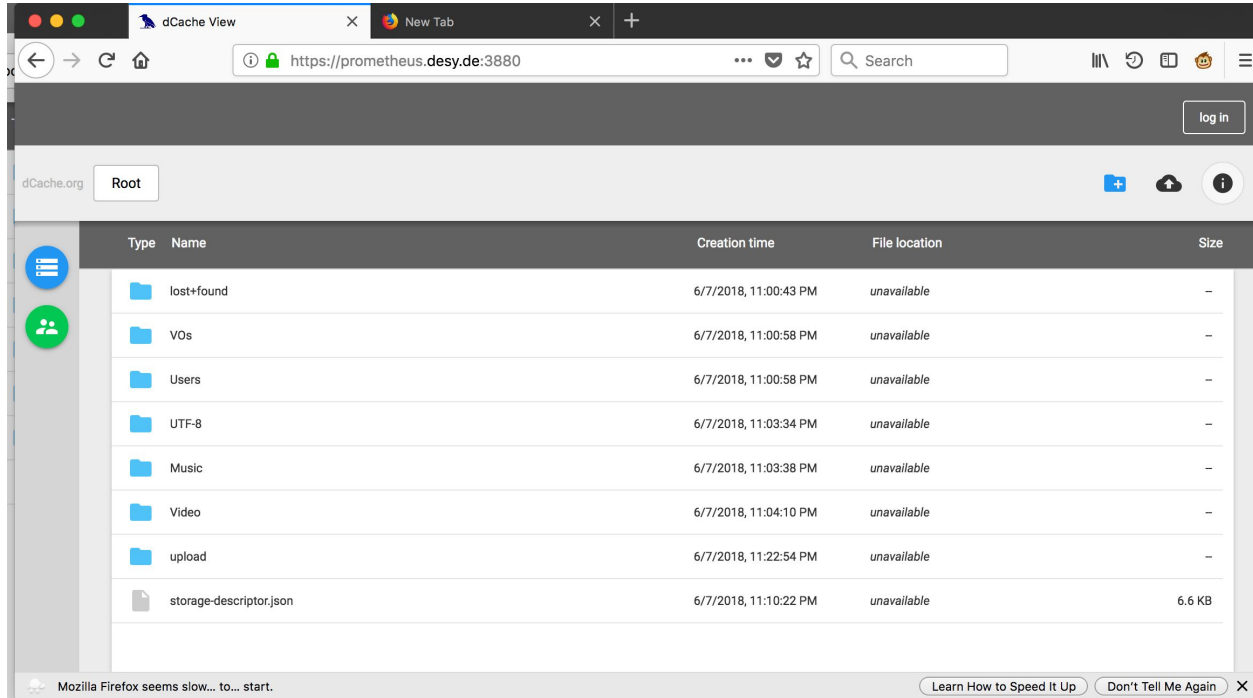
dCache 4.2.0 what is new

- Experimental support for event notification. A publish/subscribe protocol allowing clients to subscribe and get notifications of desired events:
 - Particularly of file locality changes. No need to query system “is my file staged yet?” or “Are you on tape yet?”.
 - Greatly improves efficiency w/ workload management (WM) systems like SAM or Rucio.



dCache 4.2.0 what is new

- Improvements in dCache RESTful based dCache-view interface (similar to google drive) feature.

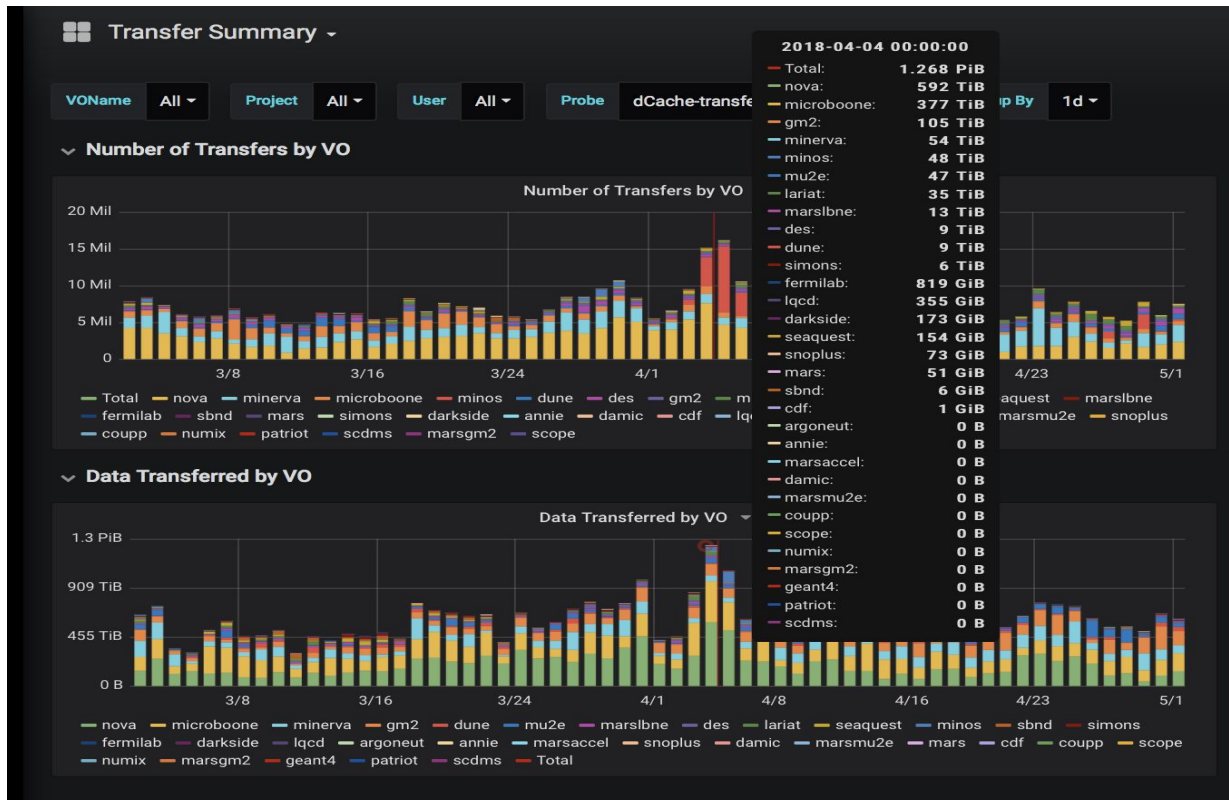


Type	Name	Creation time	File location	Size
Folder	lost+found	6/7/2018, 11:00:43 PM	unavailable	-
Folder	VOs	6/7/2018, 11:00:58 PM	unavailable	-
Folder	Users	6/7/2018, 11:00:58 PM	unavailable	-
Folder	UTF-8	6/7/2018, 11:03:34 PM	unavailable	-
Folder	Music	6/7/2018, 11:03:38 PM	unavailable	-
Folder	Video	6/7/2018, 11:04:10 PM	unavailable	-
Folder	upload	6/7/2018, 11:22:54 PM	unavailable	-
File	storage-descriptor.json	6/7/2018, 11:10:22 PM	unavailable	6.6 KB

dCache 4.2.0 what is new

- Improved RESTFul monitoring interface (and web GUI):
 - allows easy integration with external monitoring tools/frameworks (like Grafana and thus FIFEMON).
- FTP:
 - pipelining (re-use connection for multiple transfers).
 - better compatibility with globus-online.
- XRootD:
 - Support of FUSE driver.
- HTTP:
 - Https on re-direct (encrypted data transfers).
- Pools:
 - multiple checksum support (Adler32, MD5).

Public dCache: mark that date



on 04/04/18 Public dCache delivered 1.3 PiB of data in one day.

Thank you!



Data Management

Robert Illingworth

FIFE Roadmap Workshop

June 14, 2018

Upcoming SAM changes

- Experimental rewrite of the SAM station
 - Existing version is a hugely complicated C++ application
 - Contains much we no longer need; hard to make changes
 - Redone in Python; can run a project with only ~1500 lines code
 - Some functionality still missing – most notably dCache staging
 - Looking for some volunteers willing to test with real work
- Potentially allows new features
 - For example, handle restarted jobs much better
 - if we see the same job id again, mark the old one and its files as cancelled, then send the same files through again.
 - Requires the jobs to tell SAM the Condor job id; some are putting this in the description field; some aren't

Longer term planning

- We are investigating replacing some/all of SAM with Rucio
 - Rucio is an ATLAS developed product
 - Other HEP experiments and non-HEP users are or are considering using it
 - Could become a standard for the field
 - Powerful distributed data capabilities
 - Declarative rules for where files should be placed, e.g.:
 - “Ensure there are at least 2 copies of all raw data, one at CERN and one at FNAL”
 - “Ensure at least 3 copies of this dataset are available amongst all the storage sites tagged ‘europe’”
 - Doesn’t do metadata like SAM does

Plans for Rucio

- Currently setting up prototype system for DUNE
- Next step will be to consider a hybrid system of SAM metadata and Rucio file placement
 - Minimal disruption to existing users, as most don't interact directly with SAM locations
 - Tools will need some work (sam4users, F-FTS)
 - So will the SAM station
- Longer term may want to replace SAM metadata with a new, more integrated with Rucio, system
 - Fix some of the design issues with the SAM DB along the way
 - Much more disruptive; expect only upcoming experiments will change to this



Security

Mine Altunay

FIFE Roadmap Workshop

June 14, 2018

Outline

- OSG CA Shutdown
 - Future Steps for Fermilab
 - Host and User Certs
 - InCommon IGTF CA
 - Switching from CILogon Basic CA to CILogon Silver CA
 - Future Steps for Experiments
 - Non-Fermilab sites without InCommon Membership
 - DigiCert Grid CA
- Long-term plans on Certificates
 - Let's Encrypt, an online non-IGTF CA
 - Eliminating majority of certs
- HEPCloud update
 - Status
 - Future steps
- Recent talk by Jeny Teheran for Security Day:
<http://cd-docdb.fnal.gov/cgi-bin/ShowDocument?docid=5884>

- OSG CA has shutdown its services as of May 31, 2018
- Fermilab has already been working on the transition.
- Future steps are:
 - For Personal Certs, use CILogon Basic CA.
 - For Host Certs, we will purchase InCommon IGTF CA service. We will send you more information on how to acquire the host certificates
- What is InCommon IGTF CA?
 - It is an IGTF accredited CA. Will provide host certificates.
 - CILogon Basic CA cannot provide host certs, only personal certs. That is why we purchase services from InCommon IGTF CA.
 - It is provided by InCommon. in case you are curious: Cilogon Basic CA also serves the InCommon community, but only personal certs.

- We will make a contract with InCommon soon and pay for this service.
- Each site (domain) needs to individually register with InCommon and pay for it. So Fermilab cannot help collaborators at universities.
- If you are a member of an experiment, e.g. CMS or Atlas, and some of your universities are not InCommon members, then it is an issue.
 - We are working with experiments on this. We found that DigiCert Grid CA is willing to provide certs for such sites.
 - If you need a personal cert and do not have a Fermilab account, we suggest that you open an account here and start getting CILogon Basic CA certs.

List of CAs we will use

- Fermilab personal certs:
 - CILogon Basic CA (will switch to CILogon Silver CA, explain in upcoming slides)
- Fermilab Service Certs:
 - InCommon IGTF CA
- For resources (e.g. sites) outside of fnal.gov domain:
 - Either InCommon IGTF CA or DigiCert Grid CA. Both are IGTF accredited and can provide user and personal certs
 - if cannot purchase InCommon IGTF or DigiCert CA, then use Let's Encrypt CA. But Let's Encrypt is NOT IGTF accredited.
- For scientist without Fermi accounts:
 - If possible, open an account at Fermilab and get CILogon Basic CA certs
 - If not, try CERN CA, InCommon IGTF or DigiCert Grid CA (CERN CA can give personal certs, but cannot give host certs outside of cern.ch domain.)

Changes with CILogon Basic CA

- Fermilab has been getting its personal certs from CILogon Basic CA.
- This CA is operated by XSEDE and NCSA under CILogon project umbrella. CILogon also has another CA called Silver CA.
- The difference between Basic and Silver is that Silver operates at an higher assurance level and has a classical IGTF accreditation. Basic has a special accreditation called IOTA from IGTF.
- Many European sites do not recognize the Basic CA as it has a different accreditation/assurance level.
- We will switch to Silver CA in September so all our users' certs can be accepted automatically everywhere.

Changes with CILogon Basic CA

- How transition from Basic to Silver CA impact users?
 - It should be minimal impact.
 - We will be making changes to only the backend. The user frontend will remain the same.
 - In the backend, instead of sending certificate signing request to Basic CA, we will send it to Silver CA.
- This change will help our users in international collaborators a lot.

Long-term plans with certificates

- IGTF accreditation is a cumbersome and expensive step. There are only 4 CAs in the US with accreditation: DigiCert Grid CA, InCommon IGTF CA, Cilogon Basic and Silver CAs.
- We are locked into a very small CA space. However, there are hundreds of commercial CAs out there albeit without an accreditation.
- We are exploring alternatives.
- One strong contender is Let's Encrypt CA. We are talking to IGTF to create a special profile for this type of CAs, so we can use them.
- Obviously, our long term goal is to get rid of certs completely.

HEP Cloud Security Update



- We are still working on the risk assessment. But, we are on track to complete this and receive our Authority to Operate (ATO) later this year.
- HEPCloud will make a positive difference in our user's life.



FERRY Service

Tanya Levshina

FIFE Roadmap Workshop

June 14, 2018



Goal: Create account registry and quota management for Fermilab experimenters.

- Allow various services (FermiGrid Batch, dCache, NIS, POMS and others) get user's authorization, quotas and other account related information from one centralized place.
- Automate as much as possible the workflows that are initiated from Service Now as requests for an account, storage, and quota increase.
- VOMS-ADMIN and GUMS products have been dropped from the OSG software stack. Ferry, among other things, is a replacement for these services.



FERRY

 Fermilab

Frontier Experiment RegistRY(FERRY)



- **Benefits for user:**
 - Automation!
 - Quicker turnaround
 - Less error prone
- **Impact:**
 - All requests for account affiliation, grid roles, batch and storage quota modifications should be done via relevant SNOW forms.
 - A VO administrator (offline coordinator) should approve requests ONLY via SNOW, no other actions are necessary.
- **Status:**
 - Currently in integration phase with all service providers.
 - Production release in September.



FERRY



Frontier Experiment RegistRY(FERRY)

Request Catalog > Scientific Job Management > Modify user on Jenkins Cluster

Request user access or removal of the ability to edit build job configurations and trigger builds on Jenkins for an experiment.

* Add or Remove

-- None --

Affiliation (Experiment)

Requested for

Tanya Levshina

Order this Item

Submit

Request Catalog > Scientific Job Management > Modify Quota on FerraGrid

Request to increase (maybe just temporary) total number of batch slots/create accounting group/modify accounting allocation within quota

Requested for

Tanya Levshina

Affiliation (Experiment)

E-923 (NOVA)

Please Select the Batch Job Accounting Group

nova.keepup_prio

What cluster would you like to increase?

ferragrid

Are you requesting a change in the total number of slots available to the experiment? (yes/no)

yes

Additional percentage of job slots being requested:

Instructions:

If you want to specify your hierarchical quotas in percentages enter a number between 0.10 and 0.99

10

Submit

Request to add a user to Jenkins cluster



FERRY

Fermilab



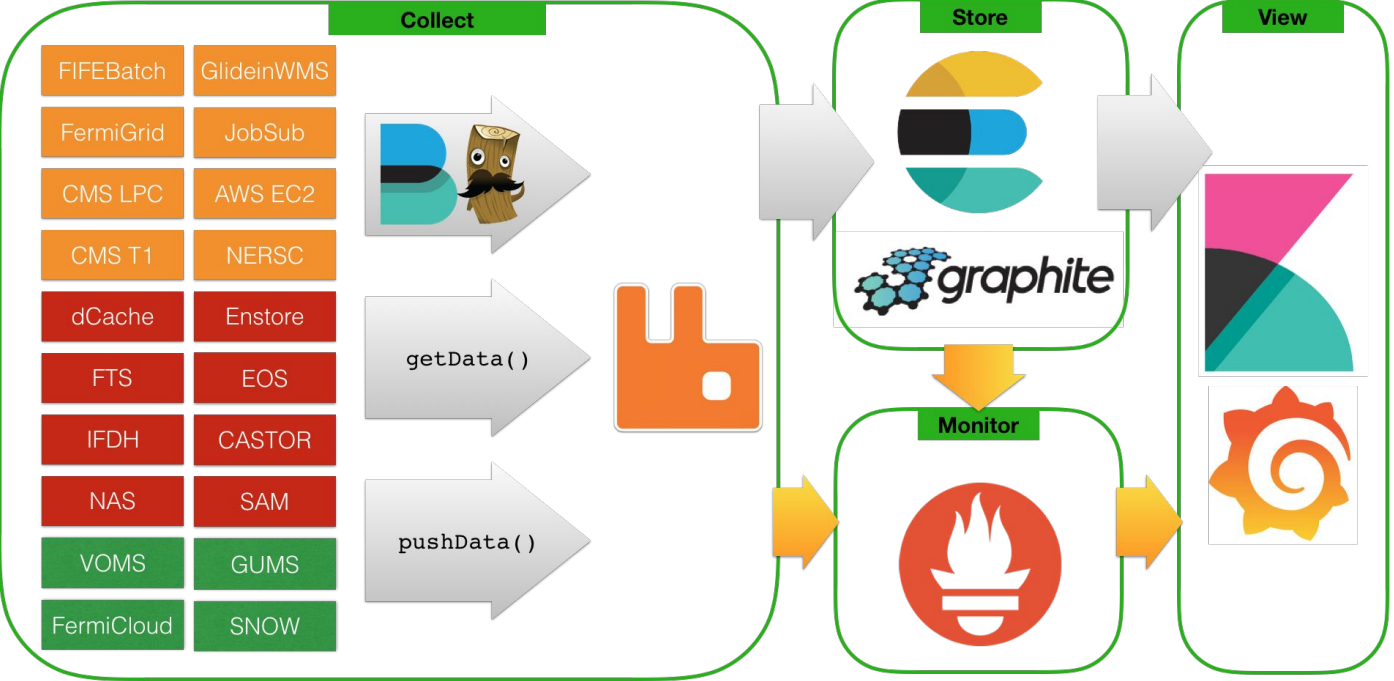
Landscape

Kevin Retzke

FIFE Roadmap Workshop

June 14, 2018

Landscape is FIFE, HEPCloud, and more



The ~~Four~~ Three W's

What's new?

- Opening access to Kibana
- Offline production shifter dashboards
- Inter-site monitoring with CERN

Why should you care?

- Kibana will allow you to show statistics about your batch jobs and transfers
- Your shifters could have concise displays of offline services

What should you change?

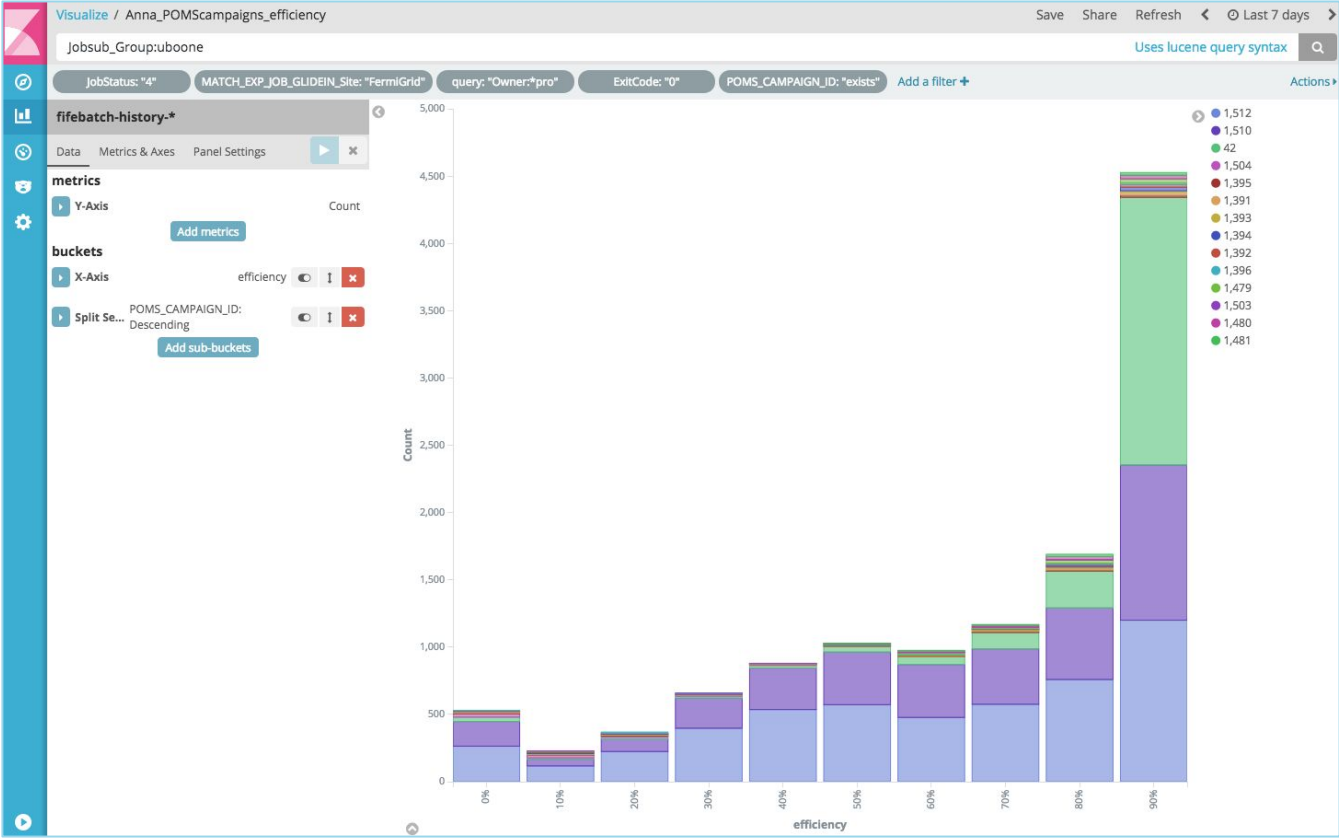
- Your approach to monitoring:
 - dynamic not static
 - interactive, not passive

Opening Read-only Access to Kibana

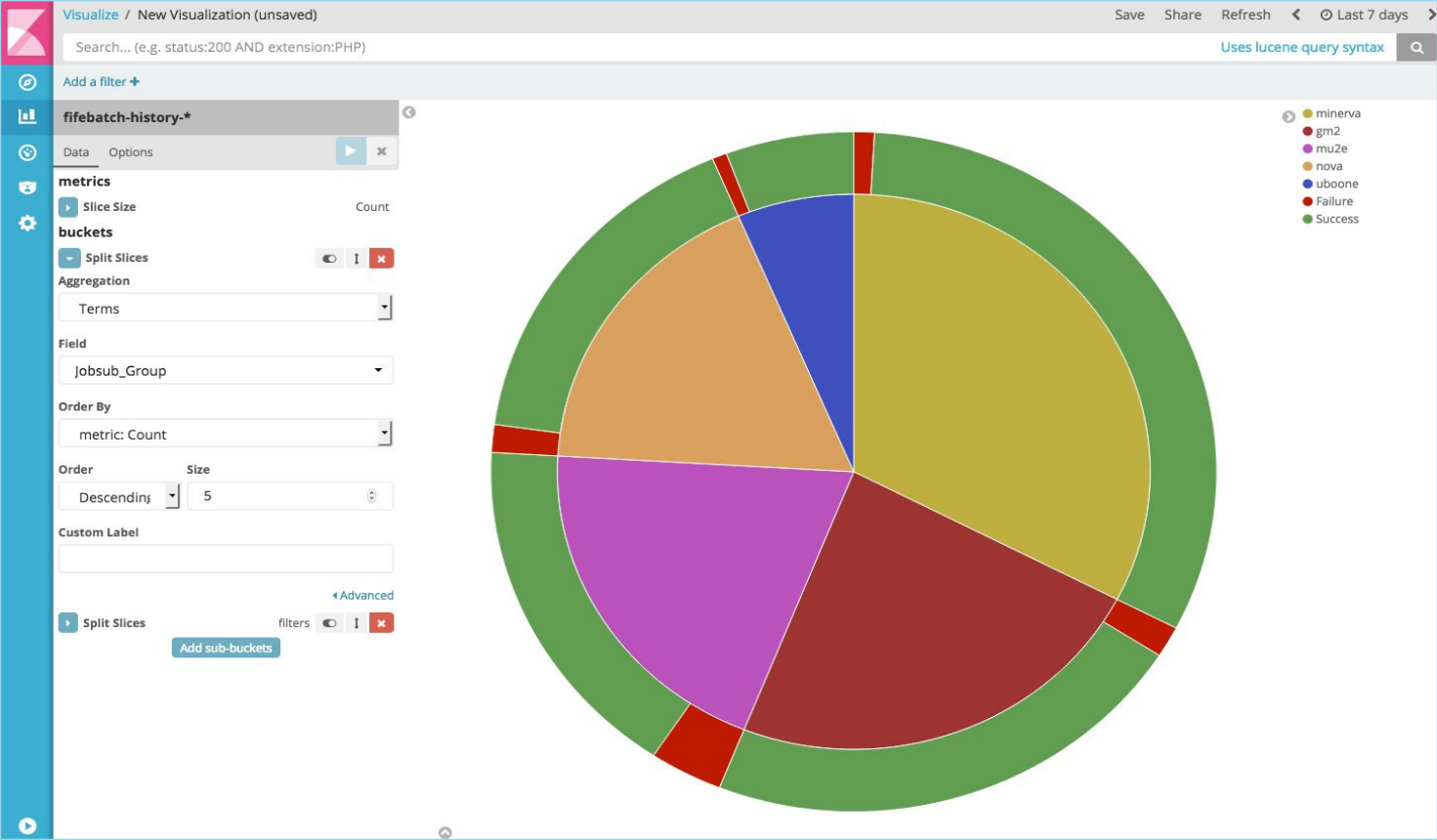
<https://landscape.fnal.gov/kibana/>

- Access complete details on jobs and transfers, fully searchable and filterable
 - Job events (submitted, started, etc) collected in real time
 - Select attributes for jobs currently in queue (idle, running, and held)
 - Final attributes and statistics for finished jobs (completed and cancelled)
 - IFDH file transfer events (start transfer, finish transfer) including user, host, source and destination paths, and transfer durations.
- Create and share custom visualizations on job efficiency, success rate, transfer times, file popularity, and more. Some examples include...

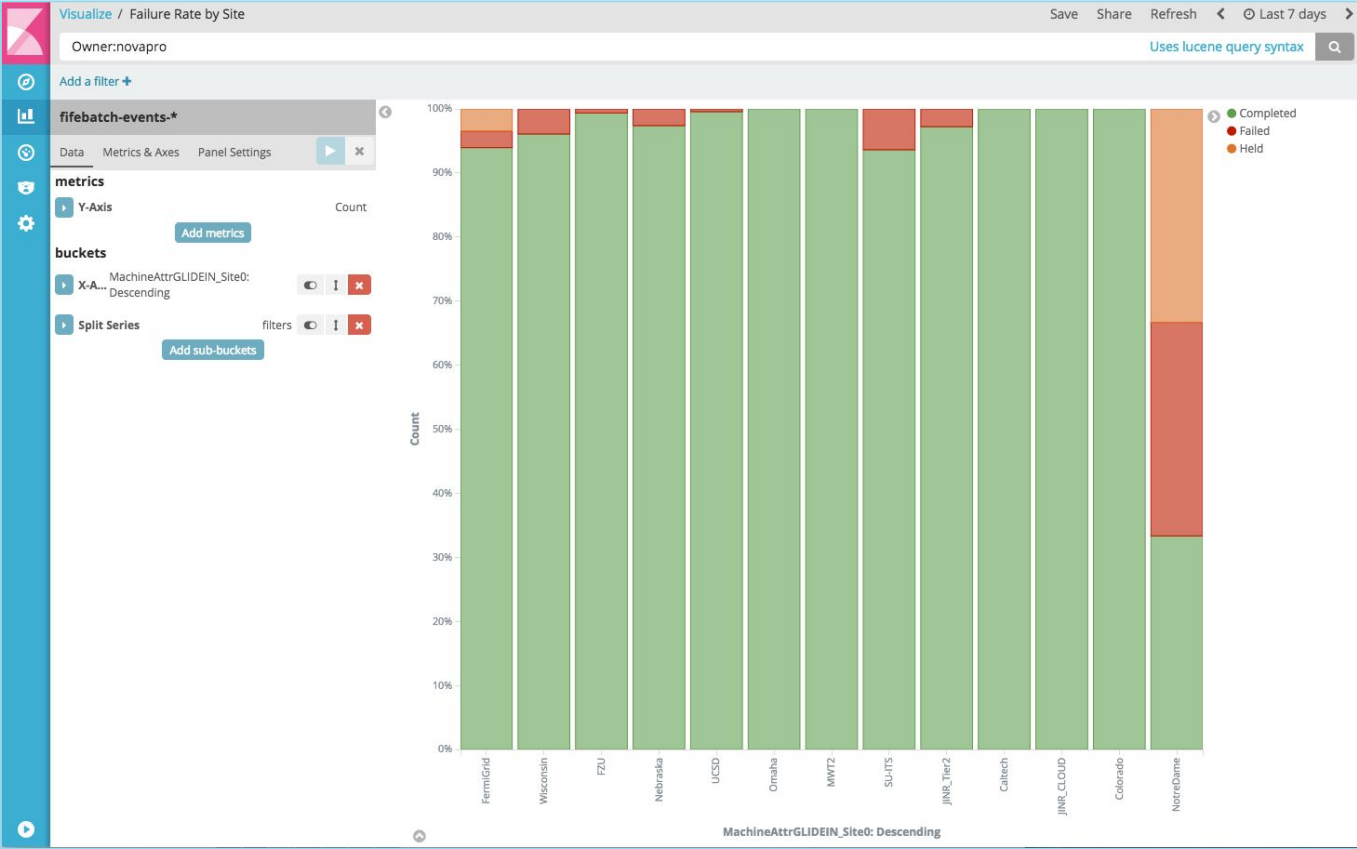
Job efficiency by POMS Campaign



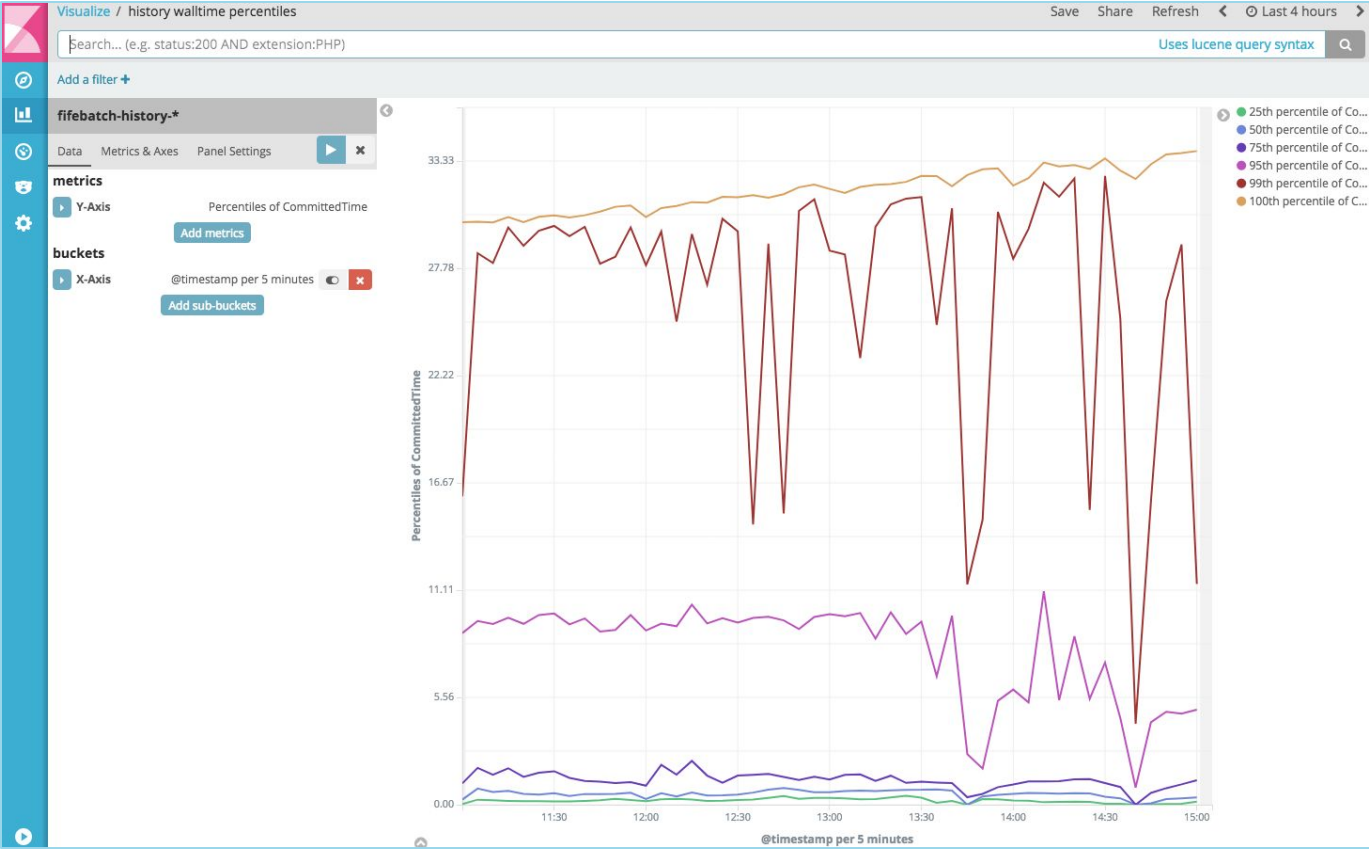
Job Success and Failure by User, Campaign, etc.



Job Success and Failure by Site



Job Walltime Percentiles

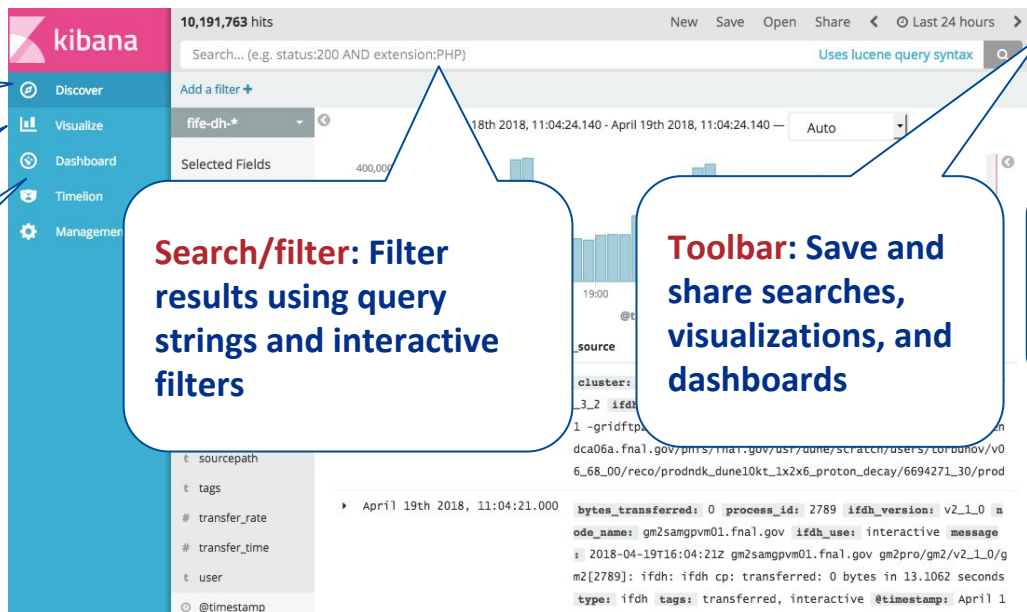


Kibana interface overview

Discover: Search and Filter raw documents

Visualize: Graph document statistics

Dashboard: Combine visualizations



The screenshot shows the Kibana interface with a search bar containing 'fife-dh-*' and a time range of '18th 2018, 11:04:24.140 - April 19th 2018, 11:04:24.140'. The interface displays a bar chart visualization and a list of search results. The search results include fields like 'sourcepath', 'tags', 'transfer_rate', 'transfer_time', 'user', and '@timestamp'. A callout box points to the search bar, and another points to the time range selector.

Search/filter: Filter results using query strings and interactive filters

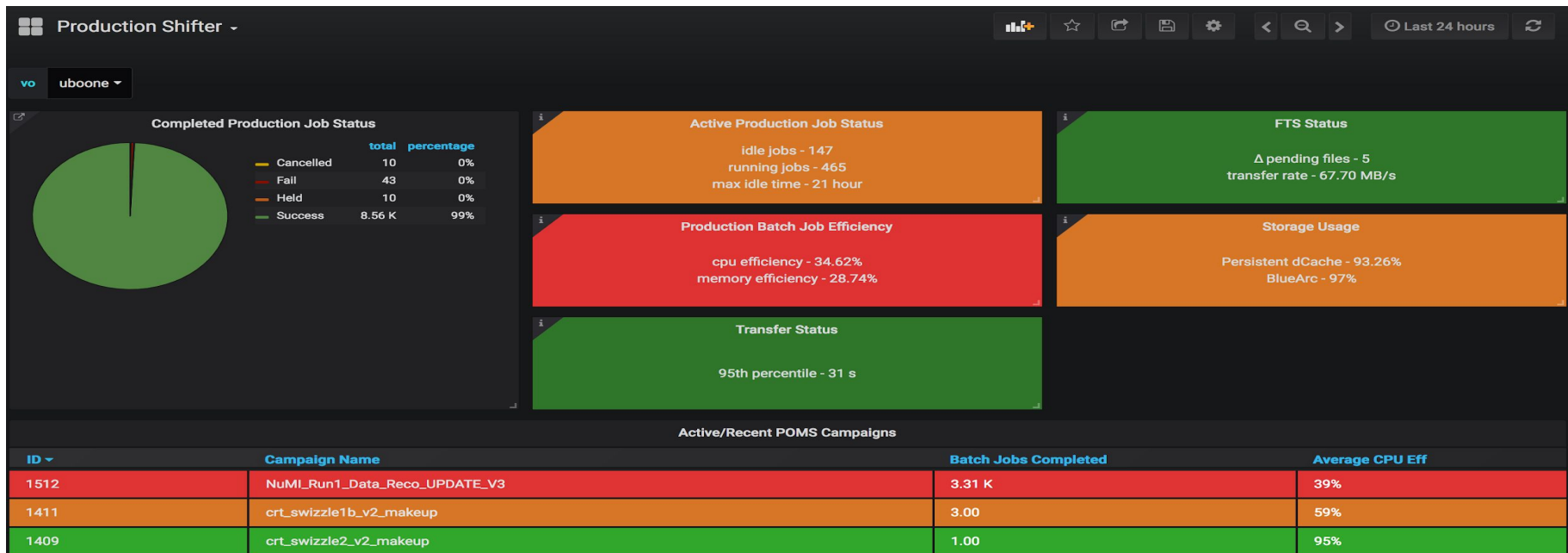
Toolbar: Save and share searches, visualizations, and dashboards

Timepicker: change time range of interest and refresh.

More: <http://cd-docdbcert.fnal.gov/cgi-bin/cert/ShowDocument?docid=6411>

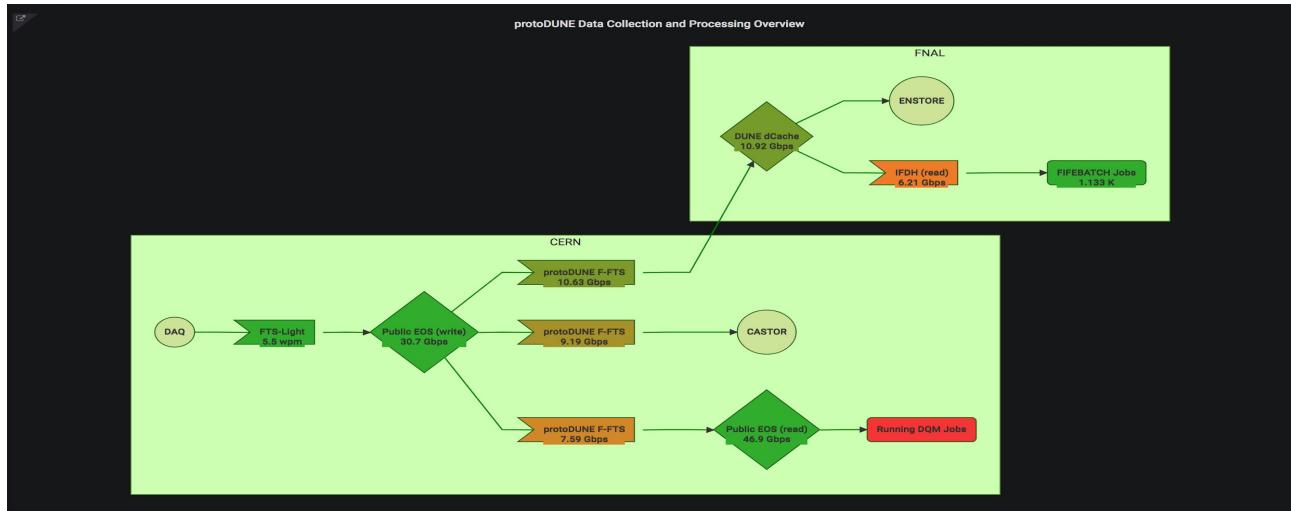
Offline Production Shifter Dashboard

- At a glance, identify if there are issues with production batch jobs, file transfers, FTS, storage, etc.
- Work-in-progress, let us know what you want to see!



Bespoke monitoring example: ProtoDUNE

- This spring we worked with protoDUNE collaborators and service providers at FNAL and CERN to provide comprehensive monitoring for intra- and inter-site data challenge.
- What can we do for you?





POMS Update

Stephen P. White & Marc Mengel

FIFE Workshop

June 2018

What's New?

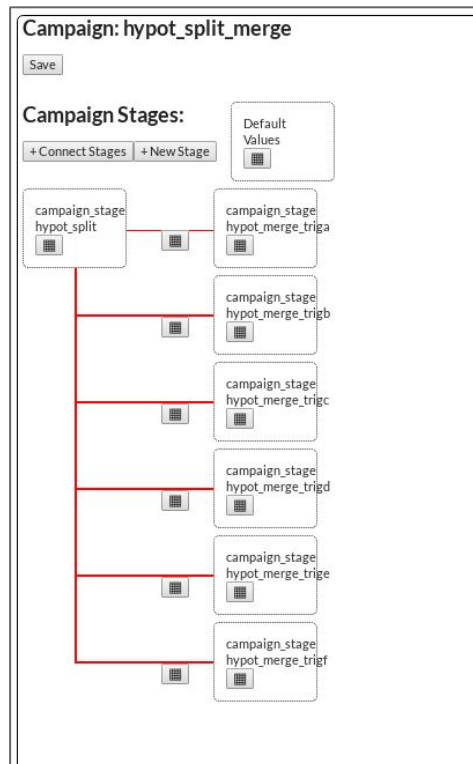
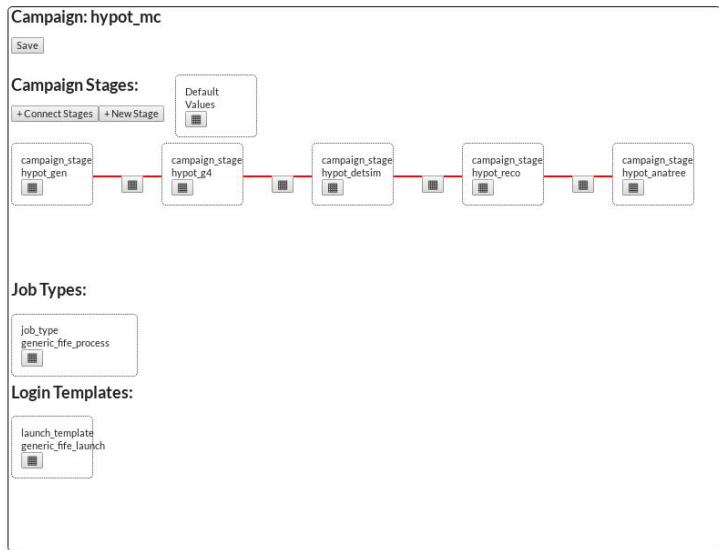
- Top-down tools
 - Campaign/Workflow editor
 - Campaign/Workflow .ini file
 - Uploader for .ini files
- Establishing accounts
- Performance
- Documentation revamp

Why?

- Previously to make a 3 stage campaign
 - Create job type(s) as needed, one screen each
 - Create each Campaign Stage one screen each
 - Mark any dependencies in the downstream stage
 - Tag each Campaign stage with the campaign name
- With Top-down tools
 - Get a sample similar workflow and either
 - Clone into GUI editor and make changes on screen
 - Download as .ini ; edit; and upload

Top Down Tools

- Campaign Editor - (Currently beta)
 - Create/edit Stages and support files, 1 screen
 - Clone from samples, or previous Campaigns
 - Customize for task at hand
 - Popups to edit properties



- .ini file for Campaign/workflow
 - whole campaign/workflow in one file
 - edit as needed (or modify in the DB with Campaign editor)
 - upload to DB with poms_client's "upload_wf"

```
[campaign]
experiment=samdev
poms_role=production
tag: mwm_test_tag
campaign_stage_list=mwm_test_1 mwm_test_2  mwm_test_17
[campaign_stage mwm_test_1]
dataset=unknown
software_version=unknown
vo_role=Analysis
cs_split_type=None
job_type=mwm test generic
param_overrides=[]
test_param_overrides=[]
completion_type=complete
completion_pct=95
launch_template=samdev-generic
[campaign_stage mwm_test_2]
...
```

Establishing POMS Access

- New, automatic POMS access with VOMS membership
 - VOMS role can be Analysis, Production or Coordinator
 - FERRY is supported too
- Triggered by existing Service desk procedures
- New experimenters are added to poms_announce
 - Yes, you can remove yourself
- Takes up to 24 hours after your added to VOMS
- Note: Manual, on the fly additions, bypassing VOMS are no longer possible.

- We are revising documentation to define, then describe, a top down approach to Campaigns
- Define all “POMS” terms used.
- Provide configuration and workflow file documentation.
 - Describing the purpose of login_setup, job_types and other table/files.
 - List all attributes available for each file, with a description of each attribute.
- Update the internal help pages, question mark links, for completeness and accuracy:

https://cdcv.s.fnal.gov/redmine/projects/prod_mgmt_db/wiki

- As number of managed jobs increases
 - Resources to track them increases
 - Concluding state of submissions from jobs more costly
- Changes so far
 - Split condor_q updates onto pomsgpvm02
 - Parallelize updates into POMS
 - Locking improvements for Postgres
 - Moving Postgres instance to its own box
- Upcoming changes
 - Not tracking individual jobs at all internally
 - Polling Landscape for Submissions
 - Links to Landscape for status pages



Continuous Integration Service

Vito Di Benedetto

FIFE Roadmap Workshop

June 14, 2018



- The CI tests use case is to run “*short*” tests to quickly provide a report on the code status
- The CI Validation use case is to run more accurate tests using the grid to run CPU intensive and long-running tests processing many (thousands ?) events
 - ✓ Currently used by DUNE, uBooNE, mu2e and GENIE
- The CI Validation can be triggered:
 - on demand to test *tags/feature branches* (DUNE and uBooNE use case)
 - periodically to test a specific production branch
 - ◆ mu2e runs it nightly to monitor the code reproducibility
 - ◆ GENIE runs it weekly to monitor the code behavior

- **Some of the available features are:**

- ✓ Allows to run multiple CI Validation workflows in the same CI build (used by mu2e to run nightly 6 CI Validation workflows)
- ✓ Allows to run multiple stages in the same job (reduce I/O and increase the efficiency)
- ✓ Provides resource usage stats (RSS memory, elapsed time, CPU%, exit code, ...) for each stage individually
- ✓ For each job allows to run:
 - ◆ *setup step* to update seed configuration, check input availability, ...
 - ◆ *finalize step* to parse logs to extract specific info, ...
- ✓ Allows to report CI Validation status to a Slack channel
- ✓ Allows to validate results using a validation procedure provided by the experiment

CI service: new features in the CI Validation

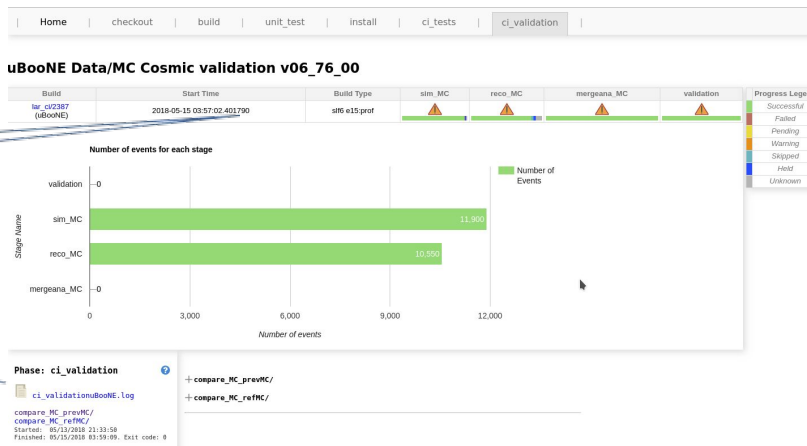
- **Some of the available features are (cont'd):**
 - ✓ The CI Validation dashboard can organize validation plots in folders
 - ◆ It supports file formats: pdf, gif, jpeg, png, ps and txt
 - ◆ Allows custom caption to describe plots
 - ◆ Allows to track history of validation statistics (efficiency, chi squares, ...) using time series

uBooNE CI Validation dashboard

The markers give access to the stage job stats

Folders with validation plots

Links to time series for validation stats



CI service: new features in the CI Validation

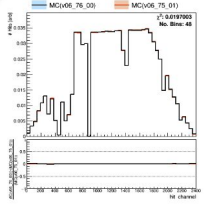
uBooNE CI Validation dashboard



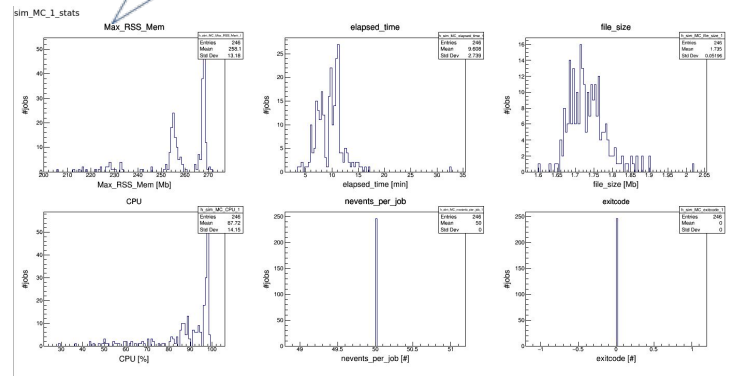
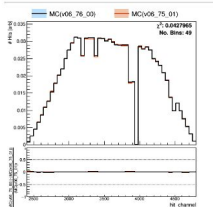
Folders with validation plots expanded

The stage marker gives access to the job resource stats

```
Phase: ci_validation
ci_validation@boNE.log
compare_MC_growMC/
ci_validation@boNE.log
compare_MC_growMC/
compare_MC_ci5prof/
Started: 05/13/2018 21:33:58
Finished: 05/13/2018 03:58:09. Exit code: 0
ChisqValues
```



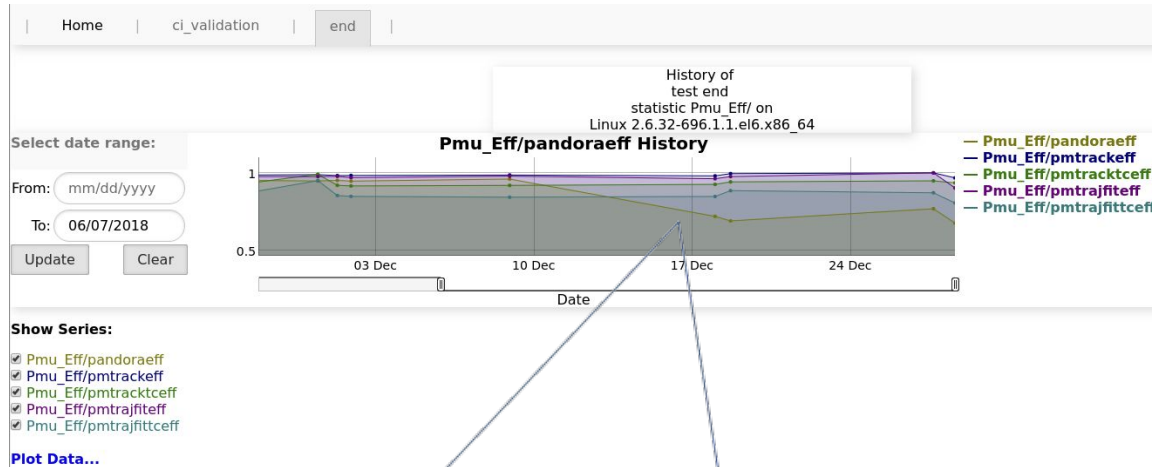
hit_channel_u. The number of 1PK hits on the U (first induction plane) by channel number, summed to try and wash out statistical fluctuations.



CI service: new features in the CI Validation

DUNE CI Validation dashboard

Validation time series



The CI Validation reported a significant drop in the track reconstruction for one of the reconstruction algorithms
Developer were quickly alerted to investigate and get it fixed

CI development: Roadmap

- Improve and extend existing functionality
 - ✓ CI users keep providing feedback and feature requests
- Provide support for authentication to access the CI validation dashboard (useful for experiment that make use of real data in the CI Validation workflow)
- Provide support for direct links to Landscape (FIFEmon/Kibana) to monitor CI Validation job status
- Provide support to compare validation plots against reference data
 - Overlay normalized plots for *current* and *reference* data
 - Use of a chi square to evaluate the comparison agreement
- Provide support for code memory/runtime profiling using a dedicated tool, we are considering “IgProf” and “Open|Speedshop”
- Improve CI test configuration to be more user friendly



CSV_QE_API: Type Safe C++ Data Retrieval

Brandon White

FIFE Roadmap Workshop

June 14, 2018

- Provides a user friendly type safe C++ interface to query PostgreSQL databases using the Query Engine REST style interface
- Can be used both inside and outside of ART projects
- Header only library for ease of inclusion into projects
- Version v1_1 is currently available via UPD
- Depends upon LibWDA v2_24_0
- Only allows for simple columnar queries; more complex queries such as joins are not supported by Query Engine (Tables may be joined by utilizing views)

Connecting to a DB with CSV_QE_API

- Databases must have the Query Engine interface enabled via a request to the Scientific Database Applications group
- All experiment conditions databases have Query Engine enabled
- Experiments will be provided a specific URI for DB access when the system is set up, with the experiment responsible for dissemination
- Concurrency concerns are handled by LibWDA under the covers, alleviating the experiment of this concern under normal conditions

Creating a Database Abstraction

First construct an object of the `query_engine` class templated on the column data types, with the DB connection arguments, and column names existing in the table of interest (must match the number of template data types)

```
#include "csv_qe_api/query_engine_api.h"

query_engine<int, double, string> qe(
    "http://ifb-data.fnal.gov:port/your_experiments/query_engine/uri_path",
    "my_db_name",
    "my_table_name",
    "int_column",
    "double_column",
    "string_column");
```

Performing Queries with CSV_QE_API

Having constructed a query_engine abstraction to your database, you may now query the specified table via SQL themed member functions of the query_engine class

```
auto results1 = qe.limit(10).get();  
// Retrieve the first 10 rows of int_column, double_column, and string_column  
  
auto results2 = qe.where("int_column", "lt", "2").get();  
// Retrieve rows of int_column, double_column, and string_column where int_column is  
// less-than 2  
  
auto results3 = qe.where("int_column", "lt", "2").order_by("double_column").get();  
// Retrieve rows of int_column, double_column, and string_column where int_column is  
// less-than 2 and order results by double column  
  
auto results4 = qe.where("int_column", "lt", "2").order_by("double_column", '-').get();  
// Same as previous example, but the order direction is reversed
```

Incremental Query Construction

Queries are not executed until the `.get()` member function of the `query_engine` class is called, allowing for queries to be incrementally constructed at runtime. The following query:

```
qe.limit(5);  
qe.where("int_column", "1t", "2");  
qe.order_by("double_column", '-');  
auto result = qe.get();
```

Is the same as:

```
auto results = qe.where("int_column", "1t", "2").order_by("double_column", '-').limit(5).get();
```

But allows for runtime mixing and matching on what to query, why to query it, and when to execute the query

Querying Array Types with CSV_QE_API

CSV_QE_API also supports querying of PostgreSQL array data types in a very similar fashion

```
// Create a query engine templated on a vector of the data type contained therein
query_engine<std::vector< int>> qe_array(db_uri, database, table, array_column_name);
auto vector_test_result = qe_array.get();

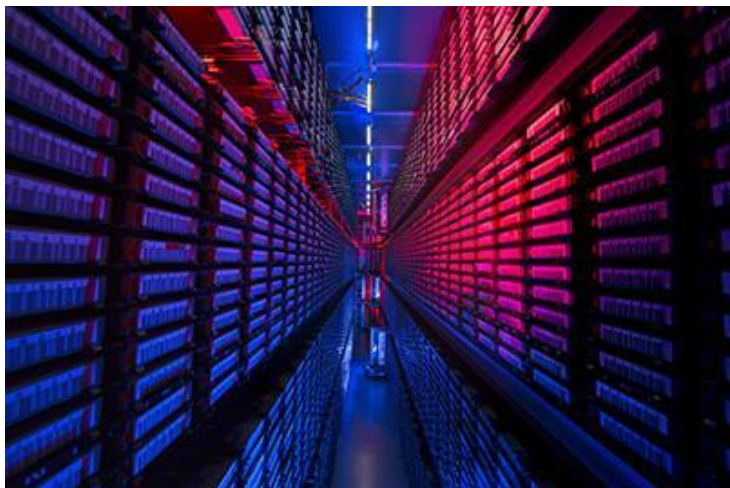
for (const auto& row : vector_test_result) {
    for ( const auto& vec_element : column<0>(row) ) { // Print each vector element
        std::cout << "Vector element: " << vec_element << std::endl;
    }
    std::cout << std::endl;
}
```

Further information on CSV_QE_API usage available via the Redmine wiki:

https://cdcvs.fnal.gov/redmine/projects/csv_qe_api/wiki

Coming Soon for CSV_QE_API

Distribution via SciSoft and CVMFS is coming in the future



Contact Brandon White with any questions, concerns or requests at bjwhite@fnal.gov

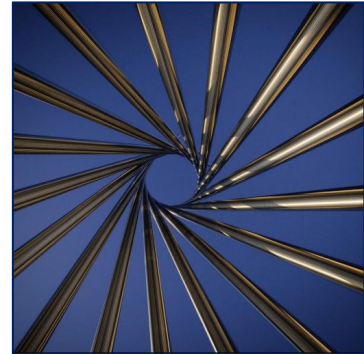


art 3

Kyle J. Knoepfel

Material for FIFE production workshop

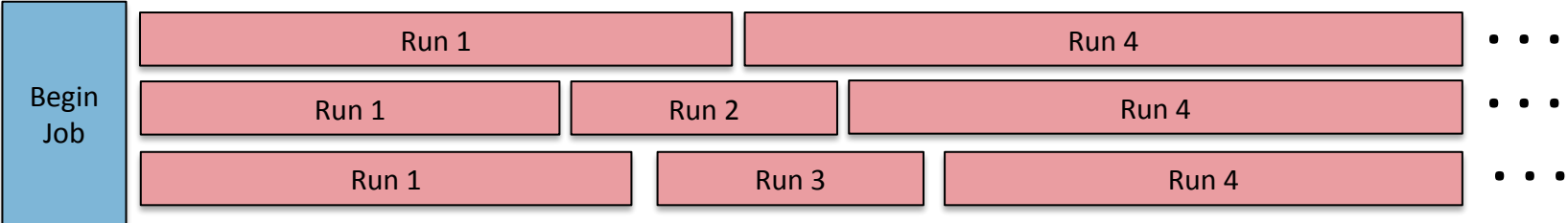
14 June 2018



art's multi-threading design

- Largely based off of CMSSW design
 - We use Intel's Threading Building Blocks (TBB)
 - Steps to be performed are factorized into *tasks*
 - Users specify the number of concurrent event loops and (optionally) the maximum number of threads that the process can use
 - Users do not explicitly create threads
 - Users **can** call TBB-provided parallel algorithms in their own code

Our goal:



Multi-threaded event-processing

- *art* 3 supports concurrent processing of events (not runs or subruns).
 - The number of events to process concurrently is specified by the **number of schedules**
 - The user can optionally specify the number of threads.
- The user ***opts in*** to concurrent processing.

(nSch, nThr)	Command
(1, 1)	<code>art -c <config> ...</code>
(1, 1)	<code>art -c <config> -j 1 ...</code>
(4, 4)	<code>art -c <config> -j 4 ...</code>
(nproc, nproc)	<code>art -c <config> -j 0 ...</code>
(1, 4)	<code>art -c <config> --nschedules 1 --nthreads 4 ...</code>

- In a grid environment, number of threads is limited to the number of CPUs configured for the HTCondor slot (*art* adjusts the number of threads).

Kinds of modules in *art* 3

- *art* guarantees that any currently-existing modules (to within some interface changes) are usable in a multi-threaded execution of *art*.
 - No multi-threading benefits will be realized with such “legacy” modules
- To take advantage of *art*’s multi-threading capabilities, users need to choose the kind of module they use:
 - **Shared module:** sees all events—calls can be serialized or asynchronous.
 - **Replicated module:** for a configured module, one copy of that module is created per schedule—each module copy sees one event at a time. Use if moving to a fully concurrent, shared module is not feasible.

General usage and guidelines

- If using an HTCondor batch system, *art 3* will ensure that its task scheduler will not use more threads than the number of CPUs requested for the job.
 - Extra steps are necessary for non-HTCondor systems.
- *art* services must be thread-safe
 - Notions such as the “current” event or module no longer have meaning in *art 3*.
 - Using TFileService requires extra serialization; *art* has instructions for how to do that.
- The *art* developers continue to work on documentation.

Guidance moving to *art* 3

- **Solve workflow issues first.**
 - You might have thread-safe modules and services.
 - If you're relying on illegal path configurations, you'll run into product dependency errors.

Recompile/rerun jobs with 1 schedule/1 thread
(default)

Add consumes statements to modules
(use -M program option for help)

Recompile/rerun jobs with 1 schedule/1 thread
and use --errorOnMissingConsumes

Recompile/rerun jobs with more than 1
schedule/1 thread

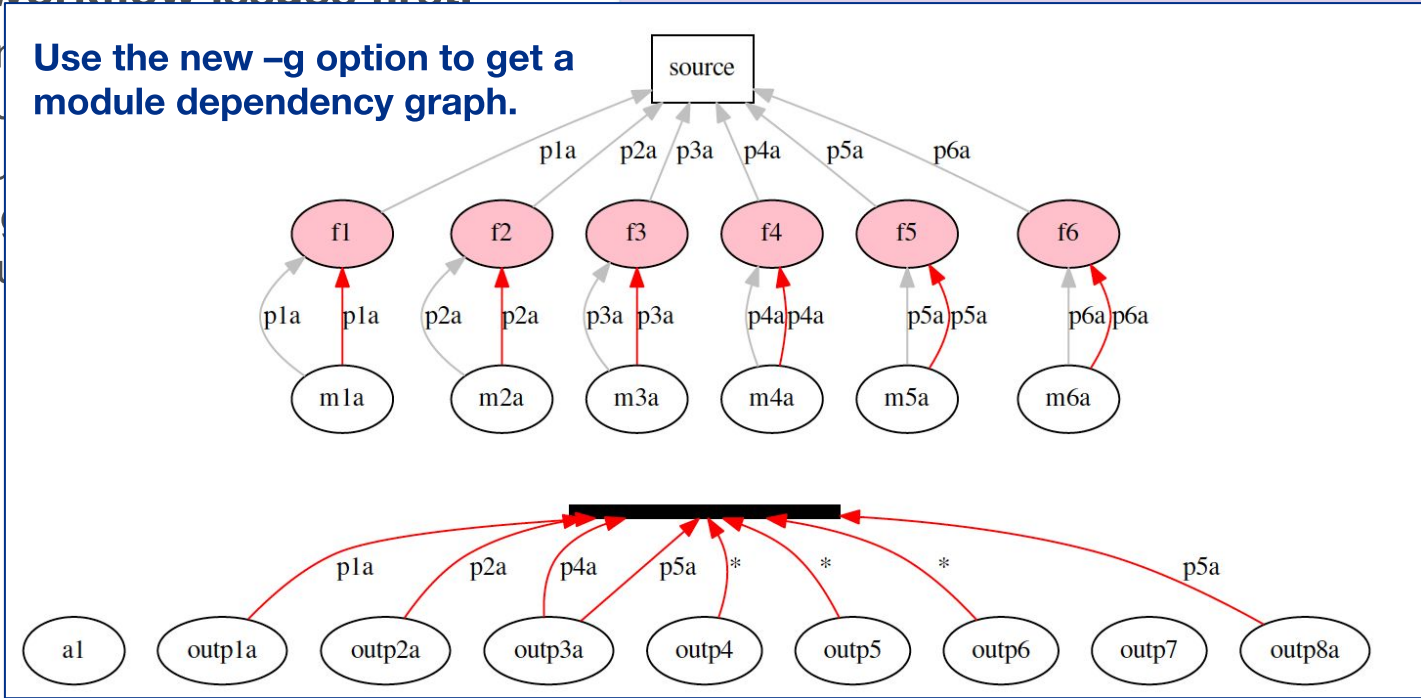
Guidance moving to *art 3*

Recompile/rerun jobs with 1 schedule/1 thread (default)

- **Solve workflow issues first.**

- You need to modify the module configuration
- If you have a configuration product

Use the new `-g` option to get a module dependency graph.



Guidance moving to *art* 3

- **Solve workflow issues first.**
 - You might have thread-safe modules and services.
 - If you're relying on illegal path configurations, you'll run into product dependency errors.
- **Determine what kind of module you need.**
 - Producer, filter, or analyzer?
 - Do you need to create (Sub)Run products?
 - Do you need to see every event?
 - Do you need to call an external library that is not thread-safe?
 - Do you have mutable data members for which operations are not thread-safe?

Guidance moving to *art 3*

- **Solve workflow issues first.**
 - You might have thread-safe modules and services.
 - If you're relying on illegal path configurations, you'll run into product dependency errors.
- **Determine what kind of module you need.**
 - Producer, filter, or analyzer?
 - Do you need to create (Sub)Run products?
 - Do you need to see every event?
 - Do you need to call an external library that is not thread-safe?
 - Do you have mutable data members for which operations are not thread-safe?
- We are working to provide guidance in dealing with such issues.
- Will work with Mu2e to demonstrate scalability.
- The SciSoft team's direct involvement with LArSoft means that we will help LArSoft as well.
- **If interested in moving to *art 3*, contact artists@fnal.gov.**



Online Computing Support

Bonnie King, Glenn Cooper

FIFE Roadmap Workshop

June 14, 2018

- Scientific Linux Systems Engineering (SLSE) Services provide online computing support from consulting to test stand/integration environments through production
 - SLSE handles operations at the OS/hardware level
 - Standardized architecture simplifies planning and implements best practices, freeing collaborators to spend more time on physics
- SLSE services should be requested early at SPPM/SCPMT to help us plan for capacity and provide engineering input