

Developer Meetings

- - April 2019
 - April 12th 2019
 - March 2019
 - March 29th 2019
 - March 1st 2019
 - February 2019
 - February 22nd 2019
 - February 1st 2019
 - January 2019
 - January 25th 2019
 - January 18th 2019
 - December 2018
 - December 13rd, 2018
 - November 2018
 - Nov 29th, 2018
 - Nov 1st, 2018
 - October 2018
 - Oct 12th, 2018
 - Oct 5th, 2018
 - September 2018
 - September 28th, 2018
 - September 21st, 2018
 - August 2018
 - August 24th, 2018
 - August 17th, 2018
 - August 3, 2018
 - July 2018
 - July 27th, 2018
 - June 2018
 - June 29th, 2018
 - June 22nd, 2018
 - June 8th, 2018
 - May 2018
 - May 4th, 2018
 - April 2018
 - April 6th, 2018
 - March 2018
 - March 30th, 2018
 - March 16th, 2018
 - February 2018
 - February 23th, 2018
 - February 9th, 2018
 - January 2018
 - January 12nd, 2018
 - December 2017
 - December 1st, 2017
 - November 2017
 - November 10th, 2017
 - November 3rd, 2017
 - September 2017
 - September 15th, 2017
 - September 8th, 2017
 - August 2017
 - August 25th, 2017
 - August 18th, 2017
 - August 11th, 2017
 - July 2017
 - July 21st, 2017
 - June 2017
 - June 23rd, 2017
 - June 9th, 2017
 - June 2nd, 2017
 - May 2017
 - May 12th, 2017
 - March 2017
 - March 17th, 2016
 - March 10th, 2016
 - March 3rd, 2016
 - February 2017
 - February 24th, 2016
 - February 17th, 2016
 - February 10th, 2016
 - February 3rd, 2016
 - January 2017
 - January 27th, 2016
 - January 13th, 2016

- October 2016
 - October 7th, 2016
- September 2016
 - September 16th, 2016
 - September 9th, 2016
 - September 2nd, 2016
- August 2016
 - August 12th, 2016
 - August 5th, 2016
- July 2016
 - July 15th, 2016
 - July 8th, 2016
 - July 1st, 2016
- May 2016
 - May 13th, 2016
 - May 6th, 2016
- April 2016
 - April 22nd, 2016
 - April 15th, 2016
 - April 1st, 2016
- March 2016
 - March 25th, 2016
 - March 18th, 2016
 - March 11th, 2016
 - March 4th, 2016
- February 2016
 - February 19th, 2016
 - February 12th, 2016
 - February 5th, 2016
- January 2016
 - January 28th, 2016
 - January 22nd, 2016
 - January 15th, 2016
 - January 8th, 2016
- December 2015
 - December 18th, 2015
 - December 11th, 2015
 - December 4th, 2015
- October 2015
 - October 23rd, 2015
 - October 16th, 2015
 - October 9th, 2015
 - October 2nd, 2015
- September 2015
 - September 25th, 2015
 - September 18th, 2015
 - September 11th, 2015
 - September 4th, 2015
- August 2015
 - August 28th, 2015
 - August 21st, 2015
 - August 14th, 2015
 - August 7th, 2015
- July 2015
 - July 31st, 2015
 - July 24th, 2015
 - July 17th, 2015
 - July 10th, 2015
- June 2015
 - June 12th, 2015
 - June 5th, 2015
 - May 2015
 - May 29th, 2015
 - May 8th, 2015
 - May 1st 2015
 - April 2015
 - April 24th 2015
 - April 17th 2015
 - April 10th 2015
 - April 3rd, 2015
 - March 2015
 - March 27th, 2014
 - March 20th, 2015
 - March 13th, 2015
 - March 6th, 2015
 - February 2015
 - Feb 20th, 2015
 - January 2015
 - Jan 14th, 2015

- Jan 7th, 2015
- December 2014
 - Dec 8th, 2014
 - December 3rd, 2014
- November 2014
 - November 12th, 2014
 - November 5th, 2014
- October 2014
 - October 29th, 2014
- September 2014
 - September 17th, 2014
- August 2014
 - August 25th, 2014
 - August 18th, 2014
 - August 11, 2014
 - August 4th, 2014
- June 2014
 - June 30th, 2014
 - June 9th, 2014
- May 2014
 - May 12th, 2014
 - May 5th, 2014
- April 2014
- April 21st, 2014
 - April 14th, 2014
- March 2014
 - March 31st, 2014
 - March 24th, 2014
 - March 17th, 2014
 - March 10th, 2014
 - March 3rd, 2014
- Feb 2014
 - February 24th, 2014
 - February 10th, 2014
- December 2013
 - December 16th, 2013
 - December 3rd, 2013
- November 2013
 - November 11th, 2013
 - November 4th, 2013
- October 2013
 - October 1st, 2013
- September 2013
 - September 23rd, 2013
 - September 9th, 2013
 - August 26th, 2013
 - August 19th, 2013
 - August 12th, 2013
- July 2013
 - July 29th, 2013
 - July 8th, 2013
 - July 1st, 2013
- June 2013
 - June 24th, 2013
 - June 17th, 2013
 - June 10th, 2013
 - June 3rd, 2013
- May 2013
 - May 20th, 2013
- April 2013
 - April 22nd, 2013
 - April 8th, 2013
 - April 1st, 2013
- March 2013
 - March 25th, 2013
 - March 18th, 2013
 - March 4th, 2013
- February 2013
 - February 11th, 2013
 - February 4th, 2013

April 2019

April 12th 2019

- Pegasus 5.0

- Site Catalog Conversion to YAML
 - mukund is mainly done
 - pushed out his changes
 - trying to make the tests green
 - Checkpointing changes to accomodate LIGO use of vanilla universe
 - Karan and Mats will explore and see if it is possible
 - cumulative stdout/stderr
 - what about time and duration values
 - since there is no DAG Node retry and job just goes on HELD state
 - Composite Events
 - Kibana dashboard needs to be updated
 - dropping ___ in the event names
 - George wants the AMQP library updated
 - Will create a JIRA item
 - Office Hours video
 - Karan will work Jasmine to upload the video
- Papers
 - RACE Paper submitted last week
 - PEARC Paper this week
- Proposals
 - Army Research
 - enabling in-situ supports for ExaScale
 - linked with what Tu is doing
 - SCEC Proposal Submitted
 - have a good chance
 - Exascale one with Michigan
 - the call will come out soon
 - Ewa , Rafael and Deborah
 - NSF GCR Proposal
 - Modelling wild fires
 - Has PRICE school input and also Deborah Post DOC
- EScience
 - Pegasus Tutorial Proposal
 - May 6, 2019: Tutorial Proposal Deadline
 - Also trying for the workflow comparison paper
 - Dynamo paper by George
- Pegasus connect discussion
 - tabled it for later when Mats is around
- HTCondor Week
 - Karan will be doing a Pegasus talk and Pegasus workshop
- Pegasus OLCF Poster
 - combine the panda poster
 - can also submit to EScience
- Ryan's work
 - Loic is moving pachyderm setup to AWS
- Loic Rafael and Tu are working on a paper for Cluster
- Software X

March 2019

March 29th 2019

- 4.9.1 Release
 - done and working on 4.9.2
- Site Catalog Conversion to YAML
 - mukund working on it
 - i still need to look at the bamboo tests
 - bamboo failing on mount scratch thing that condor thing
 - we have to fix in pegasus also. to fail on credentials in /tmp
 - check and do condor_config_val on the key and check if /tmp is in there
 - mainly affects all the users that use x509
 - LIGO has also tripped over it . Both with Pegasus and without Pegasus
- Condor vanilla checkpointing
 - karan asked him about what he is trying to do
- composite events
 - check for keys with same values
 - also do we need to pad extra keys for all events?
- Extensions to Jupyter Integration
- Pegasus Connect
 - will discuss on whiteboard on April 12th

March 1st 2019

- 4.9.1 Release
 - moving it to early next week
 - Pending Issues
 - <https://jira.isi.edu/projects/PM/versions/11891>
 - Execution environment for titan
 - service dependencies
 - PyOpen SSL
 - [Rajiv Mayani](#) please look at that and the flask dependencies
 - HPSS transfer client incorporation
 - Set the transfers to do remotely
- Office Hours
 - On Friday March 22nd on real time monitoring
- transformation catalog for 5.0
 - Mukund will work on it next
- EScience?
 - Paper
- pegasus-exitcode test
 - success message not parsed correctly
- Programmer
 - will interview the

February 2019

February 22nd 2019

- 4.9.1 Release
 - Pending Issues
 - <https://jira.isi.edu/projects/PM/versions/11891>
 - This raises the larger issue of how long we want to support externals packages

there are some packages we need to ship because of worker packages dependencies.

Consensus:
We remove mysql python externals package for 4.9.1 and 5.0.0

And also remove the dependencies from our deb and RPM builds.

 - Transfers within containers
 - We are only going to transfer from within the container till people complain
 - [George Papadimitriou](#) will add to the documentation.
 - non ascii encoding in the stdout
 - Support HPSS storage
 - The tools we use are htar and hsi

<https://docs.nersc.gov/filesystems/archive/>
- Office Hours
 - George on real time monitoring.
 - Date?
- EScience?
 - Paper
 - Tutorial submission

February 1st 2019

- 4.9.1 Release
 - ascii encoding breaks while parsing for monitoring events. monitors should have the population working and have log a warning error.
 - but we should ensure that stdout in database still gets populated
 - Karan will fix this
- New TC Format
- Shifter Support in Pegasus
 - is in 4.9 branch
- Pegasus Annual Report
 - will be working on it in coming weeks
 - will ask for input
 - next year report will be tricky . in terms of effort allocation.

January 2019

January 25th 2019

- 4.9.1 Release
 - ascii encoding breaks while parsing for monitoring events. monitors should have the population working and have log a warning error.
 - but we should ensure that stdout in database still gets populated

- YAML format for the TC
 - the line numbers should be mentioned in the errors
- GitHub commits don't trigger bamboo builds right now
 - move to webhooks?
 - slack token in bamboo.yml .
 - mats will look into it further
- SCEC for HPC Transfer certificate issue
 - Globus online certificates messed up hpc-transfer issue.
- Data Storage at NERSC
 - almost full
- Singularity container with the entry point.
 - docker singularity container conversion does not add the entry point.

January 18th 2019

- 4.9.1
 - container execution
 - data transfers happen within the container
 - python3 issue
 - vague rules to discover what python to use
 - Singularity HUB URL's updated
 - Documentation and tutorials need to be updated
 - montage examples
 - python stuff: create JIRA item
 - LIGO pull requests
 - Build pull request
 - PAM module
 - subprocess package thing
 - also related to Python3 movement
- Transformation Catalog Implementation
- Astro Py
- Shifter support at NERSC
- Panda Integration
- CENON NT
 - Rusio data pull in
 - fetching data might be easier
- Journal Paper
 - need to write something about containers

December 2018

December 13rd, 2018

- Pegasus 4.9.1 release
 - local site catalog entry creation
 - based on the pegasus version on the submit host
 - encoding issue in the stdout.
- Pegasus 5.0 Release
 - TC yaml implementation
 - mukund will create a yaml schema compatible with the TC
 - backwards compatibly
 - case by case basis
 - definitely for
 - catalogs
 - dax
 - pegasus-transfer
- SWIP Paper
 - we are in good shape
- Titan
 - under the PBS batch gahp.
- ZTF
 - the pipeline is based on docker-compose
 - peter will visit ISI with postdoc Danny in January
- Tutorial at TACC
 - karan has updated pegasus-init to work on wrangler
 - will update the tutorial notes accordingly
- OLCF accounts
 - make sure they work
 - get karan and mats can login

November 2018

Nov 29th, 2018

- Ryan
 - working on comparison paper with george on workflow systems
 - mats, karan shared neon meeting notes with Ryan
- Pegasus 4.9.1 release
 - Due for december end
 - potential issue in monitord in reference to hierarchal organization of submit directories
 - pegasus-submitdir
- ADASS Paper
 - due tomorrow
 - need to add information about sample run
- SWIP paper
 - mats and karan will work on it tomorrow afternoon.
 - cull out sections
 - add information about updated monitoring in 4.9
- OLCF Kubernetes
 - Condor is installed and configured as root
 - George tried condor log directory to lustre as condor in container has to run as user not as root
 - LOG_DIR should be /tmp
 - volumes can be attached to container to contain workflows etc
- Dynamo
 - Do dynamic scheduling
 - George thinking of using flocking
 - similar to what is done in OSG
 - non-sharedfs deployments should work

Nov 1st, 2018

- Pegasus 4.9.0 and 4.8.5 Released
 - We released it this week.
- Pegasus Business Card
 - Advocate for job postings.
 - Postdoc options
 - Programmers
 - pegasus.isi.edu/jobs
 - We should take to conferences with us
- Pegasus JAVA 8 dependence in RPM
 - there is a disconnect between RPM and common.sh
- ADASS
 - Karan working on a wpipe demo example
- New Student
 - Mukund
- Duncan started using 4.9.0 and has updated pyCBC to use singularity
 - changed our container execution model
 - all transfers done within the container now.

October 2018

Oct 12th, 2018

- Rescheduling meetings
 - New time is Thursdays 2PM starting from last week of October
- DAX API reporting
 - Perl DAX API - Rajiv
- Atlas visit
 - Wednesday we have Scientific Computing Seminar
 - Will involve writing a Pegasus code generator
 - Panda is second biggest after Condor on OSG
 - Thursday
 - Karan and George will be there.
 - Mats might be available remotely
- 4.9.0 Release
 - Mats preference is to skip the beta tag
 - Aim for the full release
 - Documentation freeze on Oct 26th
 - Try and do the builds over the weekend
- Duncan container usecase
 - cvmfs hosted container images
- Demo repository
 - panorama data and some runs from exogeni / nersc
 - Mats has two new elastic search VM's and are part of Elastic Search cluster
 - these vm's data is backed up also

Oct 5th, 2018

- Rescheduling meetings
 - Either Tuesday or Thursdays
 - Karan will circulate a doodle poll

September 2018

September 28th, 2018

- Rescheduling meetings
 - Either Tuesday or Thursdays
 - Karan will circulate a doodle poll
- Pegasus 4.9.0 Release
 - transformation selection issue
 - karan has not been able to recreate it yet.
 - will look into it more today
 - docker singularity pulls
 - container symlink
 - deprecate api's
 - modify DAX generators to indicate version/ DAX API used.
 - will look into ways on how to do it
 - one way is workflow metadata attributes
 - second is attribute to ADAG object.
 - rajiv will check how it gets stored in the metrics server
- ADASS
 - will try and do a poster with Mike at ADASS
 - deadline is Oct 8th

September 21st, 2018

- Rescheduling meetings
 - Either Tuesday or Thursdays
- Pegasus 4.9 release
 - integrity error reporting
 - pegasus-statistics reporting information about integrity errors
 - the unicorn dashboard for internal swip purposes
 - errors are appearing in the stream
 - more brainstorming required. the data is there
 - not clear whether to use grafana or kibana
 - does not have drill down functionality
 - mix of production and test workflows
 - create different queues in AMQP exchanges
 - container mount point support
 - karan is close to have that being implemented
 - transferring outputs to multiple location
 - lets say one for portal and the other for
 - list of output sites
 - good feature to add for 4.9.1
 - update --output-site option to pegasus-plan
 - pull docker images for singularity runs
 - we should do for 4.9.0
 - planner needs to tell pegasus-transfer an extra attribute.
 - add a type attribute
 - Papers
 - Github private papers repo
 - Deprecate stuff
 - perl api
 - old catalog formats
 - pegasus-plots
 - Hiring
 - <https://docs.google.com/document/d/183dkYBZ6Cpp6aYDB5MGW8Dj1DbzR11UnLPYdW7luZIA/edit>

August 2018

August 24th, 2018

- Pegasus 4.8.4 Release
 - when are we releasing?
 - next week before mats go on vacation
- error tagging

- update stampede schema to add a table called tags
- will allow us to capture number of integrity errors

August 17th, 2018

- Pegasus 4.8.4 Release
 - RPM fix ?
 - mats will manually verify
 - Karan should follow up with Stuart
- AMQP filtering
 - we are working on having filtering in built into monitor
 - nepomunk already has 33 errors identified
 - we need to db connection, pegasus-db-admin and other tools to pass properties with pegasus property prefix stripped off
- SWIP Paper
 - one reject seems to be harsh
 - we can try for HPDC also

August 3, 2018

- Pegasus 4.8.3 Release
 - singularity fix
 - <https://pegasus.isi.edu/tutorial/isi/index.php>
 - mats talked to adam at nebraska about containers.
 - the main doc book will not be updated for 4.9
- SLURM
 - Design Safe / TACC on Wrangler headnode
 - Nextflow has integration with SLURM and everything can be installed in user space
- PMC unit tests are broken
 - lets fix the tests
- Pegasus 4.9 release
 - more real life runs
 - nepomunk against ceph-s3 from one of uchicago machines
 - we need to get stats reported for integrity errors
 - larger issue of error classification
- ADASS Tutorial
 - we got into second round
 - add on exercise to run montage in the end.
- LIGO
 - Bruce group at AEI Hannover has left LSC
- Infrastructure
 - HipChat mess
 - should we move to ISI Slack
 - Public Chat feature
 - Some clients for Hipchat
 - Get a free channel from Slack
 - for all Hipchat rooms
 - what about ISI slack??
 - Github removal of old integrations
 - moving email notifications . [Rafael Ferreira Da Silva](#) will take care of it
 - we need to explore
- MINT Meeting
 - went well overall
 - issue of scoping .

July 2018

July 27th, 2018

- Pegasus 4.8.3 Release
 - VM Tutorial
 - will update pegasus-init requirements to get it working
 - main tutorial chapter will be updated for 4.9
 - because then tutorial based container may not work
 - change how docker scripts set environment
 - SCEC database loading error
- Failing Tests
 - Issue in updates to the dashboard database
- Panorama Paper
 - agreed on a re-organization

June 2018

June 29th, 2018

- Pegasus
 - 4.8.3 needs to be released because of singularity launching options
 - will wait till tutorial is updated.
 - karan will update pegasus-init with population modeling or povray option
 - 4.9
 - pegasus-statistics updated with integrity metrics
 - how to flag job errors because of integrity
 - need to figure out logic
 - value add proposition
 - maybe we should value type in the pegasus lite
 - need to implement the integrity dial
 - Start creating default local site entries to execute without local site
- ADASS Tutorial
 - Will submit today
 - Google doc shared

June 22nd, 2018

- Pegasus
 - SWIP paper submitted to escience
 - 4.8 montage tests failing
 - changes for integrity metrics in pegasus-transfer
 - updated monitord to parse events from various sources like pegasus lite output
 - mats pointed out to a bug in monitord
- LIGO
 - pip for python source package
 - update dependencies for latest packages , like pyopen ssl
 - install in the pip repository
 - pegasus-analyzer
 - interested in swip and containers.
- SCEC CSEP
 - will use containers
 - run on Comet
- 1000 genome workflow or use chimerica workflow
- ADASS Tutorial
 - montage ?
 - probably pycbc is also submitting a proposal

June 8th, 2018

- Scott Replica Catalog issue
 - Replica Catalog deletes take a long time
- Bamboo
 - bamboo emails are no longer received. so we dont come to know about workflow plan failures
- SWIP
 - monitord integrity changes. population of data from ks records working now.
 - we still need to populate data from pegasus lite records and pegasus-transfer
 - pegasus-statistics need to be updated
 - 0.1% overhead on production osg gem workflow
- Pegasus deployment at ORNL
 - we should be doing it similar to hpc-pegasus
- Pegasus Office Hours
 - next one in August
 - travels in July

May 2018

May 4th, 2018

- Pegasus 4.8.2 Release done on May 3rd
- we should consider separate user data to a separate file on pegasus-wms
- si2 meeting updates
 - some potential new users
 - ewa slides were a good overview summary
 - integrity data schema changes.
 - monitord changes need thinking

April 2018

April 6th, 2018

- Pegasus 4.8.2 Release
 - PMC bugs
 - tutorial for usc hpc
 - no longer allow + or . in the names
- Pegasus Report
 - Submitted for Ewa' review
- SWIP test run
 - discovered integrity errors in the wild
 - at colorado and university of nebraska
 - we would have not caught it before
 - e-science paper

March 2018

March 30th, 2018

- SWIP
 - pegasus-run issue, with wf restarting from scratch
 - because dagman rescue file is not there.
 - so should we update pegasus-run to look at the dagman.out file
 - so far we think it should be kept consistent with normal dagman behavior
 - to be discussed at condor week
 - mats created a Jira item for swip related statistics
 - <https://jira.isi.edu/browse/PM-1260>
 - will involve a database schema.
 - Things remaining
 - Dials to be implemented
 - stampede changes
 - pegasus-transfer changes???
- SC Tutorial Submission (April 16th)
 - <https://sc18.supercomputing.org/submit/tutorials-submissions/>
 - We should try and add exercises for containers
 - We will try for half day
 - 45 minute introduction
 - Feedback from Arizona Container Camp
 - There is interest.
 - coming up with an existing application that people understand or can relate to
 - montage - complex dax generator
 - rosetta
 - only works in nonsharedfs stuff
 - with
 - machine learning example?
 - with tensor flow?
 - requires container
 - NVIDIA has a lot of examples about machine learning
 - has to be multistep
 - and at least bag of tasks
 - Ashwin is doing some tensor flow stuff
 - on workflow.isi.edu
 - is working out of jupyter notebook
 - Genome sequencing workflows??
 - use Broad GATK sequencing workflow to use
 - SOYKB and IRRI use GATK
 - and are huge communities
 - <http://biocontainers.pro/docs/101/running-example/>
- Pegasus Report
 - we should be resolve Jira items as we fix them
 - will be also doing cumulative statistics
- Pegasus Office Hours
 - Jupyter Notebooks
 - will update the example to use namd example used for Oakridge
- Panorama Stuff
 - our multiplexing part in monitord done so far
 - however we are relying on amqp queues and routing keys for filtering
 - darshan data population
 - we need to invoke a script (pegasus-darshan) that will be invoked in the namd wrapper script, to pull the data from darshan logs on the file system and generate an ASCII output
 - Panorama.isi.edu VM
 - AMQP
 - Logstash
 - Kibana
 - Elastic Search
 - Make it do a backup every so often.

- Warns against doing it as a permanent datastore
 - Rajiv will verify
 - Influx
- Backups
 - CRASH PLAN backup for the /srv and /opt in the panorama VM
- LIGO Database locked issues
 - we need to look into the locking issues by tinkering with monitord flush intervals

March 16th, 2018

- SWIP
 - Most of the SWIP stuff is done as far as planner changes and getting the workflows running
 - we are in a position to share something
 - To do
 - sharedfs
 - Dial implementation
 - Update monitoring
 - Paper submission for EScience
- Pegasus Reports
 - new applications to attribute to pegasus grants
 - all the mike wangs work will go here
 - SCEC
 - LIGO - need to ping Duncan
- Panorama/ Pegasus workflow endpoints
 - We seems to be going towards AMQP
 - How is AMQP going to be configured
 - So far we have
 - `amqp://[USERNAME:PASSWORD@]amqp.isi.edu[:port]/<exchange_name>`
Online monitoring in kickstart
 - `amqp://[USERNAME:PASSWORD@]amqp.isi.edu[:port]/<virtualhost>/<exchange_name>`
 - Virtual Hosts
 - right now virtual host is hardcoded in monitord code. we set it to pegasus
 - global - across workflows
 - Exchanges
 - should be global across workflows
 - type direct - in panorama
 - we want them to be type -> topic instead
 - Queue
 - in panorama different queues for each workflows
 - Routing Keys
 - the routing key should be based on stampede event names
 - Events populated
 - https://pegasus.isi.edu/documentation/stampede_wf_events.php
 - We should add periodic events about states of workflows
 - SWIP integrity error events will be populated by clients

February 2018

February 23th, 2018

Eliminate support for Py2.6?

Python Dependencies

All - future

pegasus-service - [Flask](#), [SQLAlchemy](#), [Flask-SQLAlchemy](#), [Flask-Cache](#), [pam](#), [plex](#), [pyOpenSSL](#), [orderdict](#)

pegasus-monitor - [SQLAlchemy](#)

pegasus-analyzer - [SQLAlchemy](#)

pegasus-s3 - [boto](#)

pegasus-globus-* - [globus-sdk](#)

pegasus-init - [jinja2](#)

pegasus-metadata - [argparse](#)

pegasus-em - [requests](#)

PostgreSQL - [psycopg2](#)

MySQL - [MySQL-Python](#) OR [mysqlclient](#)

Note: Packages in green are available from yum.

February 9th, 2018

- SWIP
 - checksum computation will be implemented in pegasus-transfer.
 - allows us to handle the case where the input files don't have checksums in the RC
 - integrity checks are disabled now for files that dont have checksums in the RC
 - dial knob
- Tests
 - seem to be slow
 - bamboo could be moved to the new server
 - storage constraint test
- Lizard FS
 - Mats will give an update next time around
- Servers
 - Trying to do two server
 - IF we buy one server
 - Buy a storage server. That is Mats preference.
 - SoyKB workflow has
 - Compute
 - we will get a compute server first.
 - We should figure out the server and put in the request soon, and done by Feb end
- LSST
 - Tom Glanzman?
 - We will touch base on Monday with Tom and Nersc folks
- Office Hours today
 - have a presentation on containers
 - will upload on the website

January 2018

January 12nd, 2018

- AWS Batch
 - seems to be running in karan's account.
 - update documentation about aws batch
- Pegasus 4.8.1 Release
 - upto Mats whether we should tag or not.
- Pegasus Office Hours
 - Rafael will look up a new name
 - Container Presentation
 - Talk about containers
 - Blue Jeans
 - Advertising avenues
 - XSEDE workflows list
 - OSG List

December 2017

December 1st, 2017

- AWS Batch
 - Client done. still have to figure out about stdout and stderr
 - maybe we should have batch push the files and control where the jobs go in
 - also maybe each file should go to it's own stdout stderr
- Metrics for SWIP
 - Stampede
 - Metrics Server
 - Elastic Search
- Rajiv working on changing the salt configuration
- Model Integration with Wings

November 2017

November 10th, 2017

- Pegasus
 - AWS Batch
 - checked in stuff
 - jars checked in aws sub directory in the jars folder. pegasus-config classpath is updated accordingly
 - Bamboo builds
 - change in how users are handled
 - rajiv and mats worked on changing the salt configuration for the various machines
 - the major part changed was how the users are handled
 - the bamboo user got messed up and uid's were mismatching on the filesystem
 - main group for people unix accounts should be pegasus for everybody
 - only project users will have access to VM's for a particular project
 - Stewie Rebuild
 - move off stewie. the main OS needs to be updated
 - panorama
 - Rafael and Geroge will create a VM for panorama
 - CENTOS 7
 - mats will help George create VM
 - Ashwin consumers from Influx DB
 - mysql server
 - Pegasus metrics server
 - JSON vs YAML
 - initial impressions seem to favor yaml
 - YAML does have benefit of including comments
 - also YAML , JSON will result in additional lines
 - templates for site catalogs
 - LSST
 - mats will update documentation for pyglidein
 - to work with condor pool passwords thing
 - also will take mike site catalog to update NERSC entries
 - tests
 - rosetta and montage appear working again. not clear what triggered errors in first place
- SC Next week
 - Rafael and Karan are away
- AWS workshop for LIGO
- George Panorama work
 - Dakota ends up launching multiple Pegasus workflows based on it's gradient functions
 - using ensemble manager to do multiple runs
 - George will check in dakota test case and example
 - pick one approach and update documentation
 - SWIP Demo
 - think about merging stuff from panorama back to production branch
- work with ian foster and raj kettimutt on globus online
 - do multi site run
- Tudo
 - working on insitu
 - data spaces approach to have staging area
 - tudo wrote sample applications
 - evaluating on CORI using shared memory
 - burst buffers cannot be used
- Ashwin
 - analyzes influx db data
 - using statistical learning
 - python panda library

November 3rd, 2017

- Pegasus 4.8.1 release
 - 3 bugs in worker package staging.
 - pegasus-transfer PYTHONHOME unset does not work
 - hierarchal workflow handling.
 - to be discussed tomorrow
- AWS Batch
 - need to check in changes.
 - need to add options for the client and do error checking.
 - still need to figure out how to integrate in pegasus

September 2017

September 15th, 2017

- Pegasus development
 - Dashboard

- LSST might want it running out of a directory other than \$HOME/pegasus
 - No plans to tackle it right now. requirements are vague. and catch 22 situation
 - Python problem with Pegasus install
 - DAX3 problem does not work.
 - Could not be recreated
 - PyPy account should be disabled
 - pypy has a 4.3 pegasus package
 - we should remove it
 - The jobname with dagman not allowing . is fixed
- LIGO
 - Heard from Duncan. Tried out metadata stuff
- Another person at NERSC that is interested in running Condor
- AWS Batch
 - done initial development.
 - how to retrieve logs etc.

September 8th, 2017

- Pegasus 4.8.0 Release
 - went out this week
 - documentation
 - pyglidein
 - out of icecube
 - mats added a section in the documentation
 - pretty neat once it is setup
 - and works really well on machines with two factor
 - not tuned for MPI things.
 - on the submit machine a web based python thing.
 - pegasus resource profiles will work out of the box with pyglidein
- Releases
 - Post 4.8 Releases
 - changes in the debain build
 - source package has been renamed. mats removed the source part
 - changed the versioning of RPM and debian. The dev series will have the timestamp in it.
 - pegasus-version -f also has timestamp
 - Will create a separate YUM and DEB developer repositories
 - repositories will not be signed.
 - Mats is still playing setup
 - Worked a lot on Debian packaging.
- HipChat will be upgraded to Stride
- Mats updated JIRA today
- Sim Center Workflows
 - Using Condor IO thing
 - for 4.8.1 we should look at the remap thing
- SWIP Poster
 - the first review is really good
- Docker and Singularity
 - have stuff about engineering challenges
 - But not enough usage
 - Practical Aspect
- Von's Group SWAMP thing.
 - pegasus is part of trustworthy software thing?
- AWS Batch
 - AWS batch thing works
- Investigate how Dakota and Pegasus can work together
 1. Run Dakota as a job
 2. Run Dakota on submission machine
 - a. dakota calls a script that does a pegasus workflow
 3. Mix of 1 and 2.

August 2017

August 25th, 2017

- Pegasus 4.8.0 Release
 - beta3 tagged
 - monitord replay issue for rc tables against mysql server
 - Jupyter thing
 - VM updated with Jupyter
 - Docker example application
 - R builds with pegasus
 - for time being only brew builds have that disabled.
 - Condor update to the brew installation.
- Pegasus 4.9 Roadmap
 - SWIP

- lay out the changes
 - prioritize stuff for production readiness
 - the knob for integrity.
 - get into transfers.
 - signing stuff on the backburner.
 - chaos monkey tests
 - metadata things
 - aws batch support
- Pegasus Tutorial
 - George felt that Pegasus tutorial was a bit too easy.
 - it should be maybe more interactive. get the user to develop a new workflow
- Tudu will pick up Decaf work
- Dataspaces
 - do data management
- Ashwin will work on deep learning on panorama
 - use tensor flow
- Dakota
 - ini file . runs simulation and converges simulation points
 - George will be working on it
 - has a checkpointing facility

August 18th, 2017

- mats found a new hydrology user in boulder
 - based at Boulder
 - there was a magpie presentation there.
 - mats did a hosted ce tutorial
- 4.8.0beta2 release
 - tagged and sent it out.
- monitor workflow and read permissions creation
 - should only when the database is created.
 - ~/pegasus directory should be 755
- dashboard errors
 - rajiv should traverse the directory in the dashboard.
- LSST
 - cleanup issue
 - mats and karan agree on it, that it is bad application
 - we should reply to it.
 - the wrapper should copy the file and launch the job
- source a setup a script for jobs
 - has to be generically done
- registration jobs shell expansion
 - we should not do getEnv=True
- testing repo
 - stuart from LIGO asked for it.
- BOSCO
 - we have the examples updated
- Karan will remind Eliu about LIGO and Bluewaters
- Slick Jupyter Demos
 - Started up VM's
- Jupyter tutorial
 - should be integrated into the VM

August 11th, 2017

- Bamboo is finally green
- we will do a Pegasus RC1. actually a beta since we still want to address some issues.
- Rajiv fixed the build with python crypto issues
 - pyopen-ssl was updated during 4.7.x series
 - we should package only things that we are not sensitive to the versions
 - so right now pyopenssl is removed from binary builds, and all associated dependencies were removed.
- New throttling things.
 - number of jobs scale with the size of the workflows.
- SCEC all hands meeting.
- Documentation
 - Took a stab at the containers.
 - Rafael has to add a separate jupyter chapter
 - Karan will update the throttling docs
- LSST
 - Mats and Karan had a call with Tom about designing a workflow for one of the production pipelines
 - Mats and Rafael had a call with the French cluster folks (Fredrique Sutter). Fredrique works for simgrid
- Paper
 - rvGAHP paper ready for submissions
- Suraj Poster
 - Mings pass really helped

July 2017

July 21st, 2017

- VMs are down, so tests are slow, and cannot test the new features yet
 - Mats will send an email (or call) Derek to check with the VMs issue
- Try to run the Montage container test on OSG
 - TODO: Reconfigure our poll (it is not flocked yet)
- Pegasus 4.8.0
 - Bugs on the container (transformation catalog) is fixed
 - Stage in/out nodes based on the number of computing jobs on the workflow
 - TODO: add warning for errors (size of jobs)
 - Warning for category is done
 - TODO: reference implementation of a workflow using docker (1000 Genome workflow - Rafael)
 - Jupyter: add container keyword for API

June 2017

June 23rd, 2017

- Pegasus 4.8.0
 - Decaf
 - local universe jobs does not honor request_cpus , and jobs remain idle if they ask for multiple cpu's
 - karan will update pegasus to remove the request_ parameters from the local universe jobs
 - Steven Clark
 - Pegasus build issue is related to python 3 compatibility in the DAX API
- LIGO
 - Eliu plans to run on Bluewaters
 - we should confirm that he only wants to run on bluewaters.
 - they have sucky performance of getting data to the compute nodes in bluewaters.
 - set the schedd start date
- NERSC
 - Karan will do a test setup there.
- Pegasus Builds
 - failed because of detain version upgrades to build tools
 - setup tools in python complains to pegasus 4.8.0-dev

June 9th, 2017

- Pegasus 4.7.5
 - pegasus-rc-client bug fix is done
 - 4.7.5 and 4.8.0 together
- Pegasus 4.8 release
 - docker stuff is complete
 - docker tests added are green
 - karan will work on singularity next week.
 - LIGO reports pegasus lite jobs filling up /tmp . karan will check with LIGO on whether there is any environment set?
 - rafael will update his api to make it consistent with the container format
 - also will add a bamboo example.
- DECAF integration
 - karan has an idea about it.

June 2nd, 2017

- Pegasus 4.7.5
 - pegasus-rc-client bug fix to be done
- Jupyter
 - rafael will be working on it during June
- For 4.8.0
 - container
 - docker works in nonsharedfs right now.
 - work on **singularity** support.
 - clustering . clustered jobs can only refer to one container
 - symlinks - for 4.8.0 they are disabled.
 - container sharedfs example
 - we have pegasus-lite with sharedfs. automatic translation of file URL's
 - transfer refiner
 - notification email updates
 - mats updated default notification scripts. will generate svg files

- at end of workflow generate notifications that have statistics
 - monitord needs to run the remaining notifications after the workflow is done.
- makeflow integration
 - limitations for pegasus generating make flow integration
 - makeflow model
 - all files have to be on the submit host
 - how do we translate auxiliary jobs to make flow description
 - tyson at arizona.
 - add new transfer jobs
 - add new credentials
 - no postscripts there
 - monitoring
 - won't work with monitoring
 - write a new monitord.
 - maybe do an opposite translation???
 - what will be useful is to integrate with using work queue with our own dagman manager.

May 2017

May 12th, 2017

- auto scaling of stage out and stage in jobs
 - 4.8 transfer refiner will be Cluster by default.
 - auto-computation of number of stage in, stage out and cleanup jobs
 - defaults should be computed based on number of jobs at a level.
 - use a ratio or step function .
 - come up ratio ranges for auto determination
 - 1:5 for numbers of jobs < 10K (20%)
 - 1:20 for number of jobs > 20k (5%)
 - will create a JIRA item for this
- container stuff
 - close to having one example running
 - have not figured clustering jobs out yet.
 - mats agrees with the approach now. pegasus lite invokes the docker run commands.
- integrity stuff
 - will make slides
 - be specific about we have done .
 - we give them an option of running synthetic stuff
 - For
 - also define best effort part.
 - strict, off, minimal , best effort
 - how do we handle case where SHA exists.
- WDL
 - workflow definition language
 - WDL is JSON based
 - has a template approach with variable substitution
- AWS Cleanup
 - need to delete snapshots and cleanup VM's

March 2017

March 17th, 2016

- monitord stdout and stderr missing
- the VARS one. just expose the variable.
- SCEC issue
 - job managers per resource
 - got fixed by one job manager per job
 - BOSCO works partly.
- containers call from yesterday
 - dsa
- metadata
 - metadata population in postscripts
 - move metadata population to the postscripts.

March 10th, 2016

- SCEC cleanup issue

- related to Jglobus not updating and enforcing the compliance for RFC 2818 compliance
 - <https://github.com/jglobus/JGlobus/issues/159>
- LSST visit update

March 3rd, 2016

- Pegasus 4.7.4 Release
 - sent out the release
 - we did a ligo fix yesterday to pegasus transfer
- mats osg gem
 - workflow did not finish
 - pegasus-exitcode has a shortcut for a regex
 - make it more strict. whether to trigger failure in pegasus-exitcode
 - revisit how metadata population
 - trigger failure for missing records.
- SCEC RC client issue
 - Rafael will look into it for pegasus-rc-client
- containers support
 - containers on a pause right now.
- Webinar
 - lets try and schedule one for april end
 - bluejeans will be an option
 - topic will be covered new features for 4.8.0

February 2017

February 24th, 2016

- Pegasus 4.7.4 Release
 - we will tag today.
 - there is a potential monitor bug that happens on sub workflow retires only in the live mode, that Karan is unable to trace
 - ds
- containers support
 - pegasus lite launches docker wrap
 - or the other way around. because worker package has to be installed in the container in some cases
 - so double install
 - Clustered jobs
 - we want at max one container to use the clustered job.
- monitor performance
 - on OSG connect there is a difference between 4.6 and 4.7 performance replay
- monitor.log has errors indicating unable to read .out .err files.
 - we think it is a race between DAGMan and the filesystem

February 17th, 2016

- Pegasus 4.7.4 Release
 - targeted for next week.
 - LIGO ran into a prescript issue
 - pegasus lite deleted the worker package in the workflow submit directory
 - only triggered when there was a subsequent compute job.
- new transformation catalog format
 - <https://jira.isi.edu/browse/PM-1165>
- containers
 - open issue whether docker wrapper launches pegasus lite
 - or the other way around

February 10th, 2016

- Pegasus 4.7.3 Release
 - SCEC has issue with pegasus-db-admin
 - mysqldump timesout when updating their replica catalog
 - Database TC
 - remove support for Database TC
- Stewie and fisheye upgrades
 - fisheye upgrade
 - Mats agreed to do the upgrade
 - stewie runs debian 7
 - we need to upgrade it one day or later.
 - runs GridFTP and mysql
 - RabbitMQ is running there
 - MongoDB is running there
 - Catalog dependencies on stewie

- 5K limit for a new server
- OSG All Hands Meeting
 - no tutorial looks like
 - lots of pegasus users coming there
- Containers Support
 - pegasus lite invokes the docker wrap.
 - singularity support will be required.
 - container modes
 - should we support docker definition file
 - do we build on the worker nodes?
 - pull in an existing docker image from the hub
 - on the staging site
 - whether we should unload an image or not
 - we should try and cleanup
 - credential renaming has to be worked out
 - Transformation Catalog
 - how to represent container dependency in the transformation catalog

February 3rd, 2016

- Pegasus 4.7.3 Release
 - we tag later today or first thing monday
 - waiting for scott to reply
- Jupiter Notebook
 - in general jupyter the interactive interface closes if you close the tab
 - in our case it does not affect us, since we invoke pegasus-plan at the server end
 - Vicky has a workflow out of panorama that she has in jupyter as a set of the instructions
- Containers
 - karan did some exploration of docker containers via HTCondor
 - by default docker in the container runs as root.
 - means output files are written out as root
 - also the containers need to be shipped around.

January 2017

January 27th, 2016

- Pegasus 4.7.3 Release
 - 4.7.3 release.
 - condor stable release has been released.
 - we will tag next friday one way or other
 - fix monitor replay mode
 - crosscheck with rajiv on dashboard
 - centralized mysql server for master workflow dashboard
 - LIGO wants to host a mysql server for master workflow databases
 - Mats will like to see something similar
 - also look at some publish subscribe options
- Rafael give an update on the container
 - docker universe
 - htcondor support i think is mainly geared towards startds
 - preinstall software in user containers
 - another model is to let pegasus figure out data and executables
 - rafael did stuff in pegasus lite stuff
 - will have to rewrite proxy and credential environment variables
 - also how is the environment is rewritten
 - good to have a generic concept of multi-level wrappers
 - need to have a pegasus-docker-wrapper or pegasus-container-wrapper to do launch docker or singularity
 - another container technology called shifter
 - <http://www.nersc.gov/users/software/using-shifter-and-docker/using-shifter-at-nersc/>
 - OSG uses singularity because it is more friendly to launch in user space.
 - OSG is launching user jobs using singularity. the image is determined by the VO.
 - lets target pegasus lite mode first
 - little bit of data passing.
- Rafael will have a student to take forward the docker swarm stuff
 - 8 hours every week

January 13th, 2016

- Pegasus 4.7.3 Release
 - sub workflows
 - better error message for pegasus-transfer when source files don't exist
 - pegasus-kickstart
 - improve error message
 - dashboard to better separate kickstart and pegasus lite messages

- Potential SCEC issued with RV-GAHP
- results of qualtrics user survey
- Pegasus 4.8
 - swip stuff for 4.8
 - have sent emails for their use cases

October 2016

October 7th, 2016

- Pegasus 4.7 Release
 - release notes and documentation is done
 - need to follow up with Action for our build VM's
 - LIGO is not going to test 4.7 release as they are in midst of a cluster upgrade.
 - Rafael will write a blogpost about R API after the 4.7 release
- Dashboard requests 4.7.1
 - rafael and rajiv will work on getting dashboard to display the database schema version and the pegasus version
 - useful, when a new version of pegasus is deployed and .
 - Unable to read the sqllite database
 - related to users permissions on the database
- from braindump in replay mode should be able to pick up relative paths.
- brew error on macos sierra
 - brew releases are built manually
 - after the release we have to update the formula to reflect latest stable version.
- ACME workflow on MIRA
 - GitHub page to be updated with list of dependent software
 - ACME team needs to help with installation of one of the software.

September 2016

September 16th, 2016

- Builds
 - disabling RHEL5, Debian 6, Ubuntu precise. Karan will make sure in the code it works
- Pegasus 4.7.0 Release
 - reached out to LIGO. hopefully they will start testing
 - rajiv checked in dashboard changes
 - karan to write documentation for directory layout
 - rafael will update pegasus-exitcode next week.
- Pegasus 4.8.0 release
 - one of the first things will be to update the SUBDAG keyword.
- LLNL account approved for Karan
- OLCF account waiting for notarized documents to be received
- SCEC
 - concurrency limits for transfer jobs
 - prime candidate for priority stuff that will allow good interleaving of transfer jobs with the compute jobs
 - ask Scott to see if 8.5.6 condor can be released.
- ACME workflow
 - HSI client for HPSS storage.
 - Karan will reply to Jamie.
- Bluewaters HTCondor install
 - Bluewaters renewed till 2019
- Pegasus HPCC workshop on September 30th
 - karan will be there.

September 9th, 2016

- Builds
 - disabling RHEL5, Debian 6, Ubuntu precise
- Pegasus Development
 - 4.6.2 released . LIGO has updated it.
 - LIGO tripped over changes to planner submit directory behavior
 - held job reasons are recorded in the database
 - 4.7.0 release
 - went through pending items
 - targeting end of the month for the release
- proposal
 - data aware workflow management
 - no BPEL only a reference for it.

September 2nd, 2016

- Pegasus Development
 - 4.6.2 released . LIGO has updated it.
 - pegasus.dir.storage.deep true throws an error right now.
 - 4.7.0 release
 - karan looked into the HELD job
 - rajiv thinks no dashboard change required.
 - pegasus-exitcode changes will be done by rafael
 - LIGO should install 4.7.0 on dev machine.
 - SCEC production run
 - Reverse GAHP OLCF
 - once tokens are reactivated , karan will check up on rhea vrgahp and get it running
 - HTCondor on bluewaters
 - Karan opened a ticket.
 - LLNL
 - security training to be done by Karan
 - panorama
 - rafael is working on panorama demo
 - two different pegasus workflows running on 2 exogeni slices
 - and data staging server in between. shadow q has to propagate transfer priorities
 - currently it is workflow level priority. will be manually assigned.
 - 1000 genome workflow -

August 2016

August 12th, 2016

- Pegasus Development
 - 4.6.2 release
 - release notes are checked
 - tutorial documentation will be updated to include the docker tutorial
 - pegasus service init script
 - we will not include it and enable by default in the builds
 - mats will update the item accordingly
 - 4.7.0 release
 - submit directory structure
 - we need to get the depth thing fixed . Karan need to make sure if the DAGMan knob can be set automatically.
 - we should have a way to have it set for deeper
 - documentation to be set
 - pegasus-exitcode to have wait lock thing to setup it's logs
 - one option is to log only exceptions initially.
- pegasus-keg to mimic IO pattern
 - read files over and over again.
 - this way we can increase IO without increasing file size (that results in higher data transfer costs)
- DECAF WMS
 - [10:30 AM] Rafael Ferreira da Silva: <https://bytebucket.org/tpeterka1/decalf/wiki/public-docs/peterka-decaf-handout.pdf?token=c693fc8bb177...>
 - [10:31 AM] Rafael Ferreira da Silva: <https://cfwebprod.sandia.gov/cfdocs/CompResearch/templates/insert/project.cfm?proj=134>

August 5th, 2016

- Pegasus development
 - waiting for LIGO to check the support for changes for OSG, where pegasuslite URLs are converted to file URL if the staging site and compute site are same
 - Karan has a call with Duncan next week planned.
 - staging sites deep directory structure
 - mats has it working for one of the workflow.
 - <https://jira.isi.edu/browse/PM-1049>
 - automatic delayed job retries
 - the real fix should be in DAGMan. Karan will follow up with Kent. Will address for 4.8
 - postscript output redirects
 - one file per job is what we had considered earlier
 - maybe we should do it per workflow log file.
- DIPA workflow development
 - good progress there.
- Titan Setup
 - we should consider setting up it the same way as bluewaters
- Next Pegasus proposal
 - next week meeting we should iterate on items.

- Samrat issue
 - get pegasus-exitcode to look for final output files
 - checked in workflows to the pegasus repository
 - bioconductor repository
 - would be good to setup PAGE cloud VM with the workflow.
- Deter Krans Mueller
 - director of supercomputing in germany
 - supermute supercomputing cluster
 - will send a student for 3 months to ISI end of the month.
- Rafael plans to practical comparison paper
 - Gui's docker stuff.
 - do a blogpost of montage with above docker stuff.

July 2016

July 15th, 2016

- Pegasus development
 - waiting for LIGO to check the support for changes for OSG, where pegasuslite URLs are converted to file URL if the staging site and compute site are same
 - staging sites deep directory structure
 - dashboard changes for nested submit directory structure
 - fixed the on demand loading for the dashboard.
 - identify workflows that will benefit
 - LIGO
 - Splinter
 - OSG - Kink
 - put in the test cases for testing it out.
 - use the new montage dax generator
 - pull the montage dax generator via squid cache.
- Release schedule
 - Get 4.6.2 out first.
 - 4.7 probably early august.
- ALCF Mira running.
 - cobalt workflow
 - ACME workflow compilation. Waiting on Ben for the source code.
- Panorama use case
 - SNS is not enough in terms of data sizes.
 - anirban will start working on it next week.
- R Examples
 - samrat working on a bioconductor example
 - has an example workflow
 - code should be checked into github
 - samrat is working on a more advanced workflow that will be put in the examples directory also
- Gui docker nodes work on amazon ec2
 - uses docker swarm and docker machine to do setup etc
 - workflows run in condor IO mode.
- DIPA Workflows
 - waisman folks will start working on it.
- free surfer workflow
 - mats does not think there is enough uptake.
 - suchandra is working on a second version that will add more capabilities
- seismology workflow
 - rafael will check in to the repo.

July 8th, 2016

- Pegasus development
 - waiting for LIGO to check the support for changes for OSG, where pegasuslite URLs are converted to file URL if the staging site and compute site are same
 - pegasuslite signal handling
 - mats updated it. LIGO reported cases, where jobs got killed before the outputs were staged back . But the jobs themselves were not marked as failures.
 - duncan's third issue could also be related to the signal handler
 - modify kickstart to compute md5 checksums.
 - we could potentially get kickstart to validate md5 checksums
 - have an architectural idea about it.
 - gridftp currently does not expose checksumming
 - irods client has checksumming in built.
 - pegasus-init R example
 - R example will not run on OSG because of module load issues
 - all R examples will have a wrapper for the scripts
 - 4.6.2 after changes are verified.
- DIPA Workflow
 - with Waisman brain imaging pipeline that runs on Waisman cluster

- Rafael is working on a seismology workflow
- tophat workflow paper got accepted in a bio journal
- Pegasus Virtual Summer School
 - would be similar to the XSEDE ones
 - will be 1.5 hours long.

July 1st, 2016

- Mats has moved bamboo to a new RHEL7 VM
 - migrated all the tests to it.
 - there were issues with CondorC tests that are resolved now. because of path issues
- pegasus-init R
 - Rafael will integrate Samrat's R example workflow
 - Samrat is also working on a bioconductor example workflow
- rajiv made minor dashboard query changes

May 2016

May 13th, 2016

- Pegasus development
 - kickstart wrappers
 - process explosion.
 - eventually we would want it to be in the workflow.
 - handle these wrappers as credentials in the workflow.
 - what are class of files that are always required.
 - KICKSTART_WRAPPER in kickstart
 - was done for the PAPI stuff originally.
 - pegasus-init for OSG
 - pegasus-init
 - R examples?
 - rafael will do it in june.
 - job held scenarios
 - open with htcondor admin .. a job should never goto the held state
 - maybe pegasus should do quick retry for small workflows
 - for large workflows retries should happen at a longer delay
 - for workflows less than 100 nodes held duration should be small, and failures maybe should be triggered earlier
 - not for large workflows
 - revisit whether clustered jobs should be based on size of the cluster or the number of jobs
 - mats no longer likes the idea of having fixed number of transfers
 - deep directory structure for the workflows
 - can splinter move to using them?
 - right now they are condor io
 - on the data side it deep directory structure will only work
 - BOSCO SSH
 - Mats tried with condor 8.5.4 on comet.

May 6th, 2016

- Pegasus development
 - moved the submit directory creation stuff to the mapper interface
 - reorganized the code for it.
 - on the execution site for nonsharedfs case we will enable for the dashboard
 - dashboard works mostly
 - only improvement is on the file browser side. will open a JIRA item for it
 - database changes
 - for 4.7 we will add extra columns to workflow state and job state tables.
 - the dashboard needs to show the better the task metadata better for 4.7
- pegasus tutorial for virtual summer school.
 - will be based on the XSEDE tutorial
 - bluewaters will setup a VM for the tutorial.
 - Scott will do an introduction and an overview.

April 2016

April 22nd, 2016

- Pegasus development
 - 4.6.1 released today
 - had to fix bugs for symlinking not being triggered for SCEC

- dashboard for the home page should work without trailing slash
 - all other pages should work the same way . For 4.7 we should do that
 - Pegasus R example
 - rafael will work on it
 - OSG and XSEDE site catalog examples
 - Submit Directory organization
 - Relative DAGMan paths
- HTCCondor week
 - Lauren said training week
- Bluewaters training
 - 2 day training might be too long
 - we will work on pegasus training module.

April 15th, 2016

- Pegasus development
 - 4.6.1 release next week
 - pegasus-status change for new Condor changes
 - cartoon will be upgraded to 8.5.x
 - pegasus-analyzer
 - will report correctly submit failures
 - better errors for mismatch in cores/ppn requirements
 - Tag and build on Thursday.
 - pegasus-s3
 - batched uploads and downloads
 - output directory options fails if local scratch not specified
- LIGO transfer issue
 - NFS reported write as successful for a transfer job.
 - wget reported data was transferred and wget succeeded good use case for checksumming of data
 - where do checksums come from
 - for data files good placeholder in the transformation catalog.
 - SCEC had similar issues where SGT's had gotten corrupted
 - that is why SCEC put a specific job in the workflow and uses ABORT DAG on feature
- Call with Kent for adding nodes to a running DAG
- group jobs with similar errors
 - might be a python library in there
- HTCCondor Week
 - proposed a hands on tutorial
- pegasus 4.7
 - ignore integrity constraints in monitor
 - only for duplicate keys

April 1st, 2016

- Pegasus development
- Submitted tutorial for XSEDE 16
 - will include RADICAL
 - might update tutorial with BOSCO. Mats already have BOSCO to run on Comet
- Derrick Lazaro wants to build a bigger filesystem (400 TB)
 - will be backed up
 - has a commercial storage vendor in mind
 - has backed up capabilities in built (block level backup)
 - let Mats know about storage needs
 - Mats estimated our storage needs to 25-50TB
- Graduate student coming to the group mid may to july. brazilian student. currently in Florida
- Ahmad group got a EPSCoR grant
- CRAFT Meeting update

March 2016

March 25th, 2016

- Pegasus development
 - Gideon has been working on kickstart online monitoring for panorama.
 - the lib interpose monitoring requires app code to be dynamically linked to use LD_PRELOAD
 - now kickstart has a new mode, where monitoring thread will scan the proc filesystem for all processes in resource group.
 - this approach disables the PAPI counters as they need to be retrieved from app itself
 - also is working on aggregation logic
 - complicated accounting information
 - added another process called pegasus-monitor . so it is usually pegasus-kickstart-> pegasus-monitor -> application
 - can deploy without any external dependencies.

- 4.6.1 release
 - in april when karan comes back from PAGE meeting
- Condor bug on schedd evicting dagman jobs
 - LIGO noticed on other submit nodes
- mats worked with Derrick to make sure glideins work with BOSCO on comet
 - CyVerse Talk - Mats will do a hands on thing with them. Mats may do an existing tutorial.
 - raphael used the new slides.
- Pegasus workshop
 - erin will get back to us with other feedback.
 - make the intro slides more simpler.

March 18th, 2016

- Pegasus development
 - deep submit directory structure working for submit directory on PM-833 branch. however need to move to relative directory paths in the .dag file , before merging back to master
 - gideon is reworking how kickstart online monitoring work
 - working on kickstart monitor that goes through the /proc/ filesystem with the assumption all apps installed via kickstart have the same process group as pegasus-kickstart
 - pegasus workshop on campus on tuesday. it is setup <https://pegasus.isi.edu/tutorial/usc/>
 - the tutorial is setup using pegasus-init
 - will ask mats to move the XSEDE tutorial to pegasus-init
- raphael working on energy paper again
- stephan paper to HPDC got accepted

March 11th, 2016

- Pegasus development
 - R DAX API is done
 - will be proposing for CGSMD
 - Deep hierarchy structure
- LIGO meeting
 - do a local file copy against the staging site
 - having a separate staging site bogs down inter site transfers
 - metadata
 - they are interested. want monitord to transfer the stampede database to another location from the scratch submit directories
 - cannot really do it in monitord
 - can also potentially do it in pegasus-dagman
 - argument passing for sub workflows
 - will be done 4.6.1
 - jobs that work on output site directory.
 - credentials issue
 - variable substitution
 - will make use of it
 - submit directory and other directory organizations
 - are interested in using it
-
- Rosa
 - wants to do something with pegasus
- Monitord

March 4th, 2016

- Rosa
 - dispel4py Stream based workflow mapped to MPI, Storm
 - MPI 3 Failure Recovery from Node Failures
- Monitord
 - Triggered by Condor failures. Workflow killed, condor recovery did not spit out all event on recovery.
 - Need better way to test.
- DB Admin
 - Merge issues
 - rafael with confirm with gideon if there is an issue

- Bamboo
 - Rebooted for DROWN Attack
- R API
 - Unit tests done.
 - Packaging - Ship, host?

February 2016

February 19th, 2016

Pegasus development

- support for GO - mats is working on it
- dashboard shows multiple workflows with same uuid. fixed in monitord
- pegasus transfer was prepending path because of globus location
 - mats has changed the logic
- SCEC wanted to disable the stat of files that was happening automatically because of registration turned on.
 - we now have the property that can explicitly turn it off
- SCEC tripped over replica catalog insert performance.
 - rafael working on it. identified the bottleneck
- Catalog files in submit directories
 - will create a catalogs directory
 - what about file based replica catalogs and cache files etc? some of them can be large.
- Pegasus Blogs
 - SCEC
 - RVGahp?
- Website
 - highlight applications better.
- workq has a catalog server running
 - how do jobs report real time monitoring information back to monitor without rabbitmq
 - have a condor submit wrapper
 - will help us increase memory requirements in case of failures.
- PegasusLite to have pegasus-transfer invocations as kickstart records
 - kickstart

February 12th, 2016

Pegasus development

- support for GO
 - mats found a python REST API - is decent.
 - will only work on a small subset of workflows
 - only third party transfers
 - how to handle file URL's on the submit host
 - and how do we activate the end points.
 - lifetime of credentials .
 - cannot work on non shared fs mode, as what end point to use when staging to the worker nodes.
 - maybe we should look at how condor does it.
- held jobs
 - dagman added support in 8.3 where the held job reason appears in dagman.out
 - will need schema change
 - failing workflows
 - held jobs.
 - have a held job tab.
- pegasus-submitdir archive
- PMC job statistics in pegasus-statistics
 - mats and rajiv
 -

Annual Report

February 5th, 2016

Pegasus development

- 4.6.1 release
 - pegasus-glite-configure
 - change of how retries are done for transfer jobs, using requirements and dagnode retries
 - <https://jira.isi.edu/browse/PM-1049>
 - there are just 2 retries implemented for transfer jobs

- one more option is for pegasus-transfer to do better retries
 - and let the dagman retry set to 1.
 - use DAGMan influence to do in retry.
 - do more testing at our end.
 - lets change default retries for transfer jobs
 - and do this only for transfer cleanups in condor environments
 - LIGO runs
 - symlinking
 - R API
 - will target 4.6.1 and keep it similar to the python API
- 4.7.0 release
 - filesystem organization
 - Keck workshop on Pegasus on Feb 26th
 - Pegasus Annual Report
 - Pegasus GUI email
 - we will send user a direct link
 - Pegasus Announce SLES email
 - we have done on SLES 11 not on SLES 12

January 2016

January 28th, 2016

Pegasus development

- 4.6.0 release
 - Released this week
- Pegasus Website
 - new website there
 - karan will put in the old release notes.
 - Links for old documentation on the new website
 - <https://pegasus.isi.edu/docs/>
 - Rajiv has updated the docker tutorial
 - Tutorials will be moved to Pegasus website
 - Have a research link to point to Scitech website
- Gideon confirmed MoabGlite helper scripts work with stock condor
 - will also check in a tool to put in the scripts to the right locations.
- Pegasus Lite pulls in a worker package
 - should we download even by default from the worker package
 - warnings for worker package not being found.

January 22nd, 2016

Pegasus development

- 4.6.0 release
 - open items
 - <https://jira.isi.edu/issues/?filter=10952>
 - constraints algo implemented and checked in . tests worked .
 - documentation
 - karan added chapters on metadata and variable expansion
 - gideon updated execution environments
 - updated the BOSCO section about SSH
 - pegasus-analyzer exits gracefully when nothing in the stampede database
 - check if analyzer and statistics check for the version.
 - pegasus-init
 - pegasus-db-admin
 - better error message for that case.
 - karan will update tutorial to take account of default options
 - for glite style condor arguments quoting is automatically turned off
- new website.

January 15th, 2016

Pegasus development

- 4.6.0 release
 - open items
 - <https://jira.isi.edu/issues/?filter=10952>
 - Rafael almost done with Constraints cleanup algo. tests run fine on the branch
 - pegasus-bootstrap

- gideon was doing it as Jinja templates
 - will set it up a shell script. will be easier for people to update
 - documentation needs to be updated
 - map the globe
- for resource requirements add pegasus.queue keyword. update documentation to have one table. remove the documentation for priorities.
- MOAB stuff documentation. Will be considered for next major release.
- DAGMan wants to remove the functionality of running postscript in case of prescript failure
 - does not affect pegasus
- DAGMan wants to remove DAG NOOP keyword
 - was introduced for LIGO

January 8th, 2016

Pegasus development

- 4.6.0 release
 - open items
 - <https://jira.isi.edu/issues/?filter=10952>
- Condor DAGMan log messages contain HTCondor in 8.5 series
 - broke monitord
 - fixed both 4.5.4 and 4.6.0.
- 8.5.2 has DAGMan logging timestamp from condor job log also.
 - monitord has been updated for that.
- metrics reported were updated
- Globus strict checking mode.
 - gridftp + ssh version.
- Scott is working on getting the reverse GAHP stuff
- How to configure the batch_gahp

December 2015

December 18th, 2015

Pegasus development

- 4.6.0 release
 - open items
 - <https://jira.isi.edu/issues/?filter=10952>
- Reverse GAHP for Oakridge Titan
 - <https://github.com/juve/rvgahp>
 - done because cannot do incoming connections on titan
 - and also they don't want to use pilot jobs, as it is not easy to yank a job from a HTCondor queue
- Harvard Pegasus installation
 - with SLURM support.. Karan will work on this.
- We should explore remote batch GAHP stuff
 - for remote batch do
 - batch gahp --rgahp-key /give/key user@host
 - look at the remote_gahp script.
 - documentation for the batch gahp thing.

December 11th, 2015

Pegasus development

- 4.6.0 release
 - open items
 - <https://jira.isi.edu/issues/?filter=10952>
 - pegasus-db-admin
 - cleanup algorithm
 - raphael will work on it.
- pegasus-s3 cert issue
 - updated boto library to account for cacert change
 - on mac, had to disable the automatic failover
- Bypass PFN's
 - replica selectors can now order replicas. Default and regex ones updated
- monitord
 - combination of missing job terminated and exception on casting job duration as int, triggered a bug that LIGO reported.
- default behavior of planner
 - pick up pegasus.properties from cwd as a replacement for conf option
 - --sites option for * behavior , remove local from candidate sites
- pegasus-bootstrap commands
 - sets up pegasus with site catalog. and dax generators

December 4th, 2015

Pegasus development

- JDBCRC
 - should work for 4.5.3 . will work for the release
 - need to make the changes for 4.6.0
 - should consider batch inserts
 - rafael has implemented the batch inserts also
 - the database locked errors are fixed.
- Rafael is looking into how the timeouts are implemented in sql alchemy
- Mac OSX El Capitan Builds
 - Gideon fixed those. El Capitan does not allow root to modify files in /usr
 - Gideon changed the installer to install to /local
 - Upgrading the mac mini build host.
- LIGO proxy issue
 - change in how proxies are generated.
 - LIGO en-common proxies were not supported by J-Globus
 - Gideon has the patch for making the updated jar.
- Gideon has added instructions on building globus for El - Capitan
- Jobmanager-condor for obelix was updated to support both shared fs and non shared fs cases.
- metadata registration
 - information for output files is tracked.
- pegasus-metadata client . Rajiv.
- Cleanup algorithm - Rafael ?
- LIGO use case for fallback PFN for PegasusLite cases
 - they want to use existing input data for frame files, on different locations across sites
 - but have a single site catalog entry for the computation, as glideinwms provisions it
 - Karan and Mats are working on it
 - pegasus-transfer changes ?
 - sd
- LIGO running workflows across LIGO and OSG .
- Database locked errors for monitord.
- Call the 4.6 release as 5.0 release.
- Gideon working on MOAB Blahp support.

October 2015

October 23rd, 2015

Pegasus development

- Tutorial VM
 - rajiv will update dashboard screenshots and go through the Virtual machine based tutorial
- JDBCRC
 - should work for 4.5.3 . will work for the release
 - need to make the changes for 4.6.0
 - should consider batch inserts
 - sqlite supports unlimited connections
 - for write locks , 25 jobs running for write locks. after 25 and it ignores timeout settings.
 - 67 registration jobs.
 - raphael is implementing a back off
 - category for the registration jobs
 - eventually do the dagman category stuff
 - metadata registration
 - information for output files is tracked.
 - pegasus-metadata client
- concurrency limits
 - in partitionable slots this has an affect on performance
 - for 4.5.3 we will have a knob and set it to false by default.
- Dashboard and PAM problem.
 - mats will create JIRA item.
- salon working on data from MYRA
 - trying to find contention of data

October 16th, 2015

Pegasus development

- does stime include io wait time. does not appear so. the cp of 1GB file indicates that
 - so then is there a way to capture the IO wait time
- pegasus-db-admin
 - version migration for panorama works
 - metadata schema finalized

- failing jdbc RC test
- metadata population
 - metadata population from DAX working
 - metadata attributes from transformation catalog and site catalog are now incorporated, as metadata events are generated at end of site selection
 - output file sizes will be populated for files with register flag set to true.
- pegasus dashboard
 - metadata display done other than the file information that needs to be populated
- cleanup algorithm
 - will be done before raphael leaves for vacation
- website changes
- panorama changes
 - monitor change to make sure events don't get dropped
 - online monitoring spawns a thread where there is a queue that is responsible for inserting the online monitoring events into the db
 - the thread checks the database to make sure the job instance is populated.
 - CURRENTLY, it is not done for the anomaly populations.
- SNS and Acme workflow
 - maybe we can hire a student to do it
 - maybe scalarm can be used for SNS workflows
 - Ben said there is a meeting about Pegasus on Titan.
- Mats has installed wordpress on one of the machines.

October 9th, 2015

Pegasus development

- pegasus-db-admin
 - db version has been moved to string. a new column was added.
- metadata population
 - files are populated if a user specifically associates metadata with a file in the DAX or if an output file is marked for registration
 - make sure that for tasks metadata attributes are inherited from the transformation catalog.
- pegasus-metadata client
 - output format ?
 - is the client for end users
 - list files for a workflow
 - list workflow metadata
- pegasus dashboard
 - workflow level
 - task level level
 - file level metadata

October 2nd, 2015

Pegasus development

- pegasus-db-admin
 - changes discussed last week?
 - also change to string for the database version for allowing merges with panorama
 - panorama db versions should be N.x and not whole integers
- jdbc sqlite test failures
- pegasus-transfer
 - better job with grouping for ssh transfers.
- metadata population
 - planner generates the events now for associating metadata with wf, job and files
 - use case should be for a file what workflow and job created that file.
- Pegasus workshop
 - we will be using workflow.isi.edu
 - mats has created 30 training accounts on workflow.isi.edu
 - suggestions on workflow example?
 - blender rendering example..
 - pegasus-dashboard should be installed
- Sipt portal
 - back up and running

September 2015

September 25th, 2015

- Pegasus development
 - pegasus-kickstart to return record on condor_rm (SIGINT)
 - changes to data reuse algo for Chris Edlund
 - delete jobs when inplace cleanup is used for intermediate files that are not transferred to the output site.
 - use of DAGMan NOOP keyword
 - workflow test failures

- change monitor to not complain for noop jobs.
 - comma separated directories for input dir
 - automatically delete the input directory ? we all agree not a general use case.
 - pegasus-transfer grouping should be done for all protocols?
 - problem is some renames for output files
 - avi has been running workflows on OSG with pegasus lite.
 - 2 million connections over two days on SSH server
 - pegasus-db-admin error handling.
 - if it fails with error, it should not report that database has been updated. This is a bug
 - other is what to do , when 4.5 is run against
 - downgrade option
 - warn if db-admin detects database version is higher than what it is currently running, and exit with 0 exitcode.
- Pegasus IEEE article accepted
- montage workflows
 - dax generator is not maintained
 - have it as a student project to convert the DAX generator to python API.
 - they also check an overlap check
 - montage jobs have varying memory requirements
 - we should not showcase it.
- Pegasus Workshop in October
 - fallback from USC HPCC cluster required
 - whole day will be rough.
 - Mats will not be around! Going for the duke workshop.
- panorama
 - monitoring thread segfaults
 - why was the segfault happening initially
 - happening in fork system calls
 - related to starting and stopping monitoring threads
 - and how PAPI counters were updated.

September 18th, 2015

- Pegasus development
 - pegasus-db-admin updated
 - for spec added registration of flat lfn's when deep LFN are used
 - workflow tests now running.
- pegasus paper
 - will add info about galactic plane and gtfar
 - cloud challenges
 - talk about virtual clusters . precipe / wrangler
 - tie more closely to setup stuff and talk about chef/puppet and precise and wrangler.
 - gtfar
 - add them in acknowledgements
 - not much to add about cloud challenges other than image managements
- hubub conference
 - latech user who wants to run on bleaters
 - tom bishop
 - pegasus submit tutorial.
 - to do with steven...
- panorama
 - segfaults happening randomly
 - happen when the monitoring thread is started.
- craft
 - jarek
 - hubzero
 - chip design
 - instead of hubzero use open science framework - a non profit funded thing

September 11th, 2015

- Pegasus development
 - worker package tests in pegasus lite
 - pegasus lite will complain if the system architecture
 - panorama tests now work
 - maybe some problems might be masked!
 - jdbcrc
 - updated jdbcrc . for mysql and postgres deletes work differently.
 - raphael will abstract it out
 - gideon changed the way the papi counters are used in kickstart
 - earlier signals were being used for threads to report counters
 - PAPI now allows to query for counter values
- Pegasus cloud article
 - ewa is doing the final edits
- HubBub presentation
- panorama
 - darek working on getting papi counters to monitor
 - changed the job metrics table in the stampede database.

September 4th, 2015

- Pegasus development
 - worker package creation on the submit host.
 - should we include python externals directory .
 - we will put that back in. we only need boto.
 - also need to make sure it works for a RPM or deb install.
 - implement the compatibility check in PegasusLite
 - panorama tests
 - better error for input file replica selection failures
 - Scalr for openstack tests
 - action has a new openstack deployment.
 - have our two QNAPS setup on the build VM's to run workflow tests.
 - run on vmware pool.
 - SCEC shallow LFN's
 - for registration in the replica catalog.
 - put the test in 4.5 .
 - Database schema changes
 - pegasus-db-admin changes to database schema.
 - downgrades work
- The short paper
 - working on the google doc.
 - we are not actively working on ec2.
- panorama
 - adding papi counters to online monitoring.
 - pegasus-transfer explodes when signal is sent
 - online monitoring dashboard.

August 2015

August 28th, 2015

- pegasus 4.5.2 released
- worker package staging
 - planner will use a worker package from the submit side installation and use it.
- pegasus s3 tests
 - currently no s3 tests
- tests are running against 8.3.8
- cleanup algorithm update (Rafael)
 - estimate that it will be done in two weeks
 - has to work for multiple sites
- cloud computing short paper
- hub bub
- panorama and dv/dt poster and presentations . in mid september
- metadata discussion
 - google doc updated
 - leaning towards monitor populating the database
 - remove the estimated size and md5 checksum

August 21st, 2015

- pegasus 4.5.2 release
 - release notes checked in
 - db-admin changes?
 - update man pages
 - python source package
 - tests are we moving to dev branch?
 - docker problem
 - how to get around it ?
 - an issue inside docker, that is being exposed
 - we will put in a wrapper around it.
 - panorama branch is disabled
 - but tests should be fixed.
 - dark will be fixing it
 - rajiv pushed out his dashboard changes for darek. for demo at supercomputing.
- cleanup algorithm
 - Rafael will start next week
 - how will the limits be passed
- kickstart changes
- metadata schema discussion
 - next week.
 - postscript
 - dagman has plugin's

- schema
- use case
- stampede is sqlite
- pegasus-exitcode write locks.
- separate sqlite database for metadata.

August 14th, 2015

- Pegasus 4.5.1 release
 - Release notes online <https://pegasus.isi.edu/news/4.5.1> is done
- Bamboo machine troubles
 - panorama tests hung because of bamboo
 - do experiment for the case where we do condor off and see what happens to pegasus-dagman.
- Panorama tests
 - look at build #73
- pegasus-kickstart stuff
 - for interpose stuff
 - gideon investigating how to cover all cases for threads
 - wants to make sure that descriptor table is accessed in a thread safe way. in worse case
 - also is doing thread tracking, thread counters and thread lists
- directory structure organization for submit directories.
- nonsharedfs mode problem for auxillary jobs
- sudharshan cleanup algorithm
- stefan update
 - working on user models on how to submit jobs to HPC
 - what user characteristics are of submission process
- to be able to show the IO part for SoyKB
 - metrics of success
 - makespan is reduced.
 - number of service units is reduced
- what makes an application IO intensive

August 7th, 2015

- Pegasus 4.5.1 release
 - Release notes online <https://pegasus.isi.edu/news/4.5.1>
 - PEGASUS_SCRATCH_DIR is populated for all jobs in shared fs case.
- 4.6 common resource requirements
 - we are now exposing three pegasus profiles cores, nodes and ppn.
 - added logic to do specific translations for PBS and SGE
- cleanup bug fixed related to DAX transfer flag for input files
 - larger question and agreement. transfer flags for input files usually don't have any meaning.
 - transfer flag should be renamed or in the API
 - change in schema
 - at minimum we should change the DAX API's
 - transfer attribute renamed to final output?
- spaces in Pegasus URL
 - gideon feels it should be mod 20 instead
 - somewhere in documentation .
 - the planner should have more specific error message in case of spaces.
- kickstart enhancements - gideon
 - fixing edge cases in kickstart for the extended reporting
 - what can we do with the papi performance counters and see what will be used in panorama.
 - will be updated for counters.
 - gideon and darek will try and merge

July 2015

July 31st, 2015

- Pegasus 4.5.1 release
 - will release it next week
 - update the mapper documentation
 - have a link to the replica catalog
 - steven clarke cleanup issue
- resource requirements
 - update the resource requirements section for 4.6
- acme integration
 - rajiv will work with bibi to integrate it with the REST monitoring api
- kickstart changes to get papi counters
 - Only triggered if -Z option is passed
 - the paper on xsede mentioned about them reporting per threads
 - also we make better track of threads launched by the executable
 - some edge cases for the thread case

- double execve of process does not work currently
 - example: /usr/bin/env date
 - also record command line options for all sub process launched
 - in the proc record , the cmd tag
 - grabs only first 1K of arguments
- monitor amqp population
 - revert back to use the event name as the routing key for AMQP population.
- pegasus cleanup with peak storage requirements
- Panorama
 - Data analysis done..
 - ideas about writing a paper about workflow profiles
- Anomalies Detection
 - showing anomalies in dashboard and population in stampede schema

July 24th, 2015

- XSEDE Tutorial
 - 2 Posters and one tutorial
 - news item online
- Pegasus Development
 - common resource requirements PM-962
 - documentation needs to be updated
 - we have cores , hostcount
 - karan should make sure cores is translated correctly to ncpus for PBS
 - Pegasus REST API for integrating with Pegasus
 - pegasus transfer
 - checkpoint files
 - LIGO developer notion of site attribute
 - maybe we should be more clearer in the documentation
 - automatically changing parameters for memory on job retries
 - check point file for the job is a partial solution
 - monitor amqp population
 - works.. we will document it on JIRA
- Panorama
 - Darek implemented sending messages in batches from kickstart to rabbitmq
 - socket based communication between kickstart and lib interpose . was done to take of the file interleaving issue.
 - tests on obelix and exogeni indicate socket writes are atomic for panorama message

July 17th, 2015

- PMC Cpu affinity
- LIGO pegasus analyzer bug
 - has been passed to LIGO . awaiting to hear from them
- Cleanup algo
- Resource Requirements
 - common pegasus profiles
- SGE
 - change.dir should be set automatically for shared filesystem stuff
 - documented already.
- kickstart path variable to prepend.
- REST interface for monitoring for pegasus is done. Rajiv completed this week.
- extensions to the cleanup algorithm. rafael will start working .
- Pegasus 4.5.1 release
 - will be done after XSEDE.
- Pegasus XSEDE tutorial
- XSEDE Pegasus Poster
 - show a LIGO workflow for the XSEDE poster.
- Salt configuration needs to be updated
 - Student machines on salt
- panorama
 - rabbit mq installed on exogeni site.
 - darek will do message batching working.
 - gideon recommends doing it with the AMQP C API library
 - message interleaving in kickstart.
 - lot of unacknowledged messages in rabbit mq
- kickstart polling loop
- all kickstart memory values are in MB

July 10th, 2015

- PMC jobs automatic summing of maxwalltime. Should be disabled
 - In PMC case we will do a division.
- PMC CPU affinity for jobs PM-953
 - there might be a fragmentation approach.
- Pegasus REST interface
 - short cut URL end points.

- karan will send email to Lavanya.
- running on SGE cluster using GLite interface.
- harmonized pegasus profiles
- Metadata
 - will need the file implementation .
- Dashboard Panorama stuff
 - September 16th. Time series and anomaly detection.
 - Application level anomalies
 - Infrastructure level anomalies.
 - no plans for integration in production Pegasus.
- monitor profiling of monitor population.
 - we want to see how long 1000 events take to be populated in case of LIGO .
- Panorama
 - anomaly detection
 - implemented a working prototype of threshold based anomaly detection
 - kickstart sends events to rabbit mq, then monitor populates to influx db.
 - darek tool queries influx db and takes in the metadata file generated by pegasus and determines the anomaly and sends it back to rabbit mq
 - monitor then again picks up anomaly and populates it to stampede db for dashboard to display.

June 2015

June 12th, 2015

- Pegasus profiles for job/resource requirements
 - postponed till next week when mats is here
 - karan to create a list of relevant profiles
- pegasus dashboard
 - locking issue?
 - can this be related to new connection stuff or the failing tab?
 - look at connection pooling .. or maybe transactions are not being closed properly?
 - also see if there is an option for dashboard to set a read only lock when opening a connection to the databases
- panorama workflow tests
 - failing.. but merge from master was done.
 - karan to investigate
- panorama workflow dashboard
 - updated the job metrics tab for doing the polling
 - for mpi jobs the job name appears as aprun, since that is the process running on rank 0
- Job Survey paper
 - Darek sent a final version
 - will be submitting next week
- Pegasus Release timeline
 - maybe we should put on our website somewhere?
- Rafael Energy paper
 - information about building energy profile.

June 5th, 2015

- panorama usecase and metadata passing through
 - not done yet for the metadata associated with files with replica catalog
 - DONT rebase commits that have been pushed out
- job.runtime, cluster.maxruntime, maxwalltime parameters
 - how to associate profiles. have a different namespace
 - how is it expose in the DAX API
- python dependency
 - stopped support for 2.5 and 2.6
 - only affects redhat 5 systems.
 - will have to install redhat 2.6 python package on 2.5
 - setup tools for python 2.6 has to be at build time
- pegasus-dashboard updates for LIGO
- cleanup bug for intercept runs with InPlace cleanup.
- S3 storage
 - about 9TB and rising for pegasus system services backup
 - right now no backups are going to go to Glacier
 - we only keep 2 weeks of data
 - glacier is good if we want to keep 6 months of data
 - 3VM' for pegasus website , CROWD etc
 - database on stey and obelix
 - qnaps /nfs/ccg3 and /nfs/ccg4
 - Big ticket items of 9TB backup bucket in S3
 - need to keep 2 backups in S3
- HubBub talk.
 - abstract

- talk by Jack Donagata.

May 2015

May 29th, 2015

- Bamboo test failures
 - condor-c tests working now. changed the site catalog for those
 - rhel5 json module
 - pegasus-transfer will do a proper check and complain for missing json module
 - mats will update documentation accordingly
- Python Dependencies
 - New python dependency 2.6 from 2.4
 - newer versions of Fedora uses Python 3
 - Fedora will keep python 2.x support till 2020.
 - maybe have a dynamic bash wrapper across python code to pick the right python version
 - have a tool called pegasus-python??
- concurrency limits
 - apply to bamboo machine and our other workflow hosts.
 - throttle number of grid jobs per categories of jobs. that is what SCEC wants and cannot be done.
 - unless negotiation can be employed for grid universe jobs.
 - define own throttles in compute jobs
- pegasus-dashboard
 - LIGO has an issue with no authentication URL rendering.
- quoting for environment
 - implemented. changed both for environment and +remote_environment
- docker universe support
 - should work out of the box with condorio
- new dagman default values
- pegasus-statistics
 - show bad put?
- LIGO OSG
- Documentation
 - 10 minutes using pegasus-docbook
 - using new pipeline it uses 3 minutes
 - the hyperlinks don't work
 - include that into pegasus website template
 - In PHP we tell Google not to index old version
- panorama

May 8th, 2015

Bamboo test failures

- montage tests are failing because of the remote service being down
- documentation tilte is messed up. gideon will look at it

pegasus-transfer new format

- mats has come up with a new JSon format.
- backward compatibility with the old format
- create dir and cleanup jobs will be different

Metatdata

- google doc shared with people
- next steps are panorama use case for calling out
- ssh cleanup . JGlobus library does not implement ftp

LIGO on XSEDE

- have started using PMC
- data management

Python builds

- always check the python version.
- if we ship our own python modules, then we may have to

Bamboo build machine

- build and test plan (running concurrently)
- also we can run docker stuff
- automate the salt setup of bamboo agents
- maintain one OS. Can action give us a beefier VM?
- we have too many documentation builds running ?
- VW with bamboo agent and use docker

- workflow tests are a separate issue
 - they don't load the bamboo machine
 - that is more related to a big condor pool.
 - workflows tests will run always out of bamboo.
- mats and rajiv will work on it for the VM stuff.

Getting new SSL certificates

- *.isi.edu is screwed up in firefox

Metrics Server fixes

- google maps update broke the web UI.
- somehow all the colors were used in the trends ?

May 1st 2015

- Pegasus 4.5 release
 - not heard back from SCEC and LIGO
 - mats checked in the example
 - will add release slider
- Variable Expansion
 - pretty much done
 - right now we have \$()
 - we will change with \${env-variable}
 - have more helpful error message
- pegasus-kickstart
 - file does not exist. now gives a proper error
- XSEDE poster due next week
- Monitoring Service API
 - donald is almost done.
- PMC with PegasusLite
 - PMC job by default runs on the shared filesystem
 - tasks in PMC are pegasus lite tasks
 - if a task does randomio, then on shared fs might be tricky
- brazilian student contacted about pegasus application for real workflows.
- mats will be doing the transfer events for panorama next week

April 2015

April 24th 2015

- Pegasus 4.5 release
 - release candidate today rc2
 - updates to pending items
 - job throttling added to optimization guide.
 - release notes are online <https://pegasus.isi.edu/news/4.5.0>
 - waiting for db-admin unit tests to be checked in.
 - pegasus-cleanup checking
 - pegasus-lite-local.sh add some path before starting.
- rest monitoring API
 - we have not heard back from lavanya yet
 - PNNL acme stuff
- pegasus 4.6 release
 - common pegasus-transfer , pegasus-cleanup and pegasus-createdir
 - APP_PATH_PREPEND addon
 - pegasus worker package staging
 - planner calls out to common script to determine the worker package
 - if it does not exist , we build a default worker package on the fly
 - add extra logic to the untar job in the
 - pegasus-gridftp modification for ssh ftp.
 - software eggs
- panorama
- metadata for 4.6

April 17th 2015

- Pegasus 4.5.0 Release
 - rc1 working for hub
 - LIGO trying it out.. wanted to change checkpoint files. need to hear back on the dashboard changes.
 - SCEC ? waiting to hear from Scott

- <https://jira.isi.edu/issues/?filter=10851>
- pegasus-db-admin sqlalchemy issues? for updating tables?
- pass through implemented for Glite to PBS
- verification of update to pegasus version on running workflows
 - mats thinks his testing should do the trick.
- Pegasus Dashboard for bamboo user
 - URL - <https://cartman.isi.edu:5000>
 - Authentication - Uses PAM Authentication
 - Admin Users - mayani, vahi, ryнге, juve, rafsilva, darek, deelman
- Cedars visit
 - SGE cluster
 - we have 3 potential SGI cluster users Cedars, Vision group at ISI and maybe Rutgers (that will be replaced with SLURM)
- Lavanya REST API
 - <http://pegasus.isi.edu/wms/docs/4.6.0cvs/rest-api-monitoring.php>
- Pegasus 4.6 release
 - variable expansion thing figured out
 - argument strings in dax, profile values in the dax
 - site catalog.
 - replica catalog file based one.
 - need to now make changes in various parsers
 - predefined environment variable
 - metadata
 - LIGO Dibbs .. ability to do data reuse based on metadata attributes
 - panorama - pegasus - aspen interface
 - iplant
 - they want in the IRODs
 - S3 tags.
 - mats wants a better idea of what it looks like in the ideal world.
 - file management on scratch directory, submit directory also?
 - implementation of the REST API
 - implementation for held job tracking
 - Panorama requirements
 - influx db monitoring , into pegasus-transfer.
 - pegasus-transfer sends messages to rabbit mq about file size transferred
 - pegasus aspen interface (modelling tool) . apsen is a C++ library.. pegasus planner querying the aspen models for each node.
 - command line tool pegasus-aspens
 - planner needs to send application parameters, and all the metadata for the node.
 - gets back a list of attributes , memory and usage, and convert them internally into pegasus profiles
 - this can be a generator of metadata.
 - application model which is a file and a machine model
 - timeseries data . monitoring data about the dashboard, anomalies
 - there is a CEP thing that anirban is developing and will determine anomalies.
 - dv/dt requirements
 - prediction service
 - pegasus will query the prediction service

April 10th 2015

pegasus cleanup

- gideon removed a bunch of stuff
- will be completing the cleanup
- pegasus-plots will be deprecated in the release notes for 4.5 release and removed for 4.6

pegasus RC1

- built now.
- should have created a 4.5 branch and then done a tag
- pegasus-halt (is it prototype)
- pegasus-run on already running workflow
- pegasus-db-admin missing import
- mats will delete the rc1 branch

pegasus 4.5.0 release

- karan will add options for pass through text for Glite options.

pegasus-db-admin

- should be done soon

HPCC tutorial

- send link to Fan fli from CHLA
- vision group at ISI . former BBN people.

XSEDE paper

- submitted to xsede

- for journal paper, expand to pilot workflow systems. panda, swift coasters, big job

REST API

- rajiv will add to the docbook
- largely agree
- uuid for the top level workflow

April 3rd, 2015

- Pegasus 4.5 release
 - pegasus-db-admin
 - ds
 - planner will set auto update on pegasus-db-admin . and include
 - extra python modules being shipped mysql config and postgres config
 - right now on our build hosts we are building mysql and postgres.
 - RPM packaging adds dependencies automatically
 - openssl dependency
 - best option is database dependencies optional
 - targets 4.5.0 pre release candidate for thursday
 - pegasus-dashboard updates
 - pegasus-monitor failed for 4.4 runs
 - documentation
 - fix missing references
- REST API for monitoring workflows and jobs
 - work on it for next week.
- questionnaire
 - 15 responses in all.
- xsede paper
 - deadline on monday . 8 pages.
 - have number of cores
 - no reliable way for specifying cores on OSG
- web interface for influx db
- permanent influx db install

March 2015

March 27th, 2014

- metrics server
 - final change pushed out by donald
- REST API
 - job monitoring API for workflow and jobs
 - will work with Rajiv
 - next week friday we will have a spec out for the API
- Pegasus 4.5 release
 - resolving pegasus-db-admin issue
 - work on the documentation
 - should reach may first deadline
 - next week we will do a pre release for SCEC.
- Job submission paper
 - for xsede some sections you will remove.
 - need some major modifications regarding introduction.
 - new deadline for xsede is april 6th.
- pegasus transfer issue in google cloud vs amazon cloud
 - gsutil causes a 1 second overhead for a zero byte file. probably an authentication protocol
 - directly with wget works faster.
 - when you downloading larger files
 - huge overhead compared to 3 times in amazon.

March 20th, 2015

pegasus 4.4.2 release done

- will be deployed by LIGO

tagged release for SCEC production runs .. we will do a pre-release candidate

metrics server

- follow up on histogram page?
- gideon will deploy the changes on the production machine

pegasus-db-admin

- updates
- dashboard and stampede expunge functions.
- sql alchemy init and duplicate code. will enable foreign keys.
- SQLAlchemy init interface takes a URI.

pegasus-submit-dir

- till we come up with a better name
- can archive, move and delete

pegasus-dashboard archive option

- gideon will make changes to the dashboard schema.

transfer grouping in Pegasus

- PM-829

PM-851 kickstart invoke option for auxiliary jobs

pegasus dashboard updates

- LIGO uses for apache to use uncommon for single sign on and authentication

job submission survey short paper

- march 30 deadline

Panorama Updates

- wants to have a separate panorama branch
- mpi-exec has been merged back to master.
- similar to the adamant branch
- rabbit mq
 - has a rest interface
 - so easy to post http messages to it
 - uses small amount of memory
- long term we will have pegasus-service receive the messages instead of rabbit mq.
- we are collecting data and share with other people in collaboration
 - http location on obelix (the way we did for stampede)
- real time monitoring in kickstart
 - runtime metadata and file descriptor 3 (did for hubzero)

User Questionnaire

- still at same place as earlier
- gideon will send out a reminder

March 13th, 2015

- Metrics Server
 - deployed on the production server.
 - want to do anything on basis of distribution of files
 - donald will create a new histogram page ,
- Pegasus NSF Report
 - sent to Ewa
- Pegasus 4.4.2 release
 - karan will check in release notes today
- Pegasus Tutorial as part of HPC Workshop Series in April
- Gideon will be going to the summer school.
- Pegasus 4.5.0 release
 - Targeting May 1st release
 - local-scratch is picked up.
 - ensemble manager submission
 - will support both modes
 - bundle mode
 - public ensemble manager. there are security issues. user credentials.
 - the person who starts the service will setup the credentials
 - pegasus-analyzer fix for case where jobs eventually succeed after failures
 - pegasus-db-admin update
 - ds
 - transfer grouping of staging jobs
 - Pending items
 - <https://jira.isi.edu/issues/?filter=10851>
- User Questionnaire
 - 12 responses for

- a lot of people are interested in a workshop
 - better support for loops and branches
 - better provenance support .
- Workflows on Google and Amazon
 - google takes much longer to do data transfers.
 - non shared fs and shared fs
- metadata
- Panorama
 - Demo in September of Panorama functionality
 - getting data transfer metrics out of pegasus-transfer in structured way
 - what data we need to collect
 - for third party transfers we can do timings but not rates
 - darek is working on adding real time monitoring to pegasus-kickstart
 - pegasus transfer will communicate to pegasus-kickstart to report to a central server
 - can be a http server similar to metrics server
 - panorama is considering influx DB for real time monitoring.

March 6th, 2015

- metrics server update
 - plans to deploy the changes today. fixing last issue
 - still has to make the database schema changes required for planner file counts
 - will be done next week
- planner reports file breakdowns
- pegasus 4.4.2 release
 - it has fixes LIGO is interested.
 - most probably next week.
- pegasus-db-admin
 - reorganization of the code and the schema.
- pegasus-archive /pegasus-delete
 - rafael does not have time to work on these because of proposal work
 - will move to either gideon or mats
- pegasus-dashboard updates
 - has more LIGO requests for pegasus 4.5.0 release
 - wsgi script for root mode
- LIGO visit
 - post 4.5 we will do better organization of files on the file structure
 - Pegasus poster for LIGO meeting
- ensemble manager
 - scec folks will try it
 - monitord netlogger bugfix
- pegasus-transfer enhancements for panorama
- job submission paper in github
 - pegasus and job management systems.
- online monitoring for pegasus-kickstart
 - application sends signal to pegasus-kickstart via libinterpose
- pegasus-keg extensions
 - the pegasus-mpi-keg is a separate executable
 - extensions to the io stuff
 - will incorporate in 4.5.0
- NSF report
 - still waiting to hear from mats and scott
 - karan is still updating the metrics page.

February 2015

Feb 20th, 2015

- metrics server update
 - donald still has to deploy the changes.
- pegasus user questionnaire
 - gideon will send new links and will update
- SCEC update
 - scott has debugged his memory
- Pegasus Report
 - soykb and other iplant workflows ... part of ECSS
 - galactic plane
 - ahmeds work
- pegasus dashboard updates
 - pegasus-dashboard is started whenever bamboo is built up
 - dashboard show all states for a job now.
- pegasus-db-admin tool
 - test cases in bamboo
 - documentation
 - migration notes
 - some python errors that need to be fixed.
- 4.5 release

- still remaining
 - held jobs tracking in monitor
- job retry set to 1 and disable retries for DAX jobs
- decrease the held period from one hour when job is removed.
- improved documentation for output mappers
- ensemble manager todo's
 - we won't have ensemble manger in multiuser mode
 - support both modes (upload a tar file and finer grained control where he specifies the DAX files and the submit directory)
 - only the dashboard will run in multiuser mode
 - how do we start ensemble manager process
 - run as per user .
- copying of catalog files to submit directory.
- input directory copies based on recursive transfers as part of directory
 - it won't work in condorio mode because it flattens out
 - add type directory in the DAX schema.
- pegasus tutorial
- environment variable file substitution in site catalog, replica catalog and transformation catalog
- XSEDE Tutorial proposal and Posters

January 2015

Jan 14th, 2015

- metrics server update
 - no update from Donald still away from vacation
- Pegasus development
 - data configuration for different sites
 - working for steven
 - held jobs
 - pegasus-dashboard
 - root mode for dashboard and ensemble manager
 - gideon needs to confirm for ensemble manger
 - done for dashboard
 - pegasus-analyzer bug fix
 - pegasus-db-admin tool update
 - unit tests
 - bamboo pool will break.
 - upgrade to newer version of Pegasus
 - what happens to running workflows
 - pegasus-statistics with PMC - Mats and Rajiv
 - mats and rajiv will work on it.
 - docker based tutorial launcher
 - how to integrate in the build process
 - form
 - candidate machine
 - obelix
 - vmware colo vm
 - obelix.
- Pegasus Poster for Si2
 - will base on the previous years.
 - any particular thing we want to focus on ? or general?
- Pegasus Annual Report
 - User questionnaire - need to send out.
 - list of people to send it out to . Gideon has one.?

Jan 7th, 2015

- metrics server update
 - no update from Donald still away from vacation
- 4.4.1
 - installed on workflow
 - OSG and XSEDE submit hosts will be upgraded in 3 weeks
 - need to follow up with LIGO
- database upgrade tool integration
 - documentation and manage left
 - import error for properties
 - python test case
- support for per site data configuration
 - mostly done/ still need to figure out worker package staging for that.
- pegasus-dashboard
 - should we show all job instances for a job.

- held jobs logged by pegasus-monitor
- user questionnaire

December 2014

Dec 8th, 2014

- metrics server update
 - minor bugs in the UI... still need to be fixed, especially how the session states are handled
 - things remaining to do
 - database/server side pagination
 - figure out the scroll issue for the trend charts
 - move the trends charts from the home page to under planner and download tabs
 - rename run metrics to dagman metrics, and instead of showing the most number of times a workflow was run, we want to see the top applications for which dagman workflows were run
 - for the time bar on the top, have drop down menu for years and months
 - can the maps pin show the actual number, for example in the top downloads map thing
- monitor fixes
 - for the race issue with postscript handling PM-798
 - had to change the way stdout and stderr is populated for job_instance. It is now populated with the POST_SCRIPT_TERMINATED event happens
- pegasus-analyzer fixes
 - show the planner log when prescript for sub dax fails. PM-808
- we want to release 4.4.1 before the break.
 - has monitor fixes that LIGO requires
- tracking held jobs
 - decided to add a column in the jobstate table to capture why a job was held
- changes to pegasus-keg
 - to simulate reading in input and writing out of output files
 - will also simulate cputime and walltime
 - initially pegasus-keg will read in and write out the outputs and then do the sleep for the cpu time duration
 - removing the system information that it prints out
 - in the mpi version, the IO is solely done by the master.

December 3rd, 2014

- Update from Duncan on LIGO dashboard requirements
 - run a flask module from apache
 - let apache handle authentication
 - read only dashboard view
 - have a separate flask frontend.
 - they are ok with a command line tool to remove workflow entries
 - port collisions .. so they prefer apache to do the handling.
- failed jdbrc unit test case
- glite quoting for the environment
- pegasus-dashboard delete workflows capability
- failing workflow reporting in the dashboard
- monitor to follow condor job log
- db admin tool updates

November 2014

November 12th, 2014

- DAGMan metrics reporting
 - working and completed for 4.5.0cvs
 - planned metrics
 - exclude the metrics that never ran.
 - have a drop down menu - planned , planned and run
- RPM/ and DEB tracking for downloads
 - mats has a script that goes through the download logs to populate the server.
 - So we are tracking those now.
- Failed data reuse regex test
 - make it a planning only test case
- hierarchal workflows options forwarding
 - have a value of null/none
 - --inherit option with a comma separated list of long opts.
- higher level DAX API for sub workflows ?
 - hack to figure out the command line arguments for the planner
- Pegasus Distribute Wrapper
 - waiting to hear further from Steven

- a /bin/bash test case
- Metrics Server Updates by Donald
 - has the geo location running
- DB Upgrade tool - Rafael ??
 - version upgrades
 - specific table version
 - <https://jira.isi.edu/browse/PM-776>

November 5th, 2014

- DAGMan metrics reporting
 - already in recent DAGMan versions. can be enabled.
 - pegasus-run having the duplicate logic.
- Pegasus Distribute Wrapper
 - Initial implementation done and there is an example for Steven to try out
- Metrics Server Updates by Donald
- DB Upgrade tool - Rafael ??

October 2014

October 29th, 2014

- Upcoming Proposals
 - NEESGrid call
 - Robert Flashgun with Nirav..ASU stuff. Do some earthquake stuff
 - frank mckenna for needs type stuff
 - SCEC is part of the proposal
 - December 3rd due date
- Pegasus Development
 - monitor postsript handling
 - dynamic hierarchy stuff
 - Condor C with LIGO
 - Steven Clarke Distribute Stuff
 - pegasus-hpc-cluster (PHC)
 - DAGMan metrics
- Kenichi Workflow
 - SNS workflow
 - Training material.
- Metrics UI updates
 - Trends over times
 - Geo overlay
- Darek from Poland - A postdoc 1206
 - panorama project
- Adaptive Workflows
 - adapting workflows... they are not converging.
 - templating workflows
 - Hopper Site Catalog
 - Sample Site Catalogs

September 2014

September 17th, 2014

- Checkpointing feature
 - tested and implemented into pegasus
 - communicated with LIGO and John Veitch will test it next week.
 - will be run from a binary install
 - kickstart won't enforce non zero exit code for application exit code . we will require application codes to exit with non zero status.
- Profile and Properties documentation integration
- database schema upgrade tool
 - rafael starts working on it
- support for google storage
 - hassan writes a paper for google storage
 - compare S3 with google storage
 - parallel uploads of chunks not supported with gsutils.. relies on a very specific python module
 - ~/.botoconfig
 - uses oath token for authentication
- works paper revisions due oct 1st.
- dv/dt paper has been submitted as a CS dept tech report.
- DOE Oakridge meeting

- interface with ASPEN (analytical modeling) - domain specific language for defining code.
- combine aspen model with machine model and come up with estimates of runtimes.
- christopher riggers from RPI models parallel storage systems.
- Explore visualization stuff for pegasus-plots and dashboard?

August 2014

August 25th, 2014

- Ensemble Manager - User Authentication
 - initially gideon is working on a PAM based approach
- refactored netlogger dead code
- Workflow Checkpointing support - ongoing
- Google Compute Engine
 - related to google genomics
 - put in support for GCE transfer tool to interact with Google Storage (their S3 equivalent)
 - put in credential handling in the planner.
 - fits well with long term planning for pegasus.
- Replica Catalog Service

August 18th, 2014

- Data Reuse Partial Mode
- Service integration
- Profiles and Properties Documentation
 - Scope Column in the properties documentation (transformation, job and global)
 - in profiles documentation corresponding property key
- pegasus-service integration
 - need to integrate the documentation
- redhat 5 builds
 - partially... because of 2.4 installed version pegasus-s3 fail
- authentication mechanism
- pegasus-service-admin migrate option
- new tool pegasus-db-admin
- get a new 32 bit VM with cents 6.5
- also centos 7 VM
- add a setup task that cleans \$HOME/.pegasus in bamboo infrastructure.
- Docker Kernel Problem
 - if a docker build running and you stop the build, then the whole thing crashes
 - one solution is to upgrade the kernel version.
 - cartman OS can be changed or move the docker builds to a VM.

August 11, 2014

August 4th, 2014

- how to handle a single job wrapping around PMC
 - will add a property to turn the wrapping off.
- checkpointing for LIGO . synonym for checkpointing. user level state files.
 - create a JIRA item that explains that.
 - list the various cases that will be handled
 - a lot of times in case of eviction kill -9 is sent.
- pegasus dashboard changes
 - multi tenancy for users.

June 2014

June 30th, 2014

- pegasus-remove and pegasus-dagman. pegasus-dagman has a wait of 100 seconds before monitor is killed, when pegasus-remove is called.
- rafael will add a workflow test case for JDBCRC
- Still have to make a slider.
- Karan will work on XSEDE poster for Pegasus
- IPlant and metadata requirements.
- pegasus-dagman / monitor /condor-dagman
 - hierarchal
 - PMC
 - GRAM

June 9th, 2014

- 4.4 release
 - next week

- documentation items remaining
- JDBRC test cases and handover to SCEC
- Dashboard improvements
 - dashboard improvements
- Post Release Activities
 - integrate pegasus service back into the main codebase

May 2014

May 12th, 2014

- PM-747
 - will be used for soykb
 - test case
- Development releases
 - 4.4
 - plan for June 20th
 - automatic data dependencies
 - wrap up existing stuff
 - documentation
 - JDBRC change
 - documentation of FAQ's
 - 4.5
 - pegasus-service
 - some form of multi tenancy
 - python dependencies especially for external stuff is tricky
 - rename of dashboard database tables
 - pegasus-dashboard enhancements
 - separate the planning job from the prescript
 - checkpointing
 - software cleanup
 - transfers with hierarchies
 - leverage condor asynch transfers in pegasus lite
 - try for before christmas
 - 5 minute youtube video
 - 4.6
 - metadata
 - dax annotation
 - enhanced notifications
 - monitord
 - PMC data locality
 - globus online support ??
 - get credentials . at least do more research.
 - skipping symbolic links

May 5th, 2014

Condor week

- Lauren
 - Karan needs to provide more documentation for her
- Kent Wenger
 - dagman reporting
 - dagman metrics files is created by newer versions of DAGMan in the submit directory.
 - retry immediate parent
 - CMS has a requirement for this also. The most important thing on Kent's plate
- dynamic workflows
 - node expansion . may not be that worthwhile
- pegasus lite asynch transfers
 - using condor chirp in the pegasus lite shell script once the main computations are done. that way we can pipeline
 - does not work with partitionable slots
 - does not work with condor file io

Bamboo Test Cases

- Job got hung for a long time??

User Survey

- Developer Meeting will be moved to 1PM for

April 2014

April 21st, 2014

- - Pegasus Metrics
 - ewa sent out the report for metrics to Dan. we need to get her final version.
 - JIRA metrics
 - work log feature of JIRA - everybody does not find it useful.
 - all developers need to be diligent of putting tasks into JIRA
 - sub tasks in JIRA ???
 - how to track user feature requests
 - performance improvement
 - get the data structures upto speed.
 - timing the cleanup is also important and canceling it if it goes too long
 - SI2 Tasks
 - Support Data as first class objects
 - file movement open JIRA item
 - data flow dependencies
 - Support annotations for runtime and files sizes
 - software review of streamlined
 - remove pegasus-plots
 - remove libexec
 - remove unused example
 - archive sub directory
 - <https://jira.isi.edu/browse/PM-672>
 - tutorial VM's
 - refine and document metrics
 - we have the confluence page that captures
 - metadata registration in catalogs
 - triggers for enhanced notifications for long runtimes
 - we personally feel
 - pegasus service
 - have a release and multi tenancy
 - sort out all the python stuff.
 - reconsider moving pegasus-service back into pegasus git repo
 - documentation for integrating pegasus
 - enhance feature coverage and testing framework.
 - unit test coverage
 - adopt a model on how others can contribute to pegasus
 - document the process how people can contribute.
 - Customer Survey
 - identify questions to ask.

April 14th, 2014

- JIRA Policy Document or page
- [Pegasus Metrics](#)
- Pegasus Survey
 - Develop a list of questions .
 - Forward to Duncan CBC Group
- New Default Transfer Refiner - BalancedCluster

March 2014

March 31st, 2014

- Gideon changed the tutorial VM.
- Put in backward support for old credential handling.
- Mats started on an outline for the optimizations chapter.
- next week's developer meeting is cancelled.
- general Pegasus dependencies
 - python > 2.4 and less 3.0
 - in general, easier to build from source rather than from source RPMs
- update Pegasus README
- change the build.xml to say default build without docs. remove the dist-nodoc target. instead we will have ant dist-release as the default target
- also we should start having documentation per minor release and not per major release as we do now.

March 24th, 2014

- Pegasus 4.3.2 release done last week
- storage constraints paper - gideon, rafael and karan worked on it.

- karan worked on the hpc-pegasus setup.. has workflows running through PMC
- karan and mats have a XSEDE tutorial proposal that will be submitted today
- dv/dt paper rejected for HPDC. Will try for a middleware conference due mid may
- 4.4 release
 - checkpointing solution
 - leaf cleanup for hierarchal workflows
 - md5checksum option for guc transfers
 - we won't follow up on kickstart generating the checksums, but tracking checksums in replica catalog.

March 17th, 2014

Agenda

- XSEDE poster and tutorial proposal
 - will get it done this week. mats and karan will work on it.
- idafen will work on a workshop paper for xsede on reproducibility
 - 4 page limit
 - deadline is april 5th.
- energy simulation for SC 2014
 - measure energy when running workflows
 - try to check if energy usage changes whether data is transferred to a site, or everything is executed at one site.
- sane defaults for 4.4 for transfer jobs, pre scripts etc
 - transfer jobs
 - how many stage in jobs - 2 jobs and each job with 2 threads.
 - how many threads each transfer jobs - pegasus-transfer has a default to 2
 - pegasuslite job
 - change sls name ? property name change
 - control the number of threads
 - add a chapter called tuning workflows
 - mats will add about a section on tuning transfers.
 - setting clustering parameters.
 - changing back the default refiner to bundle???
 - cleanup job
 - change hold release time to one hour.
- new transfer refiner
 - maybe can use k means clustering ?
- leaf cleanup for hierarchal workflows
 - --cleanup leaf,inplace,none
 - tell the planner to throw a warning when
- sudharshan's paper
 - emphasize that the goal is not improving the makespan.
- 4.3.2 release
 - release notes checked in on friday
 - mats will tag after the release.
 - the service should be installed in the tutorial VM image.
- Condor Categories
 - similar to dagman categories.
 - will condor accounting groups work??

March 10th, 2014

Agenda

- Should we stage sub-workflow output files to parent workflow scratch? (related to leaf cleanup)
- Should we enable DAX jobs to have input and output uses, and distinguish between planner inputs and sub-workflow inputs?
- SUB DAG keyword to make pegasus generated subdag submit files match with dagman version always
 - From Kent, Wenger
 - Hey, I just wanted to touch base and find out whether you guys have made any progress towards making Pegasus-generated sub-DAG submit files match the "normal" DAGMan format.
 - (See <https://htcondor-wiki.cs.wisc.edu/index.cgi/tktview?tn=3891,4.>)
- data reuse edge case
 - have fix for it and have added unit test cases
- altassian licenses expiring?
- plan for a pegasus workshop / meeting for 2nd week of January 2015

March 3rd, 2014

- monitor fix for LIGO
 - pegasus plan prescripts were not logged in the database.
- checkpointing files
 - karan will create a JIRA item and send it to ligo folks for comment.
- transfer fix
- held jobs ?
- separate pegasus plan planning jobs

- throttle jobs via category.
- real full ahead planning
 - plan full ahead -
 - will help in debugging workflows
- hierarchal workflows planner arguments in the prescript wrapper shell scripts.
- final cleanup job for the workflow
- fix for iplant workflows cleanup. previously generated files whose locations are determined in the replica catalog should not be cleaned up

Workflow reproducibility (idafen)

- here for 3 months - march/april and may
- document the infrastructure that was used to generate the workflows
- created ontologies to describe infrastructure.
- precip API
 - expressed an interest in it .
 - he focuses not on how to deploy, but instead to describe the infrastructure
 - then do experiments that take in his description and deploy it using precept
- target two conferences
 - one systems
 - other semantic

Pegasus Submit Node on HPCC

- waiting on glite recommendations from condor-admin

Feb 2014

February 24th, 2014

SCEC Transfer Issues

- hpc login crashed for sceec workflows because of too many stageout jobs
- there were too many connections open at xinetd level
- also the stageout jobs were starving all the other local universe jobs in the workflows
- so the workflows were getting bunched at the stageout level
 - we solved it by moving only the transfers to the vanilla universe on shock
 - ran into credential handling backward compatibility we put in 4.4 after new credential handling.

Transfer Configuration for 4.4

- by default the number of threads will be 2
- we will expose a way via properties to increase the number if users want to have better bandwidth
- in case of any failures, pegasus-transfer will revert back on a single thread

February 10th, 2014

Postscript handling

- We have implemented a solution in PM-737 to get around condor quoting rules.

- MPI code are not kickstart wrapped

- Pegasus should indicate whether a clustered job or a kickstart job.

- DAGMan exitcode

checkpoint jobs

- 10% of runtimes

- pegasus-transfer will have to be changed

- link is set to type checkpoint

- transaction support for checkpoint

- timeout is job runtime - process
- pegasus-kickstart timeout method
- also has dv/dt implications for monitoring.

pegasus-exitcode assumes success and checks for failure

- refactored the script for unit tests as a library
- pegasus-statistics
- pegasus-analyzer (maybe some commonality)
- pegasus python library has to be included in worker package

pegasus-transfer

- threads are handled similar to pegasus-s3
- default threading
-
- expose options end to end
- initial threads to irods
- what options to set

pegasus-config will now work with a source checkout

December 2013

December 16th, 2013

- TODO: Talk about ADAMANT design

December 3rd, 2013

- 4.3.1 release
 - just need to send the announcement.
 - gideon has updated the build infrastructure in bamboo to build the release
 - to do
 - do a drupal snippet, to update the downloads page automatically.
 - dynamically render the page using the shared directory in drupal.
 - pegasus-analyzer will have a recurse option.
- identity management for pegasus service
 - portal use case
 - user authentications
 - website
 - put a token in a cookie.
 - draw bigger pictures on the identity stuff.
- Unicore Testing

November 2013

November 11th, 2013

- 4.4 Planning
 - according to proposal, we need pegasus as a service, metadata registration, enhanced notifications on long runtimes etc.
 - ligo realtime analysis?
 - scott and kent mentioned that real time analysis is a priority.
 - gstreamer interface.

- investigate streaming workflows
 - uncore testing support
 - Pegasus Tutorial on (Mats VM on oregon region)
 - Pegasus as a service
 - Ensemble Manager
 - an ensemble has no end state currently.
 - update documentation on the website
 - gideon plans to remove the upload catalog options. instead the clients will read in the properties and automatically upload.
 - NSF Cloud Proposal
 - Experiment management.... maybe does not align itself with NSF Cloud.
 - Adamant Demo
 - workflows are setup and done.

November 4th, 2013

- Tutorial format finalized for November 14th meeting. similar to software carpentry layout
- 4.4 release things
 - pegasus metadata support
 - dax schema changes
 - irods - support for metadata attributes
 - s3 objects - they can have tags associated with it.
 - transient replica catalog.
 - uncore support
 - for JIRA items move to the next one.
 - moteur support.
 - dv/dt wrapper support (probably in a separate dv/dt branch)
- move to VMWare for hosting websites
 - pegasus.isi.edu will be as a VM in a VMWARE ESX pool.
 - initially 4 VM's for Bamboo BNT
 - retire the machine for PAGE QC
 - long term we are moving to ESX

October 2013

October 1st, 2013

Pegasus 4.3 release

- dashboard is separate
- prepare rpm for ligo
- ssh submission for 4.3
- tutorial vm almost done
 - the clock issue remains. probably an issue with how virtualbox does the time.
- need to hear back from scott
- sepiddeh working on make flow compatible code generator.

September 2013

September 23rd, 2013

Software Carpentry followup

- Create a pegasus youtube channel.
- See if that can be linked from the ISI webcast page.

ISI Pegasus Workshop

- Submit host setup at HPCC
- specs are similar to workflow.isi.edu
- gideon will mail to HPCC admins today about this

Tutorial VM

- networking issue
 - persistent rules file /etc/udev/rules/70-persistent-networking.rules
 - instead of deleting it lets just disable it in our VM's
- X with virtual box guest additions for enabling copy paste
- turn on ntp
- larger virtual disk - will increase the size to 8GB
- X should just add couple of hundred MB's

Pegasus Release

- JDBC RC
- Tutorial VM

- pegasus-statistics
- pick up a release date
- tentatively next friday i.e the 4th.

September 9th, 2013

Software Carpentry

- Karan will prepare introductory slides for Pegasus.
- Talk to John about providing a Pegasus submit node.
- Rajiv will be working on the Pegasus RNASeq VM.
- John Mehringer will go first in the second day.
- Parking is in Levy structure in southwest corner.
- Inquire about shuttle from Health Science Campus.
- Still do - RNASeq module.
- Put Information about parking and HSC Shuttle.
 - Parking Center.

Pegasus Release

- waiting for Scott to do release testing.

Pegasus Lite Paper

- Karan will send the camera ready version today.

Precip

- using netlogger for logging.
- replace python logging framework
- incorporating events from the remote site
- AMQP ?
 - Getting events into a common file.
- Run montage using precip

Condo of Condos Workshop

- Laurent and Gideon have 10 minutes each.
- Bosco new name is MyHTC.

August 26th, 2013

Pegasus 4.3 release

- dagman metrics not implemented yet by kent. still in design phase.
- testing stuff
 - unit tests running in bamboo.
- add missing data dependencies
 - still checks and produces errors

Precip Logging

- getting the metrics back

Pegasus Hold

- how to get dagman stop submitting jobs
- idle jobs need to go on hold.
- we can send sigusr1 to dagman.
- need to handle hierarchal workflows.
- JDBC RC stuff

JDBC RC

- we will just update the existing version one.
- have a python based RC for Replica Catalog.

Ensemble Manager Paper

- Gideon will be working on it.

DAGMan replacement??

- Software engg stuff.

August 19th, 2013

- Pegasus 4.3 release
 - output mapper stuff implemented.
 - pegasus-statistics changes checked in by Rajiv
 - app metrics associated with the metrics report
 - pegasus.metrics.app
 - can be used for RNASeq tracking and other applications
 - the metrics UI will be able to filter on the name.
- Globus Online Support - move to 4.4 release
 - can only do certain parts of transfers.
 - for transfers from local submit host , we need to use globus connect
 - credentials issue
 - for submit host, there needs a local endpoint.
- LIGO testing ?
 - prepare a pre release RPM for LIGO

August 12th, 2013

- Pegasus Lite Paper
 - Wait for the Big Data and Science Workshop
- 4.3 Release
 - Output Mapper Submission
 - error if output site and a output mapper replica catalog specified
 - Globus Online Support in pegasus-transfer
 - OAuth tokens issue.. when to get the token
 - support for multi end point with different credentials
 - probably need to do a pegasus-globus-online
 - the client needs to be blocking .
 - SSH Submission
 - Will use RNASeq for that.
 - Boto downgrade worked.
 - did not build on RHEL 5
 - Test Suite
 - Suite of integration tests
 - checksum the files
- Ensemble Manager
 - Almost done with the first version
 - Will work on the Galactic Plane version
- General JUnit Tests for Pegasus
- Galactic Plane Paper

July 2013

July 29th, 2013

Software Carpentry

- Workflows Tutorial
 - 1 hours overview of HPCC if HPCC folks are interested.
 - Pegasus Tutorial (2 hours)
 - An info part on where to run jobs
 - OSG
 - HPCC
 - XSEDE
- Pegasus Development
 - Rajiv will complete the pegasus-statistics part
 - error messages (give more hints on what went wrong on site selection)
- Monitoring API
 - wants a jar with a simple API to monitor workflows
 - wrap it up in a jar
 - provide interface
 - portal integration
 - rest interface for the pegasus service

July 8th, 2013

- gideon has changes checked in dax2dot based on the closures and reductions
- karan has checked in the LCA approach. But does not scale for our performance test case.
- Also changed the way edges added for the create dir nodes. that will go in for 4.3.
- Precip Paper
 - deadline extended to the 19th of July.
- Posters to be made for XSEDE
- Sudharshan will make a poster on his cleanup work on Monday.
 - Sudharshan will be going on Monday to campus to present the poster around 1-3PM

- Will give a talk to CCG group Tuesday July 16th at 11:00AM
- Currently, sudharshan's algo takes 15 seconds on a 1000 node montage workflow.

July 1st, 2013

- monitor bug fix checked in
- algorithm to remove extra graph dependencies
- backups
 - we need to update the pegasus machine
 - jira, svn, website (website and svn need to move at the same time), crowd updates
 - confluence was moved to another . also coordinate with action to do the move.
 - mats already updated crowd today
 - there is secret number of conf files... apache on top of tomcat
 - update to debian machines
 - obelix, cartman and stewie, and the ccg worker nodes.
- mats has updated the bamboo tests to use new filesystem paths
- ADAS abstract
 - for galactic plane on Amazon. if accepted due in september.
- 4.3 release
 - fix error messages. see what can be done to improve them .
 - output replica catalog
 - pegasus-transfer tests.
 - updates to cleanup algorithm based on sudharshan's work ??
 - release notes will be updated to indicate the dashboards move to pegasus-services thing.
- Precip Paper
 - mats will do the zotero work.
 - submitting to cloud com in bristol uk.
 - seppideh has some data on openstack. could not get all instances started up.
 - seppideh will release the token to gideon to do an edit pass
- Cleanup Algorithm

June 2013

June 24th, 2013

- Pegasus Development
 - Monitor issue <https://jira.isi.edu/browse/PM-712>
 - karan has a fix that works for him.
 - needs to test the replay mode.
- Update on SCEC visit
 - pegasus-archive tool
 - archive everything other than the stampede db and braindump file
 - scott will try to cluster rupture variations for the same rupture in one task based on runtime estimates
 - the SGT will become 16 times bigger and post processing 8 times bigger on move to 1HZ. clustering rupture variations in scecc code will help in reducing the number of jobs in the DAX
 - Scott tried to generate a single DAX for the post processing workflow. Was unable to do so. Has generated two dax'es
- Galactic Plane
 - Cut out service. Slow times on retrieving the image from S3. Small bandwidth between S3 and EC2
 - Will need to have monitoring etc... Not fast enough for a webpage to be responsive.. will need some queuing up
 - Backups
 - Mats working on Kepler data.
 - mats tried backup with S3. does not like symlinks. will change the way backups are managed. the transfer times can be long.
- Update from Sudharshan
 - Good progress. showed some simulations
- Adamant Update
 - we are on hook for providing the interfaces in pegasus-transfer that will talk to the exo planner service
 - also provide shadow queue service, that gives estimates on jobs that will be in the queue.
 - supercomputing demo?
- Precip Paper
 - majeick si doing some experiments

June 17th, 2013

- Pegasus Development
 - the dax job handling is completed.
 - update on ligo front.
 - condor priorities for local universe jobs
 - not handled right now.
 - gideon has a ticket open for them.
 - gideon observation of s3
 - scalable but not good latency or
- Pegasus Lite Paper
 - mats is almost done with the runs. to grep through the runs to get the intermediate files in and out of S3

- not done the S3 caching for rosetta as yet. still not sure. too much work for the time remaining.
- mats did do the runs with task clustering. he got better numbers and saw a difference in case of rosetta.
- interleaving of compute jobs and transfers. may help montage.. but won't help rosetta
- whether we should include the new pegasus 4.2 features.
- Cleanup Algorithm
- Glacier Backups for NFS?
 - instead of using two qnaps, just have one and use other for duplicates
 - we need a place for backups
 - currently the QNAPS are 18TB each with raid 6. Raid 10 is a better configuration on the QNAP according to the forums. This means though we will have half the space.
 - have one qnap for scratch
 - have other qnap for storage - the storage will be backed upto glacier. right now QNAP only support S3. Support for glacier is coming.
 - ewa and richard think glacier backups are a good option.
 - there might be a purge policy required on glacier.
- Precip Paper
 - change tracking on
 - use dropbox
 - broadcast when you making a new version.

June 10th, 2013

- Pegasus Development
- change to dax handling
- fix of stdout
- regex based replica catalog.
- changes to pegasus-statistics for aggregate statistics

Pegasus Lite Paper

- compute data between s3 and local disk.
- compute costs for the runs ?
- have data outside
- local cache for the S3 client ?? could affect the rosetta cache.
- change the rosetta workflow.
- if there are a lot of small files.
- reading parts of files.
- Ewa will send her version of the changes.

Sudharshan Algorithm for Cleanup

- Greedy approach planned
- will try implementing a version and show the different executable workflows created

June 3rd, 2013

Pegasus Lite Paper

- Breakdown of the runtimes , experiments
 - In case of sharedfs, the kickstart runtimes in the breakdown file will be longer
 - for the S3 case we can calculate the S3 transfer time by calculating the difference between the cumulative runtimes
 - doing two experiments rosetta(cpu intensive) and montage(io intensive)

Pegasus Development

- Java DAX API issues
 - might be some bugs in there.

Precip Paper

- Ewa wants a link to pegasus website in the paper.
- have more logical thinking in the paper, like reliability and repeatability

- Sepideh adding some new figures to the paper.
- Maciek will provide an experiment use-case for the paper.

Stampede and Corral Annual Reports

- Karan and Mats will be working on these

Sudarshan's Project

- Going to look into providing a cleanup algorithm that meets a given storage constraint
- Will look at the static problem of inserting dependencies into the workflow to achieve a solution

PMC Paper

- on amazon
- with clustering and pmc

Shirts

- Should get the logo sample this week, once we approve then we can order shirts

dV/dT

- Rafael is working on a draft of the data collection and modeling paper
- We are planning on publishing data, will start drafting a format this week

May 2013

May 20th, 2013

Confluence is going slow. Mats is going to look.

Analytics are set up on Confluence now.

Pegasus Transfer

- Mats committed a new version that has support for 2-stage transfers

Pegasus S3 Client

- Gideon changed .s3cfg to .pegasus/s3cfg

Pegasus Lite Paper

- Mats is working on the experiments
- We have two weeks to the deadline

PMC Paper

- Experiments on Amazon comparing Pegasus, Pegasus w/ Clustering, PMC alone

Pegasus Service

- Finished setting up users and test suite
- Next is a quick-and-dirty ensemble manager implementation
- Gideon is going to commit a change to Pegasus that removes the dashboard components. They will live in the pegasus-service repository from now on.

Summer Student

- Need to think up a project. Needs to be research-oriented and relatively small.
- Cleanup? Precip?

Contacting users

- Find out if they need anything.

Examples

- Simple examples in Perl, Python and Java
- Gideon will add them to the examples in the pegasus Git repo

April 2013

April 22nd, 2013

Pegasus 4.2.1 Release

- monitor prescript handling fixed
 - pegasus-analyzer should detect prescript failures, and the prescript exitstatus should be logged in the database
 - pegasus-statistics was updated for the job instance report
- pegasus planner
 - need to confirm all checkins are complete
- do we want to get LIGO to do a test or just release?

Pegasus statistics across workflows - Rajiv

Pegasus Lite Paper

- Mats will do the runs on Amazon
- Karan will work on paper when he comes back

pegasus-hold and pegasus-release

- any difference between doing a hold on the dagman directly or pegasus-dagman
- we need to do more investigations on monitor

BOSCO

- Mats is trying to run on HPCC
- a single job is running fine.

April 8th, 2013

Pegasus 4.2.1 Release

- Work on it towards this week
- monitor prescript issue to fix

Pegasus 4.3

- pegasus statistics extensions
 - across root workflows
 - <https://jira.isi.edu/browse/PM-507>
 - condor temp file

Pegasus Posters

- One at XSEDE
- joint one with BOSCO team

Pegasus Lite Paper

- Submission to IEEE Big Data

New Programmer Hire

- expanded posting on confluence
- [New Programmer Hire](#)
- will send out to HPC Wire , RENC1 and USC SC Connect

April 1st, 2013

Pegasus Lite Paper

- Waiting on Ewa
- Not much we can do about the IEEE conference. The page limit is 8 , the current size of the paper.

XSEDE Poster

- Pegasus Poster. Karan will send update
- Also a joint Pegasus BOSCO poster
- Also as part of that we will get the MPI workflows up and running through Pegasus and BOSCO

Pegasus Development

- Bypass of staging input files for Pegasus Lite Case
- Inplace cleanup bug fixes done.
- pegasus-s3
 - gideon checked in changes of copy from one file to another
 - mats adds a pegasus transfer
- workflow cleanup nodes
 - separate cleanup node in the workflow
 - for hierarchal workflows we only delete the outermost workflow
 - what happens if no output-site specified
 - the ligo case!
- backward compatibility for LIGO

- Pegasus Dashboard
 - general javascript updates
- Generic Pegasus Slides
 - 2-3 slides.

March 2013

March 25th, 2013

- Pegasus Lite Paper Submission
 - We will try for <https://sites.google.com/site/sweetworkshop2013/>
 - Karan will move the paper to the ACM format
- Pegasus-statistics
 - Waiting on Scott to get back with the list of metrics
 - Rajiv will be working on it
- pegasus-s3 changes
 - we want to be able to copy output files from one s3 bucket to another
 - requires changes to pegasus-transfer and pegasus-s3
- final node for cleaning up remote directories
 - also related is getting the cleanup algorithm working when we bypass first level staging.

March 18th, 2013

- Mats has an RPM almost sorted out for LIGO that does not require us to have PYTHONPATH set. Instead the libraries go into standard locations
- Karan is testing this RPM at on spice-dev1 and has setup a page with instructions on how to submit a test workflow to VIRGO
 - <https://www.lsc-group.phys.uwm.edu/ligovirgo/cbcnote/aLIGOaVirgo/130307121559GeneralInstructions%20For%20Submitting%20Workflows%20on%20VIRGO>
- Statistics across root workflows
 - earlier gaurang had generated statistics for scec runs by hand... executiing queries on the msql command line
 - he does not have the queries documented anywhere
 - this is something we have talked about in context of 4.3 with Rajiv
 - will follow up with scott on wednesday's call
- 4.2.1 release
 - backward compatibility for LIGO . still to be done
 - probably next week after the pegasus annual report
 - RPM to handle native python installation
- Pegasus Annual Report
 - Karan will work on it this week
 - Try to follow the same template as earlier.

March 4th, 2013

- Sent link on DAGMan metrics to [DAGMan Metrics Reporting](#) to Ewa
- Metrics for Rob Quick's workflow
- Gideon pushed out kickstart changes
- Rajiv has pushed changes to the queries for the dashboard.
- Setup meeting with Jaime and Derrick at OSG AHM to discuss
 - remote_initialdir
 - extra attributes for glite/bosco submissions
 - mpi workflows.
- OSG Poster to be made this week. And 4.2 Release slides.

February 2013

February 11th, 2013

Direct submission of workflows to PBS

- Glite submission in Condor. We setup a VM that hosts a PBS scheduler and using that too test
- Karan prepared an example for 4.2 that can be used to submit directly to local PBS using the glite interfaces in Condor
 - the remote_initialdir / +remote_iwd does not work
 - problem for MPI codes
 - for the time being, the example prepared relies on kickstart to change the directory before launching a job
 - there is also a ssh style that allows us to use BOSCO to do remote submissions using SSH to a PBS cluster
 - that one also has the issue of remote initialdir

- jobstate.log refactoring.

- data transfer (support for globus online)

- lightweight tracing
- task stats. net link socket pegasus-kickstart . how much memory the task used and io used.
- add task stats to kickstart
- ptrace
- trace linux equivalent is system tap

- dashboard improvements
- single api for clients
- last week drop down
- performance run on large workflows.

February 4th, 2013

- CCGrid / Pegasus Lite Paper
 - Performance section
 - remove the experiments section?
 - OR
 - extra experiments section
 - have the squid proxy cache
 - find a workshop to submit the paper
- Cloud Paper
 - Ewa is working on it.
- Git HUB Migration
 - - couple of branches like monitord , pmc and dang are branches
 - - svn will be made read only .
 - - update the website with all the development information
 - - bamboo scripts
 - - documentation (long term)
 - - nightly builds
- SSH Submission
 - - gsissh submission for blue waters
 - - ssh to blue waters is required for OTP
 - - passing of parameters to PBS
 - - SSH key
 - - ssh agent.
 - - queue keyword
 - - Batch session
 - - submit jobs to HPCC
 - - Gideon will do that.
- monitord memory explosion
 - - long term for monitord
 - - pegasus-dagman replacement
- minor release 4.2.1
 - - potential monitord bug issue
 - - long term dagman replacement
- Response time for metrics page
 - - occasionally it is slow