Meetings

3/29/2012 Teleconference

Attendees: David R., Taghrid, Karan, Guarang, Ahmed, Dan G.

Topics

Dashboard

- Change invocation to use submit directory
- ISI: Add Python module to find database from braindump module

  e.g. from stampede.util import find_db

  dburl = find_db("/path/to/submit/dir")

- Feature requests:
    - Troubleshooting: invoke pegasus-analyzer? Add "see failed jobs" button.

    - Highest priority: Debugging(https://pegasus.isi.edu/wms/docs/4.0/monitoring_debugging_stats.php#monitoring_pegasus-analyzer)
    - Pie chart(http://pegasus.isi.edu/workflow_gallery/gallery/brain/run_1/82bf135e-ce7d-417d-9062-f082134e0347.php#inv_chart_div)
- Potential beta testers:
    - Stephen Cox, RENCI. Also has his own form of dashboard.
- To-do:
    - Get dashboard running through the distribution.

    - Submit directory
    - Troubleshooting
    - Link it to distribution (need release)

- Live dashboard?
    - Install Condor and have pegasus running on a machine at LBL.

    - CentOS; yum

    - All runs in workflow gallery are 3.1 (no host information)

Triana integration

Triana (http://www.trianacode.org/)

- Yang schema was important for getting things to work.
- Triana has a "streaming" model. What gets counted as a job retry is streaming of the work.
- Going to run some workflows through pegasus and their cluster.
- Going to test-drive the dashboard? Didn't talk about it much.
- Paper plans: we will see.

FutureGrid status

- re: India, Need a window when we are going to run our experiments.
- Add monitoring stuff (blipp)
- Space issues on submit VM – mount scratch drive; not done. For these experiments we don't need a shared file system.

2/21/2012 Project Meeting

- **Attendees** Ewa, Karan, Dan, Taghrid, Gaurang, Fabio and Martin

  Stampede All Hands Meeting

  Pegasus 4.0 Update
    - pegasus analyzer works with the stampede databases
    - pegasus statistics changes
      - csv format output
    - stampede database
      - new job instance
    - new schema upgrade tool to migrate users from 3.1 to 4.0 schemas
      - minor changes still required to be done.
    - nothing new on pegasus plots
  - **Future plans**
    - have a sort of a provenance trail/ point out which DAX job failed by pegasus analyzer .
    - replay workflows… more at pegasus level than at monitord level
    - dashboard to be integrated in 4.1. Can be released separately also.
  - **Stampede Dashboard**
    - Chris Brooks student also was working on the dashboard.
    - She will go through and build a whole lot of performance tests for the API.
    - David at LBNL is the main person working on the dashboard.
    - David is using JQ Plot.
    - Karan will run a real life LIGO workflow and tar it up and sent the workflow files to Dan
    - David can use it for his development runs.
    - Pointer to the stampede API
      - https://confluence.pegasus.isi.edu/display/pegasus/Pegasus+Statistics+Python+Version+4.0
    - Pointer to pegasus plots for David
      - http://pegasus.isi.edu/wms/docs/3.1/monitoring_debugging_stats.php#plotting_statistics
    - Long term maybe we can support uploads to a restful service
    - automating the dashboard via monitord or pegasus-run.
    - how to integrate multiple sqlite databases in a single dashboard instance.
      - maybe when the workflows start, a simple rest request is sent to dashboard tell what workflow to monitor
      - like a global sqllite db that resides in $HOME/.pegasus , the registers the sqlalchemy string.
      - dashboard should start with a pointer to the sqlalchemy connection string that has main sqllite db that has pointers to all the workflow db's.
      - Dan and Martin feel that it should be a web service call
      - Move dashboard to Github to coincide with Ahmed's work on periscope
      - Users may have to start up dashboard separately if we want dashboard to track multiple sqllite databases.
  - **Periscope Demonstration**
    - Some visualization of experiments on FG
    - Will require tornado instead of web.py that is being used by the dashboard.
    - backend to the dashboard , can be be made compatible with the periscope stuff. will talk to stampede database for workflow information and periscope for network monitoring. Periscope is works better with tornado??
    - Dashboard should be released before the periscope work is stitched into it
    - monitord may send rest messages to periscope brokering service.
    - Populate from FG runs the periscope information brokering service/location service.
    - What does TG use for monitoring.?
    - In general Periscope maybe able to forewarn about the failures.
  - **Online Analysis Work**
    - all of it is written in R.
    - works against the database
    - does not use the stampede api
    - looking at the rate of job failures predict whether a workflow will fail or not?
    - built probability models for different job types in a workflow.
    - online analysis will work on the same place where there is dashboard.
    - How does pegasus integrate with online analysis work? user notifications?
    - In the near term, we need to upgrade the analysis to generate some user notifications
    - Placeholder for using this, can be either in site selection or in retry hooks.
  - **General**
    - Ditched AMQP
  - **To Do**
    - Integrate dashboard with tornado work.
    - Move analysis tools to work online.
    - Analysis tools should be using the API.
    - figure out how much space can a VM get
  - **Website / Stampede Organization**
    - Organize the Stampede wiki page by the artifacts discussed below.
    - Update stampede papers on the WIKI.
    - Stampede Page should have the stampede logo.

- - WF prof should go somewhere else?
      - http://pegasus.isi.edu/documentation/wfprof
  - **Futuregrid**
    - experiments targetted for SC 2012
    - space on VM
      - m1 small - does not have much disk space
      - m1 large - also varies from india and sierra
    - place where the vm's are stored in FG have about 80GB's
    - in eucalyptus there might be a way to externally mount a filesystem using bit bucket?
    - submit host in VM
    - first set of experiments
  - **Experiments**
    - ISI Failure once it starts (wrong executable) / change DAX .
    - task failure after start/ kill process (Martin)
    - task failure (Martin)
    - task failure / hang suspend process
    - Data Transfer failures . ISI
      - jobs don't bring the data back/ write out the data
    - Host shutdown (Martin)
    - job monitor failure (Martin)
    - network issues
    - disk filling up .
    - Futuregrid does not have a storage service.
      - Martin has some physical nodes that we can get
    - On India
      - eucalyptus on xen
      - openstack on xen
      - every month there is a maintenance in FG . First Tuesday every month.
    - Larger story for the paper
      - stop early . early detection of the failures.
      - induce failures, models of failures.

  - **Papers**
    - SC 2012 paper for the FG experiments
    - Journal Stampede Paper (completing the picture and putting various parts together)
    - Special issue of workflows due May 1st.
    - How does the stampede work relate to provenance ? Maybe a future paper in reference to that. Martin is working with Beth on a similar thing. Performance and Provenance interlap
  - **Future Plans**
    - Workflow Gallery - make it available. Useful for others to use? everybody should look at the wiki.
    - do instrumenting and include it in kickstart . tie it down with work martin has been doing. Will be useful for workflow debugging and re-running
    - some instrumentation at compiler level? sort of pegasus make . Potentially a lot of work in that.
    - do experiments on FG where data is staged using gridftp instead of condor io . martin has accelerators that will increase the bandwidth on internet2 or calit
    - **Artifact from the stampede projec**t
      - Dashboard
      - General SQL schema
      - Formalized the netlogger yang model
      - Analysis functions / Stampede DB API
      - Anomaly Detection work ? Can that be resurrected. Priscilla did not transfer. So not sure.
      - Connection with periscope ?
      - Also periscope is integrated with gridftp. We should use it.
  - **For wei**
    - how to adapt scheduling of workflows, as errors happen, diskspace used up or hosts problems.
  - **Funding opportunities.**
    - At LBNL a lot of dynamic workflows from a single domain.
    - DOE oscar - more for research not for software development
    - opportunities for application tie in for dynamic environments
    - potentially for NSF . STCI will not work. It will be more experimental. More to go in for Si2
      technologies for processing - extreme scale science requires multi scale and multi processors . some doe workshop.
      - potential tie in with ann data scheduling stuff
  - **From Dan Email**
    - 1. Finding: Collaborative extreme-scale science requires a range of services to manage the collaboration itself. These may include machine-accessible lists of collaborators with their roles, authorities and duties, project planning tools, effort and financial reporting services, meeting scheduling and management services. The lack of uniformity, and indifferent quality of many of the existing services is a major impediment to collaborative science.
    - Recommendation: Develop Office-of-Science guidelines for centrally supported services, with special focus on the Software-as-a-Service mode of delivery. The central support and any essential research and development, may be provided by ASCR in concert with science programs or by commercial services where appropriate.

12/1/2011

- **Attendees** Ewa, Karan, Dan, Taghrid, Monte and Martin
- Taghrid working on a proposal . 500K? over 3 years. Taghrid will send a version to Ewa to look at by end of next week.
- Doing some experiments on Future Grid
  - Dan needs to get an account on FG or already has
  - Dan says Taghrid should get an account.

- - Jens runs on Future Grid. Periodogram workflow on Future Grid.
    - Dan has some modifications to kickstart in mind.
    - He will put in a request in Pegasus JIRA .
  - Paper Deadline
    - List of experiments to be done.
    - HPDC too close . Maybe some workshop at HPDC that has a later deadline.
    - SC submission is a possibility (April 27th)
  - Status of the dashboard
    - Same as what was shown in SC.
    - Stephanie will be doing some scalability tests against the database API. Write a performance API test suite.
  - Integration with Periscope
    - Dan is interested in doing it. But not sure whether LBNL will have time to do it.
    - He feels the right time to look at periscope will be when they start doing runs on FG.
  - Follow on to Stampede Proposal
    - Ewa and Dan want to do something.
    - STCI proposals are now pushed into Si2.

3/3/2011

Attending: Dan, Taghrid, Chris, Fabio, Karan and Prasanth

Agenda

- Stampede DB Changes for 3.1
  - https://confluence.pegasus.isi.edu/display/stampede/Stampede+DB+Changes+for+Pegasus+3.1
- AMQP Incorporation
  - Taghrid and Dan are working on the paper. She's updating the analysis and Dan has monitord streaming messages to AMQP, optionally binary-encoded using BSON. We should plan a transition to this version of monitord. As I mentioned in an earlier email, the command-line options have changed somewhat. With that exception, all current functionality (as far as I know) remains unchanged. Is there a test suite? Examples of current and new options attached.
- Galactic Plane Plotting

Notes:

- Fabio to work on a test suite for monitord .
- Fabio is ok with monitord amqp changes.
- Prasanth showed the plots
  - Chris suggested ticks every 100 for larger workflows etc.
  - Chris will give more feedback
- Dan asked about the service API for the plotting dashboard. Chris will be working on it.

7/29/2010

Attending: Dan, Chris, Gaurang, Fabio, Priscilla, Gideon, Karan and Prasanth

- Fabio has incorporated latest changes from Monte. Still to check in code to the SVN
- Right now tailstatd does not work on restart of workflows.
- AMQP incorporation. Replacement for the broker protocol. Fabio and NL is going to support both in tailstatd and netlogger.
- Filtering would be incorporated in the nl_loader end for time being instead of broker.
  - Dan says should be easy in Mongo DB loader.
- Dan might call in to one of the FG sessions at TG.
- Chris says no easy way to figure out if there are multiple workflows in the DB
- Super Computing Plans
  - Dan says demo and talks in the Berkeley and USC Booths.

To Do's

- Gaurang will modify pegasus-run to use new version of tailstatd
- Karan and Fabio to finalize the braindump entries
- Prasanth will send a link to the current interface talking to the Restful API.
- Prasanth will get in touch with Chris students for the REST API update.

7/15/2010

Provided by Dan

Attending: Dan, Chris, Gaurang, Fabio, Priscilla, Gideon (I think)

Agenda items:

- deployment (svn) integration
- more datasets to Monte
- command line tools using new API
- tailstatd progress and plans
- Priscilla's progress on analysis
- annual report

Notes for each item:

- for netlogger, resolved to use svn link (to trunk), update for nightly pegasus builds and specific (or tagged) revision for releases.
  - to-do: For Priscilla's code, she is currently not using svn and will consult Martin on whether to check in under pegasus or use a similar procedure to netlogger

- more datasets exist, but untested and suspicion is they'll break tailstatd.
  - to-do: This is mostly a debugging task for ISI team, but any working datasets can be forwarded to Monte for correctness and performance testing of DB.

- ISI team wasn't aware new API was ready.
  - to-do: Monte will send out docs, and they will attempt to get the "status" command switched over as an initial test.
  - to-do: There was a question as to whether the "percent done" result of the status command could be properly calculated with current database, this question will be answered (by ISI team).
  - to-do: ISI team will also look at whether "percent done" is a common metric that should be included in the API – and try to think of other metrics that it is worthwhile to bake-in to avoid repetition across tools.

- two things have to be done for tailstatd: (1) test & debug against "other" cybershake and ligo runs, (2) add ability to handle sub-workflows.
  - to-do: Resolved to do them in that order, given that (1) is a bit of a blocker for Monte to test new datasets (whereas without (2) he just has to run manually per-subdirectory).

- Priscilla has working code, but not clear on how to tie it into the rest of the project. Currently reading out of a database directly. This needs further discussion.

- Annual report is due soon.
  - to-do: Fabio will send out new version of annual report tonight (Thursday).
  - to-do: Dan will work on adding his sections by next week
  - to-do: Chris volunteered to do an edit pass

7/01/2010

Attendees: Priscillia,Marcos, Fabio, Karan, Gideon, Monte and Prasanth

Agenda

- Update from Fabio and Monte on the runtime DB population
  - Successfully finished it.
  - sqllite db location - by default in the submit directory. It will be derived from the workflow basename
  - location of sqllite db for workflow of workflows
    - not clear as yet. there is merit for both approaches.
    - one per sub workflow
    - one per workflow of workflows

- checkin of netlogger code in pegasus svn
  - current version works for pegasus.
  - will check in incrementally only.

- update from Prasanth on using the REST API
  - Chris was not available. Prasanth to send an email.

- Annual Report -
  - Dan on vacation. Chris is not on the call
  - Priscillia will look at the document that Dan sent for NSF reporting and compose the email for Martin's group.

6/03/2010

Attendees:Dan, Monte, Priscillia, Fabio, Karan, Gideon and Gaurang

Agenda

- plan for a db-agnostic Python API for retrieving workflows, jobs, and tasks (Monte)
- adding a "task.id" to the task events in the logs
- eScience paper deadline: July 16th. I'm on vacation starting June 28th, returning July 14th.

Note

- The DB API is different from the REST API
  - What is the difference between REST and Python API ? To be ironed out.
  - The REST API will be used by the GUI.
  - The REST API should be similar to the python API.
  - Monte will do this after the tailstatd and loader have been integrated.

task id

  - It will be renamed to task_sequence_id
  - identified the task number in the job.

- will be of numeric type.
- PRE SCRIPT and POST Script will have predefined numeric id's -1 and -2 respectively.

Changes to the schema

- - Added a workflow state table
  - task_sequence_id in the task table.

Paper Deadline

- - Hard to make the E Science deadline.
  - Maybe target CCGrid and HPDC

- For paper what sort of experiments we need to do
  - SCEC run
  - Maybe LIGO run?

Milestones

- - In July we work out what sort of runs.
  - By end of June we will have tailstatd and netlogger integrated to start doing the runs and collecting data.

Anomaly Detection

- - do some soft failures.
  - make the filesystem unreadable.
  - things taking too long to finish?
  - SCEC outliers?
    - There were some workflows that ran faster.
    - Possible causes were that ruptures were larger than others.

Anomaly Detection Code

- - The Grid 2007 code that was written for anomaly detection should be written again in python
  - Longer term Dan wants to replace Broker with a message bus using AMQP

5/20/2010

F2F meeting at USC
Attendees: Dan, Karan, Gaurang, Martin, Chris, Fabio, Marko, Prasanth, Gideon, Monte, Ewa

Updates

- Tailstatd Progress (Karan, Fabio)
  - Linking Pegasus SVN to Netlogger SVN: Instead of including a checkout into Pegasus SVN we should link the two and track one of the Netlogger releases. That way we can just change the tag we are tracking to upgrade to a new version of Netlogger.
  - Discussion of adding information to logs about hostnames and files. Currently tailstatd does not do this, but it is needed to test NL loader.
  - We would like to have tailstatd output data directly to sqlite. Monte says that when loader is complete the integration should be trivial.
  - Still need to track multiple workflows, handle wfs without kickstart, handle rescue dags and reruns/restarts.

- Pegasus-Analyzer (Karan, Fabio)
  - Command line tool for troubleshooting. Uses jobstate.log. Provides output helping to diagnose issues.
  - This was a long-standing request from LIGO.
  - Can be migrated to using the STAMPEDE database when it becomes available.
  - Also helps us determine what users want to see out of database.

- Pegasus GUI (Karan, Prasanth)
  - We stopped developing on this for now because we want to integrate with the STAMPEDE RESTful API.
  - Trying to imbed Flex widgets in GWT. Would like to get some of Jim's code and try to integrate it.
  - Discussion about using GWT for GUI. Chris and Dan don't like it because it requires too many skills and the project doesn't require all the stuff GWT provides. Pegasus uses it because we have Java skills, but not much PHP or Python.
  - Walkthrough of GUI.
  - Discussion about how user can view submit/log files through GUI. Requires GUI to run on submit host. Not required as long as GUI can show the user exactly where the files are.

- Netlogger Update (Dan)
  - Added nl_broker
  - SQLite, MongoDB, and PostgreSQL loaders working
  - Want to have tailstatd connect directly to nl_loader
  - Looking to add support for RabbitMQ in the future
  - Updated workflow metrics page on wiki with SQL queries. The queries are more-or-less instantaneous.
  - Discussion about connecting tailstatd directly to nl_loader
  - Discussion about loading performance. Should be faster than workflow.
  - Discussion about query performance. Need to do experiments. Need indexes, but indexes slow inserts.
  - Also added a perfsonar interface to netlogger

- Anomaly Detection (Martin)
  - Have been working on anomaly detection code

- How to query anomaly detection from GUI? Anomaly detection data should go somewhere. It is generating events. The events can go into a database.

- Location Service (Martin)
  - Maps guids to metadata about the workflow. Allows you to query for information about what is running where. Gets information from broker.
  - Discussion about how this fits in the architecture. Is it an analysis service? Output goes to users, and also to other services (e.g. the anomaly detection service).

- USF Update (Chris)
  - REST API 1.0 completely finished
    - Backend works with MongoDB and emits JSON (would like to return XML in the future)
    - Need to update backend to work with SQL DB
    - Thinking about adding xpath-like queries to API (Martin: is similar to perfSONAR metadata service. We should use that design.)
  - Two students working on FLEX objects
    - pi chart, stacked area, working on bar chart
    - stacked area works with realtime data (needs units)
  - Discussion about queries again (MongoDB vs SQL)
    - Need to compare performance of both using some example data (Dan: maybe write a paper on it)
    - Selecting example data (Gideon has broadband, montage, and epigenome)
    - Selecting representative queries

Reporting

- Need to send out reports to funding agencies soon

Outstanding Issues

- Changes to stampede schema
  - Not tracking when workflows start and end
  - Job duration is not an integer
  - Workflow parent->child relationships
- Granularity of SQLite db files
  - One per submit dir, one per workflow, on per sub-workflow?
  - Discussion of mapping of databases to workflows and users
- Handling rescue DAGs
  - What is a rescue DAG?
  - What is the issue? When a user submits a rescue dag, is that a new workflow? (NO) Need to skip old events to prevent duplication.
- Issues in parsing workflow data in tailstatd
- Timestamps in NL events
  - ts is currently ignored, some events may not have timestamps available
  - Discussion to sort out all the issues about what events correspond to in the real world
  - How to map real events to netlogger events
- Job.map is not loaded. Should we add that to Stampede? (Not right now, but keep it in mind)
- Discussion about getting events as they happen by sending them directly from the worker node.
- Discussion about changing event names to match actual events
- What is the archive going to store? Raw events, or processed data?
  - Martin needs to get raw events.
  - The broker can get the events from tailstatd and rebroadcast them to loader, log file, Martin's stuff, and anything else.

Near-term Plans

- Martin: Anomaly detection, start+end event matching
- Need to test: tailstatd->broker->loader->db and db->api->gui
- Including Python and other things in Pegasus to get it installed for LIGO
- Get some users to help us test: Bio people, Brian, Kevin
- Paper for e-Science, poster for SC?
- For Paper we will be running the SCEC workflows
- We need to have the tailstatd parsing multiple workflows.

Broker Architecture

- Tailstatd will send a stream of events to broker.
- Subscribers subscribe to the Broker with the level of information they want.
- There needs to be a client for subscribing to the broker.
- Broker should collect some metrics about events. Not clear if broker generates events or we have a subscriber.
- nl_loader may send heartbeats back to broker or another broker for metric/instrumenting purposes
- Any monitoring of Stampede Broker etc will happen via Nagios.
  - Nagios by default provides load information, what services are up, whether webpage responds.

Relation to Future Grid

- Plan to use in the INCA deployment
- Inca will generate netlogger events and push it via the broker into netlogger.
- ISI deployment and configuration of netlogger and broker etc.

4/07/2010

Face to Face meeting at USF
Attendees: Chris, Monte, Jim , Karan, and Chris students

1. Jim has a restful api sitting on top of Mongo DB that can answer question about state of the workflow. The API can be used to populate the GUI that Prasanth has working at http://butterfly.isi.edu:8080/StampedeX/
   a. The API right now only talks to Mongo DB backend. But Jim and Monte will work together to present this API as the API for any backend. Hence, the GUI part is independent of whether the data is loaded in sqllite/mysql or mongo db.
2. Chris team have a mockup of what they think the dashboard should look like. They will send out the details of it. We then can decide how Prasanth's GUI needs to be updated to incorporate that functionality/layout
3. Chris team will be developing charting solutions in Flex. Chris and Jim feel flex is good for visualizing data especially for the larger scale workflows like SCEC has. As far as possible we should try to have a single dashboard, with embedding flex objects in GWT where required rather than two dashboards being developed. One in GWT and one in Flex
   a. An open question is about embedding flex objects in GWT.
      It looks like it is possible
      http://lgrammel.blogspot.com/2010/02/i-am-working-on-project-that-integrates.html

Action Items

- Fabio and Monte to get the live population of the sqllite DB working i.e the DB gets populated while the workflow is running.
- Document the Restful API. Jim will send details out when he is done. We can use this as the first version of the API, and suggest additional queries that we want the API to handle.
- Karan will send out the new stampede format logs to Jim, so that Jim can tweak his code to parse the new netlogger events. He has been till now using the SCEC runs from last year.
- Once Jim and Monte are agreed on the API , Prasanth can start using the API to populate the trouble shooting interface. Initially the data may reside on the Mongo DB. But that is temporary till Monte and Jim have the API talking to the other backends
- Prasanth and Chris's team will be exploring how to embed flex in GWT.

2/25/2010

Attendees : Ewa, Dan, Chris, Martin, Gaurang, Karan, Fabio, Raphael, Gideon

Gaurang suggested STAMPEDE : Simple Tools for Archiving, Monitoring Performance and Enhanced DEbugging

Fabio to upload tar/txt file of events to the Wiki. Send netlogger events to socket..

First cut to try doing sqlite uploads from listener.

Change the way tailstatd writes to Local DB. Use Netlogger api to write netlogger events and then write to Mongo-DB or SQLite.

Stampede NMI Builds

Get account for DAN at ISI and NMI

Write ant scripts to pull netlogger pieces from SVN and add to pegasus builds..F

Do benchmarks on the SCHEMA to run queries efficiently.