PDAC Meeting 2016-11-03

Date

03 Nov 2016

Attendees

- Gregory Dubois-Felsmann
- Kian-Tat Lim
- Unknown User (xiuqin)
- Brian Van Klaveren
- Unknown User (jalt)
- Greg Daues
- Trey Roby
- John Gates
- Fritz Mueller

Goals

- Review progress toward PDAC completion at the end of the month

Discussion items

Time	Item	Who	Notes

	NCSA / infrastructure	Unknown User (jalt)	• Going to announce on Monday that `lsst-dev01` will be in full operation in two weeks
			• Looking at the sizing model this month, looking for ways to improve it
			  • Fritz Mueller: Andy Hanushevsky has been looking at it too and may have some suggestions
			• Looking at DRP/Princeton needs for QA
			• Discussion of handling 100 sq. deg. of HSC public-data
			  • Comparison of this to the Data Access group's need for further scaling tests
			  • QA system appears to need the interactive creation of user databases
			  • Discussion of whether the QA system can be hosted on the same cluster as the S82/WISE PDAC
			• All hosts in the PDAC cluster are functioning

	SLAC / Data Access - Qserv	Igor Gaponenko	• Stripe 82 catalog databases have been loaded into Qserv; simple tests seem to be working
			  • Have not tried performance / large-scale tests yet
			  • Lots of changes to the loading procedure were required in order to make this efficient & correct
			  • Procedure is now highly parallel
			• Next step: produce a writeup on loading, especially for Unknown User (speckins)'s use
			  • Should be possible to experiment with loading additional databases without affecting S82 data
			• Working on better file organization for the image data
			  • Looking at following Unknown User (jalt)'s proposal for data immutability (see ~~RFC-249~~ - Common Dataset Organization and Policy IMPLEMENTED)
			• Ran into some problems with respect to time stamps / time zones in creating Docker containers (see ~~DM-8119~~ - Incorrect time zone settings within Qserv Docker containers DONE)
			  • Interpreting logs when containers have a different time zone than the host OS - KTL says that ~~DM-1203~~ - Set single time for all project software & common machines DONE - Set single time for all project software & common machines Done is meant to be the current policy on log timezones and formatting; if NCSA can now support UTC-only systems, that would be preferable.
			  • Interpretation of time stamps in the database data - KTL says that database columns are meant to be explicit in the schema (usually in the column name itself) as to which time scale they are in (TAI or UTC) and should never be in local time. But while this appears to have been followed in the baseline schema, it doesn't seem to have been written down. MySQL (MariaDB) appears to store TIMESTAMP columns as integers containing YYYYMMDDHHMMSS interpreted as UTC but they appear to ignore leap seconds and do not accept 60 as a seconds value.
			  • Unknown User (jalt): NCSA is very happy to go to uniform UTC time zone settings in production services

	SLAC / Data Access - DAX	Brian Van Klaveren	• Ready to redirect `dbserv` at the Qserv cluster
			• Qserv is now up, so we should try to do this today
			• Still working on getting `metaserv` up today and tomorrow
			• John Gates made progress on `imgserv` and we will make some images available today
			• John Gates: raw images and cutouts are now working
			  • Need Jim Bosch's changes to `afw` deployed in order to get codas working, may take a couple of days
			• IPAC input: please prioritize getting Qserv going behind {{dbserv}}
			• Everybody is OK with giving first priority to `dbserv`-Qserv and `imgserv`-raw/calexp
			• Some discussion with Unknown User (ymei) of whether DAX is giving accurate reports of float vs. double column types
			  • Conclusion: prioritize getting `metaserv` up to produce accurate answers in preference to fixing the immediate issue

	IPAC / SUIT	Trey Roby	• Working on getting multiple types of searches attached to `dbserv` from the UI. Have demonstrated full round trips from the UI to the current temporary `dbserv`. Really need the actual Qserv service in order to continue to make progress.
			  • These search processors are already running on the PDAC Firefly deployment
			• Load-balancing server is up and running
			  • Planning to put the DAX services behind the load-balancer as well
			• Planning to do direct DAX queries from an IPython notebook as a proof-of-concept, with Firefly visualization in the notebook
			• Have started the creation of the PDAC Firefly application page
			  • Will bring in the IRSA light curve viewer once the images are available

Action items

- [x] Igor Gaponenko: change the Docker container production/configuration procedures in such a way that internal timezone setting of the containers were inherited from their host environment. The implementation of this will be conducted in a context of ~~DM-8119~~ - Incorrect time zone settings within Qserv Docker containers DONE . Implement this before the next meeting.