PDAC Meeting 2017-11-30

Prototype Data Access Center (PDAC) meetings take place every other Thursday at 9:00 Project Time on the BlueJeans infrastructure-meeting channel: https://bluejeans.com/383721668

Date

30 Nov 2017

Attendees

Usual attendees

- Paul Domagala
- Gregory Dubois-Felsmann
- Igor Gaponenko
- Fabio Hernandez
- Simon Krughoff
- Kian-Tat Lim
- Fritz Mueller
- Xiuqin Wu

Goals

Discussion items

Key: AGENDA TOPICS, DISCUSSION, FOLLOWUP

Time	Item	Who	Notes
	status of spontaneous reboot problem	Unknown User (pdomagala)	ref. ⚠ IHS-606 - Jira project doesn't exist or you don't have permission to view it. • All but three nodes recovered after the mass firmware update. (Hot news: looks like two more are back now.) • Hoping but not certain that this will solve the problem. • Expecting all nodes to be available again for work today. • Brief discussion of state of Qserv fault-tolerance ○ Qserv already supports failover to replica data on other nodes - this has been tested. But this requires the data to be present. ○ Automated system for *distributing* replicas across a cluster not yet available.
	data ingest experience and Qserv master performance	Igor Gaponenko	• Pleased with the performance of the NVMe filesystem. Not yet able to distinguish its performance *in the Qserv context* from that of the SATA SSD RAID filesystem, though lower-level tests showed a clear difference. ○ Suspecting that the NVMe advantage mainly shows up in heavy parallel read situations and we are not really exploring that corner of the phase space. • WISE loading and indexing is now complete through the first-year NEOWISE(r) data. ☐ Igor Gaponenko will produce a short writeup summarizing the tables loaded and the small differences from their materialization at IRSA. • Schema annotation not yet loaded for WISE. ○ Needed for `metaserv` ○ Will review the (excellent) WISE documentation for guidance. ☐ Gregory Dubois-Felsmann and Fritz Mueller will work on preparing the milestone testing.

	/scratch space retention policy	Unknown User (pdomagala)	Re: ⚠ IHS-613 - Jira project doesn't exist or you don't have permission to view it.
			
			• There has been a request for a "trash can" functionality allowing a short period for recovering data after it has been removed by the automated cleanup system.
			  ○ The current retention period is 180 days **based on `mtime`** (not `atime`).
			  ○ The suggestion is to move data at some time *before* 180 days (e.g., 150 days) to a parallel directory tree on the same filesystem, and then deleting it from there at the end of the full 180 days.
			  ○ Simon Krughoff: people are getting surprised by the period when it expires, but the length of the period is not in dispute.
			  ○ The new feature would primarily aid people who access the data regularly enough that they would be likely to notice the missing file during the "trash retention" period.
			  ○ Unknown User (pdomagala), Kian-Tat Lim: If the data are being accessed regularly and are needed for more than 180 days, perhaps they shouldn't be in `/scratch` in the first place?
			• Fairly lengthy discussion. Rough conclusion:
			  ○ This feature could be implemented without any immediate impact on space requirements, and the sign of the longer-term impact on space is difficult to estimate.
			  ○ The constraining factor is the NCSA staff time required to implement the feature and its relative priority compared to other things. For now the NCSA team management has determined that this is a "nice to have", not a priority, and that it will not be implemented in the near future. This can always be re-discussed at need among T/CAMs.

	Science Platform (LSP) Workshop at IPAC 04 Dec 2017 to 08 Dec 2017	Gregory Dubois-Felsmann	• The agenda is available on Confluence. Please have a look.
			• Will be talking to the team T/CAMs and science leads about what the end-of-week outputs should be - we want to have definite goals for artifacts that will be produced during the week.
			  ○ They may still be reformatted into proper documents in the following days, but there should be some immediate written outputs to force decision-making.
			• Discussion with Fabio Hernandez about availability of workshop materials for those who cannot attend in person or "live" remotely.
			  ○ A lot of the work will not be in the form of presentations, but we will try to capture whiteboard photos, notes, etc. and make them available promptly.

	Next week's meeting	Unknown User (pdomagala) / Gregory Dubois-Felsmann	• There *will* be an Infrastructure call next week, although many of the usual suspects will be in the LSP Workshop. It may be brief.

Action items

☐

Attachments

File	Modified

No files shared here yet.