

PDAC Meeting 2017-02-09

Date

09 Feb 2017

Attendees

- [Gregory Dubois-Felsmann](#)
- [Unknown User \(xiuqin\)](#)
- [Trey Roby](#)
- [Unknown User \(jalt\)](#)
- [Fritz Mueller](#)
- [Fabio Hernandez](#)
- [Kian-Tat Lim](#)
- [Brian Van Klaveren](#)
- [Igor Gaponenko](#)
- [John Gates](#)
- [Simon Krughoff](#)
- [Greg Daues](#)
- [Zeljko Ivezić](#)
- [David Shupe](#)

Goals



- Review status as of the initial exposure of PDAC to a science user

Coordinates

- <https://bluejeans.com/383721668/>

Discussion items

Time	Item	Who	Notes
15 min	NCSA /Infrast ructure	Unknown User (jalt)	<ul style="list-style-type: none">• Machine configuration adjustments<ul style="list-style-type: none">◦ SUI machines: mount <code>/datasets</code> and set up <code>suiadmin</code> home directory - done earlier this week◦ Qserv czar: mount <code>/datasets</code> - done this morning◦ SUI request for SMTP access from Tomcat server nodes<ul style="list-style-type: none">▪ SMTP is already available internally at <code>mail.ncsa.illinois.edu</code> a/k/a <code>smtp.ncsa.illinois.edu</code> - OK to go ahead and use it.

15 min	SLAC /Data Access	Fritz Mueller et al.	<ul style="list-style-type: none"> Operational status & monitoring webserv API developments: <ul style="list-style-type: none"> metaserv - Brian Van Klaveren <ul style="list-style-type: none"> Working on updating the API, working on a preview. Will be deployed as API "v1". Will have new approach to the "level" field, involving a discovery scheme similar to the schema browser. Uploaded a link to Python code displaying the API. Discussion about future directions; will be picked up again at the Monday 11 PT meetings imgserv - cutout service - John Gates <ul style="list-style-type: none"> Rows for proposed APIs I11 and I12 added to API page. Kenny Lo is working on implementing them: <div style="border: 1px solid orange; padding: 10px; margin: 10px 0;">  DM-9115 - Jira project doesn't exist or you don't have permission to view it. </div> <ul style="list-style-type: none"> Also working on writing up a description of a "v1" API for <code>imgserv</code>, for the next major round of work. SUIT feedback: please focus on the single-epoch by-ID implementation, it is a blocker for the light-curve viewer. Question about mosaicing and downsampling coadd patches (skipped - will defer to Monday DAX-SUIT meeting) WISE data loading <ul style="list-style-type: none"> Igor Gaponenko can start this now. Request for <code>/datasets</code> on Qserv czar derives from this. Remaining problem with Stripe 82 dataset: there is a small hole between the NCSA and IN2P3 sections of the processing in the catalogs. Igor Gaponenko is working on this: <div style="border: 1px solid orange; padding: 10px; margin: 10px 0;">  DM-8241 - Jira project doesn't exist or you don't have permission to view it. </div> <ul style="list-style-type: none"> Discussion of deploying a test of replication
15 min	IPAC /SUIT	Unknown User (xiuqin) et al.	<ul style="list-style-type: none"> User guide page Development status <ul style="list-style-type: none"> Light curve viewer - working at IPAC on WISE data, currently completing adapting it to the LSST interface to the Stripe 82 data. Blocked on the single-epoch, by-ID cutout service (one feature of the LC viewer is providing a scrollable display of postage stamps around the selected object in each epoch).
5 min	Science requirements	Gregory Dubois-Felsmann & Zeljko Ivezić	<ul style="list-style-type: none"> Brief discussion of the current 1 arcmin "overlap distance" in the Qserv partitioning scheme. This is documented in LDM-135, section 3.3.6: <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p><i>To avoid [the need for data exchange among shard servers], each partition can be stored with a pre-computed amount of overlapping data. This overlapping data does not strictly belong to the partition but is within a preset spatial distance from the partition's borders. Using this data, spatial joins can be computed correctly within the preset distance without needing data from other partitions that may be on other nodes.</i></p> <p><i>Overlap is needed only for the Object Catalog, as all spatial correlations will be run on that catalog only. Guided by the experience from other projects including SDSS, we expect to preset the overlap to ~1 arcmin, which results in duplicating approximately 30% of the Object Catalog.</i></p> </div> Gregory Dubois-Felsmann expressed concern about whether the specific value of this limitation is well understood in the LSST science community, and suggested that we should do outreach no later than this year to ensure that it is and to collect any feedback that might affect our design. Private outreach to various collaborations followed by a presentation at the August All-Hands? Brief discussion of alternatives to true two-point queries, such as binning (by GROUP_BY). Any alternative strategies like this that emerge should themselves be added to our reference queries. Zeljko Ivezić suggested that in advance of the AHM we work to identify specific science use cases that require spatial joins, and understand those queries in more detail, before making any public presentation. He asked Gregory Dubois-Felsmann to file a ticket to record the need for this work. Fritz Mueller noted that, although the 1 arcmin value is documented (and under DM-level change control - gpdf) in LDM-135, it does not appear to be visible in any higher-level documents.

Action items

- ☐ [Gregory Dubois-Felsmann](#) File JIRA ticket regarding the chunk overlap radius science-requirements issue