Data Access / Apps Dependencies 2015-08-04

Date

04 Aug 2015

Attendees

- John Swinbank, Jim Bosch, Serge Monkewitz, Jacek Becla

Goals

- Discuss Inter-team dependencies for Data Access / Application

Discussion items

Butler

- two features needed on hcs side:

    a. multiple parent repositories (usecase: combine official data with L3 data brought in by user, produce one output)
    b. data from different cameras

- plan is to do the design for "multi-parent repositories" in W16, with help from Apps, primarily Paul Price. Then tentatively planning quick implementation starting ~Feb 2016 (again with help from Paul)
- butler phone call every 2 weeks, considering Thu 11:00 pacific, first one Sept 7

Spatial queries

- Needed for images in the short term, later might want it on catalogs too
- This might require additions to butler API
- providing output from spatial queries - maybe as a separate metadata

Organizing input data for processing

- maybe reorganize data by objects (all pixels for a given object per file, or multiple objects per file)
- this is "implementation detail" that butler should hide
- would be useful to get hints what will be processed next. This might come from middleware pipeline orchestration

Long term planning

- will want to optimize image loading for processing (seeDLP-523, scheduled for W18). So we need a milestone on DataAccess side: optimize image access patterns after W18
- Need to identify more such things. Jim will update dependency page in Confluence

Short term ingest integration

- Hack existing ingest and dump data to cvs, or load to mysql and dump csv from there
- If we want to integrate with firefly, we need to ingest images as well (at Bremerton), and implement some basic spatial search for images