DM Leadership Team Virtual Face-to-Face Meeting, 2020-02-25 to 27

Logistics

Date

25 Feb 2020 – 27 Feb 2020

Location

- This meeting will be conducted on BlueJeans: nobody is expected to travel to participate in it.
- The Workroom in Tucson has been reserved for those in Tucson.

Browser	Room System	Phone Dial-in
https://bluejeans.com/293724745/	1. Dial: 199.48.152.152 or bjn.vc 2. Enter Meeting ID: 293724745 -or- use the pairing code	Dial-in numbers: • +1 408 740 7256 • +1 888 240 2560 (US Toll Free) • +1 408 317 9253 (Alternate Number) Meeting ID: 293724745

Slides

- Please upload slides to Confluence in advance of your presentation (to avoid issues with screen-sharing, etc).

Attendees

- Robert Gruendl
- Colin Slater
- Eric Bellm
- Unknown User (mbutler)
- Yusra AlSayyad
- Jim Bosch
- Kian-Tat Lim
- Gregory Dubois-Felsmann
- Unknown User (npease)
- Zeljko Ivezic
- Wil O'Mullane
- Tim Jenness
- Frossie Economou
- Unknown User (gcomoretto)
- Simon Krughoff
- Jeff Kantor (off and on)
- John Swinbank
- Leanne Guy

Apologies

- Wil O'Mullane flight on Ceb 27 now at 12:00 MST(was 17:00) so it looks like i may be dialed in from car and airport on the last day

Agenda

Day 1: 2020-02-25			
Time (Project)	**Topic**	**Coordinator**	**Notes**
Moderator: Leanne Guy			
09:00	Welcome	Wil O'Mullane	• Introductory remarks • Review agenda and code of conduct
09:10	Project news	Wil O'Mullane	• We've not had a DMLT call since 2020-02-10 — what project-level news has happened since then? --- • LCR for Auxiliary Telescope naming now submitted — please take a look! • Chris Stubbs' Observing Run Debrief. • Open questions about how the “lessons learned” (vis-a-vis simple technical fixes) from AuxTel are disseminated to the wider project. • Leanne Guy is driving the verification effort; aiming to get all priority 1a requirements verified for this summer's reviews. ◦ She will be coming to those responsible for milestones to plan working requirements testing into milestone ☑ Frossie Economou — file an RFC about the possibility of renaming the LSP to “VERA”. 02 Mar 2020

09:30	Gen3 middleware update	Tim Jenness	• Slides.
			• Last report on Gen3 middleware status was demo mid-December, 2019. In the ensuing couple of months, management of these efforts has transitioned from Fritz Mueller to Tim Jenness.
			• Status and forward trajectory update.
			--- 	
			• Extensive discussion of what “Gen3 feature parity” means.
			  ◦ Effectively, it is *possible* to move all processing jobs that we currently use Gen2 for to Gen3.
			  ◦ It is not necessary that all tasks be converted, or that the registry schema be stable.
			  ◦ Demonstrated by e.g. running a DRP processing, AP processing, etc.
			• Need to define intermediate goals and milestones for middleware development, e.g. support for OCPS.
			• Support from Project Management (ie, Wil O'Mullane) for Tim to effectively tell us what /who he needs and then the project will support that.
			☑ Robert Gruendl — formulate a definition for DM-DAX-12 (Gen3/2 feature parity) & DM-DAX-13 (Gen2 deprecation). 16 Mar 2020
			Update: work in progress on Roadmap to Deprecation of Gen2 Butler page.

10:30	Break		
Moderator: Wil O'Mullane			

11:00	Image display in support of observing	Wil O'Mullane	• Lauren Corlies (EPO) will join
			• Firefly has been used in support of early LATISS operations, and has thrown up some problems; no doubt Robert Lupton can expand on those.
			• What is DM's response?
			• Should consider:
			  ◦ future Firefly development plans (Gregory Dubois-Felsmann) - notes
			  ◦ the report of the Image Display WG (DMTN-126, Yusra AlSayyad).
			  ▪ Scope was to understand use cases and make suggestions for potential tooling.
			  ▪ *Not* to design "one overall image display system".
			  ▪ Yusra provided the meeting with a summary of the results; the interested reader is referred to DMTN-126.
			  ▪ Invitation for DM team to take part in an Astrowidgets workshop.
			  ▪ Chris Waters is currently using Ginga/Astrowidgets extensively.
			  ▪ Simon Krughoff has been in regular contact with Eric Mandel, JS9 developer; he has been very responsive.
			  • Mandel has provided a Docker image, which may make it easier to deploy JS9 in the browser.
			  • There is a JS9/electron as a desktop app that could potentially be a replacement for DS9 if we want to unify user experience.
			  ◦ the possibility of including Ginga and/or JS9 in the Nublado environment.
			  ◦ DM-PORTAL milestone in 2021 is to restart portal development - the portal is a little different to this perhaps (firefly blends these) - we should discuss.
			  ▪ If we restart end 2021, then early in 2021 should we do a technology survey, possibly in conjunction with EPO. When DM-PORTAL was added no prior package for exploration was added (nor milestone)
			--- 	
			• Wil O'Mullane requests a "bake off" between Ginga and JS9 in the LSP JupyterLab environment.
			  ◦ This means making these tools available to observers and seeing what they like.
			  ◦ As well as a technical evaluation from the LSP team as to what they can support.
			  ◦ Wil regards this as high priority, but not as high as the EFD.
			  ◦ Tension between "notebook CI" and this means it is unlikely to be available for observers until late Spring.
			  ◦ Will get a date for that on Thursday.
			• This bakeoff is not the same as the portal decision.
			☑ Yusra AlSayyad/ John Swinbank — capture feedback from observers about their interactions with Firefly (this might go in the WG report, or it might be a separate document). 16 Mar 2020
			☑ Frossie Economou — get JS9 and Ginga integrated into the LSP JupyterLab environment (date TBC). 04 May 2020
			☑ Yusra AlSayyad — make contact between Frossie Economou and Ginga developers. 09 Mar 2020
			☑ Wil O'Mullane — get Gregory Dubois-Felsmann access to the summit deployment before the 10 March observing run. 02 Mar 2020

12:00	How do we manage calibration products?	John Swinbank	• Slides. • What is the model for managing calibration products during the operational era? • For example, are calibration products versioned through git repositories (as they are during construction)? Are they exclusively managed through the Butler? At what points can data be "ingested" to a data repository? • My understanding is that the middleware and calibration products teams have built an impressive toolbox of technologies that can be used to implement whatever data management policy we want, but that nobody has yet written down what that policy should be... and many people have different, incompatible, implicit policies in their heads. --- • Raw calibration data is not included here; all raw data is kept together in Butler repos in the Data Backbone regardless of purpose. Not; all raw data together no matter purpose • The logic to choose which master calibration products to use when processing data is currently undefined. • We need to define the available technical solutions, but more so we need to define processes and procedures, not just the technical design. • Would a more active product owner or an assistant help with pushing this definition, without needing a working group? • The next Ops Rehearsal has dealing with master calibrations as an explicit part. • Middleware has been defining one strategy via RFC on obs_lsst_data for pipeline-generated products that are converted to human-readable forms for acceptance and curation in a git repo. Simon and Tim are working on a DMTN describing this. ☑ Tim Jenness — Extend the DMTN to solve whole problem or at least describe the future questions to be answered; leverage/delegate to Chris Waters to avoid overload. 01 Apr 2020 ☑ Leanne Guy — Consider how to engage with calibration product data management from product owner side. 01 Apr 2020 – Scheduled for DM Leadership Team Virtual Face-to-Face Meeting, 2020-05-12/14
12:30	Break		

Moderator: John Swinbank			

13:00	Draft proposal for image capture simplification	Robert Gruendl ,Kian-Tat Lim	• Slides
			• Current proposal is being discussed here.
			• This proposal attempts to alter the means by which image creation and forwarding would occur, replacing elements central to the overall data management and prompt processing systems.
			
			--- 	
			
			• At the moment we write two FITS files out for (almost) every observation: one from the CCS, one from DM code. The proposal is effectively to eliminate the latter.
			• The proposal would have some impact on the Tony Johnson / Camera Team schedule for the CCS Image Writer; we would provide some support.
			• How will we use the channels freed by not running the DAQ client over DWDM?
			• Are we still transferring data twice, or can prompt processing data be written to the DBB?
			  ◦ Should not be necessary to do this twice given the lack of crosstalk correction.
			  ◦ But the data representation might not be ideal for DBB.
			  ◦ There will be a copy at the base for the OODS.
			  ◦ System could be evolved after it is up and running.
			  ◦ Agreed that duplicate transfers would be “non ideal”.
			• Is there a potential for another catchup buffer in this architecture?
			  ◦ No; the camera is not maintaining a buffer of images.
			• Catchup is pulling data from the DAQ, and reconstructing a FITS file? Is it using the same code as was used to write the data to start with?
			  ◦ Not clear yet; still has to be investigated.
			  ◦ Catchup should definitely be based on CCS Image Writer, rather than Forwarder, code.
			• Slide 5, step 5 — Unknown User (mbutler) suggests this should also include backups.
			• In order to recognize the benefits, we should make a decision and move on this soon, before development effort gets spent elsewhere.
			• Proposal is that Steve P. would work on this, in conjunction with Tony Johnson; would need the latter's buy in.
			  ◦ K-T reports Tony is keen on this idea.
			  ◦ Need to check that Steve is happy with this idea.
			• Existing CCS codebase is “not horrible“, but is complex.
			  ◦ It is currently not publicly available. Licensing unclear.
			• What is the decision making process?
			  ◦ Needs to go through an LCR, likely updating LSE-309.
			• Proposal to start working on this ASAP, including LDF work on ComCam image capture; needs careful offline planning.
			
			☑ Kian-Tat Lim — LCR proposal for image capture simplification. 09 Mar 2020

14:00	Incremental template generation in LOY1	Eric Bellm	• Slides. See also DMTN-107.
			• Alert science in early operations would be enhanced by incremental template generation prior to DR1.
			• How much effort would be required of the construction project?
			• Do we have estimates of the operations effort to run it during LOY1?
			
			--- 	
			
			• Incremental: what does it mean?
			  ◦ We make a template once in year one, and then we don't modify it after it has been made. We don't keep adding to existing templates.
			  ◦ We are aware of coverage & overlap issues here.
			• How many images are needed per filter to make a template?
			  ◦ 3 is the number Eric likes, but there is some ongoing discussion.
			  ◦ 3 is consistent with requirements on image noise.
			  ◦ Do not expect any form of DCR correction in year 1.
			• No disagreement with Eric's estimate of pipeline & workflow development.
			• Ops plans are not yet clear enough to speak directly to Eric's plans for Execution and QA, but they sound plausible.
			• Computing impact:
			  ◦ Absent templates, what happens when images arrive at the LDF (assuming no alert production).
			  ◦ Enough single frame processing to return telemetry to the scheduler.
			  ◦ Template generation would be run at end-of-night.
			  ◦ Eric says prompt-processed-style PVIs would be sufficient for incremental template generation; don't need the more elaborate DRP system.
			• How do we manage the impacts of some data being available to users before project-provided data products? How do we prevent our own users from scooping us? What products are we producing? How do we prevent everybody trying to use the LSP to access the data and do their own reductions?
			  ◦ Some of this can be controlled with throttling.
			  ◦ Agreed to return to this topic at a future meeting.
			• We will LCR expanding the construction scope as proposed by Eric.
			• Then it will be Bob's call as to how the Operations team reacts.
			
			☑ Wil O'Mullane — schedule a discussion about rolling out data products and capabilities to users without having them scoop the project or swamp our resources. 23 Mar 2020
			
			☑ Eric Bellm — submit an LCR describing changes to the construction plan to enable incremental template generation. 30 Mar 2020

14:30	Close		

Day 2: 2020-02-26

Moderator: Robert Gruendl

09:00	SDM standardization update	Yusra AlSayyad	• Slides.
			
			Big questions:
			• What is the process for updating the DPDD?
			  ◦ Covered by project level change control.
			  ◦ There are many DPDD update tickets; need to prioritise getting them done.
			  ◦ Speed of development for Pipelines vs. DPDD changes is very different; impedance mismatch.
			  ◦ It is not necessary that the DPDD list everything described in the SDM; it's also possible to queue up DPDD updates on `master` rather than baselining them as they arrive.
			• What is the "missing link" between the SQL schema and consumers (Qserv, etc)? Is it Felis?
			  ◦ Who is maintaining Felis since BVan left DAX?
			  ◦ Suggestion that a testing framework is necessary.
			  ◦ Hsing-Fang may have the best sense of what is the next most useful utility to be added to the Felix toolkit, and she would be in the best place to make this happen – consensus that Hsin-Fang will be the Felis maintainer.
			• Changes to `BaselineSchema.yaml` should be change controlled.
			• Need to write a technote on what the schema is, where it's used, where it's going, etc. Some tension between providing enough visibility into what's happening without overly constraining or overloading the people who are doing the work. Agreed that Wil would do this as a compromise.
			
			☑ Leanne Guy — produce a plan for interaction between the DPDD and the concrete SDM schema. DM-23818 - Getting issue details... STATUS 04 May 2020
			
			☑ Fritz Mueller — find somebody to update the online schema browser. 06 Apr 2020 (Igor Gaponenko assigned)
			
			☑ Kian-Tat Lim — arrange for the schema browser to be removed, until & unless the action to update it comes true. 09 Mar 2020
			
			☑ Colin Slater — ensure change control policy for BaselineSchema.yaml is documented. In progress on DM-23614 - Getting issue details... STATUS 06 Apr 2020
			
			☑ Wil O'Mullane — write a technote describing his understanding of schema management DM-23658 - Getting issue details... STATUS . 06 Apr 2020

09:30	Parquet data products	Colin Slater	• We should be clear on our overall strategy for Parquet data products, including:
			  ◦ Are we committed to support Parquet (or more generally a columnar data format) as a user facing format for LSST catalog data products.
			  ◦ if so, how do we slice/tile the data within the files?
			  ◦ How do we make these available? Bulk download? By sky region?
			  ◦ What is the strategy on using catalog data in Parquet files for backup or disaster recovery.
			  ◦ Who controls the schema for Parquet data products?
			  ◦ Who validates the generated data against the schema?
			• We should also decide which documents, and how, need to be updated to reflect the decisions taken above.
			
			--- 	
			
			• See also the RFC-662 - Getting issue details... STATUS ticket
			• Slides (on Google | captured PDF)
			• Notes from Gregory Dubois-Felsmann
			• We note that providing a service backed by Parquet files is just one possible use of Parquet.
			  ◦ Refined scope for this session: do we store the data that we make available in Qserv in Parquet files?
			• The DAX team view Qserv partitioning as an internal tuning parameter, rather than something that should be exposed through public data products.
			• Move for a hierarchical representation like e.g. healpix, independent of either Pipelines or DAX representation.
			  ◦ We already use HTM for reference catalogs.
			• Worried about making a one-size-fits-all approach to download — likely need both filesystem and object storage.
			  ◦ Also should consider a CDN.
			• Note that IRSA userbase very much wants bulk download, and almost all catalogs are available in this way.
			  ◦ Some concerns about agency views on data rights.
			  ◦ We recognize that at this is a likely upscope, which we should identify.
			  ◦ We should not refer to this as a “bulk download service”.
			
			☑ Robert Gruendl — prepare a technote defining the meaning of “bulk download”. 26 May 2020
			
			☑ Unknown User (mbutler) & Gregory Dubois-Felsmann — identify existing requirements, or suggest new requirements, for a user-facing ”bulk-download“ service (but not under that name). 05 Jun 2020

10:30	Break		
Moderator: Kian-Tat Lim			
11:00	Networks Status and Planning	Jeff Kantor	• Summit • Summit – Base • Base – LDF --- • Full-bandwidth testing is pending availability of the forwarders; these are not currently being procured due to uncertainty over the post-crosstalk-descope data acquisition design. ◦ Do not regard this lack of testing as a major risk. • LSST Security Summit is coming up in April. Agenda unknown (until then, talk of encryption is just speculation). • Query whether there should be a full VNOC at the summit, given that it is likely to be staffed during the night. • Query as to whether international partners need VNOCs. • VNOC is a small set of servers, directly measuring aspects of network performance (dropped packets, etc), and providing a facility to document network events, together with a transmission of that information to a central collecting point, which then publishes to web portals.

11:30	APDB update	Fritz Mueller	• Slides
			• Cassandra has been chosen for evaluation as a potential platform for implementing the APDB
			• Hardware has been procured and deployed at NCSA to support this evaluation.
			• Report on progress of this effort and possibly early findings.
			
			--- 	
			
			• ap_proto is a simple simulation of the AP pipeline; it approximates what the pipeline is supposed to do, but without science logic.
			  ◦ https://github.com/lsst-dm/l1dbproto
			  ◦ This is the same text fixture as was used for SQL system evaluations.
			• Current hardware provides 1–3 months of experimentation; then another couple of months of cloud experimentation; should have a costing on the Cassandra system sometime in the summer.
			  ◦ Should report on this at the next DMLT.
			• Use caution when comparing absolute values between the SQL and Cassandra results presented.
			• The DAX group will push the Cassandra investigation as far as they can, but will jump to a custom solution if they find it to not be viable.
			
			☑ Fritz Mueller — report on progress on Cassandra / APDB to the DMLT. 23 Mar 2020 (Report deferred to next DMLT due to reassignment of Andy S to middleware task.)

12:00	Future operations rehearsals	Robert Gruendl	• Slides • Brief discussion of the plan for Ops Rehearsal #2, which is coming up soon. • Longer term discussion. What are our future operations rehearsals? Are they being scheduled to reflect particular hardware deliveries or other capabilities, or based on the calendar? Are we really treating them as “operations rehearsals”, or are we misusing this word to mean “integration exercise”? --- • We should be clear that making data available “through the LSP” means more than just having it accessible on a filesystem through a Butler. • Expectation is the rehearsal terminates after running pipelines and simple QA; no data being made available for community inspection. • Note that “prompt processing” in these slides are in scare quotes for a reason — they are not LDM-148 Prompt Processing Service processing, but just data processing that takes place soon after data has been acquired. • Kubernetes cluster at the base is about a week away. • Keen to run what verification we can during the ops rehearsals. • Some consensus on moving operations rehearsals away from hardware delivery dates, not least because hardware become available will almost certainly be immediately pressed into use. • Only hard part in terms of Gen3 middleware is making data incrementally available. ◦ Ie, incremental visits arriving, contrasted with a complete data release. • John Swinbank would be a good point of contact for information on and coordination of pipelines activities. ☑ Wil O'Mullane (with Bob Blum) — coordinate schedule for Ops Rehearsal #2 with the LATISS team to make sure that we aren't disrupting LATISS engineering work. 16 Mar 2020 ☑ Robert Gruendl & John Swinbank — agree on pipelines availability for OR#2. 16 Mar 2020
12:30	Break		
Moderator: Simon Krughoff			

13:00	Public access to data after the 2 year proprietary period	Eric Bellm	• We should develop and advertise a clearer plan for how non-Data Rights holders can access data release(s) that are no longer proprietary.
			  ◦ Bulk access through a cloud host?
			  ◦ Unauthenticated API or Portal access?
			  ◦ Something else?
			  ◦ More if they pay?
			• Have to make sure this is consistent with Ops project thinking.
			• Notes from Gregory Dubois-Felsmann
			
			--- 	
			
			• This discussion is in part a response to discussions that arose at the AAS meeting around access to Rubin Obs. data.
			• Can we make a specific statement acknowledging the challenges involved in providing public access to Rubin data?
			• Even *coming up with a plan* here is outside our formal scope, and it's clearly not a day-one problem for Operations.
			  ◦ Should the DMLT be doing *anything* here, even though we care?
			  ◦ Broadly: no, although we shouldn't do anything that'll make it harder to solve this problem in future.
			
			☑ Wil O'Mullane — write a paragraph for the SAC describing the DMLT's professional opinion on how we might make old data releases available in operations, should we be asked to do so. Done ... DMTN-144 30 Mar 2020

13:30	Progress on Conda packaging	Kian-Tat Lim	Slides
			
			--- 	
			
			• See DMTN-110, DMTN-138.
			• It will be possible to support a non-conda-forge channel for packages which require Rubin-specific patches.
			• This does not reduce the (current) two installation mechanisms to one. It does change the lsstsw mechanism.
			  ◦ eups distrib / newinstall process will remain the same, but it will shift more packges to the Conda environment.
			• Who is the customer of this work? Who will maintain it in the long term?
			  ◦ Product owner is not well defined; perhaps it's K-T.
			  ◦ Not clear who will maintain it into operations.
			• What is the meaning of the drop-dead-date?
			  ◦ The toolset becomes available and used within lsstsw.
			
			☑ Wil O'Mullane, John Swinbank, Leanne Guy — understand who the maintainer of (Conda?) packaging is in the operational era. 16 Mar 2020
			
			☑ Leanne Guy — determine product owner for Conda packaging. 16 Mar 2020

14:00	~~How do we process data from Cerro Pachón in flexible ways at the Data Facility?~~	Robert Lupton	• Robert Lupton was unavailable.
14:30	Close		

Day 3: 2020-02-27

Moderator: Gregory Dubois-Felsmann

09:00	Plans for the next half-cycle	John Swinbank	• We'll next meet in only three months, so rather than a full cycle plan, let's talk about our goals for that period.
			• Each group please provide (~10 minutes total):
			  ◦ A brief retrospective on what's happened since our last meeting.
			  ◦ Plans for the next three months.
			• Architecture (Kian-Tat Lim)
			  ◦ OCPS is the new name for OCS Driven Batch; doc updates coming in S20.
			  ◦ Prompt Services requirements coming from the Commissioning Team primarily at the moment.
			  ▪ Prompt Services covers a bunch of things, not just Prompt Processing; includes Header Service, OODS, etc.
			• DM Science(Leanne Guy)
			  ◦ validate_drp redesign effort is currently looking at MetricTask.
			  ◦ Not committing to ingesting HSC RC2 data to Qserv every month, but everybody agrees this would be a good idea.
			• Alert Production (John Swinbank)
			  ◦ Aim to use G3 middleware for any LDF-supported AP pipeline runs.
			  ◦ Keen to make decisions about the future of the Alert Filtering Service soon.
			• Data Release Production (Yusra AlSayyad)
			  ◦ Tests have been performed on satellite trail rejection.
			  ◦ The uncertainty on Tony Tyson's claim that we may lose 30% of images is that it's not clear how different future satellite constellations will look from precursor data, and we have some technical concerns with some of the analysis which has been performed to date.
			  ▪ HSC has a narrow field of view, and a relatively small survey time allocation; just been lucky it's not seen any so far.
			• DAX (Fritz Mueller)
			  ◦ Fritz has been involved with the team working on the DAQ.
			• Data Facility (Unknown User (mbutler))
			  ◦ Many members of the DMLT extend thanks to Michelle and the NCSA team in the current difficult situation over the LDF.
			  ◦ Concerns about Qserv disk lifetime; Michelle is pressing ahead with procurement.
			• SQuaRE (Frossie Economou)
			
			☑ Leanne Guy — follow up with Kian-Tat Lim about the Prompt Services Product Owner role w.r.t Commissioning needs. 16 Mar 2020
			
			☑ Leanne Guy — present status on RC2 ingest to Qserv at May DMLT. 05 May 2020

	Wrap up	Wil O'Mullane (If not boarding flight)	Actions and next meetings.
			
			• Seattle 2020-05-12/14
			  ◦ This meeting will go ahead in person.
			  ◦ But people who want to opt out for either domestic or environmental reasons will be assured of a good remote connection.
			• Virtual, 2020-11-16/19
			  ◦ **Note this is one week later than previously planned.**
			  ◦ This meeting will be **virtual**.
			• Tucson, 2021-02-22/25.
			  ◦ MCR booked - does not seem to clash with anything
			• In future, we expect the February meeting to be a regular in-person meeting, with virtual meetings in May and November.
			• There may also be an all-hands in Chile in 2021.
			
			☑ Wil O'Mullane — confirm dates for February 2021 DMLT meeting. 09 Mar 2020

11:00	Close		

Attached Documents

File	Modified
PDF File 2020-02-25 — DMLT — Swinbank — Calibration Products.pdf	Feb 25, 2020 by John Swinbank
PDF File Image Capture Simplification.pdf	Feb 25, 2020 by Kian-Tat Lim
PDF File Bellm_Incremental_Templates_200225.pdf	Feb 25, 2020 by Eric Bellm
PDF File Conda Packaging Status.pdf	Feb 26, 2020 by Kian-Tat Lim
PDF File DM Middleware Updates, 2020-02-25.pdf	Feb 26, 2020 by Tim Jenness
PDF File DMLT-F2F-20200226-OPSRehearsals.pdf	Feb 26, 2020 by Robert Gruendl

PDF File APDB Update.pdf	Feb 26, 2020 by Fritz Mueller
PDF File SDMUpdateFeb2020.pdf	Feb 26, 2020 by Yusra AlSayyad
PDF File 2020-02-27 — AP S20.pdf	Feb 27, 2020 by John Swinbank
PDF File DAX mid-S20 Plans.pdf	Feb 27, 2020 by Fritz Mueller
Microsoft Powerpoint Presentation next 3 months .pptx	Feb 27, 2020 by mbutler
PDF File DRP-S20b-Activities.pdf	Feb 27, 2020 by Yusra AlSayyad
PDF File Arch S20B Plans.pdf	Feb 27, 2020 by Kian-Tat Lim
PDF File dmlt_2020_feb.pdf SQuaRE S20A	Feb 27, 2020 by Frossie Economou
PDF File next 3 months .pdf	Feb 27, 2020 by mbutler
PDF File DM Science Plans S20B.pdf	Feb 27, 2020 by Leanne Guy
PDF File Kantor Networks.pdf	Apr 01, 2020 by Jeff Kantor
PDF File Feb 2020 DMLT - Columnar Data Products.pdf	Nov 20, 2020 by Gregory Dubois-Felsmann

Download All

Action Item Summary

Task report

Looking good, no incomplete tasks.

Pre-Meeting Planning

Suggested topics for discussion

Topic	Requested by	Time required (estimate)	Notes
Gen3 Middleware Update	(by request of the DMLT)	1 hour?	Last report on Gen3 middleware status was demo mid-December, 2019. In the ensuing couple of months, management of these efforts has transitioned from Fritz Mueller to Tim Jenness. Status and forward trajectory update.
Validation of Parquet data products	Fritz Mueller Leanne Guy	60 minutes	• Who controls the schema for Parquet data products? • Who validates the generated data against the schema? Following discussion with Fritz on this topic, I (Leanne Guy) propose to expand this session to address our overall strategy for Parquet, including: • Are we committed to support Parquet (or more generally a columnar data format) as a user facing format for LSST catalog data products. • if so, how do we slice/tile the data within the files? • How do we make these available? Bulk download? By sky region? • What is the strategy on using catalog data in Parquet files for backup or disaster recovery. We should also consider which documents, and how, need to be updated to reflect our strategy
Management of data during the operational era	John Swinbank	30 minutes	• What is the model for managing data and data products during the operational era? • For example, are calibration products versioned through git repositories (as they are during construction)? Are they exclusively managed through the Butler? At what points can data be “ingested” to a data repository? • My understanding is that the middleware and calibration products teams have built an impressive toolbox of technologies that can be used to implement whatever data management policy we want, but that nobody has yet written down what that policy should be... and many people have different, incompatible, implicit policies in their heads.

SDM Standardization Update	Leanne Guy		
Networks Status and Planning	Jeff Kantor	30 minutes	• Brief overview of Summit, Summit - Base, and Base - LDF Network status and schedule
Public access to data after the 2 year proprietary period	Eric Bellm	30 minutes?	We should develop and advertise a clearer plan for how non-Data Rights holders can access data release(s) that are no longer proprietary. Bulk access through a cloud host? Unauthenticated API or Portal access? Something else? More if they pay?
Status of drill down tool for analysis of pipeline outputs	Leanne Guy	30 mins+	What is the status of the drill down tool being developed by the external consulting firm? A demo would be good (if possible) . (No longer necessary given demos Tim has already shown)
Image display	Leanne Guy	15 mins+	HSC map is now working with DESC DC2 data. I would like to float and discuss the idea of using HSCMap for multi-band image display
Incremental Template Generation in LOY1	Eric Bellm	30 minutes	Alert science in early operations would be enhanced by incremental template generation prior to DR1. This would be new scope–how much effort would be required of the construction project, and do we have estimates of the operations effort to run it during LOY1?
Future operations rehearsals	John Swinbank Robert Gruendl	30 minutes	How are we scheduling future operations rehearsals? Are they tied to particular hardware deliveries / system capabilities being available, or are they purely time based? What can we nail down now, to enable them to be used in planning V&V activities? Outline for second rehearsal is in current LDM-643.
Progress on Conda packaging	Leanne Guy	15 mins	Conda packaging updates
Image display, in particular in support of observing	John Swinbank (by request of the DMLT)	1 hour	Firefly has been used in support of early LATISS operations, and has thrown up some problems. What is DM's response? Should consider future Firefly development plans, the report of the Image Display WG (DMTN-126), and the possibility of including Ginga and/or JS9 in the Nublado environment.
Draft Proposal for Image Capture Simplification	Robert Gruendl	1 hour?	Current proposal is being discussed here. This proposal attempts to alter the means by which image creation and forwarding would occur replacing elements central to the overall data management and prompt processing systems.
APDB Update	Fritz Mueller (by request of the DMLT)	30 mins?	Cassandra has been chosen for evaluation as a potential platform for implementing the APDB, and hardware has been procured and deployed at NCSA to support this evaluation. Report on progress of this effort and possibly early findings.