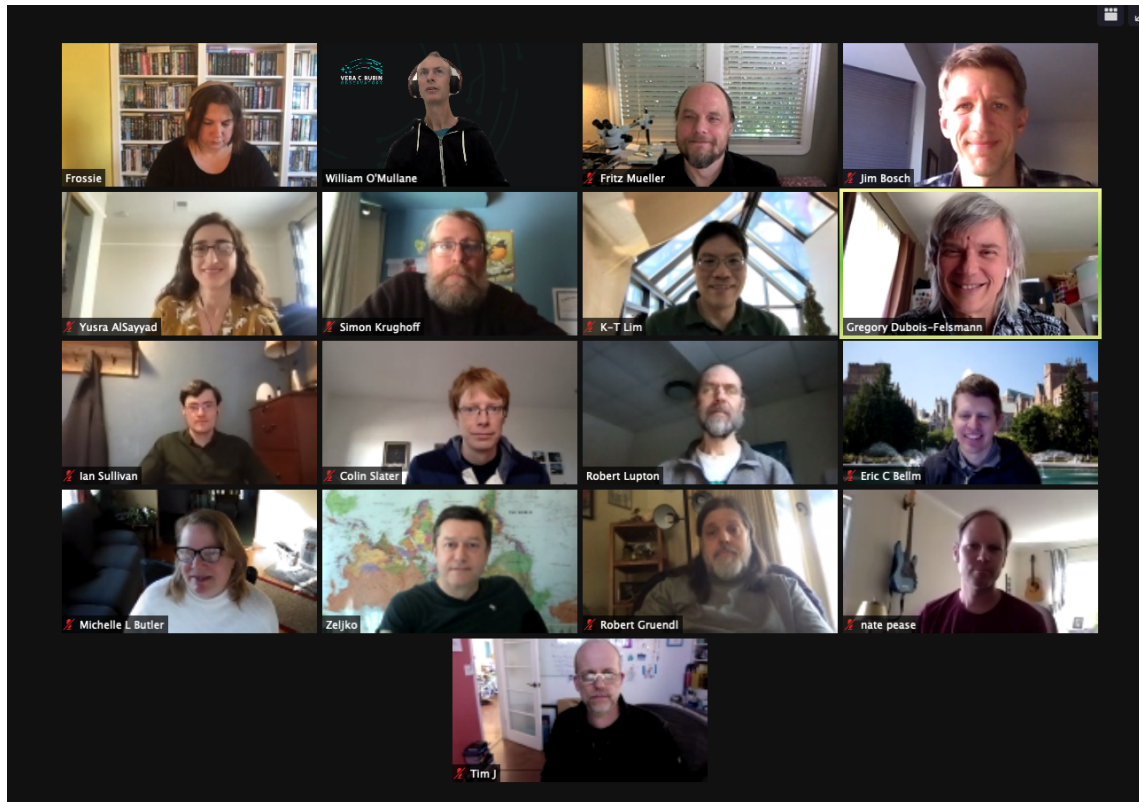


DM Leadership Team Virtual Face-to-Face Meeting, 2021-02-23/24



- Logistics
 - Date
 - Location
 - Attendees
 - Day 1, Tuesday 23 February 2021
 - Day 2, Wednesday 24 February 2021
- Proposed Topics
 - Attached Documents
- Action Item Summary

Logistics

Date

23 Feb 2021 – 24 Feb 2021

Location

This meeting will be **virtual**; details of the teleconferencing system are to be determined.

Join Zoom Meeting

<https://gemini.zoom.us/j/92361521530?pwd=aVJsMTNENE42UzgyUW4vSjVEd0Nsdz09>

Meeting ID: 923 6152 1530

Password: 219352

Attendees

- Wil O'Mullane
- Yusra AlSayyad
- Simon Krughoff
- Kian-Tat Lim
- Unknown User (npease)
- Eric Bellm
- Ian Sullivan
- Zeljko Ivezic
- Robert Gruendl
- Colin Slater
- Robert Lupton
- Jim Bosch
- Fritz Mueller
- Gregory Dubois-Felsmann
- Tim Jenness
- Unknown User (mbutler)

Day 1, Tuesday 23 February 2021

Time (Project)	Topic	Coordinator	Pre-meeting notes	Running notes
Moderator: Yusra AlSayyad	Notetaker: Simon Krughoff			
09:00	Welcome	Wil O'Mullane	<ul style="list-style-type: none"> • Introductory remarks • Review agenda and code of conduct 	Slides for Quarterly https://docs.google.com/presentation/d/1g6GrtisnIqMvY75t1C4Epx_n2JZCUhOF0zS9bGoSyok/edit?slide=id.gbe9b77a59e_1_0
9:15	Project news and updates	Wil O'Mullane		<ul style="list-style-type: none"> • FY2 • Subject to change: End of construction is currently 08/2023. Beginning of Ops is 10 /2023. The difference is construction schedule reserve • Updated date for ComCam on Telescope – July-ish. Subject to change. Uncertain for on sky, but hopefully end of year. New milestones upcoming
9:45	Ops Rehearsal	Wil O'Mullane	<p><i>Moved from 9am PT Wed</i></p> <ul style="list-style-type: none"> • Quick check in on OR#2 for commissioning - we put this off for Gen3 and ComCam move to summit etc. <ul style="list-style-type: none"> ◦ Now we have Gen3 repos - ◦ ComCam is up on Floor3. ◦ Should we do something to ensure we are ready for the integration on Floor 3 ? <ul style="list-style-type: none"> ▪ EFD , ComCam, Hexpod , cable wrap .. ▪ Or a repeat of last time but with Gen3 	<ul style="list-style-type: none"> • We will do another Ops rehearsal. It will be Gen 3 focused. Schedule TBD. ComCam is now on the summit, so we have options. RHL: it will be good if we can make it work for us to learn things we need to learn.

<p>10:00</p> <p>9:30</p>	<p>Provenance WG</p>	<p>Frossie Economou</p>	<p>Moved from 2pm PT.</p>	<ul style="list-style-type: none"> • RHL: What do you mean by logs? FE: Everything. Not sure how to rendezvous camera logs, with pipeline logs e.g. • GPDF: These are process logs, not science log books? FE: Correct • KSK: We are not doing hardware provenance. • KTL: What about provenance of experimental hardware? GPDF: We will be linking to the maintenance management system (MMS). That should tell you what hardware is where and when, but we are assuming that system will be fully functional without input from this working group. • KTL: It's possible that the YAML description in obs_lsst could get out of sync with the on summit reality if these are not intentionally linked up • TJ: Detector serial numbers are directly encoded in the image headers, so no need to go back to the MMS for that. RHL: But raft serials are not • RHL: There are systems that may be dropped that the WG could raise as important sources of provenance that should go forward. <p>✓ Gregory Dubois-Felsmann : Will help with word-smithing in the report to call out the need to preserve any camera metadata /telemetry that is not ending up in the EFD. Due by end of Provenance WG tenure.</p> <ul style="list-style-type: none"> • FM: Sizing model? KTL: If there is a recommendation to get rid of heavy footprints, it will help since they are currently in • FE: Source IDs are a concern. See the report with recommendations. • GPDF: Do we need to not worry about compressed PVI's? KTL: We still need to worry. We are assuming a compression factor for those. • RHL: When is the report due out? FE: By next DMLT F2F. RHL: I'd like to see a draft. <p>✓ Frossie Economou Will circulate Provenance WG document when draft is ready.05 Apr 2021</p>
-------------------------------------	--------------------------	---	---------------------------	--

10:00	Container builds	Frossie Economou	<p>Moved from 11:45amPT</p> <p>Looking for efficiencies in:</p> <ul style="list-style-type: none"> • Container builds across RSP, Prompt Processing, DRP, Telescope & Site • Non-service non-stack packaging 	<ul style="list-style-type: none"> • KTL - should we built T&S software with stack? which with same tech or at same time ? <ul style="list-style-type: none"> ◦ Merlin would like the entangled part released daily .. ◦ It would be good to notice if something on T&S breaks when we update stack .. problem is finding out later. • Tim notes in chat that PanaDA now using lsst_distrib containers. • RHL - painted ourselves in a corner - we should now ask on the mountain we want to be able to developer type things – need to get together and decide how to achieve that .. feels the docker method is more aimed at stable releases. <ul style="list-style-type: none"> ◦ Not entirely agreed - containers do not make it slower necessarily.. ◦ Conways law - currently Science Platform, T&S Software etc all under one lead now - not in ops. ◦ Pick a day in march to try to settle some of this .. seed some tech notes .. • JB - also need to look at release and patches - rather than using the bleeding edge . <ul style="list-style-type: none"> ◦ this was suggested some people recoil from it ◦ only way we currently test T&S and Stack is on the summit .. need some tests (Notebook can be driven by nublado) <ul style="list-style-type: none"> ▪ KTL intent was to use NCSA test stand for this - but AuxTel was not running there ▪ RHL says it may be Tucson test stand . • Colin - whats the rigidity of T&S versioning situation <ul style="list-style-type: none"> ◦ Fundamental T&S issue with XML - interface change requires all to be built .. ◦ But it may be the way we are packaging this with stack which is the problem ◦ TimJ points out Nublado is used for observing which was not supposed to happen - Nublado was meant to insulated from this. ◦ Go back and create cleaner interfaces .. • KTL - T&S does not have a build engineer in ops - do we we actually have anyone to do any out put • Interfaces or Monolith .. <ul style="list-style-type: none"> ◦ some background given on lack of control system and why notebooks came in to being • GPDF - could use cookie between two containers - but htis is the discussion we should have in the meeting <p>Frossie Economou Doodle poll a day in March for a discussion on container builds with coordinated group DM and T&S .. 04 Mar 2021</p>
10:30	Break			
Moderator: Simon Krughoff	Notetaker: Wil O'Mullane			
11:00	Campaign Management	Kian-Tat Lim		<p>DMTN-181 draft note on campaigns</p> <ul style="list-style-type: none"> • FM - s this agnostic to which things are colocated on site (slide 5) <ul style="list-style-type: none"> ◦ assumes single site .. but workflows should be single site but campaigns across sites • Frossie- is campaign offline or data drip <ul style="list-style-type: none"> ◦ would include data drip .. • GPDF - slide 4 - pipeline not mentioned in definitions is workflow a pipeline execution <ul style="list-style-type: none"> ◦ some discussion on that TimJ clarified (sorry missed it)

- GODF Marking intermediate products bad and keeping them out of downstream processing was a critical feature of the _BaBar_ workflow system.
 - JB - It is important - dataids may not be enough then -
 - exclusions may change for new campaign
 - may be better to record it more explicit to the campaign ..
 - TJ - not automated enough - no plan for using generated metrics in down stream . The provenance records which inputs were used.
 - KTL some up front setup from humans - pipelines should take care of some of the
 - FE - is there an interface that could be used by Computers as well as humans ? DO humans have to be involved in nightly
 - yes and no – some human needs to configure the nightly processing .. no one should have to push a button
 - Implementation is important - do not want to encourage very manual intervention
 - RHL - exclusions from looking at a flat are not different to QA systems flagging things .. its a continuum. We will find things in the middle and say we should have flagged this - can we go back and figure out which products are affected ?
 - TJ - we absolutely need a way to know that if Flat X is suddenly determined to be bad then we can find all products that used X and redo them
 - if you change the campaign its a new one - you dont have to run all of it again of course (if gen3 allows you to use partial workflows.
 - RHL want to continue the processing with some new exclusion .. not necessarily a new campaign - it may be semantics .
 - TJ Regarding provenance, you could imagine that we could rebuild a graph from the provenance of that product and just remake that product - JB agrees

- TJ - does this cover which data was processed ?
 - not this system
 - holes in processing would be important though ..
 - KTL thinks perhaps this would be an addon "linter" ..
 - GPDF - if campaign is immutable and runs for 8 months need something which tracks
 - could include campaigns
 - I think I'd like semantic changing of versions to just be one case of semantic versioning of software releases, with all config changes recognized as essentially the same as a change to the software.
 - JB - levels now defined work if we push them down to level of workflow
 - I think I'd like semantic changing of versions to just be one case of semantic versioning of software releases, with all config changes recognized as essentially the same as a change to the software.
 - Iterations add workflow in the middle of pre existing workflows ..
 - progress tool needs to consider how the chunking was done
 - RHL this implies semantic versioning
- WOM not all campaigns are equal .. some we restart and interact with some we must always start from the beginning.
- FE - do not want two ways to get the processing done - manual and automatic, if we do not put pressure on automation it will not happen. We would not want to be manually processing data after 10 years ..
- KTL - trying to find clear places where automation can be added. Would like more pressure to get automated ..
 - Not yet an assignment of who will build this - can not promise the scope
 - So want a framework that could be used
 - FE - when we have made things simple before its bitten us - we should define what we need and get a high level developer to implement it - not just anyone can pick it up.

- RHL - does not understand the worry about (lack of) automation - will do small QA runs and we will inject, KT has defined mechanisms that allow that.
 - Frossie once you provide tooling for humans it takes effort and a better system could have been built to automate it
 - SK .e. if the name of the person running the pipeline changes the configs would evaluate to true, but if the size of the aperture for aperture correction changes it would evaluate as false
 - TJ - problem say good seeing - want the graph builder to understand it right ? or is a user making a query to do the query and get the list of raws .. what about 10million input files - calculate metrics which are used in down stream tasks .. if its not wanted we do not need to build it
 - JB assume we were building this ..
 - KT an interface for a list of exclusions is only an option not the main way - there should be a selector function built in
 - RHL offers HSC example in deep fields - TJ does not think anyone is AGAINST exclusion lists
 - JB we already have a system for this which is not using an explicit list ..
 - YA In chat pointed out the metrics may be selected on if stored as datasets - they do not have to be in the registry.
- GPDF - .e. if the name of the person running the pipeline changes the configs would evaluate to true, but if the size of the aperture for aperture correction changes it would evaluate as false
- TJ in favor of exclusion lists - there are multiple levels as part of the observatory, everything else should go through single frame processing , some you will never want in a coadd
 - its more complex - ideally it would always come from the metrics ..
 - different selector for each coadd (YA)
- CS - easier to think how it all interacts when we have a design for straw system ordinarily .. particular mapping of campaign to workflow below it.
 - list of BPS commands put in a form
 - how you come up with that list is another intellectual/science problem
 - but that is exactly the part that needs to be worked out
 - RHL suggests do this concretely for HSC ..

				<ul style="list-style-type: none"> • TODO <ul style="list-style-type: none"> ◦ Seems fleshing out the external tools (slide 7) would be useful ◦ tooling to generate the BPS lists and exclusion lists ◦ provenance from previous campaign to come up with new one .. ◦ concrete example with HSC .. ◦ Overarching architecture
12:30	Break			
Moderator: Fro ssie Economou	Notetaker: Kian-Tat Lim			
13:00	Alert Distribution & Brokering	Eric Bellm and Leanne Guy	<ul style="list-style-type: none"> • Status report on the SAC review of the Broker proposals • Discussion of a proposal for a "hybrid" alert distribution system (dmtn-165.lsst.io); implications for the alert filtering service and alert DB 	<p>Broker selection</p> <ul style="list-style-type: none"> • Got 9 proposals in Dec out of 15 letters of intent • All wanted full stream • Most likely will want at least 7 rather than 5 • Do MOUs come from Ops project? • SLAC might have more bandwidth outbound to support more brokers • Could also relax latency, cut contents, or provide streams in the cloud with user-pays • Can we support user-pays at SLAC? Difficult, not metered • Make use of "smart networking fabric" across borders? Possibly • Are there support costs or other issues that might be hidden? • Should be discussed on the Ops side of the fence • 10 Gbit baseline might be per-node, so achieving larger could be reasonable • Conclusion: Don't forestall any SAC moves to expand the list of brokers <p>Hybrid alert concept</p> <ul style="list-style-type: none"> • Unlikely to be able to build a usable end-user Alert Filtering Service by the end of construction • Previous options: descope AFS and leave to community brokers or outsource • Instead, use hybrid alerts: small notification packets with separate large downloads • Small = ~200 bytes/alert; expect minimal overhead (not VOTable) • Full alert backing store can be the same as the Alert Database • Alert Database is archive of all the alerts independent of filtering • Can set rate limits per user • Direct access to notification stream and full alerts would be restricted to data rights holders • Advantages: <ul style="list-style-type: none"> ◦ More users ◦ Bring in outside data ◦ Filters in any language/system ◦ No monitoring of performance /security ◦ Rate limit can be user-managed ◦ No on-project processing ◦ Don't need to handle user filters • To ensure equity/access, need bootstrapping to get people running easily • Couldn't brokers do this? Some have mentioned it but none do now <ul style="list-style-type: none"> ◦ Project-provided gives perception of stability • Full-stream brokers might also use it <ul style="list-style-type: none"> ◦ Extra latency is probably not large (but latency to insert into Alert Database might be a problem) • DMTN-165 <p>Wil: Initial thought: we have community brokers, don't need another filtering service, don't have effort to build it, so descope everything</p>

				<p>Don't try to do hybrid alerts? If A&A is needed, complexity goes up; could possibly be farmed out to others (Antares?)</p> <p>Discussion in chat about VOEvent serialization (there is one in JSON) and transformability into web pages (like XSLT for XML-to-XHTML)</p> <p>May still need internal filtering of alerts, but that would likely be before publication to stream and database</p> <p>Frossie:</p> <ul style="list-style-type: none">• Could leverage RSP interfaces for A&A• Possibly leverage Butler Registry server (signed URLs)• Could write a template for boilerplate of subscribing to notification stream• Not clear that there is a lot more code needed• Maybe spend a day with RSP team to determine how much <p>Zeljko: Why is code running on RSP OK and in AFS a problem? Running in independent containers works for sandboxing but is less efficient</p> <p>Colin: This could be better than many of the actual broker proposals if presented as one</p> <p>Gregory: What are the computational requirements for processing the notification stream? Doesn't seem huge, but need to calculate</p> <p>Possible EPO synergy</p> <p><input checked="" type="checkbox"/> Eric Bellm Frossie EconomouDiscuss how RSP interfaces can be used to enable the hybrid Alert model and determine how much extra coding is needed 30 Apr 2021</p>
14:00	Close			
Day 2, Wednesday 24 February 2021				
Moderator: Will O'Mullane		Quick check session. Notetaker: Ian Sullivan		

9:00	Codifying Slack etiquette	Kian-Tat Lim	<ul style="list-style-type: none"> • Thread usage • @ on every message ("DESC-style"?) • @channel usage • Do we need to write things down? If so, is the Dev Guide or Community or somewhere else (DMTN?) the best place? 	<ul style="list-style-type: none"> • Should we be prescriptive or suggestive? <ul style="list-style-type: none"> ◦ Should document expected behavior to help new users • Could include in name of channel the expected culture • Encourages use of text snippets instead of massive code blocks • GPDF: It can help if the original poster solicits replies to be in a thread • FE: Impossible to mandate culture <ul style="list-style-type: none"> ◦ the dev guide is really useful in that it also lays out the expected culture <ul style="list-style-type: none"> ▪ Team leads point new users to it during onboarding ▪ Problem is that we are having people join (at ~20%) without onboarding ◦ Threading is controversial, people are afraid of missing things <ul style="list-style-type: none"> ▪ YA: Uses a thread spool emoji to encourage threading ▪ SK: Threads can also be to focus the conversation between a couple people, without 50 people chiming in <ul style="list-style-type: none"> • want it in the channel, to still keep the conversation public ◦ The support channel is special, FE makes sure to re-read everything there every week to make sure nothing was missed ◦ IS: problem has often been non-project people joining DM channels and unwittingly breaking cultural norms <ul style="list-style-type: none"> ▪ JB: the people who follow the dev guide aren't a problem, it's people who join from outside that are not likely to ever look at the dev guide. ▪ JB: A message people receive when they first join would be more helpful ▪ WOM: A welcome message when you first join would work, but that only comes when you first join Slack ◦ WOM: DEI discussion brought up the Tavern in an unfavorable light, could make clear that ◦ FE: Putting the standards in the dev guide allows us to own it ◦ FE: LSST slack is essentially now a US astronomy Slack <ul style="list-style-type: none"> ▪ We are outnumbered ▪ Worried this will become a big problem for the support channels ◦ FE: Could consider making Rubin-Ops only slack ◦ WOM: We should create a dev guide page documenting Slack culture. <ul style="list-style-type: none"> ▪ New users can be sent a link to that page <ul style="list-style-type: none"> • MG: reference DMTN-155 and include Melissa Graham in drafting the text • FE: If we have to rename channels (such as support channels), those should go through RFC <div> <input checked="" type="checkbox"/> Write RFC for renaming channels to make support nature clear Frossie Economou 01 Jun 2021 </div> <div> <input checked="" type="checkbox"/> Write a Slack user guide for the dev guide Kian-Tat Lim 29 Mar 2021 </div>
------	---------------------------	--------------	---	--

9:20	QA plots/site	Colin Slater	<p>Slides</p> <p>Leanne and Colin wish to have publicly accessible QA plots (from pipe_analysis etc ..).</p> <p>There is a docker image for the site :</p> <ul style="list-style-type: none"> this could be spun up on a login node or it could (should?) be deployed on the cluster using ArgoCD etc ... <ul style="list-style-type: none"> but then it will not be public <ul style="list-style-type: none"> unless we deploy on google and push the data to it .. then what ever does the pushing should be deployed properly 	<ul style="list-style-type: none"> TJ: The plots are just reading from the filesystem <ul style="list-style-type: none"> Will be more development in the future SK: Plans to move to Kubernetes, but this is just a python process for now KT: The page is serving the plots, so there isn't/shouldn't be a link to the plots themselves WOM: If security locks down commissioning data even more, then these plots could be left public KT: While NCSA services are behind the VPN, these are non-privileged containers and don't have to be <ul style="list-style-type: none"> MB: These all access GPFS and so are indeed behind the VPN FE: if you want authentication, we need to consider adding these to the science platform <ul style="list-style-type: none"> Is this a one-off, or a template for many future services? <ul style="list-style-type: none"> WOM: these things never remain a one-off, especially if they're public. MB: It's not just A&A, it depends on which nodes it operates on WOM: this is currently at NCSA, but will also need to be at IDF, USDF, IN2P3 JB: need to carve out room for this to run on DP0 without using the remote Butler KT: It can run as a service, does not need remote Butler for now. We can set up something now in containers as long as it is temporary. WOM: Temporary containers are OK as a prototype CTS: what is the path forward? <ul style="list-style-type: none"> FE: could possibly run this at NCSA without a VPN under some circumstances FE: need to decide if this is a roadmap, and if it is plan how that leads to the Science Platform TJ: If it's in DP0, we need a proper web service MB: short term version is just for the developers, so maybe it should remain behind the firewall CTS: is the only way out of the firewall a new kubernetes cluster? <ul style="list-style-type: none"> WOM: that's a topic for an extended discussion <p>✓ Focused meeting to plan way forward for publicly accessible QA plots not behind NCSA firewall 03 Mar 2021 Colin Slater Fr ossie Economou Kian-Tat Lim Unknown User (mbutler)</p> <p>✓ Set up meeting for longer term development process for QA plot service C olin Slater29 Mar 2021</p>
------	---------------	--------------	---	---

9:40	Handing over the Community platform to the Ops.	Melissa Graham for Leanne Guy	<ul style="list-style-type: none"> DM is not scoped to provide support in construction. Now that we have CET funded in pre-ops and they are building a model for community engagement in operations, I'd like to hand the management and evolution of the community platform to the OPS-CET DMLT_CommunityPlatform.pdf 	<ul style="list-style-type: none"> KT: Is Jim Annis a moderator? MLG: Yes FE: can we use "deliver to operations" not "handover to operations" <ul style="list-style-type: none"> DM/DP will continue supporting the service, emphasis that it is for in-project use as well as community service CET has authority to design front page, assign moderators, etc..., need to maintain private groups/private topics for internal communication tool for DM We could write this division of responsibilities down in a tech note RHL: It is hard to transmit project knowledge to CET. support channels are said to be essential, but are not sustainable <ul style="list-style-type: none"> MLG: continued participation of DM expertise is essential RHL: How does the DM side of support scale? <ul style="list-style-type: none"> WOM: That is an operations issue, and not a DM issue <ul style="list-style-type: none"> CET is meant to be the curator and first line of defense against science questions. That means DM/DP does not have to monitor Community, but CET may call on specific experts to answer hard questions when they come up. FE: We should discuss this when Leanne is present FE: We will test this with DP0 <p>✓ Write tech note dividing responsibilities for maintaining Community. Melissa Graham (Leanne Guy) Frossie Economou 01 Jun 2021</p>
Moderator: Kia n-Tat Lim	Notetaker: Yusra AlSayyad			

10:50	Update on DMTN-139	Gregory Dubois-Felsmann	<p>At the 1-11-2021 while reviewing DMLT ticket DM-15198 Gregory offered to give an update.</p> <ul style="list-style-type: none"> Slides (PDF) 	<p>TJ : Isn't the exposure table just a table with all the fits headers and not a derived quantity? GPDF: I don't remember that being the case. KT: Confirms. DPDD doesn't even have a schema for the exposure table. GPDF: I'm more worried about the visit table, which has metrics.</p> <p>TJ: Could we do a join of all the squash metrics and a link to the exposures? GPDF: Yes, and can we map that onto CAOM2? We previously acknowledged that CAOM2 was awkward to use in production, but I think we can still use it afterward. This is lookback. TJ: Ah, It's not the concept of visit, but the actual post-processed visit.</p> <p>FE: Big Big fan of cutting down data models. There are way too many now, and it's hard to deal with them. I would love to go to the two that Gregory has suggested. KT: As long as there are two and not more that... dynamic observing.</p> <p>Slide: 9 Metadata creation and loading workflows. DRPs. K-T: This also relates back to campaign management activities. GPDF: If we can extract gen3 into the obsCore data model, we can do that on the fly and get a respectable image browser. There's tooling that can use that.</p> <p>Slide 10: Nightly Processing WO: It is looking like we will be asked to hold > 6hr. CS: The metadata is useful for interpreting the alert stream itself. It'd be weird if the metadata for the alert weren't available until 24 hours later. GPDF: Yes, we should be able to at least record that we HAVE TAKEN an observation. People would be able to deduce that from the alerts anyway. KT: Some metadata is OK to release. The fact that we took a picture, what airmass or seeing is not a problem. WO: People are ONLY worried about the pixels. Everything else is OK. KT: So now we're back to <6h release of the metadata. FE: It's frustrating, but we understand what they're worried about. We should get ahead of it and say: here's what we want to do without releasing the pixels. What makes me nervous is that'll anyone can call Wil and tell them that they have a draconian solution to this problem he has to implement.</p> <p>GPDF: I know Frossie has opinions on this. FE: Having a mode where we can serve static files (avoiding computation which is required by the baseline image service design) KT: if everyone's looking at the same asteroid, we can cache it. There was also that statement that "If people are using the pixels and not the catalogs, then we failed." Notetaker's editorial: boo WO: That's basically an FTP server. FE: noooooooooo</p> <p>GPDF (running out of time): Frossie, take a look at the tap slide.</p>
-------	--------------------	---	--	--

11:20	APDB Update	Fritz Mueller	<ul style="list-style-type: none"> APDB Cassandra scale experiments are concluding; summary report on recommended design and hardware requirements. Coordination discussion w/ AP team: when/where /how will at-scale Cassandra APDB be integrated into ongoing AP development efforts? Slides 	<p>KT: sometimes, I found string reps of numbers compress better than binary reps. It ends up being more verbose but compressing better. AS: In this case, it's not numbers that are the problem. KT: Sounds not worth investigating.</p> <p>YA: Reconciling this what you said earlier about pandas being faster than afwTable. So converting to pandas dataframe is slow. But, converting to afwTable is slower? AS: Yep EB: The AP pipeline uses pandas and I'm not comfortable committing to refactoring at this point. AS: You can save money by having a smaller cluster. CS: It matters where you're doing it because it determines what you build into the AP cluster vs. the APDB cluster. Not fair to include the timing of the client when measuring the scaling of the DB server cluster. FM: I just want to bring to your attention that this conversion takes time, and you should be aware that it's there. [Notetaker's aside: There is lower hanging fruit on the client side]</p> <p>IS: It would be nice to see scaling with a factor 10x higher source density for the performance of outliers. AS: I'm worried about averages. The total number in the database is what's important.</p> <p>FM: Cassandra has been holding up as a horizontal scaling strategy. \$ In cloud vs. vendor: 1-2 years is the breakeven point The good thing is both are fine. The bad thing is it's a tough decision. MB: Make sure you include the test systems too! Also, can you grow this, or do you need 12 right away? FM: It'll only save us half a year at most. But I'm open to your advice.</p> <p>AS: Code is on a separate branch that has diverged significantly. EB: We took some action with Michelle and co and are using Postgres at NCSA. It's working out. We're working on gen3 migration issues and schema changes that we can put in. We have a way forward that's functional. It wasn't obvious that even during commissioning, would we hit the scale where we need Cassandra. How and when should we test this. Maybe we can separate it from when we go on sky timeline.</p> <p>FM: It sounds like you're OK with your Postgres solution at NCSA. We'll evolve that with the Cassandra API over the next few months. Do you think Postgres will get you through commissioning? EB: evolves with how long commissioning, but have to check with commissioning and the team.</p> <p>Wil: How big is it? 25-30TB per year? total 300TB KT: Archive it KT: Google has a way to snapshot your disks. Oh no google? OK, well, still archive it.</p>
12:30				
Moderator:	Notetaker: Frossie Economou			

13:00	Team status		<ul style="list-style-type: none"> • 9 minutes each • By coordinates ? <ul style="list-style-type: none"> ◦ UW 47.65, -122.30 Slides ◦ Princeton 40.34, -74.68 slides ◦ Urbana 40.11,-88.19 slides ◦ SF 37.76, -122.43 Arch ◦ Palo Alto 37.43, -122.15 slides ◦ Gibraltar 34.14 , -5.35 ◦ Tucson 32.20, -110.96 DM Science Plans & S QuaRE Update 	<p>* UW</p> <p>- Was the bug detected with the fake source injection pipeline an aperture correction bug? Yes.</p> <p>- How well are things working in Decam HITS and bulge fields? Technote for bulge data. Differencing artifacts remaining. Please ask Ian for more numbers.</p> <p>* Princeton</p> <p>- Please leave Jim alone on Focus Friday</p> <p>- What is the limit factor of the 3 iterations of HSC-R2? Waiting time is less for Gen3. Having more reruns would help "slightly"</p> <p>- Is there any way to get Leanne's group to look at Gen3 outputs? Triage process requires a lot of experience from Lauren.</p> <p>* NCSA</p> <p>- Tucson teststand PDUs delayed till April -- RHL in chat</p> <p>- No change in development, NCSA still working with Tony on main camera -- Wil in response to slide</p> <p>* Architecture</p> <p>- Build engineer ad out, please share -- Wil</p> <p>* DAX</p> <p>- No questions</p> <p>* SQuaRE</p> <p>- Where are we with the Science Platform landing page? squareone under early development, released next halfcycle</p> <p>* Science</p> <p>- Has Focus Friday really impacted Stack Club? [Resumption of earlier discussion about better cover for Stack Club by going back to assigning people to mind it - Wil will discuss slide phrasing]</p> <p>* Wrapup</p> <p>- Thanks to all for a good meeting</p> <p>- Let's use Slack instead of Zoom chat for side discussions next time [but then we can get distracted -- KT]</p> <p>- Action review</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Gregory Dubois-Felsmann to discuss the application of the OCPS/UWS to the image services with Kian-Tat Lim and Simon Krughoff 24 Mar 2021 <input checked="" type="checkbox"/> Eric Bellm to get together with Frossie Economou / SQuaRE and Colin Slater and Gregory Dubois-Felsmann to discuss hybrid alert model architecture 14 Apr 2021

14:10	Wrap up			next DMLTs: <ul style="list-style-type: none"> • 2021 June 8-10 - Clash with Penn State Stats • PCW not in person • 2021 October 26-28 - Clash with ADASS ... move? • 2022 February 15-17
14:30	Close			

Proposed Topics

Topic	Requested by	Time required (estimate)	Notes
Update on DMTN-139	Gregory Dubois-Felsmann		At the 1-11-2021 while reviewing DMLT ticket DM-15198 Gregory offered to give an update.
Alert Distribution & Brokering	Leanne Guy Eric Bellm	1hr	<ul style="list-style-type: none"> • Status report on the SAC review of the Broker proposals • Discussion of a proposal for a "hybrid" alert distribution system (dmtn-165.lsst.io); implications for the alert filtering service and alert DB
Verification plan through end construction	Leanne Guy	0.5-1hr	
Provenance WG	Leanne Guy	(depends on progress to report)	<ul style="list-style-type: none"> • Progress & Status
Campaign Management	Leanne Guy Kian-Tat Lim		Plans for campaign management
Codifying Slack etiquette	Kian-Tat Lim		<ul style="list-style-type: none"> • Thread usage • @ on every message ("DESC-style"?) • @channel usage • Do we need to write things down? If so, is the Dev Guide or Community or somewhere else (DMTN?) the best place?
APDB Update	Fritz Mueller	0.75-1hr	<ul style="list-style-type: none"> • APDB Cassandra scale experiments are concluding; summary report on recommended design and hardware requirements. • Coordination discussion w/ AP team: when/where/how will at-scale Cassandra APDB be integrated into ongoing AP development efforts?
Handing over the Community platform to the Ops.	Leanne Guy	30 mins	<ul style="list-style-type: none"> • DM is not scoped to provide support in construction. Now that we have CET funded in pre-ops and they are building a model for community engagement in operations, I'd like to hand the management and evolution of the community platform to the OPS-CET
QA plots/site	Colin Slater Leanne Guy Unknown User (mbutler)	20-30 min	<p>Leanne wishes to have publicly accessible QA plots (from <code>pipe_analysis</code> etc ..).</p> <p>There is a docker image for the site :</p> <ul style="list-style-type: none"> • this could be spun up on a login node • or it could (should?) be deployed on the cluster using ArgoCD etc .. <ul style="list-style-type: none"> ◦ but then it will not be public <ul style="list-style-type: none"> ▪ unless we deploy on google and push the data to it .. ▪ then what ever does the pushing should be deployed properly
Ops Rehearsal	Wil O'Mullane	10-20min	<p>Quick check in on OR#2 for commissioning - we put this off for Gen3 and ComCam move to summit etc.</p> <ul style="list-style-type: none"> • Now we have Gen3 repos - • ComCam is up on Floor3. • Shoud we do something to ensure we are ready for the integration on Floor 3 ? <ul style="list-style-type: none"> ◦ EFD , ComCam, Hexpod , cable wrap .. ◦ Or a repeat of last time but with Gen3

Container builds	Frossie Economou	30min-1hr	Looking for efficiencies in: <ul style="list-style-type: none"> Container builds across RSP, Prompt Processing, DRP, Telescope & Site Non-service non-stack packaging
------------------	----------------------------------	-----------	---

Attached Documents

File	Modified
PDF File Bellm_Rubin_alerts_210223.pdf	Feb 23, 2021 by Eric Bellm
PDF File DMLT_CommunityPlatform.pdf	Feb 24, 2021 by Melissa Graham
PDF File Campaign Management KTL.pdf	Feb 24, 2021 by Kian-Tat Lim
PDF File Slack Etiquette KTL.pdf	Feb 24, 2021 by Kian-Tat Lim
PDF File Sullivan-AP-Status-February-2021.pdf	Feb 24, 2021 by Ian Sullivan
PDF File DMLT-imageArch-20210224-v1.pdf Added screenshot	Feb 24, 2021 by Gregory Dubois-Felsmann
PDF File Arch S21B Plans.pdf	Feb 24, 2021 by Kian-Tat Lim
PDF File NCSA next 3 months (2-2021).pdf	Feb 26, 2021 by mbutler
PDF File DMLT vF2F DM Science 2021-02-23_24.pdf	Feb 26, 2021 by Leanne Guy
PDF File DAX Mid S21 Status_Plans.pdf	Feb 26, 2021 by Fritz Mueller
PDF File APDB Update.pdf	Feb 26, 2021 by Fritz Mueller
PDF File dmlt_2021_feb_status.pdf SQuaRE update	Feb 26, 2021 by Frossie Economou
PDF File DRPActivities202102.pdf	Feb 26, 2021 by Yusra AlSayyad
PNG File Screenshot 2021-02-24 at 22.57.53.png	Apr 10, 2021 by Wil O'Mullane

[Download All](#)

Action Item Summary

Description	Due date	Assignee	Task appears on
<input type="checkbox"/> Frossie Economou Will recommend additional Level 3 milestones for implementation beyond just the DAX-9 Butler provenance milestone. 15 Mar 2022	15 Mar 2022	Frossie Economou	DM Leadership Team Virtual Face-to-Face Meeting, 2022-02-15 to 17
<input type="checkbox"/> Kian-Tat Lim Convene a meeting with Colin, Tim, Robert, Yusra to resolve graph generation with per-dataset quantities (likely based on Consolidated DB work). 18 Mar 2022	18 Mar 2022	Kian-Tat Lim	DM Leadership Team Virtual Face-to-Face Meeting, 2022-02-15 to 17
<input type="checkbox"/> Frossie Economou Write an initial draft in the Dev Guide for what "best effort" support means 17 Nov 2023	17 Nov 2023	Frossie Economou	DM Leadership Team Virtual Face-to-Face Meeting - 2023-Oct-24

<input type="checkbox"/> Convene a group to redo the T-12 month DRP diagram and define scope expectations Yusra AlSayyad 30 Nov 2023	30 Nov 2023	Yusra AlSayyad	DM Leadership Team Virtual Face-to-Face Meeting - 2023-Oct-24
<input type="checkbox"/> Gregory Dubois-Felsmann Complete DMTN-105 defining the goal for "Prompt Products Release Ops" 11 Dec 2023	11 Dec 2023	Gregory Dubois-Felsmann	DM Leadership Team Virtual Face-to-Face Meeting - 2023-Oct-24
<input type="checkbox"/> Frossie Economou - you have some text for PSTN-017 02 May 2024	02 May 2024	Frossie Economou	DMLT Meeting - 2024-04-22
<input type="checkbox"/> DMLT Meeting - 2024-04-22 UKDF part in data facilities for PSTN-017 22 May 2024	22 May 2024		DMLT Meeting - 2024-04-22
<input type="checkbox"/> Richard Dubois USDF part in data facilities for PSTN-017 and distrib processing ? 22 May 2024	22 May 2024	Richard Dubois	DMLT Meeting - 2024-04-22
<input type="checkbox"/> Fabio Hernandez FrDF part in data facilities for PSTN-017 22 May 2024	22 May 2024	Fabio Hernandez	DMLT Meeting - 2024-04-22
<input type="checkbox"/> Cristián Silva - section on summit/data acquisition for PSTN-017 22 May 2024	22 May 2024	Cristián Silva	DMLT Meeting - 2024-04-22
<input type="checkbox"/> Tim Jenness - section on middleware for PSTN-017 22 May 2024	22 May 2024	Tim Jenness	DMLT Meeting - 2024-04-22
<input type="checkbox"/> Richard Dubois who from SLAC?		Richard Dubois	DMLT Meeting - 2024-04-29
<input type="checkbox"/> Cristián Silva if you come to SLAC lets have plenty of photos of locks and racks etc ..		Cristián Silva	DMLT Meeting - 2024-04-29