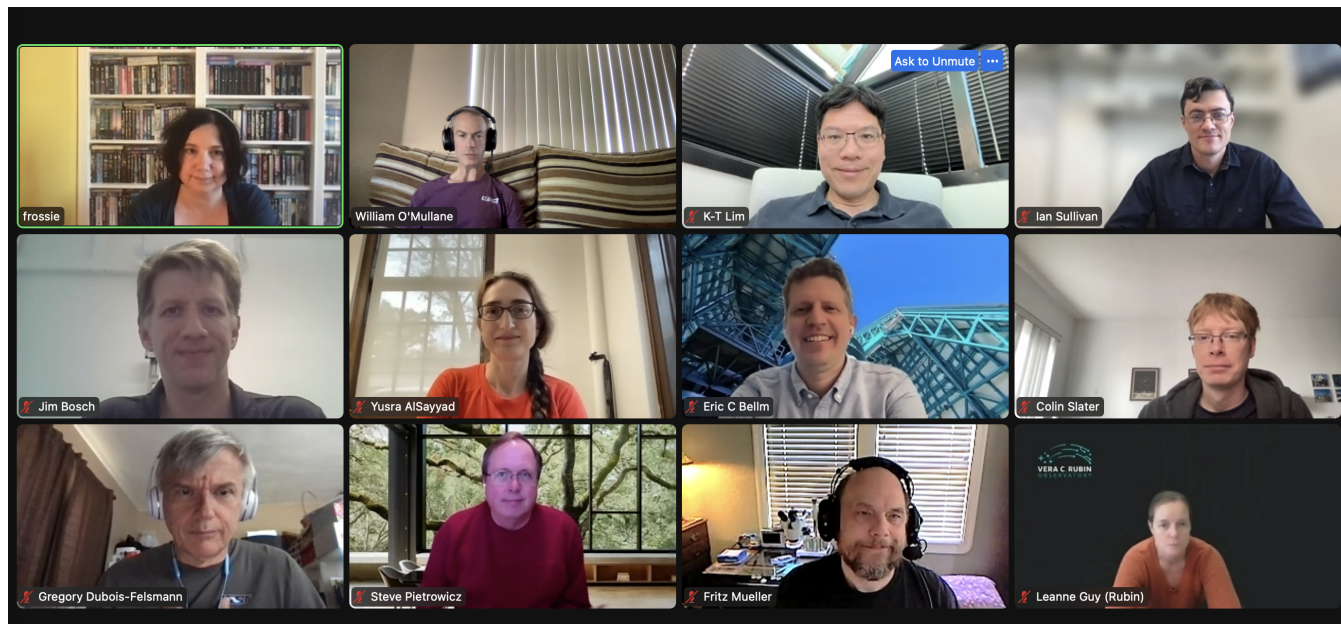


DM Leadership Team Virtual Face-to-Face Meeting - 2022-10-18



- Logistics
 - Date
 - Day 1, Tuesday
 - Day 2, Wednesday
- Proposed Topics
 - Attached Documents
- Action Item Summary

Logistics

Date

2022 October 18 and 19

Location

This meeting will be **virtual**;

Join Zoom Meeting (DMLT link)

Meeting: <https://noirlab-edu.zoom.us/j/93412065536?pwd=QWhvb29kbVI5NEZNVFMrR0dzN0RBZz09>

Meeting ID: 934 1206 5536

Password: 752226

Attendees:

All technical sessions of DMLT vF2F will be open to DM members. Should there be need for a closed session it will be marked "DMLT only".

Day 1, Tuesday

Time (Project)	Topic	Coordinator	Pre-meeting notes
----------------	-------	-------------	-------------------

Moderator: Wil O'Mullane		Notetaker: Frossie Economou	
09:00	Welcome - Project news and updates	Wil O'Mullane	
09:30 (DMLT only)	BPS plugin support /Processing USDF /Summit	Colin Slater Tim Jenness	Decide whether we have an opinion on summit batch and developer-driven processing (parsl vs condor vs panda). We are happy to support 2 bps plugins (probably not 3).
10:00	User Generated Data Catalogs	Fritz Mueller	Report on a general design and pieces required for support of user generated data products in the RSP, so work planning can begin.
10:30	Break		
Moderator: Frossie Economou		Notetaker: Kian-Tat Lim	
11:00	Commissioning Cluster, Yagan, and do we have what we need?	Cristián Silva	Primary objective: Do we need to buy more nodes to increase the cluster's cores and ram? Secondary objective: What do we need to run in Yagan

11:30	OPS planning and milestones	Will O'Mullane	<div><div>keysummarytypecreatedupdateddueassigneereporterprioritystatusri</div><div><div><div><div><div><div></div></div><div>Jira project doesn't exist or you don't have permission to view it.</div></div><div><div>View these issues in Jira</div></div></div></div></div></div>
12:15	Meeting free weeks 2023	Will O'Mullane	<div>Agree dates for meeting free weeks 2023<ul style="list-style-type: none">• Spring :JTM - March 14-16 so March 20, Or April 3 or April 10 (around Easter)• Summer: June 12 - 16• Autumn: Sept 25 - 29• Winter: Dec 25 - Jan 5</div>
12:30	Break		
Moderator: Lianne Guy	Notetaker: Will O'Mullane		

1:00	Prompt Processing migration to the USDF	Eric Bellm	slides
1:45	close		

- NB: IVOA meeting "DAL" session, mainly on ADQL, is 13:30-14:30 PT on Tuesday 2022-10-18, partially overlapping the planned "Prompt Processing migration to the USDF" session.

Day 2, Wednesday

Moderator: Will O'Mullane	Notetaker: Yusra AlSayyad
---	---

9:00	<p>Status and plans:</p> <p>DAX</p> <p>SQuARE</p> <p>DRP</p> <p>IT DevOPs</p> <p>Arch</p> <p>AP</p> <p>NCSA</p> <p>DM Science</p>	<p>Fritz Mueller</p> <p>Frossie Economou Yusra</p> <p>AlSayyad</p> <p>Cristián Silva</p> <p>Kian-Tat Lim</p> <p>Ian Sullivan</p> <p>Steve Pietrowicz</p> <p>Leanne Guy</p>	<p>>>> presentors = ['Leanne', 'Cristian', 'Frossie', 'KT', 'Fritz', 'Stephen', 'Ian', 'Yusra']</p> <p>>>> random.shuffle (presentors)</p> <p>>>> presenters.remove ('Fritz')</p> <p>>>> presenters.append ('Fritz')</p> <p>(no?)</p> <p>['Fritz', 'Frossie', 'Yusra', 'Cristian', 'KT', 'Ian', 'Stephen', 'Leanne']</p> <p>DAX F22/S23 Status_Plans.pdf</p> <p>Cristian Silva - DMLT vF2F October 2022.pdf</p> <p>Arch F23A Status and Plans.pdf</p> <p>DM Science Retrospective and Plans F22AB</p> <p>NCSA DMLT F2F Oct 2022.pdf</p>	<ul style="list-style-type: none"> • K-T: Is it possible to pay CADC? • Frossie: No, but it is possible to trade some of our systems with some other systems. • YA: Been announcing since March, but if you didn't get the memo with Scarlet Lite's computational performance improvements, we're not worried about having resources to run it. • Wil: Does this mean we can retire that risk? • YA: No, because that risk was about ALL algorithms, and just because we're not worried about Scarlet anymore doesn't mean that we're not worried about things like galaxy measurement. We still have more algos than we have compute for, and there will be hard tradeoffs to be made. <p>Wil: And Andres is going to give a talk in Spanish for us next week.</p> <ul style="list-style-type: none"> • Colin: The backup link, that's microwave from the summit to where? • Cristian: Just summit to base. It'll be invisible. When the fiber goes down, we change the priority to wireless. It's only for control data, not science data. • Wil: We had a network meeting with SLAC last week. Are they OK with the routers? • Cristian: Yep. We need a list of what needs to be reachable from Rubin. • Wil: The question from (didn't catch the name?) about the suitability of the routers for installation? • Cristian: Oh, we didn't get to that. We gave him the specs. It's a router behind our routers at the summit. Another at SLAC. one for the 100, one for the 40 <p><input checked="" type="checkbox"/> Richard Dubois Confirm on your side that SLAC is OK with the router specs? 03 Nov 2022</p> <p>Wil: Cristian, we should talk with Kevin and give him a delivery time so that he can put in P6 a date that it will be in place. Then we can communicate with NSF about something happening. We have a milestone but need an activity.</p> <ul style="list-style-type: none"> • Wil: Good about the Postgres. Tiago was showing me this serverless Postgres he set up, and I told him that you have something like that with Kubernetes that he can use. • Richard: Do you need help from the Fermilab DBA folks to set up the Postgres on Kubernetes? • Cristian: Sure! The servers are already running. Who do I write? • Richard: <not taker didn't catch the name>. • Wil: We'll want the help once everyone is adding their schemas. • K-T: to the extent that they are using the butler, it's just one schema. I think it's up to Steve and me to migrate the SQLite DB for the OODS over to Postgres. Got bogged down. • Steve: Yeah, I'm in contact with Bruno about that. • Frossie: K-T, What was the metrics thing? • K-T: Thinking about how to get metrics into Sasquatch. To my knowledge, the way it's done down is outside of the pipeline. • Frossie: Yeah, but Angelo is working on it now. • Richard: Jenkins workers running in the cloud... what's the need to move it to SLAC? • K-T: no requirement, but it would allow us to do things Merlin asked for. He wants to run pipeline tests against live butler repos. We never had it at NCSA, even though we could have. • Frossie: Those jobs were run at the DF so that we could run deep tests with the production system. • K-T depends on how much we want to own everything • Cristian: Is LFA replication bi-directional? • K-T: No • Cristian: we could use our multi-site replication! • Wil: Let's punt this to a future discussion. <p><input checked="" type="checkbox"/> Kian-Tat Lim Convene a meeting following up vF2F discussion about what to do with Jenkins 24 Nov 2022</p>
10:30			Break	
Moderator: Wil O'Mullane	Notetaker: Ian Sullivan			

11:00	Needs for "tactical" databases	Robert Lupton	See RHL's rough notes	<ul style="list-style-type: none"> • FE: There are many producers and many consumers of metrics. Each has different needs <ul style="list-style-type: none"> ◦ Elephant in the room is the consolidated database. Metrics were going to be one of the things persisted with the Butler ◦ Stuff that needs to go from USDF to the summit needs to go through Sasquatch ◦ Stuff that needs to be presented to the user needs to be in the consolidated database • RHL: thinks there could be other ways of moving stuff from USDF to the summit <ul style="list-style-type: none"> ◦ this is a proposal for what we need, which might be what the consolidated database may be • FE: We have tooling other than Chronograf on the summit, also have notebooks • KTL: Rucio is another method to transfer (calibration) datasets from the USDF to the summit <ul style="list-style-type: none"> ◦ Any time series-only data should go in the database • JB: interested in looking into the database keys that are explicitly time oriented. <ul style="list-style-type: none"> ◦ WOM: should follow up outside this meeting • FE: Concern about camera data that is not in DM-land. • RHL: Interested in implementation here. The camera team has tools, how do we use them? <ul style="list-style-type: none"> ◦ WOM: We can port the tools, and make the queries talk to influx DB. • FE: Kafka plus influx is Sasquatch. Is it true that camera telemetry scalars are going in the EFD? <ul style="list-style-type: none"> ◦ RHL: Almost ◦ WOM: will be down there soon, and can assist with the port if needed. • KTL: This mostly deals with the keys, not the values. Those will change over time, and will require talking with the scientists • FE: I don't see how this schema gets defined without GPDF. We need data engineers • GPDF: I don't need to be involved in the definition of each table, but yes to deciding what keys are time varying etc. • WOM: Once we have postgres at the summit, we should start putting schemas in it • JB: What kind of skills does the person setting up the schema at the summit need to have? <ul style="list-style-type: none"> ◦ WOM: that is what I am trying to work out. ◦ CS: The need is for someone to understand the conditions at the summit ◦ WOM: That's Patrick, Tiago, Erik, Robert <ul style="list-style-type: none"> ▪ Question is whether we can make use of general knowledge of people available at Fermilab <p>✓ Kian-Tat Lim Determine whether people at Fermilab can assist with setting up schemas for the summit 02 Nov 2022</p> <p>✓ Robert Lupton Identify point of contact on commissioning side for summit schema 26 Oct 2022</p> <p>✓ Robert Lupton Kian-Tat Lim Kick off meeting between commissioning, Fermilab, and summit (Carlos) for summit schema 09 Nov 2022</p>
-------	--------------------------------	---------------	---------------------------------------	---

11:30 (DMLT only)	Wrap Up /Actions /AOB/	Wil O'Mullane		<ul style="list-style-type: none"> YA: Campaign management reorganization (slides) <ul style="list-style-type: none"> LG: Why are Colin and Yusra interim product owner and group lead? <ul style="list-style-type: none"> YA: if someone signs up for the campaign management team, I want to leave these roles open to them. LG: might be good to write current, not interim YA: hoping that someone will step up and want to do the group lead role. Agree current is a better term KT: pilot/co-pilot rotation sounds like a great idea, but people are not necessarily replaceable in what they were working on before. Is it expected that people will pause other work, and it will be restarted when they return from campaign management? YA: I don't want people to feel trapped in this role, which can easily happen. RD: Gotta run. We discussed this at the retreat so I'm ok with it. Remember there can be members from Europe. We also worried that 100% would indeed burn people out. EB: There may be distinctions between how this looks for DRP/AP/commissioning. <ul style="list-style-type: none"> WOM: Frossie has some org charts, we will write it up with more detail GPDF: very supportive. Looks a lot like the mechanisms we set up eventually with BABAR. 100% agree this tends to burn people out; there is a small subset of people who enjoy it. <ul style="list-style-type: none"> agree with EB that there will be significant differences in rhythm between AP and DRP LG: question for Colin: <ul style="list-style-type: none"> do you see different people in V&V rotating through as well? CS: This is not a position I intent to sit in to hold it, would be happy for someone else to step up as product owner, but it is too important to leave empty at first LG: I think it would be very helpful to have V&V team members rotate through, to spread experience. <ul style="list-style-type: none"> WOM: really like the idea of more people having experience organizing /leading. Builds resilience FE (slides) <ul style="list-style-type: none"> This plan that YA proposed came out of DPP retreat last month. <ul style="list-style-type: none"> Campaign management now part of Data Production under YA, next to Algorithms&Pipelines Qserv now in Data Services under FE FE: format of this meeting <ul style="list-style-type: none"> Propose that we upload status slides in advance, use the timeslot for questions and discussion (standup style) FM: finds it useful to write/present plan in response to the discussions from DMLT, would find it less useful to prepare in advance JB: slides written to be read are different than ones written for presentation, since that allows greater clarification WOM: We used to reserve some time at the end for discussion WOM: suggestion is to start last session later, so that people have time to finish their presentations. Provide drafts in advance. Provide time for reading the presentations, and schedule time (~1hr) for discussion afterwards.
12:30	Close			

Proposed Topics

Topic	Requested by	Time required (estimate)	Notes
Prompt Processing migration to the USDF	Eric Bellm	45 minutes?	Prefer scheduling after lunch Tuesday
Meeting free weeks 20223	Wil O'Mullane	15m	Agree dates for meeting free weeks 2023 <ul style="list-style-type: none"> Spring :JTM - March 14-16 so March 20, Or April 3 or April 10 (around Easter) Summer: June 12 - 16 Autumn: Sept 25 - 29 Winter: Dec 25 - Jan 5
User Generated Data Products	Fritz Mueller	30m	
Commissioning Cluster, Yagan, and do we have what we need?	Cristián Silva	30m or less	
Processing frameworks at USDF	Colin Slater	30m	Re: developer-driven processing. Closed-session please. CTS: I think this is the same as what Tim is proposing below
OPS planning and milestones	@womullan	1h	See the epics we have and what milestones make sense - block milestones with epics.
BPS plugin support	Tim Jenness		Decide whether we have an opinion on summit batch and developer batch (parsl vs condor vs panda). We are happy to support 2 bps plugins.

Attached Documents

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"/> Tim Jenness - section on middleware for PSTN-017 22 May 2024	22 May 2024	Tim Jenness	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