DM Leadership Team Face-to-Face Meeting, 2018-05-22 to 24



Full res on flickr

Logistics

Date

22 May 2018 - 24 May 2018

Location

- Petersen Room, Allen Library, University of Washington.
- Room 485 (fourth floor).
- Room is available 09:00 on 22 May to 13:00 on 24 May.

Accommodation

- A hotel block booking is available at the Residence Inn by Marriott Seattle University District.
- Contact Erin Carlson to secure your place.
- Bookings must be made by 2018-04-20. ** The hotel block is now full, but Erin Carlson may be able to provide assistance **
- See details here.

Travel

- From Seattle-Tacoma International Airport, the Link light rail will take you to the University of Washington Station.
- The Allen Library (see location, above) is a ten minute walk from the station.
- Ride-share services (Lyft, Uber, etc) will likely cost around \$30-40 from the airport to the University District.

Per Diem

| Total | Breakfast | Lunch | Dinner | Incidentals | First & Last Day of Travel | |
|---------|-----------|---------|---------|-------------|----------------------------|--|
| \$74.00 | \$17.00 | \$18.00 | \$34.00 | \$5.00 | \$55.50 | |

Attendees

Traveling

- Wil O'Mullane (arriving SEA Monday May 21 at 16:54)
- Frossie Economou (arrive SEA Monday 14:36, depart SEA Thursday 14:40)
- Tim Jenness (arrive SEA Monday 14:36)

- Kian-Tat Lim (arrive SEA Monday 19:44, depart SEA Thursday 21:15)

- Leanne Guy (arrive SEA Sunday 14:36, depart SEA Thursday 17:05)
 Unknown User (xiuqin) (Arriving 6:47pm Monday, Alaska-589)
 Fritz Mueller (Arrive SEA Monday 14:10, depart SEA Thursday 09:10)
 Vaikunth Thukral (Arrive SEA Monday 16:25, depart SEA Thursday 17:10)
- Jim Bosch (arriving 19:06 Sunday on UA2151, departing 22:50 Thursday on UA1695)
- Unknown User (gcomoretto)
- Simon Krughoff
- Yusra AlSayyad (arriving 5/20 21:55 UA1807, departing 5/26 11:50 UA1730)
 Unknown User (mbutler) (5/21-5/23 only)
- Margaret Gelman
- Jeff Kantor (5/22 and 5/23 only, arriving 5/21 3pm, leaving 720am 5/24)
- Michael Wood-Vasey (5/20-5/23: arr SEA Sun 20:58 dep SEA Wed 22:23)
- Robert Gruendl
- Gregory Dubois-Felsmann (also attending DM-SST 5/21 and IVOA InterOp starting 5/27)
- ...

Local

- John Swinbank
- Eric Bellm
- Colin Slater
- Zeljko Ivezic

Remote

• ...

Regrets

- Mario Juric
- Robert Lupton

Agenda

| | Day 0: 2018-05-21 | | | | |
|-------|--|--------------------------------------|---|--|--|
| Time | Торіс | Chair | Notes | | |
| 09:00 | SST Meeting | Zeljko Ivezic / L eanne Guy | Agenda provided by DM System Science Team. | | |
| 17:00 | Close | | | | |
| | | | Day 1: 2018-05-22 | | |
| 09:00 | Welcome | Wil O'Mullane | Review outstanding actions Review & adjust agenda https://github.com/lsst/wom_presentations/blob/master/DMLT-May-2018/DMLT-WOM-MAY-2018.pdf | | |
| 09:30 | Review outstand ing DMLT action items and tickets | Wil O'Mullane | Action items Jira tickets | | |

| 10:00 | Compre ssion WG | Wil O'Mullane / Robert Gruendl | Final close out of this working group. Please review the WG report in advance of the meeting. https://dmtn-068.lsst.io/ Includes plans & requirements for tuture work in this area. LossyCompress_WG_20180522.pdf The proposal is that we can store compressed versions of PVIs for easy access for (e.g.) producing postage stamps. Note that we are currently sized to store 30 days worth of (uncompressed) PVIs. Anything older than that needs to be recomputed (by imgsery). We, as DM, meed to develop a "menu" of -3 options for science collaborations, including: Time to produce calibrated image from raw: Compression level; Time range of compressed images stored. This menu can likely be appended to the end of MRC-325 - Jira project doesn't exist or you don't have permission to view it. John Swinbank: Ensure pipelines code for regenerating images on demand is included in LDM-151, that an equivalent service is included in the appropriate DAX design documentation, and that both of them are included in the DM construction plan. 19 Jun 2018 Wil O'Mullane, Kian-Tat Lim: Provide "menu" of compression of options as an addition to RFC-325. 19 Jun 2018 Fritz Mueller: Develop a complete (to the best of our current understanding) list of data products that have to be accessible through DAX services, and a plan for extending that list in future. 19 Jun 2018 | | |
|-------|---|--|---|--|--|
| 10:30 | Break (Re | freshments | Provided) | | |
| 11:00 | QA WG | John Swinbank | Report on current status. Slides. | | |
| 11:30 | Summit– Base ITC & network s status update | Jeff Kantor | Provide a status of the Summit and Base IT infrastructure installation and networks. Kantor Networks and Base.pdf | | |
| 12:00 | Status of HEALPi x/MOC | Unknow n User (xiuqin) / Gregor y Dubois- Felsmann | Per RFC-441 HEALPix and MOC should be adopted as standard LSST data products. What is the plan for and status of implementation of this? Should we draft some requirements ? Slides Gregory Dubois-Felsmann: Discuss potential HiPS/MOC plans with the SST. 19 Jun 2018 | | |
| 12:30 | Lunch (Pro | Lunch (Provided) | | | |
| 13:30 | Next-to- databas e processi ng | Leanne Guy / Fri tz Mueller | Follow up on actions from DM Leadership Team Face-to-Face Meeting, 2018-03-08/09. Colin Slater: Publish technical note on next-to-database processing use cases. 06 Jul 2018 Wil O'Mullane: Following the publication of Colin Slater's technical note on next-to-database processing, decide what further meetings or discussions are required. 26 Jun 2018 | | |

| 14:00 | Pipeline s output metadata | Kian- Tat Lim | Plans for having LSST Science Pipelines provide the metadata that is required by DAX & SUIT as part of standard processing. Slides | | | | |
|-------|--|---|---|--|--|--|--|
| | | | Gregory Dubois-Felsmann: Refresh Confluence page describing the mapping of the LSST data model to CAOM, then propose it for RFC. 19 Jun 2018 | | | | |
| | | | Kian-Tat Lim: Map Butler v3 schema to CAOM. Delegated in | | | | |
| | | DM-14555 - Jira project doesn't exist or you don't have permission to view 19 Jun 2018 it. | | | | | |
| | John Swinbank / Yusra AlSayyad: Schedule work during F18 to construct a task which translates pipeline outputs to a form consistent with the DPDD. 31 May 2018 | | | | | | |
| | | | Gregory Dubois-Felsmann: Provide "meta-schema" of information which must be provided in pipelines output tables. 19 Jun 2018 | | | | |
| 14:30 | Nightly reporting | Wil O'Mullane | Status of LCR-1203 Its impact on DM requirements Other reporting tools needed (? e.g. real time alert stream diagnostics) Wil O'Mullane: Shepherd LCR-1203 to a useful conclusion. 24 Jul 2018 | | | | |
| 15:00 | Break | | | | | | |
| 15:30 | Moderni zing the Science Pipeline s codebase | Bosch • Jim Bosch writes: or There is currently a lot of ad-hoc, grass-roots effort going into cleaning up the stack and its documentation that has at best inconsistent support from the DMLT, and this is both distracting developers from scheduled work and leaving them frustrated because the effort they're donating via free/science time isn't making a big dent. DMLT needs to agree to and present | | | | | |
| 17:00 | Close | | | | | | |
| 18:00 | | | o join the DM Project Manager & the local UW LSST team for drinks at Ravenna Brewing, 5408 26th Ave NE (25 minutes walk from the inutes by bus). | | | | |
| 19:30 | Dinner at | BaBar, <mark>268</mark> | 5 NE 46th St (10 minutes walk from Ravenna Brewing). | | | | |
| | | | for dinner, please provide Wil O'Mullane with \$34 (your dinner per diem). | | | | |
| | | | ning for dinner, please let John Swinbank know in advance so we don't wait for you. | | | | |
| | The resta | urant will ac | commodate all the dietary needs that we know of; please make your needs known to the server. | | | | |
| | - | | Day 2: 2018-05-23 | | | | |
| 09:00 | EFD interface | Frossie Econom | Discuss LCR-1344 and the path forward with starting to exercise DM interfaces to the EFD. | | | | |
| | s & requirem | ou | We agreed to appoint Simon Krughoff as product owner of the DM-EFD. | | | | |
| | ents | | Margaret Gelman: Organize a meeting with T&S to decide whether we extract from the EFD databases or from DDL. 19 Jun 2018 | | | | |
| | | | Simon Krughoff: Develop written use cases/requirements for DM-EFD. This is being handled in | | | | |
| | DM-15012 - Jira project doesn't exist or you don't have permission to view it. | | | | | | |
| | | | | | | | |

| 09:30 | Generati on 3 Middlew are next steps | Fritz Mueller / Jim Bosch | • Fritz : | What's the current status of Gen 3 Butler & SuperTask as at the end of S18? When will they be ready for widespread use? What's our migration plan? What resources are needed to continue middleware development into F18 and beyond? Fritz slides (F18 Supertask status) ✓ Fritz Mueller: Organize "mini review" of RFC-484 - Jira project doesn't exist or you don't have permission to view | | |
|-------|---|--------------------------------------|--|--|--|--|
| | | | | Fritz Mueller / Jim Bosch / Yusra AlSayyad / John Swinbank: Discuss staffing plans for the F18 BG3 group. 25 May 2018 Fritz Mueller: Provide example QuantumGraph to the Data Facility. 19 Jun 2018 Initial example (illustrative of QG API) was sent. Andy S. awaiting feedback/questions about the the API from NCSA. | | |
| 10:30 | Break (Re | freshments | Provid | Jed) | | |
| 11:00 | Test specifica tions in Jira | Unknow n User (gcomor etto) | | Can DM adopt the technology being developed by the Systems Engineering group? GCM_TestApproach.pptx | | |
| 11:30 | Status of docume ntation for this summer' s reviews | Wil O'Mullane | | | | |
| 12:30 | Lunch (Pr | ovided) | | | | |
| 13:30 | Codeba se consiste ncy | Kian- Tat Lim | Portal, Firefly, Qserv, Header Service, "L1" Image Ingest, Workload/Workflow, Data Backbone, Jellybean, SQuaSH: these use different languages, different dependencies, different skills than Science Pipelines, Task Framework, Butler. Their code is generally not reviewed, understood, or maintained outside of their own groups. What standards and processes are we and shoul we be applying to these to ensure that they function well and are maintainable for the lifetime of the project? Slides Kian-Tat Lim: Figure out which (project and/or DM level) documents need to be updated to better capture standards & processes for development within DM and elsewhere. | | | |
| | LIT-337 - Jira project doesn't exist or you don't have permission to view and it. | | | | | |
| | | | | DM-14570 - Jira project doesn't exist or you don't have permission to view 19 Jun 2018 it. | | |

| 14:15 | Operatio ns rehearsa Is | Margare t Gelman | We have multiple "operations rehearsal" milestones scheduled through construction: LDM-503-09 (2019-01-04) LDM-503-11 (2019-10-31) LDM-503-12 (2020-01-31) LDM-503-13 (2020-11-30) LDM-503-15 (2021-10-29) LDM-503-16 (2022-05-31) We should start planning for what's involved in these activities. Slides Margaret Gelman: Circulate a skeleton plan for an operations rehearsal based on AuxTel. 19 Jun 2018 Margaret Gelman: Arrange for Felipe to construct realistic (in terms of file structure, headers, etc) images based on TS8 data. 19 Jun 2018 |
|-------|---|--------------------------------------|---|
| 15:00 | Break | | |
| 15:07 | Photo at t | he Fountair | - out the door turn right and its on the left |
| 15:30 | Internal on-disk data represen tations. | Simon Krughoff | Currently, the standard processing results in FITS files containing the afwTable catalogs from the processing. These are then transformed to Apache parquet format for use in QA endeavors. This means we are storing the same information in two redundant formats. I would like to explore how we treat the various file formats. Do we treat some as transient and others as persistent? Can we settle on a single format that will serve multiple needs with the understanding that some transformation may be necessary to serve the community? Depending on the above, are there impacts on the sizing model that need to be taken into account (realizing that the sizing model is specifically for operations)? |
| 16:00 | Product tree | Unknow n User (gcomor etto) | Based on the different documents available (and in preparation), I would like to share some thoughts, maybe a proposal, for a categorization of the different products that can be identified (at the moment) inside DMS. (GCM_ProductTree.pptx) |
| 16:30 | Future software releases | Unknow n User (gcomor etto) | What's the plan for upcoming releases? Will the next release be 15.1 or 16.0? What's the timescale? What documentation and/or characterization needs to be prepared to accompany it? Are future releases time based, or linked to specific events in commissioning/l&T/etc? Note that we now have LDM-503 milestones described as "Science Pipelines Release" preceding major rehearsals. Do these releases include just Pipelines or also Firefly, Qserv, Platform, Data Facility,? If those are not included in the stack release, what are their release plans? GCM_Releases.pptx Next release will be 16.0.0, coming in June. No changes to standard release process (modulo addition of extra "0"). |
| 17:00 | Close | | |
| 17.00 | 01036 | | Day 3: 2018-05-24 |

| 09:00 | Plans for F18 | John Swinbank | Each T/CAM (or science lead, or their nominee) to give a ~10 minute presentation of plans for F18. System Management (Wil O'Mullane) System Science (Leanne Guy) Architecture (Kian-Tat Lim) Science Pipelines: Alert Production (John Swinbank) Calibration Products (John Swinbank) Data Release Production (Yusra AlSayyad) SUIT_F18plan (Unknown User (xiuqin)) DAX (Vaikunth Thukral) Data Facility (Margaret Gelman) SQuaRE (Frossie Economou) | | | | |
|---|---|---|---|--|--|--|--|
| | | | Margaret Gelman: Discuss infrastructure for notebook aspect in PDAC at infrastructure meeting. 31 May 2018 | | | | |
| | | | Colin Slater: Provide Frossie Economou with details on outstanding data products documentation request tickets. 19 Jun 2018 | | | | |
| | | | Unknown User (xiuqin): Discuss requirements for spectral data visualization with key stakeholders (primarily Patrick Ingraham and | | | | |
| | | | Robert Lupton). 19 Jun 2018 DM-15475 - Jira project doesn't exist or you don't have permission to view it. | | | | |
| | | | Kian-Tat Lim: Work within a framework developed with Margaret Gelman to develop a list of all the various data products generated or managed by DM. 19 Jun 2018 Wil O'Mullane: Check prerequisites for LDM-503-10 and adjust dates (likely one year later). 19 Jun 2018 | | | | |
| Milestone "owners": LDM-503-05: Eric/John LDM-503-06: Margaret (to move to more appropriate date?) LDM-503-07: John (should be changed to say "making test stand data available through a LDM-503-08: Michelle LDM-503-08b: Margaret LDM-503-09a: John LDM-503-09: Margaret (to move to more appropriate date) | | LDM-503-05: Eric/John LDM-503-06: Margaret (to move to more appropriate date?) LDM-503-07: John (should be changed to say "making test stand data available through a Butler".) LDM-503-08: Michelle LDM-503-08b: Margaret LDM-503-09a: John | | | | | |
| 10:30 | Break (Re | freshments | Provided) | | | | |
| 11:00 | Plans for F18 (continu ed) | John Swinbank | Resolve outstanding cross-team issues. Agree on final plan. | | | | |
| 12:30 | Review action items & plans for next meeting | Wil O'Mullane | Expected to be 2018-11-05/07 in Princeton. Also lets open diaries for the next meeting or two after that - Tucson and Illinois perhaps. Provisional dates for next-but-one meeting 26/27/28 February likely in Tucson. Provisional dates for next-but-two meeting 21/22/23 May likely at NCSA. | | | | |
| | | | | | | | |

Pre-Meeting Planning

Suggested topics for discussion

- Check on status of requirements for next to database processing, per discussion at DM Leadership Team Face-to-Face Meeting, 2018-03-08/09.
- State of Butler Gen 3 and next steps how are we preparing to make it work with Pegasus ? ٠
- Status of documentation for JSR - https://docs.google.com/spreadsheets/d/15WZbLHcHZGEf4GzzSnfmiklfnCw6Z8lwNDwP4Qvi2Dw/edit? usp=sharing Technical topics
- - HEALPix/MOCs
 - ° Metadata (units, UCDs, descriptions, linkages, etc.) by the pipelines
 - Nightly reporting
 - The relationship between DBB and LSP user file storage space. This was brought up at the LSP workshop December 5, 2018. Several of us think there may need to be a working group for this.

| Торіс | Requested by | Time required (estimate) | Notes |
|--|---------------------------------|---------------------------------|---|
| Update on status of Summit - Base ITC and Networks | Jeff Kantor | 20 minutes plus questions | Provide a status of the Summit and Base IT infrastructure installation and networks. |
| Modernizing the Science Pipelines codebase, or agreeing how not to. | Jim Bosch | 60-90 minutes | There is currently a lot of ad-hoc, grass-roots effort going into cleaning up the stack and its documentation that has at best inconsistent support from the DMLT, and this is both distracting developers from scheduled work and leaving them frustrated because the effort they're donating via free/science time isn't making a big dent. DMLT needs to agree to and present a united and coherent policy and plan for modernization so we aren't fighting our own devs (and each other) on this. I will come with at least one rough proposal, but it may be too ambitious to be tenable, and I'd really like other members of DMLT to come armed with additional proposals. |
| EFD requirements and interfaces | Frossie Economou | 30 mins | Discuss https://project.lsst.org/groups/ccb/node/2470 and the path forward with starting to exercise DM interfaces to the EFD, |
| Test Specs in jira/MD with Unknow n User (gcomoretto) | Wil O'Mullane | 30 mins | Look at JIRA test director and how SE are using it. Also how might we use this to make our test specs, test reports and traceability. There is also some terminology difference to be covered (test spec = test run etc.). |
| Final Close out of Compression WG with Rob ert Gruendl | @womullan | 20 mins | Assuming those interested will have looked at the report this would be highlights and potential work to be done in the future. |
| The DMS Product Tree | Unknown User (gcomoretto) | 20 mins | Based on the different documents available (and in preparation), I would like to share some thoughts, maybe a proposal, for a categorization of the different products that can be identified (at the moment) inside DMS. |

Attached Documents

| File | Modified |
|---|---|
| PDF File KTL Metadata DMLT 2018-05.pdf | May 21, 2018 by Kian-Tat Lim |
| PDF File KTL Processes DMLT 2018-05.pdf | May 22, 2018 by Kian-Tat Lim |
| PDF File Kantor Networks and Base.pdf | May 22, 2018 by Jeff Kantor |
| PDF File LossyCompress_WG_20180522.pdf | May 22, 2018 by Robert Gruendl |
| PDF File 2018-05-22 — DMLT QA WG.pdf | May 22, 2018 by John Swinbank |
| Microsoft Powerpoint Presentation OpsRehearsalsv2.pptx | May 23, 2018 by Margaret Gelman |
| JPEG File DMLT-UW-2018.jpg | May 23, 2018 by Wil O'Mullane |
| PDF File SST Plans F18 .pdf | May 24, 2018 by Leanne Guy |
| PDF File 2018-05-24 — AP — F18.pdf | May 24, 2018 by John Swinbank |
| PDF File KTL Arch Epics DMLT 2018-05.pdf | May 24, 2018 by Kian-Tat Lim |
| PDF File DMLT-RFC-441_201805-v3.pdf | May 24, 2018 by Gregory Dubois-Felsmann |
| Microsoft Powerpoint Presentation SUIT_F18plan201805.pptx | May 24, 2018 by xiuqin |
| PDF File DAX_DMLT_05_2018.pdf | May 24, 2018 by Vaikunth Thukral |

| Microsoft Powerpoint Presentation DMLT_F2FMay2018_LDF.pptx | May 24, 2018 by Margaret Gelman |
|---|---------------------------------|
| Microsoft Powerpoint Presentation F18 Supertask Update.pptx | May 28, 2018 by Fritz Mueller |
| Microsoft Powerpoint Presentation GCM_Product-Tree.pptx | May 29, 2018 by gcomoretto |
| Microsoft Powerpoint Presentation GCM_Releases.pptx | May 29, 2018 by gcomoretto |
| Microsoft Powerpoint Presentation GCM_TestApproach.pptx | May 29, 2018 by gcomoretto |
| | |

Download All