

# 2021-08-17 Science Metric Development Agenda and Meeting notes

## Date

17 Aug 2021

## Location

Browser	Room System	Phone Dial-in
<a href="https://bluejeans.com/690803707">https://bluejeans.com/690803707</a> or <a href="https://ls.st/rsv">https://ls.st/rsv</a>	1. Dial: 199.48.152.152 or <a href="https://bjn.vc">bjn.vc</a> 2. Enter Meeting ID: 690803707 -or- use the pairing code	Dial-in numbers: <ul style="list-style-type: none"><li>+1 408 740 7256</li><li>+1 888 240 2560 (US Toll Free)</li><li>+1 408 317 9253 (Alternate Number)</li></ul> Meeting ID: 690803707

## Attendees

- Colin Slater
- Keith Bechtol
- Erik Dennihy
- Jeffrey Carlin
- Simon Krughoff
- Robert Lupton

## Regrets

- Leanne Guy (medical leave)



## Metric Tracking Dashboard

URL:

- DC2 and HSC RC2 monthly reprocessing (Gen3 and Gen2/3, faro): <https://chronograf-demo.lsst.codes/sources/2/dashboards/75>
- Nightly CI (Gen3, faro): <https://chronograf-demo.lsst.codes/sources/2/dashboards/70>

## Discussion items

Item	Who	Pre-meeting notes	Notes and Action Items
News, announcements and task review	Leanne Guy	<ul style="list-style-type: none"><li>The Director's Review in preparations of the Joint Status Review is planned for 7-9 Sep 2021. There will be a strong focus on SITCom. We should present performance metric verification status.<ul style="list-style-type: none"><li>chronograf dashboard</li><li>pipeline Characterization Metric Report</li><li>example per-detector metric (e.g., 5 sigma depth)</li><li>documentation</li></ul></li><li>International in-kind contributor Nacho Sevilla is interested in starting to get prepared for commissioning SV related work. What is our approach to starting to incorporate such in-kind contributions?<ul style="list-style-type: none"><li>What do we need? What do we want?</li><li>Concept to do direct imaging survey with AuxTel</li><li>Getting faro to a stable place</li><li>What is the usage in 6 month time</li><li>Capability to defining work packages. Specific dataset and specific tests to run and come back with an analysis.<ul style="list-style-type: none"><li>Example: is stellar locus stable, e.g., in HSC RC2? Can you make sense of the results in science context.</li></ul></li></ul></li></ul>	<ul style="list-style-type: none"><li>Plan is to use labels for sandbox activities</li><li>Postpone discussion of failure modes</li><li>Cronograf dashboard now set up to show metrics computed with faro for both DC2 and HSC RC2. Can select between native Gen3 processing and Gen2Gen3 conversion. This is good milestone.</li><li>Nightly dashboard. There is one successful run on 11 August? Thought we disabled this job?<ul style="list-style-type: none"><li>validate_drp_gen3. It was breaking. Returning all Nans. Replace with RC2 subset? Is there value in having another small dataset regularly processed, e.g., ci_imsim.</li><li>We used to do CFHT, DECam, and HSC running nightly.</li><li>According to jenkins, still running on validate_drp_gen3 nightly. jenkins is happy, but NaN values are not being displayed in cronograf.</li></ul></li><li>Just got faro running as part of ci_imsim</li><li>About to deploy checking biases on the mountain. Robert will put ticket number in the minutes. We will have notebooks associated with CP verify package. Many calibration products. Deliver as notebook. Idea that this will be run every afternoon/night to check that that calibrations products were generated.</li></ul>

Bugs, issues, topics, etc of the week	All	<ul style="list-style-type: none"> <li>• Congratulations to <a href="#">Erik Dennihy</a> who has completed his first faro ticket</li> <li>• How do we enable faster access to large numbers of metrics in data repos (e.g., tens of thousands of per-detector metric values could be generated in a single night)?</li> <li>• Some errors were introduced to faro in <div>  <b>DM-31381</b> - Jira project doesn't exist or you don't have permission to view it. </div> </li> </ul> <p>• See Slack threads <a href="#">here</a> and <a href="#">here</a>.</p> <ul style="list-style-type: none"> <li>◦ Homogenization of run methods is high-priority item <div>  <b>DM-31061</b> - Jira project doesn't exist or you don't have permission to view it. </div> </li> <li>• What guidance do we provide for developers regarding executing faro on real data (and ensuring success) before merging changes? We need to make sure we aren't merging breaking changes – these breakages are often not evident from scons/Jenkins. (see conversation in <div>  <b>DM-31382</b> - Jira project doesn't exist or you don't have permission to view it. </div> </li> </ul> <p>)</p> <ul style="list-style-type: none"> <li>◦ It seems we need an efficient way to actually run a pipeline against actual data repository to test the pipeline yaml files and runQuantum methods across multiple analysis contexts</li> </ul>	<ul style="list-style-type: none"> <li>• We will need a technical solution to access large numbers of metric values. Summary task can write an aggregate statistic (e.g., mean), but could also create astropy table for example. Another solution would be putting results in relational database (but this functionality is not planned to be soon, no concrete plan to implement, not going to happen in next 3 months). <ul style="list-style-type: none"> <li>◦ Keith to repeat test with more recent weekly</li> <li>◦ We might need something ad hoc. Might need to compare with metadata.</li> <li>◦ Colin has this as high priority</li> </ul> </li> </ul> <p>✓ Keith Bechtol to report on timing tests of loading thousands of metric measurements into memory with more recent weeklies 31 Aug 2021</p> <ul style="list-style-type: none"> <li>• Dan has PR as short term solution to consistency of run methods</li> <li>• Integration test. Executing faro on real. <ul style="list-style-type: none"> <li>◦ Need small CI dataset for faro. Execute faro against some dataset.</li> <li>◦ Ideally, we would start from an already processed dataset</li> <li>◦ CI HSC has coadd outputs. 2-3 hours.</li> <li>◦ Is there a processed ci_hsc dataset. We could set up a cron job?</li> </ul> </li> <li>• Multiple potential solutions to have faro run as part of routine pipeline testing and avoid causing breakages on master, i.e., make sure that jenkins run fails if faro is broken <ul style="list-style-type: none"> <li>◦ Could faro pipelines be merged with DRP pipelines? (obs_lsst/pipelines/imsim/DRP.yaml)</li> <li>◦ Can we just add faro to ci_hsc pipelines?</li> </ul> </li> <li>• It is possible to run specific analysis contexts of faro (e.g., per detector, per tract) separately <ul style="list-style-type: none"> <li>◦ Hsin-Fang suggests that we separate out according to the stage of analysis to make more modular and possible to scale to larger datasets</li> </ul> </li> <li>• Adding faro to ci_imsim and ci_hsc would make it so that faro is executed as part of jenkins</li> </ul>
Reprocessing status and metrics review	<a href="#">Jeffrey Carlin</a> <a href="#">Keith Bechtol</a>	<ul style="list-style-type: none"> <li>• <a href="#">Hsin-Fang Chiang</a> would like us to add faro pipelines to the <a href="#">DRP yaml pipeline definition</a>. This means that faro will be run automatically as part of monthly reprocessing.</li> </ul>	(Hsin-Fang will join the meeting at 11:30)
Development status	<a href="#">Leanne Guy</a>	<ul style="list-style-type: none"> <li>• Fall 2021 epic <div>  <b>DM-30748</b> - Jira project doesn't exist or you don't have permission to view it. </div> </li> <li>• Backlog epic: <div>  <b>DM-29525</b> - Jira project doesn't exist or you don't have permission to view it. </div> </li> </ul>	
AOB	<a href="#">Leanne Guy</a>	Next meeting 24 Aug 2021	

## List of tasks (Confluence)

Description	Due date	Assignee	Task appears on
-------------	----------	----------	-----------------

<input type="checkbox"/>	<a href="#">Leanne Guy</a> to talk to Science Pipelines (Yusra) about when do this transfer <b>19 Oct 2021</b>	19 Oct 2021	<a href="#">Leanne Guy</a>	<a href="#">2021-09-28 Science Metric Development Agenda and Meeting notes</a>
<input type="checkbox"/>	<a href="#">Leanne Guy</a> arrange to disucss at a future meeting if there are metrics from PDR3 & this paper that we might want to include in faro. <b>26 Oct 2021</b>	26 Oct 2021	<a href="#">Leanne Guy</a>	<a href="#">2021-08-31 Science Metric Development Agenda and Meeting notes</a>
<input type="checkbox"/>	Colin to ask about capturing ideas for improvement to the stellar locus algorithm <b>30 Nov 2021</b>	30 Nov 2021		<a href="#">2021-11-09 Science Metric Development Agenda and Meeting notes</a>
<input type="checkbox"/>	<a href="#">Colin Slater</a> to make a preliminary draft agenda for a workshop to clarify visualization use cases for science verification and validation		<a href="#">Colin Slater</a>	<a href="#">2022-04-19 Science Metric Development Agenda and Meeting notes</a>
<input type="checkbox"/>	<a href="#">Jeffrey Carlin</a> to review metric specification package organization and the relationship to formal requirements documents		<a href="#">Jeffrey Carlin</a>	<a href="#">2022-04-19 Science Metric Development Agenda and Meeting notes</a>
<input type="checkbox"/>	<a href="#">Keith Bechtol</a> Schedule a time to have focused discussion on verification package, potentially next status meeting		<a href="#">Keith Bechtol</a>	<a href="#">2021-09-14 Science Metric Development Agenda and Meeting notes</a>
<input type="checkbox"/>	<a href="#">Keith Bechtol</a> to make a ticket to better understand mapping of these camera and calibration products characterization efforts to verification documents and the focus of these efforts. Discuss with the SCLT		<a href="#">Keith Bechtol</a>	<a href="#">2021-09-14 Science Metric Development Agenda and Meeting notes</a>