# Vera C. Rubin Observatory
## Rubin Observatory Operations
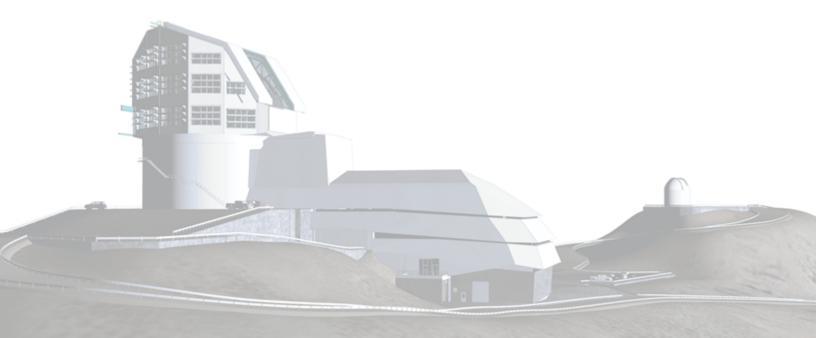
# Resource Allocation Committee

**Abhijit Saha, William O'Mullane, Melissa Graham**

**RTN-084**

**Latest Revision: 2025-02-21**

**D R A F T**

# Abstract

A draft vision for the Rubin Observatory's Resource Allocation Committee.

# Change Record

| Version | Date | Description | Owner name |
|---------|------|-------------|------------|
| 0 | YYYY-MM-DD | Unreleased. | William O'Mullane |
| 1 | 2025-02-18 | Release. Replace notes with draft vision. | Melissa Graham |

*Document source location:* `https://github.com/lsst/rtn-084`

# Contents

# Resource Allocation Committee

## 1   Introduction

This is a **draft** document to describe the future Resource Allocation Committee (RAC).

The RAC will be a standing committee that considers proposals for computational resources beyond the default allocation from users of the Rubin Science Platform (RSP) deployed at the US Data Access Center (DAC) in the Cloud (data.lsst.cloud).

The potential quantities of the resources that the RAC will allocate, and the process by which the RAC will operate, are being developed in this draft document.

**The contents of this draft document will change.**

### 1.1   Terms

**Data rights holder:** an individual with Rubin data rights as defined by RDO-013.

**User**: anyone with an RSP account at data.lsst.cloud. Data rights are required in order to hold an RSP account.

**Science community member:** anyone using the Rubin data for scientific analysis. This includes individuals without data rights, who cannot hold an account at data.lsst.cloud, but who might be analysing public or post-proprietary Rubin data.

## 2   Resources

It is a requirement defined by the Science Requirements Document that the "Data Management System will also provide at least 10% of its total capability for user-dedicated processing and user-dedicated storage" (LPM-17). This does not mean that the RAC's job is simply allocating this 10% among users. For example, user resources for data access and analysis that are provided via the Cloud are essentially shared among the simultaneous active users, whereas

the batch processing system is reserved for users with data processing projects that were approved by the RAC.

The following describes some potential resources that the RAC might be responsible for allocating.

## 2.1 Notebook servers

Currently, the Notebook Aspect allows users to self-select a small, medium, or large server.

- Small (1.0 CPU, 4 GB RAM)

- Medium (2.0 CPU, 8 GB RAM)

- Large (4.0 CPU, 16 GB RAM)

If, in the future, larger server options exist, the RAC might be responsible for approving user proposals to access them.

## 2.2 Batch processing

This refers to parallel processing (asynchronous) job submission. The user batch facility is focused on supporting a large variety of smaller needs for the science community (DMTN-202). For example, image reprocessing with specialized algorithms for specific scientific analyses.

User batch processing will be available by Data Release 1 (DR1). Access to user batch processing will be allocated by the RAC.

## 2.3 In-kind contributions

One of the goals of the Rubin In-Kind Program[1] is to augment the available resources for data- and compute-intensive use cases for the science community.

---

[1] https://www.lsst.org/scientists/in-kind-program

Towards this end, some in-kind contributions are Independent Data Access Centers (IDACs) which will offer their own separate computational resources[2].

In some cases, these IDACs might request that their resources be allocated to members of the Rubin science community by the RAC.

## 2.4  Bulk downloads

Users of the RSP can already download data subsets (image and table data), but limit estimates remain largely to-be-determined (TBD) and might depend sensitively on user load. The Data Previews will be used to quantify and optimize user experience with respect to data transfers.

If, in the future, bandwidth for very large (bulk) downloads of Rubin data by individual users is available, then it would be a resource allocated by the RAC. Note that data transfers to IDACs (as opposed to individual users) would be considered separately by Rubin Observatory.

# 3  Membership

**At first**, while the available resources are small and demand is undersubscribed, the RAC could predominantly consist of Rubin staff with the technical expertise, with some community participation (both for oversight, as well as community liaison) to approve feasible requests.

**In the future**, as demand and/or supply increases, the RAC should also expand to include a combination of Rubin staff and science community members with relevant technical and scientific expertise to evaluate and rank proposals. Any IDACs which request that the RAC allocate their resources should be invited to appoint a RAC member.

# 4  Proposals

Over time, the types and volume of resources to be allocated will grow and evolve. The "Call for Proposals" will capture the changes, and the proposal materials to be submitted will be

---

[2]https://www.lsst.org/scientists/in-kind-program/computing-resources

redesigned to capture the growing complexity.

**At first**, while the available resources are small and demand is undersubscribed, proposals might simply be a single paragraph to motivate and justify the request. At this stage, the submission mechanism can be correspondingly simple.

**In the future**, as demand and/or supply increases, proposals for larger or more competitive resources will require correspondingly longer scientific justification and technical feasibility. A management system for proposals and reviews (similar to, e.g., NOIRLab's Time Allocation System[3]) would be needed.

## 4.1 Timescales

**At first**, while the available resources are small, and if the demand is undersubscribed, the RAC could simply start off reviewing proposals as they come in (i.e., fast-turnaround, or on a rolling basis).

**In the future**, as demand and/or supply increases, the proposal timescales for larger or more competitive resources would move to a quarterly (or semesterly) basis to provide time for review.

Calls for proposals and their deadlines would be broadly advertised to the science community.

## 4.2 Critera

Rubin data rights are required to access and analyze proprietary Rubin data, and to hold an RSP account at data.lsst.cloud, as described in RDO-013. Thus, Rubin data rights are also be required to submit a proposal to the RAC[4].

**At first**, while the available resources are small and demand is undersubscribed, the RAC might apply a very simple criteria of basic feasibility or even engage individual users via email to make sure their request is understood.

---

[3]`https://time-allocation.noirlab.edu`

[4]A future exception might be made if, for example, an IDAC provides access to public or post-proprietary Rubin data *and* offers their computational resources to science community members regardless of Rubin data rights *and* they request that the RAC allocate these resources.

**In the future**, as demand and/or supply increases, criteria for larger or more competitive resources would be established by the RAC that includes weighted grading based on scientific impact and technical needs.

Requests from students for completion of degree-related projects should be weighted higher.

## 4.3   Dual-anonymous review

**At first**, this might not be necessary – and it might take time to implement.

**In the future**, best practices for unbiased reviews should be adopted.

# 5   Allocations

It is likely that the allocations would be by quarter (3 months) or semester (6 months). If resources are undersubscribed, or a scientific need for longer-term allocations is identified, the RAC might consider granting long-term allocations.

The technical implementation of how approved resources will be applied to users' RSP accounts remains to-be-determined.

**At first**, it might be as simple as a toggle that allows a user account access to shared batch processing resources, or to select a larger server.

**In the future**, resources might be more specifically allocated. For example, some processing is "pre-scheduled" (so as not to overburden resources). It might be that programs are assigned processing priority ("nice values") based on the RAC ranking or technical need (i.e., daily processing on the Prompt Products might require a higher priority).

# A   References

**[RDO-013]**, Blum, R., the Rubin Operations Team, 2020, Vera C. Rubin Observatory Data Policy,
URL `https://ls.st/RDO-013`,
Vera C. Rubin Observatory RDO-013

**[DMTN-202]**, Dubois-Felsmann, G., 2021, Use cases and science requirements on a user batch
facility, URL `https://dmtn-202.lsst.io/`,
Vera C. Rubin Observatory Data Management Technical Note DMTN-202

**[LPM-17]**, Ivezić, Ž., The LSST Science Collaboration, 2018, LSST Science Requirements Docu-
ment, URL `https://ls.st/LPM-17`,
Vera C. Rubin Observatory LPM-17

# B  Acronyms

| Acronym | Description |
| --- | --- |
| OPS | Operations |
| RAC | Resource Allocation Committee |
| RTN | Rubin Technical Note |