Vera C. Rubin Observatory
Rubin Observatory Project Office

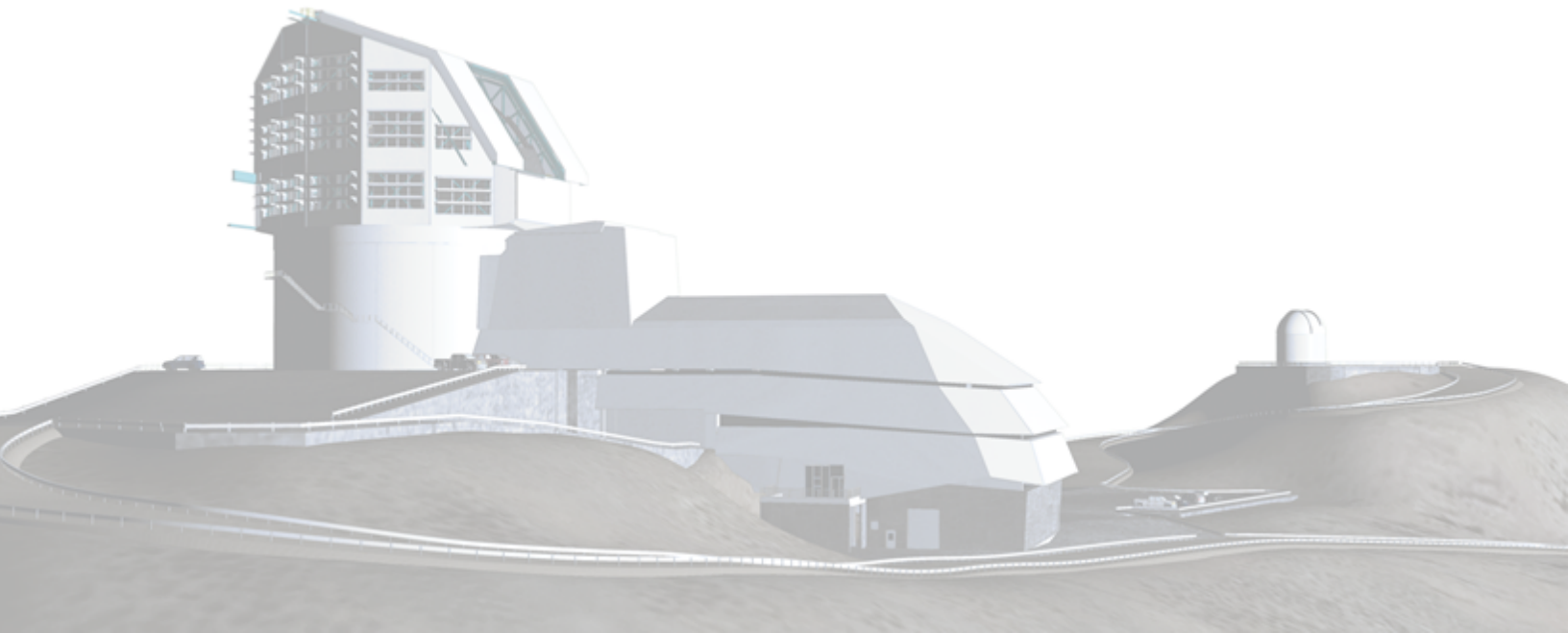
Summit Computing Cluster

Cristián Silva and Robert Lupton

ITTN-061

Latest Revision: 2023-05-16

DRAFT



Abstract

Description of the deployment of the summit computing cluster.

Change Record

Version	Date	Description	Owner name
1	YYYY-MM-DD	Unreleased.	Cristian Silva

Document source location: `https://github.com/lsst-it/ittn-061`

Contents

Summit Computing Cluster

1 Introduction

The original concept for the Commissioning Cluster was to put it at the base in La Serena ((Lim et al., LDM-148, Sec. 9)), but there are a number of problems with this:

- Data taken on the summit would have to be transferred to the base, and ingested into a separate butler
- We would have to install, and maintain, a batch processing system at the base (including managing the butler registry)
- Some of the functionality provided by the Commissioning Cluster is needed to monitor the survey data quality, and it is not clear that the summit-base link is reliable enough to guarantee that the results of this analysis will always be available at the summit. Depending on the definition of 'degraded mode' this may or may not be a problem

Accordingly, we propose that the Commissioning Cluster nodes be moved to the summit, to augment a pre-existing cluster, yagan. We note that much of the most compute-intensive work during commissioning is expected to be carried out at USDF (Dubois & O'Mullane (RTN-021)). The benefits of this modified topology for the IT group include:

- Reduced administration overhead
- Limit security constraints to a single location
- Simplifies configuration deployment
- Promotes summit independence

The big picture of the current topology is the following:

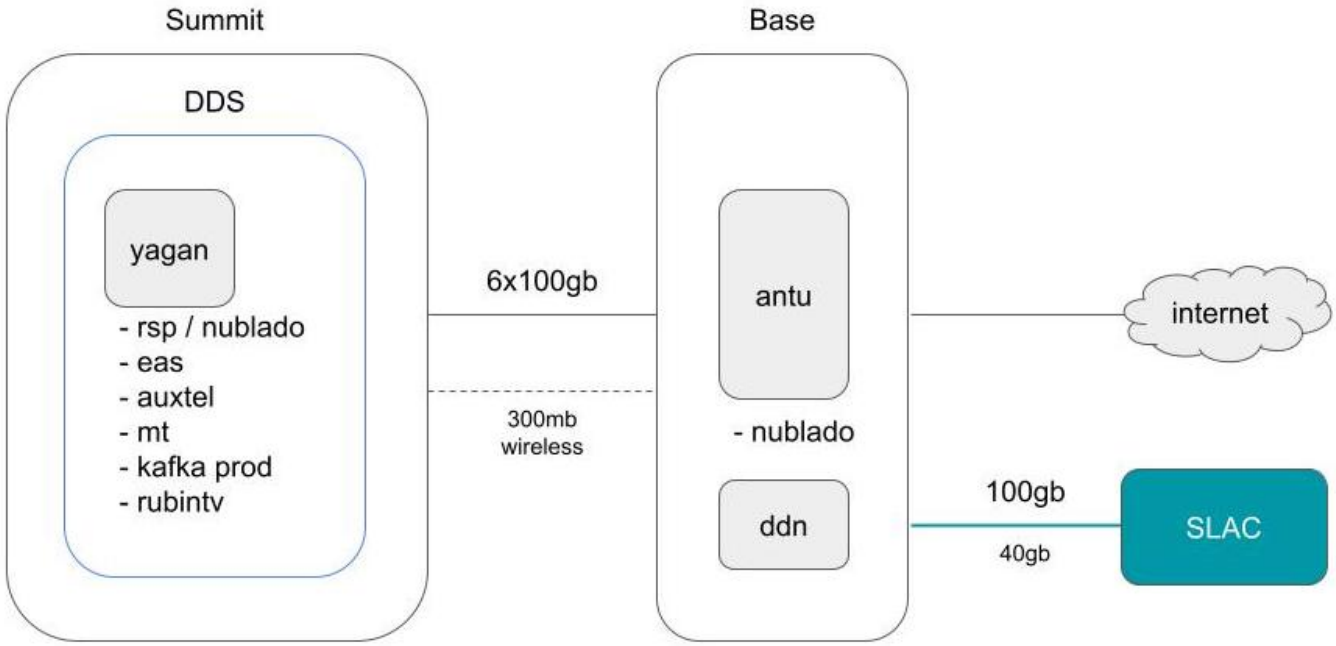


FIGURE 1: topology

2 Current State

2.1 Hardware

The commissioning cluster runs in the antu nodes, with a capacity of 832 cores and 3.2TB of Ram. Details of its hardware configuration can be found in ITTN-014 Computing Infrastructure

The servers composing the antu cluster are a mix of different Dell models, previously deployed as forwarders and DTNs by NCSA. There's also a DDN unit of about 700TB of storage, and it is currently providing NFS mounts for Nublado

FIGURE 2: antu cluster

The computing cluster at the summit is called yagan. It has a capacity of 576 cores and 2.2TB of Ram.

FIGURE 3: clusters at the summit

The servers in the yagan cluster have the same model and specs. This hardware model is shared with several other services and clusters at the summit, so they all work as a big pool of spares in case a node fails.

2.2 Software

All clusters are provisioned using Rubin Devops Stack (puppet, ipa, etc)

The antu cluster runs:

- Nublado

The yagan cluster runs:

- Rubin Science Platform
- EFD (in migration process)
- EAS CSCs
- Auxtel CSCs
- Kafka Producers
- MT CSCs
- RubinTV

Regardless of the several components running in yagan, its cpu usage stays below 10% and the RAM usage is never above 5%

3 Planned State

Each rack at the summit has a maximum of 48RU, 8RU are used by network equipment such as switches and fiber headers. Therefore each rack could potentially host 40 servers of 1RU each.

The kubernetes clusters at the summit are all of 1RU with 64 cores; hence each rack could deliver up to 2500 cores. However, given the current commissioning and summit cluster size, the maximum size of the new computing cluster would be 1400 cores.

4 Storage

Storage plays a crucial role in commissioning a cluster. The storage system should be fast and reliable to ensure that data can be accessed quickly and without interruption.

Rubin DevOps team prefer to have a single storage endpoint at the summit, which can simplify management and maintenance of the storage system. This can also help ensure consistency across the cluster and make it easier to manage access control and permissions for different users and applications. However, it's important to ensure that the chosen storage solution can meet the performance and capacity requirements of the cluster as well as any security and compliance requirements.

The storage endpoint of the summit will be the LFA, hence will host all the LSSTCam data, the data products, and all the systems needing disk space , so all of these components should be factored in when determining the total storage capacity required for the cluster.

The following is an estimation of storage sizing:

LSSTCam:

15MB/CCD * 200 CCD/exposure * 2 exposure/visit * 1000 visit/day (including calibs) * 30 days = 180TB

Data products:

LSSTCam Size * 15 = 2700TB (considering that images will be uncompressed)

EFD:

As per DM-28554: 25TB

Total of usable disk space: 3000TB (3PB)

5 Software Needed to Support Commissioning Cluster Operations

- Access to the OODS-butler
- A batch-processing system, ideally the one adopted by USDF (PanDA over slurm?)

6 Open Issues

Once we have moved the hardware to the summit, there will still be a number of things which still need to be addressed:

- The relationship of the Commissioning Cluster with the camera diagnostic cluster.
- The degree to which the Commissioning Cluster can rely on a network connection to the USDF.
- Commissioning Cluster's needs for disk now that it's not co-located with the CDC.

A References

[RTN-021], Dubois, R., O'Mullane, W., 2022, *Data Facilities Transition Plan*, RTN-021, URL `https://rtn-021.lsst.io/`,
Vera C. Rubin Observatory Technical Note

[LDM-148], Lim, K.T., Bosch, J., Dubois-Felsmann, G., et al., 2018, *Data Management System Design*, LDM-148, URL `https://ls.st/LDM-148`

B Acronyms

Acronym	Description
CC	Change Control

CCD	Charge-Coupled Device
DDN	Data Delivery Network
DM	Data Management
EAS	Environmental Awareness System
EFD	Engineering and Facility Database
IT	Information Technology
ITTN	IT Technote
LDM	LSST Data Management (Document Handle)
LFA	Large File Annex
MT	Main Telescope
NCSA	National Center for Supercomputing Applications
NFS	Network File System
OODS	Observatory Operations Data Service
PMO	Project Management Office
PanDA	Production ANd Distributed Analysis system
RAM	Random Access Memory
RTN	Rubin Technical Note
USDF	United States Data Facility