Alternative science operations approach for LSST

William O'Mullane

2019-04-28

1 Introduction

We are currently experimenting with Google and Amazon services for science platform and processing. These services are priced to deliver compute and storage - our current model at the LDF is also service oriented but is not proceed in the same manner making comparisons difficult. It has been difficult to get an alternative cost model together. An initial approach to a cloud costing was outlined in DMTN-072, this approach was to try to cost the hardware and compare to cloud pricing.

In this document a radical restructuring of LSST operations is explored - a technology stack underpinned by commodity services which could be provided by commercial providers or computing centers. Here we look first at how we would run something like this -we can then leave one free variable which is the cost of the underlying compute and storage services. This will both help to sanity check the LDF costing and potentially allow us to have a ball park for assessing commodity provider offers.

2 Plugable service oriented architecture

In the Kavli workshop in Vegas (Feb 2019) we took a long term view to astronomy archives and data processing. We suggested a layered service model as depicted in Figure 1[1]. Our requirements are no longer unique and we have access to a wealth of open source software, commodity hardware, and managed cloud services (offered by commercial providers and federally-funded institutions) that are well positioned to meet the needs of LSST Momcheva et al. (2019); Bektesevic et al. (2019).

We took Figure 1 and made a more LSST oriented version in Figure 2. This is pretty close to how we are currently but we do not treat the compute and storage as pure services.

[1]The full document is here `https://petabytestoscience.github.io/PetaBytes-2019-04-26.pdf`

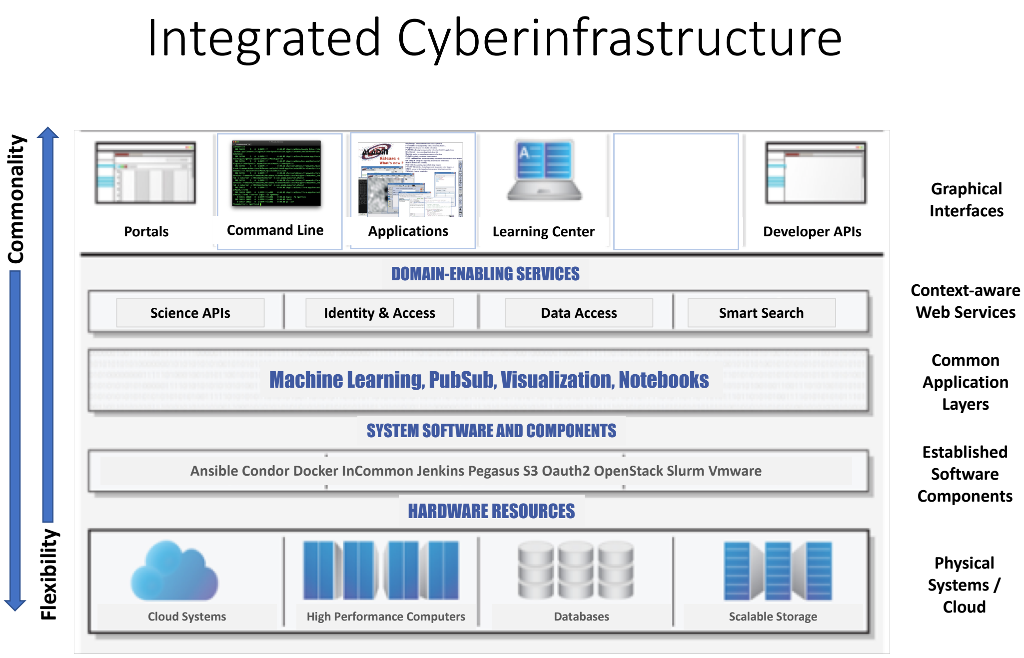

FIGURE 1: An example a cyberinfrastructure built on an Infrastructure as Code design model. Note that while this example does not have astronomy-specific tooling, our recommendations highlight the importance of developing astro-specific layers that are fully accessible to scientists in both the application and the graphical interface layers.

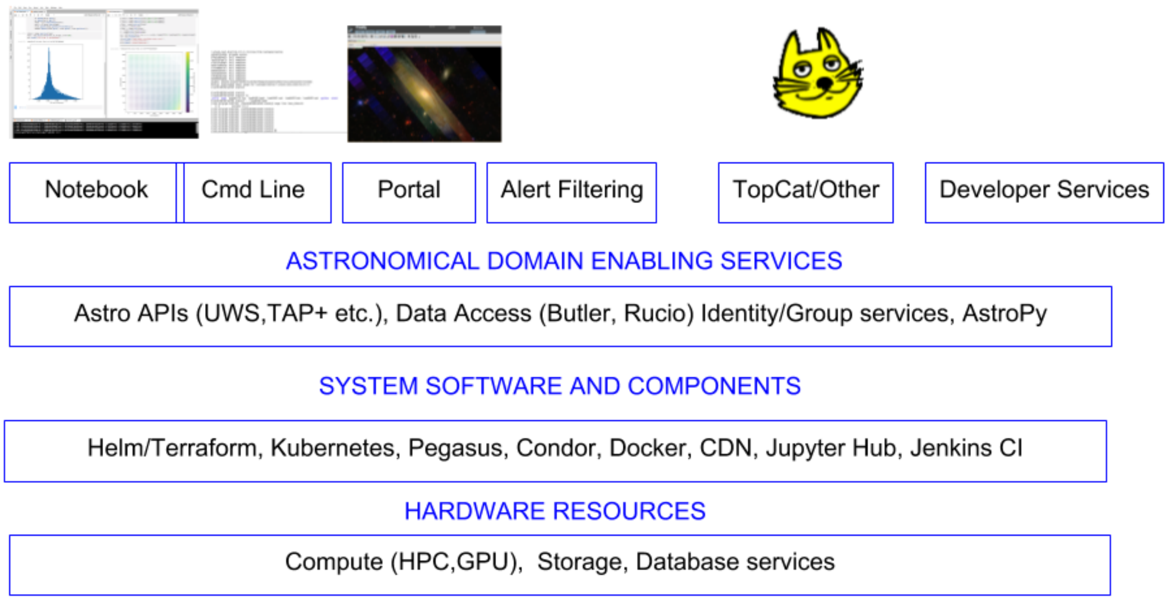

FIGURE 2: An example LSST cyberinfrastructure built analogous to the CI model shown in Figure 1.

3 Science operations

The role of science operations within LSST is to deliver LSST's science products: the science images, the alert stream, the annual data releases, the science software, and the Science Platform. In the current ops proposal not all groups required to do this are under control of Science operations.

Figure 3 gives a slightly augmented view of the science operations department.

3.1 Other implied changes to the current operations proposal

Notably missing from Figure 3 is QA. Currently QA is spread across three departments - the suggestion here is to place all QA activities under the survey performance department. Consolidation of the QA activities in one department may allow for some personnel saving.

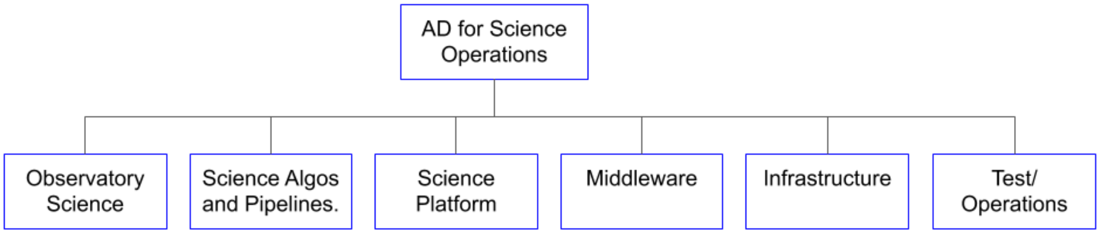

FIGURE 3: Possible configuration of Science Operations Department for operations of LSST

The data release team in science operations would require a verification scientist (this may be 0.5FTE) while the SDQA and Semantic scientists may move to QA in survey science.

All data facility work, be it with a partner or in commercial cloud should be firmly under science operations - hence there is no LDF department and no associate director for LDF.[2]

4 Conclusion

A restructuring of operations would give more transparent cost, allow for a better comparison to commodity pricing for many services and would yield considerable savings.

A References

References

Bektesevic, D., Mehta, P., Juric, M., et al., 2019, In: American Astronomical Society Meeting

[2]This is in line with AMCL recommendations

Abstracts #233, vol. 233 of American Astronomical Society Meeting Abstracts, 245.05, ADS Link

Momcheva, I., Smith, A.M., Fox, M., 2019, In: American Astronomical Society Meeting Abstracts #233, vol. 233 of American Astronomical Society Meeting Abstracts, 457.06, ADS Link

[DMTN-072], O'Mullane, W., Swinbank, J., 2018, *Cloud technical assesment*, DMTN-072, URL `https://dmtn-072.lsst.io`,
LSST Data Management Technical Note

B Acronyms and glossary items

Acronym	Description
AMCL	Aura Management Council for LSST
CI	cyberinfrastructure
DM	Data Management
DMTN	DM Technical Note
LDF	LSST Data Facility
LSST	Large Synoptic Survey Telescope
NSF	National Science Foundation
OPS	Operations
QA	Quality Assurance
SDQA	Science Data Quality Assurance
US	United States
cloud	A visible mass of condensed water vapor floating in the atmosphere, typically high above the ground or in interstellar space acting as the birthplace for stars. Also a way of computing (on other peoples computers leveraging their services and availability).
cyberinfrastructure	Sometimes denoted CI, A term first used by the US National Science Foundation (NSF), and it typically is used to refer to information technology systems that provide particularly powerful and advanced capabilities.
software	The programs and other operating information used by a computer.