



Rubin Science Platform @LAPP

- Dominique Boutigny
- [Sabine Elles](#)
- Marie Paturel

ESCAPE LAPP meeting - 25th October 2021

Rubin Science Platform (RSP)

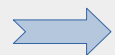
We define the *LSST Science Platform (a.k.a RSP)* as a set of web applications and services made available to the scientific community to access, visualize, subset, and perform next-to-the data analysis of the LSST dataset.

It represents the integrated set of services that will be offered to LSST users.

LSST Science Platform Vision Document

See Dominique's presentation (WP5 meeting - October 2021)

<https://drive.google.com/file/d/1wnF4v4E8fFgyIBUJfdKC0hiU7vISj5Eu/view?usp=sharing>



Purpose of this presentation :

- RSP@LAPP status
- RSP computing overview
- Ongoing work

Rubin Science Platform @LAPP

- RSP installed in collaboration with the LAPP IT Department
- 2 virtual machines hosted on LAPP Proxmox hypervisor
- lsst01 and lsst02 :
 - 16/32 GB (RAM), 8 CPU, 160 GB (disk)
 - OS : Ubuntu 20.04
 - specific VLAN

Hardware

- RSP@LAPP deployment

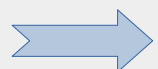
done in collaboration with Dominique, B. Gounon (CC-IN2P3)

⇒ 3 platforms are available in France (Dominique, CC and LAPP)
and F. Jammes (LPC) (Kubernetes expert)

Main difficulties :

new tools (Helm, ArgoCD) / Kubernetes ⇒ very sharp learning curve

RSP@US was still in development phase



RSP@LAPP platform is stable now & ready for testing / evaluation

Rubin Science Platform @LAPP

RSP computing overview



Kubernetes :

Open source orchestrator for deploying containerized applications

- scalable, reliable distributed system, self-healing & API oriented
- set of pods/containers, services, secrets, ...
- builds its own local secured network
- based on site specific configuration files (yaml)
- ...



rsp-lapp config. files
rsp-lapp.in2p3.fr
identifier

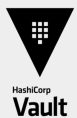


HELM : the package manager for kubernetes

- provides tools to define, install and upgrade a K8S application
- defines K8S application configuration using charts



ArgoCD : automated deployment, update and monitoring of a K8S application
(Continuous Deployment tool based on Git)



Vault : management of tokens, passwords, certificates and encryption keys

CILogon / Git : user authentication

Rubin Science Platform @LAPP

RSP computing overview



Kubernetes :

Open source orchestrator for deploying containerized applications

- scalable, reliable distributed system, self-healing & API oriented
- set of pods/containers, services, secrets, ...
- builds its own local secured network
- based on site specific configuration files (yaml)
- ...



HELM : the package manager for kubernetes

- provides tools to define, install and upgrade a K8S application
- defines K8S application configuration using charts



ArgoCD : automated deployment, update and monitoring of a K8S application
(Continuous Deployment tool based on Git)



Vault : management of tokens, passwords, certificates and encryption keys

CILogon / Git : user authentication

RSP@LAPP

forked git repository

merge done periodically

official git repository

⇒ local fork ?

US Vault

⇒ IN2P3 Vault

git user Id

≠ from LAPP user Id

⇒ IN2P3 CILogon ?

Rubin Science Platform @LAPP

The K8S components - ArgoCD display

The screenshot displays the ArgoCD web interface for the Rubin Science Platform. The interface is organized into a sidebar on the left and a main grid of application cards on the right.

Sidebar Filters:

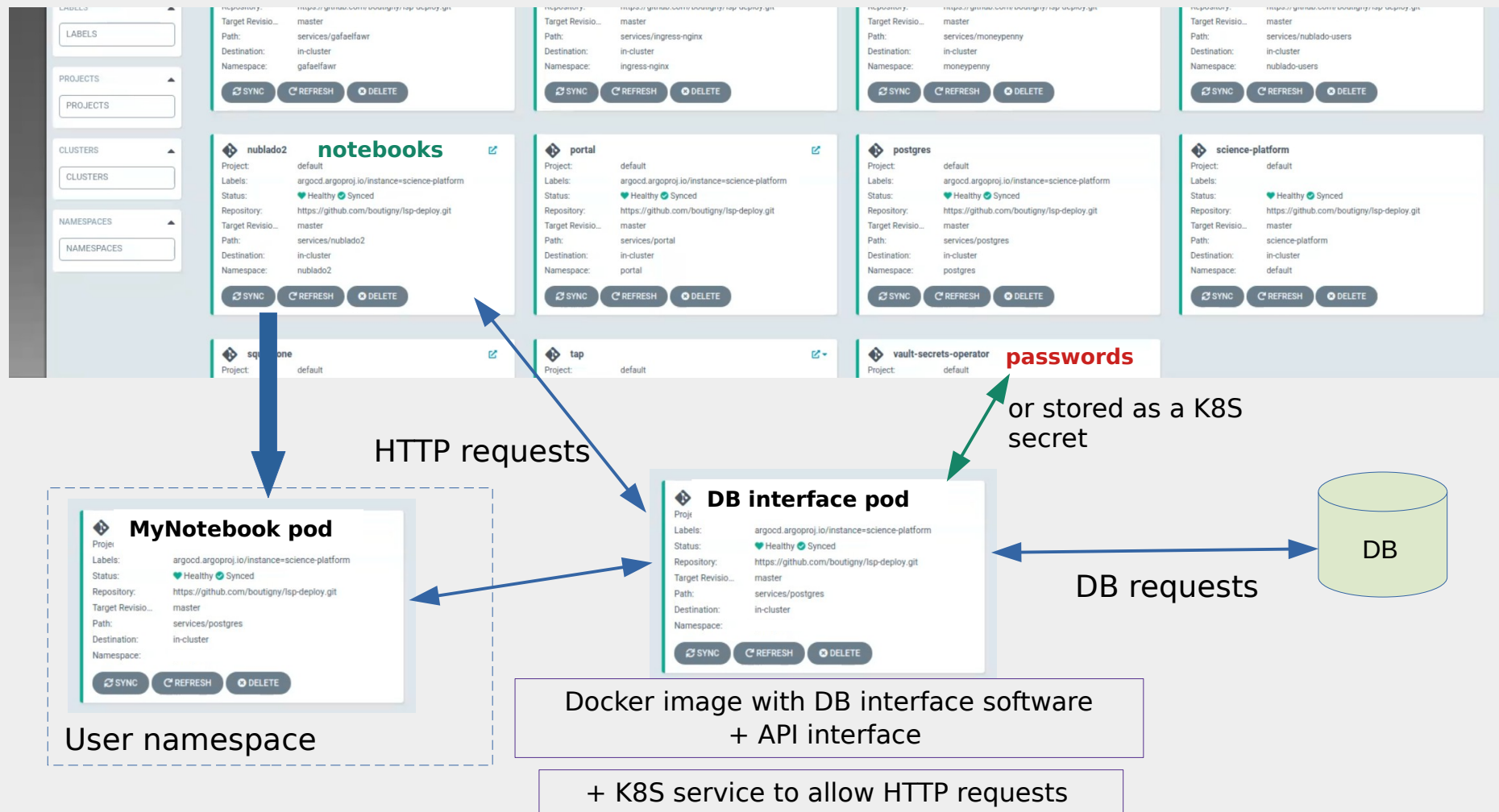
- SYNC STATUS:** Unknown (0), Synced (15), OutOfSync (0).
- HEALTH STATUS:** Unknown (0), Progressing (0), Suspended (0), Healthy (15), Degraded (0), Missing (0).
- LABELS:** A search box labeled "LABELS".
- PROJECTS:** A search box labeled "PROJECTS".
- CLUSTERS:** A search box labeled "CLUSTERS".
- NAMESPACES:** A search box labeled "NAMESPACES".

Application Cards (Grid):

Application Name	Project	Labels	Status	Repository	Target Revision	Path	Destination	Namespace
argo	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/wf	in-cluster	argo
argocd	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/argocd	in-cluster	argocd
cachemachine (docker mngt)	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/cachemachine	in-cluster	cachemachine
cert-manager (certificates)	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/cert-manager	in-cluster	cert-manager
gafaelfawr (authentication)	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/gafaelfawr	in-cluster	gafaelfawr
ingress-nginx (serveur HTTP)	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/ingress-nginx	in-cluster	ingress-nginx
moneypenny (user env.)	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/moneypenny	in-cluster	moneypenny
nublado-users (notebooks user)	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/nublado-users	in-cluster	nublado-users
nublado2 (notebooks)	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/nublado2	in-cluster	nublado2
portal	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/portal	in-cluster	portal
postgres (local RSP DB)	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	services/postgres	in-cluster	postgres
science-platform (RSP)	default	argocd.argoproj.io/instance=science-platform	Healthy Synced	https://github.com/boutigny/lsp-deploy.git	master	science-platform	in-cluster	default
squareone (RSP UI)	default							
tap	default							
vault-secrets-operator (passwords)	default							

Rubin Science Platform @LAPP

The K8S components - add a DB interface pod



The DB pod can be easily replicated (K8S replicate) in case of heavy use

RSP backend services : batch computing services

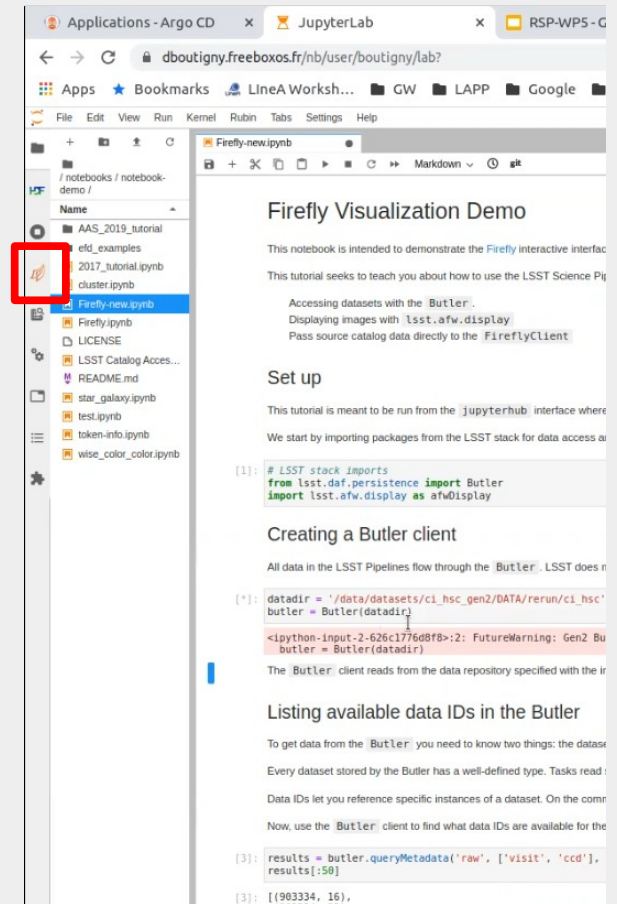
Dask interface evaluation

provides advanced parallelism for analytics

- python language
- Dask API's integrates with NumPy, pandas,...

provides a job scheduler

- from laptop to computing cluster
- provides a job parallelisation mechanism



- ✓ Dask is already used by LSST at CC-IN2P3
- ✓ the RSP@LAPP Dask component works locally
(parallelisation is done directly on RSP VM machines)

ongoing work : submit the Dask jobs to CC-IN2P3 starting
from the RSP@LAPP platform

Rubin Science Platform @LAPP

General remarks :

- it is possible to tune/modify the K8S cluster directly, without HELM, by using the K8S commands while logged on lsst01/lsst02 VM
 - ⇒ usefull to debug the cluster or to test new feature
- the biggest challenge with deploying a K8S cluster is to find the log file that contains the relevant information
 - ⇒ tool to collect application logs on Kubernetes / Kibana ??

RSP@LAPP remarks :

- the platform is in a specific VLAN
 - ⇒ it is not visible from the outside world
- the lsst01/02 capacity can be increased temporary if necessary
- the users are identified by git, not by the usual LAPP userId,groupid
 - ⇒ no access to the LAPP local data directories
- K8S training course is ongoing (F. Jammes - LPC Clermont)

LAPP IT department planning
Spring 2022

➡ if you're interested in testing the RSP@LAPP platform...

Questions ?