



REANA & Compute4PUNCH

Status & Update

Manuel Giffels, 26.02.2024

What is REANA?

reana

[Home](#) [Examples](#) [Get Started](#) [Documentation](#) [News](#) [Roadmap](#) [Contact](#) [Blog](#)

reana

Reproducible research data analysis platform

Flexible

Run many computational workflow engines.



Scalable

Support for remote compute clouds.



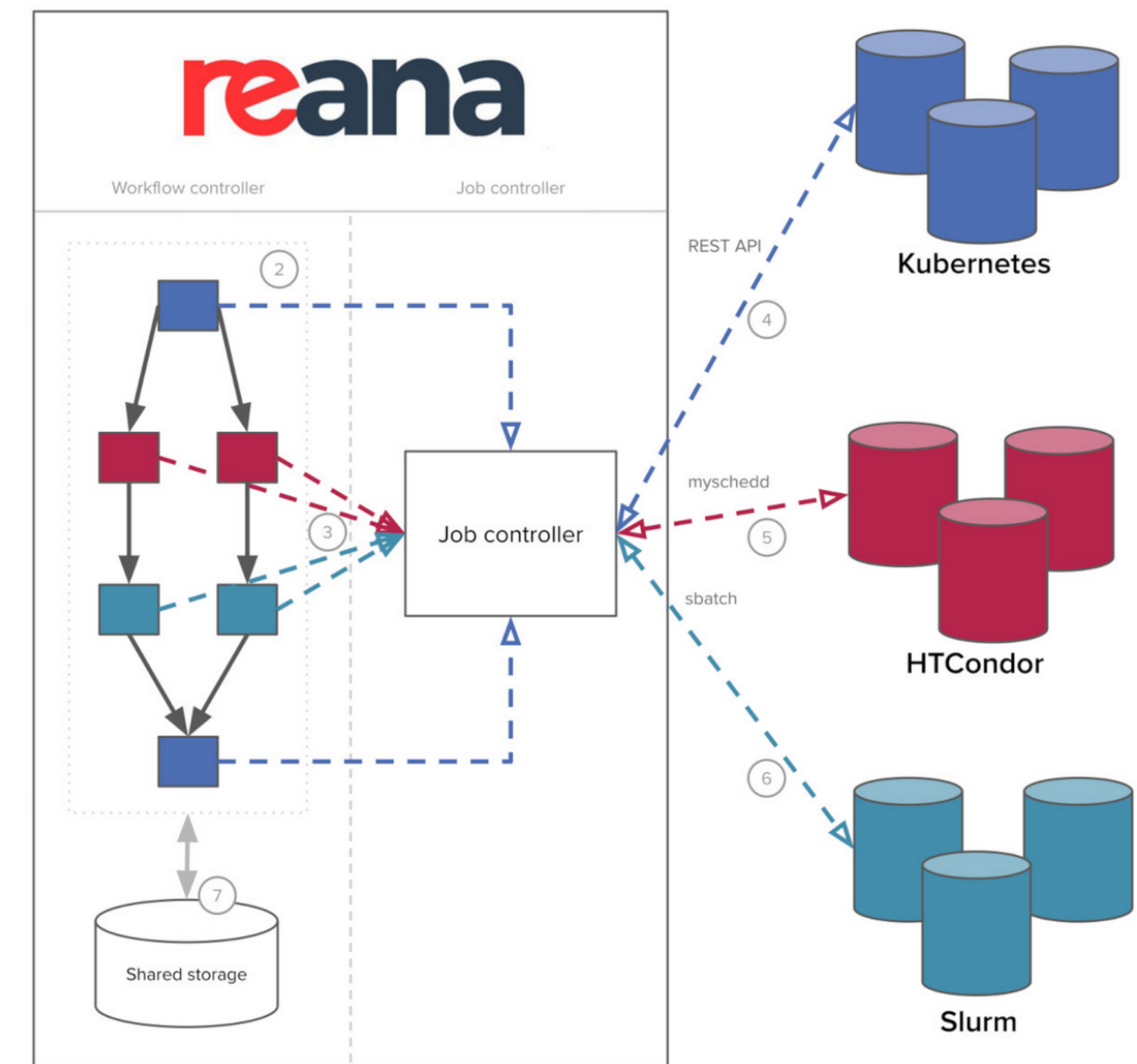
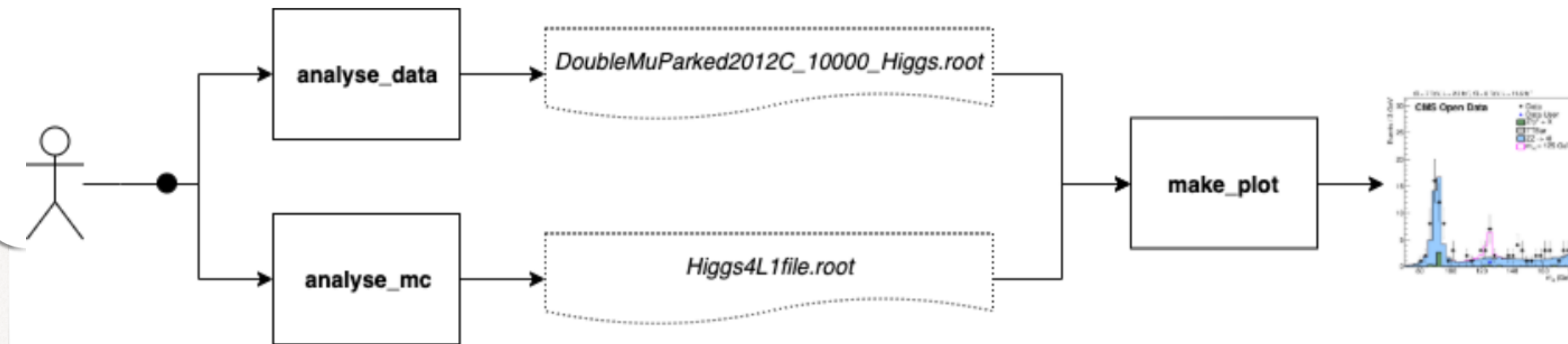
Reusable

Containerise once, reuse elsewhere. Cloud-native.



Free

Free Software. MIT licence. Made with ❤ at CERN.



What is REANA?

reana

[Home](#) [Examples](#) [Get Started](#) [Documentation](#) [News](#) [Roadmap](#) [Contact](#) [Blog](#)

reana

Reproducible research data analysis platform

Flexible

Run many computational workflow engines.



Scalable

Support for remote compute clouds.



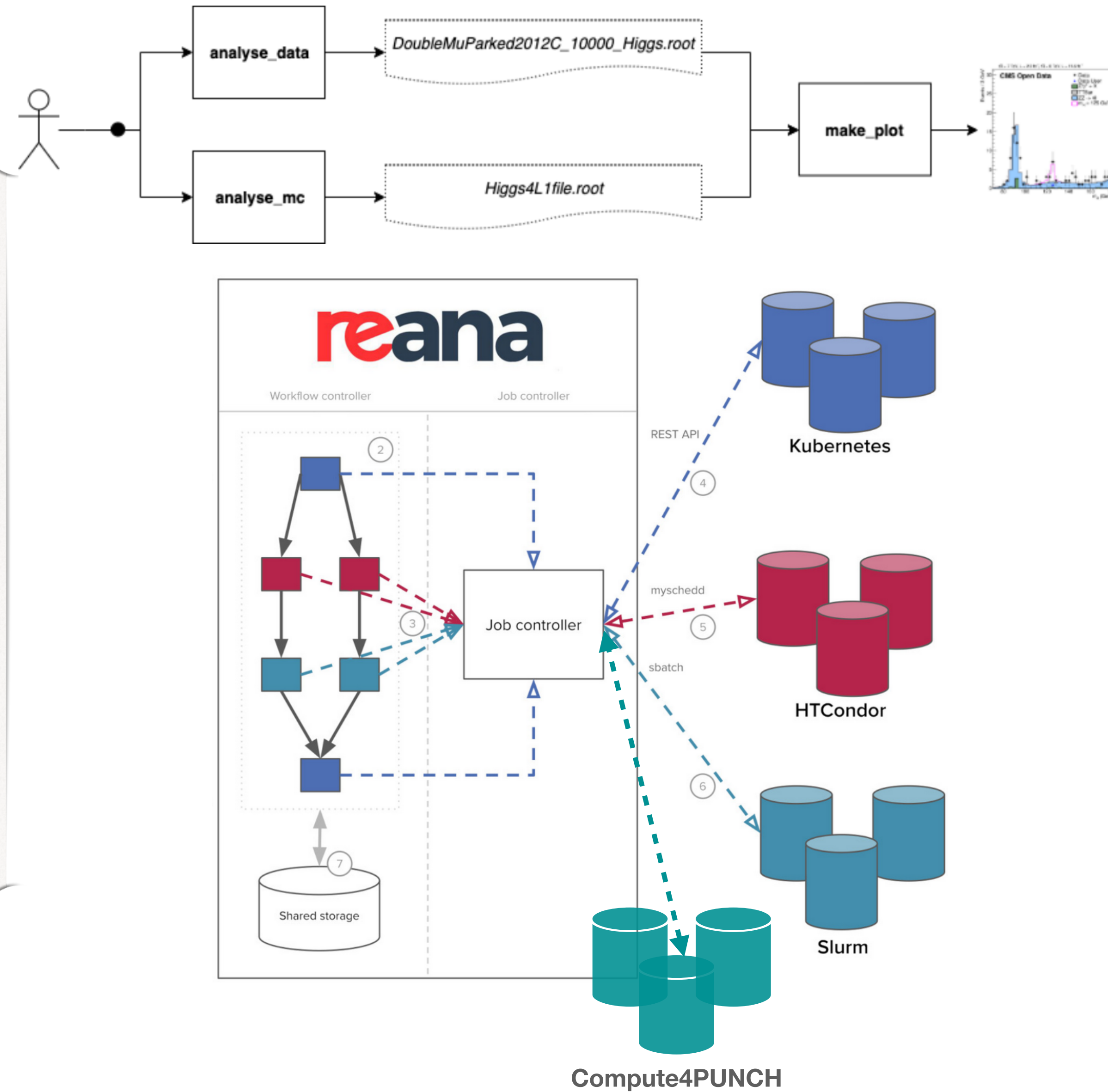
Reusable

Containerise once, reuse elsewhere. Cloud-native.



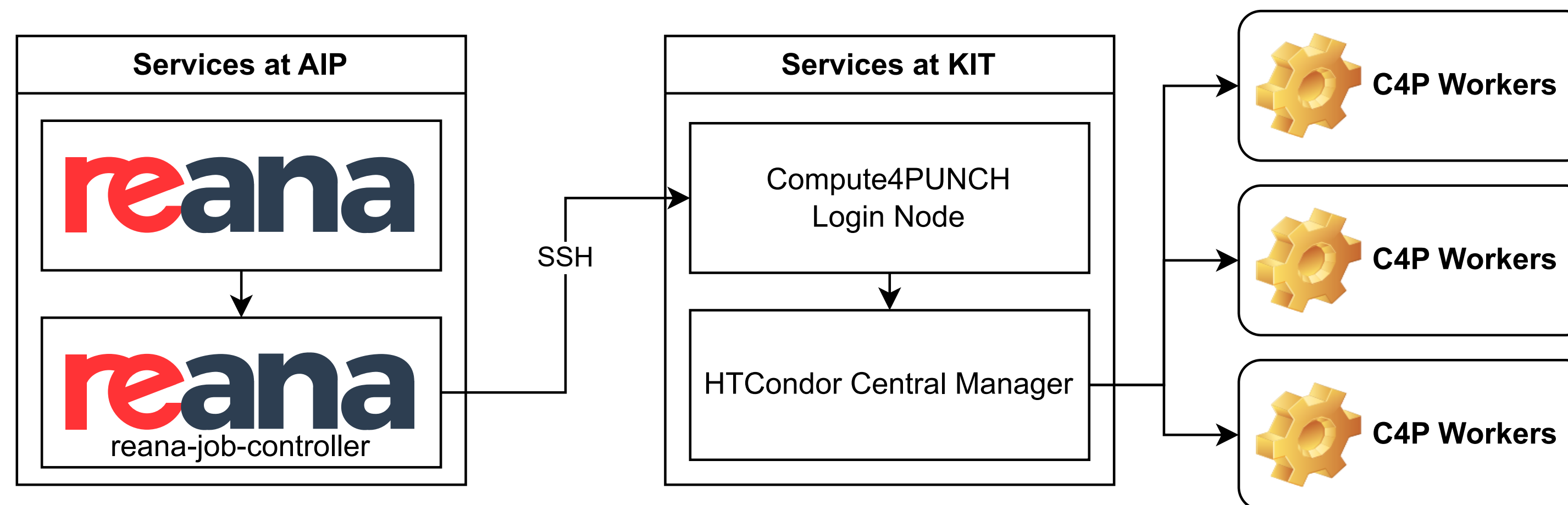
Free

Free Software. MIT licence. Made with ❤ at CERN.



How to get your jobs on Compute4PUNCH?

- Actual job submission in REANA is handled by a reana-job-controller per workflow (K8S Pod)
- There is now a prototype implementation of the Compute4PUNCH job controller available ([reana-job-controller/pull/430](https://github.com/reana/reana-job-controller/pull/430))
- Prototype utilizes a tokenized (mytoken) ssh connection to the c4p-login node at GridKa to submit, watch and handle jobs on Compute4PUNCH
- Needs a replacement in the future by using e.g. HTCondor's remote submission



How to access Compute & Storage4PUNCH

- An access token is necessary to access Compute & Storage4PUNCH
- Luckily REANA already comprises the management of arbitrary secrets
- Benoit already implement a convenient tool to create and upload mytokens to REANA
- Token can be accessed in the REANA job controller, used for ssh access to c4p-login

Fine for the initial token, but how to deal with renewals?

- C4P's HTCondor will take care of the token renewals

Secret management commands ¶

secrets-add

Add secrets from literal string or from file.

Examples:

```
$ reana-client secrets-add --env RUCIO_USERNAME=ruciouser

$ reana-client secrets-add --file userkey.pem

$ reana-client secrets-add --env VOMSPROXY_FILE=x509up_u1000
                                --file /tmp/x509up_u1000
```

secrets-delete

Delete user secrets by name.

Examples:

```
$ reana-client secrets-delete RUCIO_USERNAME
```

secrets-list

List user secrets.

Examples:

```
$ reana-client secrets-list
```

First REANA workflow on Compute4PUNCH

reana



✓ reana-cern-open-data-tutorial #40 
Finished an hour ago

finished in 10 min 17 sec
step 2/2

⚙ Engine logs >_ Job logs 📁 Workspace 📄 Specification

Step analysis finished in 8 min 5 sec compute4punch wlcg-wn:latest \$ mkdir -p results && source /cvmfs/sf...

```
-----  
| Welcome to ROOT 6.30/02                               https://root.cern |  
| (c) 1995-2023, The ROOT Team; conception: R. Brun, F. Rademakers |  
| Built for linuxx8664gcc on Dec 20 2023, 18:55:48          |  
| From heads/master@tags/v6-30-02                        |  
| With g++ (GCC) 12.1.0                                   |  
| Try '.help'/'?', '.demo', '.license', '.credits', '.quit'/'.' |  
-----
```

```
Processing code/df103_NanoAODHiggsAnalysis.C+...  
Info in <TUnixSystem::ACLiC>: creating shared library /srv/./code/df103_NanoAODHiggsAnalysis_C.so  
Warning in <RTaskArenaWrapper>: CPU Bandwith Control Active. Proceeding with 8 threads accordingly  
Info in <TCanvas::SaveAs>: ROOT file results/higgs_4mu.root has been created  
Warning in <TCanvas::Constructor>: Deleting canvas with same name: c  
Info in <TCanvas::SaveAs>: ROOT file results/higgs_4el.root has been created
```

- Deployed on a private REANA instance
- Modified helm chart deployment required
- Multiple parts of REANA deployment need to be replaced
- Successfully run the CERN open data tutorial via REANA
- Storage4PUNCH not yet involved

Discovered Drawbacks so far

- All files end up in REANA workspace by default, also temporary ones. Your REANA instance will be fill-up quickly for data intensive workflows.
→ Large files need to be deleted after migration to Storage4PUNCH?
- You cannot specify compute requirements in a workflow yet, like how many cores, memory or disk space you need.
- You cannot force your workflow to run on a dedicated compute site yet.
- You cannot delete failed workflows, you can only hide them. It is part of the concept of REANA, but it is annoying for testing.
- In the current setup the entire REANA workspace needs to be duplicated on the c4p-login node, would be nice to sync only the actually needed input files.