



**KubeCon**



**CloudNativeCon**

Europe 2018

# Multi-Cloud Federated Kubernetes at CERN

**Clenimar Filemon**  
clenimar@lsd.ufcg.edu.br

**Ricardo Rocha**  
ricardo.rocha@cern.ch





**Founded in 1954**

**Fundamental Science**

What is 96% of the universe made of?

What was the state of matter just after the Big Bang?

Why isn't there anti-matter in the universe?





LHCb

ATLAS

CERN Meyrin

SUISSE  
FRANCE

CMS

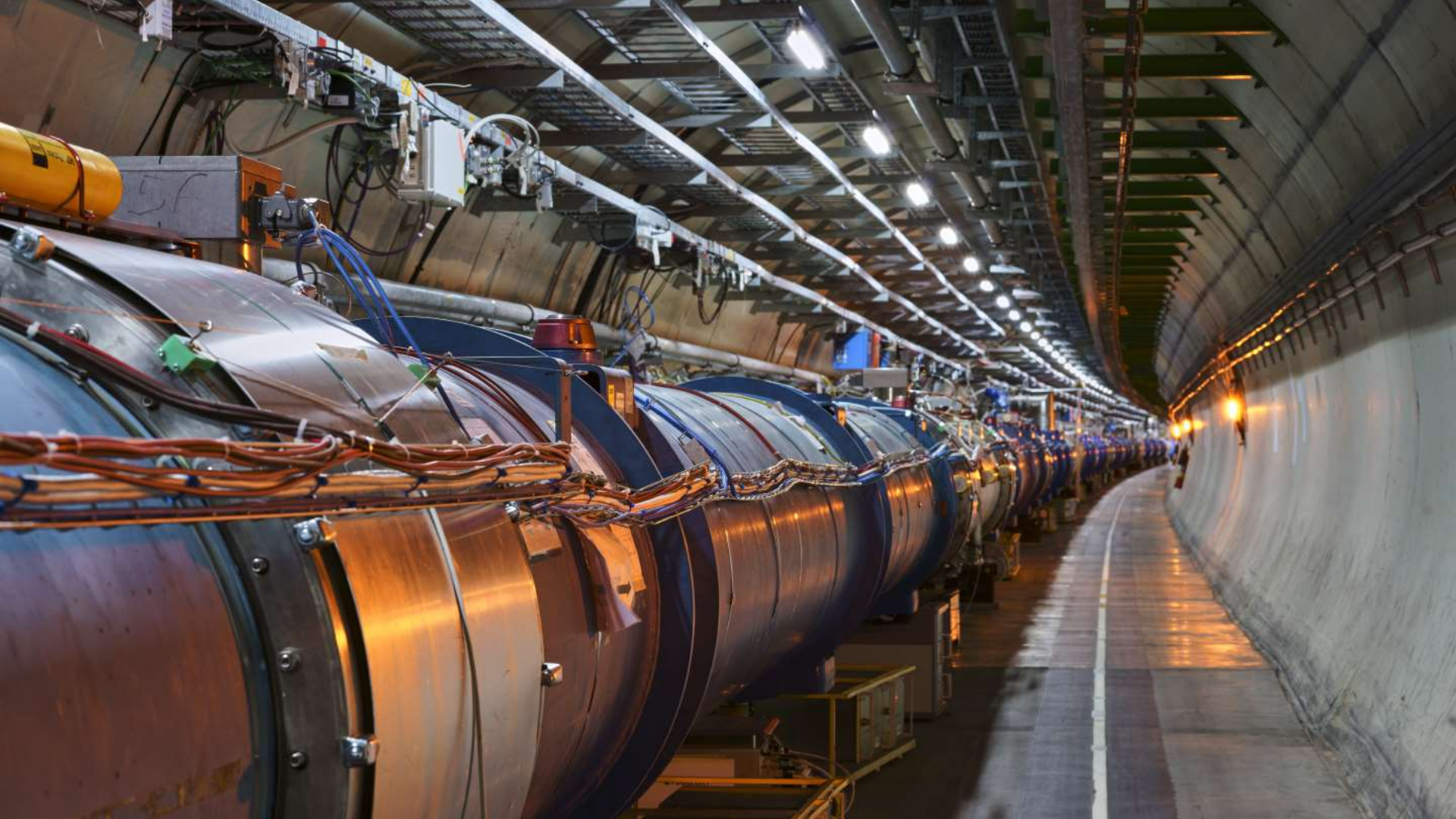
CERN Meyrin

SPS = 7 km

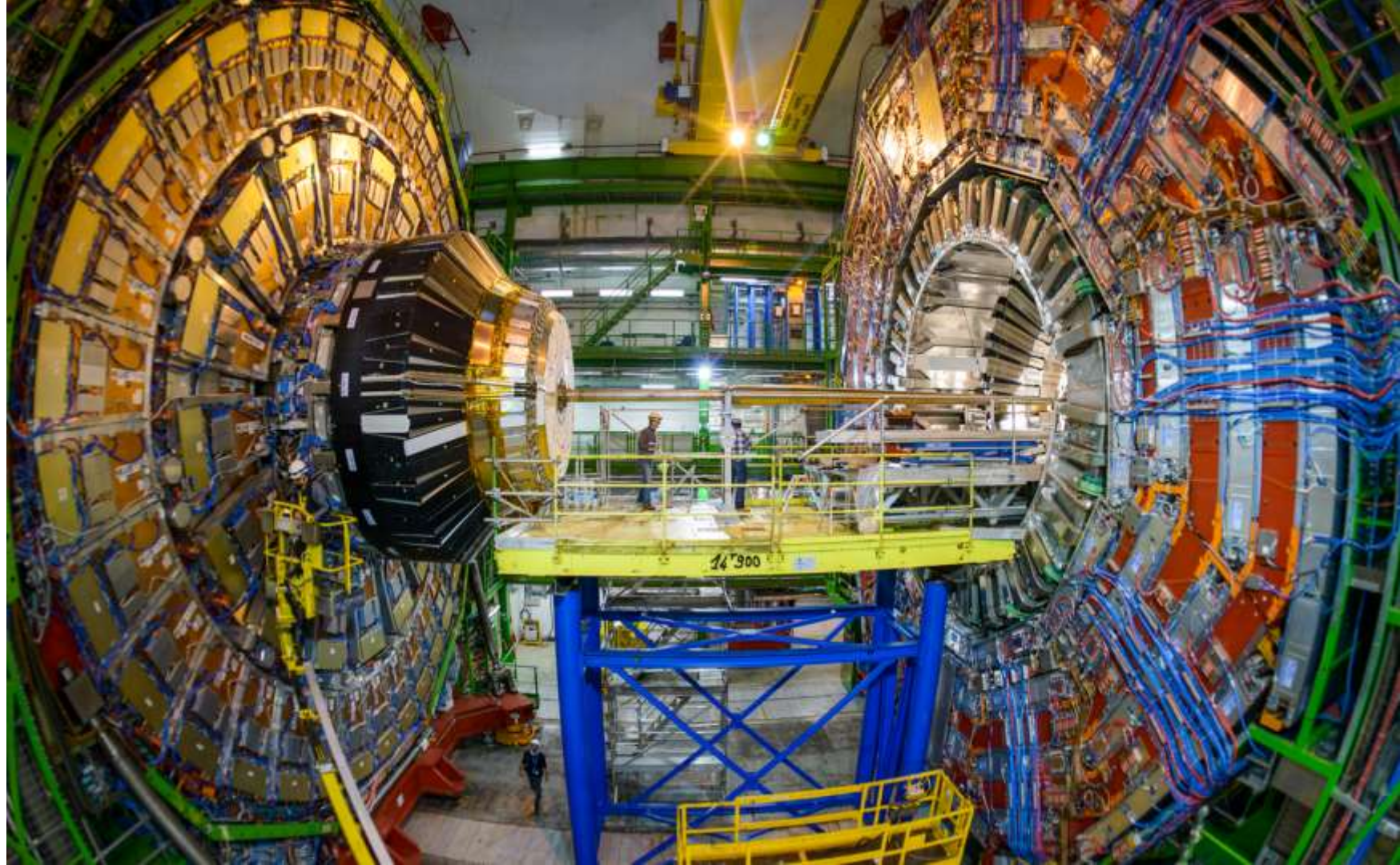
ALICE

LHC = 27 km











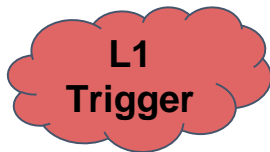


## Collisions

~40 MHz

~ 1PB/sec

Huge Data



Still Big



## Hardware Filter

~ 100 kHz



Still Big



## Software Filter

~ 1 kHz

~ 1-10 GB/s

## Raw Data







**320 000 Cores**

**4 300 Projects**

**250 Petabytes**

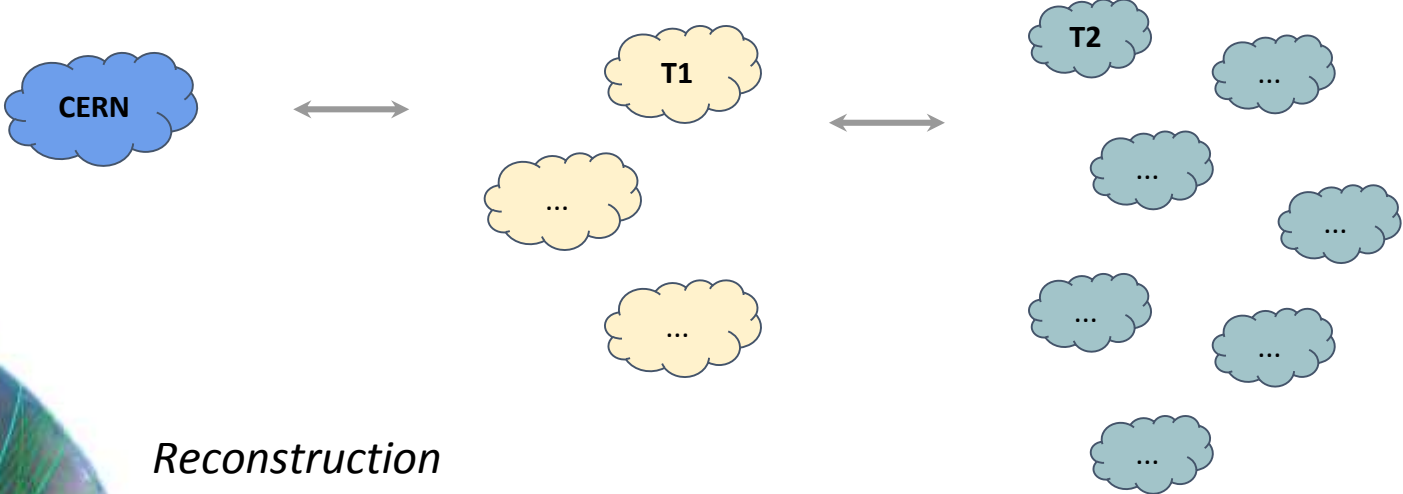
**3 300 Users**

**10 000 Hypervisors**

**210 Kubernetes Clusters**

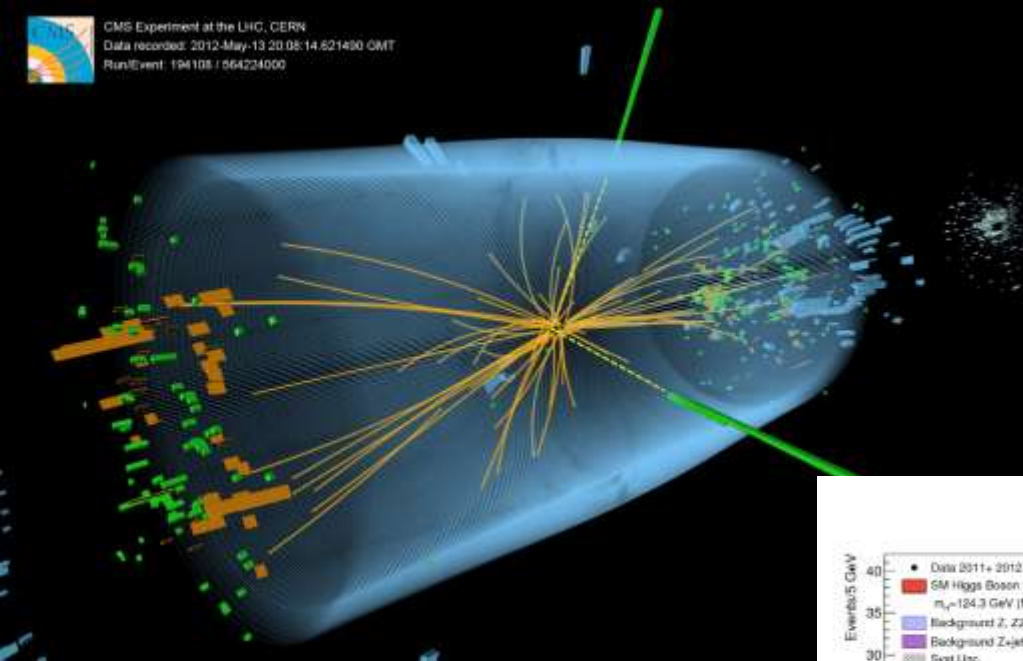


# Distributed Computing



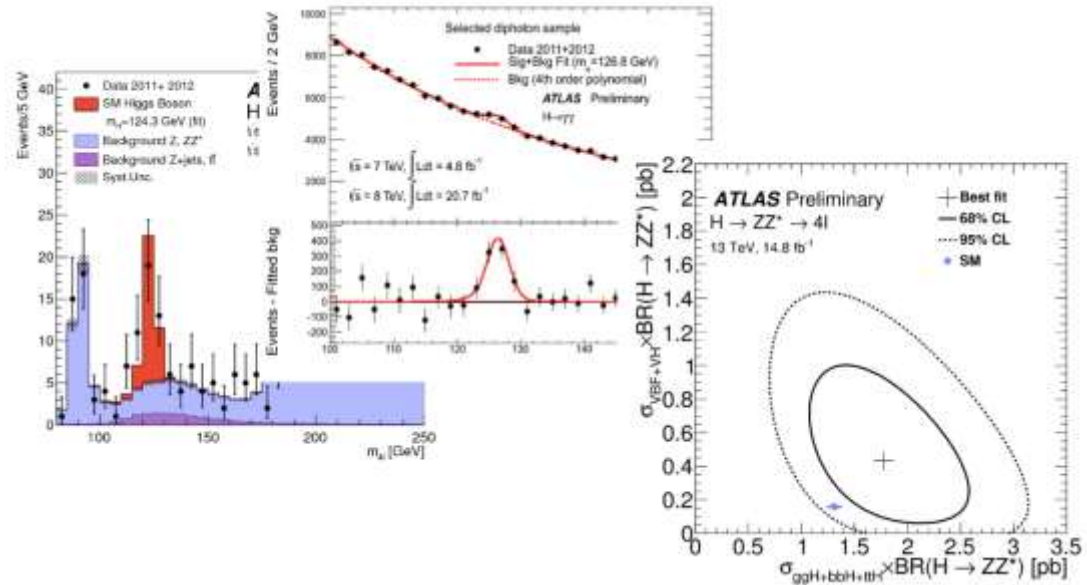
*Reconstruction*  
*Calibration*  
*Simulation*  
*Analysis*

**200+ Sites**      **~400 000 Jobs**  
**700 000 Cores**      **~30 GiB/s**



CMS Higgs Event, May 2012

ATLAS Higgs Analysis, May 2012





# Motivation for Federation

Periodic load spikes

*International Conferences, Reconstruction Campaigns*

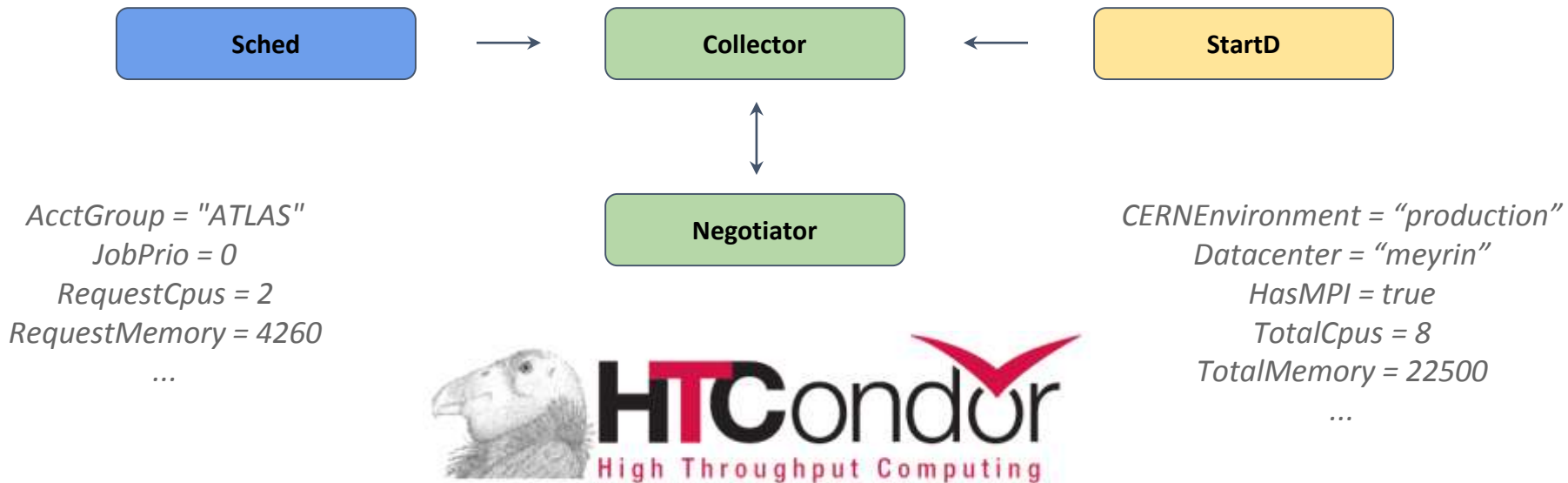
Simplification

*Monitoring, Lifecycle, Alarms*

Deployment

*Uniform API, Replication, Load Balancing*

**Use Cases:** CERN Batch System, RECAST Analysis



*Matchmaking with ClassAds*

*Fair Share*

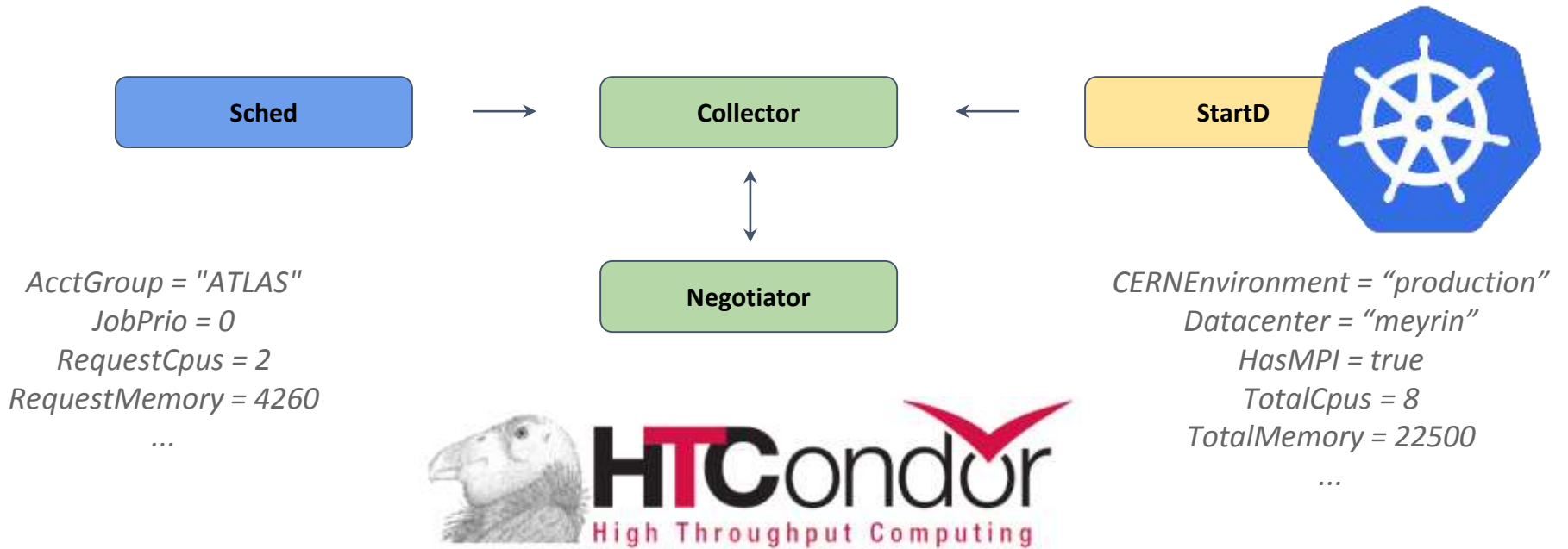
*Preemption*

*Extensive Experience in HEP*

*Running Virtualized*

*External Storage and Networking*





*Matchmaking with ClassAds*

*Fair Share*

*Preemption*

*Extensive Experience in HEP*

*Running Virtualized*

*External Storage and Networking*

**Host**



Sched

Collector



Negotiator

*kubefed init fed --host-cluster-context=condor-host ...*



StartD

...



StartD

...



StartD

...



TS-Systems

kind: DaemonSet

...

hostNetwork: true

containers:

- name: condor-startd

image: ../cloud/condor-startd

command: ["/usr/sbin/condor\_startd", "-f"]

securityContext:

privileged: true

livenessProbe:

exec:

command:

- condor\_who

Host



Sched

Collector



Negotiator

kubefed init fed --host-cluster-context=condor-host ...

kubefed join --context fed tsystems \  
--host-cluster-context condor-host --cluster-context tsystems

# REANA / RECAST

*Reusable Analysis Platform*

Workflow Engine (Yadage)

Each step a Kubernetes Job

Integrated Monitoring & Logging

Centralized Log Collection

Workflow Monitor for b907a91b-7f02-4b3e-afa3-e9298b888e2d

Status: **Active**

Workflow Visualization



Log

Messages from the request processor will appear below

```
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:acquiring lock for workflow b907a91b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:running analysis on workflow b907a91b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:acquiring workflow lock for http://reana.org/workflow-engine/
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:processing ['conda-forge', 'gromacs', 'cleaner', 'config-from-server', 'b907a91b-7f02-4b3e-afa3-e9298b888e2d']
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:running analysis on workflow b907a91b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:setting up for context {'workflow_spec': {'name': 'root', 'context': {'workflow_spec': {'default':
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:program started: workflow b907a91b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:preparing workflow b907a91b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25 15:57:13.000 [WARNING] workflowengine.processor:no input archive specification, skipping archive
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:setting up entry point workflowengine.processor.workflowengine.workflowengine
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:checking package verifiles for context: {'workflow_spec': {'name': 'root', 'context': {'workflow_spec': {'default':
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:connecting backend by the stateful backend
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:setting up backend system:stateful backend with url: 'http://localhost:3000/'
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:backend:backend base settings: {'backend': {'name': 'stateful', 'url': 'http://localhost:3000/'}}
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:single workflow:running workflow on context {'workflow_spec': {'name': 'root', 'context': {'workflow_spec': {'default':
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:single workflow:workflow streamers in path dir
2018-04-25 15:57:13.000 [INFO] workflowengine.processor:single workflow:workflow streamers in path dir
```

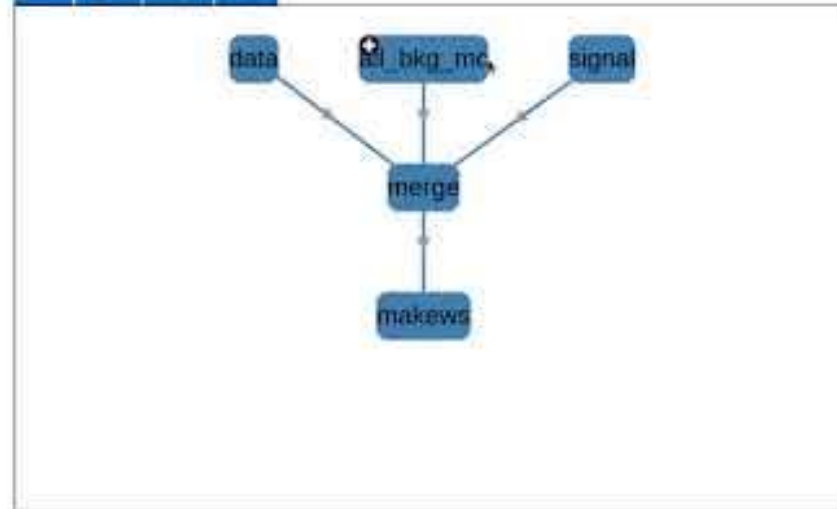


## Workflow Monitor for bf95b4cf-1baa-4320-b4c2-2a2bd5bc6cac

Status: **ACTIVE**

### Workflow Visualization

[refresh](#) [refresh](#) [collapse](#) [expand](#)



2018-05-25 23:37:07

<https://www.youtube.com/watch?v=jNyd97LiTXk>

# Credits



*CERN OpenStack Cloud and Batch teams (Spyros Trigazis and all)*

*Lukas Heinrich, REANA / RECAST*

*Kelsey Hightower*

# Thank You



Great Community, Amazing Tools