



KubeCon



CloudNativeCon

Europe 2018

Multi-Cloud Federated Kubernetes at CERN

Clenimar Filemon
clenimar@lsd.ufcg.edu.br

Ricardo Rocha
ricardo.rocha@cern.ch





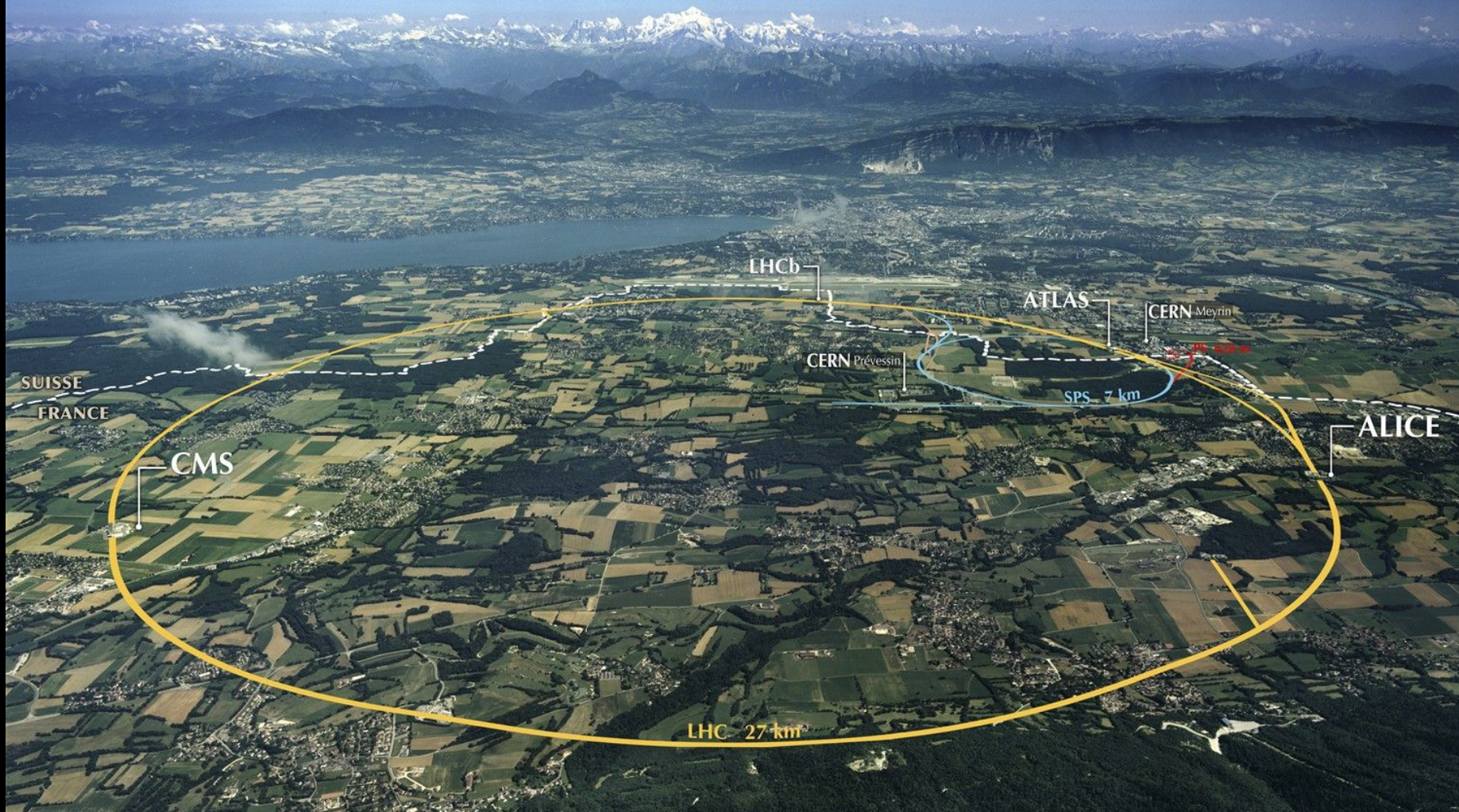
Founded in 1954

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

CERN Prévessin

SPS 7 km

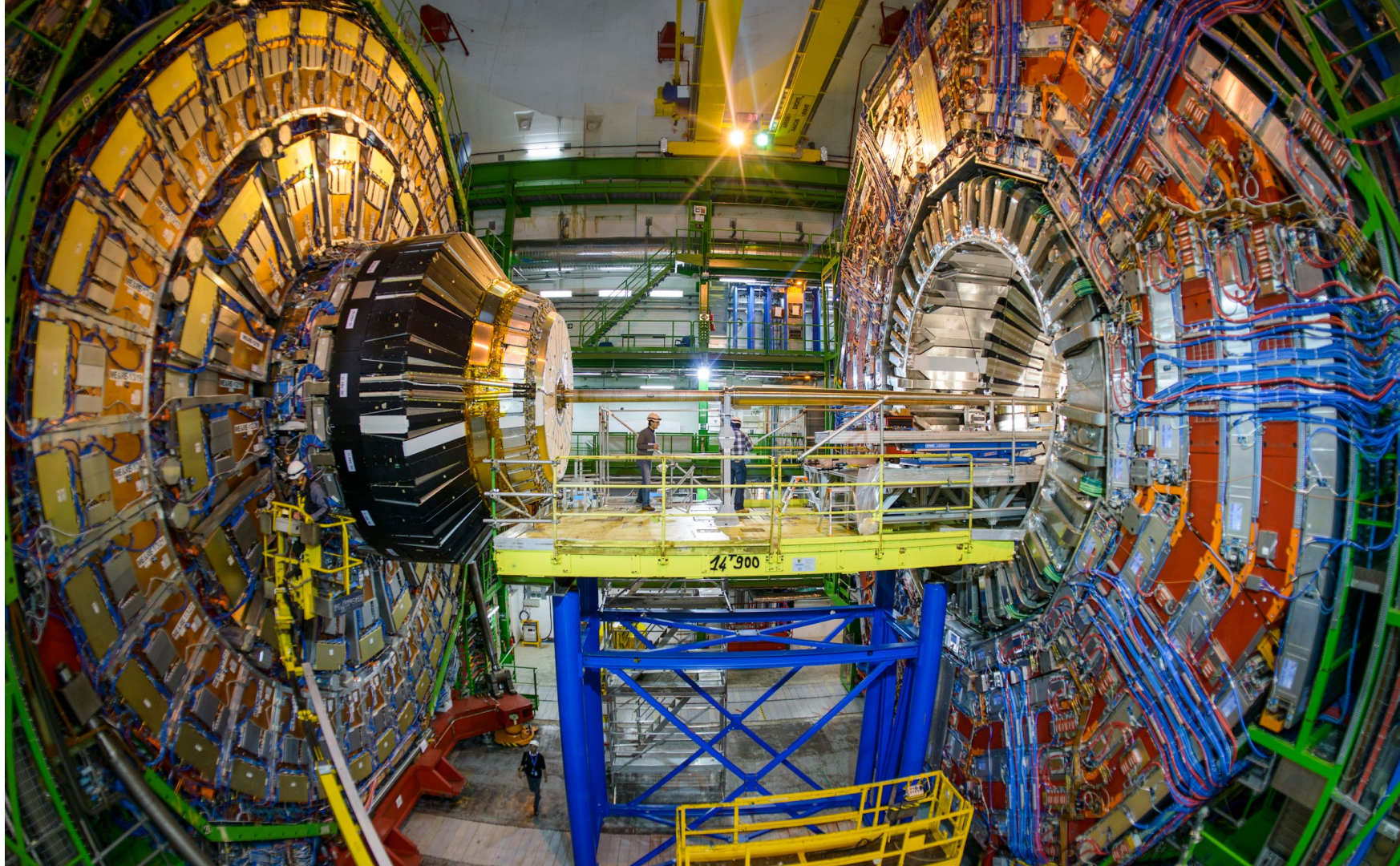
ALICE

SUISSE
FRANCE

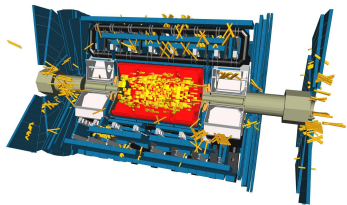
CMS

LHC 27 km







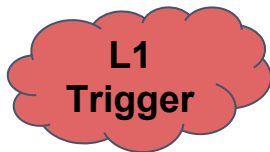


Collisions

~40 MHz

~ 1PB/sec

Huge Data



Still Big



Hardware Filter

~ 100 kHz



Still Big

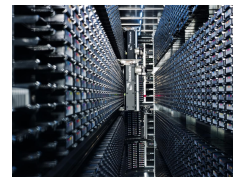


Software Filter

~ 1 kHz

~ 1-10 GB/s

Raw Data





320 000 Cores

4 300 Projects

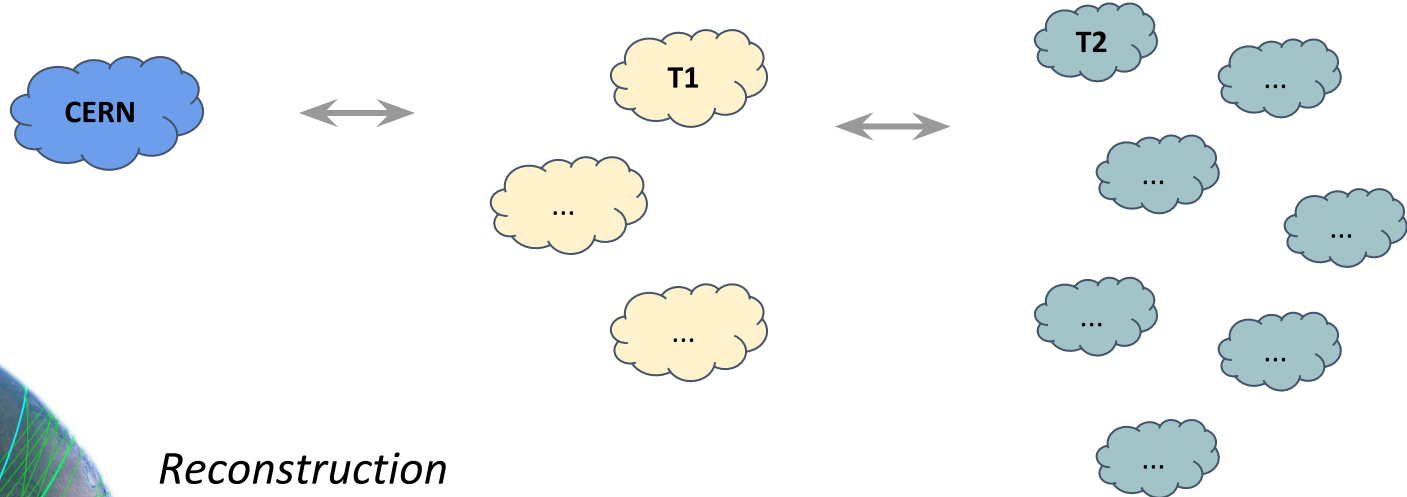
250 Petabytes

3 300 Users

10 000 Hypervisors

210 Kubernetes Clusters

Distributed Computing



Reconstruction

Calibration

Simulation

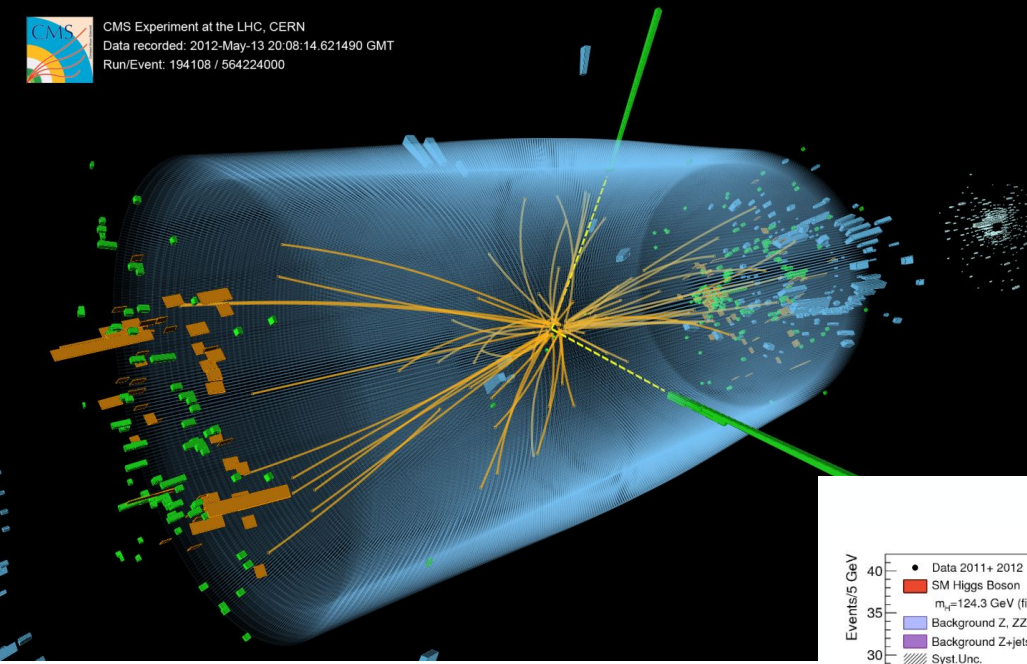
Analysis

200+ Sites

~400 000 Jobs

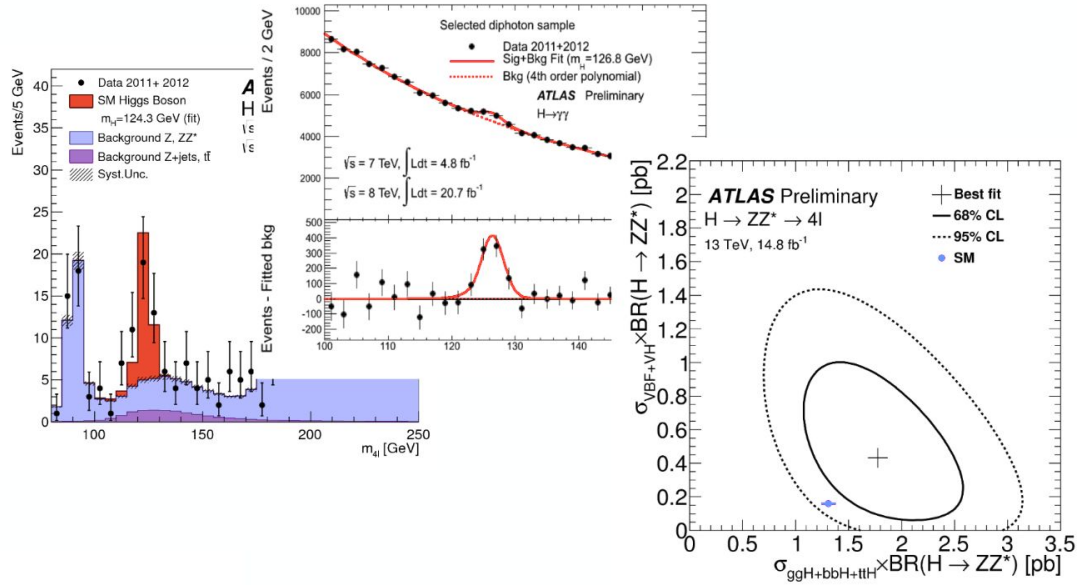
700 000 Cores

~30 GiB/s



CMS Higgs Event, May 2012

ATLAS Higgs Analysis, May 2012



Motivation for Federation

Periodic load spikes

International Conferences, Reconstruction Campaigns

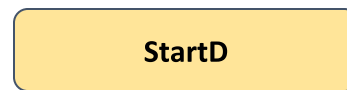
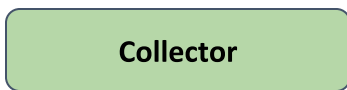
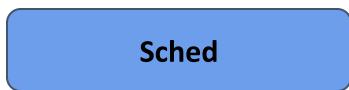
Simplification

Monitoring, Lifecycle, Alarms

Deployment

Uniform API, Replication, Load Balancing

Use Cases: CERN Batch System, RECAST Analysis



AcctGroup = "ATLAS"
JobPrio = 0
RequestCpus = 2
RequestMemory = 4260
...

CERNEnvironment = "production"
Datacenter = "meyrin"
HasMPI = true
TotalCpus = 8
TotalMemory = 22500
...



Matchmaking with ClassAds

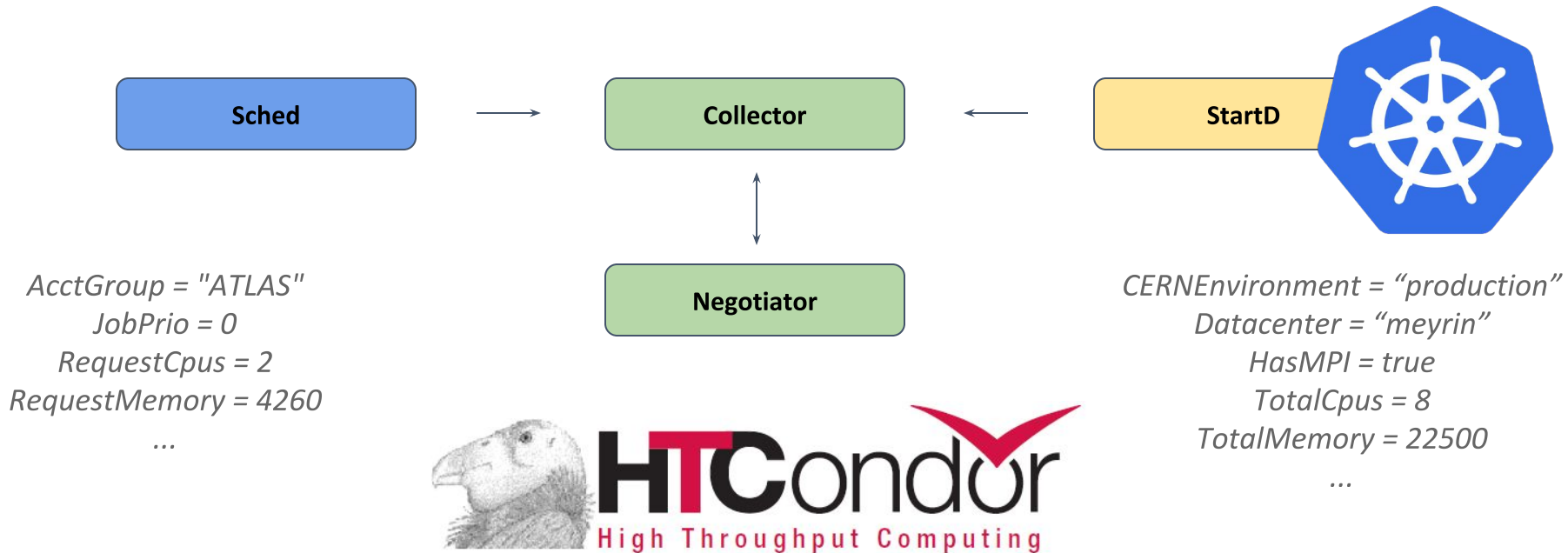
Extensive Experience in HEP

Fair Share

Running Virtualized

Preemption

External Storage and Networking



Matchmaking with ClassAds

Fair Share

Preemption

Extensive Experience in HEP

Running Virtualized

External Storage and Networking

Host



Sched

Collector

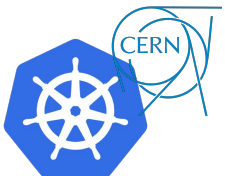


Negotiator

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

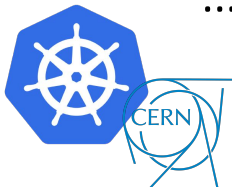
StartD

...



StartD

...



StartD

...



T Systems

Host



Sched

Collector



Negotiator

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

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

[redraw](#) [relayout](#) [collapse](#) [expand](#)



Last seen: 2018-04-25 15:57:13

Log

Messages from the request processor will appear below.

```
2018-04-25T15:53:30. INFO:wflowbackend.process:acquiring context from wflowserver for wflowid b907a91b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25T15:53:30. INFO:wflowbackend.backendtasks:running analysis on worker: wflow-nonint-b907a91b-7f02-4b3e-afa3-e9298b888e2d-76skn
2018-04-25T15:53:30. INFO:wflowbackend.backendtasks:acquiring wflow context from http://wflow-server.default.svc.cluster.local
2018-04-25T15:53:30. INFO:wflowbackend.process:processing ['setupFromURL', 'generic_onsuccess', 'cleanup', '--config-from-server', 'b907a91b-7f02-4-
2018-04-25T15:53:30. INFO:wflowbackend.backendtasks:running analysis on worker: wflow-nonint-b907a91b-7f02-4b3e-afa3-e9298b888e2d-76skn
2018-04-25T15:53:30. INFO:wflowbackend.backendtasks:setting up for context {'u'shipout_spec': {'u'suser': 'u'root', 'u'host': 'u'wflow-shiptarget.default-
2018-04-25T15:53:30. INFO:wflowbackend.backendtasks:prepared workdir workdirs/b9/07/a9/1b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25T15:53:30. INFO:wflowbackend.backendtasks:preparing workdir workdirs/b9/07/a9/1b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25T15:53:30. WARNING:wflowbackend.backendtasks:No input archive specified, skipping download
2018-04-25T15:53:30. INFO:wflowbackend.process:setting up entry point wflowyadageworker.backendtasks:run_workflow
2018-04-25T15:53:30. INFO:wflowyadageworker.backendtasks:running yadage workflow for context: {'u'shipout_spec': {'u'suser': 'u'root', 'u'host': 'u'wflow-
2018-04-25T15:53:30. INFO:wflowyadageworker.backendtasks:connecting backend: py:sharedstatekube:backend
2018-04-25T15:53:30. INFO:yadage.utils:setting up backend py:sharedstatekube:backend with opts {'optsyaml': '/yadageconfig/backendopts'}
2018-04-25T15:53:40. INFO:wflowyadageworker.backendtasks:yadage base settings {'updateinterval': 30.0, 'dataarg': 'workdirs/b9/07/a9/1b-7f02-4b3e-a-
2018-04-25T15:53:40. INFO:wflowyadageworker.simple_workflow:running workflow on context {'u'shipout_spec': {'u'suser': 'u'root', 'u'host': 'u'wflow-shipt-
2018-04-25T15:53:40. INFO:wflowyadageworker.simple_workflow:no workflow parameters in init dir
2018-04-25T15:53:40. INFO:wflowyadageworker.backendtasks:load workflow from source
```


Terminal

File Edit View Search Terminal Help

Every 1.0s: kubectl --context fed get job

rocha250: Wed Apr 25 17:59:06 2018

NAME	DESIRED	SUCCESSFUL	AGE
wflow-job-0bf2ee91-0387-4e6b-a357-f2dc309aa9da	1	0	51m
wflow-job-5e3c5079-027d-4444-8c74-c1eae51d68c8	1	1	35s
wflow-job-9a49ac80-6792-46e6-9782-109ab0bd151f	1	1	51m
wflow-job-d6b84b74-c944-4539-8b30-f4be0115f559	1	1	36s
wflow-job-d8dcd03e-3b51-437a-b960-17422bc444b5	1	1	35s

Terminal

File Edit View Search Terminal Help

Every 1.0s: kubectl --context atlas-recast-x get job

rocha250: Wed Apr 25 17:59:06 2018

NAME	DESIRED	SUCCESSFUL	AGE
wflow-job-05a48de3-fdaa-4a25-bb3d-6e65730410d9	1	1	4m
wflow-job-0767bf09-8a30-4866-8aa7-21a4b59adac9	1	1	1m

wflow-job-0ee8

File Edit View Search Terminal Help

Every 1.0s: kubectl --context atlas-recast-y get job

rocha250: Wed Apr 25 17:59:06 2018

NAME	DESIRED	SUCCESSFUL	AGE
wflow-job-05dc99b4-18a7-4a0f-a1ad-79d127d42de3	1	1	1m

wflow-job-0eca84

Terminal

File Edit View Search Terminal Help

Every 1.0s: kubectl --context atlas-recast-z get job

rocha250: Wed Apr 25 17:59:06 2018

NAME	DESIRED	SUCCESSFUL	AGE
wflow-job-560db6	1	1	2m
wflow-job-601816	1	1	2m
wflow-job-61ba2e	1	1	3m
wflow-job-0f54b731-25d4-4298-9c33-eb7ba7d5c4bb	1	1	4m
wflow-job-6faf7b	1	1	1m
wflow-job-1979e7b3-3dc7-4ffd-b443-0303c26d13eb	1	1	4m
wflow-job-6fce76	1	1	4m
wflow-job-2cdda822-3b49-49e8-ab41-494e90abc568	1	1	4m
wflow-job-796a66	1	1	2m
wflow-job-3169fd59-d69c-4b27-9f0b-ba07ecf71e72	1	1	3m
wflow-job-7a1aa6	1	1	1m
wflow-job-3243737b-4911-42e8-896c-326722283631	1	1	1m
wflow-job-3cb26d1c-b9a7-492c-bc7a-118f7969d889	1	1	1m
wflow-job-7e7eft	1	1	1m
wflow-job-4aa2a473-e82e-459a-9323-35371d907871	1	1	1m
wflow-job-4b27dba5-ddc6-46d3-99f2-6457affe413d	1	1	1m
wflow-job-5a02c36f-fd98-4e1d-983a-399186a61121	1	1	1m
wflow-job-5e3c5079-027d-4444-8c74-c1eae51d68c8	1	1	35s
wflow-job-6f0cbb42-3c35-454f-a868-47a6cad3481b	1	1	3m
wflow-job-787f8526-25da-4117-996d-628aec16fe91	1	1	3m
wflow-job-9d5b47e0-d3be-4dae-8c3f-dd70f62a2fa4	1	1	4m

File Edit View Search Terminal Help

Every 2.0s: kubectl --context fed get clusters

NAME	AGE
atlas-recast-x	5d
atlas-recast-y	5d
atlas-recast-z	5d

Credits



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

Lukas Heinrich, REANA / RECAST

Kelsey Hightower

Thank You



Great Community, Amazing Tools

```

File Edit View Search Terminal Help
Every 1.0s: kubectl --context fed get job                                rocha250: Wed Apr 25 17:59:06 2018

NAME          DESIRED  SUCCESSFUL  AGE
p12           1         0            6h
wflow-job-0bf2ee91-0387-4e6b-a357-f2dc389aa9da  1         0            51m
wflow-job-5e3c5079-027d-4444-8c74-c1eae51d68c8  1         1            35s
wflow-job-9a49ac80-6792-46e6-9782-109ab0bd151f  1         1            51m
wflow-job-d6b84b74-c944-4539-8b30-14be0115f559  1         1            36s
wflow-job-d8dc0d3e-3b51-437a-b960-17422bc44b5  1         1            35s

```

```

Terminal
File Edit View Search Terminal Help
Every 1.0s: kubectl --context atlas-recast-x get job                    rocha250: Wed Apr 25 17:59:06 2018

NAME          DESIRED  SUCCESSFUL  AGE
wflow-job-05a48de3-fdaa-4a25-bb3d-6e65730410d9  1         1            4m
wflow-job-0767bf09-8a30-4866-8aa7-21a4b59adac9  1         1            1m
wflow-job-0eec040d-3cc7-410f-a61f-5da96b674d3c  1         1            3m
wflow-job-178998f2-6e8d-4c62-9dbb-80674fcac266  1         1            1m
wflow-job-1b04c55d-ec0d-4a3b-ba90-043a9570d330  1         1            4m
wflow-job-32648046-fa23-4435-b827-cf94a5351d40  1         1            34s
wflow-job-32a1f4f3-80bc-4650-ba3b-59087f06ab07  1         1            4m
wflow-job-3bd70c21-98be-4b52-abae-4c3a0b3d9178  1         1            3m
wflow-job-3dc192c8-35db-4672-9196-5067feb00977  1         1            4m
wflow-job-41bd7b4b-6fc9-43ca-ab86-b24997570e57  1         1            4m
wflow-job-6315c7ab-83c6-4327-89aa-372e67258c60  1         1            2m
wflow-job-6e6f05de-ae48-4a9a-81c7-520a939e1e95  1         1            3m
wflow-job-6f51ccf7-0100-4769-9255-ab1196604938  1         1            1m
wflow-job-73ed6deb-5f9c-46f6-8b2b-cd86878d160a  1         1            3m

```

```

Terminal
File Edit View Search Terminal Help
Every 1.0s: kubectl --context atlas-recast-y get job                    rocha250: Wed Apr 25 17:59:06 2018

NAME          DESIRED  SUCCESSFUL  AGE
wflow-job-05dc99b4-18a7-4a0f-a1ad-79d127d42de3  1         1            1m
wflow-job-0eca84f3-7599-42e8-976e-2537424d37a4  1         1            1m
wflow-job-1373465c-9c82-4c71-91a9-626a691959be  1         1            3m
wflow-job-18e38cd1-6bea-407b-806f-80932dad9508  1         1            1m
wflow-job-2ca4735c-4c65-4f7a-b5fa-43186292c73b  1         1            4m
wflow-job-4712bc1c-6a94-42f2-96ed-1d6558306859  1         1            3m
wflow-job-560db075-96c1-4304-bb59-558de1e73860  1         1            3m
wflow-job-60181635-745e-4b00-9203-249903f9a894  1         1            3m
wflow-job-61ba2027-e11f-4d35-aae9-241539782e0e  1         1            1m
wflow-job-6fa77b4a-4a67-46c0-b072-18f581abd81  1         1            3m
wflow-job-6fce076a-fb80-4999-bb2e-f03a2cf0b28  1         1            1m
wflow-job-796a8002-6c91-489a-80de-2f807d0ec160  1         1            2m
wflow-job-7a1aa0d5-7ed1-425a-bbd3-6f6b5a455278  1         1            1m
wflow-job-7e7efb4a-fcbd-4e33-b5f8-bdebafbf99  1         1            4m

```

```

Terminal
File Edit View Search Terminal Help
Every 1.0s: kubectl --context atlas-recast-z get job                    rocha250: Wed Apr 25 17:59:06 2018

NAME          DESIRED  SUCCESSFUL  AGE
wflow-job-05835422-1aa8-4d53-b72e-e73b7b8b728d  1         1            2m
wflow-job-0f54d731-25d4-4298-9c33-e07ba7d5c4bb  1         1            3m
wflow-job-19794703-3dc7-4ff0-b149-0303c2cd13eb  1         1            4m
wflow-job-2cdda822-3b49-49e0-ab41-494e90ab5c68  1         1            1m
wflow-job-3169fd59-d69c-4b27-9f0b-ba07ec7f1e72  1         1            4m
wflow-job-3243737b-4911-42e8-896c-326722283631  1         1            2m
wflow-job-3cb26d1c-b9a7-492c-bc7a-1187f969d889  1         1            3m
wflow-job-4aa2a473-e82e-459a-9323-35371d907871  1         1            1m
wflow-job-4b27dba5-ddc6-4e03-99f2-6457af6e413d  1         1            1m
wflow-job-5a02e3f1-fd99-4e1d-903b-309106a61121  1         1            1m
wflow-job-5e3c5079-027d-4444-8c74-c1eae51d68c8  1         1            35s
wflow-job-6f0cbb42-3c35-454f-a868-47a6cad3481b  1         1            3m
wflow-job-707f8526-25da-4117-996d-628aac16fe91  1         1            3m
wflow-job-9d5b47e0-d3be-4dae-8c3f-dd70f62a2f44  1         1            4m

```

yadage Workflow Service - Chrome

Not secure | https://atlas-recast-x-lbe5fd5xynhp-minion-0.32453/monitor/b907a91b-7f02-4b3e-afa3-e9298b88e2d

Yadage Service Home Examples Workflow Submissions Launch Workflow theinric_dummy (ATLAS) Logout

Workflow Monitor for b907a91b-7f02-4b3e-afa3-e9298b88e2d

Status: ACTIVE

Workflow Visualization

redraw layout collapse expand

Last seen: 2018-04-25 15:57:13

Log

Messages from the request processor will appear below.

```

2018-04-25T15:03:39. INFO:wflowbackend.backendtasks:acquiring context from wflowserver for wflowid b907a91b-7f02-4b3e-afa3-e9298b88e2d
2018-04-25T15:03:39. INFO:wflowbackend.backendtasks:running analysis on worker: wflow-nonint-b907a91b-7f02-4b3e-afa3-e9298b88e2d-76skn
2018-04-25T15:03:39. INFO:wflowbackend.backendtasks:acquiring wflow context from http://wflow-server.default.svc.cluster.local
2018-04-25T15:03:39. INFO:wflowbackend.process:processing ['setupsFromURL', 'generic_onsuccess', 'cleanup', '--config-from-server', 'b907a91b-7f02-4-
2018-04-25T15:03:39. INFO:wflowbackend.backendtasks:running analysis on worker: wflow-nonint-b907a91b-7f02-4b3e-afa3-e9298b88e2d-76skn
2018-04-25T15:03:39. INFO:wflowbackend.backendtasks:setting up for context {'shipout_spec': {'u': 'root', 'u': 'root', 'u': 'wflow-shiptarget.default-
2018-04-25T15:03:39. INFO:wflowbackend.backendtasks:prepared workdir workdirs/b9/07/a9/1b-7f02-4b3e-afa3-e9298b88e2d
2018-04-25T15:03:39. INFO:wflowbackend.backendtasks:preparing workdir workdirs/b9/07/a9/1b-7f02-4b3e-afa3-e9298b88e2d
2018-04-25T15:03:39. WARNING:wflowbackend.backendtasks:No input archive specified, skipping download
2018-04-25T15:03:39. INFO:wflowbackend.process:setting up entry point wflowyadageworker.backendtasks:run_workflow
2018-04-25T15:03:39. INFO:wflowyadageworker.backendtasks:running yadage workflow for context: {'shipout_spec': {'u': 'root', 'u': 'root', 'u': 'wflow-
2018-04-25T15:03:39. INFO:wflowyadageworker.backendtasks:connecting backend: py:sharedstatekube:backend
2018-04-25T15:03:39. INFO:yadage.utils:setting up backend py:sharedstatekube:backend with opts {'optsyaml': '/yadageconfig/backendopts'}
2018-04-25T15:03:40. INFO:wflowyadageworker.backendtasks:yadage base settings ('updateinterval': 30.0, 'dataarg': 'workdirs/b9/07/a9/1b-7f02-4b3e-a-
2018-04-25T15:03:40. INFO:wflowyadageworker.simple_workflow:running workflow on context {'shipout_spec': {'u': 'root', 'u': 'root', 'u': 'wflow-ship-
2018-04-25T15:03:40. INFO:wflowyadageworker.simple_workflow:no workflow parameters in init dir
2018-04-25T15:03:40. INFO:wflowyadageworker.backendtasks:load workflow from source
2018-04-25T15:03:40. INFO:wflowyadageworker.backendtasks:load workflow from source

```