



KubeCon



CloudNativeCon

Europe 2018

Multi-Cloud Federated Kubernetes at CERN

Clenimar Filemon
clemenar@lsd.ufcg.edu.br

Ricardo Rocha
ricardo.rocha@cern.ch





Founded in 1954

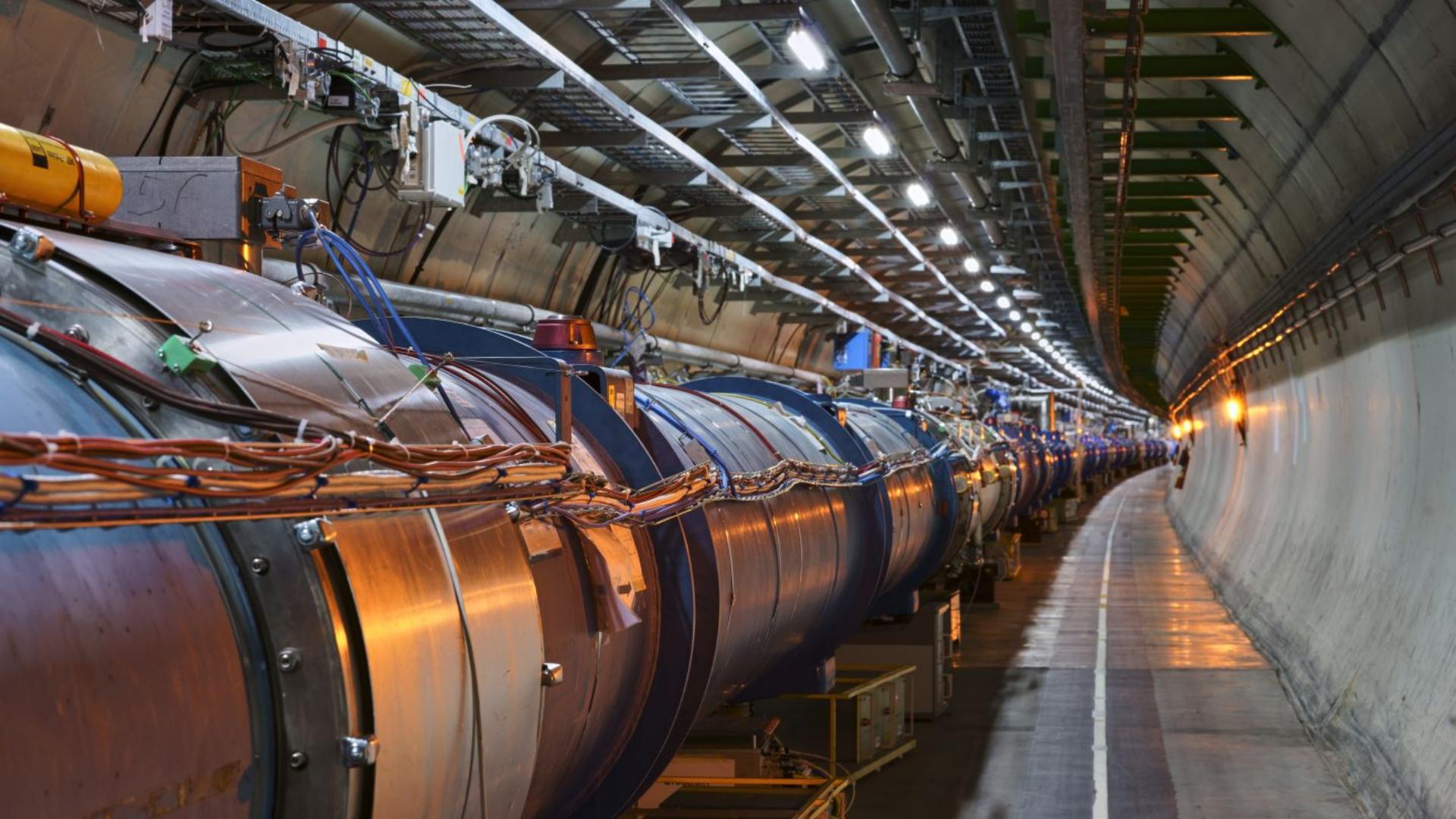
Fundamental Science

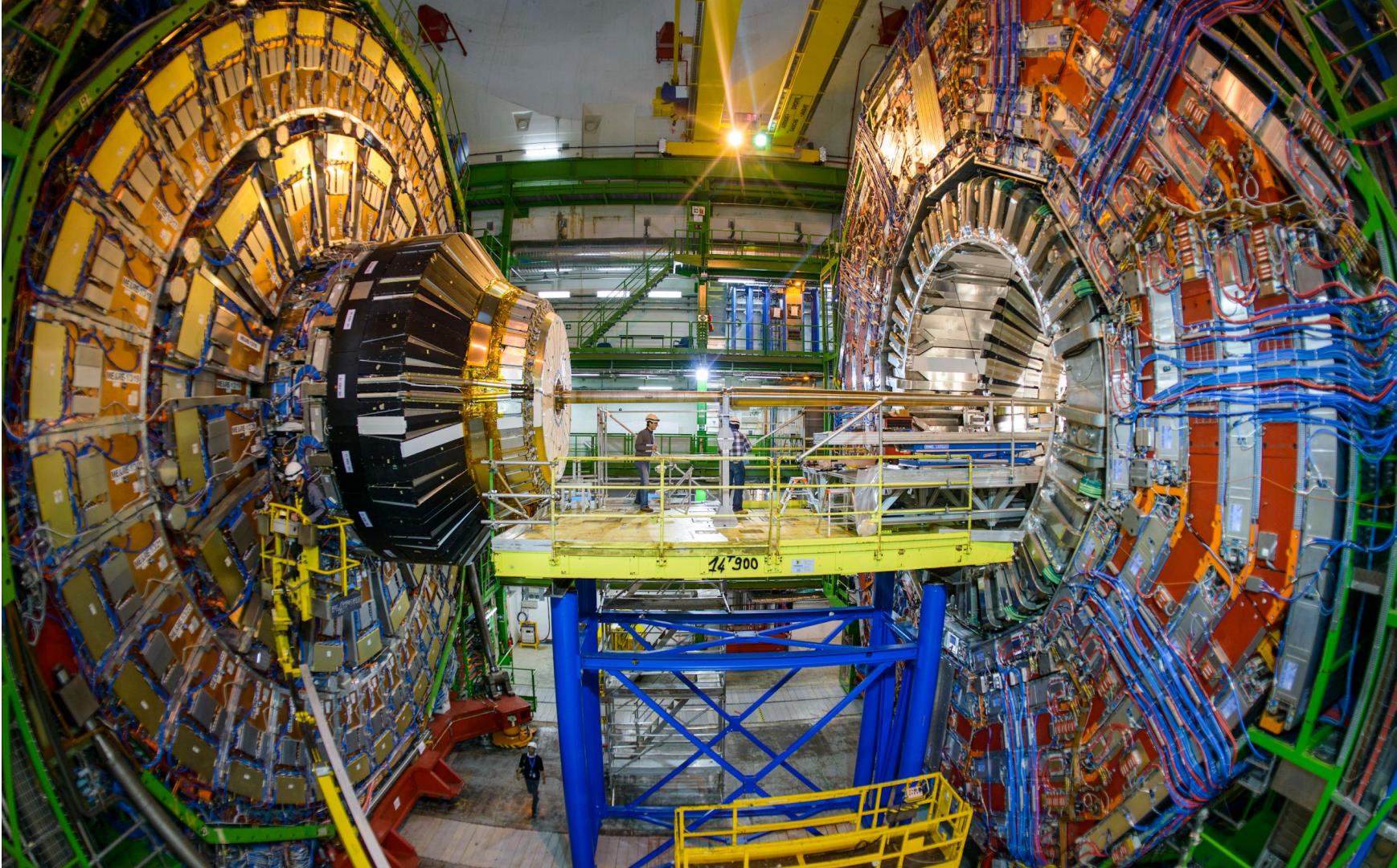
What is 96% of the universe made of?

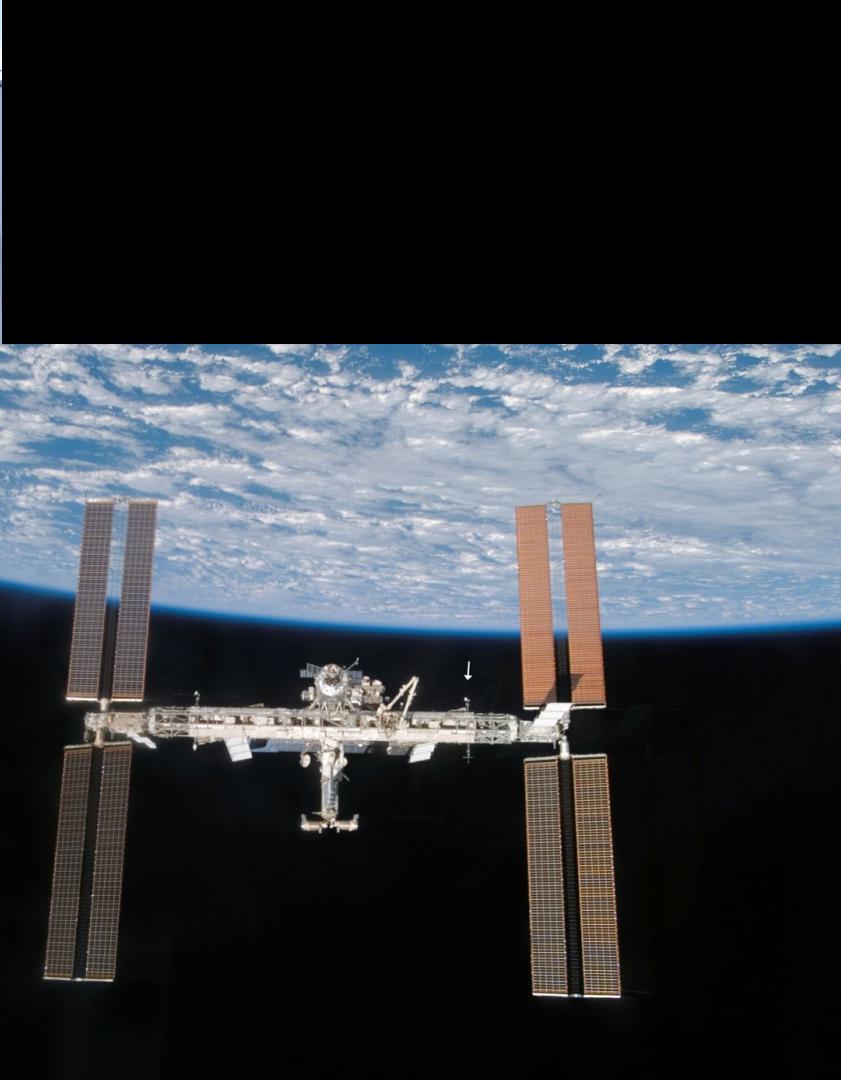
What was the state of matter just after the Big Bang?

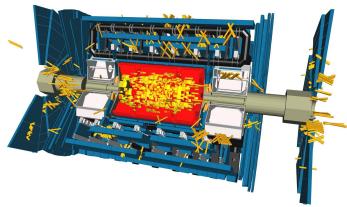
Why isn't there anti-matter in the universe?









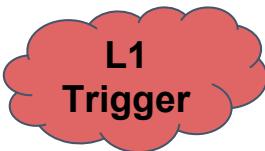


Collisions

~40 MHz

~ 1PB/sec

Huge Data



Hardware Filter

~ 100 kHz

Still Big



Software Filter

~ 1 kHz

~ 1-10 GB/s

Still Big

Raw Data





320 000 Cores

4 300 Projects

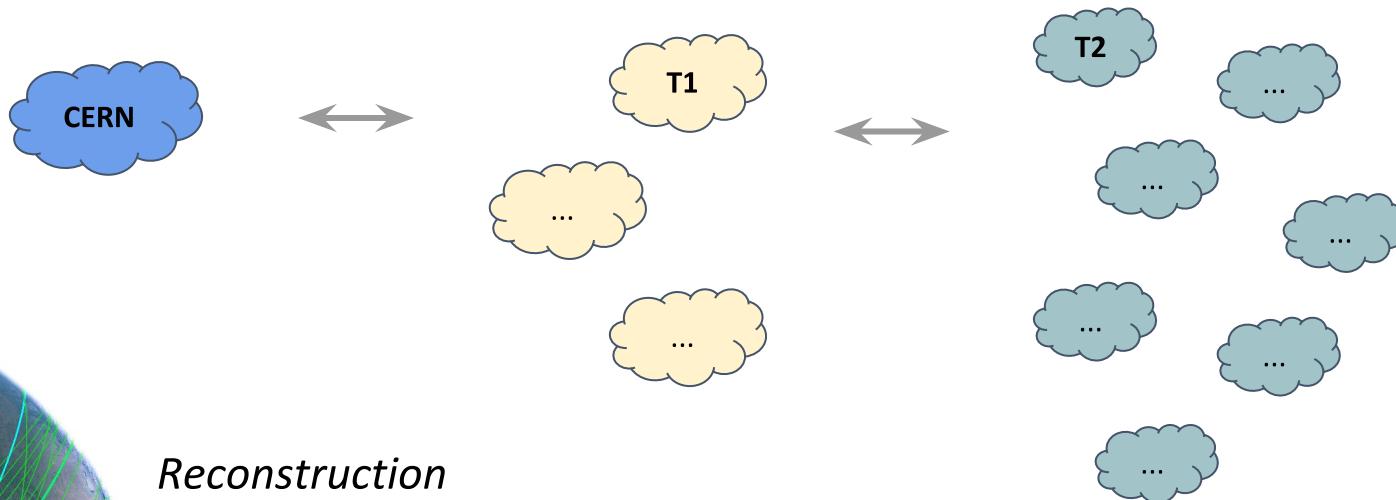
250 Petabytes

3 300 Users

10 000 Hypervisors

210 Kubernetes Clusters

Distributed Computing



200+ Sites

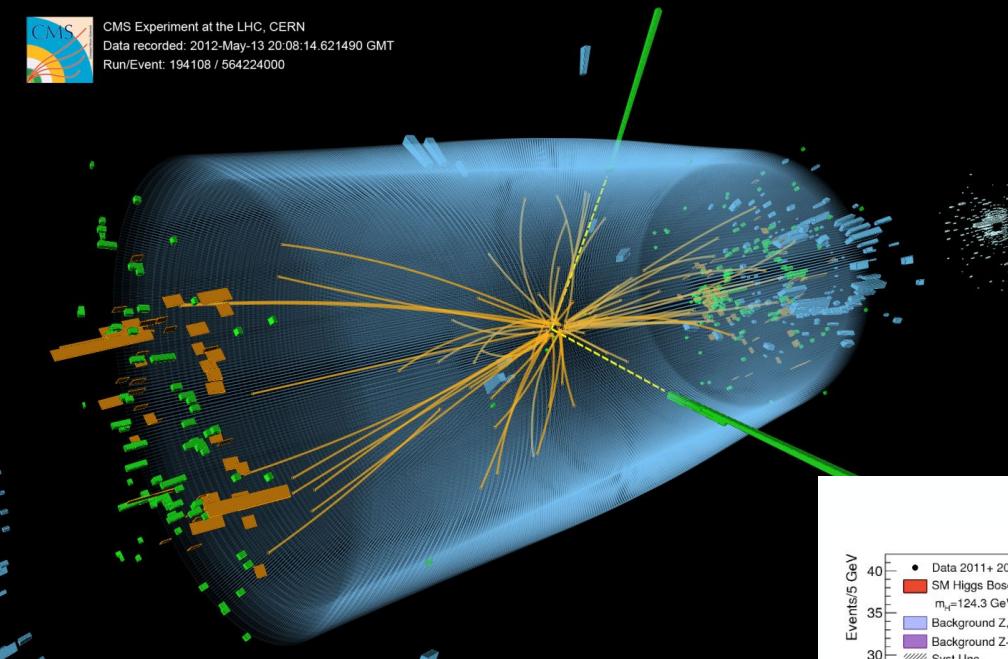
~400 000 Jobs

700 000 Cores

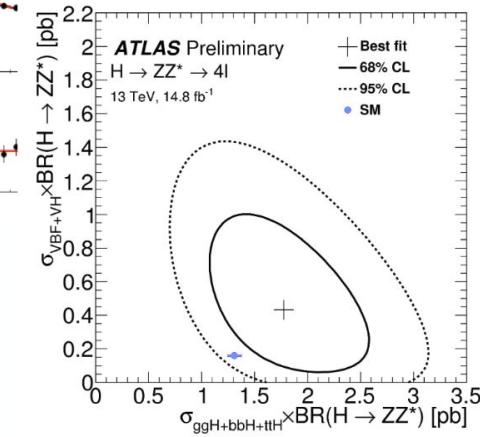
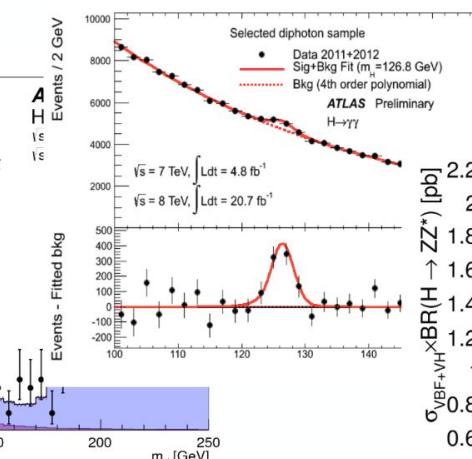
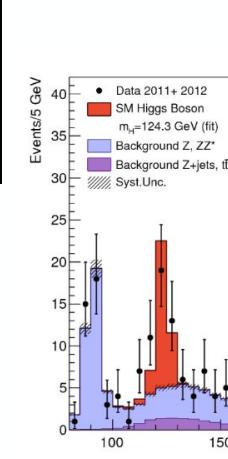
~30 GiB/s



CMS Experiment at the LHC, CERN
Data recorded: 2012-May-13 20:08:14.621490 GMT
Run/Event: 194108 / 564224000



CMS Higgs Event, May 2012



Motivation for Federation

Periodic load spikes

International Conferences, Reconstruction Campaigns

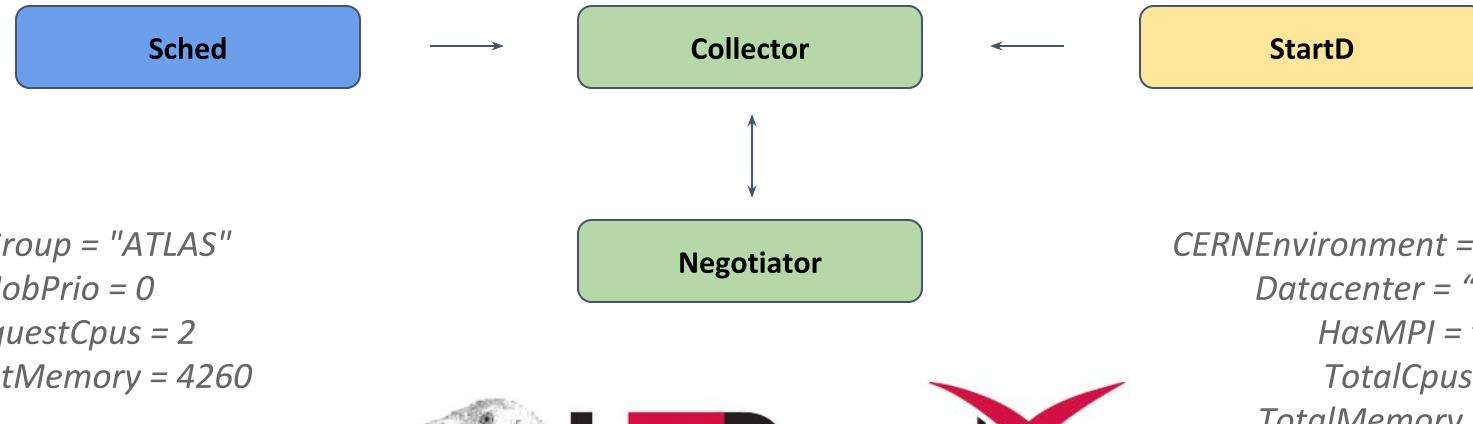
Simplification

Monitoring, Lifecycle, Alarms

Deployment

Uniform API, Replication, Load Balancing

Use Cases: CERN Batch System, RECAST Analysis



Matchmaking with ClassAds

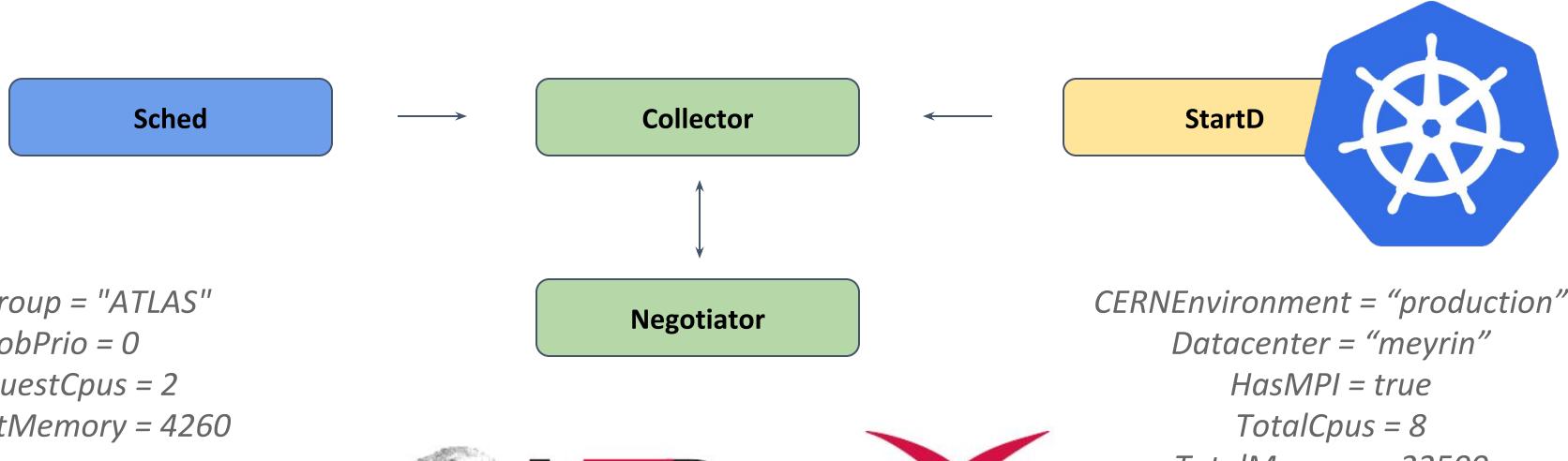
Fair Share

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

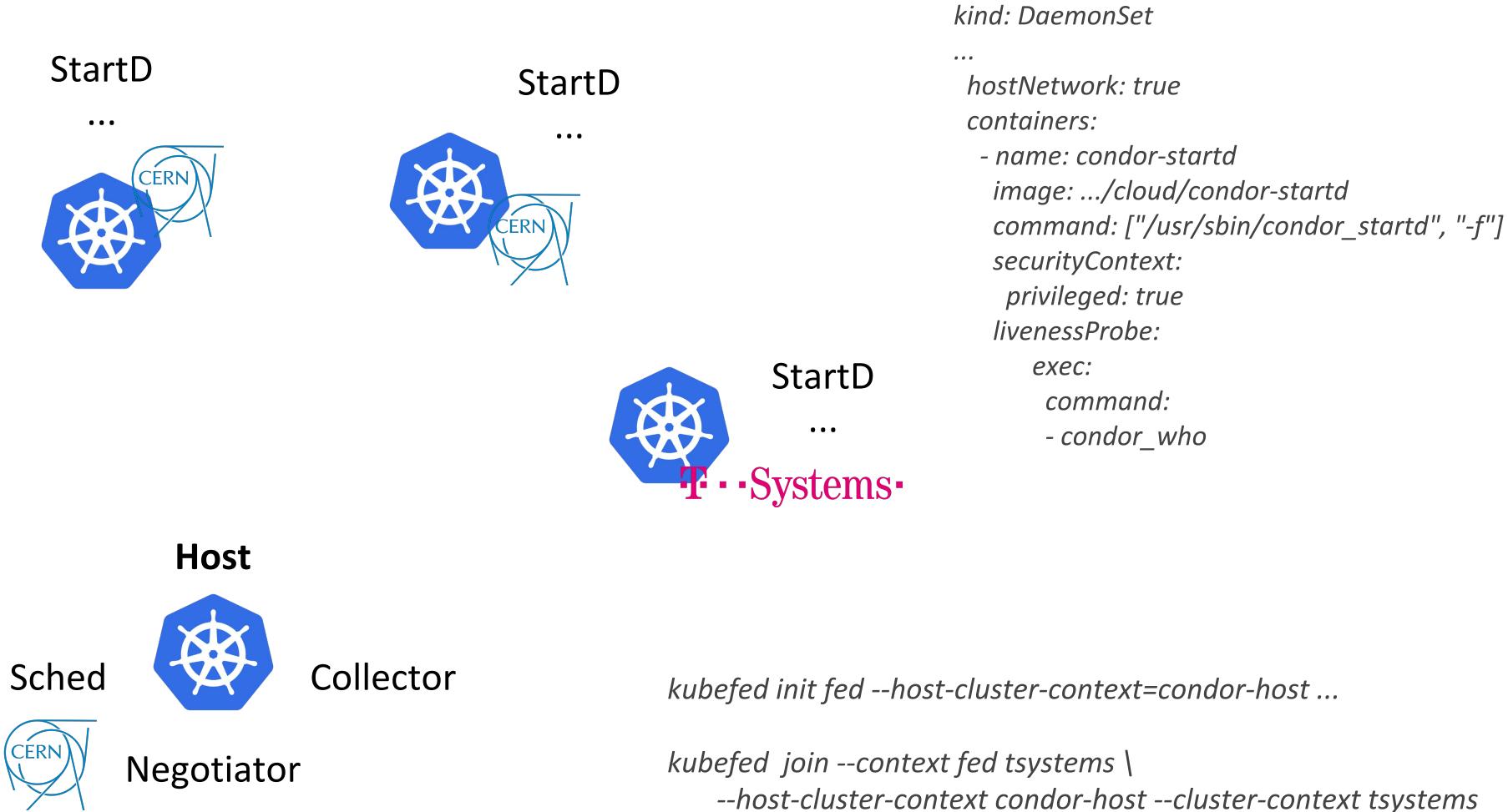


Collector



Negotiator

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



REANA / RECAST

Reusable Analysis Platform

Workflow Engine (Yadage)

Each step a Kubernetes Job

Integrated Monitoring & Logging

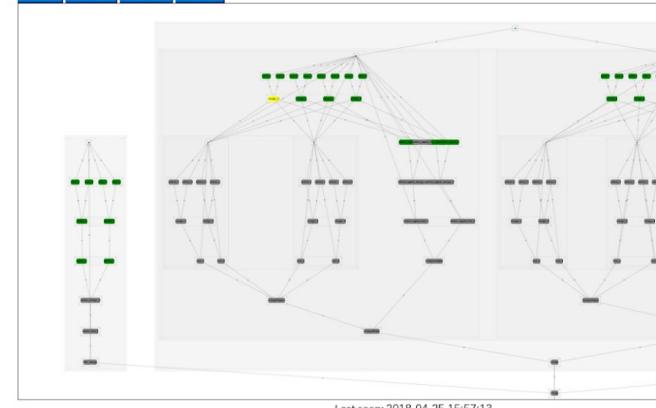
Centralized Log Collection

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.

```
2018-04-25T15:53:39.. INFO:wflowbackend:process:acquiring context from wflowserver for wflowid b907a91b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25T15:53:39.. INFO:wflowbackend:backendtasks:running analysis on worker: wflow-nonint-b907a91b-7f02-4b3e-afa3-e9298b888e2d-76skn
2018-04-25T15:53:39.. INFO:wflowbackend:backendtasks:acquiring wflow context from http://wflow-server.default.svc.cluster.local
2018-04-25T15:53:39.. INFO:wflowbackend:process:processing ['setupFromURL', 'generic_onsuccess', 'cleanup', '-config-from-server', 'b907a91b-7f02-4b3e-afa3-e9298b888e2d-76skn']
2018-04-25T15:53:39.. INFO:wflowbackend:backendtasks:running analysis on worker: wflow-nonint-b907a91b-7f02-4b3e-afa3-e9298b888e2d-76skn
2018-04-25T15:53:39.. INFO:wflowbackend:backendtasks:setting up for context {u'shipout_spec': {u'user': u'root', u'host': u'wflow-shipitarget.default'}}
2018-04-25T15:53:39.. INFO:wflowbackend:backendtasks:prepare:workdir:workdirs/b907a91b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25T15:53:39.. INFO:wflowbackend:backendtasks:preparing workdir workdirs/b907a91b-7f02-4b3e-afa3-e9298b888e2d
2018-04-25T15:53:39.. WARNING:wflowbackend:backendtask:No input archive specified, skipping download
2018-04-25T15:53:39.. INFO:wflowbackend:process:setting up entry point wflowyadageworker.backendtasks:run_workflow
2018-04-25T15:53:39.. INFO:wflowyadageworker:backendtasks:running yadage workflow for context: {u'shipout_spec': {u'user': u'root', u'host': u'wflow-shipitarget.default'}}
2018-04-25T15:53:39.. INFO:wflowyadageworker:backendtasks:connecting backend: py sharedstatekubeb:backend
2018-04-25T15:53:39.. INFO:yadage.util:setting up backend py:sharedstatekubeb:backend with opts {'options': {'yadageconfig/backendopts': ''
2018-04-25T15:53:40.. INFO:wflowyadageworker:backendtasks:yadabe base settings {'updateinterval': 30.0, 'dataarg': 'workdirs/b907a91b-7f02-4b3e-afa3-e9298b888e2d-76skn'}
2018-04-25T15:53:40.. INFO:wflowyadageworker:simple_workflow:running workflow on context {u'shipout_spec': {u'user': u'root', u'host': u'wflow-shipitarget.default'}}
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
```

<https://github.com/recast-hep>

<https://github.com/diana-hep/yadage>

<https://github.com/reanahub>

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-cleae51d68c8	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-d8dc03e-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-0eeecTerminal

File Edit View Search Terminal Help

wflow-job-1bd^{Every 1.0s: kubectl --context atlas-recast-y get job}

rocha250: Wed Apr 25 17:59:06 2018

NAME	DESIRED	SUCCESSFUL	AGE
wflow-job-32a1	1	1	1m
wflow-job-3bd7	1	1	1m
wflow-job-3dc1	1	1	1m

wflow-job-41bcTerminal

wflow-job-6315Terminal

wflow-job-6effwflow-job-18e38cFile Edit View Search Terminal Help

wflow-job-6f5jwflow-job-2cac47^{Every 1.0s: kubectl --context atlas-recast-z get job}

rocha250: Wed Apr 25 17:59:06 2018

NAME	DESIRED	SUCCESSFUL	AGE
wflow-job-60181fwflow-job-05835422-1aa8-4d53-b726-e73b7b8b728d	1	1	2m
wflow-job-61ba2ewflow-job-0f54b731-25d4-4298-9c33-eb7ba7d5c4bb	1	1	3m
wflow-job-6faf7lwflow-job-1979e7b3-3dc7-4ffd-b443-0303c26d13eb	1	1	4m
wflow-job-6fce7ewflow-job-2cd822-3b49-49e8-ab41-494e90abc568	1	1	1m

NAME	DESIRED	SUCCESSFUL	AGE
wflow-job-796a6wflow-job-3169fd59-d69c-4b27-9f0b-ba07ecf71e72	1	1	4m
wflow-job-7a1aaewflow-job-3243737b-4911-42e8-896c-326722283631	1	1	2m
wflow-job-7e7eflwflow-job-3cb26d1c-b9a7-492c-bc7a-118f7969d889	1	1	3m

NAME	DESIRED	SUCCESSFUL	AGE
wflow-job-4aa2a473-e82e-459a-9323-35371d907871	1	1	1m
wflow-job-4b27dba5-ddc6-46d3-99f2-6457afffe413d	1	1	1m
wflow-job-5a02c36f-fd98-4e1d-983a-399186a61121	1	1	1m

NAME	DESIRED	SUCCESSFUL	AGE
wflow-job-5e3c5079-027d-4444-8c74-cleae51d68c8	1	1	35s
wflow-job-6f0ccb42-3c35-454f-a868-47a6cad3481b	1	1	3m
wflow-job-787f8526-25da-4117-996d-628aec16fe91	1	1	3m

NAME	DESIRED	SUCCESSFUL	AGE
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

