



INDIGO - DataCloud
Better Software for Better Science

Geographically distributed Batch System as a Service: the INDIGO-DataCloud approach exploiting HTCondor

Speaker: Vallero Sara (INFN Torino)

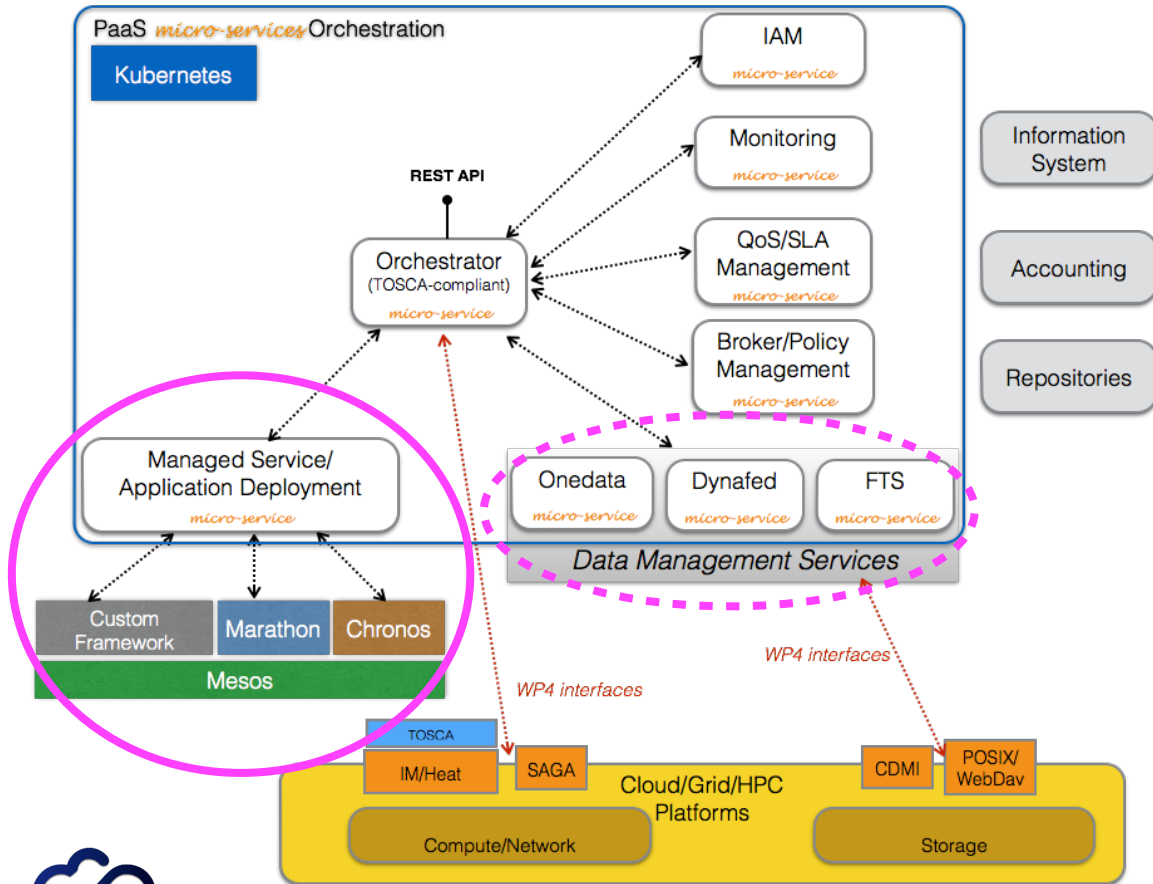
CHEP 2016 - October 10/14 - San Francisco (CA)

Talk outline

- goals of the *Batch System as a Service* work-package within the INDIGO-DataCloud project
- description of the software tools used: Docker, Mesos, Calico
- first solution using Marathon implemented
- first studies of geographical deployment:
 - network topology
 - services distribution and provisioning model
 - smoke tests and first benchmarking results

Our mission

Key components of the INDIGO PaaS framework.



The goal:

- keep delivering a well consolidated computational framework, while **complying to modern computing paradigms**
- ease system administration to all levels (hardware/applications)
- provide a smooth end-user experience

Batch System as a Service:

- **automatically and dynamically deploy a complete batch system cluster** (with appropriate user interfaces) **in highly-available and scalable configurations**
- use **HTCondor**:
 - widely used within the scientific community
 - cloud aware
- networking:
 - HTCondor shared port mode
 - **overlay networks**
 - **span the cluster over multiple sites** (Connection Broker)
- storage:
 - scalability, performance and reliability
 - CVMFS (using Parrot)
 - integrate with the INDIGO Data Services (Onedata, Dynafed, FTS)

