INDIGO - DataCloud

RIA-653549

Recenti sviluppi in INDIGO DataCloud

18/5/2016

Better Software for Better Science.

Giacinto Donvito
INFN-Bari
INDIGO-DataCloud Technical Director

INDIGO-DataCloud is co-founded by the Horizon 2020Framework Programme

Key Dates

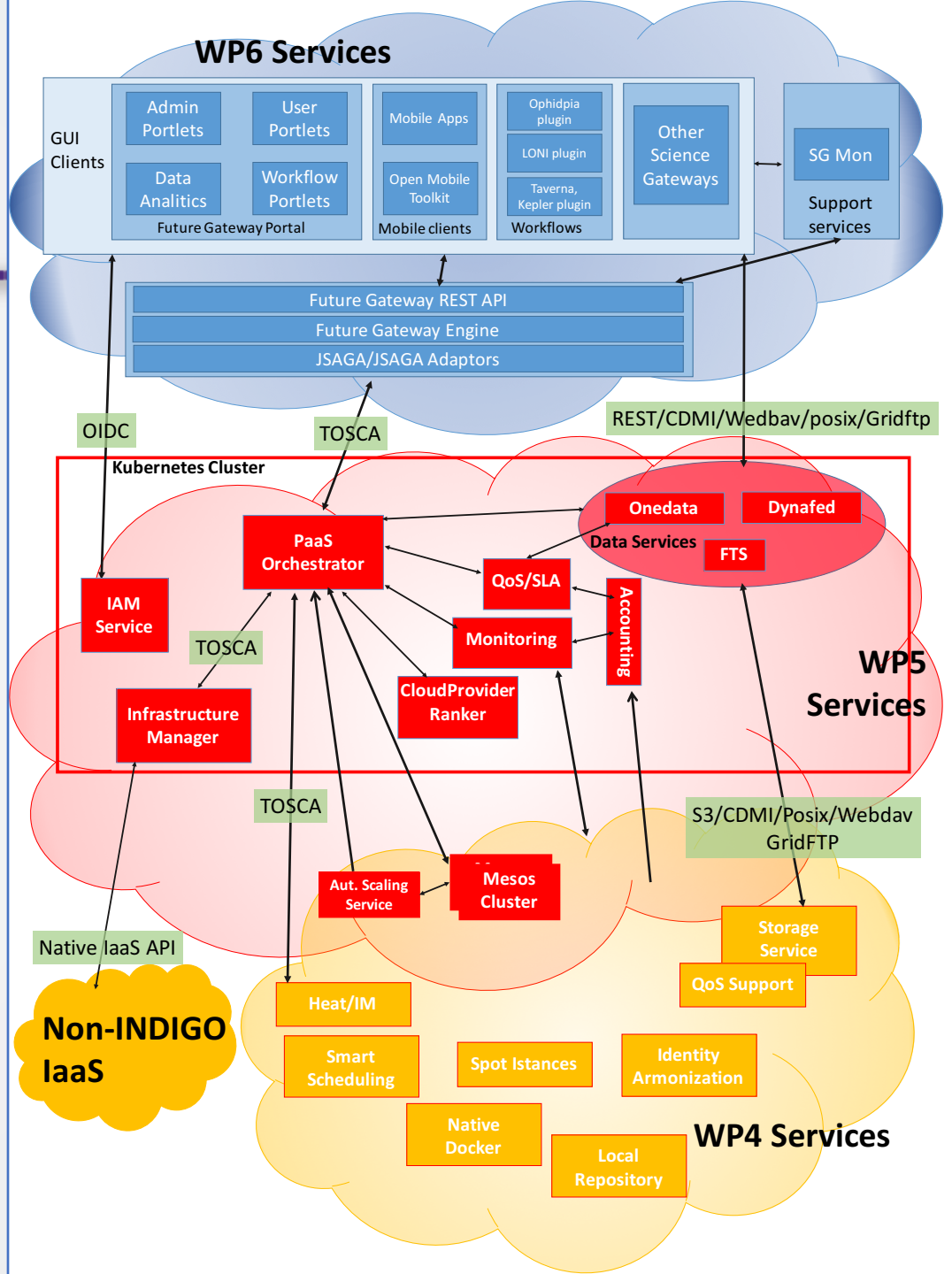
- INDIGO-DataCloud started on April 1st, 2015.
- Kick-off meeting on 22-24/4/2015.
- The first beta release was ready by March 2016, with first demos shown at: EGI Conference Nov 2015 (Bari), CloudScape 8-9 March 2016, Brussels and on 4-5/4/2016 in Amsterdam at the "INDIGO Champions meet Developers" meeting.
- 4-6 May 2016, Frascati: All-Hands, Collaboration Board and Technical Board meetings.
- **First public release: 1st August, 2016**
- **Mid-term review by the EC** scheduled on 19-20 September 2016 in Bologna.
- Second public release due by March 2017.
- The project will end on September 30th, 2017.

- See the path for our releases at https://www.indigo-datacloud.eu/indigo-roadmap.

Some updates

- Full definition of the INDIGO Architecture, recently submitted to arXiv (http://arxiv.org/abs/1603.09536) and to a peer-reviewed journal.
- More than 120 ”user stories” collected.
- Submission of articles and presentations to many conferences (as an example, we recently submitted 11 abstracts to CHEP2016 only).
- 10+ concrete implementations being worked on right now by our scientific community champions, together with developers and high-level experts.
  - For instance: use cases by EuroBioImaging, LifeWatch, Large Binocular Telescope, Cherenkov Telescope Array, Elixir, Climate modeling, High-Energy Physics, EMSO, WeNMR/INSTRUCT.
  - With a number of publications in the pipeline.
- Several external collaborations and contributions to upstream projects.

INDIGO "Simplified" Architecture



The INDIGO solutions

- **The INDIGO architecture can be seen as providing:**
  - **Site-level solutions**
  - **Data solutions**
  - **Automated solutions**
  - **User-level solutions**
- All of them integrate in a consistent global framework. Frequently a given solution spans multiple INDIGO WPs.
- There are many details “behind the scenes”. They are essential and addressed by our Work Packages. But let’s focus here on a bird’s eye view from a practical perspective.

Site-level solutions

- New scheduling algorithms for open source Cloud frameworks.
  - Both **fair-share scheduling** and **spot instances**
- Full support for containers, *with or without Docker*.
- Dynamic partitioning of batch vs. Cloud resources.
- Storage QoS and data lifecycle support.
- Support for external infrastructures.
- Automated synchronization of dockerhub repos with the local repository of open source Cloud frameworks.
- Improved automation at IaaS level, based on TOSCA

Data solutions

- Integrated local and remote Posix access for all types of resources (bare metal, virtual machines, containers).
- Transparent mapping of object storage (e.g. Ceph, S3) to Posix.
- Transparent data caching and replicas.
- Transparent gateway to existing filesystems (e.g. GPFS, Lustre).
- Webdav, GridFTP, CDMI, web, fuse access.
- Dropbox-like functionalities, based on ownCloud (target: September 2016).
- Linux, Mac OS, Windows desktop support.

Automated solutions

They are typically based on TOSCA templates used to specify resource requirements, dependencies, and configuration of the services/applications (sample templates for common use cases are provided by the project).

- Selection of resources across multiple Cloud providers (e.g. depending on data location or resource requirements).
- Support for application requirements in Cloud resource allocations (e.g. for what regards InfiniBand or GPUs).
- Dynamic instantiation, automated monitoring and elasticity of long-running services.
- Dynamic instantiation, automated monitoring and elasticity of batch systems, front-ends included.
- Support for custom frameworks for porting arbitrary applications to the Cloud, with automated monitoring and scalability.
- Support for big data analysis applications (Ophidia, Spark).
- Mesos clusters transparently spanning multiple data centers (not in the first release).

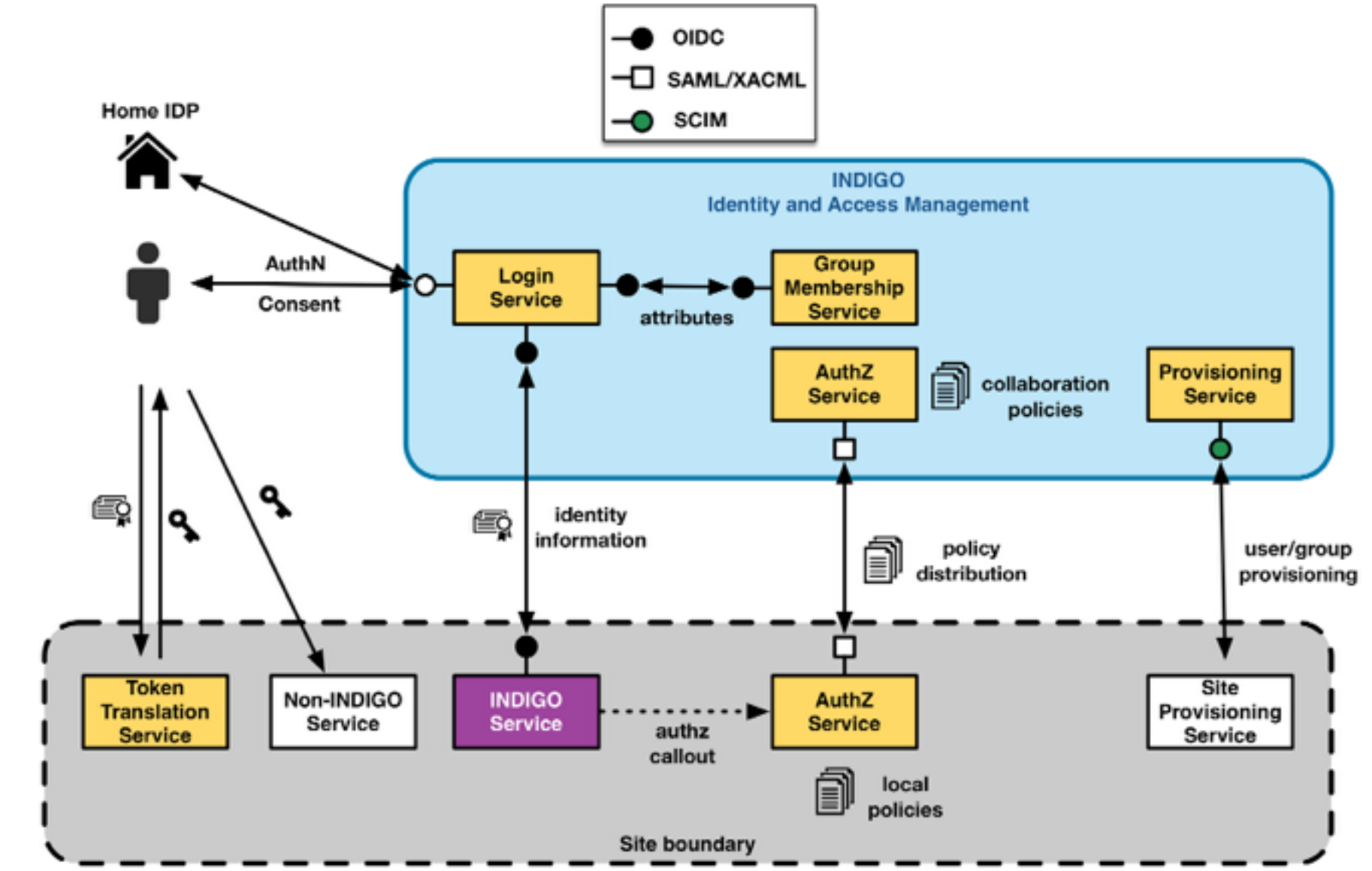
User-level solutions

- Customizable / programmable portal (gateway) engine integrated with the features mentioned in the previous slides.
- Sample portals delivered for selected applications ("all-in-one", "plug and play" bundles).
- Mobile toolkit to access INDIGO features on mobile devices.
- AAI architecture integrated at all levels supporting X.509, SAML, OpenID Connect.

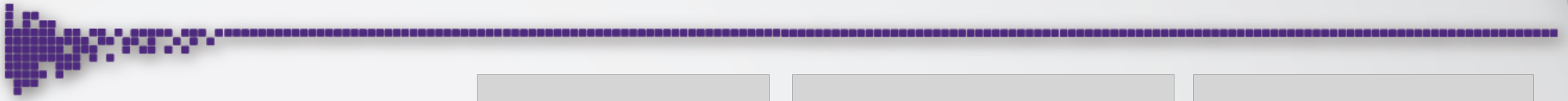
The 1st release: a starting point

- With the 1st release (1st August 2016), we'll provide both key components and their first concrete applications to use cases.
- We'll demo several of these in the coming weeks and months (including also the EC review in September).
- However, **this should only be a starting point**. After the first release is out, **we need to step up**:
  - Installations of the INDIGO software and application use cases first at the pilot level (WP3) and then at production level (data centers, e-infrastructures, adoption by user communities)
  - Communication activities
  - Exploitation activities

INDIGO AAI architecture



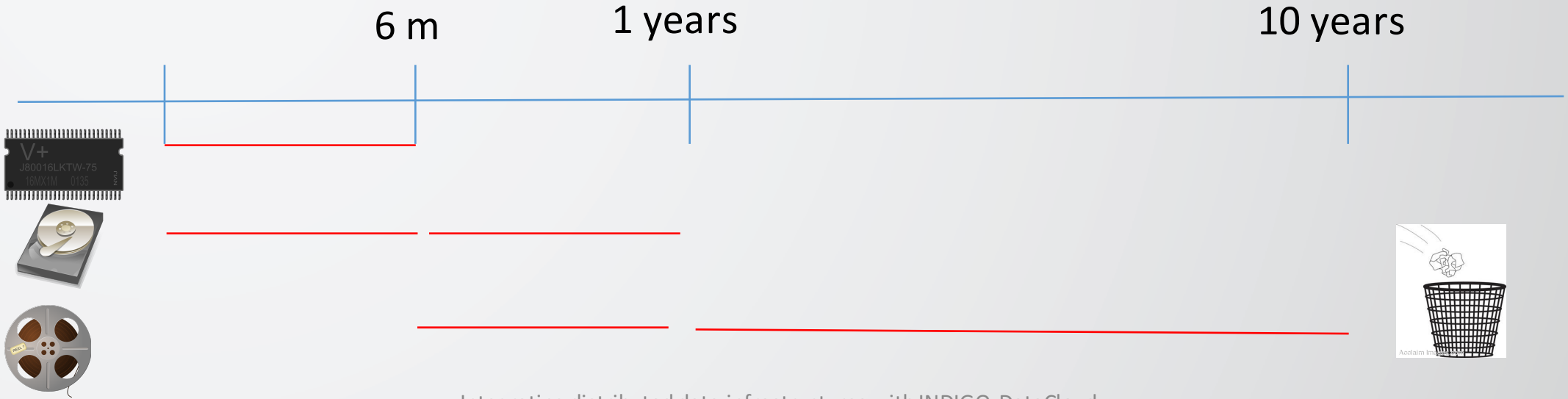
Storage Quality of Service and the Cloud



Amazon	S3		Glacier
Google	Standard	Durable Reduces Availability	Nearline
HPSS/GPFS	Corresponds to the HPSS Classes (customizable)		
dCache	Resilient	disk+tape	TAPE

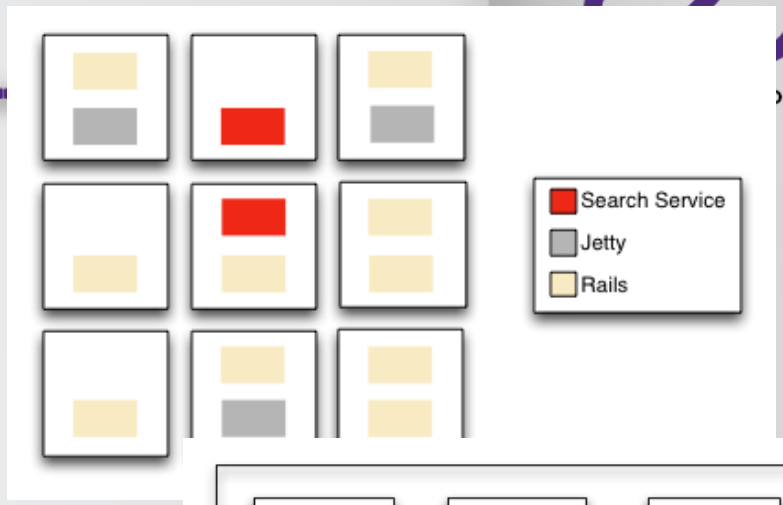
Next step : Data Life Cycle

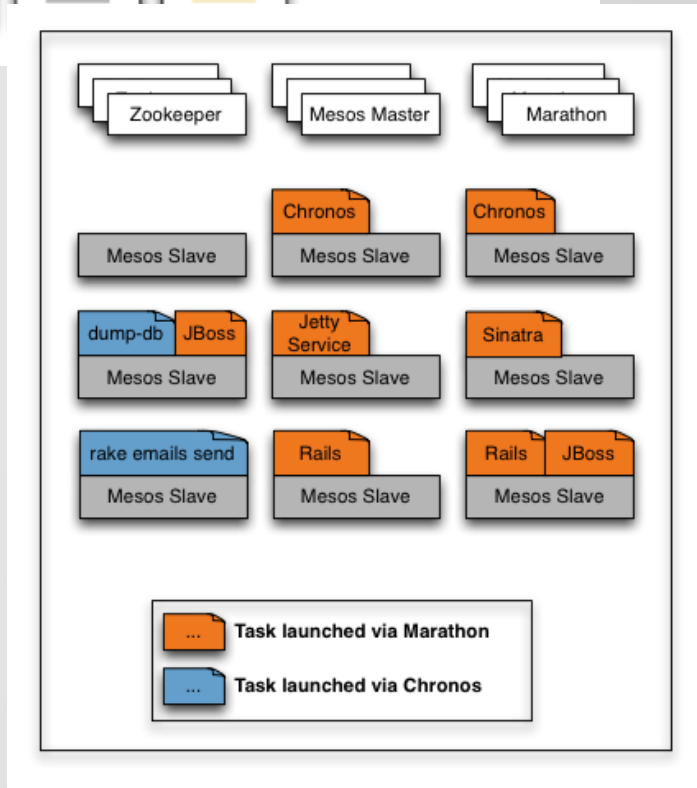
- Data Life Cycle is just the time dependent change of
  - Storage Quality of Service
  - Ownership and Access Control (PI Owned, no access, Site Owned, Public access)
  - Payment model: Pay as you go ; Pay in advance for rest of lifetime.
  - Maybe other things



On Mesos / Marathon /Chronos

- Marathon/Chronos run on top of a Mesos cluster
- Enables scale-out of services by deploying additional containers on spare nodes.
- Marathon will be used to instantiate the Long Running Services and keep them monitored and running
  - It will take care of recover in case of failure
- Chronos will be used in order to deal with application execution:
  - It is able to deal with dependency/re-executing the application in case of failure
- The Mesos cluster will be used in order to share the resources among different use-cases





Managed Services Deployment and Applications Execution through Mesos

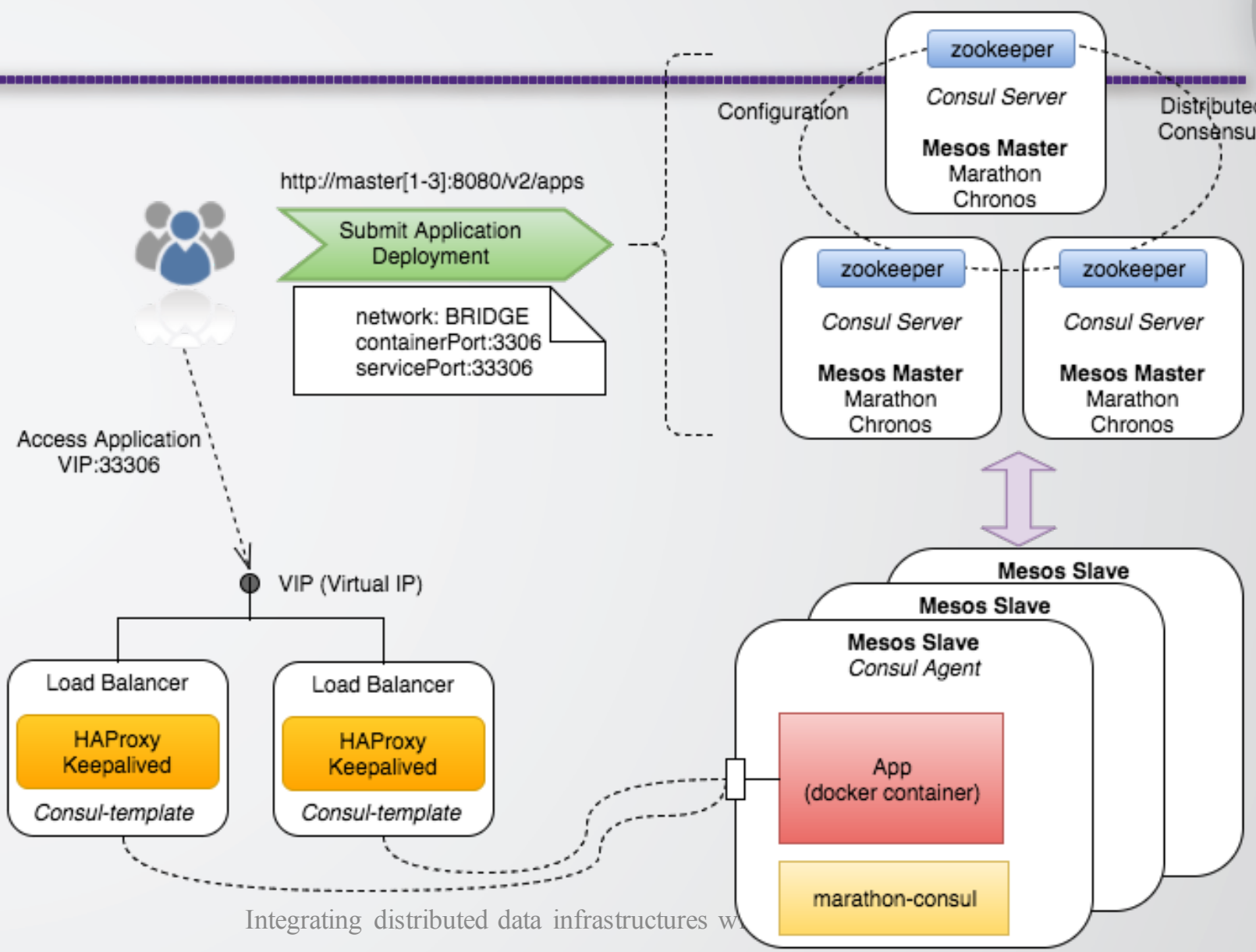
- **Mesos** is able to manage cluster resources (cpu, mem) providing *isolation* and *sharing* across distributed applications (frameworks)
- **Marathon** and **Chronos** are two poweful frameworks that can be deployed on top of a Mesos Cluster.
- INDIGO PaaS will use:
- Marathon to deploy, monitor and scale **Long-Running services**, ensuring that they are always up and running.
- Chronos to run user **applications** (jobs), taking care of fetching input data, handling dependencies among jobs, rescheduling failed jobs.

Status of the prototype

- Automatic deployment through ***Ansible*** recipes embedded in TOSCA and HOT templates
  - All the services run in docker containers;
- **High-availability** of the cluster components:
  - Leader election among master nodes managed by Zookeeper;
  - HA Load-balancing;
- **Service discovery** through **Consul** that provides also DNS functionality and health checks;
  - services are automatically registered in Consul as soon as they are deployed on the cluster
- The external access to the deployed services is ensured through HAProxy **load-balancers in HA** (using keepalived and a floating IP);
- Work in progress to enable Calico for the Network configuration
- **SSL** support and basic **authentication** enabled for Marathon/Chronos

Service Deployment and application execution

We are now working on adding a Calico network configuration

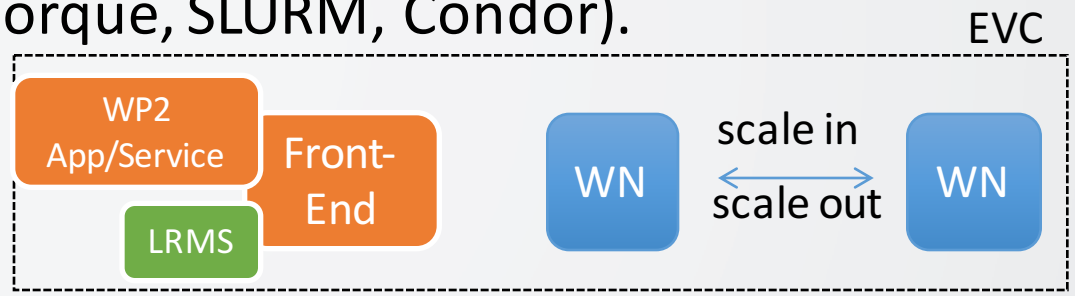


Mesos: First Release

- Mesos cluster fully integrated in the PaaS:
  - INDIGO **Orchestrator** will be able to interact directly with Marathon/Chronos to manage user requests for the deployment of Long-Running Services or the execution of applications
  - The cluster metrics will be collected by INDIGO **Monitoring** pillar
  - The cluster elasticity will be automatically ensured by INDIGO **Scaling Service**
  - The **authentication** mechanisms will be extended through the integration with INDIGO **IAM**
  - Data access and management will be integrated with the **Data Services** provided by INDIGO

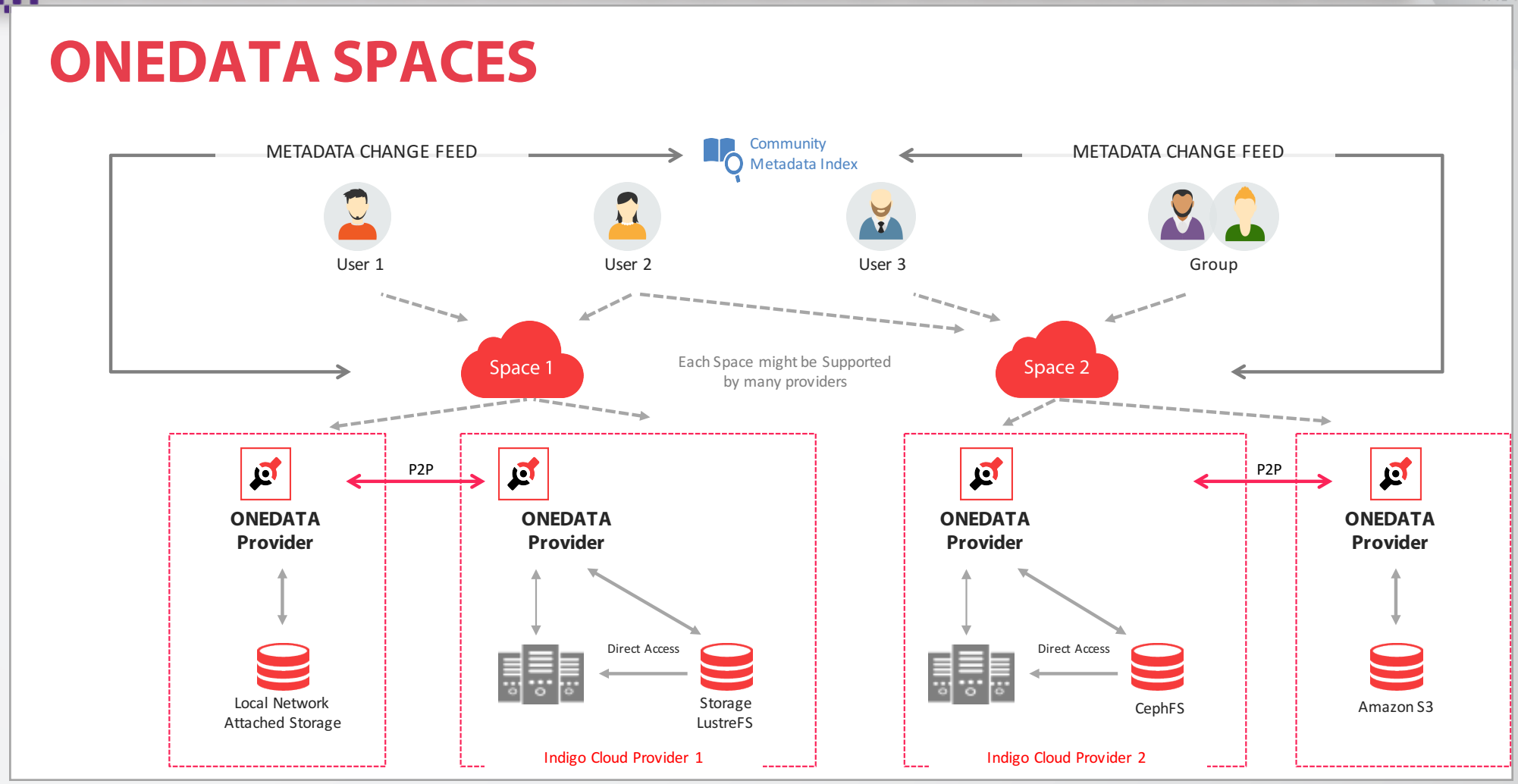
Automated Scaling Service - Goal

- On-demand provision of computing resources (horizontal elasticity).
- Two levels
  - Level 1: Customized Elastic Virtual Clusters (EVCs) supporting multiple LRMS (Torque, SLURM, Condor).

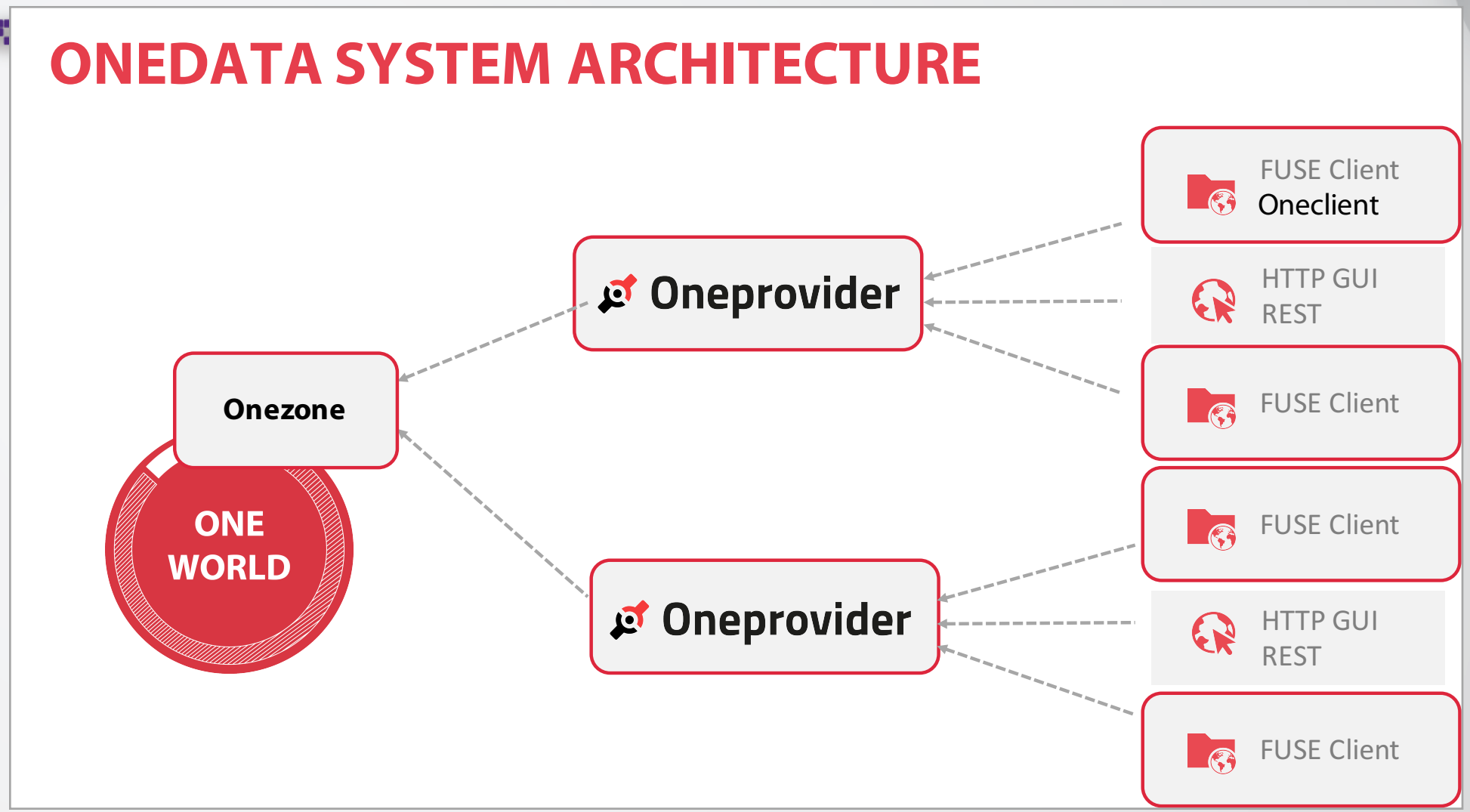


  - Level 2: Elastic Mesos Clusters

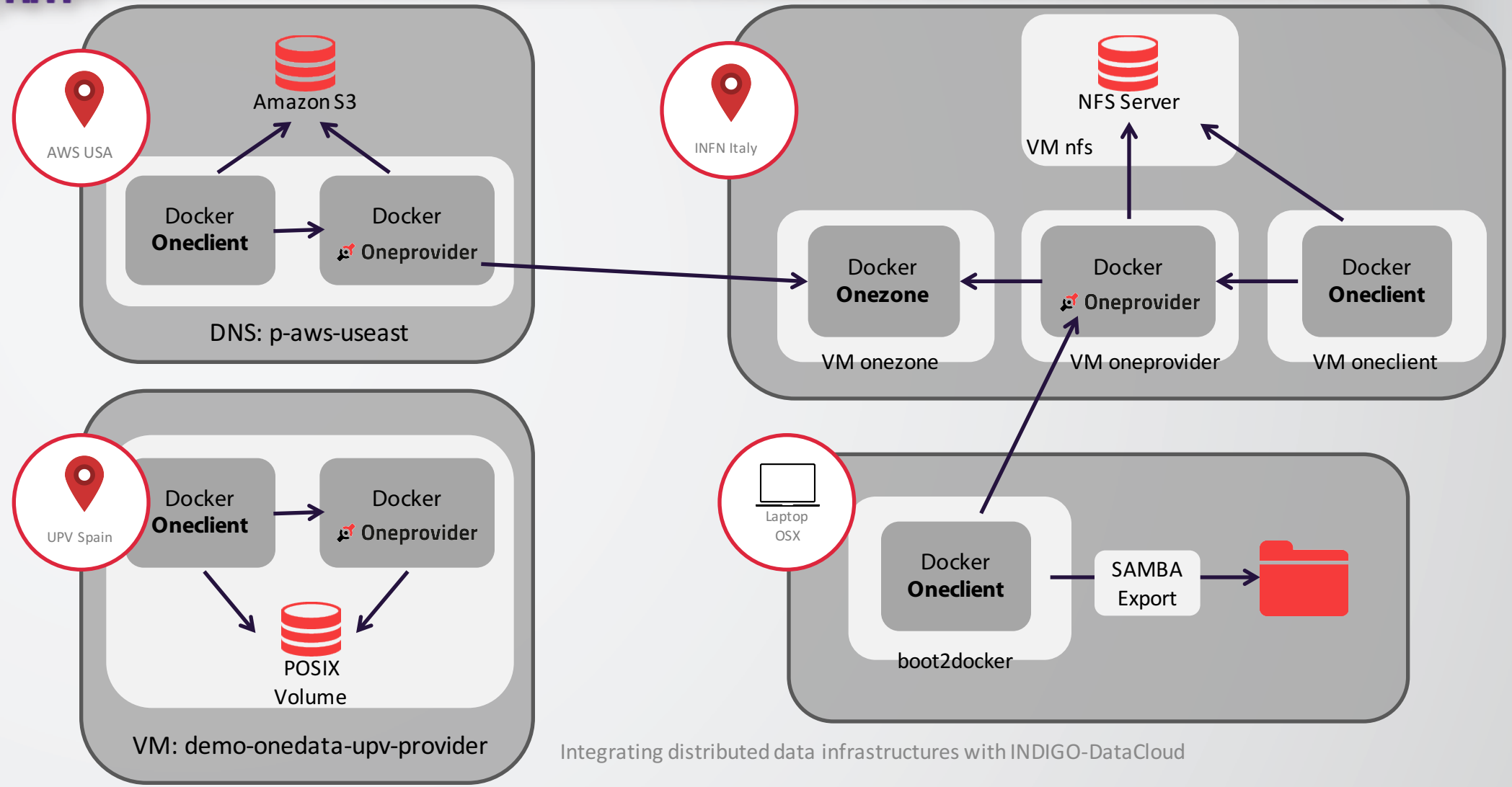
Data Management: Onedata



Data Management: Onedata



Data Federation



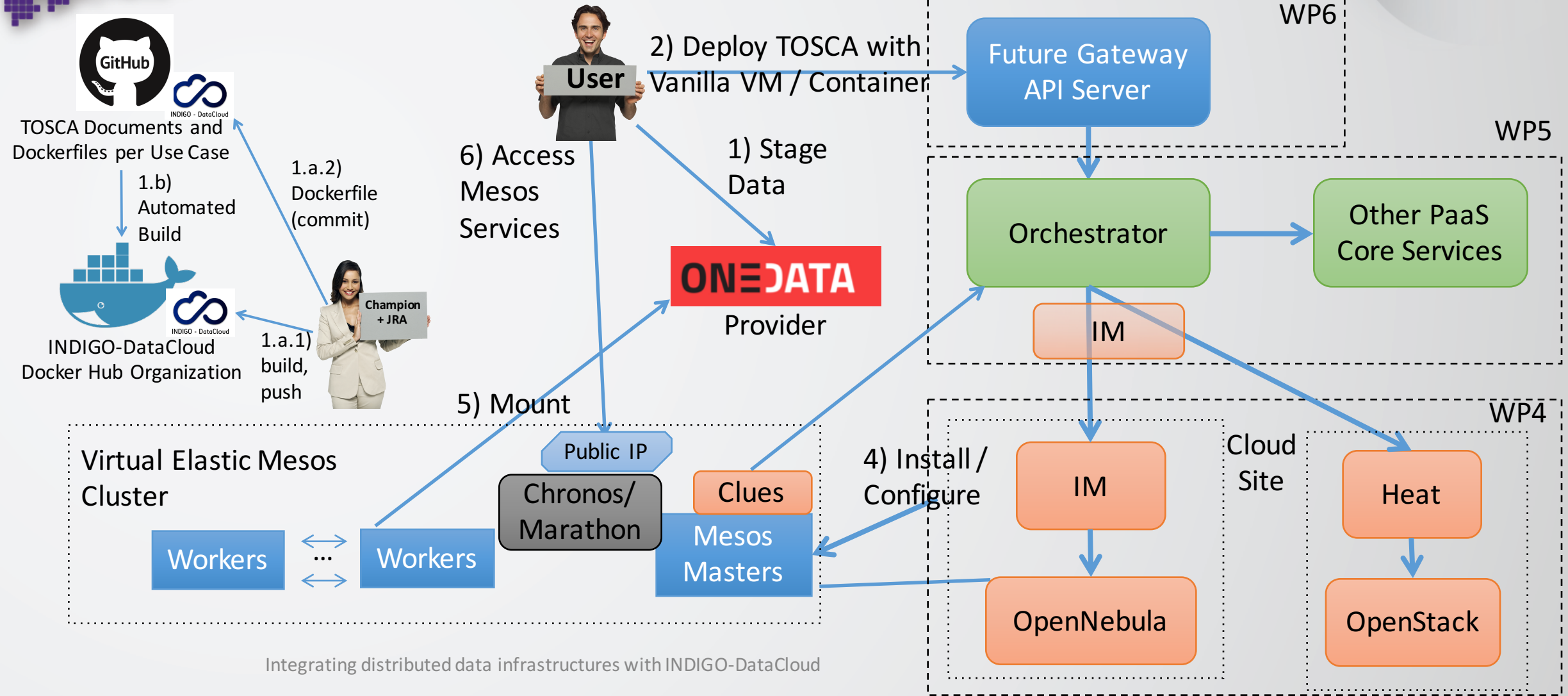
Data Management: Onedata

- Access tokens based on long running macaroons
- Support for POSIX, S3 and Ceph storages
- Provides CDMI and POSIX access to the data
- Support for Zones
- Fully dockerized
- Batch configuration and depolyment

A possible Phenomenal-INDIGO integration scenario

- Phenomenal already rely on a very rich set-up exploiting Mesos
- INDIGO is able to provide a customizable environment where an a priori complex cluster could be deployed in an automatic way:
  - Using a specific TOSCA Template build with the expertise of the INDIGO PaaS developers
- INDIGO could provide to Phenomenal:
  - (Automatic) Resource provisioning exploiting any kind of cloud environment (private or public)
    - Reacting on the monitoring the status of the services istantiated
  - Advanced and flexible AAI solution
  - Advanced and flexible data management solution
  - Advanced scheduling across many cloud provider based on:
    - SLA/QoS, Data location, availability monitoring and ranked with highly flexible rules
  - Easy to use web interface both for the end users and for the services admin/developers

Phenomenal exploiting INDIGO



Exploitation (1)

- There are currently *many* activities going on with communication and collaboration with external entities, such as:
  - RDA (Research Data Alliance), specifically with the QoS and Data Lifecycle WG and Database WG.
  - SNIA (Storage Networking Industry Association).
  - National Research Networks such as GARR.
  - EC Projects such as Thor, PhenoMeNal, Beacon.
  - IBM (on TOSCA), Yahoo! (on spot instances).
  - OpenNebula (containers).
  - OpenStack (at several levels, including the newly formed OpenStack Scientific WG).
  - Yandex.
  - ESA.
  - Dissemination at the EC level (e.g. in preparation for the ICRI2016 Conference, October 2016)
  - Collaborations with non-EC institutions such as Nectar in Australia, Ohio State University and Lawrence Livermore National Lab in the US.
  - EGI Communities not directly participating to INDIGO.

Exploitation (2)

- At the INDIGO AHM (4-6/5/2016): a meeting dedicated to describing strategy and exploitation plans of the INDIGO software by the 4 INDIGO industrial partners.
- Followed by a meeting dedicated to possible exploitation of INDIGO solutions in the ESA/Copernicus context (both with ESA and with external industrial partners).
- This week, meetings with:
  - GARR & some GARR partners (follow-up from the GARR Workshop).
  - Phenomenal (EU project on data processing and analysis pipelines for molecular phenotype data generated by metabolomics applications).
- **Today and tomorrow**, some demos at the INFN CCR workshop.
- We will organize dedicated training / dissemination sessions or workshops in the future.
  - Videos are already planned (and some of them are already available, see for example https://www.youtube.com/watch?v=sEDBZFZjrvE, https://www.youtube.com/watch?v=UtbFAhvQZ40).

Some Use Cases (from the AHM)

The HEP Contest: End-User Data Analysis

- Within HEP-Workloads is worth to distinguish among distinct types. In this contest we focus on End-User Data Analysis related activities

- LHC users (physicists) usually perform three main steps analysing detector data
  1. Check content of small/medium data-sets to develop analysis strategy
  2. Run analysis algos over large sized data-set to skim the whole available statistic
  3. Process previously skimmed data for further reduction/selection/data visualization (iteratively)
- Notes:
  - Executed code (a job) could be either simple scripts or a more complex payload built over the 'official experiment framework'
  - Data can be local o remote wrt computational resource
  - Data (even the skimmed one) are meant to be shared/accessed by a group of collaborators

More concretely: The workflow to enable

- Being able to instantiate a cluster on demand, where to run
  - ❑ HEP analysis applications: Experiment software plus user specific plugin
  - ❑ 'Just' user scripts
- Depending on the step of the analysis flow, the cluster might be used to
  - ❑ Run a batch system
  - ❑ Join the Experiment pool (fetching workloads, aka jobs)
- **A key point is the data access**
  - ❑ We want to read local & remote data coming from the LHC detectors
- But also provide support to data sharing:
  - ❑ Reduced data from a previous run must be available to collaborators
  - ❑ Working group can be spread around the world
- Finally:
  - ❑ The aim **is to transparently span over multiple IaaS** to maximize the computational power

WP2 use case analysis (WP2 + high-level experts)

Wednesday 14:15-16:00 & 16:30-18:30

EUROBIOIMAGING: https://project.indigo-datacloud.eu/attachments/902/download

LIFEWATCH: https://project.indigo-datacloud.eu/attachments/887/download

LBT: https://project.indigo-datacloud.eu/attachments/896/download

CTA-Archive: https://project.indigo-datacloud.eu/attachments/1052/2Slides_INAF-CTA.pdf

ELIXIR: https://project.indigo-datacloud.eu/attachments/886/download

update: https://project.indigo-datacloud.eu/attachments/1019/download

EMSO: https://project.indigo-datacloud.eu/attachments/983/download

CMCC/ENES: https://project.indigo-datacloud.eu/attachments/1076/INDIGO_WP2_CMCC-ENES.pptx

WeNMR:
https://project.indigo-datacloud.eu/attachments/995/download
https://project.indigo-datacloud.eu/attachments/998/download
https://project.indigo-datacloud.eu/attachments/1001/download

DARIAH: https://project.indigo-datacloud.eu/attachments/1028/download

EGI FedCoud: https://project.indigo-datacloud.eu/attachments/1034/download

From Gleb (provisional link):

https://drive.google.com/file/d/0B8XsRtHRPjWxdnpXZEpHOVcyR2s/view?usp=drive_web

OTHER COMMUNITIES (HEP):

From Isabel on HEP Pheno:

https://project.indigo-datacloud.eu/attachments/1031/download

From Daniele on HEP and quantum chemistry:

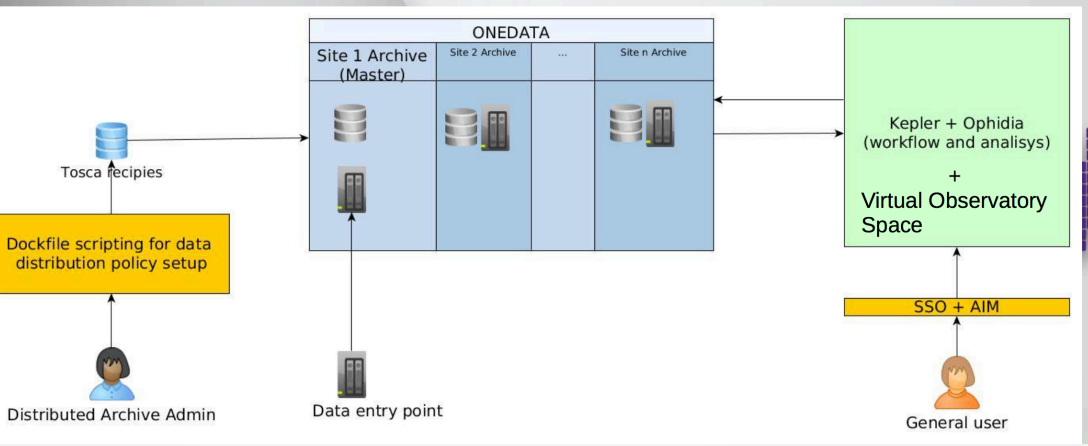
https://project.indigo-datacloud.eu/attachments/1058/WP2-spiga-hep.pdf

Defining WP2 applications in INDIGO

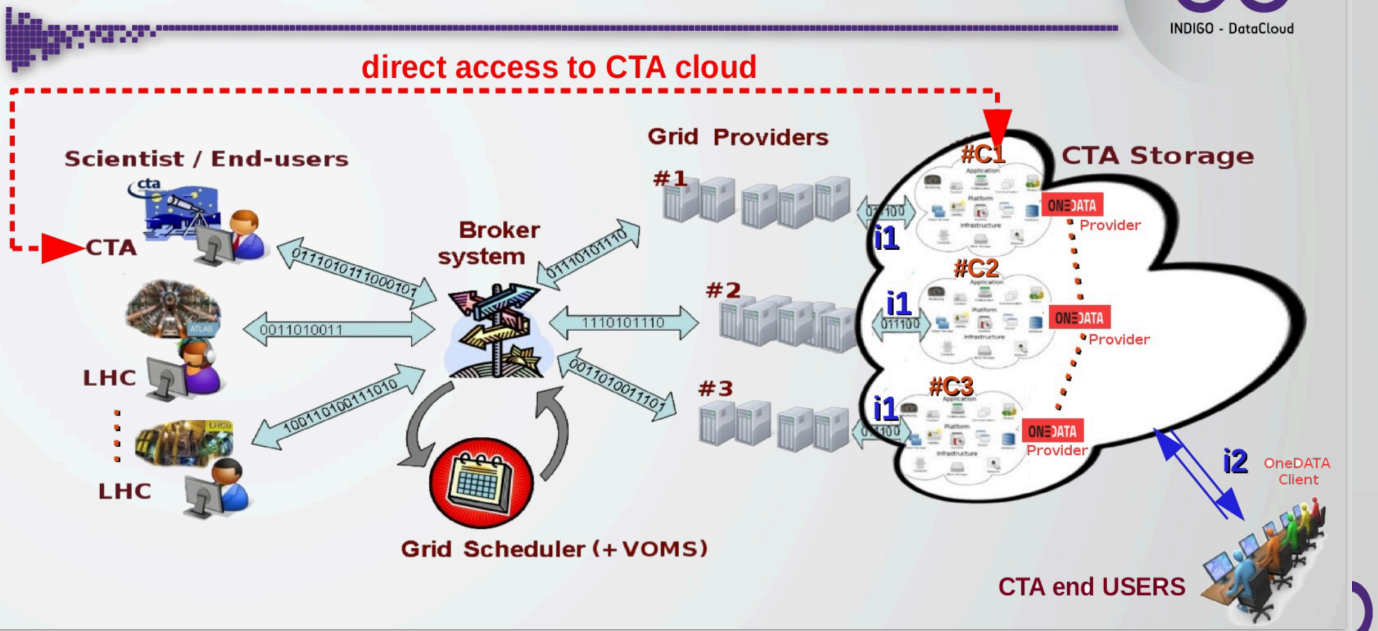
Defining_WP2_applications_in_INDIGO_v3.pptx

AAI WP2 Perspective

LBT Distributed Archive use case (1)
-Distribution over federated archives-

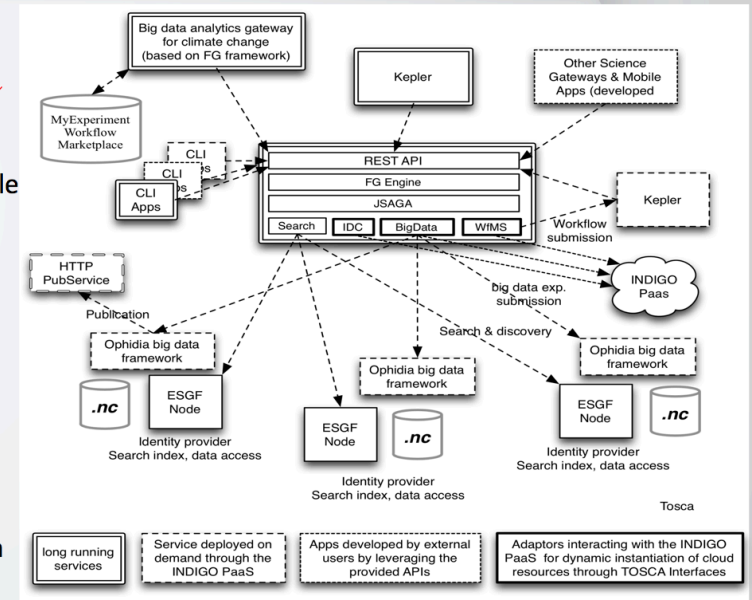


C.T.A. Distr. Archive proposed prototype!



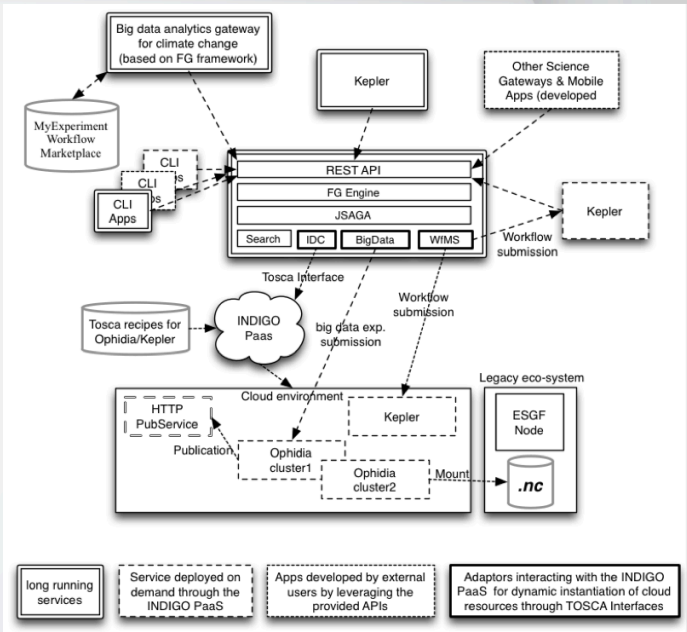
Use case 1: CMIP* distributed analytics use case (I)

- Distributed exp. for climate data analysis
- Two-level workflow strategy to orchestrate large scale experiments
- Interoperability with ESGF is a mandatory requirements
- Access through different clients
  - Gateway
  - Kepler
  - CLI
- Interactive and batch scenarios



More use cases

- Dynamic instantiation of an Ophidia cluster and Kepler WfMS
- Automated deployment through IM/TOSCA interfaces
- Data locality aspects will be considered as the analysis will run "close" to the data
- Interoperability aspects with ESGF to be taken into account to ensure a seamless integration of the INDIGO components into the research community eco-system



Project Status

DEMOS

- *Mercoledì* primo Coffee Break
  - **AAI** (Andrea Ceccanti/Presented by Daniele Spiga)
  - **Mesos Chronos Autoscaling** (Marica Antonacci)
- *Mercoledì* secondo Coffee Break
  - **Onedata** (Lukasz Dutka)
- *Giovedì* primo Coffee Break
  - **Mesos - Jenkins** (Cristina Aiftimiei)
  - **Synergy** (Presented by Lisa Zangrando)
- *Giovedì* secondo Coffee Break
  - **AAI** (Andrea Ceccanti/Daniele Spiga)
  - **Onedata** (Lukasz Dutka)