DODAS: How to effectively exploit heterogeneous clouds for scientific computations

Daniele Spiga INFN (spiga@infn.it)

on behalf of DODAS Team

eosc-hub.eu

@EOSC_eu

EOSC-hub receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 777536.

Outline

- Introduction to DODAS
  - Scientific context and Motivations
- DODAS and INDIGO DataCloud project
  - Design and principles
- Architecture
- DODAS and Compact Muon Solenoid Experiment
  - Results
- **DODAS a Thematic Service in the context of EOSC-hub project**
  - The Alpha Magnetic Spectrometer Experiment

DODAS from 10 Km

Dynamic On Demand Analysis Service: DODAS

- A service for **generating over cloud resources an on-demand, container based solution to:**

- **Instantiate a standalone HTCondor batch system**
- **Seamlessly integrate an existing HTCondor pool**

- **Instantiate cluster Big Data processing : Spark, Hadoop**

Allows to utilize "any cloud provider" with almost zero requirements

HEP Community and Compact Muon Solenoid experiment

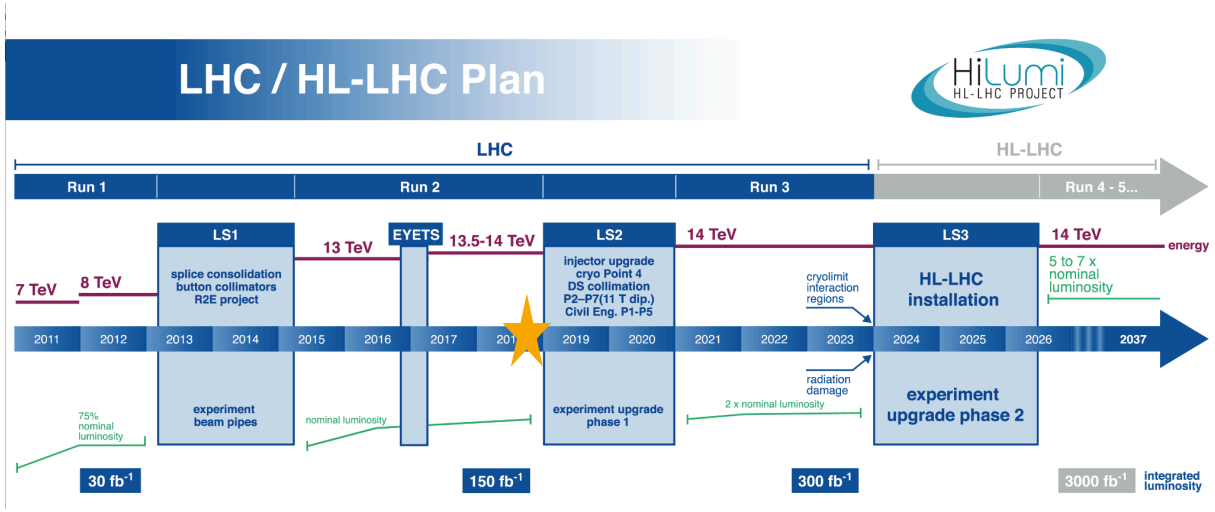
LHC Luminosity increases

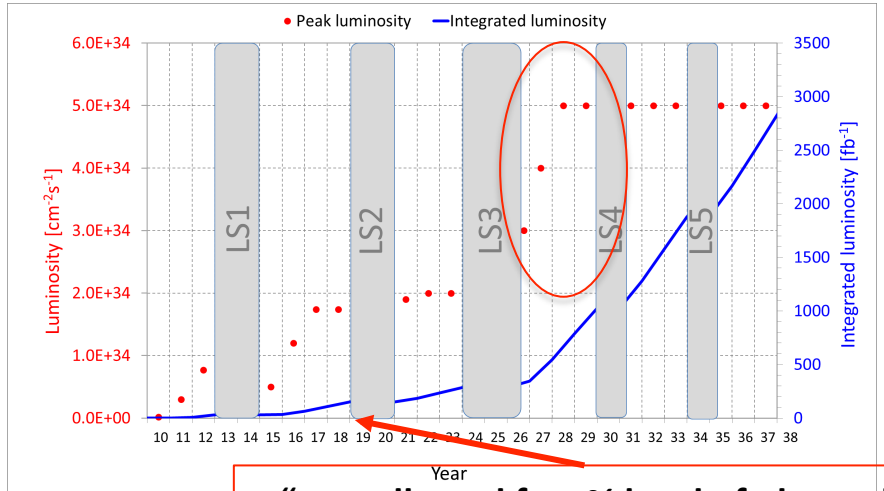
- Trigger rate on 2026+ = up to 10 kHz
- Moreover, 2026+: new detectors
  - More DAQ channels, Bigger events

Expect a factor 3x10x2 ~ 50 (?) of complexity →huge computing challenge

- Experiments are looking at combining several solutions
  - Software, Event model, **infrastructure.**

→ Need to be able and use _whatever_ we are offered as computing resource





"we collected few % level of planned LHC data"

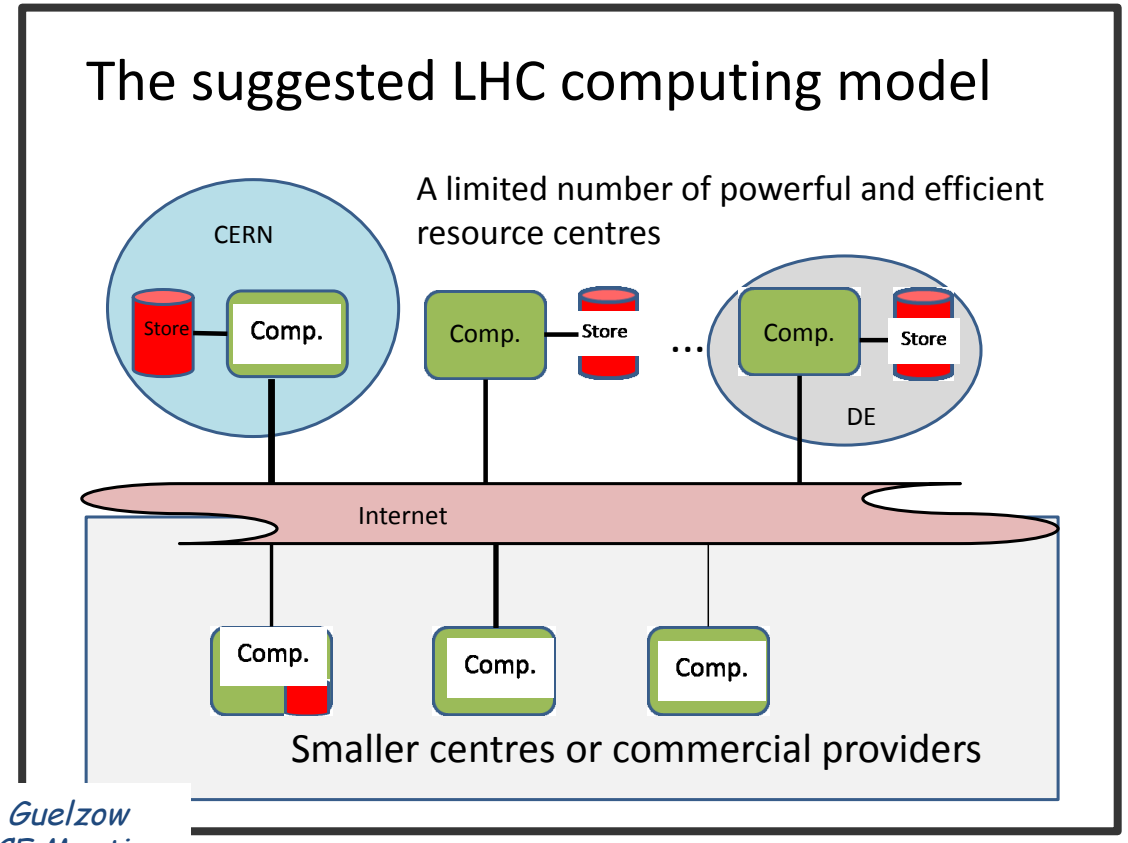
A Service for HEP data processing

In this context, DODAS has been developed to allow

- ***Opportunistic resources integration***
  - Targeting both private and pubblic cloud providers
- ***Extension of already existing facility***
  - Both to absorb peaks of usage and to generate mission specific facilities
- ***User-friendly computing infrastructure generation and management***
  - of computing cluster to analyse research data

Look to the future:

- ***DODAS design compliant with future paradigms***



V. Guelzow
SCF Meeting
12/5/2017

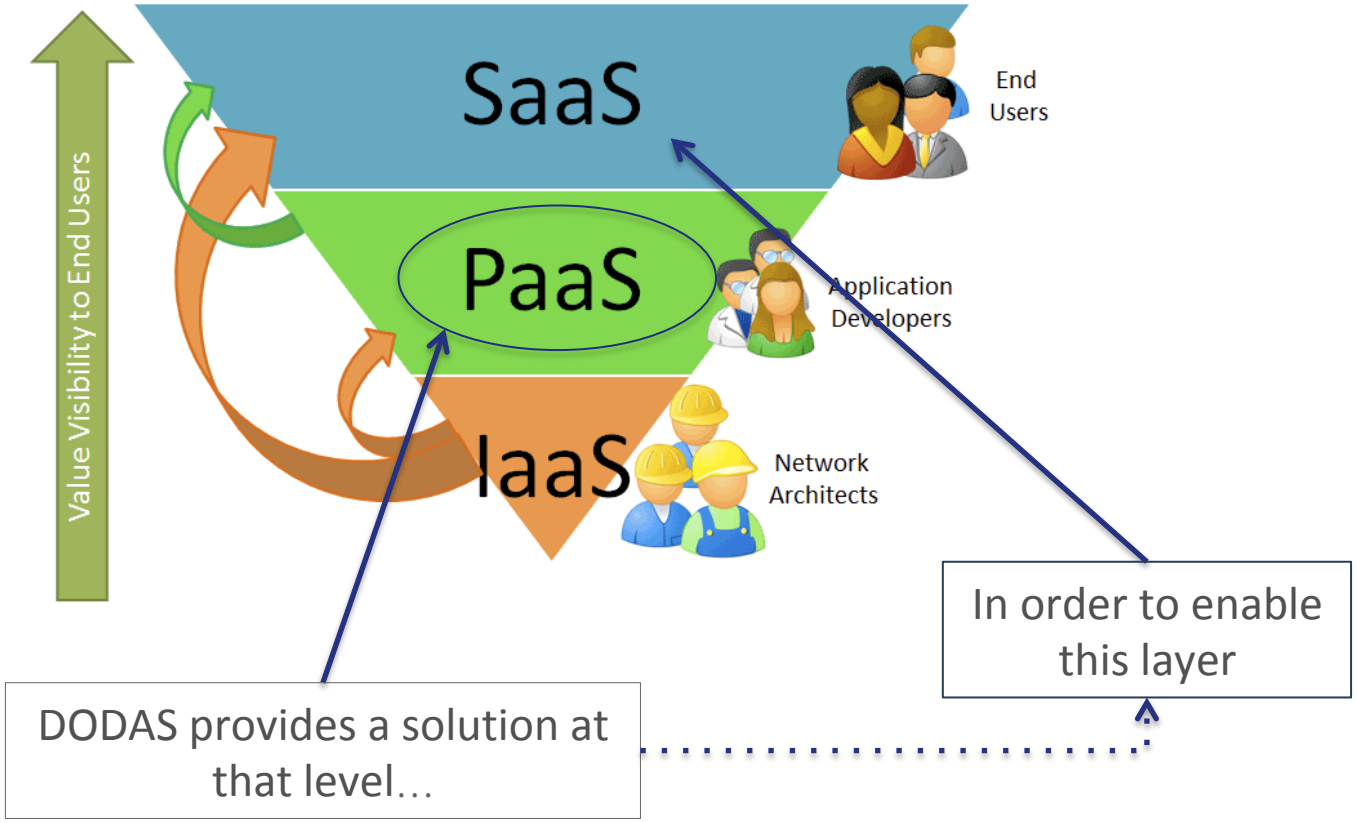
The Genesis of DODAS: INDIGO-DataCloud

INDIGO - DataCloud
Better Software for Better Science

What matters at the end… are the applications.

The INDIGO Main Objective:

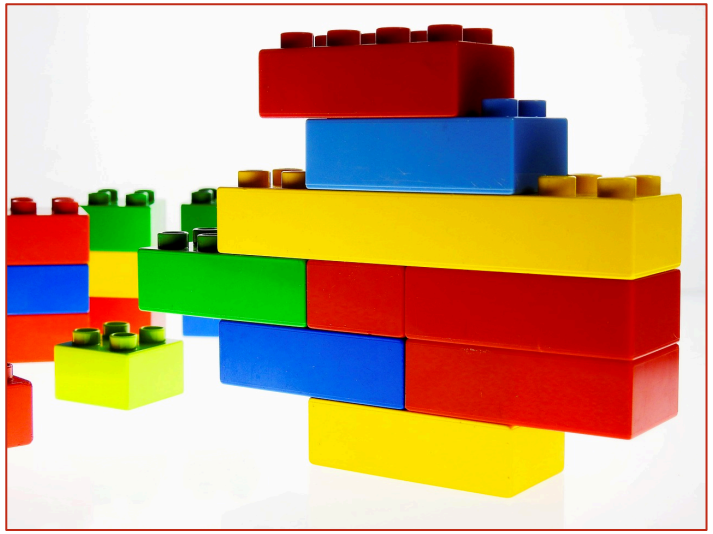
- To **develop software components and solutions** to facilitate (or simply make possible) the **exploitation of distributed cloud and storage resources** through **public or private infrastructures**.
  - Tailored to science and targeting **multi-disciplinary scientific communities**



Building services using INDIGO's Brick Box

There is a huge set of tools and solutions available, but **there is NOT a one-size-fit-all solution**

- Open, Standard-based, flexible and extensible **building blocks**
  - Each use case can compose and customize



The rationale behind DODAS implementation

Allow the setup of complex services and application as friendly as creating a single Virtual Machine.

- Batch systems, databases, Big Data platforms etc etc…

- Automate setup of **intricate workflows with intricate configurations**
  - Dynamic Cluster with multiple software dependences, interactions with experiment specific internal/external services..

- **Abstract underlying infrastructures**
  - **To avoid learning Cloud APIs** for any IaaS to exploit, as in the case of **Hybrid Cloud model**

DODAS Architectural Pillars

Abstraction

Definition of resource requirements via extensible **TOSCA** templates

Full independence from specific Cloud infrastructures **[Infastructure Manager]**

Automation

Automated software and service configuration via **Ansible** roles

Automated horizontal scalability of the instantiated resources **[Clues]**

Self-healing

Self-healing for applications and services **[Mesos/Marathon]**

DODAS Architectural Pillars (cont)

Multi-cloud support

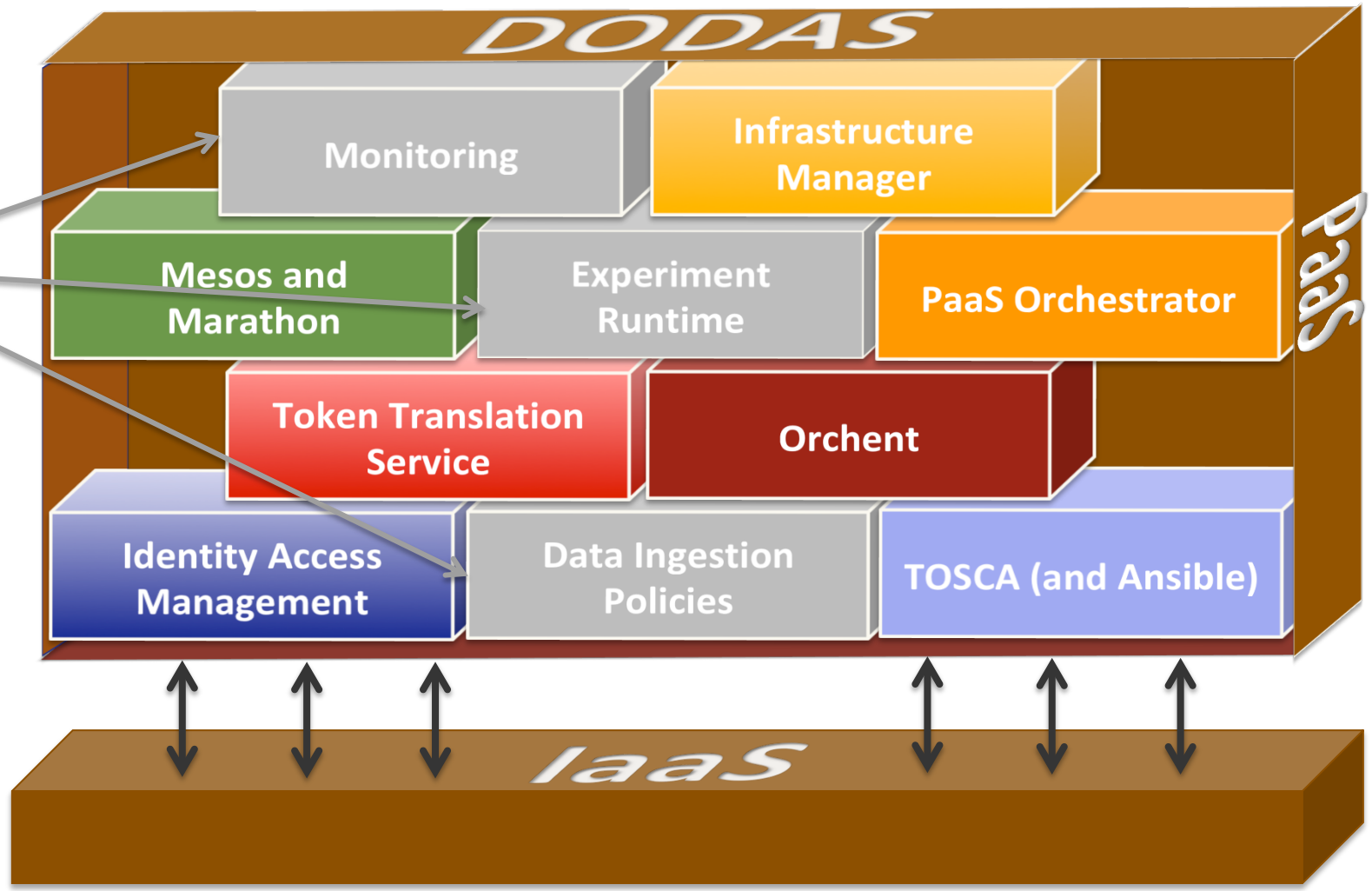
Automatically address multiple and heterogeneous Cloud infrastructures **[PaaS Orchestrator]**

Flexible and federated authentication mechanism

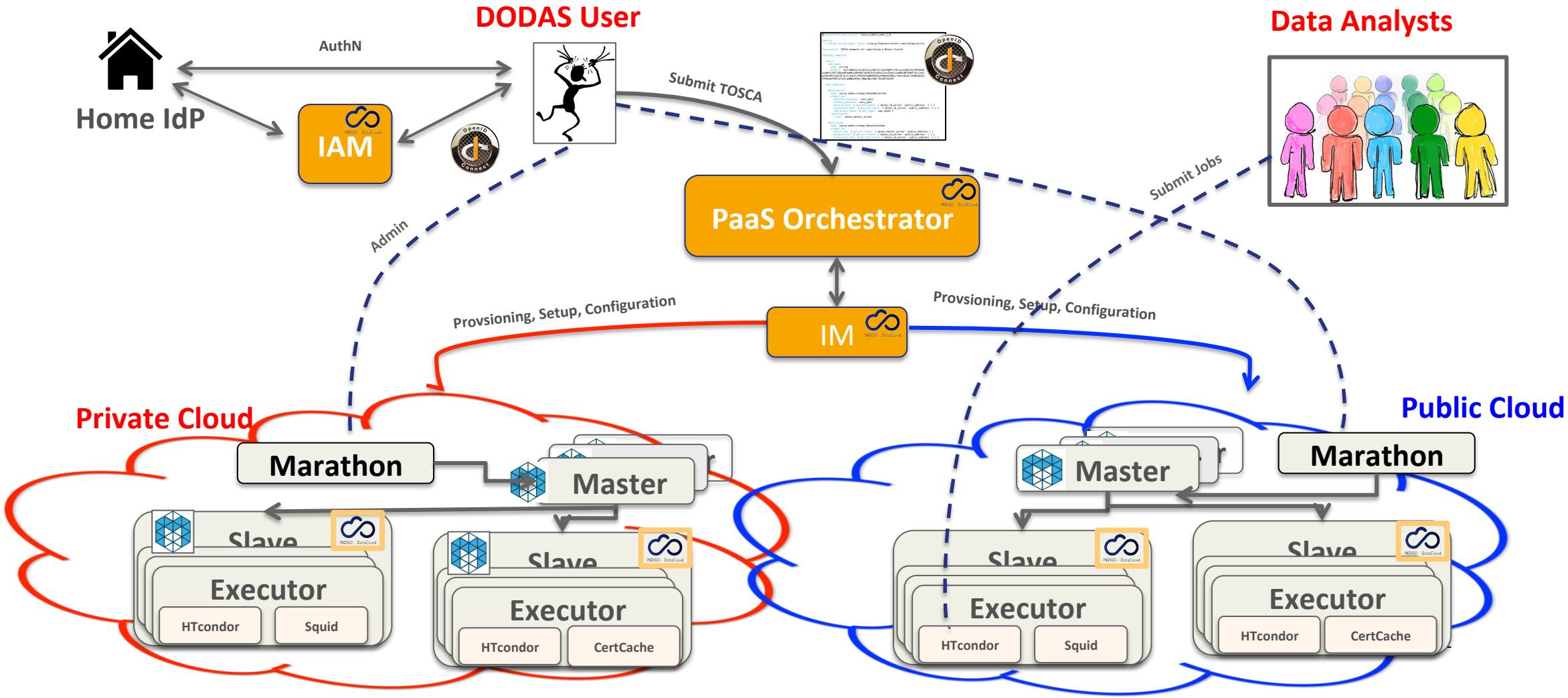
- Support to multiple authentication methods **[Identity Access Management]**
  - OpenID Connect, SAML2.0, LDAP, Local (Username/Passwd)
- Identity harmonization
  - Guarantees a single identity irrespective of authentication method
- Support for services relying on heterogeneous (and legacy) AuthN mechanisms

DODAS Lego Architecture

Specific DODAS Extension Integrated with existing building blocks



A insider look at DODAS



A key aspect: this is not a community specific Service

- Everything is ***experiment agnostic***
  - Flexible enough to support multiple and diverse use cases
- **Highly Customizable:** By design E-PaaS provides **three major handles**

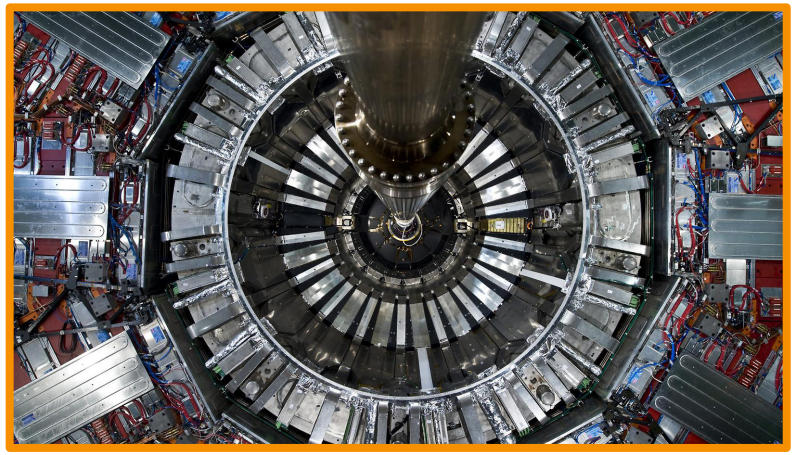
To support user tailored computing environments

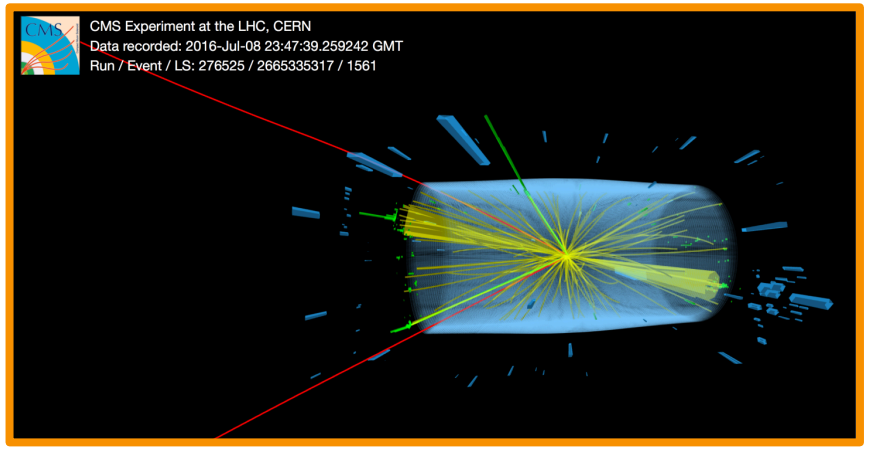
To automate configuration and deployment of custom services and/or dependencies

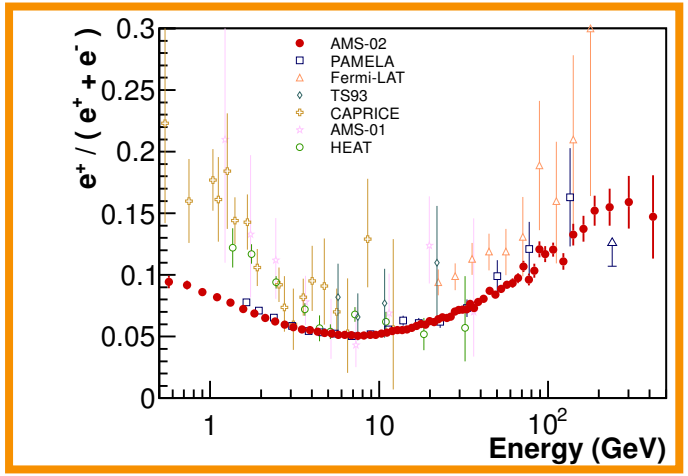
To define input parameters and customize the workflow execution

→ All of this embedded in a human readable YAML file

Currently Supported Communities

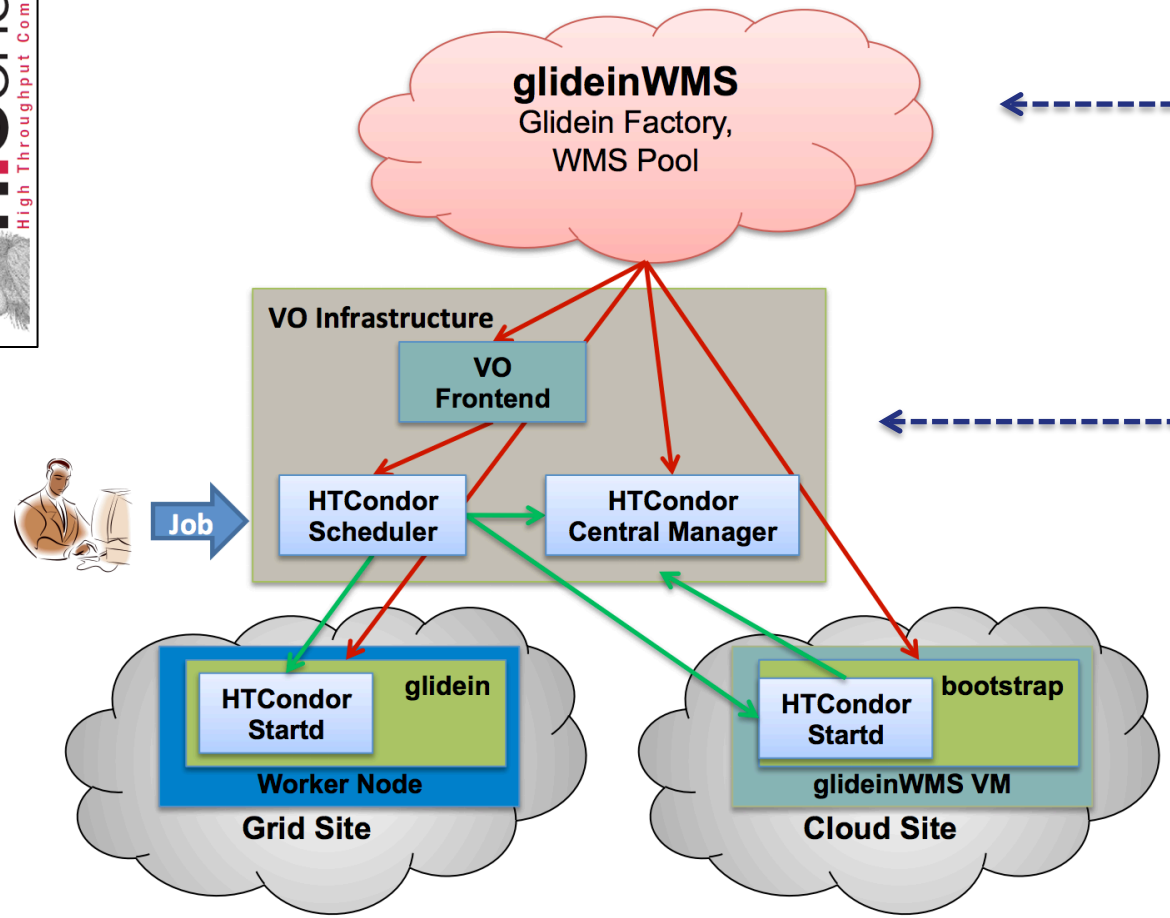






CMS Job Submission Infrastructure

HTCondor Global Pool



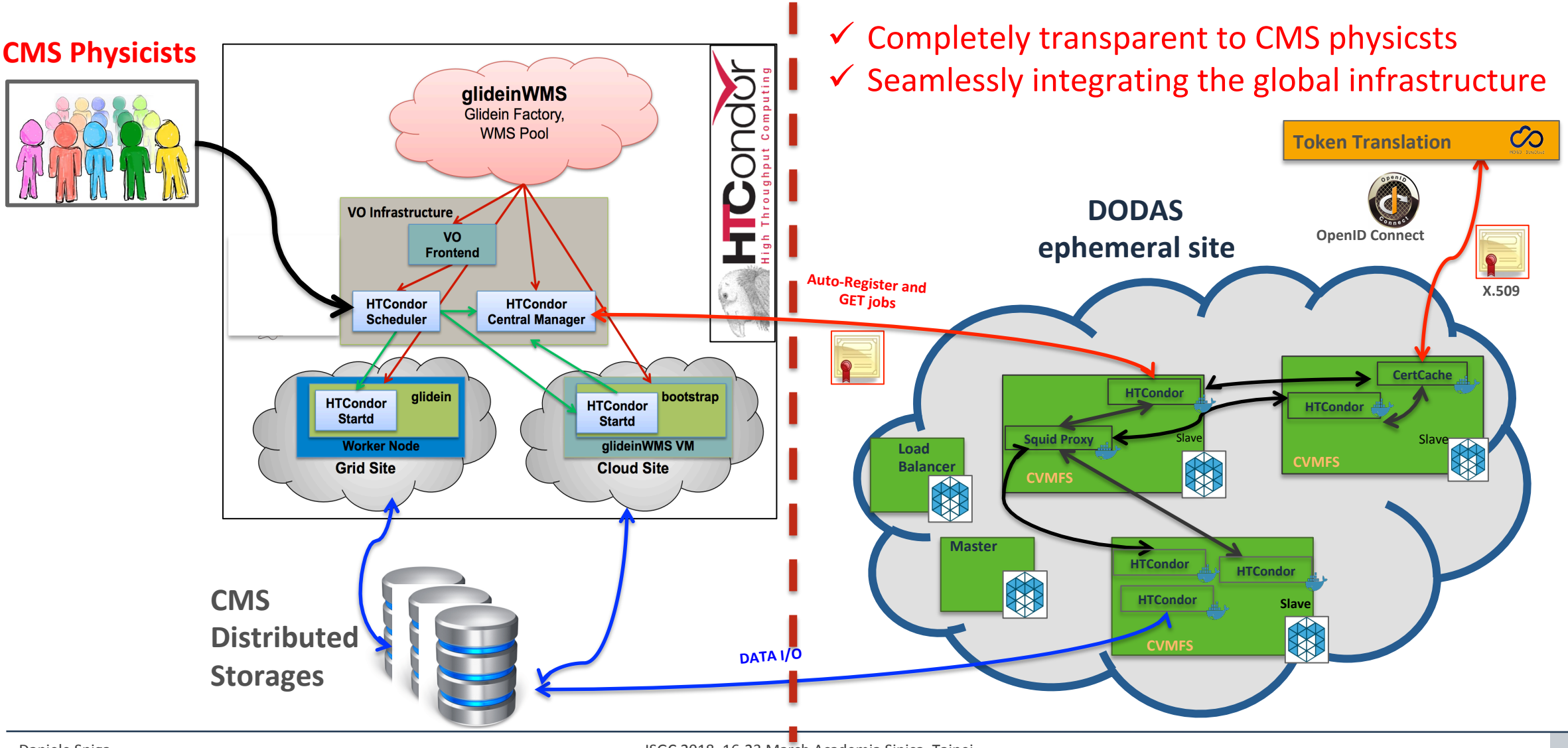
- In the first stage of matchmaking, glideinWMS frontend matches jobs to their desired sites and requests the glideinWMS factory to send glideins (properly configured condor tar ball)

- The 2nd stage of matchmaking is when a job gets matched to a slot once the condor starts on the worker node and makes itself available in the pool

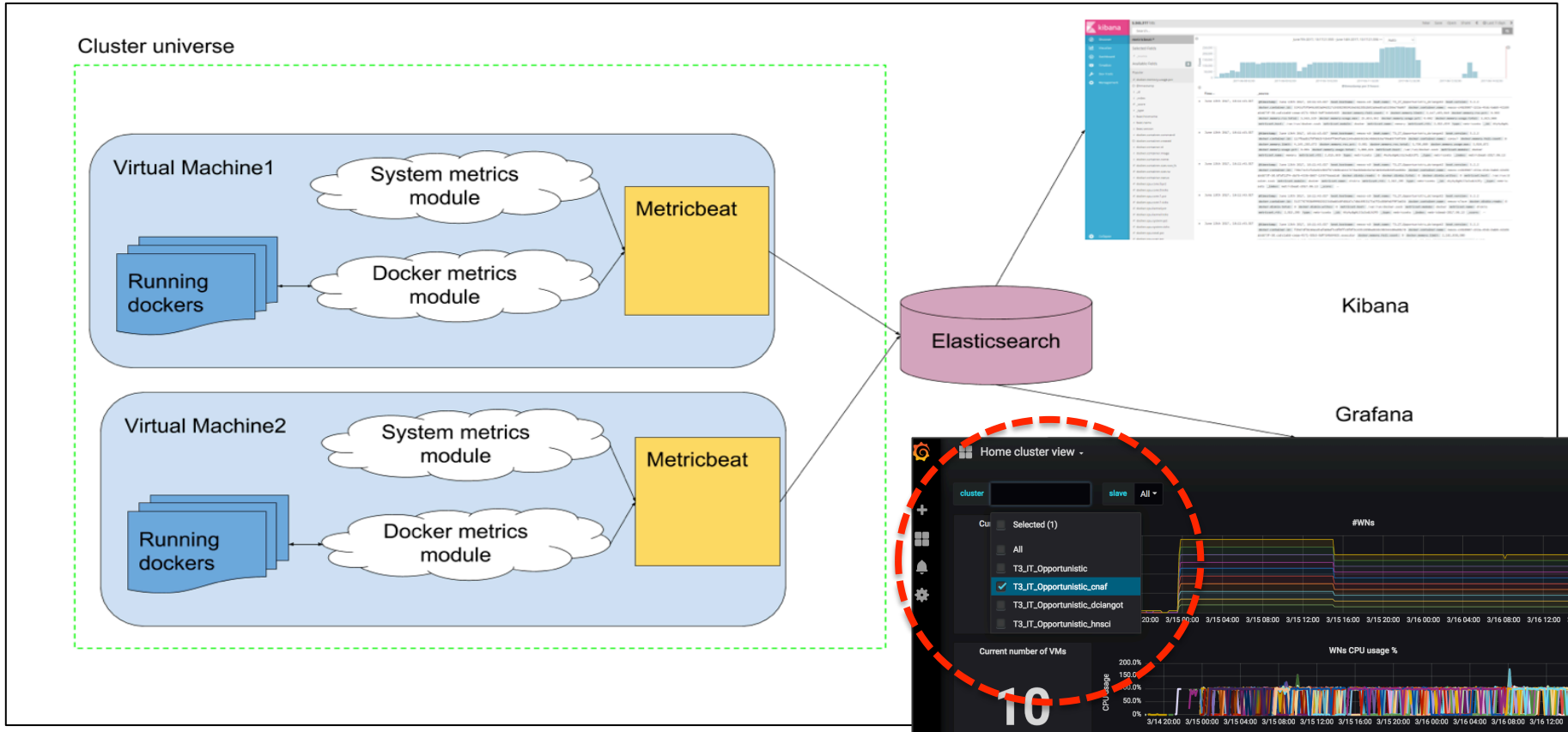
- Glidein pulls in the job and then GLExec is used to switch to central production or analysis user's credentials

DODAS in the CMS Global Pool

- ✓ Completely transparent to CMS physicsts
- ✓ Seamlessly integrating the global infrastructure



Monitoring System

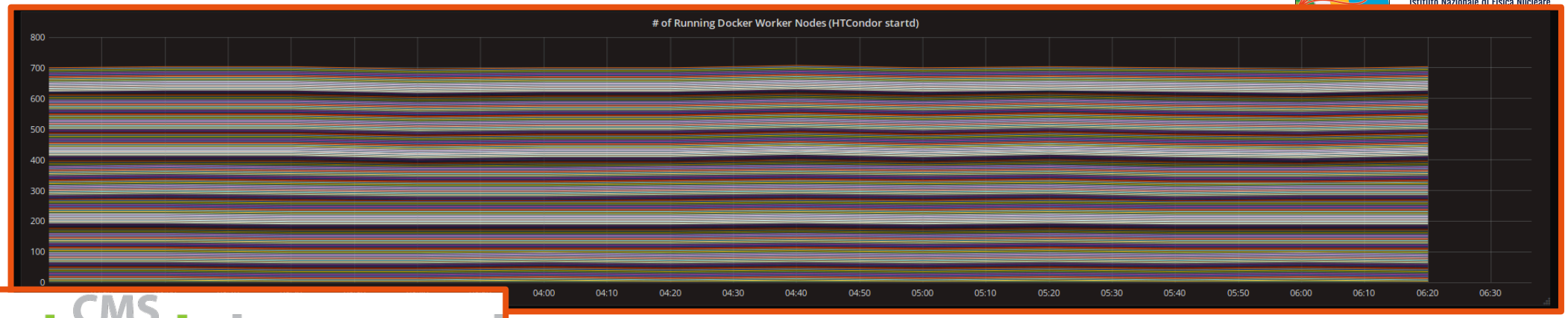


Support both

- **Central** Elasticsearch and kibana/grafana
- **Per Cluster** automated ES setup

Some results

A Scale Test:
3 TB RAM
1.5K Cores
30 TB Disc



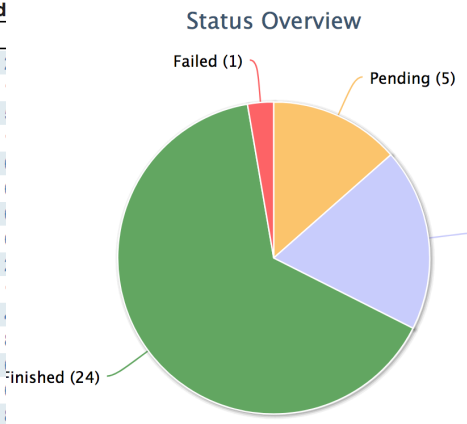
CMS dashboard

Submit

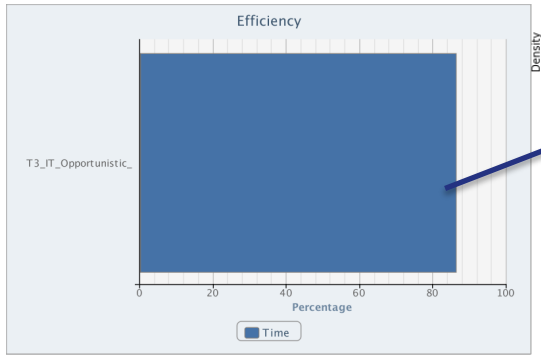
Start » Jobs

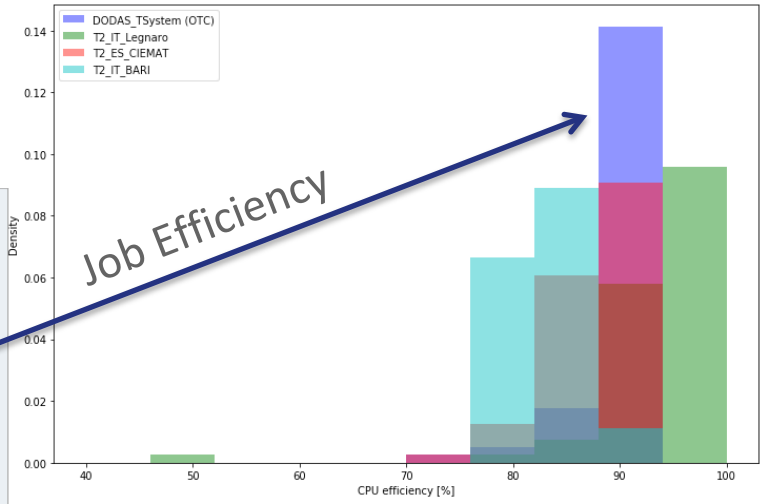
Table | Chart | Show

name	Current Status					Grid
	Sub	Pend	Run	Term	Done	Canc
T3_US_TACC	3280	0	0	3280	377	0
T3_US_NERSC	2029	0	0	2029	934	0
T3_IT_Opportunistic	580	0	0	580	9	22
T2_RU_PNPI	187	0	0	187	0	1
datagrid.cea.fr	1	0	0	1	1	0
T2_TR_METU	4	4	0	0	0	0
T2_ES_IFCA	15	15	0	0	0	0
T3_UK_London_RHUL	219	213	0	6	0	0
T2_RU_ITEP	119	1	1	117	92	5
T3_US_TAMU	2987	0	3	2984	2957	0
T2_PL_Warsaw	450	0	3	447	30	8
T3_IT_Perugia	65	0	3	62	46	0
T2_IT_LegnaroTest	189	0	9	180	167	13
T3_GR_IASA	78	4	9	65	65	0
T3_US_NotreDame	297	7	9	281	236	37
T3_IN_TIFRCloud	590	1	14	575	566	0
T2_PK_NCP	111	4	17	90	87	0
T2_UK_SGrid_Bristol	135	7	17	111	82	0
T3_US_Rutgers	1742	2	22	1718	13	0
T2_CN_Beijing	520	2	22	496	346	3
T3_TW_NCU	11865	8	28	11829	758	1
T3_US_UMiss	5086	0	35	5051	367	0
T3_IT_Opportunistic_C	96	0	35	61	48	0
T2_FI_HIP	549	0	48	501	373	5



No overhead introduced by DODAS architecture!





Success Stories so far

The project starts **in June 2016**: The CMS use case in the context of INDIGO

- **Spring 2017** selected as solution to generate CMS ephemeral site using a 20k$ Microsoft Azure Grant
  - [See: **Dynamic extension of INFN-CNAF Tier1 Data Center]**
- **Mid 2017** DODAS in Helix Nebula project [will run for the whole 2018]
  - Extensively used on TSystem IaaS Provider

A.Falabella et al

- **Summer 2017 selected as Thematic Services in EOSC-hub project**
  - In this context has been prototyped the integration with AMS computing workflows

[See: **Harvesting dispersed computational resources with Openstack: a Cloud infrastructure for the Computational Science community**

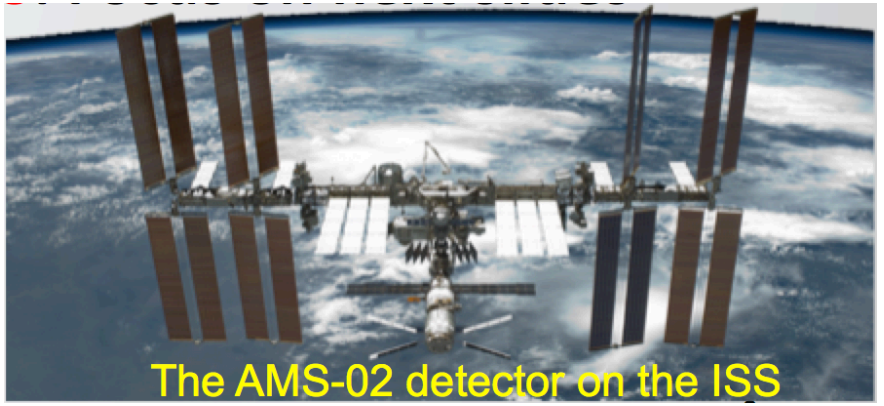
M. Mariotti et al

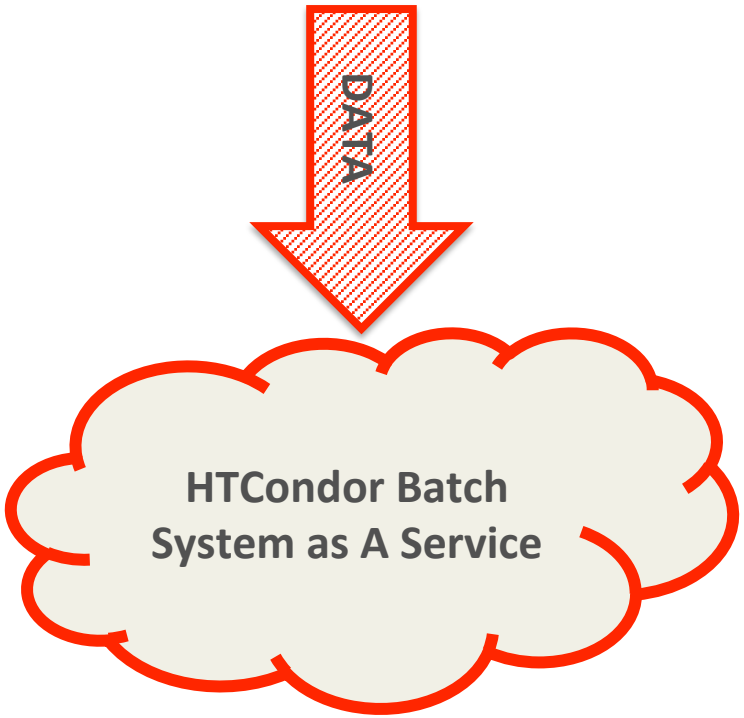
EOSC-hub DODAS Thematic Service : AMS integration

- AMS computing model **has not been** originally **designed to cope with the cloud computing paradigm**
- AMS software environment is **not highly portable**
- **Cost** for deploying AMS environment on "new" **resources is too high**
  - Setting-up environment requires huge effort wrt obtained gain

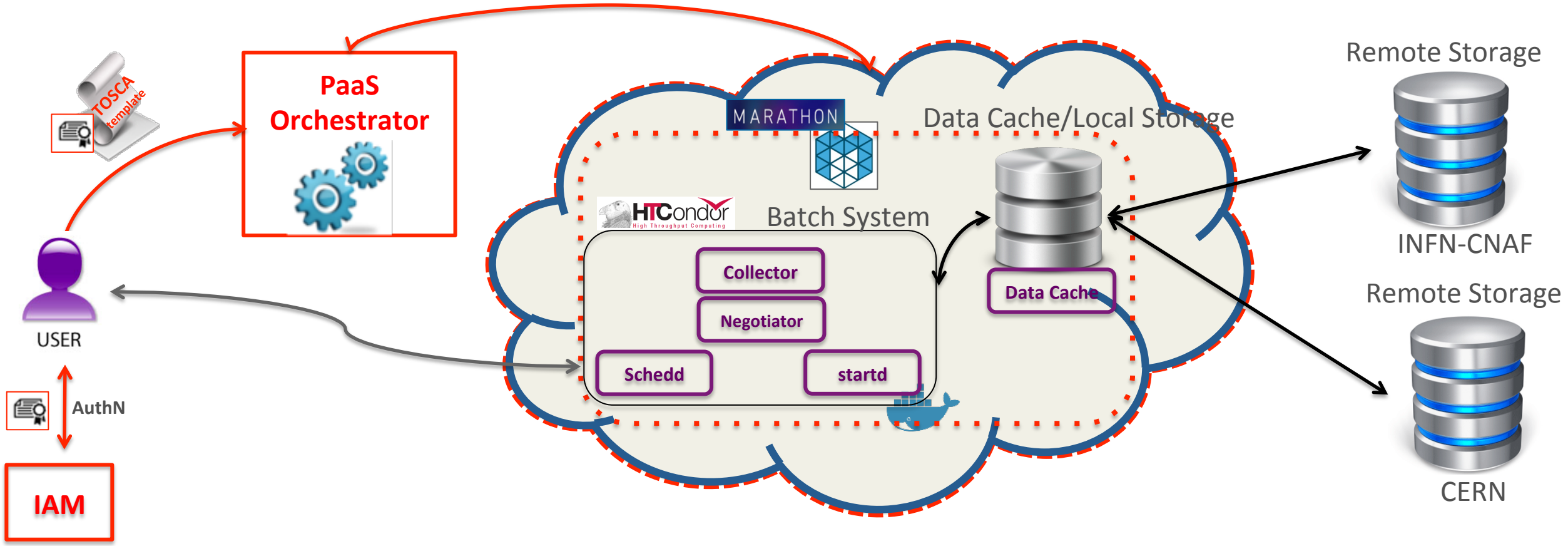
All this appears to be a limiting factor in the scope of the data analyst daily work

- it does not result trivial to exploit modern computing infrastructures such as cloud IaaS





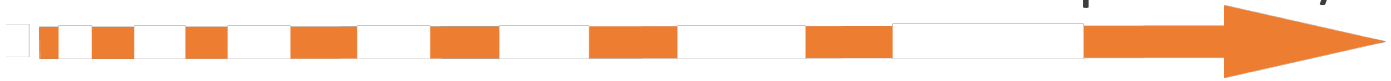
Architectural Schema



The Future

- Thematic Service under EOSC-hub
  - **Attract new scientific communities** seeking to easily exploit Clouds
    - For Instantiate batch system and Machine Learning frameworks
  - **Integrate EOSC-hub service** (EOSC-hub portfolio services)
    - E.g. Accounting, Monitoring etc etc…

- Extend and integrate **data ingestion solutions**
  - Dynamic on-demand smart data cache to optimize I/O



eXtreme DataCloud

Data Management for extreme scale computing

- **Extend the BigData support** e.g. adding more frameworks and libraries

Conclusions

- "Dynamic On-Demand Analysis Service" **seamless integrated into existing HTCondor Global Pool of CMS**
  - Currently in production phase
- **Thematic Service** under the EOSC-hub Project
  - AMS prototype has been prepared
  - http://www.eosc-hub.eu/
- **Successfully tested several scenarios**, both private and public providers
  - DODAS demonstrated reduction of costs through setup and operational efficiency increase
- DODAS experts are available for consultancy and training interested communities.
  - Contact spiga@infn.it

Thank you for you attention

Contact

spiga@infn.it
http://www.eosc-hub.eu/

eosc-hub.eu @EOSC_eu