
# The DODAS Thematic Service

**Daniele Spiga, spiga@infn.it**

eosc-hub.eu

@EOSC_eu

**Dissemination level**: Confidential
Disclosing Party: Project consortium
Recipient Party: European Comission

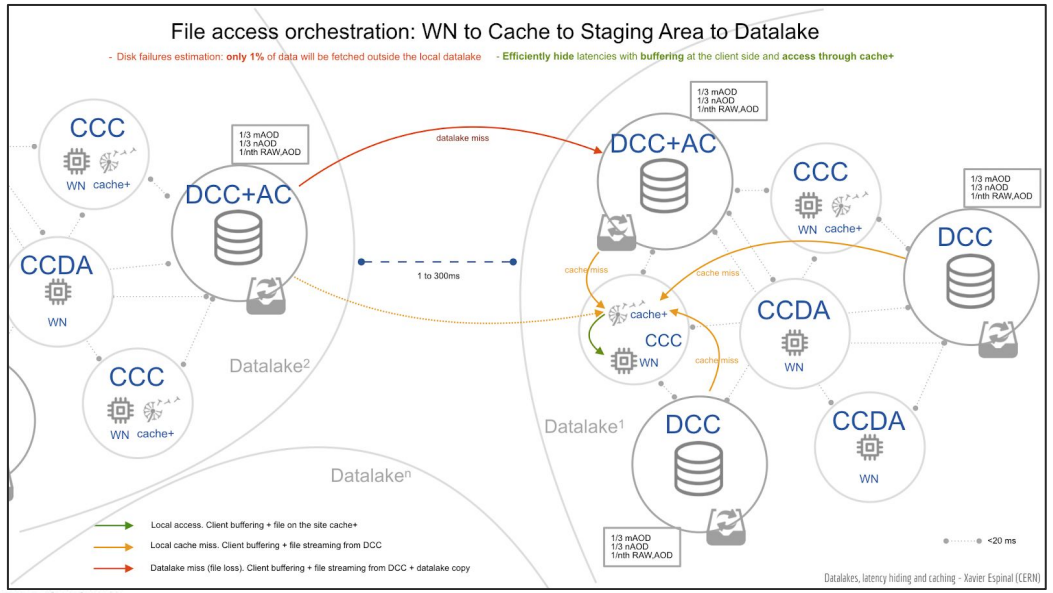
EOSC-hub receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 777536.

# Outline

- A quick overview of the the scientific case
- Technical motivation, gaps to fill
- Integration with EOSC-hub services
- Impacts:
  - On scientific communities
  - On Infrastructure providers
- Demo

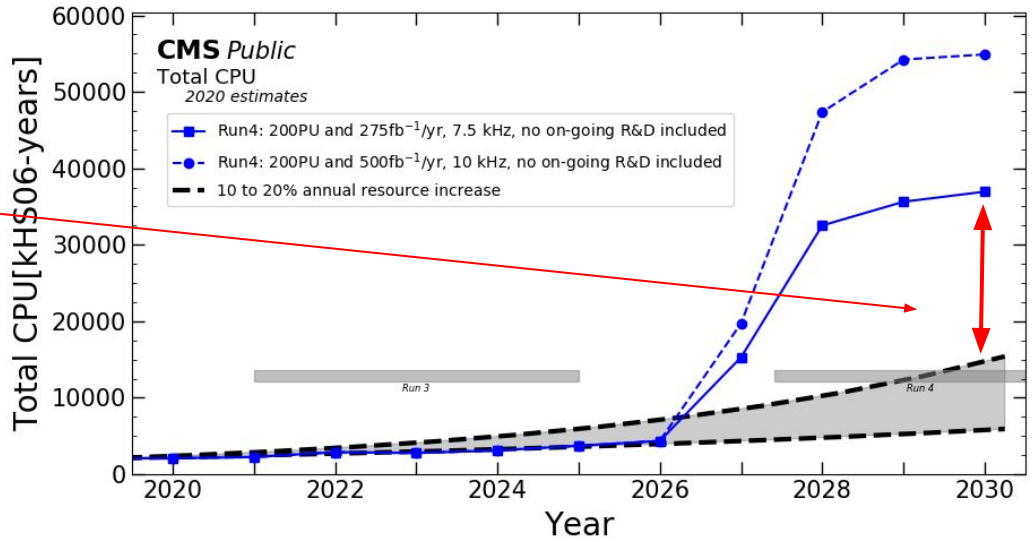
# Scientific case: High Energy Physics - The initial One

To provide new models to integrate **compute resources** for the future needs of High Energy Physics experiment:
**The Compact Muon Solenoid Experiment**




*Terminology:*

*AC - Archive center Defined as Tape or tape-equivalent-QoS enabled center able to archive custodial data.*
*DCC - Data and computer center providing disk-equivalent QoS storage*
*CCC - Compute center with cache*
*CCDA - Compute center without cache: relies on accessing all data via the network from either a CCC or a DCC*




Experiment agnostic compute solution

- **Compute Center with Cache**
- **Compute Center without Caches (diskless)**

**A suitable solution to enable data analysis on DataLakes**

From Data Lake Straw Proposal @ DOMA

# Scientific platform: Motivations

- A solution designed with the goal to enable users **to create and provision infrastructure deployments, automatically and repeatedly**, on “any cloud provider” **with almost zero effort**.

- Implement the ***Infrastructure as Code*** **paradigm**: driven by a templating engine to specify high-level requirements . Declarative approach **allows to describe “What” instead of “How”**
  - Let the underlying system to abstract providers and automatically instantiate and setup the computing system(s)

- Allows to instantiate **on-demand container-based clusters**

To lower the bar of sys-admins skills for accessing cloud resources
- automate the complete flow

To facilitate the modernization process
- brings the cloud to the scientific communities

To bridge scientific experiment data and distributed computing
- builds composable and portable clusters

# Integrations with EOSC-hub

**The DODAS lego blocks Platform**
**heavily relies on EOSC-hub portfolio of services**

## AAI:

- **INDIGO-IAM** is used to implement authentication and authorization
  - allows for a federated model ( i.e. **Egi Check-In** )
  - supports capabilities based AuthZ

## Compute Resources orchestration

- **INDIGO-PaaS Orchestrator** is used to coordinate the provisioning of virtualized compute resources on both private and public Cloud,
  - integrates the **Infrastructure Manager** to interface the clouds

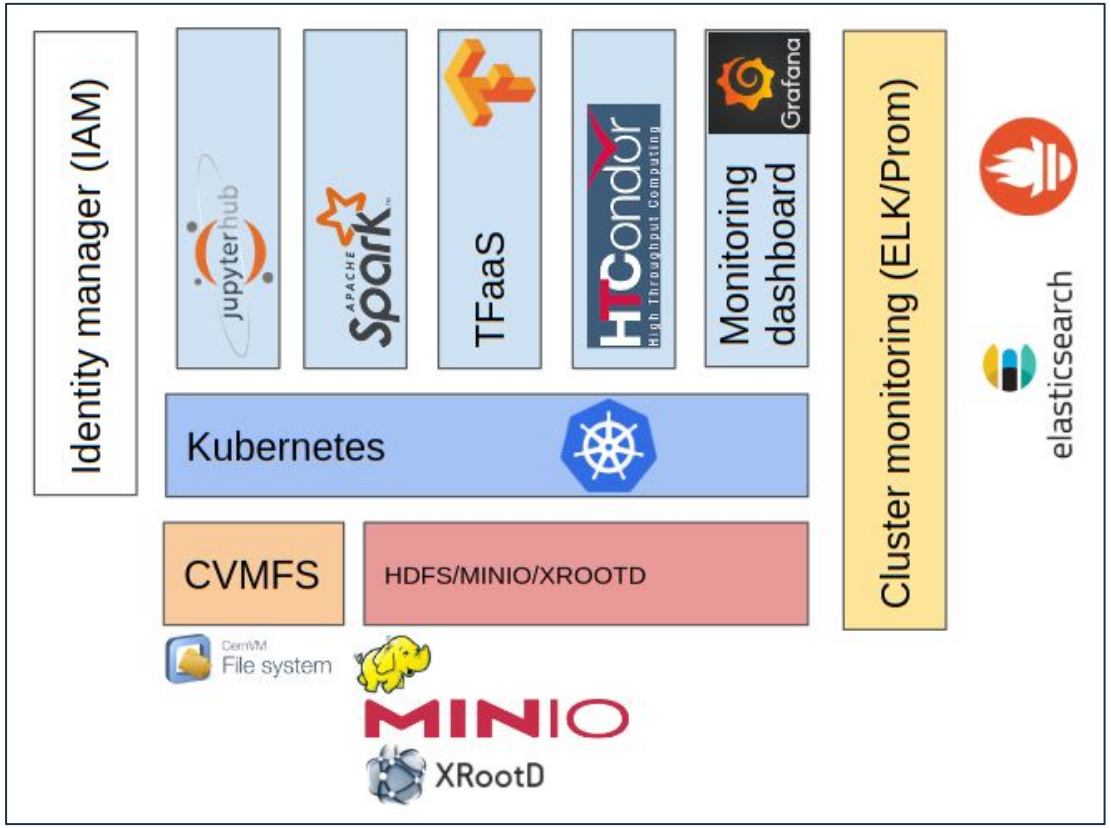
## Data Access

- **XrootD** is used to implement caching mechanisms for remote I/O optimization
- **Onedata** has been also supported

## Software Distribution

- **CVMFS** is the adopted service for distributing software and libraries and user configurations across distributed clusters

**The pillars to provide**




- **Highly Customizable** to accommodate needs from diverse communities
  - Built on top of modern industry standards

# Scientific or societal impact: highlights on Communities

**CMS users analysis: hh—>2b2tau**

- Skimming minAOD (Data & MC) produce flat-ntuple

**Dr. F. Brivio**
**Dr. C. Amendola**

OpenData

**DODAS generated batch system creation and implementation and CMS Open Data 2010 VM Monte Carlo generation example**

2018 CERN SUMMER STUDENT PROGRAMME REPORT

Felipe Navarro

*Supervisor:* Kati Lassila-Perini

September 4, 2018

**The Alpha Magnetic Spectrometer measures Charged Cosmic Rays (0.1 – 2000 GV) in space since 2011, May 19th**

1. AMS collaboration previously published B, C and O fluxes only as a function of energy and time-integrated. **This new analysis, has been performed using the ntuples produced running on DODAS;**
2. Electrons and positrons fluxes, as a function of time have been already published with 27 days time granularity. **A new analysis, and using the ntuples produced on DODAS, is extending the time range and producing the electron (positron) fluxes on a daily (weekly) basis;**

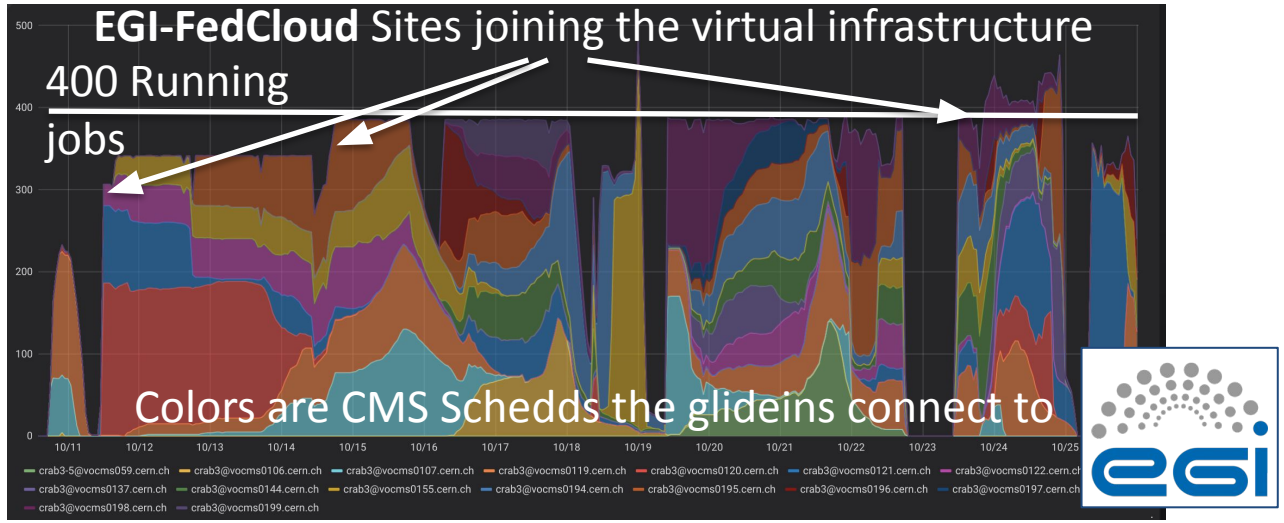
**Dr. M. Duranti**
**Dr. V. Formato**

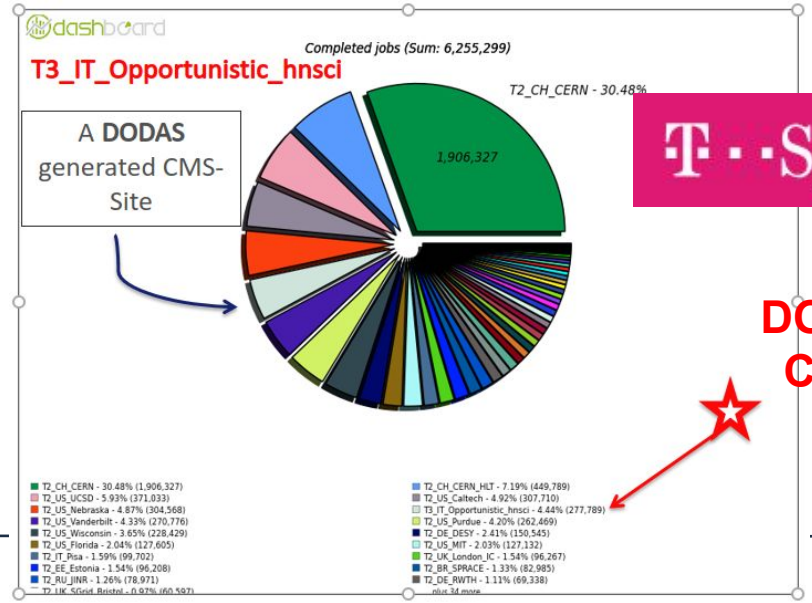
The LAT instrument onboard of Fermi gamma-ray science telescope (Atwood et al 2009) observes the sky in the gamma rays range between 30MeV - 300 GeV since August 2008.

DODAS clusters allows to reduce the computation time, scaling with cluster size. This implies a faster turnaround in the data analysis steps **(VERY PROMISING)** **Dr. Sara Cutini**

**High Energy Physics ; Astro Particles Experiments**

# Scientific or societal impact: highlights on Infrastructures

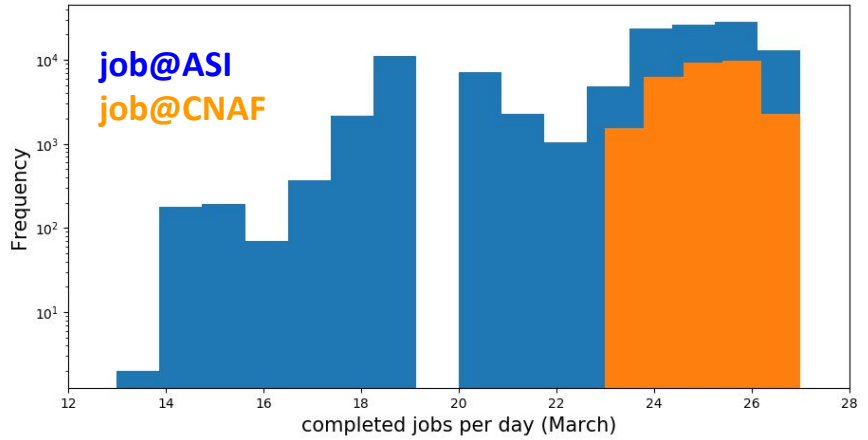


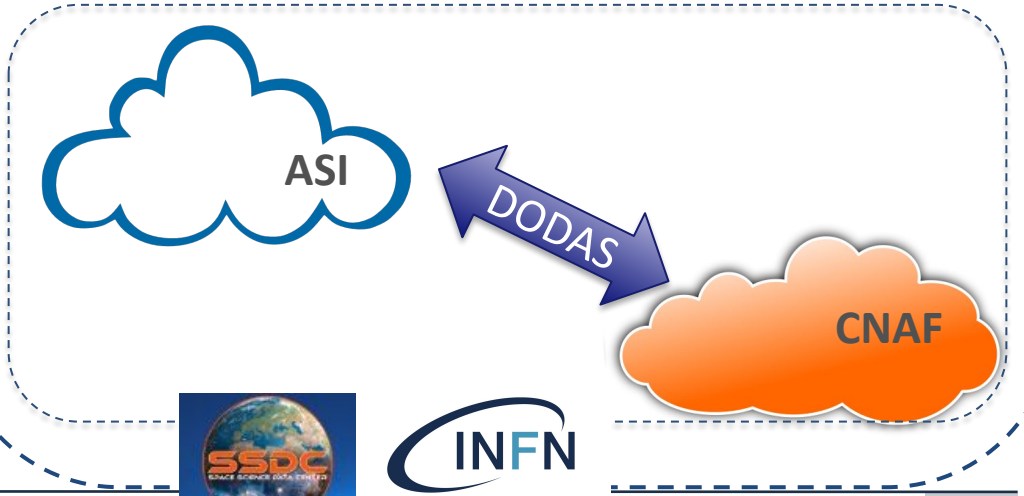




T-Systems

**DODAS within top 6 CMS Tier2s for 10 days**

**The Space Scientific Data Center (SSDC) of the Italian Space Agency (ASI) host an AMS farm.**

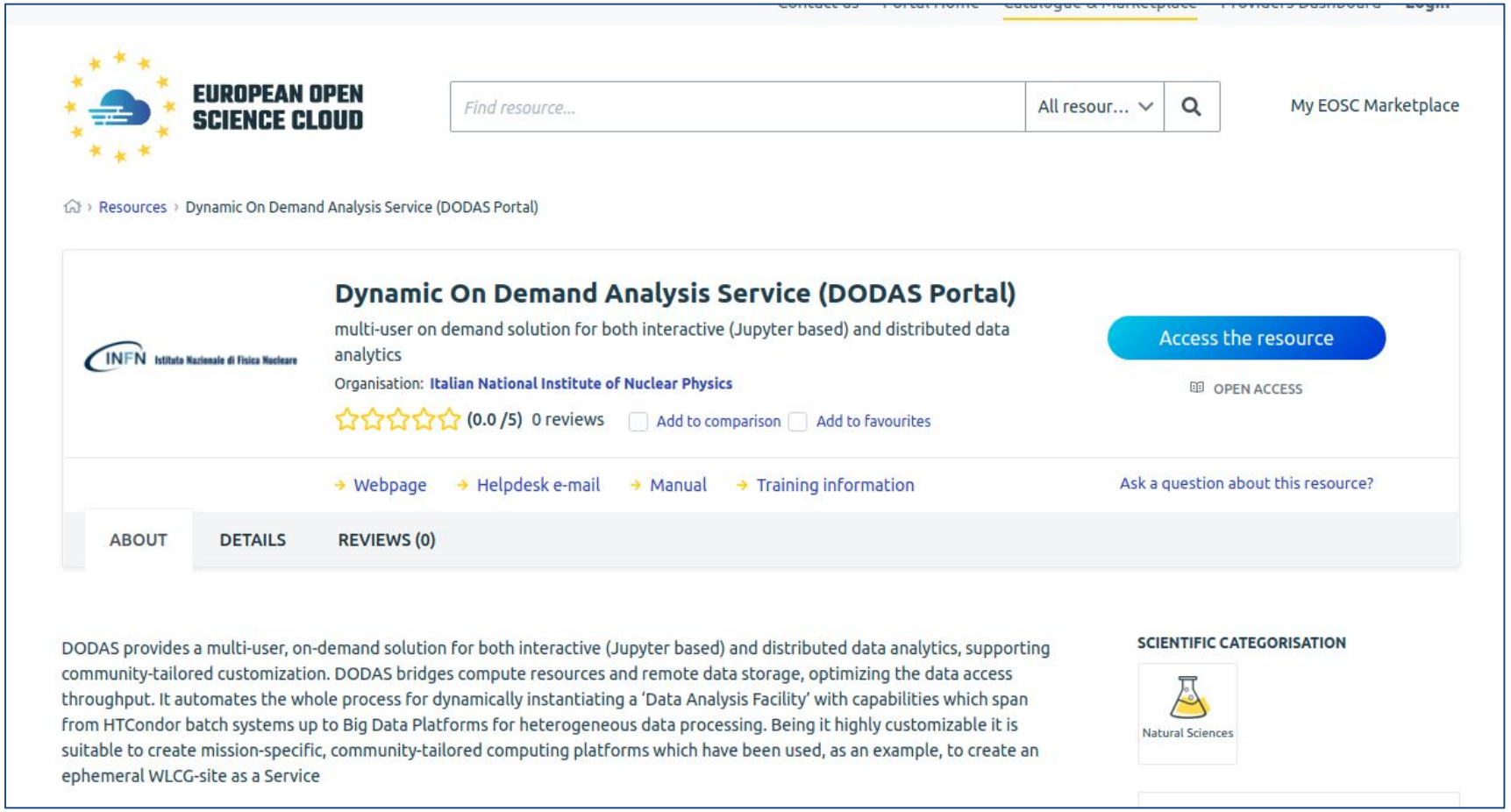



SSDC

INFN
Istituto Nazionale di Fisica Nucleare

# DODAS and the EOSC Service portal




DODAS Will

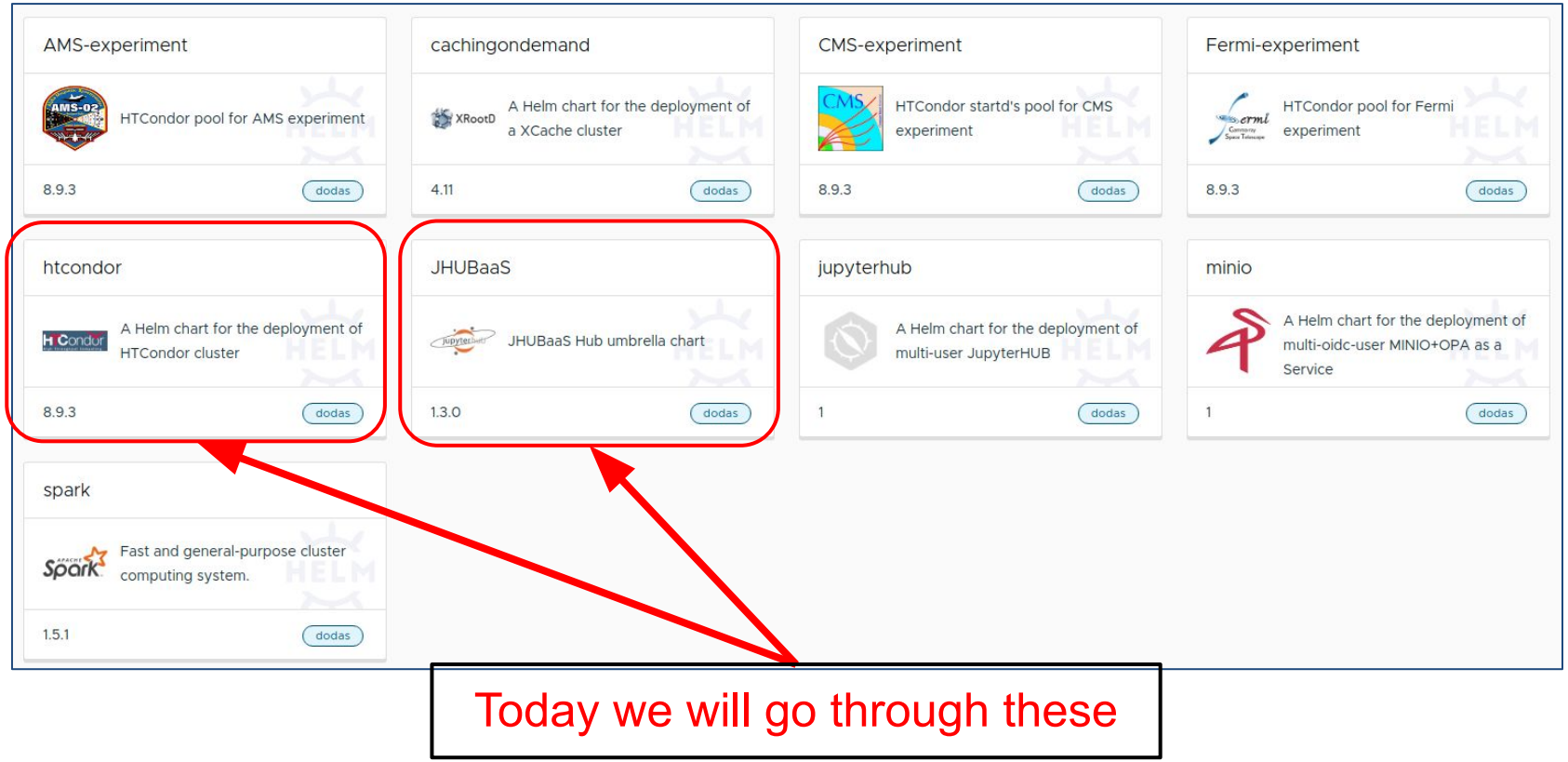
- keep operating under the **EGI-ACE project**
- will continue to support the adopter in the context of the **INFN-Cloud National project**

# Introduce the demonstration

## Today's Demo By Steps

- Very quick overview of the underlying declarative approach
  - **to show how DODAS interacts with any cloud providers**
- Walkthrough the DODAS Catalog
  - **to show authN/Z and how to choose your DODAS preferred app**
- let's instantiate a generic batch on demand
  - **how to instantiate a DODAS app**
- Interacting with instantiated services
  - **how to use the app, i.e. launching a job**

**DODAS (composable) Catalog. Demo instance here** (proper user authorization required)



# Backup

EOSC-hub

2017

First prototype developed in the context of INDIGO-DataCloud project.
Started the first integration with CMS Experiment at CERN

2018

DODAS become a Thematic Service of EOSC-hub. Consolidated CMS and developed a intense exploitation plan

Start the integration with AMS Experiment nowadays in production

2019

Scale test over HNSciCloud

Google Compute Engine

2020

Start the integration with Virgo : work in progress

Establishing synergies with ESCAPE Project and WLCG (DOMA)

Start the integration with FERMI-Lattice Experiment... today the first real analysis is ongoing

# Thank you
# for your attention!

*Questions?*


eosc-hub.eu @EOSC_eu