

# INDIGO-Datacloud: a Cloud-based Platform as a Service oriented to scientific computing for the exploitation of heterogeneous resources

*Thursday, 13 October 2016 11:15 (15 minutes)*

INDIGO-DataCloud (INDIGO for short, <https://www.indigo-datacloud.eu>) is a project started in April 2015, funded under the EC Horizon 2020 framework program. It includes 26 European partners located in 11 countries and addresses the challenge of developing open source software, deployable in the form of a data/computing platform, aimed to scientific communities and designed to be deployed on public or private Clouds, integrated with existing resources or e-infrastructures.

In this contribution the architectural foundations of the project will be covered, starting from its motivations, discussing technology gaps that currently prevent effective exploitation of distributed computing or storage resources by many scientific communities. The overall structure and timeline of the project will also be described.

The main components of the INDIGO architecture in the three key areas of IaaS, PaaS and User Interfaces will then be illustrated. The modular INDIGO components, addressing the requirements of both scientific users and cloud/data providers, are based upon or extend established open source solutions such as OpenStack, OpenNebula, Docker containers, Kubernetes, Apache Mesos, HTCondor, OpenID-Connect, OAuth, and leverage both de facto and de jure standards.

Starting from the INDIGO components, we will then describe the key solutions that the project has been working on. These solutions are the real driver and objective of the project and derive directly from use cases presented by its many scientific communities, covering areas such as Physics, Astrophysics, Bioinformatics, Structural and molecular biology, Climate modeling, Geophysics, Cultural heritage and others. In this contribution we will specifically highlight how the INDIGO software can be useful to tackle common use cases in the HEP world. For example, we will describe how topics such as batch computing, interactive analysis, distributed authentication and authorization, workload management and data access / placement can be addressed through the INDIGO software. Integration with existing data centers and with well-known tools used in the HEP world such as FTS, Dynafed, HTCondor, dCache, StoRM, with popular distributed filesystems and with Cloud management frameworks such as OpenStack and OpenNebula as well as support for X.509, OpenID-Connect and SAML will also be discussed, together with deployment strategies. A description of the first results and of the available testbeds and infrastructures where the INDIGO software has been deployed will then be given.

Finally, this contribution will discuss how INDIGO-DataCloud can complement and integrate with other projects and communities and with existing multi-national, multi-community infrastructures such as those provided by EGI, EUDAT and the HelixNebula Science Cloud. The importance of INDIGO for upcoming EC initiatives such as the European Open Science Cloud and the European Data Infrastructure will also be highlighted.

## **Tertiary Keyword (Optional)**

Computing middleware

## **Secondary Keyword (Optional)**

Distributed data handling

## **Primary Keyword (Mandatory)**

Cloud technologies

**Primary authors:** SALOMONI, Davide (INFN CNAF); DONVITO, Giacinto (INFN-Bari); CAMPOS PLASENCIA, Isabel (Consejo Superior de Investigaciones Cientificas (CSIC) (ES)); MARCO, Jesus (Universidad de Cantabria (ES)); OLIVEIRA GOMES, Jorge (LIP Laboratorio de Instrumentaco e Fisica Experimental de Particulas); GAIDO, Luciano (Universita e INFN (IT)); DUTKA, Lukasz (Cyfronet); PLOCIENNIK, Marcin (PSNC); HARDT, Marcus (Karlsruhe Institute of Technology); FUHRMANN, Patrick (DESY); BARBERA, Roberto (Universita e INFN, Catania (IT))

**Co-authors:** ITALIANO, Alessandro Italiano (INFN-CNAF); Mr LOPEZ GARCIA, Alvaro (CSIC); DORIGO, Alvise (Universita e INFN, Padova (IT)); CECCANTI, Andrea; KRYZA, Bartosz (ACC Cyfronet-AGH); WEGH, Bas (KIT); BERNARDT, Christian (Deutsches Elektronen-Synchrotron (DE)); MICHELOTTO, Diego (INFN CNAF); AIFTIMIEI, Doina Cristina (INFN); RONCHIERI, Elisabetta; GIORGIO, Emidio (Istituto Nazionale Fisica Nucleare (IT)); FATTIBENE, Enrico (INFN - National Institute for Nuclear Physics); FANZAGO, Federica (Unknown); SCHAER, Frederic (CEA); MOLTÓ, Germán (Universidad Politécnica de Valencia); ALOISIO, Giovanni (Unknown); ANDRONICO, Giuseppe (INFN SEZIONE DI CATANIA); BLANQUER, Ignacio (UPV); SITERA, Jiri (Unknown); PINA, Joao (LIP, Lisbon); SCHWARZ, Lionel (IN2P3); ZANGRANDO, Lisa (Universita e INFN, Padova (IT)); MATYSKA, Ludek (Unknown); FARGETTA, Marco (INFN Catania); VERLATO, Marco (Universita e INFN, Padova (IT)); ANTONACCI, Marica (INFN Bari); MOURA DAVID, Mario Jorge (LIP Laboratorio de Instrumentaco e Fisica Experimental de Particulas); SGARAVATTO, Massimo (Universita e INFN, Padova (IT)); VELTEN, Mathieu (CERN); VILJOEN, Matthew James (STFC - Rutherford Appleton Lab. (GB)); OWSIAK, Michal (PSNC); URBANIAK, Michal; CABALLER, Miguel (UPV); MULAC, Milos (Unknown); ORVIZ FERNANDEZ, Pablo (Universidad de Cantabria (ES)); MILLAR, Paul; SOLAGNA, Peter (EGL.eu); HONORE, Pierre-Francois (Istituto de Fisica Corpuscular (IFIC) UV-CSIC); BRITO DA ROCHA, Ricardo (CERN); BRUNO, Riccardo (INFN Catania); FIORE, Sandro (Unknown); VALLERO, Sara (Universita e INFN Torino (IT)); TANEJA, Sonia (Universita e INFN, Bologna (IT)); JANKOWSKI, Stanislaw (PSNC); BAGNASCO, Stefano (I.N.F.N. TORINO); DAL PRA, Stefano (INFN); REYNAUD, Sylvain (CNRS); BELL, Tim (CERN); SZEPIENIEC, Tomasz; LLORENS, Vincent (CNRS); SAPUNENKO, Vladimir (INFN-CNAF (IT)); SUSTR, Zdenek (Czech Technical University (CZ))

**Presenter:** FUHRMANN, Patrick (Deutsches Elektronen-Synchrotron (DE))

**Session Classification:** Track 6: Infrastructures

**Track Classification:** Track 6: Infrastructures