LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS
partículas e tecnologia

Computing @ LIP

LIP Distributed Computing and Digital Infrastructures group

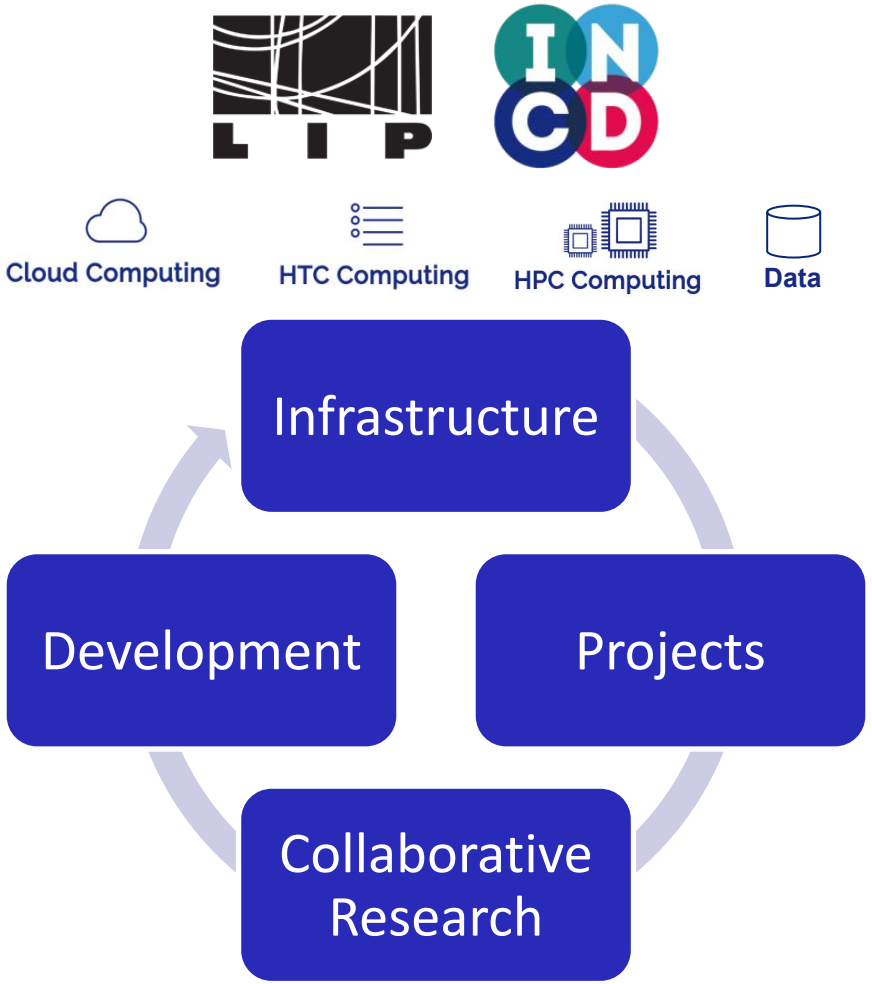
Delivering IT services
Computing and data

For **LIP**

1. **Delivering internal IT services to LIP**
   - Supporting research, innovation, education, outreach and administrative activities at LIP.
2. **Participation in national and international projects, initiatives and digital infrastructures**
   - Performing Research and Innovation

Via **INCD**

3. **Provide compute and data oriented services to Portuguese academic and research community**
   - INCD is a spinoff from the LIP participation in projects on computing and digital infrastructures.
   - Is the basis for the LIP computing services.
   - INCD itself is a legal entity, private non-profit association that joins LIP, FCT and LNEC.
   - Research infrastructure in the FCT roadmap of research infrastructures (until 2022).



Competences
Projects & Initiatives

The group has participation in projects, infrastructures and initiatives in areas such as:

- Digital infrastructures
- High Performance Computing
- High Throughput Computing
- Cloud Computing
- Federated/distributed computing
- Software Quality Assurance
- Software management
- Linux containers
- Data repositories

LIP core IT

Infraestrutura Nacional de Computação Distribuída

computing and data services for the academic & research community

LIP | FCT-FCCN | LNEC

Infrastructures and initiatives

EGI · IBERGRID · WLCG Worldwide LHC Computing Grid · EUROPEAN OPEN SCIENCE CLOUD · RNCA REDE NACIONAL DE COMPUTAÇÃO AVANÇADA

Projects

EOSC SYNERGY · EGI-ACE · C-SCALE · EOSC Future · BIG HPC · EURO² · interTwin · DT-GEO · AI4 · eosc · iMagine · pollen repositório

Human resources

- **LIP staff**
- **INCD staff**
- **Collaborators**

- **7x PhDs**
  - 2x LIP staff
  - 2x INCD staff
  - 3x Collaborators
- **8x Engineers**
  - 7x LIP staff
  - 1x INCD staff
- **3x Technicians**
  - 3x LIP staff

	Staff salaries
LIP (computing projects)	**88%**
LIP (own funds)	2%
INCD (mainly projects)	10%

- **High sustainability risk**
- **Services on best effort**

LIP staff

Catarina Gonçalves	Communication	LIP Lisboa - Communication, dissemination, outreach
Hugo Gomes	Technician	LIP Lisboa - Web development, IT support, events, multimedia, communication
Carlos Manuel	Technician	LIP Lisboa - Design, web development, events, multimedia, communication
Davi Parma	Technician	LIP Lisboa - Web development for internal and administrative services
João Martins	Researcher	LIP Lisboa - Fabric mgmt, storage, computing, HPC, grid, virtualization, support
João Pina	Researcher	LIP Lisboa - WLCG Tier-2, software management, user support, grid
Jorge Gomes	Researcher	LIP Lisboa - Projects mgmt, network mgmt, computing, sw development, security
Mário David	Researcher	LIP Lisboa - Cloud computing, containers, sw quality assurance, development
Nuno Dias	Researcher	LIP Lisboa - Security, data protection, network services, desktops, laptops
José Aparício	Engineer	LIP Lisboa - Datacenter, networks, notebooks, desktops, hw maintenance, support
Miguel Viana	Engineer	LIP Minho - Cloud, containers, software integration and validation, monitoring
Samuel Bernardo	Engineer	LIP Lisboa - Development, sw quality, AAI, computing, cloud, containers, DevOps
Zacarias Benta	Engineer	LIP Minho - HPC, fabric mgmt, computing, virtualization, support, sw development

Communications

Core Computing Services

INCD + Collaborators

Catarina Ortigão	Administration	INCD - Administrative and managerial support
César Ferreira	Engineer	INCD - HPC/HTC, fabric mgmt, computing, virtualization, containers, support
João Machado	Researcher	INCD - Data repositories, open science, data services, software development
António Esteves	Researcher	University of Minho - Application of machine/deep learning techniques
António Pina	Researcher	University of Minho - Application performance analysis, parallel programming
José Rufino	Researcher	Polytechnic Institute of Bragança - Parallelization strategies for GPU algorithms

INCD

LIP

LIP Computing Projects	Source	Start	End	Funding
EGI-Engage	EU	2015	2017	108 500€
INDIGO-DataCloud	EU	2015	2017	503 625€
DEEP-Hybrid-DataCloud	EU	2017	2020	362 500€
EOSC-Hub	EU	2018	2021	338 687€
EOSC-Synergy	EU	2019	2022	433 000€
INCD SAICT/2016 (FCT infrastructures roadmap)	FCT	2017	2022	223 000€
EuroCC	EU/FCT	2020	2022	347 051€
BigHPC	FCT	2020	2023	249 561€
EGI-ACE	EU	2021	2023	196 238€
Contract FCT for catchall research data repository	FCT	2022	2023	19 999€
Support for the Tier-2 WLCG (FCT CERN fund)	FCT	2022	2024	29 999€
EOSC-Future	EU	2022	2024	160 375€
iMagine	EU	2022	2025	222 125€
DT-Geo	EU	2022	2025	542 875€
AI4EOSC	EU	2022	2025	350 250€
interTwin	EU	2022	2025	342 812€
EuroCC 2	EU/FCT	2023	2025	146 000€
EOSC-beyond	EU	2024	2027	182 750€
ENVRI-Hub NEXT	EU	2024	2027	225 302€

Only WLCG support project in 7 years

INCD Projects	Source	Start	End	Funding
INCD SAICT/2016 (Research Infrastructures Roadmap)	FCT	2017	2022	2 323 048€
EOSC-Synergy	EU	2019	2022	125 000€
EGI-ACE	EU	2021	2023	78 029€
C-Scale	EU	2021	2023	75 227€
iMagine	EU	2022	2025	50 414€

INCD Protocols	Source	Start	End	Funding
RICA protocol	FCT	2018	until runs out	150 000€
RNCA protocol 2021	FCT	2021	2021	20 000€
RNCA protocol 2021 - datacenter housing Lisbon	FCT	2021	2021	163 000€
RNCA protocol 2022	FCT	2022	2022	97 972€
RNCA protocol 2022 - datacenter housing Lisbon+UTAD	FCT	2022	2022	243 000€
RNCA protocol 2023	FCT	2023	2023	80 000€
RNCA protocol 2023 - datacenter housing Lisbon+UTAD	FCT	2023	2023	243 000€
Institutional scientific employment position	FCT	2023	2029	285 000€

Before 2021 the INCD housing costs including the LIP WLCG Tier-2 were supported by FCT without a formal protocol.

Breakdown (excluding upcoming projects)

LIP over last 7 years

• Human resources + travel	3 394 748 €	Equipment LIP (LA)	200 000 €
• Equipment (CERN fund T2 project)	22 499 €	Yearly maintenance	12 000 €
• Overheads	1 051 694 €		
• Total	**4 576 597 €**		

INCD over last 5 years

• Human resources + travel	484 053 €
• Equipment	2 026 975 €
• Overheads	82 167 €
• Other (protocols + VA)	1 135 493 €
• Total	**3 728 690 €**

National Distributed Computing Infrastructure

Services: scientific computing, data processing and other data oriented services

Target: scientific and academic community, infrastructures, R&I projects, SMEs

Promote: shared resources, advanced computing and data services for research

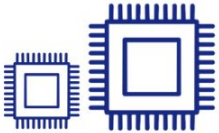
Interface: international digital infrastructures (EGI, IBERGRID, WLCG, EOSC)

Cloud Computing

cloud computing

HTC Computing

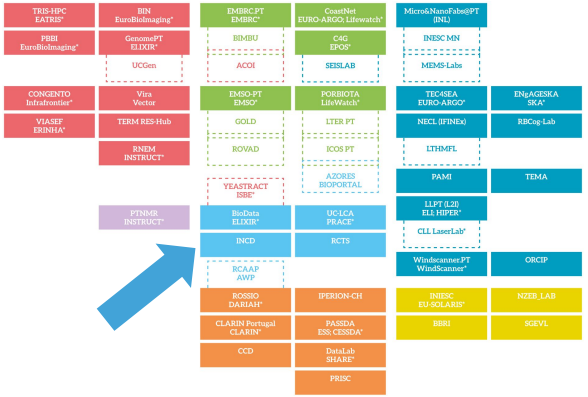
high throughput computing (GRID)



HPC Computing

high performance computing

FCT Roadmap of Research Infrastructures



- Biological and Medical Sciences
- Environmental Sciences
- Physical Sciences and Engineering
- E.Infrastructures
- Social Sciences and Humanities
- Energy
- Materials and Analytical Facilities

2013 — **10 RIs**

INCD submitted to the RI roadmap

2014 — **40 RIs**

The proposal for INCD passes the scientific evaluation and enters the FCT infrastructures roadmap.

2015

The legal entity "Associação INCD" is established.

2016

The Overall budget is split over 40 RIs. An INCD project with reduced budget is submitted.

2017

Roadmap project starts. Establishing INCD basis.

2018

First INCD collaborators are contracted.
First international public tender.

2019 — **56 RIs**

Revised FCT infrastructures roadmap

2021

Infrastructure project is extended until October 2022.

2022

INCD joins RNCA. Project extended to December 2022. Final tender. Roadmap project finishes.

EoI 100 RIs

INCD P2020 project - 01/SAICT/2016 nº 22153

The project funded by the roadmap of research infrastructures finished in December 2022, the results from the evaluation of the project were recently announced by FCT.

The planned scientific objectives were fully achieved. The results show great scientific quality, namely in terms of publications in international refereed journals. The project contributed to the training of young researchers and to the international projection of the team involved.

Advanced Computing Network - RNCA

RNCA
REDE NACIONAL DE COMPUTAÇÃO AVANÇADA

Infraestrutura Nacional de Computação Distribuída

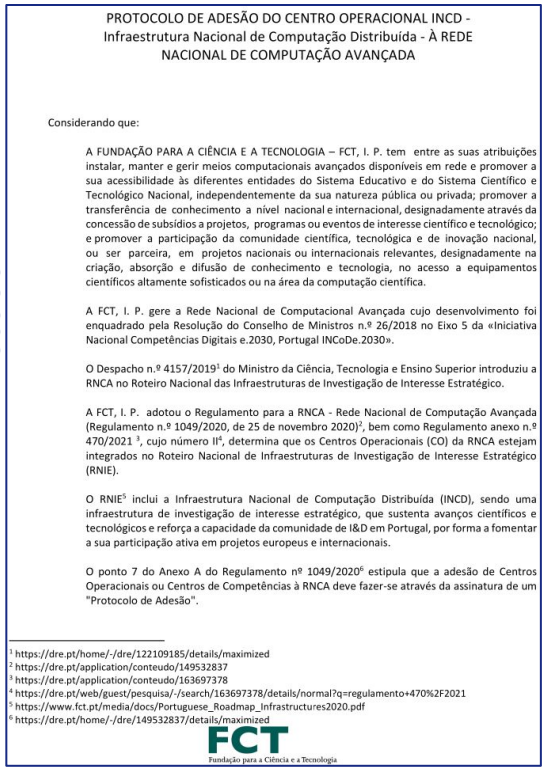
Minho Advanced Computing Center

LCA Laboratory for Advanced Computing

UNIVERSIDADE DE ÉVORA

FCCN UNIDADE DA FCT
Tecnologia para o Conhecimento

RNCA Protocol



PROTOCOLO DE ADESÃO DO CENTRO OPERACIONAL INCD - Infraestrutura Nacional de Computação Distribuída - À REDE NACIONAL DE COMPUTAÇÃO AVANÇADA

Considerando que:

A FUNDAÇÃO PARA A CIÊNCIA E A TECNOLOGIA – FCT, I. P. tem entre as suas atribuições instalar, manter e gerir meios computacionais avançados disponíveis em rede e promover a sua acessibilidade às diferentes entidades do Sistema Educativo e do Sistema Científico e Tecnológico Nacional, independentemente da sua natureza pública ou privada; promover a transferência de conhecimento a nível nacional e internacional, designadamente através da concessão de subsídios a projetos, programas ou eventos de interesse científico e tecnológico; e promover a participação da comunidade científica, tecnológica e de inovação nacional, ou ser parceira, em projetos nacionais ou internacionais relevantes, designadamente na criação, absorção e difusão de conhecimento e tecnologia, no acesso a equipamentos científicos altamente sofisticados ou na área da computação científica.

A FCT, I. P. gere a Rede Nacional de Computacional Avançada cujo desenvolvimento foi enquadrado pela Resolução do Conselho de Ministros n.º 26/2018 no Eixo 5 da «Iniciativa Nacional Competências Digitais e.2030, Portugal INCoDe.2030».

O Despacho n.º 4157/2019[1] do Ministro da Ciência, Tecnologia e Ensino Superior introduziu a RNCA no Roteiro Nacional das Infraestruturas de Investigação de Interesse Estratégico.

A FCT, I. P. adotou o Regulamento para a RNCA - Rede Nacional de Computação Avançada (Regulamento n.º 1049/2020, de 25 de novembro 2020)[2], bem como Regulamento anexo n.º 470/2021 [3], cujo número II[4], determina que os Centros Operacionais (CO) da RNCA estejam integrados no Roteiro Nacional de Infraestruturas de Investigação de Interesse Estratégico (RNIE).

O RNIE[5] inclui a Infraestrutura Nacional de Computação Distribuída (INCD), sendo uma infraestrutura de investigação de interesse estratégico, que sustenta avanços científicos e tecnológicos e reforça a capacidade da comunidade de I&D em Portugal, por forma a fomentar a sua participação ativa em projetos europeus e internacionais.

O ponto 7 do Anexo A do Regulamento nº 1049/2020[6] estipula que a adesão de Centros Operacionais ou Centros de Competências à RNCA deve fazer-se através da assinatura de um "Protocolo de Adesão".

[1] https://dre.pt/home/-/dre/122109185/details/maximized
[2] https://dre.pt/application/conteudo/149532837
[3] https://dre.pt/application/conteudo/163697378
[4] https://dre.pt/web/guest/pesquisa/-/search/163697378/details/normal?q=regulamento+470%2F2021
[5] https://www.fct.pt/media/docs/Portuguese_Roadmap_Infrastructures2020.pdf
[6] https://dre.pt/home/-/dre/149532837/details/maximized

FCT Fundação para a Ciência e a Tecnologia

INCD yearly addendum to the protocol

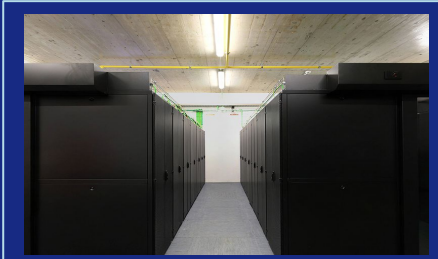
- Advanced computing resources for applications approved in the FCT Advanced Computing Projects Calls;
- Development and operation of a research data repository;
- **Computing resources that satisfy the computational processing commitments assumed with CERN, or support scientific research and technological development projects associated with official scientific collaborations or recognized by CERN, as well as technology-based projects in the field of applied sciences in collaboration with activities supported by the same laboratory;**
- Computing resources that satisfy the national participation in the international EGI, EUDAT and IBERGRID infrastructures, as well as the participation in the EOSC initiative;
- Computational resources that support the **research and development activities of LIP** and LNEC;

INCD operations centers in 2022

INCD-A @ LNEC in Lisbon
HPC / HTC / Cloud / Federation
6000 CPU cores
5 Petabytes online raw
100 Gbps
Includes the WLCG Tier-2

INCD-B @ REN in Riba-de-Ave
(DECOMMISSIONED in 2023)
HPC / HTC
2600 CPU cores
384 Terabytes raw
1 Gbps

INCD-L @ LIP in Lisbon
Tape storage
1 Petabyte backups
10 Gbps



INCD-D @ UTAD in Vila Real
(BEING DEPLOYED)
HPC / HTC / Cloud / Federation
5000 CPU cores + IB HDR200
4 Petabytes online raw
10 Gbps

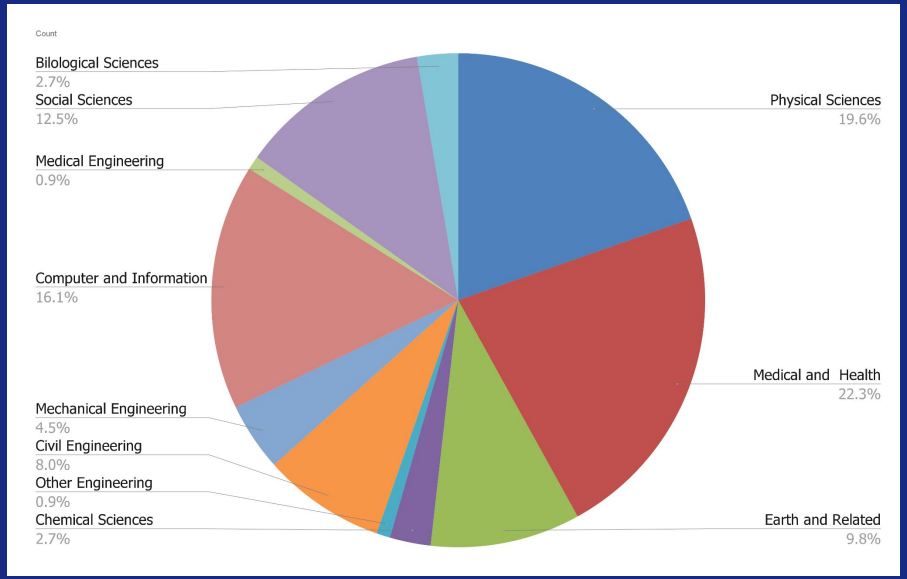
INCD-C @ UC in Coimbra
(BEING IMPROVE)
Tape storage expansion
20 Petabytes
10 Gbps

INCD supported projects in 2022

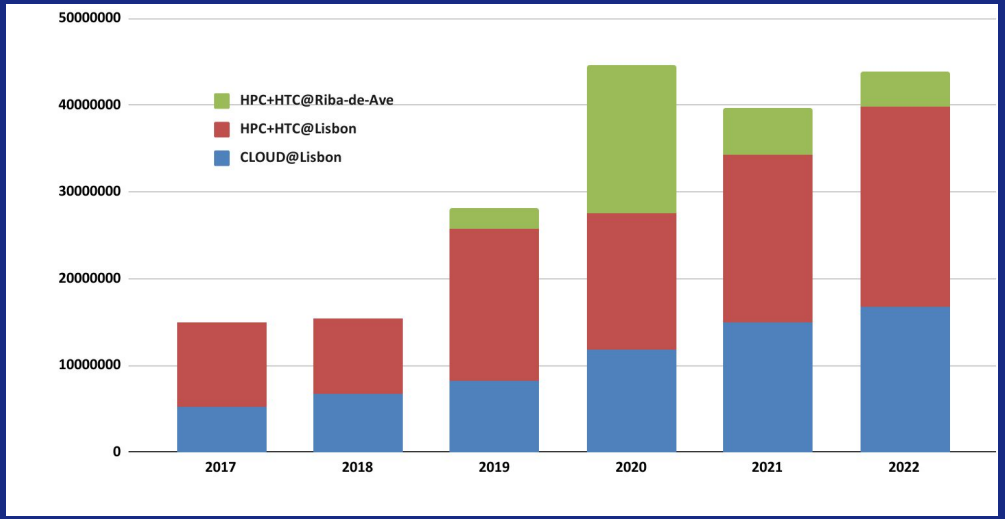
FCT, european and other projects supported: → **92**

Organizations, research units and infrastructures: → **53**

Percentage of supported projects per research area in 2022



Processing Time in Hours since 2017



Total usage in 2022 over 43.854.000 hours

Projects outcomes 2017 - 2022

	2017	2018	2019	2020	2021	2022	Total
Papers direct	52	79	59	97	89	126	**502**
Conference papers	24	30	47	25	32	22	**180**
Books	0	0	0	2	0	0	**2**
PhD Thesis	8	4	3	7	9	5	**36**
MSc Thesis	10	5	17	12	19	23	**86**
Conference posters	9	12	2	24	39	8	**94**
Patents					2	1	**3**
Curated datasets				1	2	19	**22**
Total	**103**	**130**	**128**	**168**	**192**	**204**	**925**

LIP in **IBERGRID and EGI**

LIP responsibilities and activities:

- **IBERGRID and EGI provide the backbone for WLCG**
- **Infrastructure operations coordination at Iberian level and interface with EGI operations**
- **Software management for the EGI and IBERGRID federations**
- **National technical contact point**
- **Security contact for Portugal**
- **Support to user communities**
- **Developing and operating core services e.g. software repositories for the EGI federation**
- **Integration of thematic and/or user services**

IBERGRID Iberian conferences since 2007

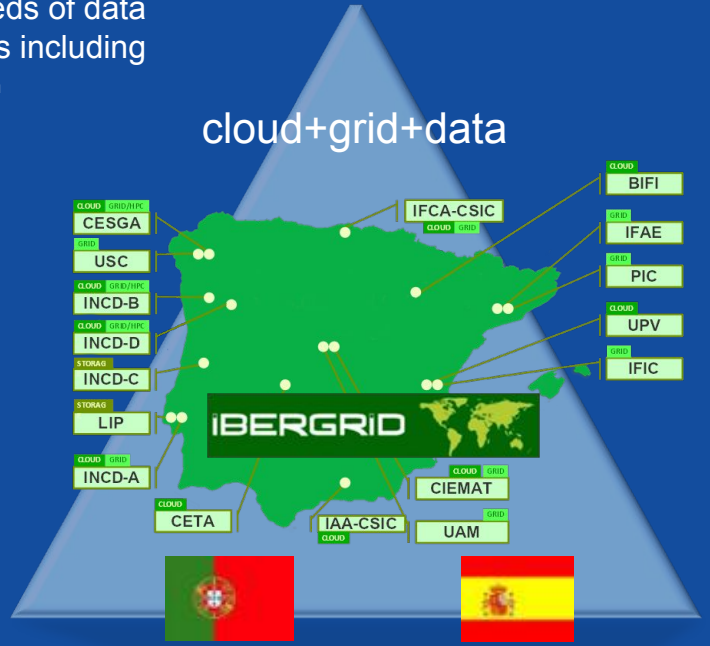
WLCG
Worldwide LHC Computing Grid

EGI

36 EU funded projects
1790[1] Enabled articles
7.1B HTC CPU hours consumed
29 EGI Council participants
84.000 Users
407M Computational jobs
70M Cloud CPU hours consumed

Federating compute and storage from hundreds of data centres including WLCG

cloud+grid+data



The challenge of **Computing for the LHC**

The **MoU for Collaboration in the Deployment and Exploitation of the WLCG** was signed in 2006 by CERN, GRICES and LIP.

GRICES was the office of international relations of the ministry of science.

Under the MoU **LIP operates Tier-2 computing and data facilities for ATLAS and CMS** integrated in the WLCG.

The European Organization for Nuclear Research (CERN)

and

Gabinete de Relações Internacionais da Ciência e do Ensino Superior
Laboratório de Instrumentação e Física Experimental de Partículas

declare that they agree on this Memorandum of Understanding for collaboration in the deployment and exploitation of the Worldwide LHC Computing Grid.

Done in Geneva
on 10 April 2006

Done in Lisbon
on 25 July 2006

For CERN

ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE - CERN CSO

Jos Engelen
Chief Scientific Officer

For GRICES

Maria Virginia de Magalhães Corrêa
GRICES Director

For LIP

Gaspar Pereira de Morais Barreira
LIP President

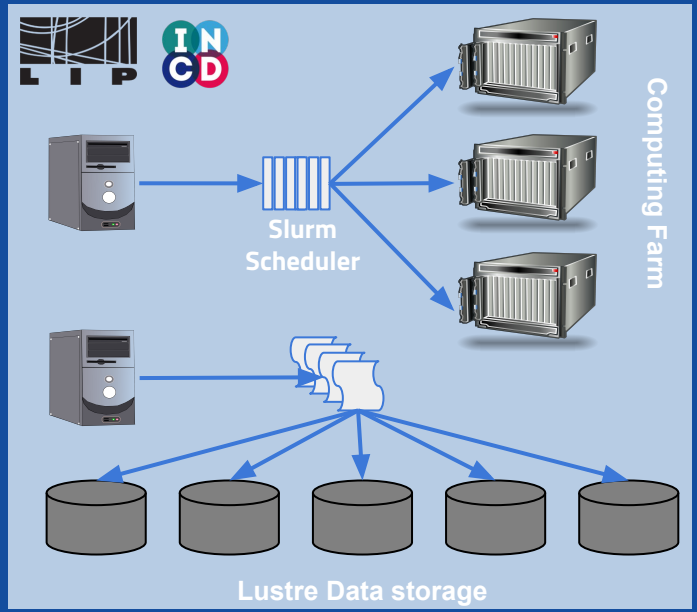
Portuguese Tier-2
ATLAS and CMS

The **Tier-2 / Tier-3** uses the INCD infrastructure and is **operated by the LIP** computing team.

- Based at the INCD Lisbon site
- Shares the site Slurm and Lustre systems

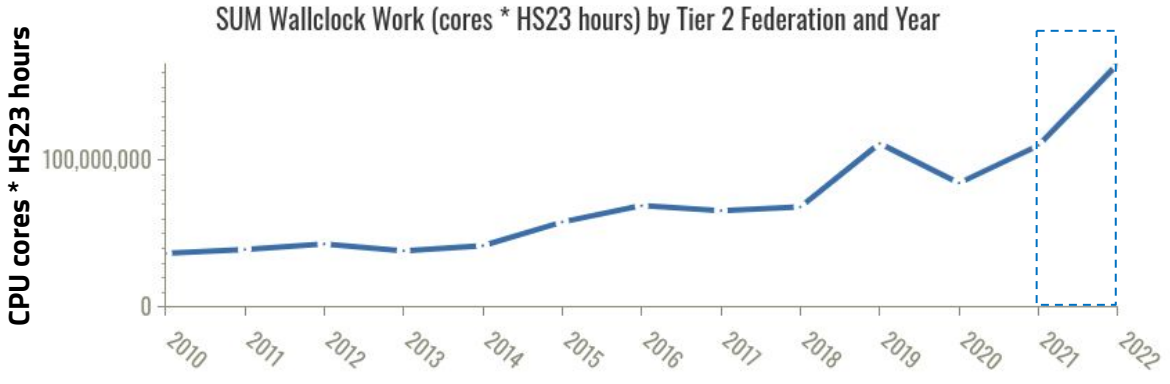
ARC-CE with Slurm as scheduler

XRootd Webdav StoRM SRM with Lustre as Underlying Storage

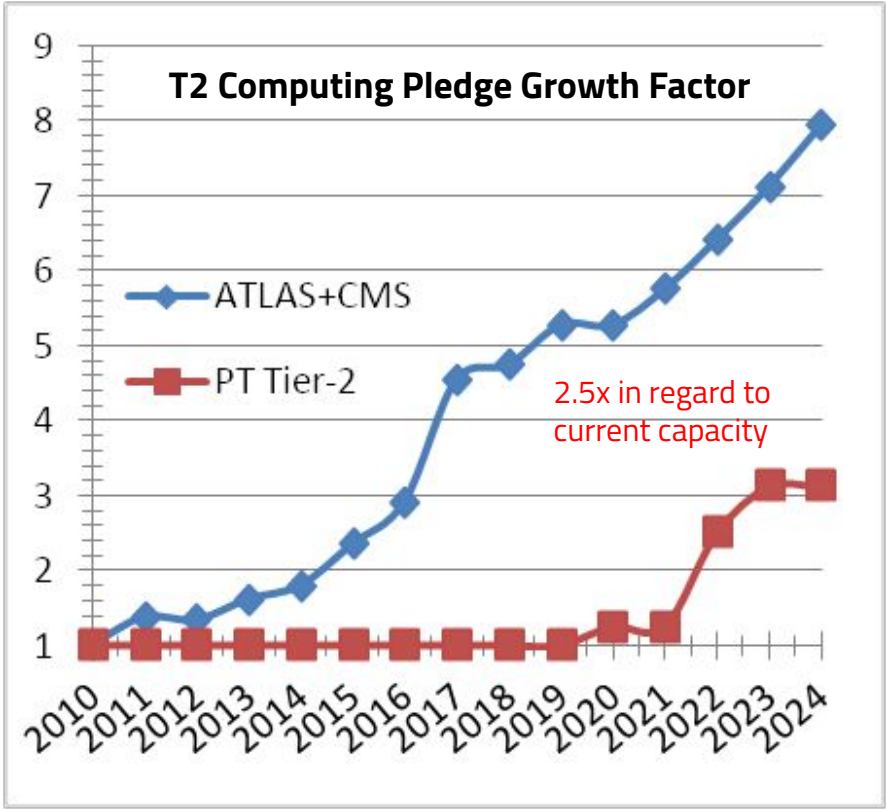


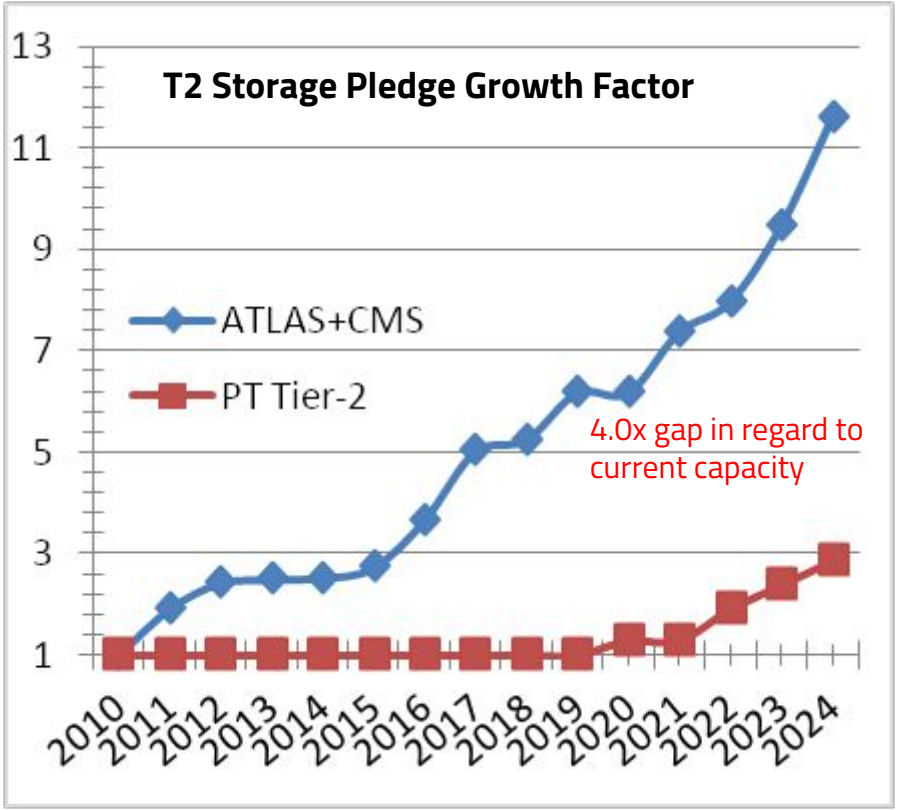
Since 2010:

- 22,077,100 jobs
- 1,029,090,773 HS23 hours
- 50% ATLAS / 50% CMS

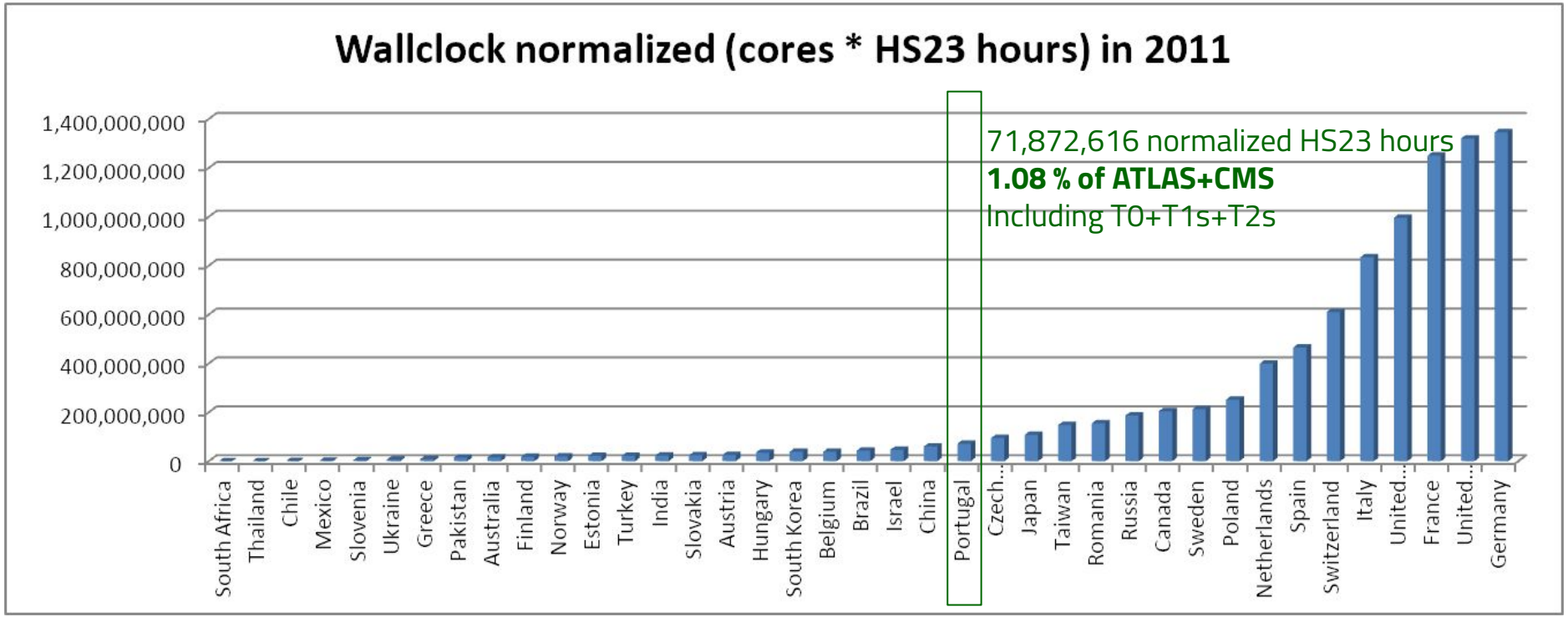


CERN / WLCG - Portuguese Tier-2 pledge since 2010





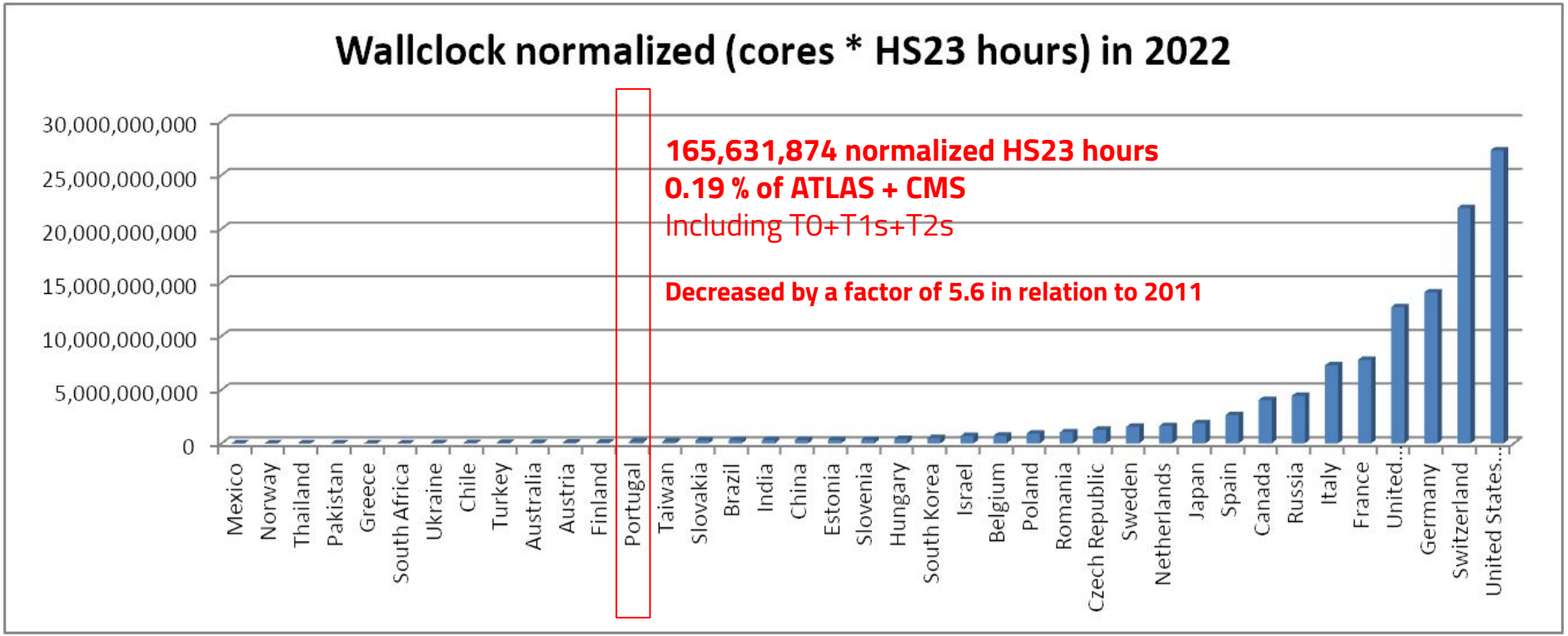
CERN / WLCG - Portuguese Tier-2 accounting (in 2011)



The Portuguese Tier-2 started in 2009 with a capacity of 90% of a nominal Tier-2

https://accounting.egi.eu/wlcg/countries/normelap_processors/COUNTRY/Year/2010/1/2023/7/lhc/onlyinfrajobs/

CERN / WLCG - Portuguese Tier-2 accounting (in 2022)



https://accounting.egi.eu/wlcg/countries/normelap_processors/COUNTRY/Year/2010/1/2023/7/lhc/onlyinfrajobs/

SWOT

Strengths

- Expertise in scientific computing, software integration, management and quality assurance, etc.
- Participation in international research e-infrastructures and initiatives (WLCG, EGI, IBERGRID and EOSC).
- Operating the Portuguese WLCG Tier-2 under the CERN LHC computing MoU.
- Participation in European projects.
- Founding member of INCD and key technological partner.
- Partnership with FCT-FCCN and LNEC and collaboration with other organisations through INCD.
- Previous participation in the FCT roadmap of research infrastructures of strategic interest through INCD.
- Participation in the Portuguese Advanced Computing Network (RNCA).

Weaknesses

- Lack of compute and storage resources to address user needs.
- Capacity wise becoming irrelevant at national level.
- Lack of sustainability with many activities being supported on a voluntary and/or best effort basis.
- Highly overworked team.
- Very large dependency on projects to pay IT staff salaries.
- Heavy administrative burden compromising the effective use of the human resources.
- End of the FCT infrastructures roadmap.
- No clear path and role in the national computing landscape.
- No perspectives for real funding beyond participation in EU projects.

Opportunities

- Participation in activities related to High Performance Computing. ↓
- Participation in open data and digital repositories related activities. ↓
- Potential for public sector applications.
- Possible evolution of the RNCA model.

Threats

- Lack of hardware capacity for ATLAS, CMS and other experiments.
- Competitive market makes difficult contract and retain IT personnel.
- Lack of sustainable funding for human resources.
- Exacerbated focus on supercomputing at national and European level.
- Increasingly higher competition in projects, funding and infrastructure.
- INCD sustainability and future are highly uncertain.

Final remarks

- HEP is highly compute and data intensive and has very large specific requirements
  - Computing infrastructures and support are both essential and an integral part of HEP activities and they have been extremely beneficial to the wider research community and society.
- The infrastructure capacity for HEP activities available to LIP needs to be significantly improved
  - If not the computing support for LIP activities will be likely irreversibly compromised.
- HPC systems are not a replacement for the current dedicated facilities such as the WLCG Tier-2s
  - They are only suitable to provide opportunistic capacity for certain compute intensive applications (e.g. Monte Carlo)
- Portugal must fulfill the LHC computing MoU
  - Does not make sense to fund participation at CERN and not fund the means necessary to take full advantage of that participation (funding for matching research projects and supporting infrastructures including computing).
- There is *life* beyond High Performance Computing
  - Other computing paradigms must also be supported, distributed computing and data access is essential in large collaborative endeavors like HEP experiments.

LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS
partículas e tecnologia

Thanks!

Discovery through science

Innovation through technology

Sharing with People

Projects and activities

- DT-GEO
  - Digital Twin of geophysical extremes dealing with geohazards earthquakes, volcanoes, and tsunamis.
  - Software and Service Quality assessment
  - udocker integration with workflow managers in HPC
  - Application containerisation
- InterTwin
  - Common approach to the implementation of DTs applicable across scientific disciplines
  - Software release and management
  - Quality and validation for applications and services
- iMagine
  - Imaging data and services for aquatic science
  - Federated computing infrastructure
  - Supporting the DEEP AI platform service
- EGI-ACE
  - Advanced computing for EGI
  - Quality assurance for the EGI middleware distributions for grid (UMD) and cloud (CMD).
  - Integration and supporting of thematic services
  - Integration of HPC resources in EGI
- BigHPC
  - Simplify management of High Performance Computing infrastructures for BigData and parallel applications.
  - SQA, DevOps, monitoring and containerisation.
- AI4EOSC
  - Advanced services for AI, ML and DL models and applications in the EOSC.
  - Software quality, data FAIRness
  - Integration of udocker for serverless computing
- EuroCC 2
  - Awareness and communication
  - Training and skills
  - Interaction with academia and public sector
- Contract with FCT - data repositories
  - Pilot for a national research data repository
  - Integration and service provisioning
- c-Scale
  - Federate European EO infrastructure services
  - Integrating and supporting EO use cases
- EOSC-Future
  - IT Service management for EOSC services

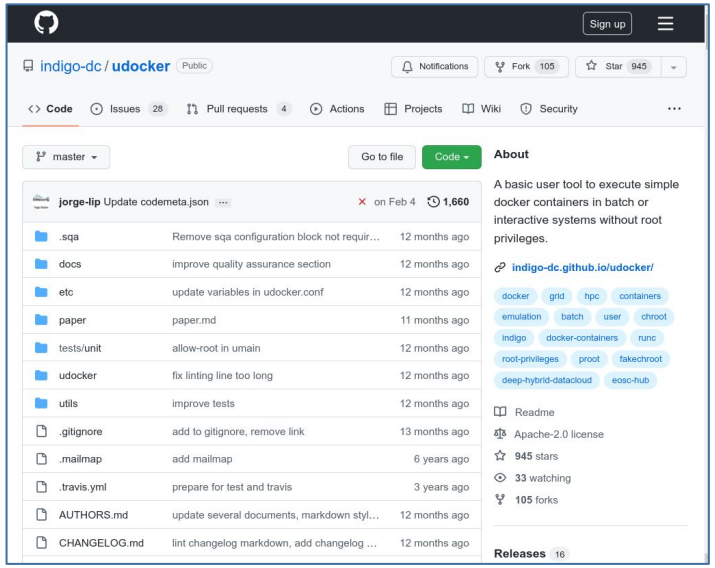
Software from LIP
udocker

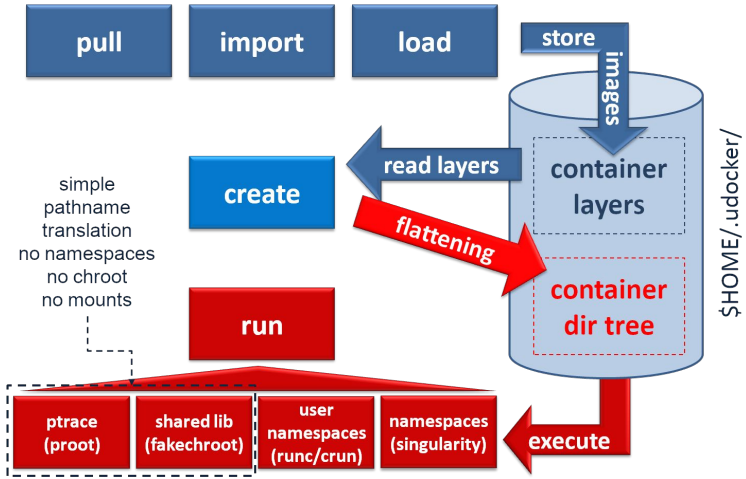
UDOCKER

User tool to execute docker containers in user space. Developed at LIP:

- Fully user space.
- No root privileges required to use or install.
- Does not require compilation.
- Download and execution of docker containers by non-privileged users.
- Suitable for Linux batch systems and interactive clusters managed by other entities such as grid infrastructures.
- Does not require Linux namespaces.

https://github.com/indigo-dc/udocker



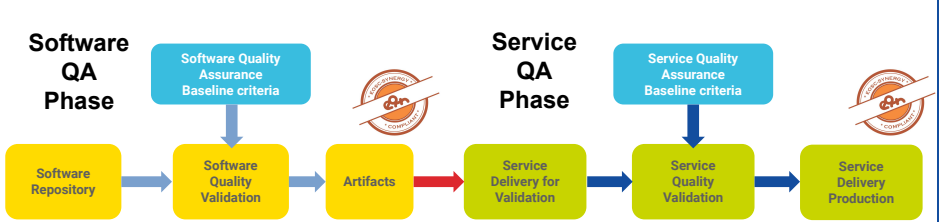


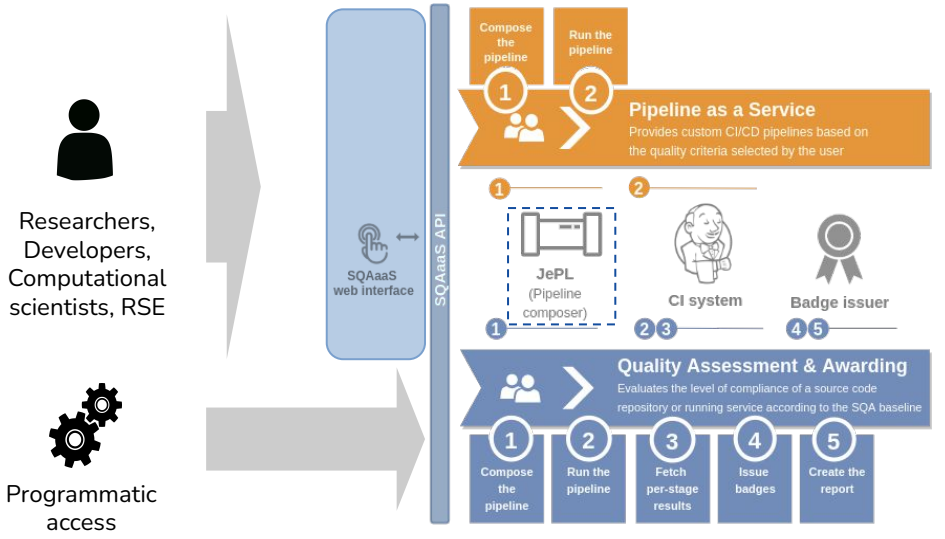
Software from LIP **SQAaaS**

Quality Assurance as-a-Service platform (SQAaaS)

- Enables the on-demand creation of CI/CD pipelines making quality verification and validation easily accessible to developers.
  - The **Pipeline as a Service** building block allows you to compose and test customized CI/CD pipelines in accordance with reference criteria.
  - The **Quality Assessment & Awarding** building block analyses, the level of compliance to the quality baselines.
- Integrates a wide range of quality verification tools that are made easily available through a friendly web interface.

https://sqaaas.eosc-synergy.eu

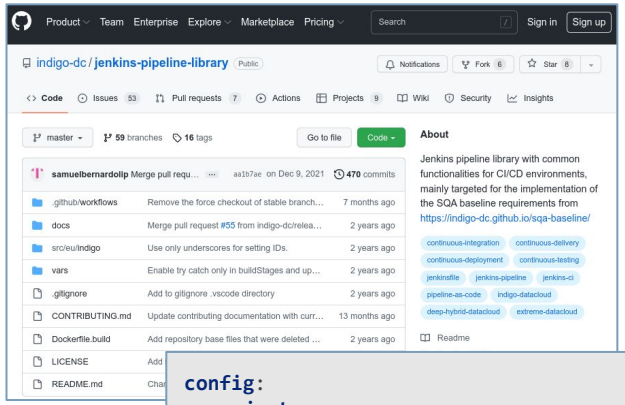




Software from LIP
JePL

Jenkins Pipeline Library (JePL)

- The library that powers the SQAaaS platform.
- Especially suitable for complex setups, you can use directly the JePL instead of the SQAaaS.
- Tech-savvy users tend to favor code over a graphical interface for the task of managing their CI/CD pipelines.
- JePL uses pipeline descriptions written in YAML.
- Just add JePL to your software repository and build your software or service quality assurance using YAML descriptions to benefit from the full set of features.
- JePL implements the software and service baselines maintained by EOSC-Synergy.



config.yml

```yaml
config:
  project_repos:
    myrepo:
      repo: 'https://github.com/myorg/myrepo'

sqa_criteria:
  qc_style:
    repos:
      myrepo:
        container: myrepo-testing
        tox:
          tox_file: /myrepo-testing/tox.ini
            testenv: stylecheck
```

```
|-- .sqa
|    |-- config.yml
|    |-- docker-compose.yml
  |-- Jenkinsfile
```

docker-compose.yml

```yaml
services:
  myrepo-testing:
    image:
"indigodatacloud/ci-images:python3.6"
    hostname: "myrepo-testing-host"
    volumes:
      - type: bind
        source: ./myrepo
          target: /myrepo-testing
```

https://github.com/indigo-dc/jenkins-pipeline-library