



SKA French Regional Center Network infrastructure

CEA/IRFU Workshop on computing
strategy

D. Chapon, D. Corre (CEA/Irfu)

September, 23rd 2026

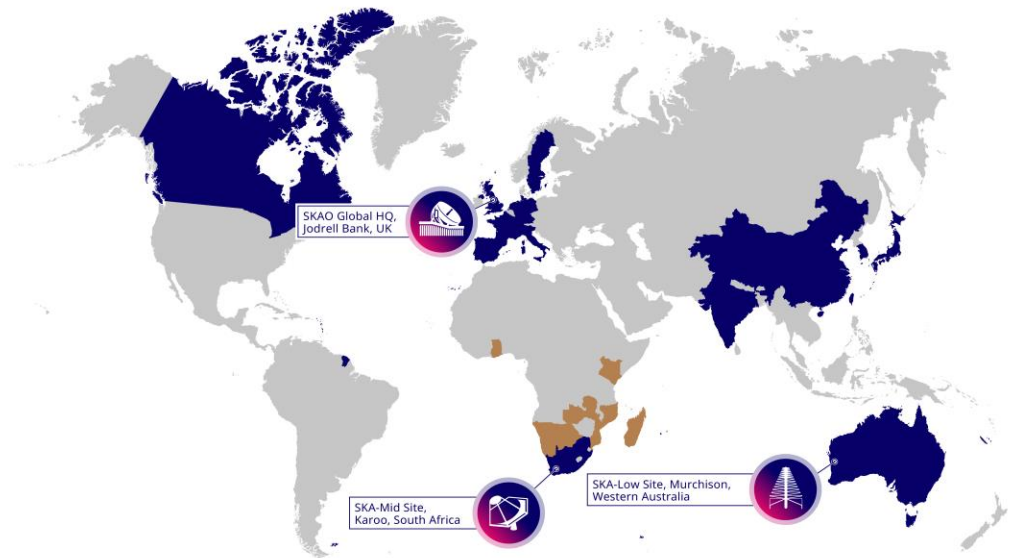


The SKA in brief



- SKAO: an intergovernmental organisation for astrophysics and fundamental Science for the next 50 years.

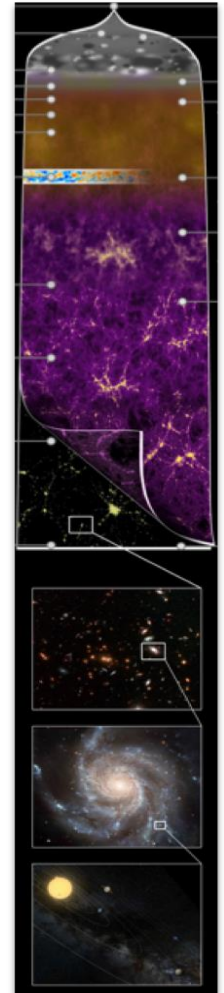
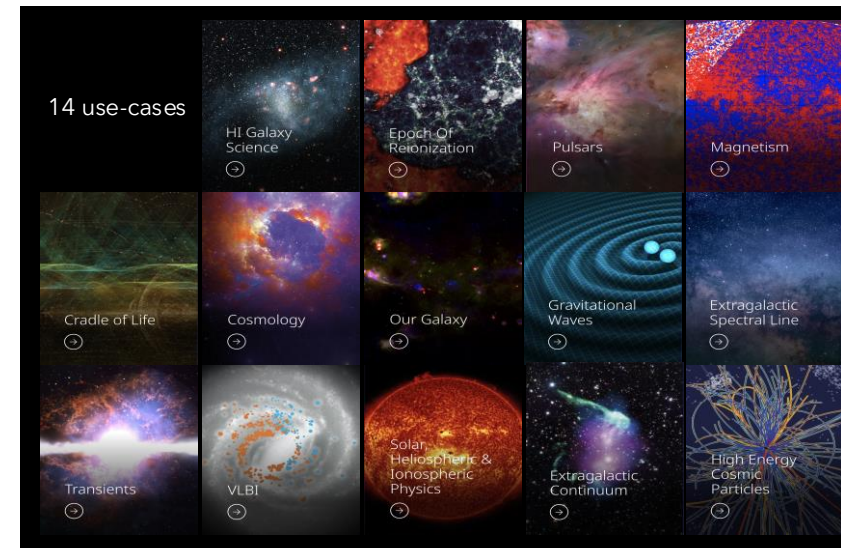
- A huge cyber-infrastructure for Big Science in the XXI^e century, allowing to radically change our understanding of the Universe
 - A wide variety of scientific use-cases
 - A wide range of computing challenges, involving massive data



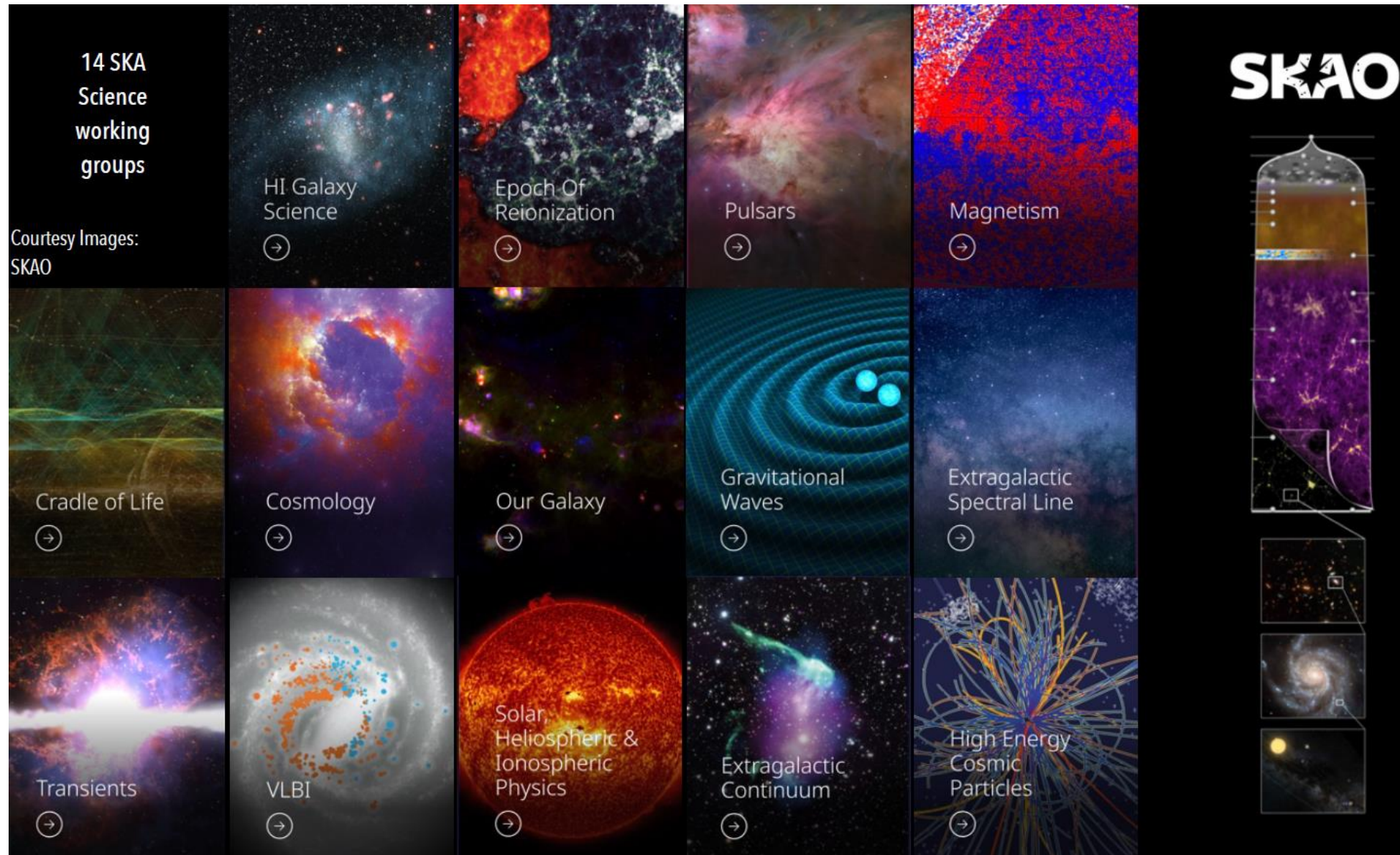
■ SKAO Partnership - includes SKAO Member States* and SKAO Observers (as of August 2025)



■ African Partner Countries



A very large range of Science cases



The SKA infrastructure

Australia : SKA low

131k antennas log /512 stations, Beamforming
European precursors :LOFAR, Nenufar (Nançay)



50MHz



350MHz

South Africa: SKA mid

197 dishes
64 existantes (MeerKAT)



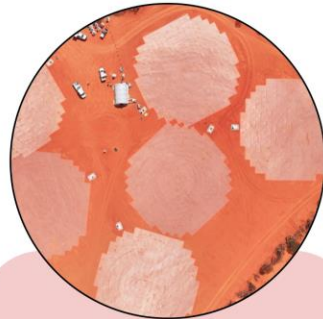
15GHz

- World's largest radioastronomical observatory (continental-scale)
- Operation will start in 2028, for the next 50 years
- Huge expected scientific outcome, covering almost all key questions in astrophysics
- Paramount technological challenges, from hardware to software

SKA milestones

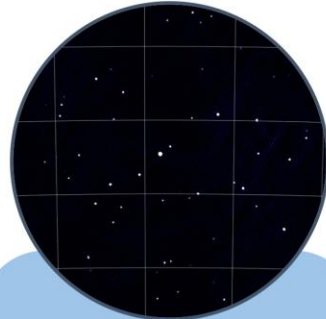


Credit: I. Heywood, SARAO



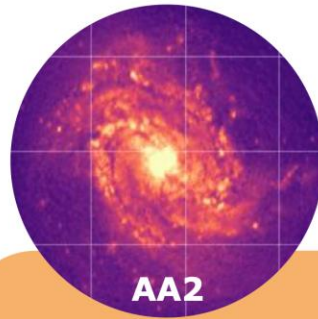
Construction

- Building antennas, dishes, roads etc!
- Followed by Assembly, Integration and Verification

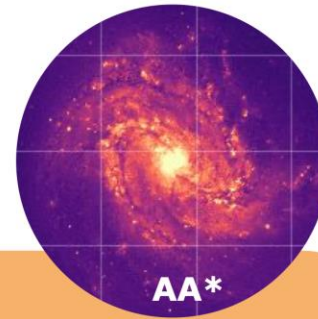


Commissioning

- SKAO activity
- Collaborative across system verification and science commissioning



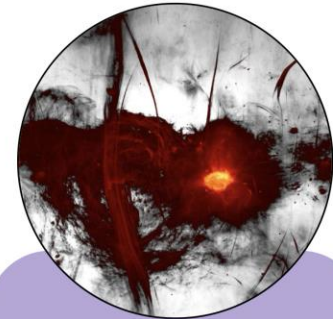
AA2



AA*

Science Verification

- A full dress rehearsal of the end-to-end system for every mode of operation
- Once modes and pipelines are working, the community can submit target ideas
- Data will be publicly available for scrutiny
- Build trust and fostering an early science return



Cycle 0

- Shared-risk PI projects
- SRCNet resources ready for user
- Proprietary periods



Now



Now



Now

Soon

First half 2027

First half 2029

First half 2029

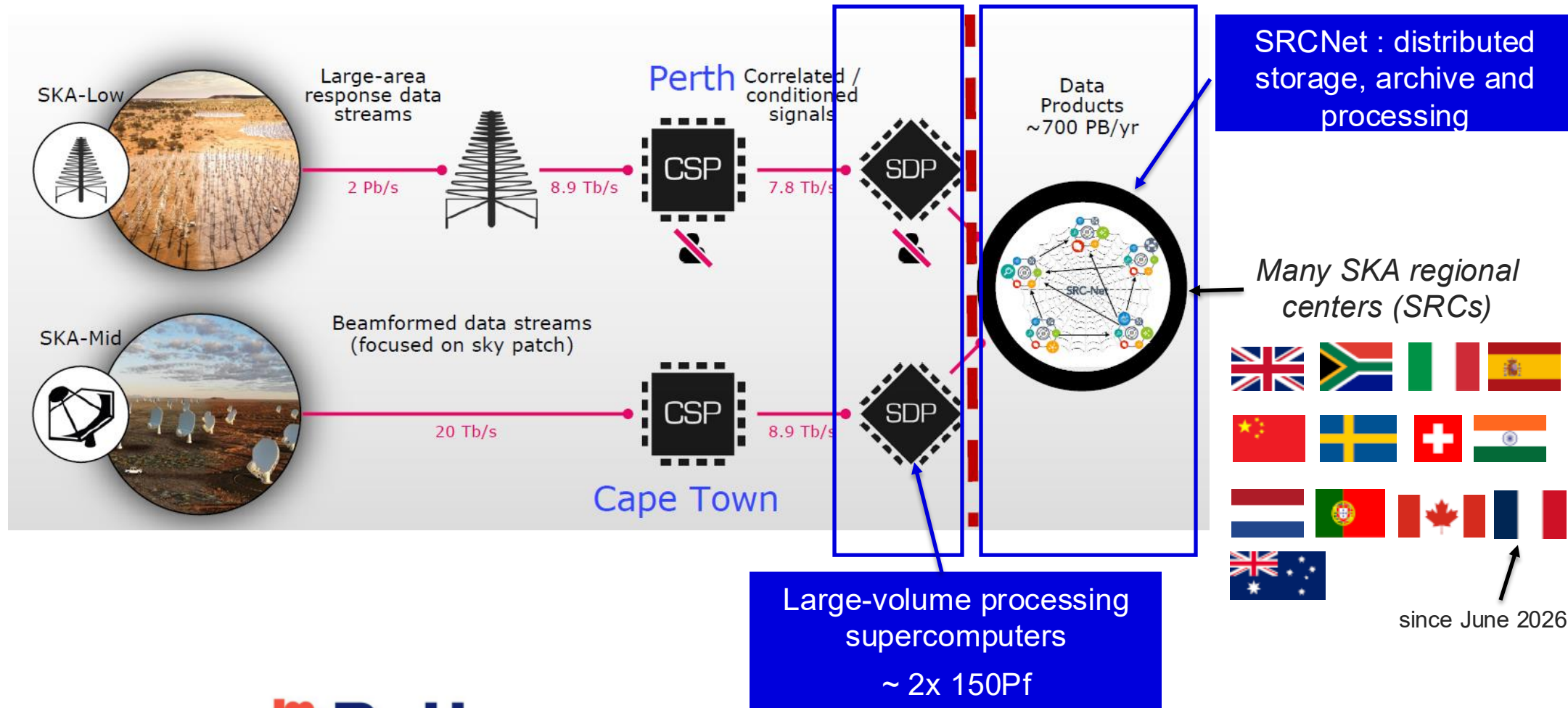
First half 2031

2030

2032

SKA : data processing challenges

- Scientific observatory with an unprecedented amount of data to be produced (and analysed ?): **350 Pbytes/year**



Bull/SKAO contract signed - 11/2025

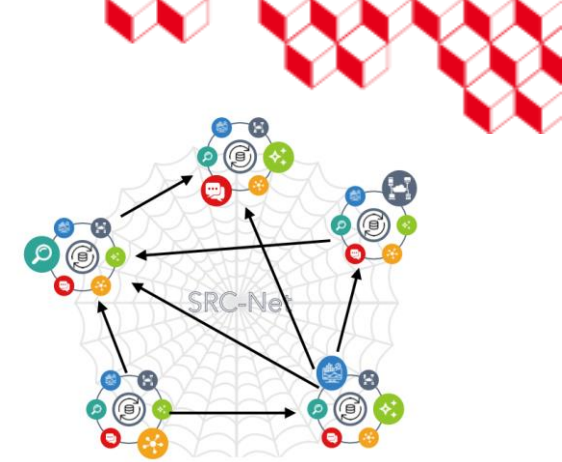
SRCNet : future resource requirements

Compute

- Steady increase up to 15 PFLOPS
- Uncertainties on memory price (2028+) €€€

Storage

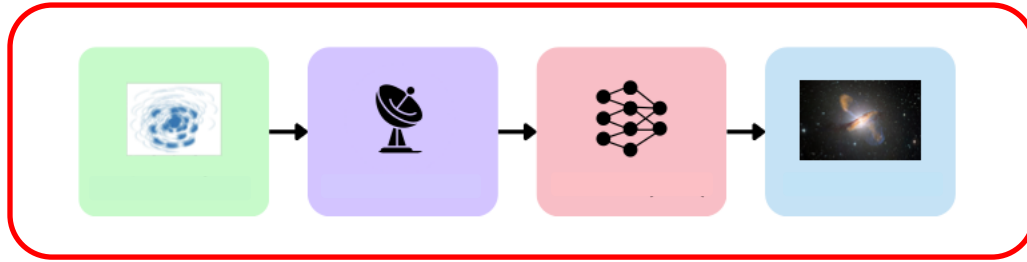
- Steep ramp starting in 2028



		2026	2027	2028	2029	2030-2035
Compute	Req.	0.5 PFLOPS	2.5 PFLOPS	5 PFLOPS	10 PFLOPS	15 PFLOPS
	Pledged	0.8 PFLOPS	4 PFLOPS	TBD	TBD	TBD
Storage	Req.	8PB	8PB	20	100 PB disk + 20PB (tapes)	+250 PB/yr + 500 PB/yr (tapes)
	Pledged	8.6 PB	30.5 PB	TBD	TBD	TBD

France : 3% of global resources

SRCNet: scientific value chain

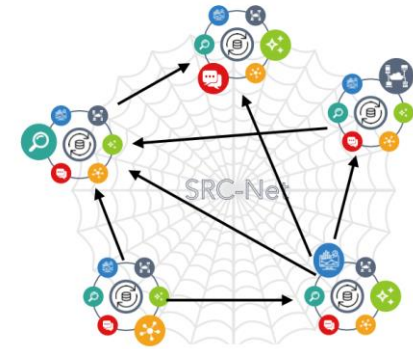


Wide variety of scientific workflows:

- Streaming data/time series for transient
- Large-scale surveys
- Hyperspectral data (Interstellar medium, Epoch of reionization)
- ...

Complexity :

- AI/HPC-coupled, simulation/reconstruction coupling
- Data denoising with ML



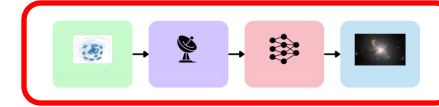
Workflow deployment

- Portability & Performance portability on heterogeneous architectures
- Deployability (packaging/containerization)
- Reproducibility/benchmarking

Services

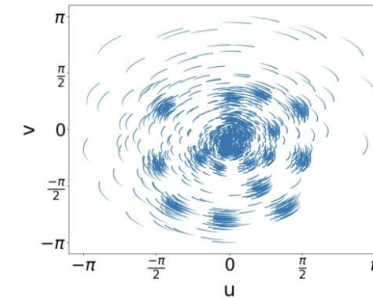
- Data logistics/compute orchestration/storage
- Post-processing/visualization
- Monitoring/operations

AI-based data processing at scale

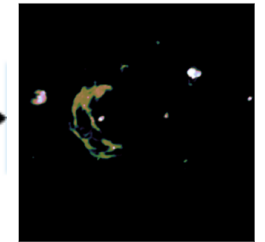
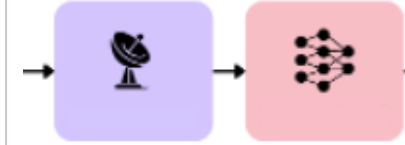


Making AI-based workflows ready for Science

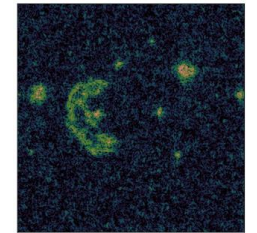
- ML-based workflows for accurate image reconstruction
- Physics-informed : instrument model, measurement statistics, etc
- Validated uncertainty quantification
- **Transverse cross scientific domains at Irfu**



Radio-interferometric measurements



Reconstruction



Error map



... at scale for Exascale ready workflows

- Strong involvement in NumPEX's work on HPC-AI-coupled workflows
- Development of scalable ML-based analytics for Exascale, focus on radio-interferometric image reconstruction (core components in SKA workflows)
- **Distributed computing (multi-GPU) with ML-based denoisers** (e.g. diffusion models, DRUNet, etc) based on DeepInverse
- **Next with SKA: towards large-scale real image reconstruction from radio-interferometric measurements.**
- Execution of scientific workflows on distributed compute/storage infrastructures (Exa-AToW)



Deployable scientific workflows

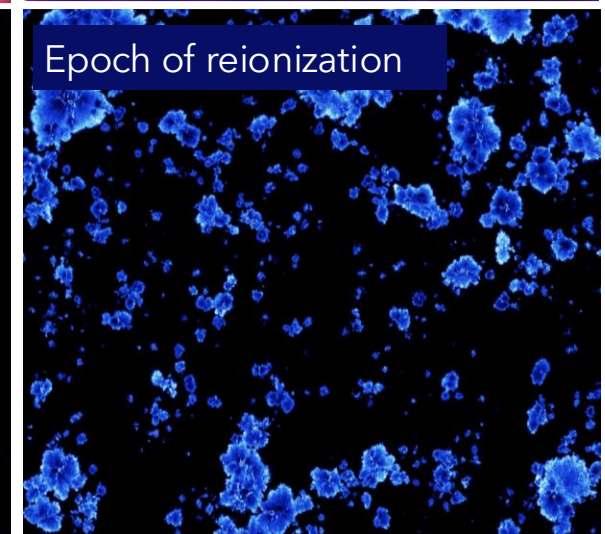
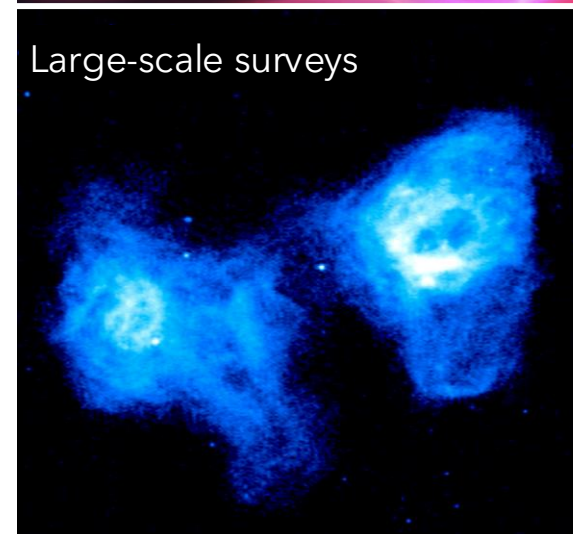
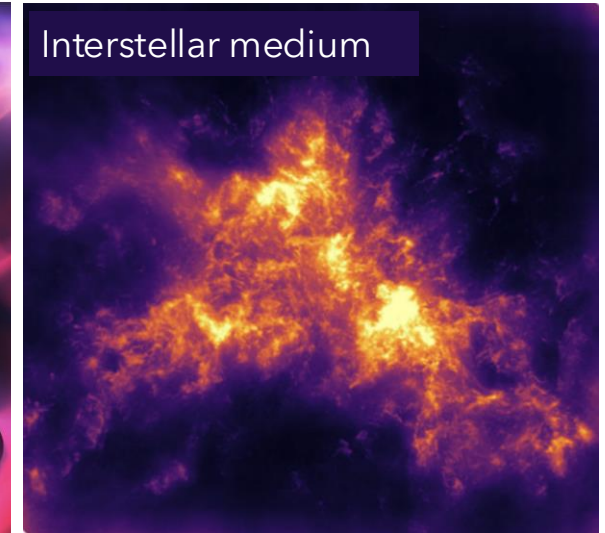
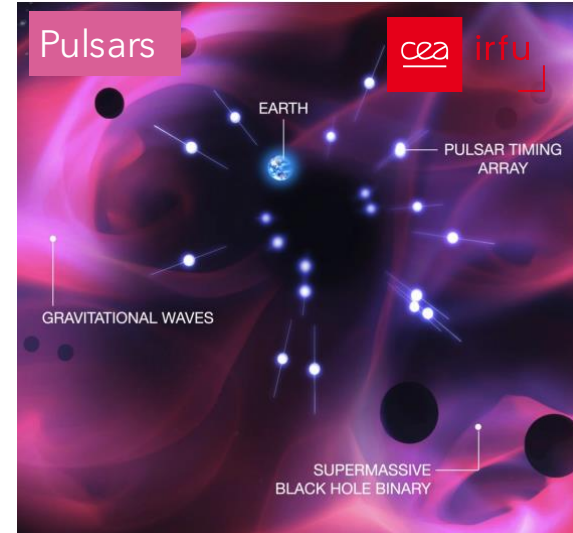
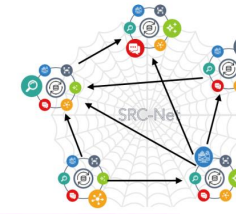
ASTRA project (2026-2030)



LPENS
LABORATOIRE DE PHYSIQUE
DE L'ÉCOLE NORMALE SUPÉRIEURE



- **Goals:** work focusing on 4 mature radio astronomy pipelines to provide deployable workflows on heterogeneous distributed architectures to those that will be implemented in SKA.
- **Method:** provide a generic framework for workflow development integrating the federation of the data storage/compute/execution services as well access to ressources (*e.g. authentication, data logistics, etc*)
- Deployment demonstrator on the first FR-SRC node



The SRCNet infrastructure



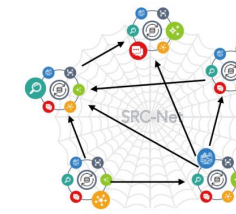
■ Major constraints :

- Exa-scale data volumes are too large to centralise.
- Computation must occur close to the data : « *Moving software to the data* ».
- Complex scientific workflows (large scale data analysis, HPC, AI).

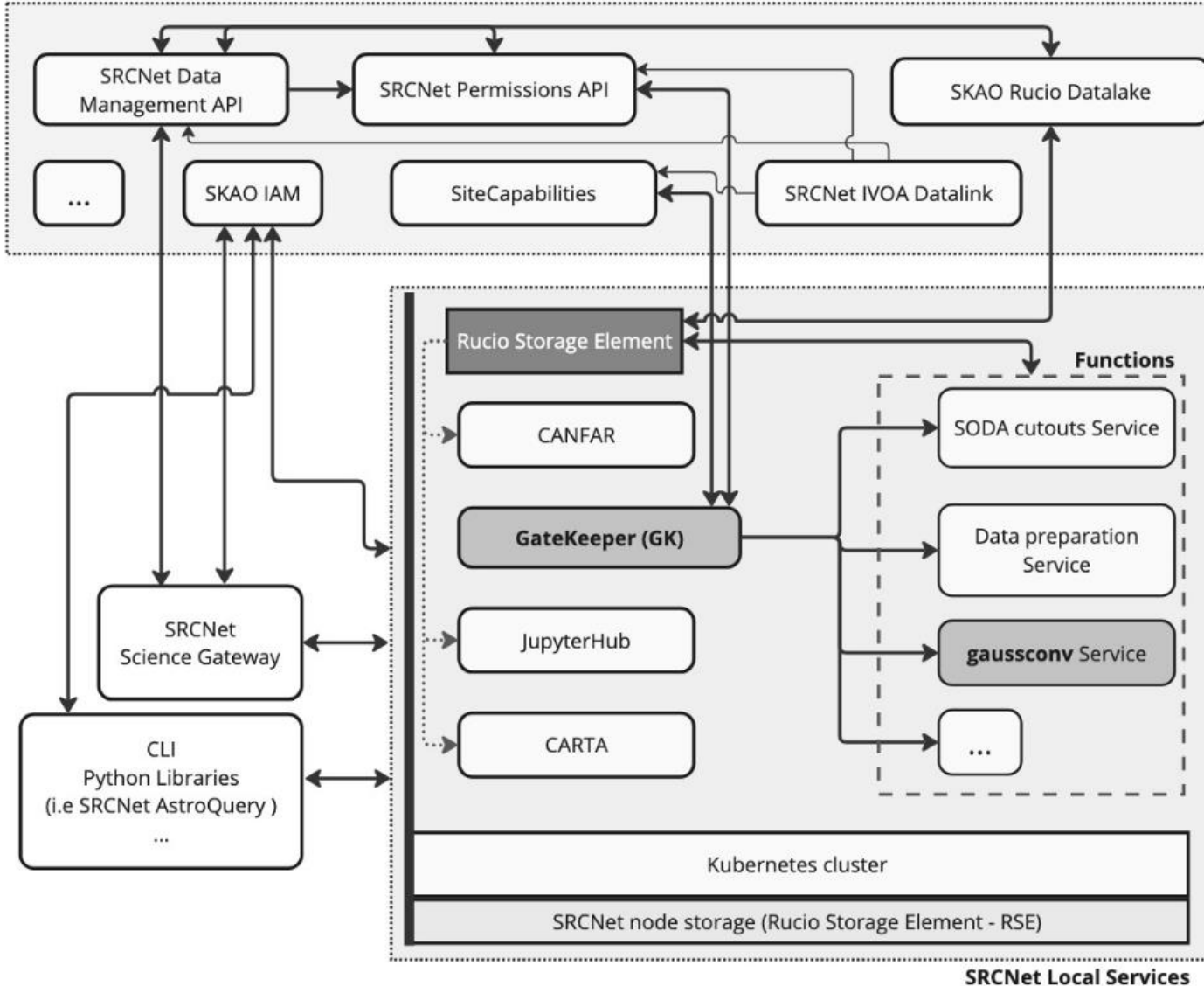
■ Infrastructure implications :

- **Distributed heterogeneous** regional centres geographically disseminated (storage, computation, data analysis) under local governance : *local SRCNets spread over 14 countries.*
- **Coordinated** through common standards (API) and global services (identity, data distribution) : *Global SRCNet.*
- **Federated** ecosystem in which both *global* and national *site-local* services must interoperate seamlessly.

The SRCNet infrastructure



SRCNet Global Services



➤ Global services :

- Identity
- Authorisation
- Data distribution
- Service catalogues / discovery

➤ Local services :

- Storage resources
- Compute resources
- Data analysis

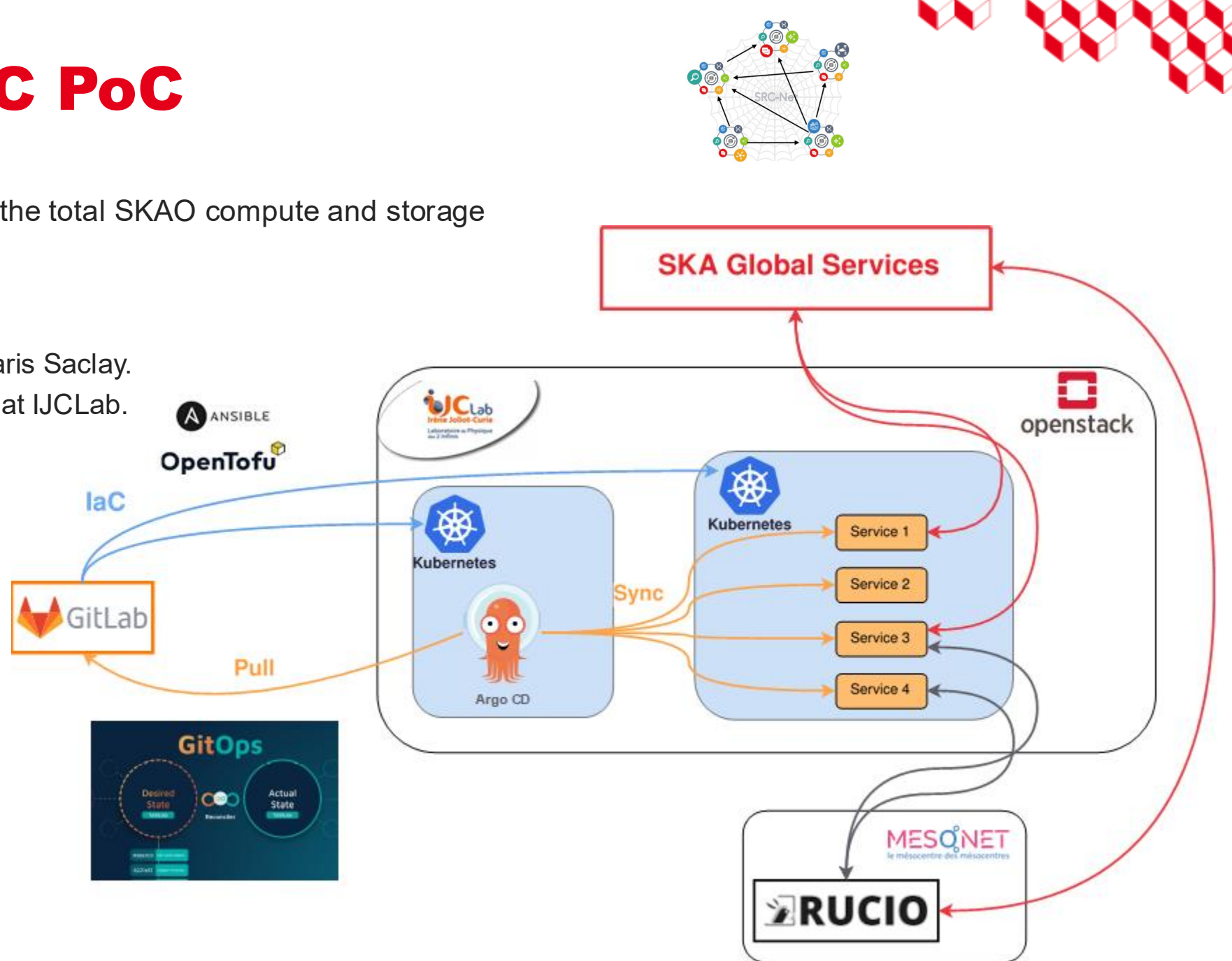
➤ Serverless computing :

- Kubernetes clusters
 - Rapid deployment
 - Scaling
 - High availability
- Microservices

x14 countries

The French SRC PoC

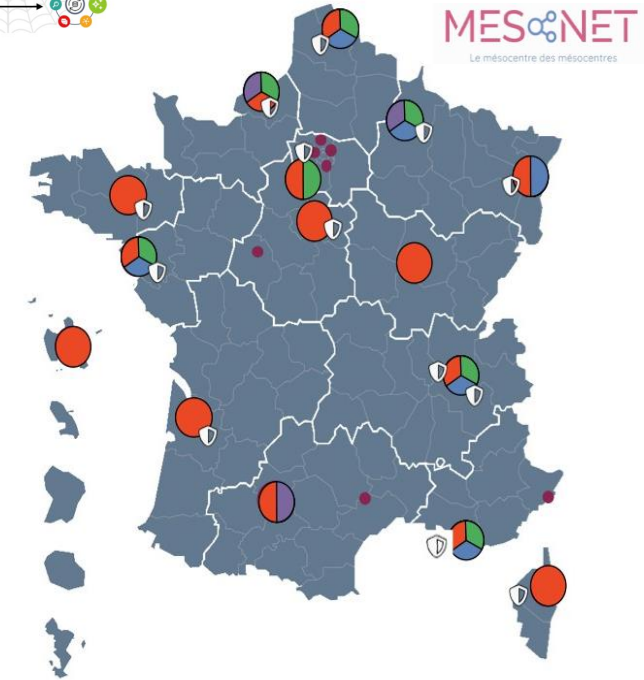
- France is engaged to host **3%** of the total SKAO compute and storage resources.
- **Proof of Concept (PoC):**
 - CEA Irfu / IJCLab / Université Paris Saclay.
 - VirtualData OpenStack platform at IJCLab.
 - Rucio storage from MesoNet.
 - GitOps approach.
 - Fully automated deployment :
 - Kubernetes clusters
 - SKA local services
 - Takes few minutes.
 - On-going work.



The French SRCNet infrastructure

Foreseen PoC scaling :

- Distributed data storage among facilities in regional centers (Mesocentres) using Rucio.
- Distributed HPC / AI workloads on :
 - existing heterogeneous infrastructures
 - Self-deployed Kubernetes clusters in regional centers clouds
- Centralised management

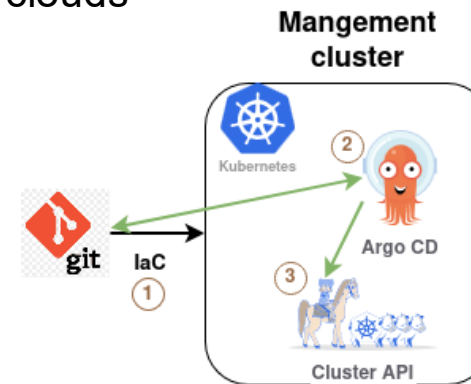


The French SRCNet infrastructure



Foreseen PoC scaling :

- Distributed data storage among facilities in regional centers using Rucio.
- Distributed HPC / AI workloads on :
 - existing heterogeneous infrastructures
 - self deployed Kubernetes clusters in regional centers clouds
- Centralised management :
 - Git repositories as source of truth

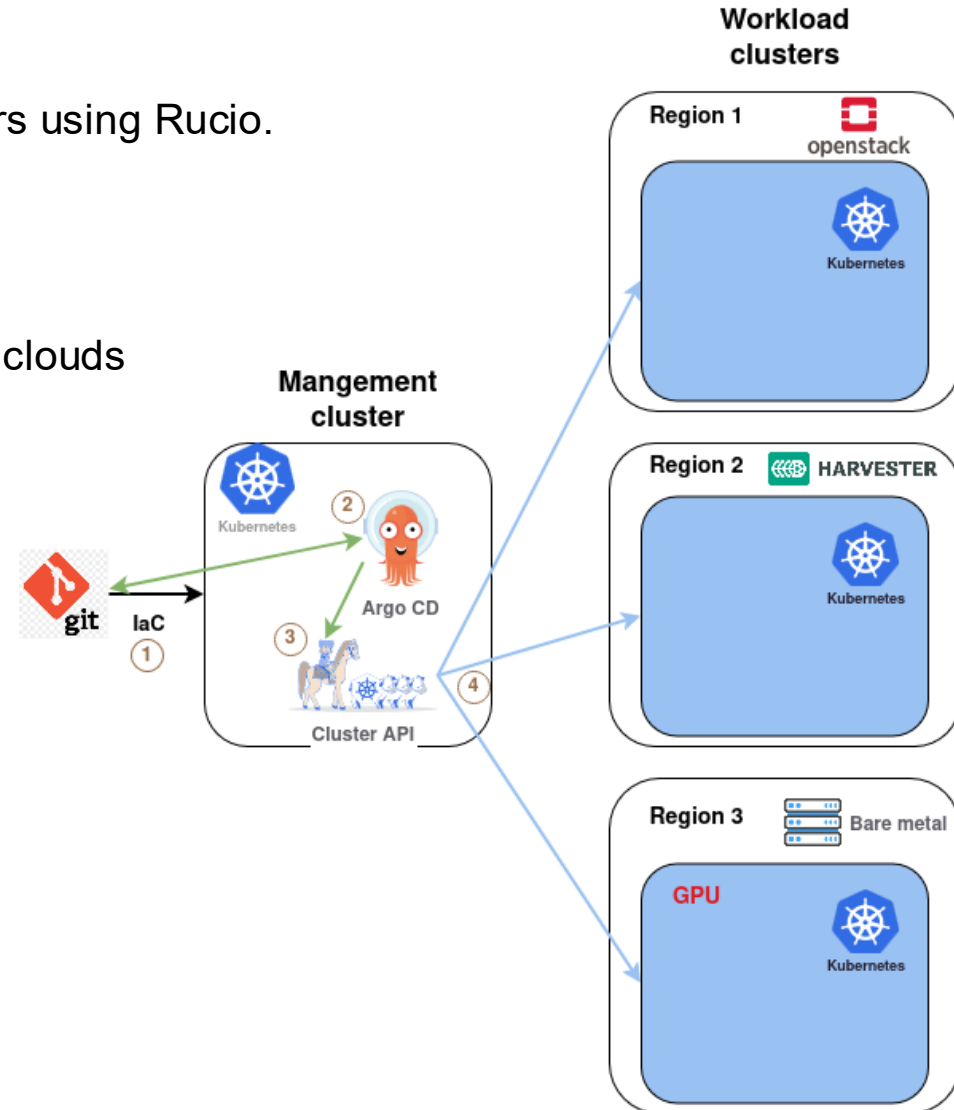


The French SRCNet infrastructure



Foreseen PoC scaling :

- Distributed data storage among facilities in regional centers using Rucio.
- Distributed HPC / AI workloads on :
 - existing heterogeneous infrastructures
 - self deployed Kubernetes clusters in regional centers clouds
- Centralised management :
 - Git repositories as source of truth
 - Kubernetes clusters : deployment automation on heterogeneous cloud providers (via Cluster API)

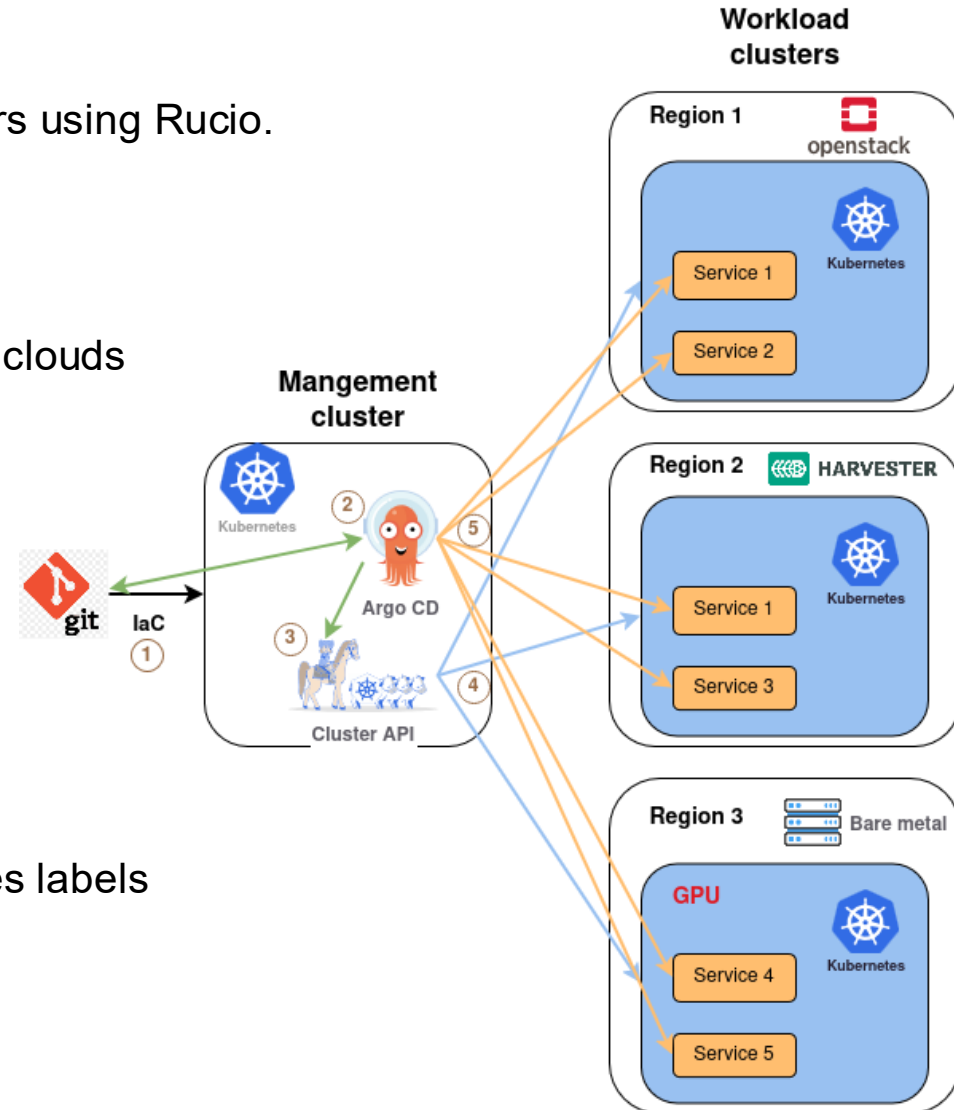


The French SRCNet infrastructure



Foreseen PoC scaling :

- Distributed data storage among facilities in regional centers using Rucio.
- Distributed HPC / AI workloads on :
 - existing heterogeneous infrastructures
 - self deployed Kubernetes clusters in regional centers clouds
- Centralised management :
 - Git repositories as source of truth
 - Kubernetes clusters : deployment automation on heterogeneous cloud providers (via Cluster API)
 - Applications / Services:
 - Deployment automation on the Kubernetes workload clusters (via Argo CD)
 - Target region specificities, like GPU, using Kubernetes labels

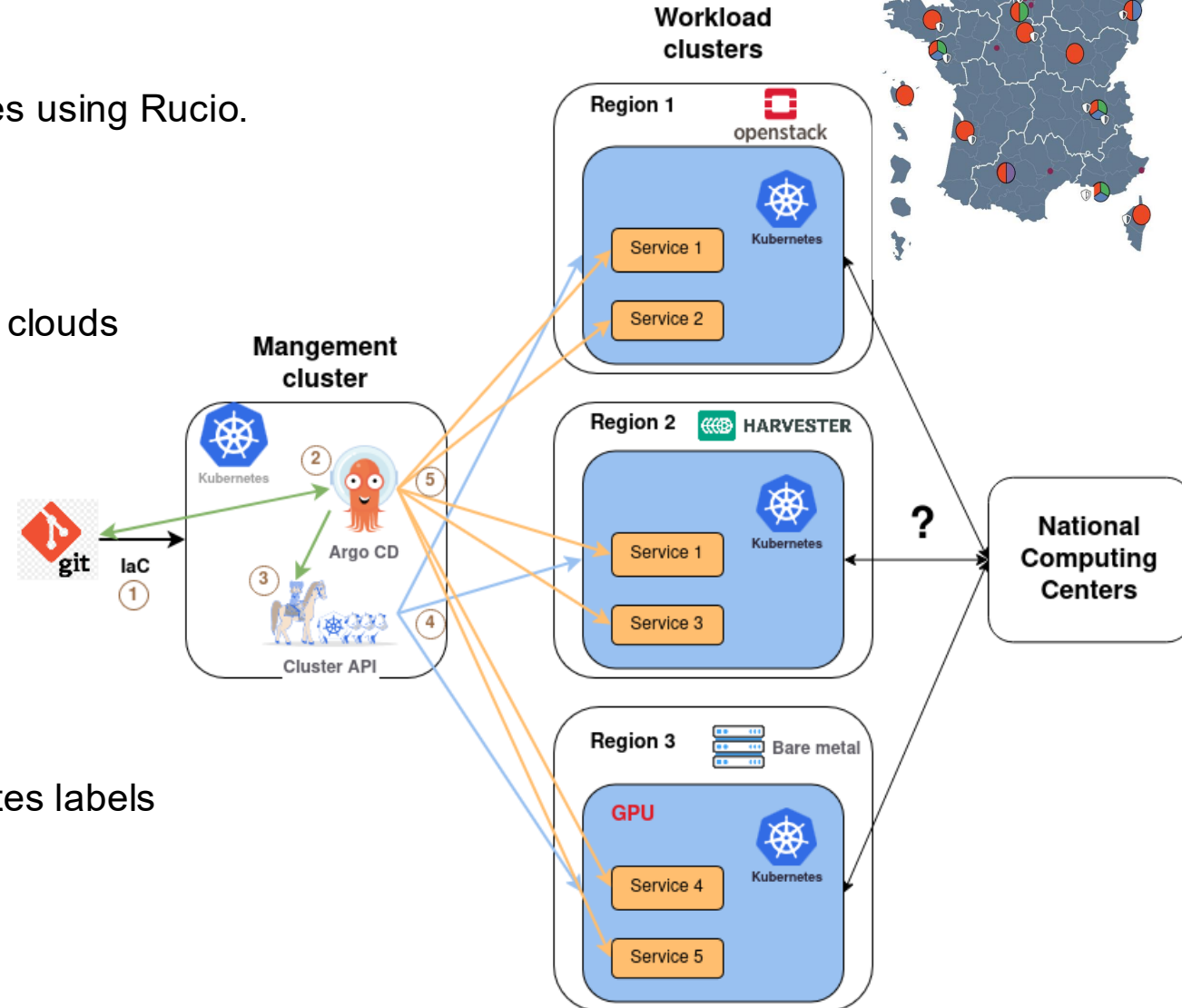


The French SRCNet infrastructure



Foreseen PoC scaling :

- Distributed data storage among facilities in regional centres using Rucio.
- Distributed HPC / AI workloads on :
 - existing heterogeneous infrastructures
 - self-deployed Kubernetes clusters in regional centres clouds
- Centralised management :
 - Git repositories as source of truth
 - Kubernetes clusters : deployment automation on heterogeneous cloud providers (via Cluster API)
 - Applications / Services:
 - Deployment automation on the Kubernetes workload clusters (via Argo CD)
 - Target region specificities, like GPU, using Kubernetes labels
 - Interface with National computing centers:
 - On-going investigation (**PanDA**, slinky, Volcano, ...)



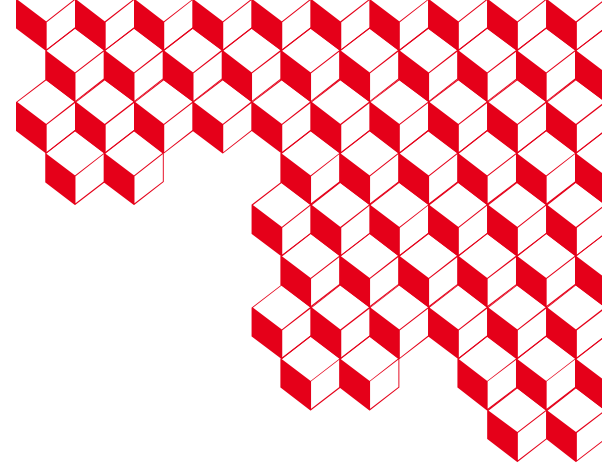
French SRCNet infrastructure feedback

- Processing and analysing such massive amount of data becomes more frequent (SKA, LHC, CTA, etc).
 - **Federation of distributed regional centers**
- Building complex infrastructure architecture requires a wide range of expertise (CI/CD, IaC Automation, distributed HPC/ AI workloads, deployment of distributed Kubernetes clusters, microservices orchestration, Federated Identity management, etc)
 - **Teams of experts**
- Analysing exa-scale data requires complex softwares and algorithms to maximise the scientific outputs.
 - **Professional developers**
- Such infrastructure facilities and skills diversities are rarely available within a single institution.
 - **Infrastructure mutualisation and teams collaboration**
- Current and future experiments will greatly beneficiate from a generic platform federating HPC, storage and cloud resources allowing execution of complex workloads on heterogeneous infrastructure.

Take-away messages

- SKA will produce a **massive amount of data** (700Pb/y), with overwhelming computing challenges
- Scientific return resides in **mastering the computing chain from the workflow design to their deployment**:
 - Developing and mastering complex (AI/HPC-based) data analysis pipelines, with scientifically validated tools
 - A change of culture to :
 - **Accelerate the development of scientific AI at scale**, which requires collaborations across communities
 - **Adopt professional development process** and tools for deployability/reproducibility/sustainability
 - **Cultivate the required technical expertise to maintain/evolve/operate a distributed compute/storage infrastructure.** (Irfu must capitalise on the experience learned from SVOM, SKA, continue close collaboration with other labs/institutions)
 - **Mandates a close relationship between the scientific teams and the computing centers (GENCI, Mesocentres)**
- These challenges are not limited to SKA : particle physics, materials, climate, geosciences, etc ...
 - **Tackling these computing challenges need collaboration at the French / European level.**

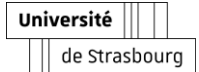
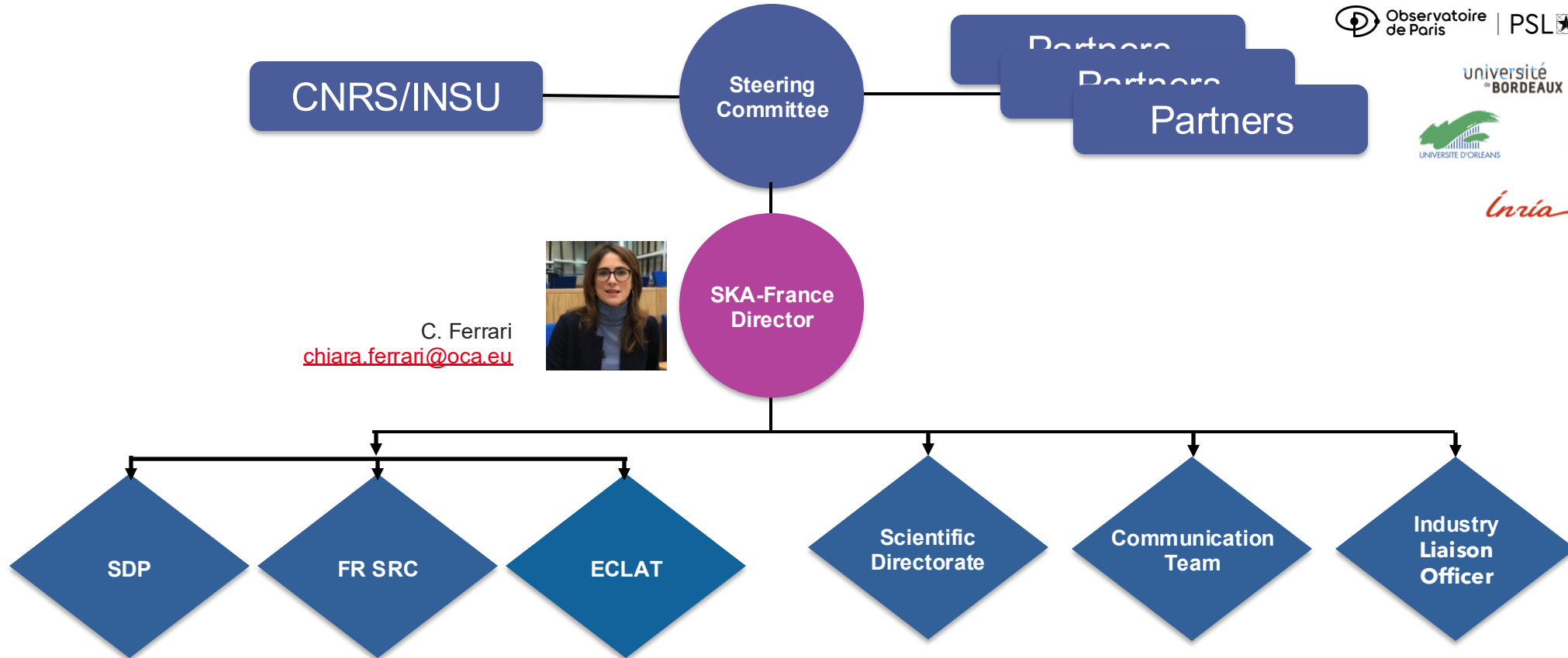
**Thank you for your
attention !**



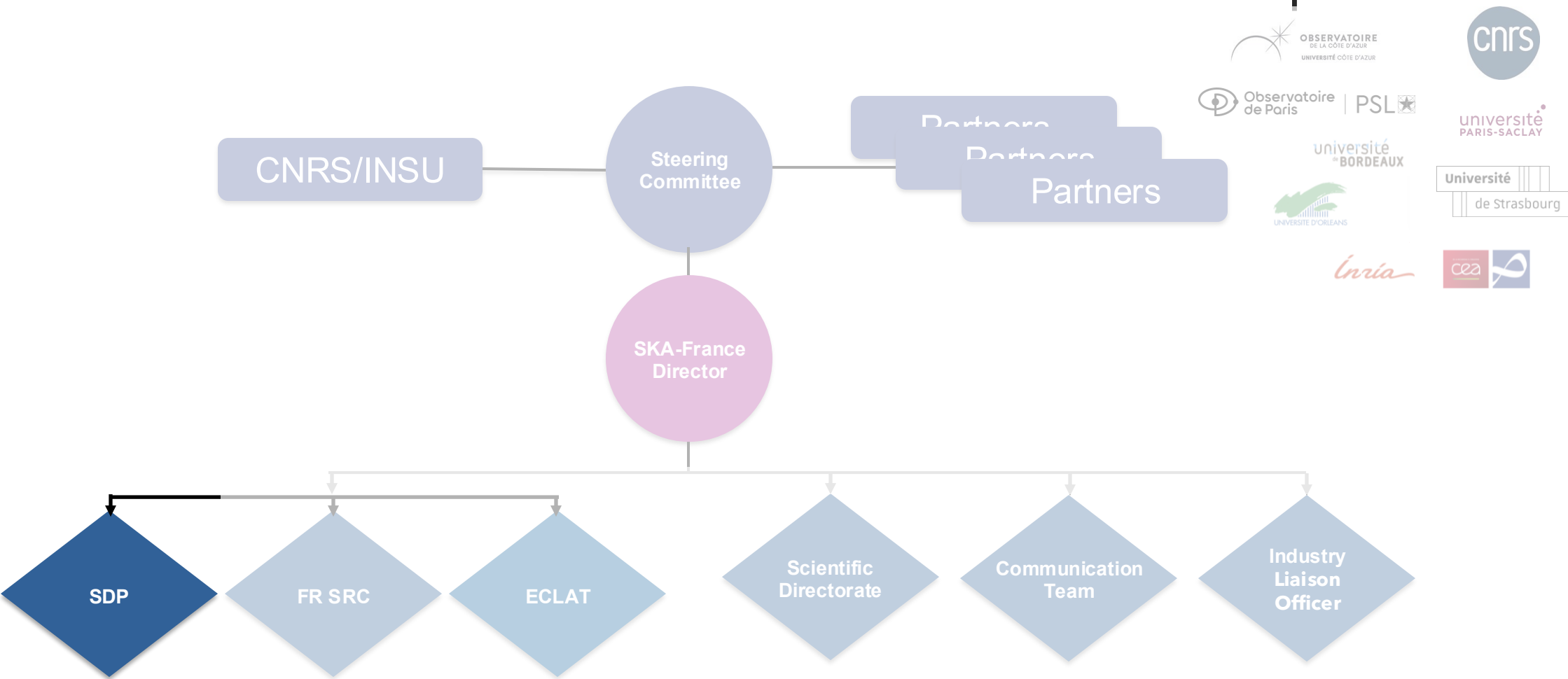


Backup

SKA-France: une coordination nationale depuis 2016



SKA-France: une coordination nationale depuis 2016



SKAO

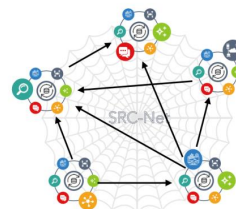
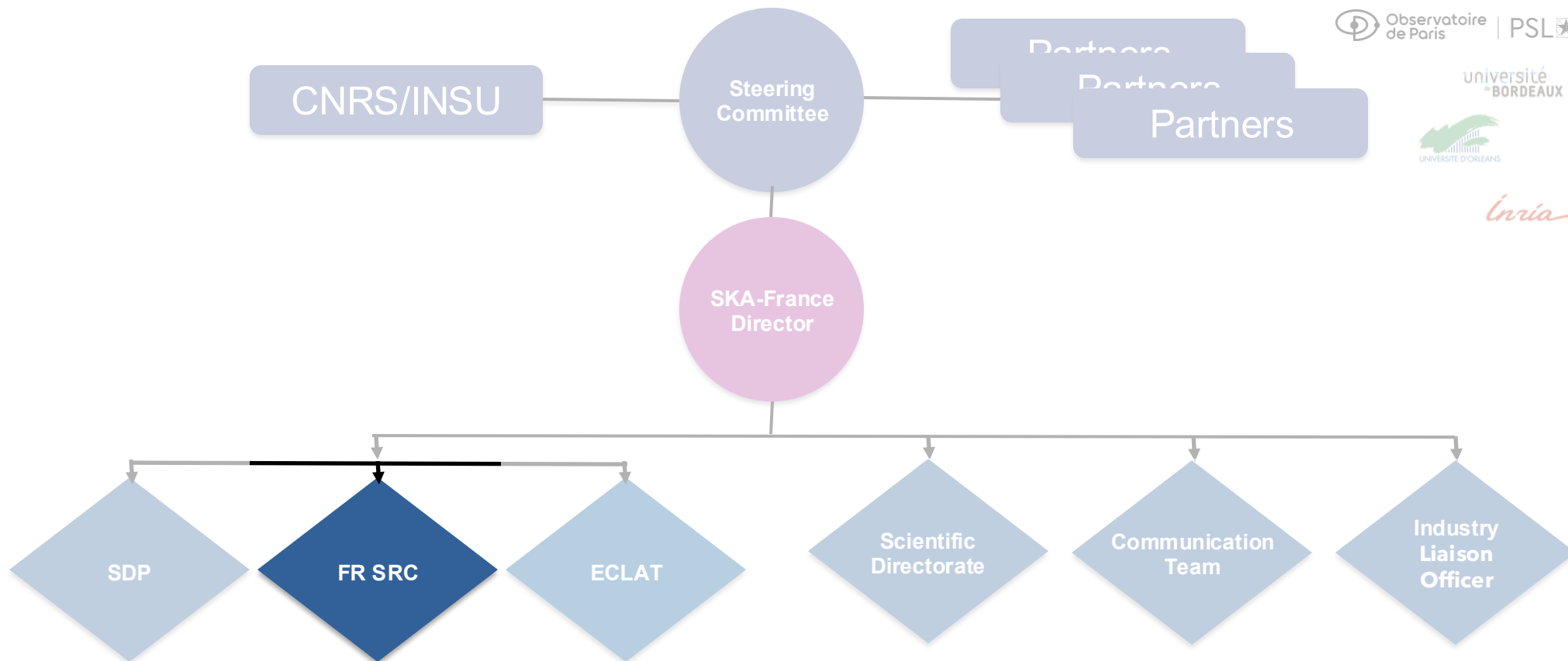


S. Mignot
shan.mignot@oca.eu



- Développement, déploiement et pilotage de workflows complexes et diversifiés, qui généreront des défis de type exascale
- Optimisation d'exécutions sous contraintes de ressources pour une "science sobre"

The French SRCNet



OBSERVATOIRE
DE LA CÔTE D'AZUR
UNIVERSITÉ CÔTE D'AZUR



Observatoire
de Paris | PSL

université
BORDEAUX

université
PARIS-SACLAY



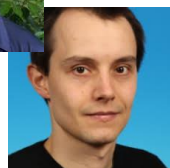
Université
de Strasbourg

Inria



J. Wagg
jeff.wagg@oca.eu

fr | SRC



D. Chapon
damien.chapon@cea.fr



université
PARIS-SACLAY



Université
de Strasbourg

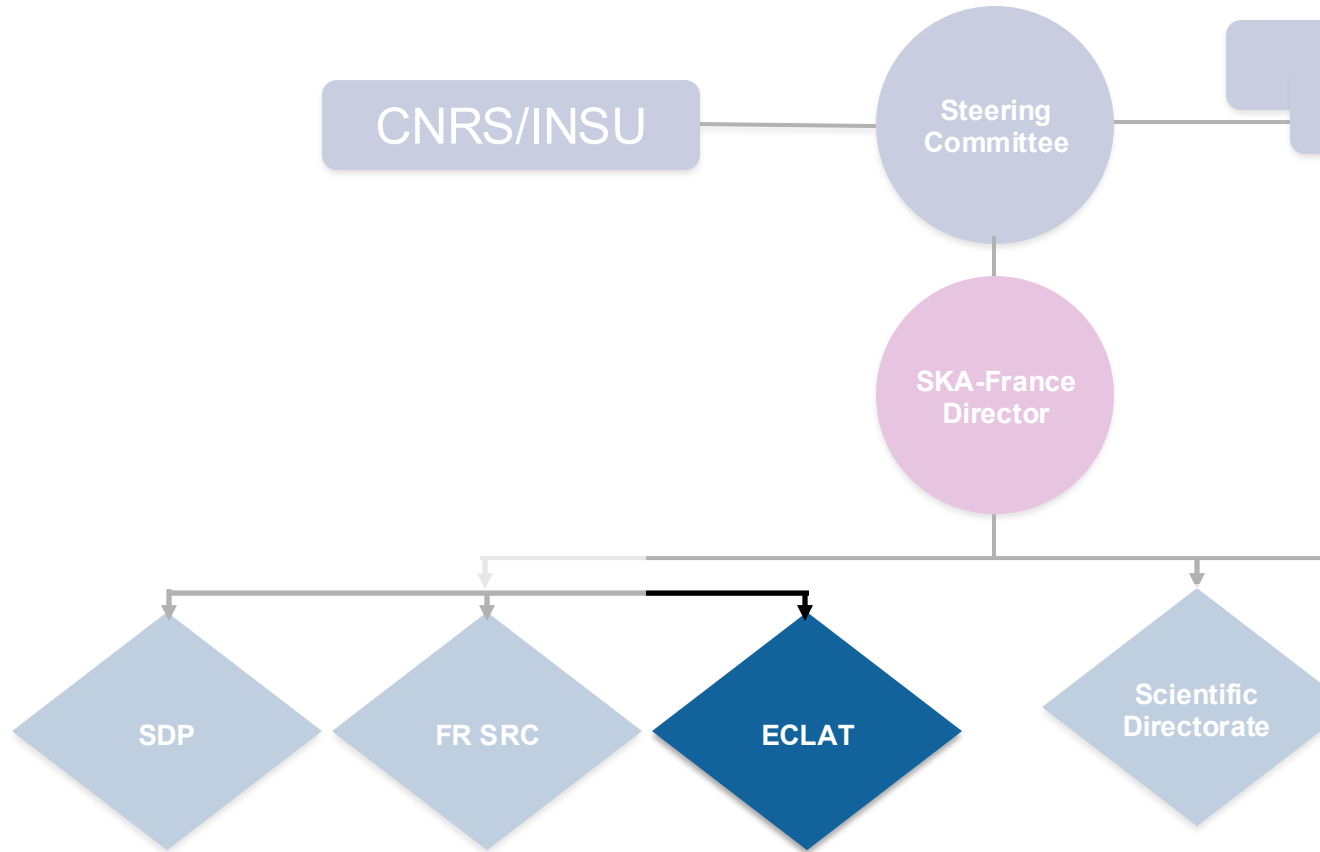
Observatoire
de Paris | PSL



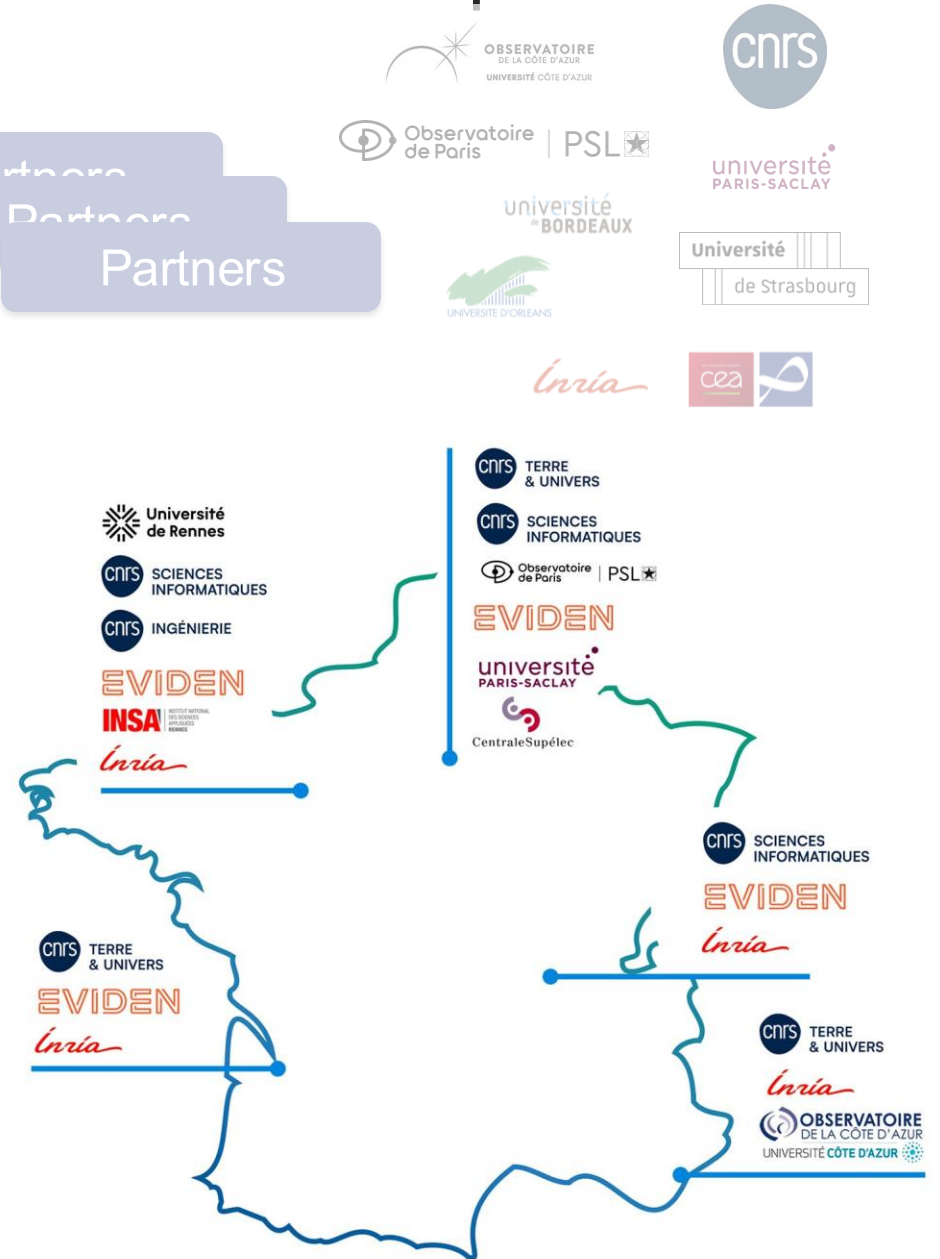
PEPR EXPLORATOIRES
APPEL À PROJETS
2024
PROGRAMME NumPEx « APPEL À PROJETS
NUMÉRIQUE POUR L'EXASCALE »

DOCUMENT PRÉSENTATION PROJET
ASTRA

SKA-France: une coordination nationale depuis 2016



D. Gratadour
damien.gratadour@obsspm.fr

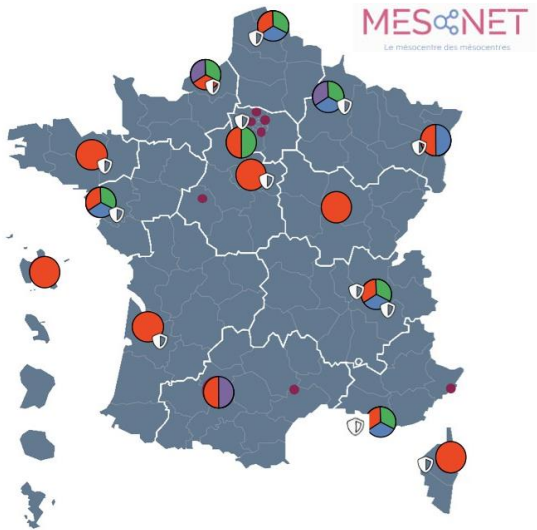


fr | SRC : liens avec les infrastructures numériques



SKA Observatory
Jodrell Bank
Lower Withington
Macclesfield SK11 9FT
United Kingdom
+44 (0)161 306 9600
www.skatelescope.org

9 February 2023



Philippe Lavocat
CEO
GENCI
6 bis rue
Auguste Vitu
75015 PARIS

Dear M. Lavocat

I write to you as Director General of SKAO in support of your proposal Jules Verne, which is an expression of interest to host and operate an Exascale supercomputer for the EuroHPC Joint Undertaking in the scope of the EUROHPC-2022-CEI-EXA-01 call.

I understand that France, through GENCI as Hosting Entity and CEA as Hosting Site, is setting up a consortium with the Netherlands (represented by SURF) for the second Exascale system in Europe. The proposal has the ambition to deploy an HPC system capable of 1 EFlops sustained performance on one partition before the end of 2025, integrating European sovereign components and made around a modular and balanced architecture (scalar and accelerated compute partitions, tiered data centric storage architecture, unified system administration, execution, and allocation...) toward the convergence of HPDC, HPDA and AI.

■ ■ ■

Yours sincerely

Prof. Philip Diamond, SKAO Director-General

MÉMORANDUM D'ENTENTE « Stratégie pour l'interopérabilité des infrastructures numériques scientifiques » Version du 24 avril 2025

Entre
La coordination SKA-France
et
projet MesoNET
et
projet PEPR Numpex
et
L'infrastructure de Recherche Data Terra
et
L'infrastructure de Recherche Climeri
et
le Très Grand Centre de calcul du CEA « TGCC »
et
L'Institut du développement et des ressources en informatique scientifique « Idris »
et
Le Grand Equipement National de Calcul Intensif « GENCI »

ci-après individuellement la « Partie » et collectivement les « Parties »