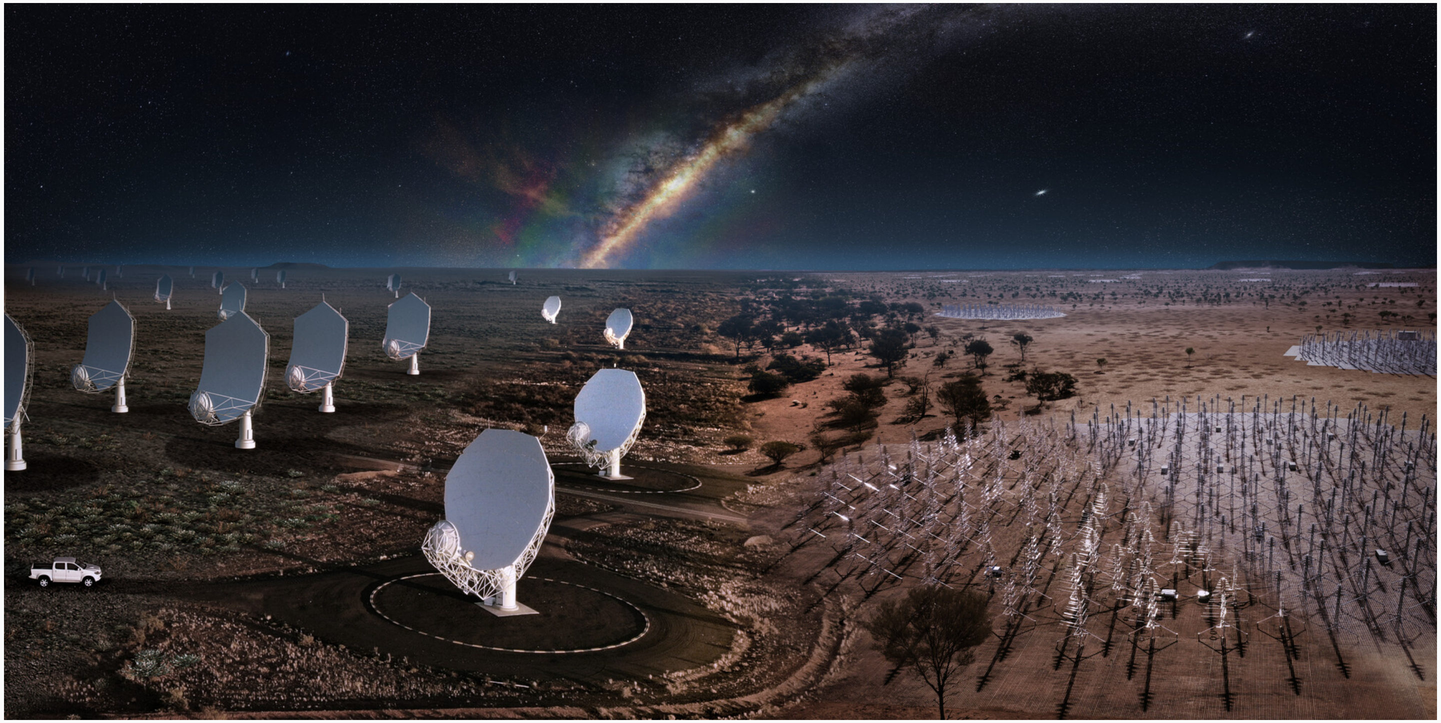


Extreme scale dataflows in the continuum for next gen giant observatories

Damien Gratadour

cnrs

l'Observatoire de Paris PSL

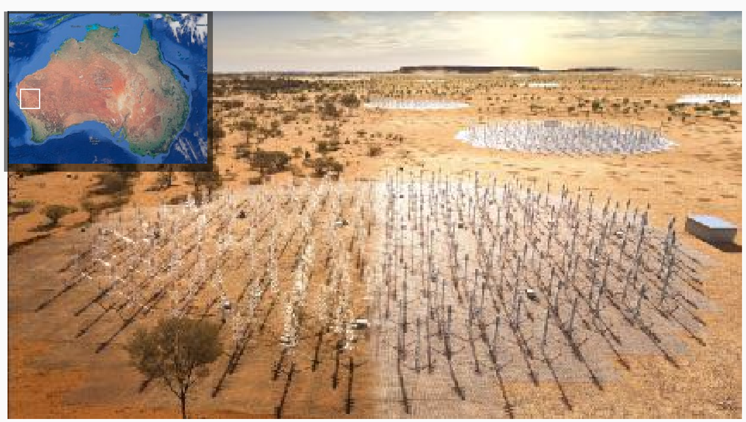
Extreme Computing Lab for Astronomical Telescopes ECLAT

SKAO: the largest (radio-)telescope

1 observatory: 2 telescopes (**Australia & South Africa**) + Headquarters (**U.K.**)

A giant software observatory, streaming data globally

50 MHz — 350 MHz — 15 GHz →

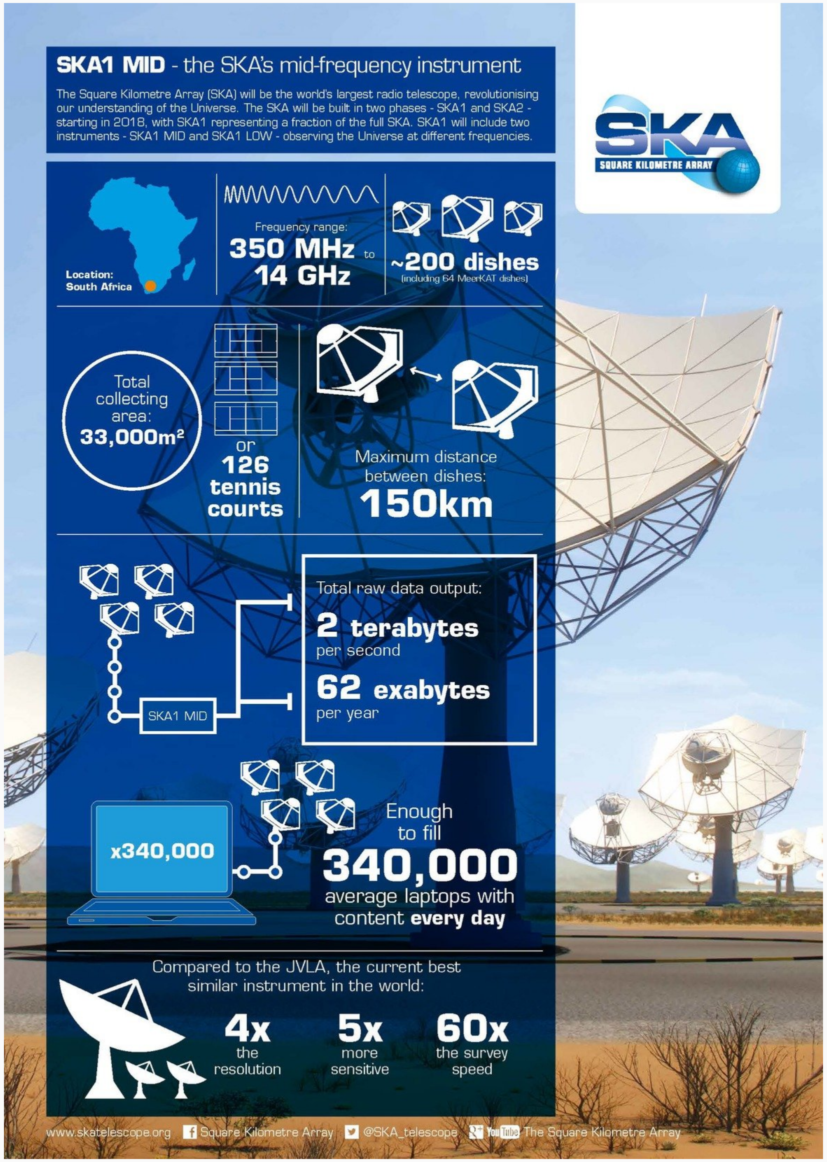


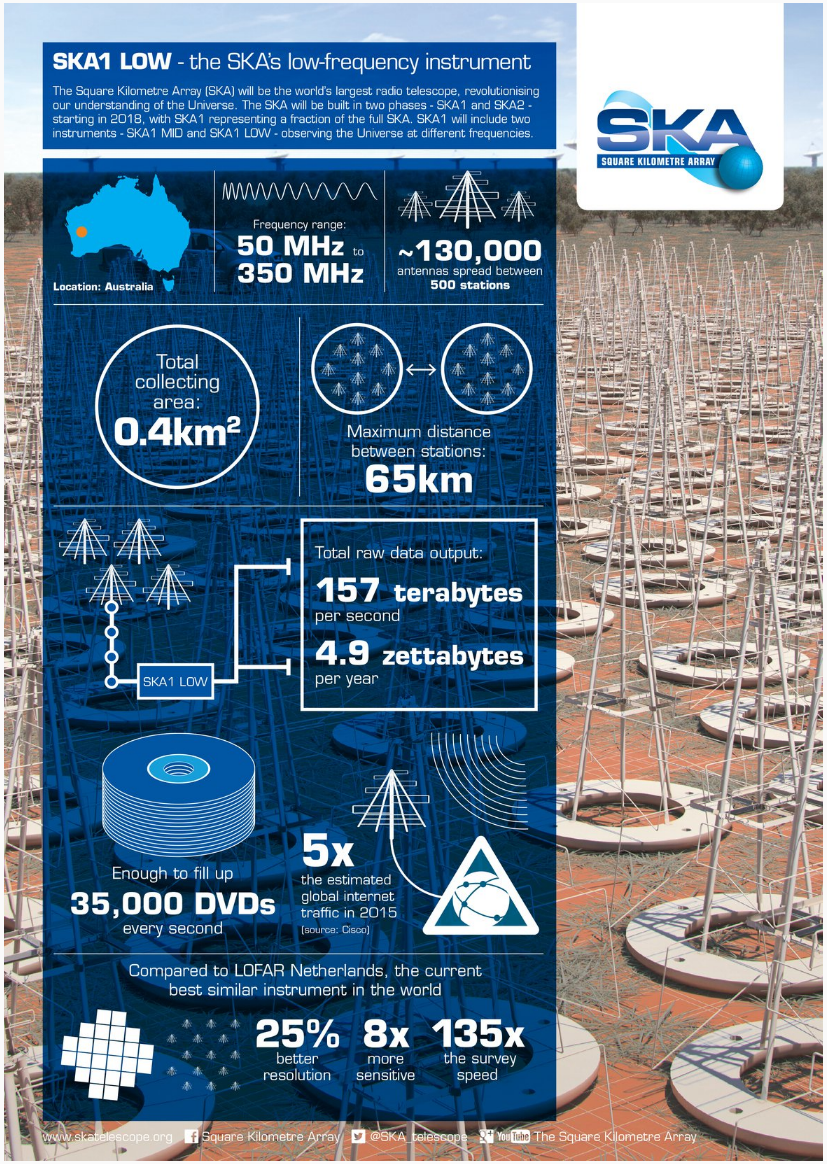
SKAO: unraveling the unknown

A truly Global infrastructure

End users communities across the globe

SKAO: Strawman view

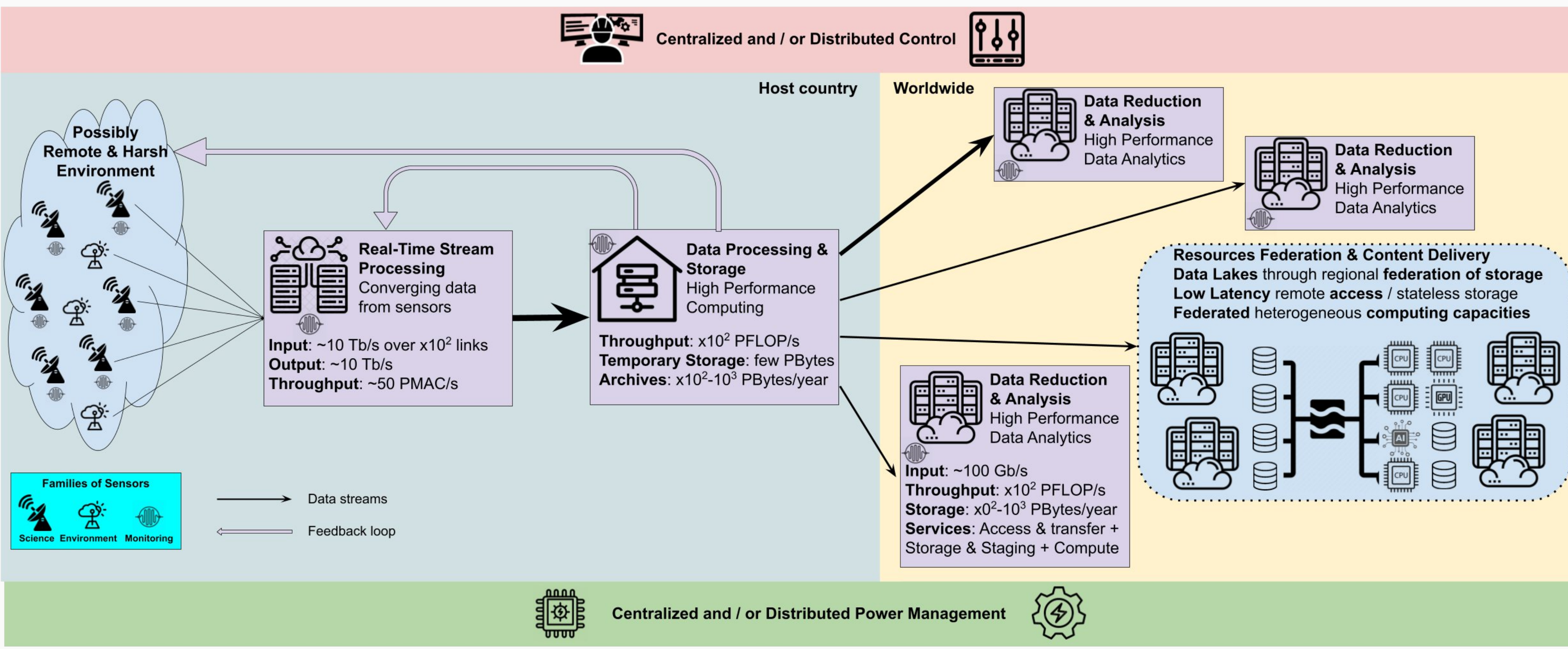




Cyber Continuum for SKA

Hierarchical architecture: system of systems

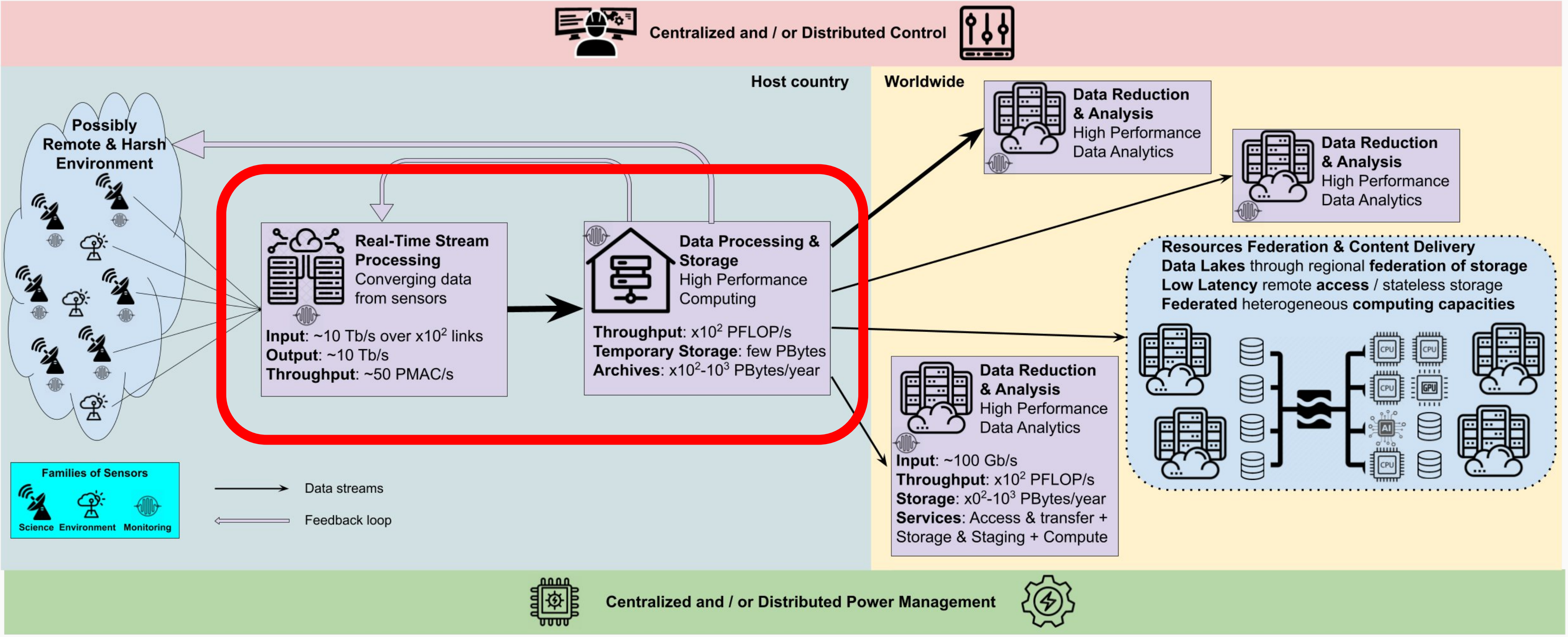
- Large amount of distributed & heterogeneous sensors
- Real-time stream engine for raw data convergence
- State-of-the-art datacenter for processing, storage and distribution
- Distributed network of national HPC facilities for content delivery to the users



Edge-to-HPC computing for SKA

Collect, Converge and Reduce data streams from distributed sensors

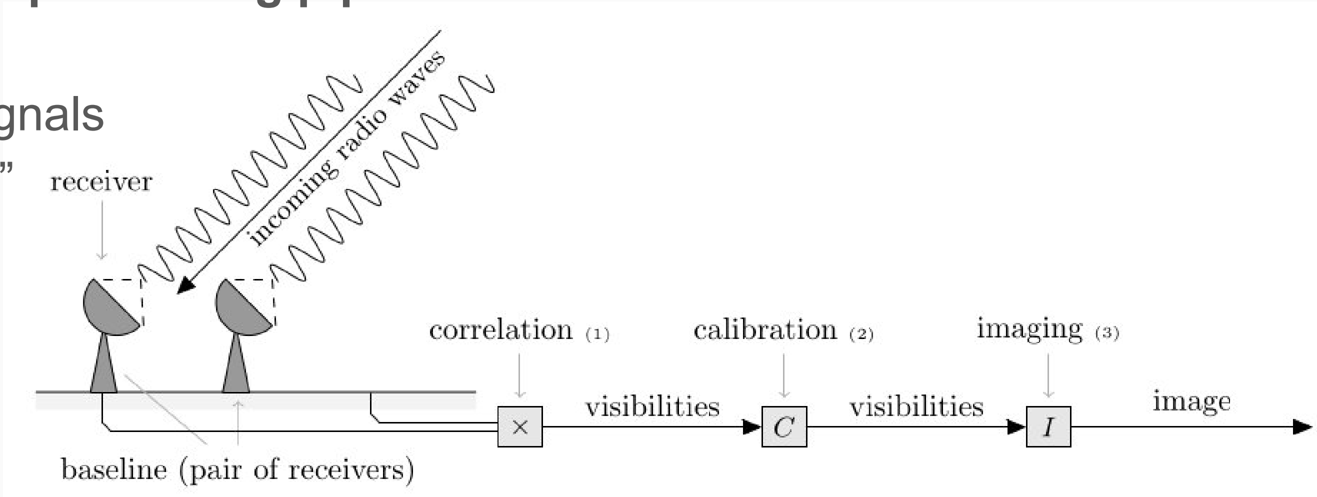
- In situ & Online data processing with centralized HPC systems
- Reduce continuous 10 Tb/s stream to 350 PB/year of data products
- Affordable / Adaptable / Frugal / Resilient
- Duplicated in two host countries (with centralized control in UK)



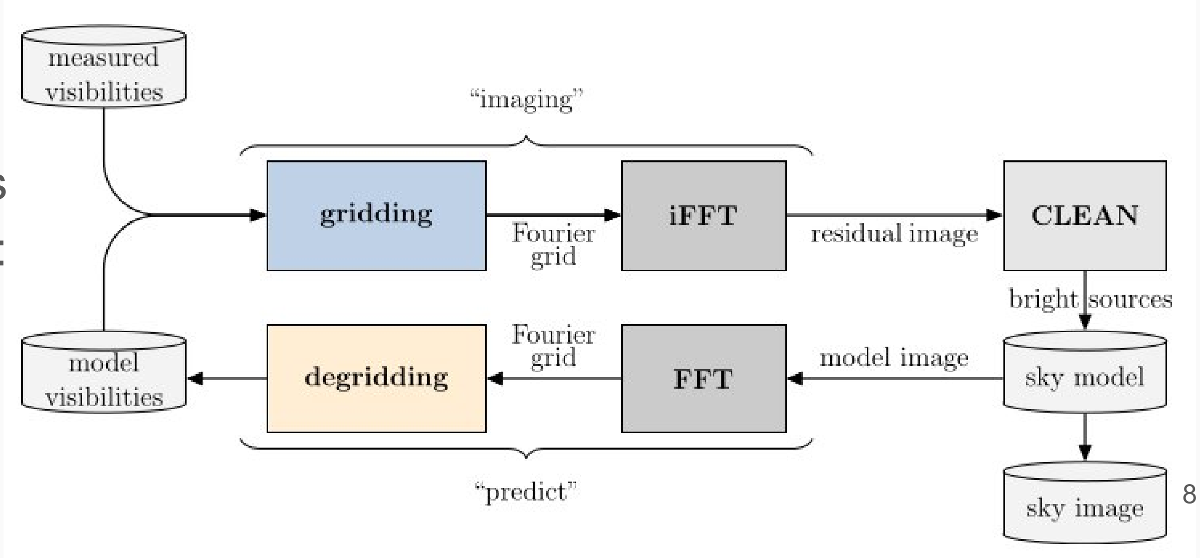
SKA computing challenges

Typical acquisition and processing pipeline

- Combine pairs of signals to create “visibilities”



- Iterative algorithm to perform image synthesis incl. calibrations + deconvolution

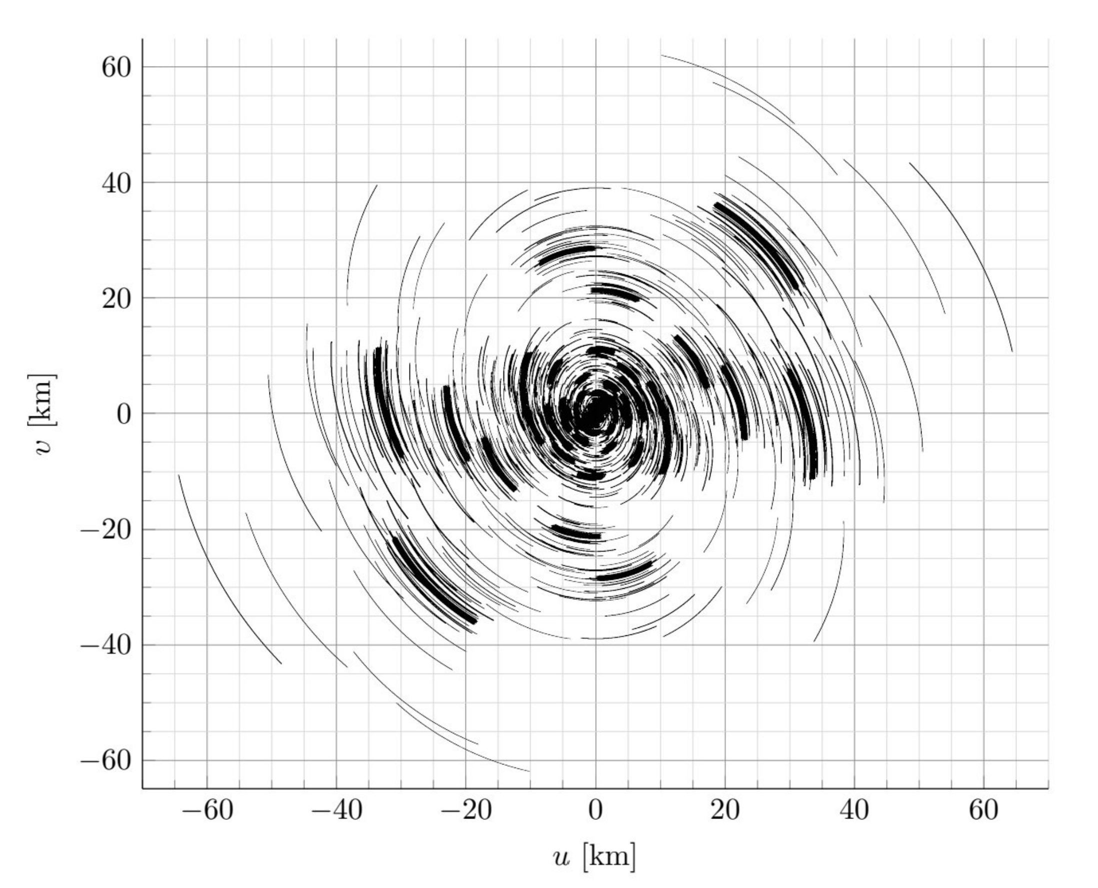


- Multiple targets (planets galaxies, cosmic dawn): variable obs. scenario = variable workload

SKA computing challenges

Imaging:

- Earth rotation synthesis: increasing frequency plane coverage over extended observation time
- Both iterative (deconvolution) and incremental (observing time)
- Trade-off between online processing (buffer based) and batch processing (storage)



Common challenges for SKA & CERN-LHC

Large collection of distributed & heterogeneous sensors

3 families of sensors:

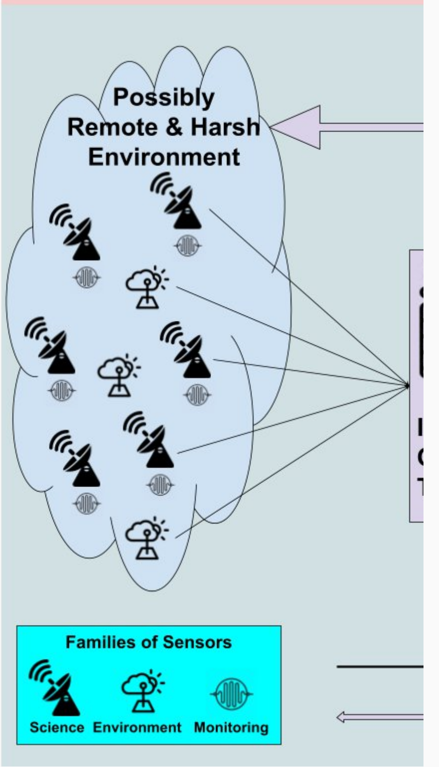
- Science: detect signals from object of interest
- Monitoring: monitor hardware (detector, components, etc..) state
- Environment: measure environmental conditions (humidity, temperature, etc..)

Distributed over large area:

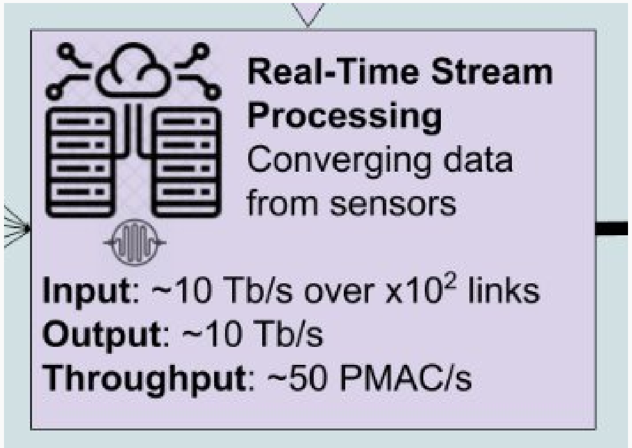
- x10^2 km2 for SKA

Data convergence challenge:

- Use monitoring and environmental sensors information for:
  - Calibration
  - Health / status monitoring & anomaly prediction
- Require a loopback system to optimize science data quality & analysis
- Embedded analogue computing & digitalization with sensors



Common challenges for SKA & CERN-LHC



Real-Time Stream Engine to produce raw data

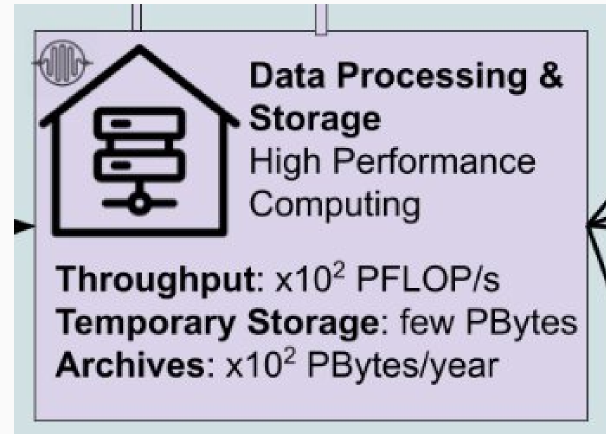
Data frontend: converge signals from x10^2 to 10^3 links

- ~10Tb/s to 1 PB/s of data as input
- I/O bound computing (~50 PMAC/s)
- **Aggregation, convergence, filtering & selection of data**
- **No reduction & storage for SKA** and 10^3 reduction for CERN, as output

Constrained environment (and remote for SKA):

- Energy consumption
- Complexity & Cost
- Reliability & Maintenance

Common challenges for SKA & CERN-LHC



Raw Data Processing, Storage & Distribution

- Data intensive compute facility (10 Tb/s of input data + pre-Exascale Throughput)

Hierarchical storage strategy

- R/W to storage at high bandwidth
- Large storage capacity (x10^2 - 10^3 PBytes / year)
- Short term buffers versus data archives

Heterogeneous workloads: from I/O bound to compute bound

- In the case of SKA: Low arithmetic intensity, Iterative process
- In the case of LHC: Compute intensive steps for reconstruction and IO bound steps for analysis

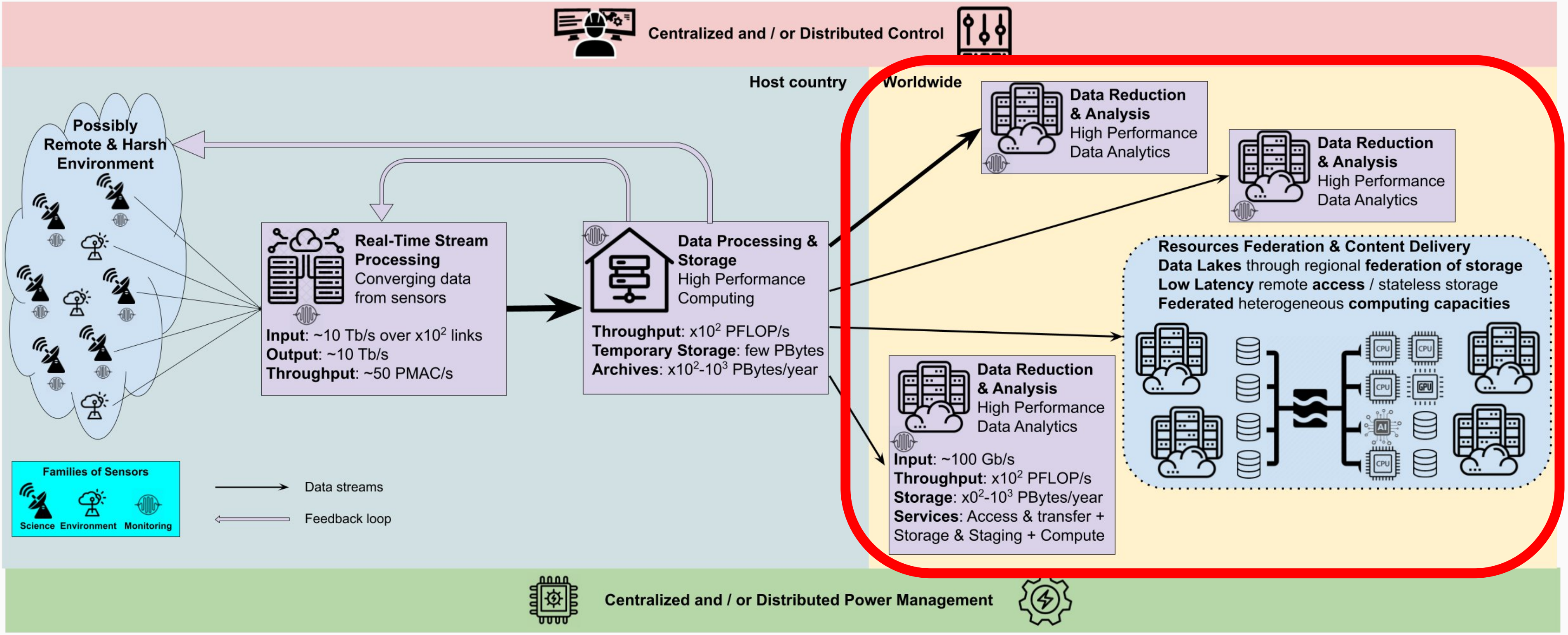
Output data products distributed globally

- Worldwide multicast over 100Gb/s links

HPC-to-Cloud computing for SKA

Federate resources to analyze distributed data

- Rely on external resources (regional centers), possibly at continental level
- Federate: compute, data logistics, storage, wide-area workflows
- Increasing use of AI for many science programs
- Access patterns, provenance, resources accounting, power management

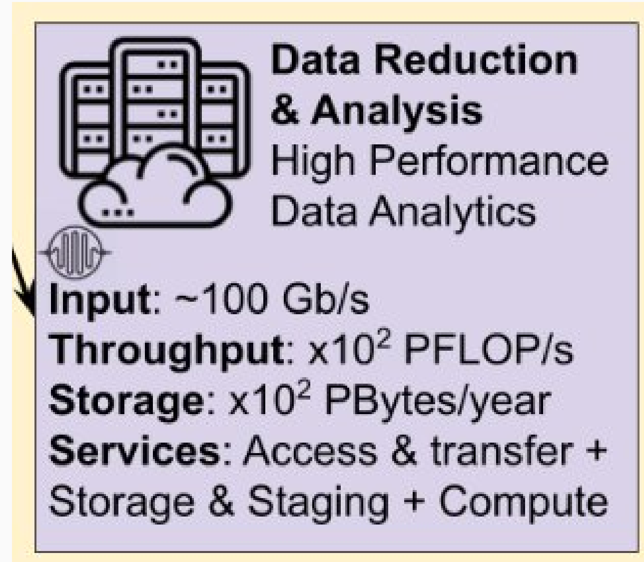


Common challenges for SKA & CERN-LHC

Local Facilities for Data Reduction

Relying on “external” facilities for analysis work

- National or Regional facilities
- No control on design and operations
- Throughput needs depend on workloads. Typical requirement is x10^2 PFLOP/s
- Accounting of resources usage



Handling massive amounts of data on shared facility

- Standardized strategy for data handling (access & transfer together with storage & staging)

Workloads need to run on a variety of heterogeneous environments

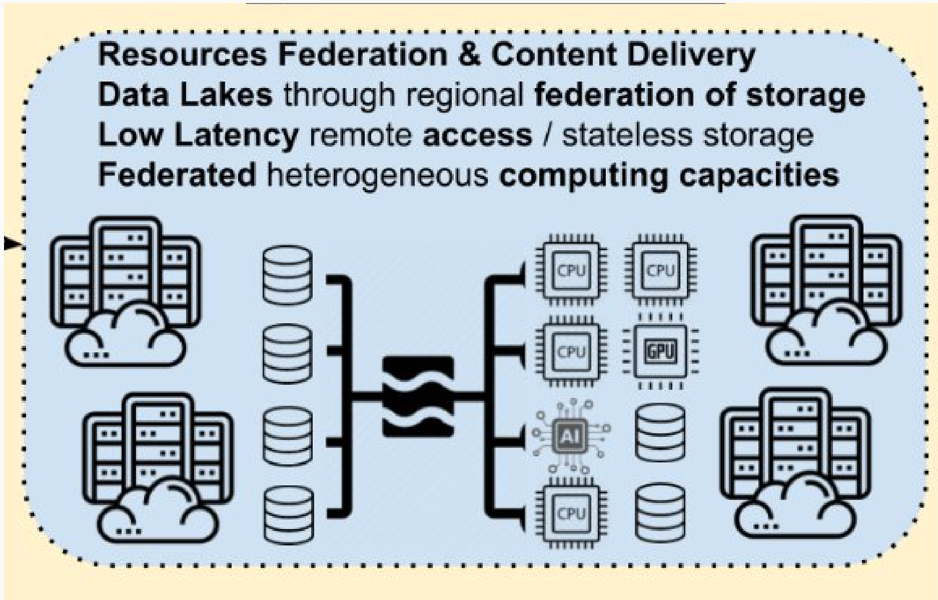
- Portability is key
- Ability to optimize performance on a variety of environments

Common challenges for SKA & CERN-LHC

Regional Facilities for Data Reduction

Federated computing to process data lakes

- Provenance (workflow and data)
- Global accounting
- Resources preemption strategy



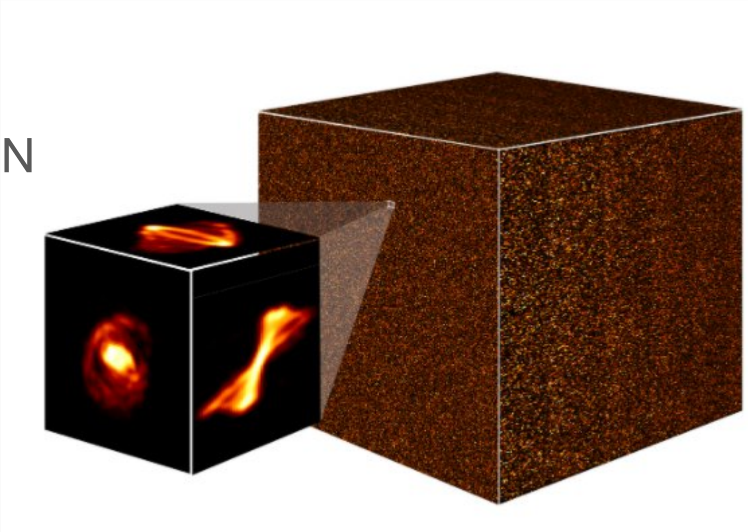
Content delivery

- Remote access and latency hiding
- Secure access

Resources federation

- Services, runtimes and algorithms for federated & heterogeneous computing

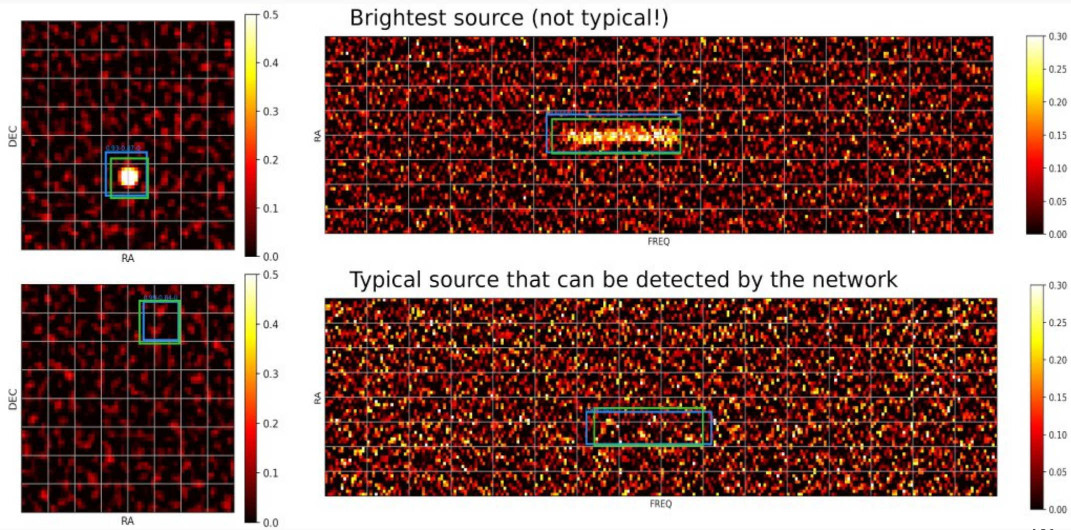
HPC-to-Cloud computing for SKA



A French Success: SKA Data Challenge 2

- Source detection and classification using DNN
- Merging catalogs produced by 2 models
- ~1TB of data,: 25 GPU hours on Jean Zay
- Score: sensitivity + precision

- copyright Photothèque CNRS/Cyril Frésillon



Challenges across the continuum

Facilities operations

- Multiscale **system of systems**
- **Intercontinental control** strategies
  - Including “owned” and “shared” facilities
- **x10 years typical lifetime**
  - Continuous integration of emerging & non-conventional technologies
  - Preserve operations

Facilities management

- **Limited power envelope**
  - Access to power grid
- **Cost containment**
  - Mostly relying on taxpayers money
- **Optimized operations**
  - Dynamical cyberinfrastructure, including reconfigurable HPC

What sustainable means ?

“meeting the demand of current generation without putting the demands of future generations at stake”

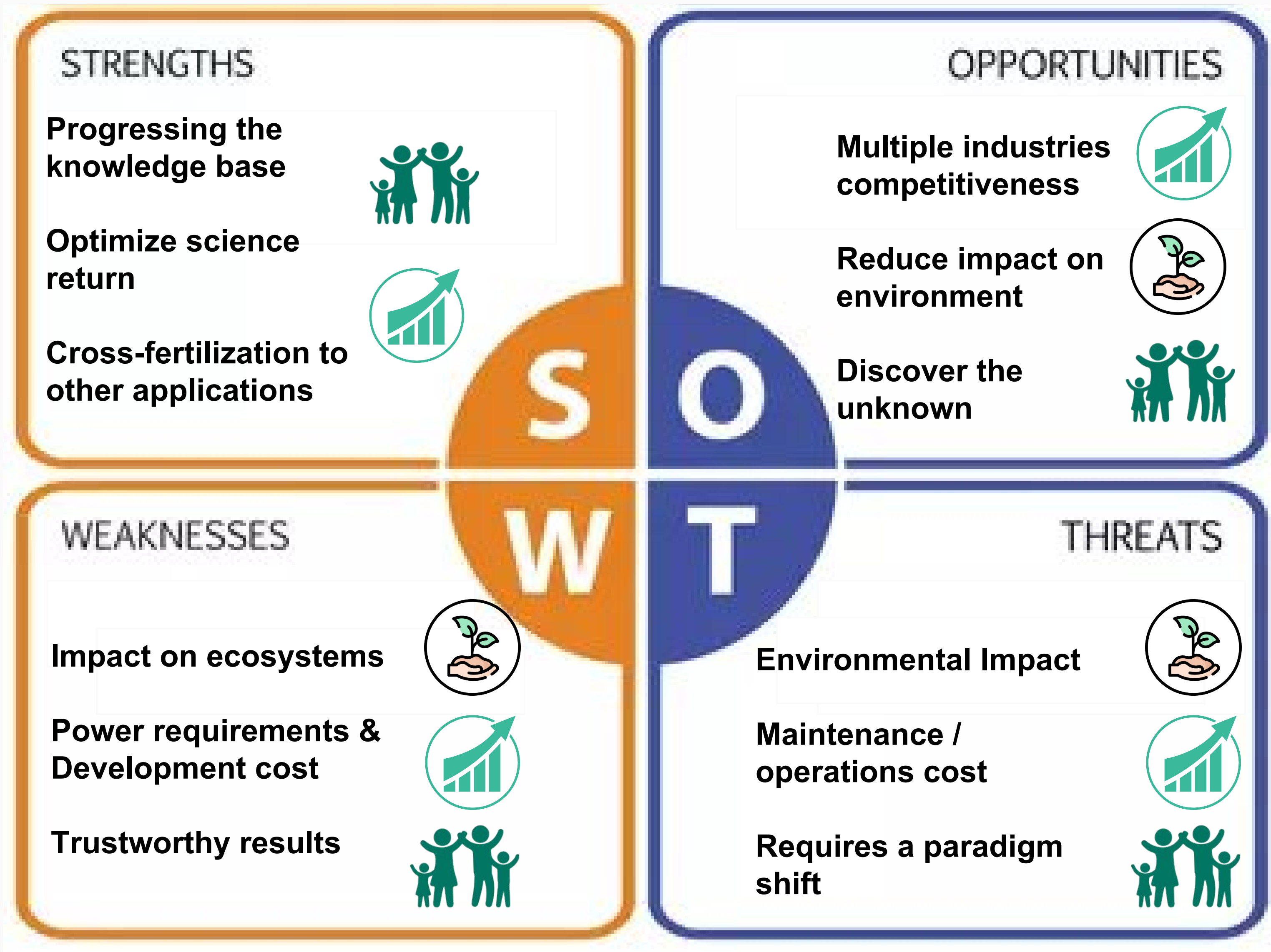
Can be analyzed as the convergence between:

- **Economic**
- **Social**
- **Environmental**

ICT developments can:

- make contributions on all fronts
- come with negative impacts

A SWOT analysis on sustainability



Horizontal challenges, addressed sustainably

SKA is representative of a new class of global infrastructures

- Cover the full edge-to-HPC–to-cloud continuum, with strong inter-dependencies
  - **Interoperability** (functional + non-functional) is a key requirement
- “Operational” cyber-infrastructure (as opposed to digital twins)
  - **R.A.S** (Reliability, Availability, Serviceability) of all components is mandatory
  - **High efficiency** of all components is required

Need to address the continuum as a whole

- Improving one aspect (e.g. energy efficiency) of one component (e.g. HPC) impacts others (e.g. sensors / edge)
- Covering many domains: access to energy, TCO, predictive maintenance, etc..
- Harnessing many technologies: AI, cyber-security, HPC, big data, etc ...

Let’s do it sustainably !

All aspects of sustainability represent both opportunities and challenges

- Close partnerships with industry
- Maximize positive societal impact
- Minimize environmental impact

The pathway to sustainability

- Tackling inter-dependencies between 3 dimensions:
  - **Societal**: Enhance access to affordable services and knowledge
    - Cross-fertilization across application domains

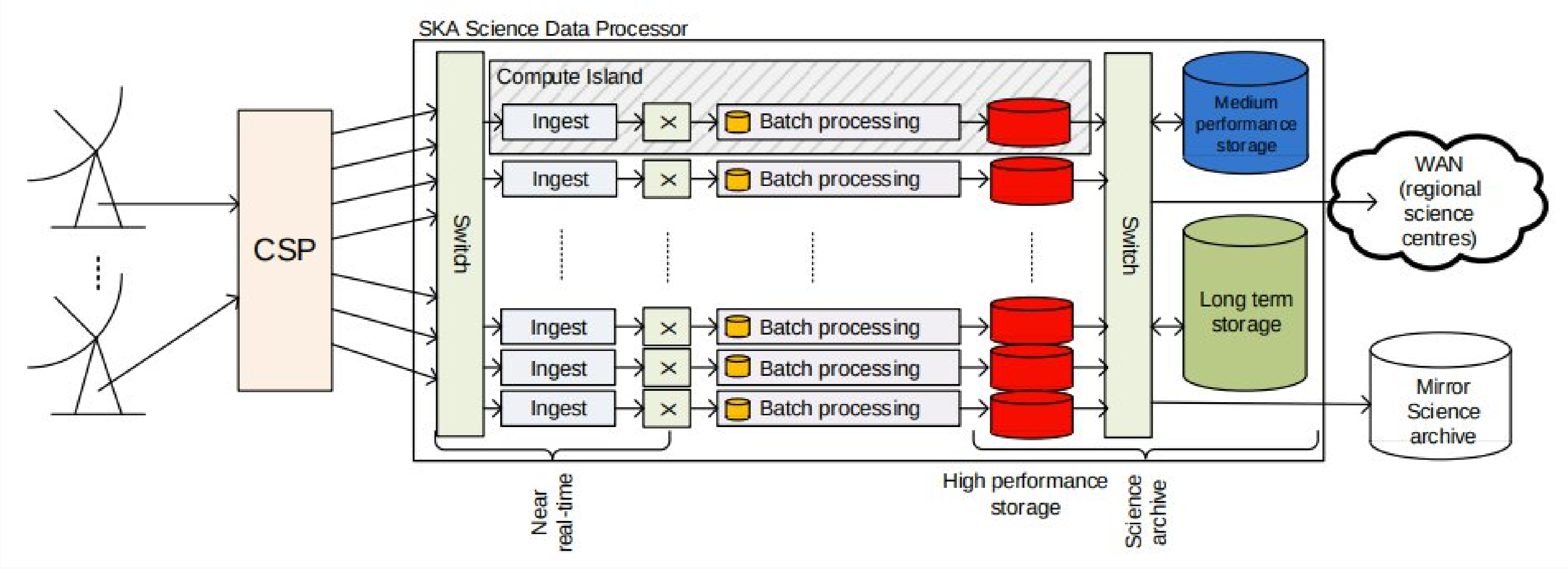
  - **Environmental**: Global Energy efficiency, across the continuum
    - Key responsibility: reduce impact of major infrastructures
    - Operational constraints: Maintain operations (thus European leadership) over long lifespan

  - **Economic**: joint public-private development of sovereign technologies
    - Strengthen industry across the continuum: increase business opportunities, create jobs
    - Reduce development & maintenance cost

SKA SPC co-design

Co-design, development and validation of building blocks

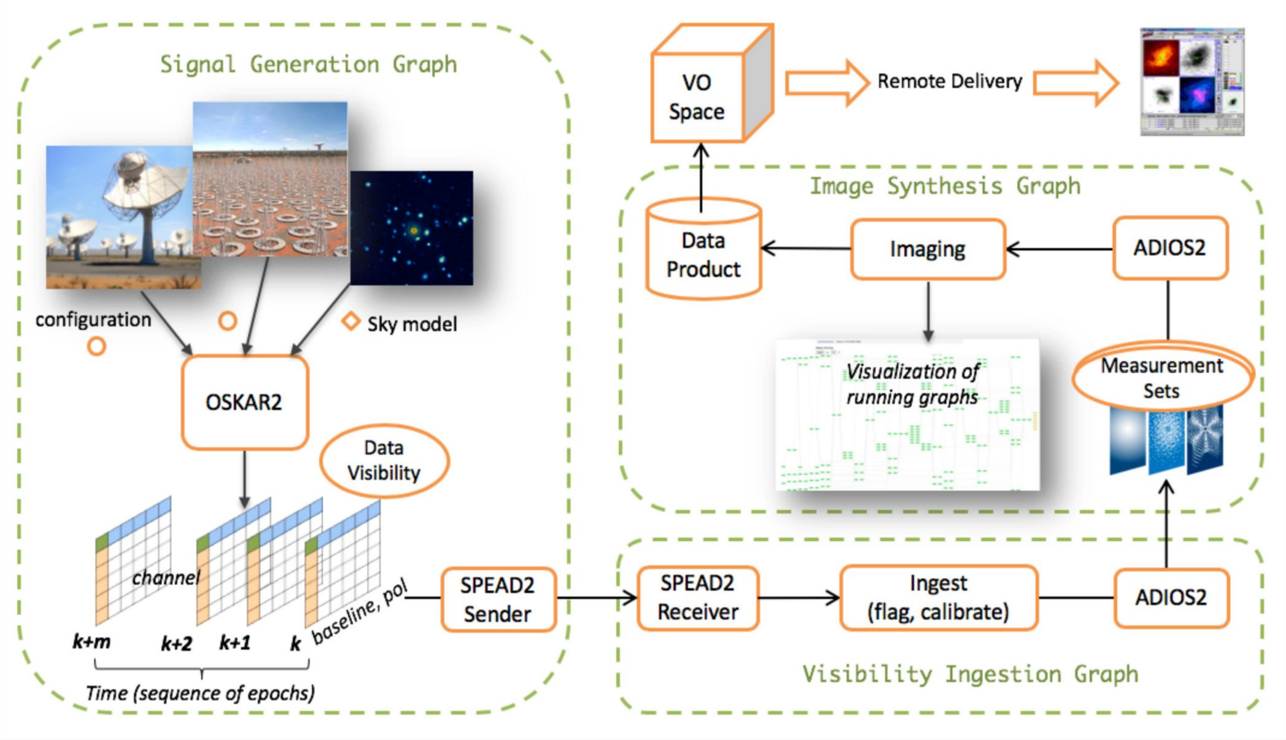
- In-situ / online stream processing & data pressure management
  - Incl. Multiple concurrent workflows (fast + low res versus HiFi image synthesis) in iterative + incremental pipeline
- Coupling DNN & large scale HPC workflows
- Data life cycle management (provenance & layout) as core topic (interoperability with regional centers component)



SKA SPC co-design

Demonstrating the SKA online workflow at full scale

- Addressing Edge-to-HPC component
- Going beyond Gordon Bell finalist study from Wang et al. 2020
- From data generation to Scientific Data Products
- Including **frugality and resilience** as sustainability indicators



Supporting initiatives

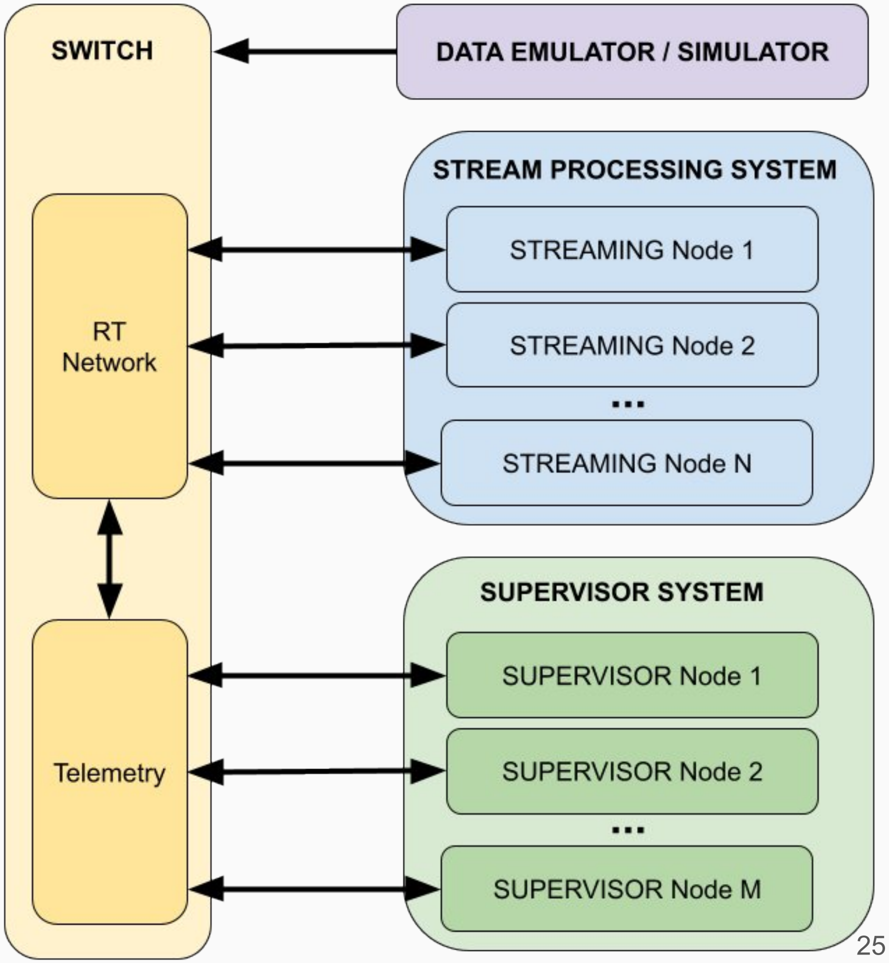
ECLAT: ***Laboratoire Commun*** **CNRS-Inria-Atos** (head: D. Gratadour)

- Support structure for French contribution to SKA
- In kind contributions from partners, incl. **INSU, INS2I, INSIS**, multiple **Inria** teams together with **Atos R&D** and business dev.
- Truly multi-disciplinary and trans-sectoral collaboration

Supporting initiatives

STREAMS: a continuous integration platform

- ~1M€ hardware budget
- built around x**10 Tb/s backbone**
- Significant donations from vendors (NVidia, Graphcore)
- Strong contributions from industry partners (Thales, REFLEX CES)
- Collaboration with IDRIS (host), GENCI and other partners
- Integration in national ecosystem (e.g. PEPR ORIGINES)
- Additional partnerships being discussed ...



Supporting Initiatives

Fostering disruptive R&D in Europe

- TransContinuum initiative (TCI) from ETP4HPC
- 8 associations in Europe covering the whole compute continuum (inc. BDVA, HiPEAC, 5G IA, etc ..)

TCI
TransContinuum Initiative

- Working on white paper with CERN on Big Science Infrastructures
- Contribution to ETP4HPC Strategic Research Agenda
- Working closely with experts and advisers to progress EU roadmap on cyber-infrastructures using **SKA challenges as pilot for future global needs**

SKA & NumPEx (Exascale national program)

SKA: opportunity to address wide range of topics (ExaFlops & ExaBytes)

- Although France is getting back onboard “late”, strong science community, federated through ECLAT can provide support
- Highly challenging but far reaching visibility
- Designing solutions for a 50 years lifespan: sustainability as key indicator
- Strong link with industry is desired: return on investment

Transverse topics:

- Cybersec (AAI) together with data lifecycle management (provenance & layout)
- Frugality (power efficiency) together with interoperability (full continuum) & resilience (continuous operations)

Significant contributions from / to NumPEx:

- ExaDoST & ExaDiP clearly contributing to French roadmap:
    - Construction phase of observatory + scientific exploitation

That’s it for today !