Extreme scale dataflows in the continuum for next gen giant observatories

Damien Gratadour

cnrs

l'Observatoire de Paris PSL

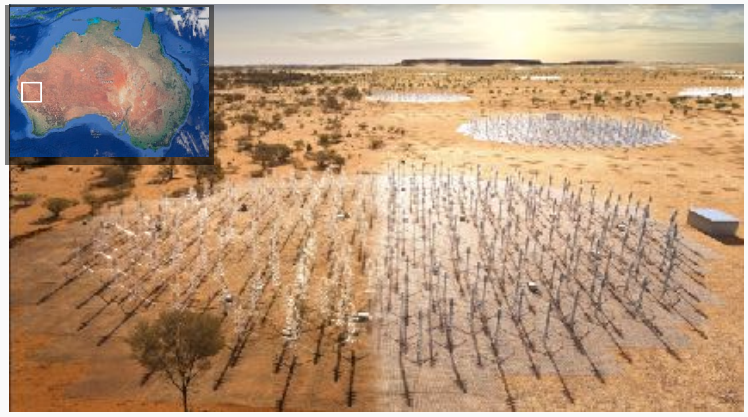
Extreme Computing Lab for Astronomical Telescopes ECLAT

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 →



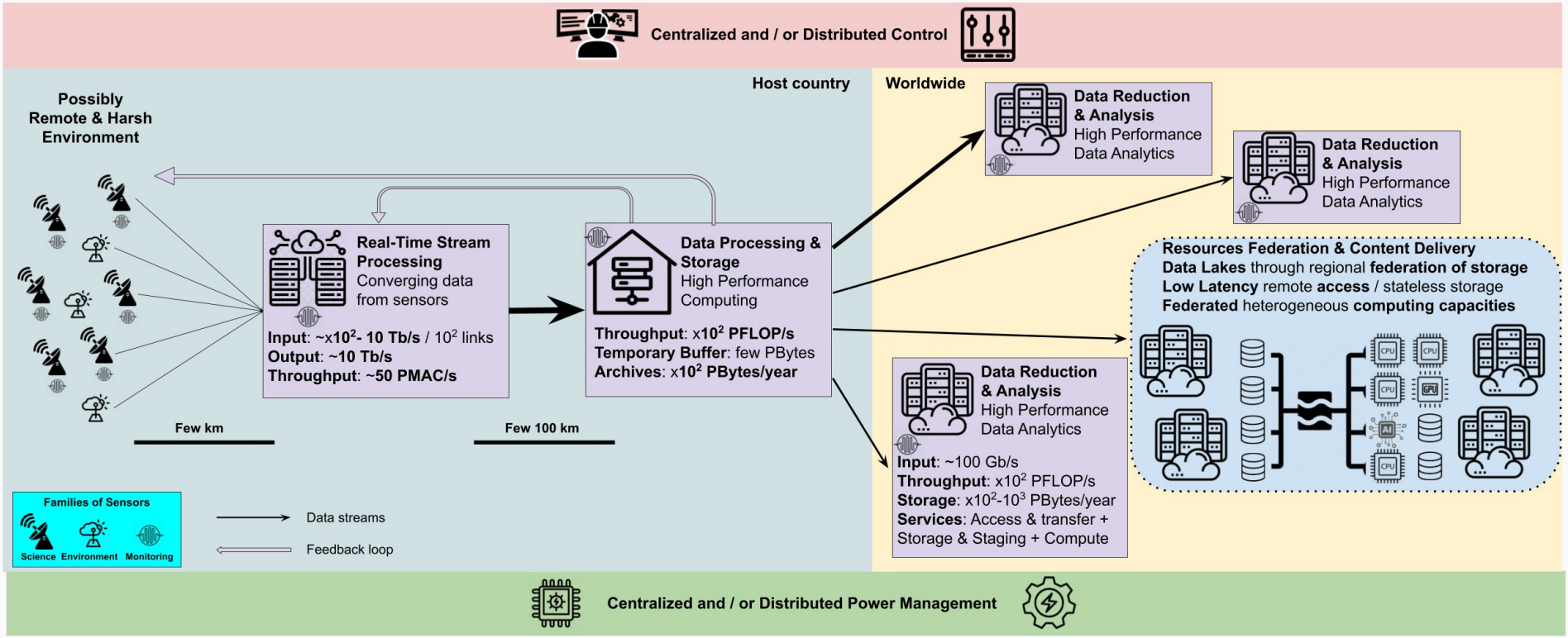
SKAO: unraveling the unknown

A truly Global infrastructure

End users communities across the globe

Cyber Continuum for SKA

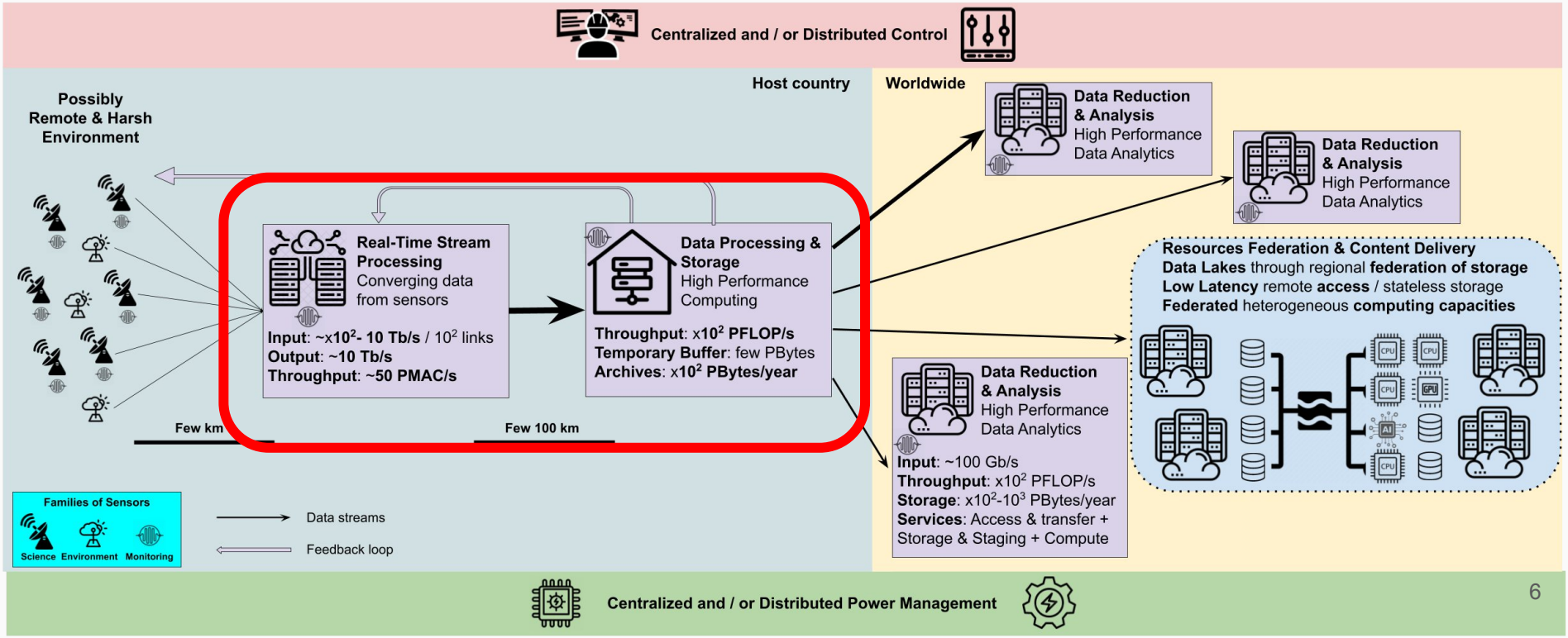
Hierarchical architecture: **system of systems**



Edge-to-HPC computing for SKA

Collect, Converge and Reduce data streams from distributed sensors

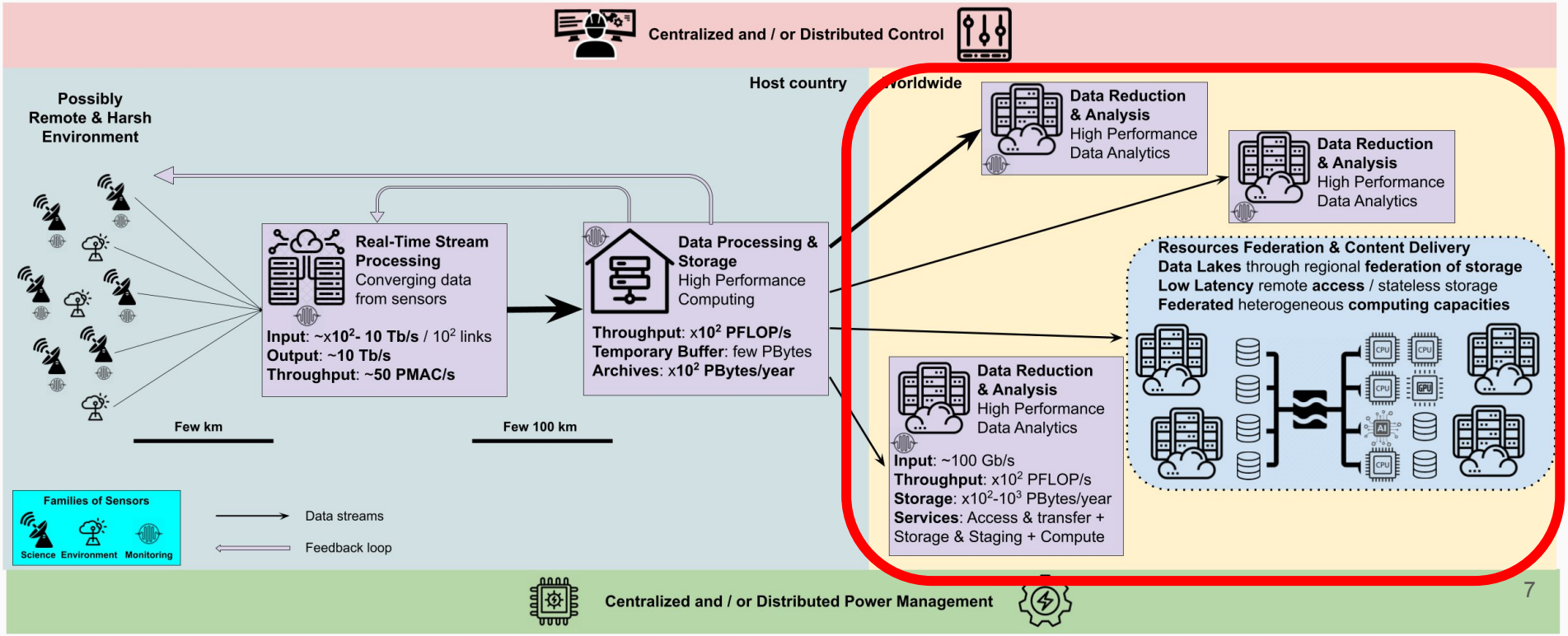
- In situ & Online data processing with centralized HPC systems
- Reduce continuous 10 Tb/s stream to 350 PB/year of data products
- Affordable / Adaptable / Frugal / Resilient
- Duplicated in two host countries (with centralized control in UK)



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



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

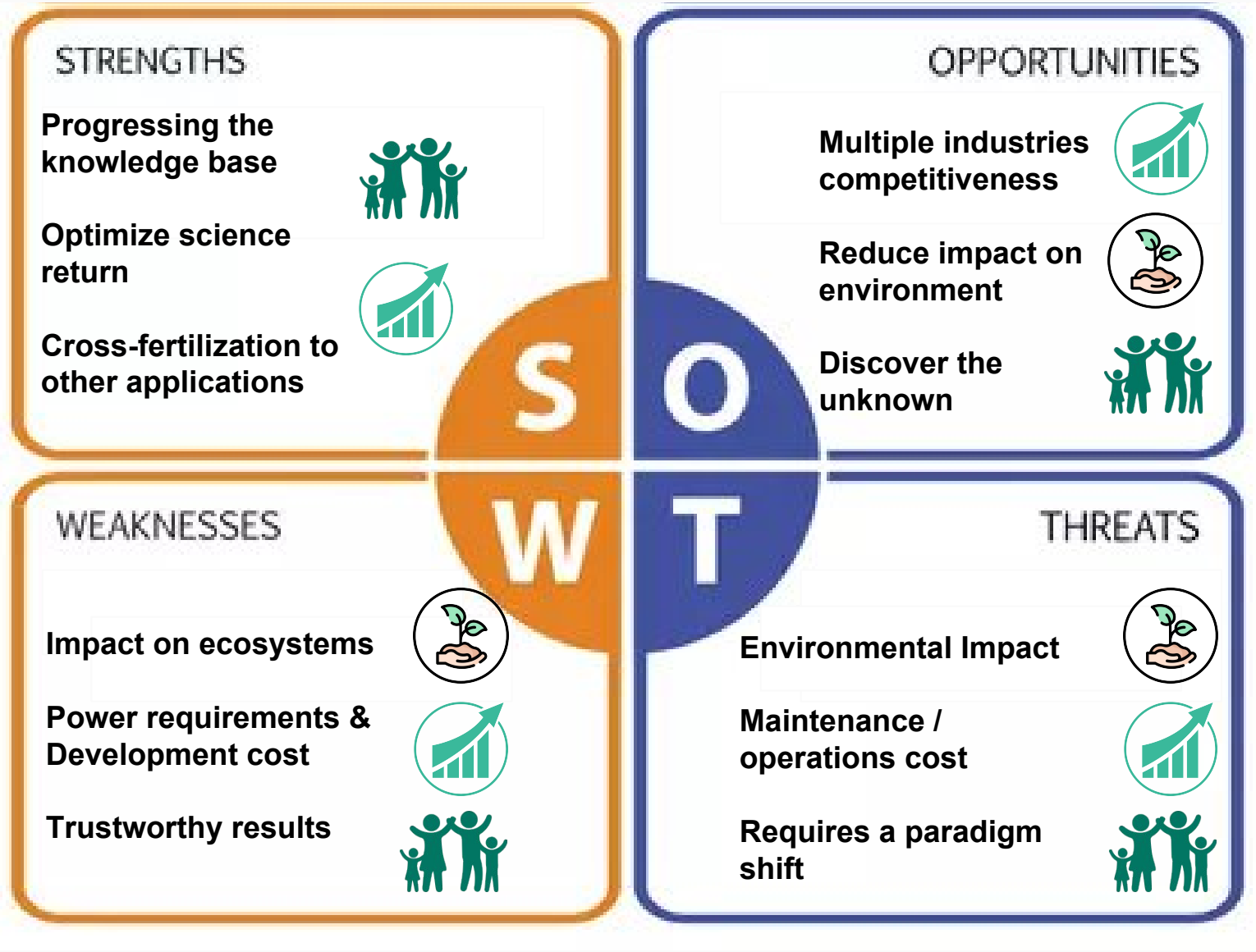
- **Limited power envelope**
  - Access to power grid
- **Cost containment**
  - Mostly relying on taxpayers money
- **Optimized operations**
  - Dynamical cyberinfrastructure, including reconfigurable HPC

Key considerations

- The case for **global** infrastructures: **transcontinuum by design**
  - New major challenge in many strategic areas
  - Cover the full edge-to-cloud continuum, with **strong inter-dependencies**
    - **Interoperability is a key requirement**
  - "**Operational**" cyber-physical infrastructures (as opposed to digital twins)
    - **Reliability** 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) impact others (e.g. sensors / edge)
  - Covering many domains: access to energy, TCO, supply chain, RAS, predictive maintenance, etc..
  - Harnessing many technologies: AI, cyber-security, HPC, big data, etc ...

A SWOT analysis on sustainability for SKA



The pathway to sovereign sustainability

- Tackling inter-dependencies between 3 dimensions:

  - **Societal**: Enhance access to affordable services and knowledge
    - Cross-fertilization across application domains

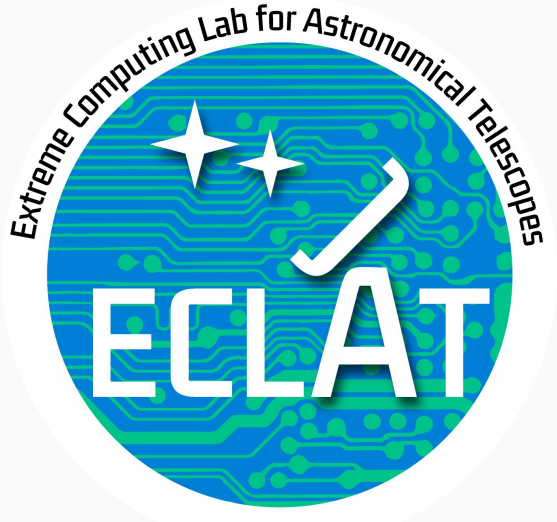
  - **Environmental**: Global Energy efficiency, across the continuum
    - Key responsibility: reduce impact of major infrastructures
    - Operational constraints: Maintain operations (thus European leadership) over long lifespan

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

Supporting initiative

ECLAT: *Laboratoire Commun* CNRS-Inria-Atos

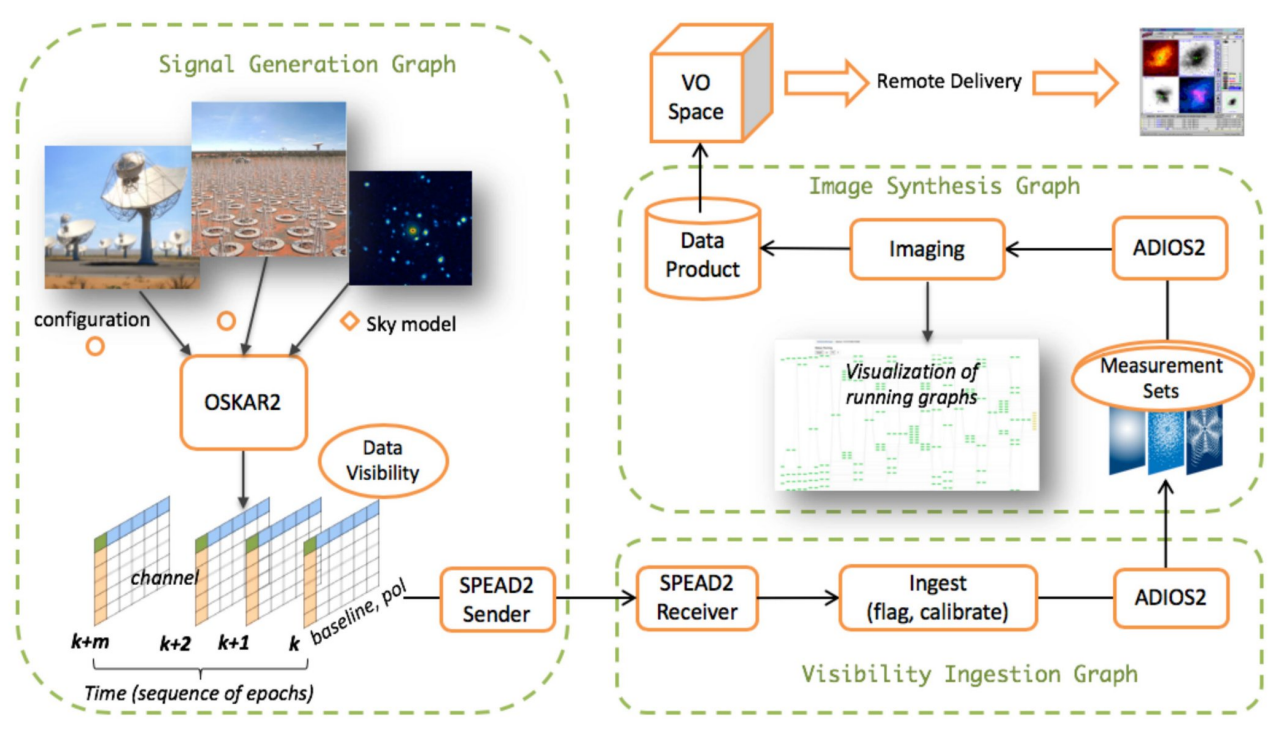
- Support structure for French contribution to SKA
- In kind contributions from partners, incl. **INSU, INS2I, INSIS**, multiple **Inria** teams together with **Atos** (now Eviden) **R&D** and **business dev**.
- Truly multi-disciplinary and trans-sectoral collaboration



Supporting initiative

NumPEx: High Performance Numerics for the Exascale

- Large exploratory R&D project (see F. Tessier's talk, this conference)
- Strong involvement including core activities (WP co-lead)
- Identified as a strategic application, building demonstrators at scale
- Including **frugality and resilience** as sustainability indicators



Horizontal challenges, addressed sustainably

Big science requires unprecedented (and exciting !) ICT breakthroughs

- Integrate / leverage emerging HPC / HPDA technologies
  - Across the infrastructure continuum
  - At all scales
  - Maximize science return
  - Converge design and operation / maintenance models: continuous integration
- AI has a key role to play
  - Across the infrastructure and at all scales
  - From producing science to managing the infrastructure
  - Change of paradigm calling for new AI methodologies
- And there are more ...

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

That’s it for today !