Powering giant instruments for Big Science

Damien Gratadour

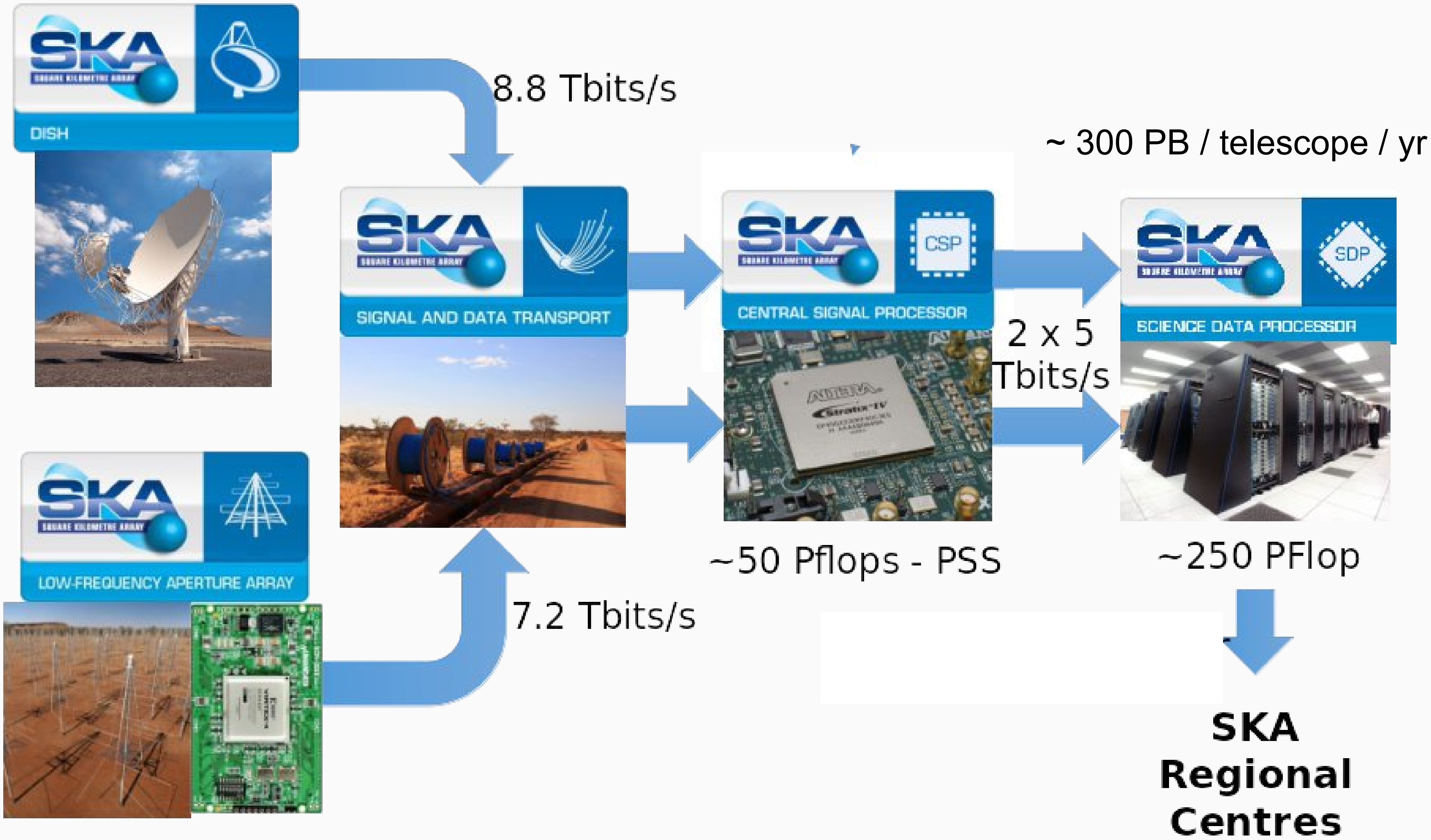
l'Observatoire de Paris
PSL

SKAO: a massive infrastructure for radio-astronomy

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

A broad range of science cases

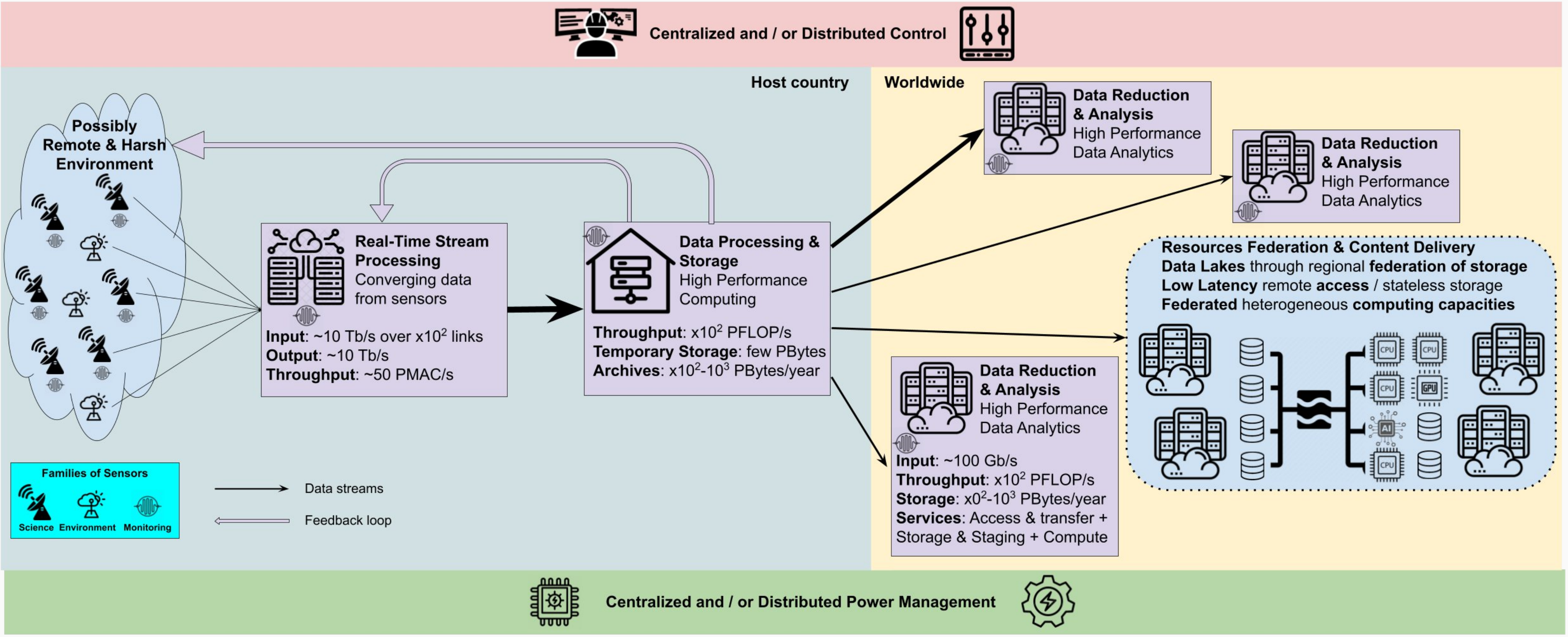
SKAO: A Giant Software Observatory



Cyber Continuum for SKA

These giant scientific infrastructures share common challenges

- 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



Challenges across the continuum (1)

Large collection of distributed & heterogeneous sensors

3 families of sensors:

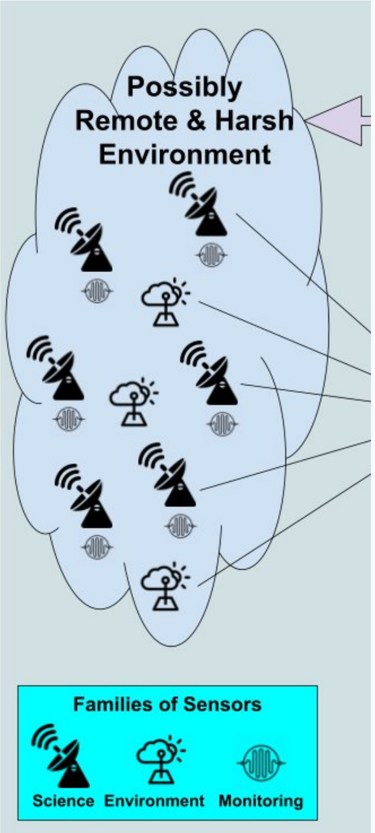
- Science: detect signals from object of interest
- Monitoring: monitor hardware (detector, components, etc..) state
- Environment: measure environmental conditions (humidity, temperature, etc..)

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

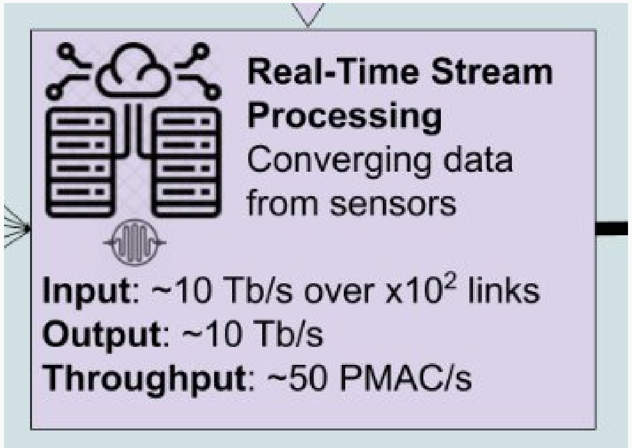


Challenges across the continuum (2)

Real-Time Stream Engine to produce raw data

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

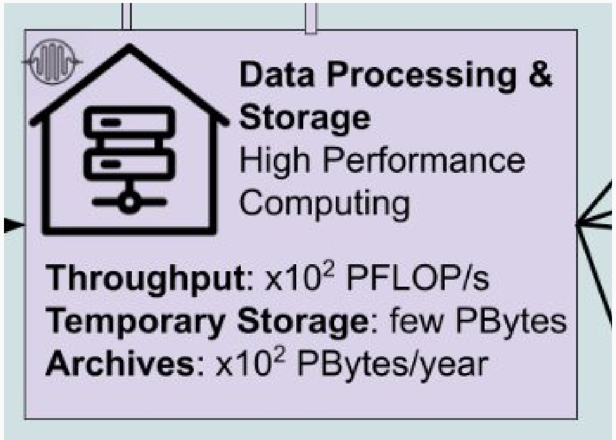
- ~10Tb/s as input
- I/O bound computing (~50 PMAC/s)
- **Aggregation, convergence, filtering & selection of data**
- **No reduction & storage**



Constrained and remote environment:

- Energy consumption
- Complexity & Cost
- Reliability & Maintenance

Challenges across the continuum (3)



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 PBytes / year)
- Short term buffers versus data archives

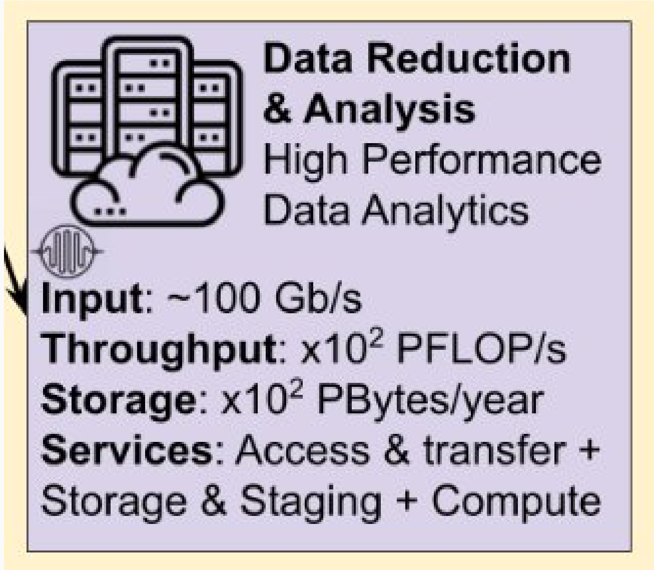
Heterogeneous workloads: from I/O bound to compute bound

- Low arithmetic intensity, Iterative process

Output data products distributed globally

- Worldwide multicast over 100Gb/s links

Challenges across the continuum (4)



Local Facilities for Data Reduction

Relying on "external" facilities for analysis work

- National or Regional facilities
- Not controlled by observatory
- Throughput needs depend on workloads. Typical requirement is x10^2 PFLOP/s
- Accounting of resources usage

Handling massive amounts of data on shared facility

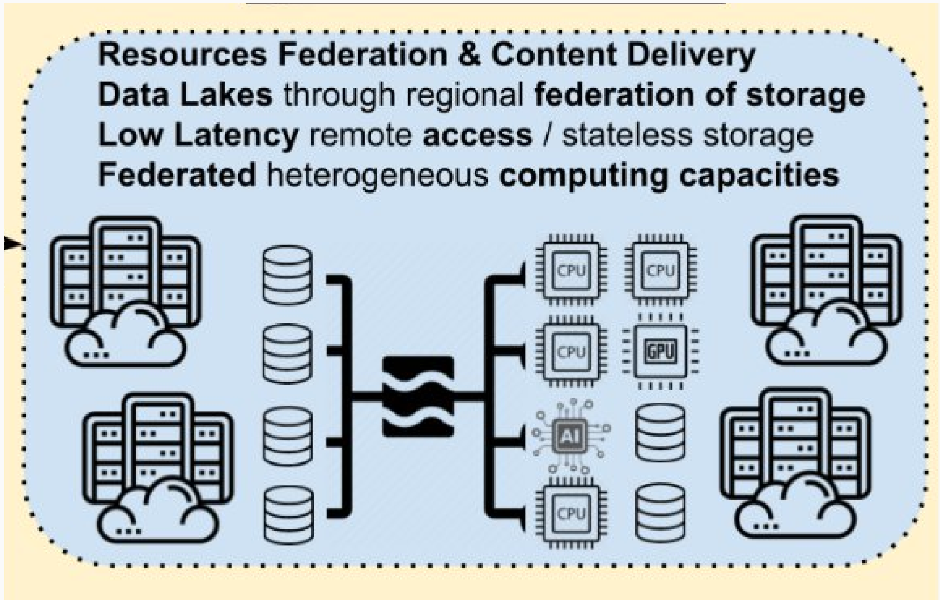
- Standardized strategy for data handling (access & transfer together with storage & staging)

Workloads need to run on a variety of heterogeneous environments

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

Challenges across the continuum (5)

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

Challenges across the continuum (6)

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

Sustainability is (also) a driver

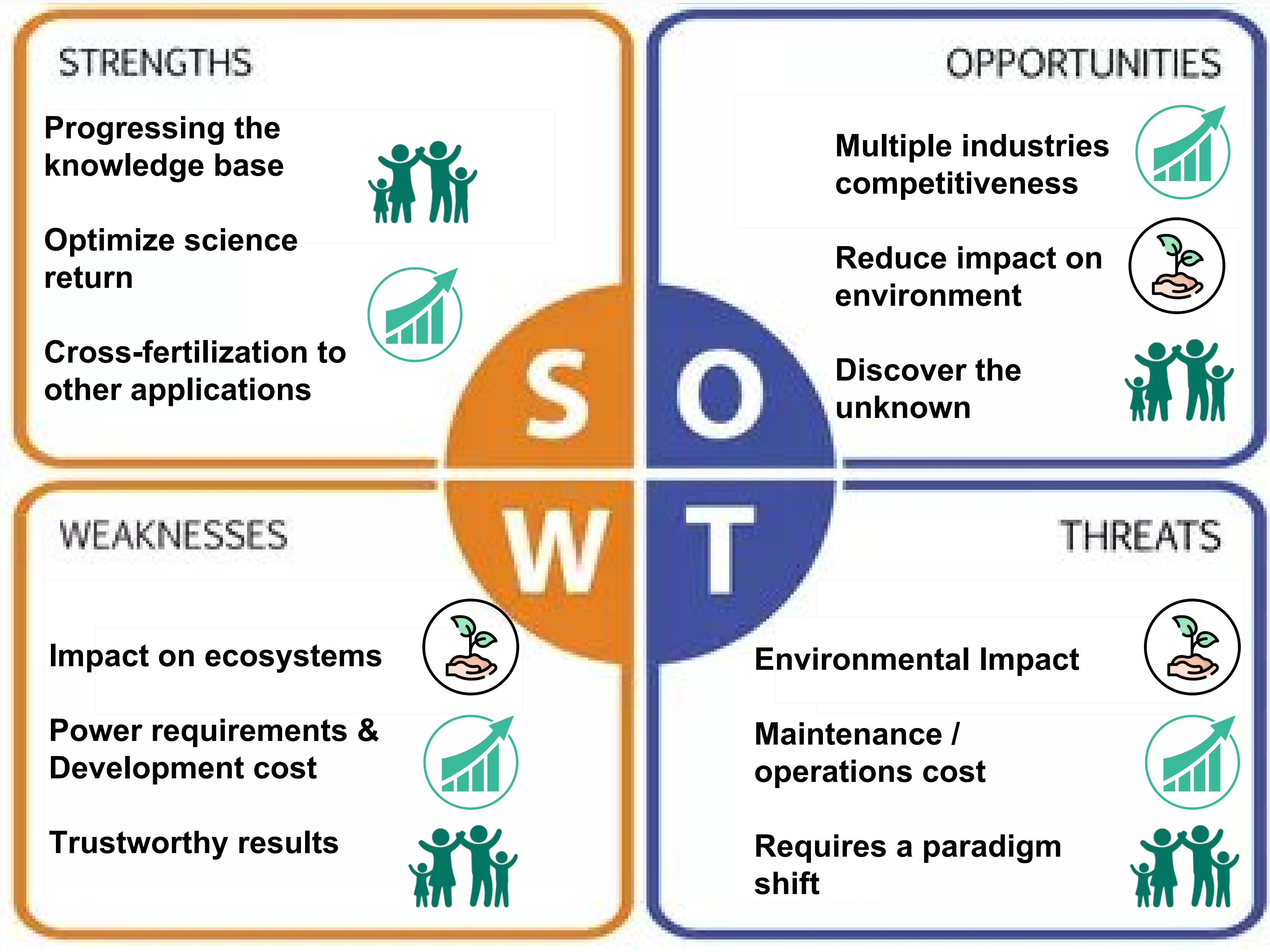
What most stakeholders want:

- **I am a human** can you:
  - Make me understand the world ?
  - Make my life better ?
    - progress the knowledge base, open new opportunities
  - Mitigate / null the environmental impact ?
- **I am a tax-payer** can you:
  - Be resources effective ?
  - Make your activity economically viable ?

What the scientific community wants:

- **(Big) Science is a lengthy process with a long term perspective**
  - Train the next generation of scientists, disseminate results at large
- **The required infrastructures are giants:**
  - Cost effectiveness to preserve other smaller scale research infrastructures
  - Building them is tough but not enough: operations across 50 years is key

A SWOT analysis on sustainability



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