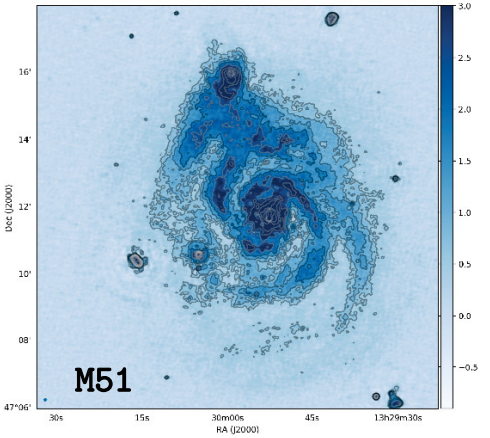
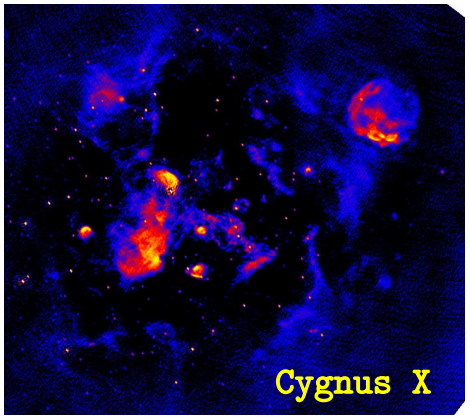
Data processing across borders:

Lessons learned & Future perspectives

Outline:

- SURF/SURFsara (NL e-infra)
- Distributed workflows
- Success stories
- Modern technology (cloudified layers)

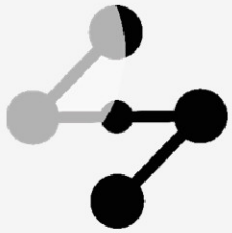




SURF{sara, net, market} + NLeSC

*** SURF is a Cooperation with 100+ members**

SURFsara services: clear international focus



High bandwidth network

Cartesius
Lisa
HPC Cloud
Research Cloud
Grid
Spark
Kubernetes
>80K cores

dCache
SURFdrive
ResearchDrive
Object Store
Preservation
EPIC PID
iRODs

> 50 PB

Support
Consultancy
Training
Visualisation
Optimisation
Collaboration

* Data center runs on 100% green energy

EGI EUDAT GÉANT EPICO EOSCpilot e-dci PRACE ETP 4 HPC HELIX NEBULA THESCIENCECLOUD QUTech INTEL-PCC SOIL SURF SARA

* HPC (compute intensive): PRACE-RI & HPC Europa (transnational access)

* HTC (I/O intensive = RA) : EGI (federated access via Virtual Organizations)

Distributed Workflows & Requirements

SURF SARA

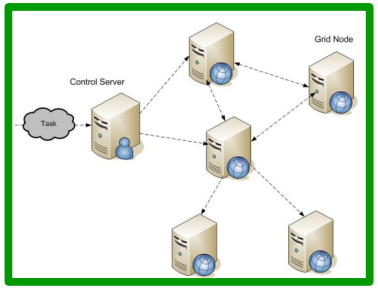
* A computational workflow consists of several layers

* Each layer provides a set requirements that together create a set of foundational requirements

Layers:

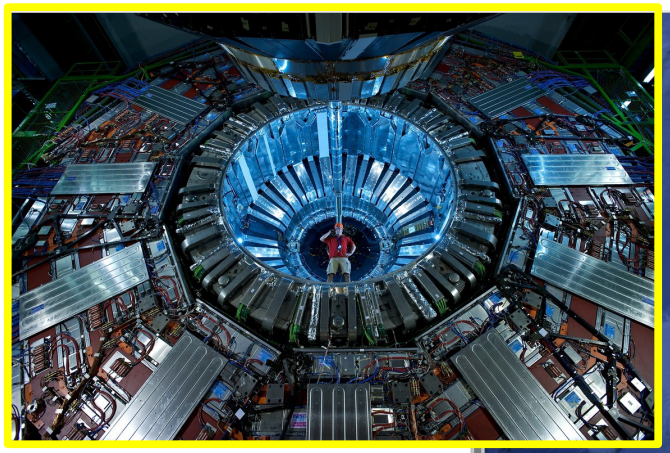
(REQ) Users (Astronomer | KSP | Observatory)
(REQ) Services & Applications (community providers)
(REQ) e-Infrastructure (compute, storage, network providers)
|
|→ " foundational requirements, but may differ dep. on usage model "

* Work on distributed compute & storage
  → coordination across sites (required)
  e.g. access, monitoring, accounting



SURFsara & Grid for HEP

* CERN developed the Grid for the WLCG project



WLCG: 42 countries, 170 compute centers, I/O intensive
& embarrassingly (small) parallel tasks (2M/day)
→ 1M cores and 1 Exabyte of data storage

SURFsara & Generic Grid compute

HEP has proven that distributed processing works !

* Organizations aimed at promoting Grid for generic science: EGI, OSG, NGI's

Grid service components @ SURFsara for generic science:

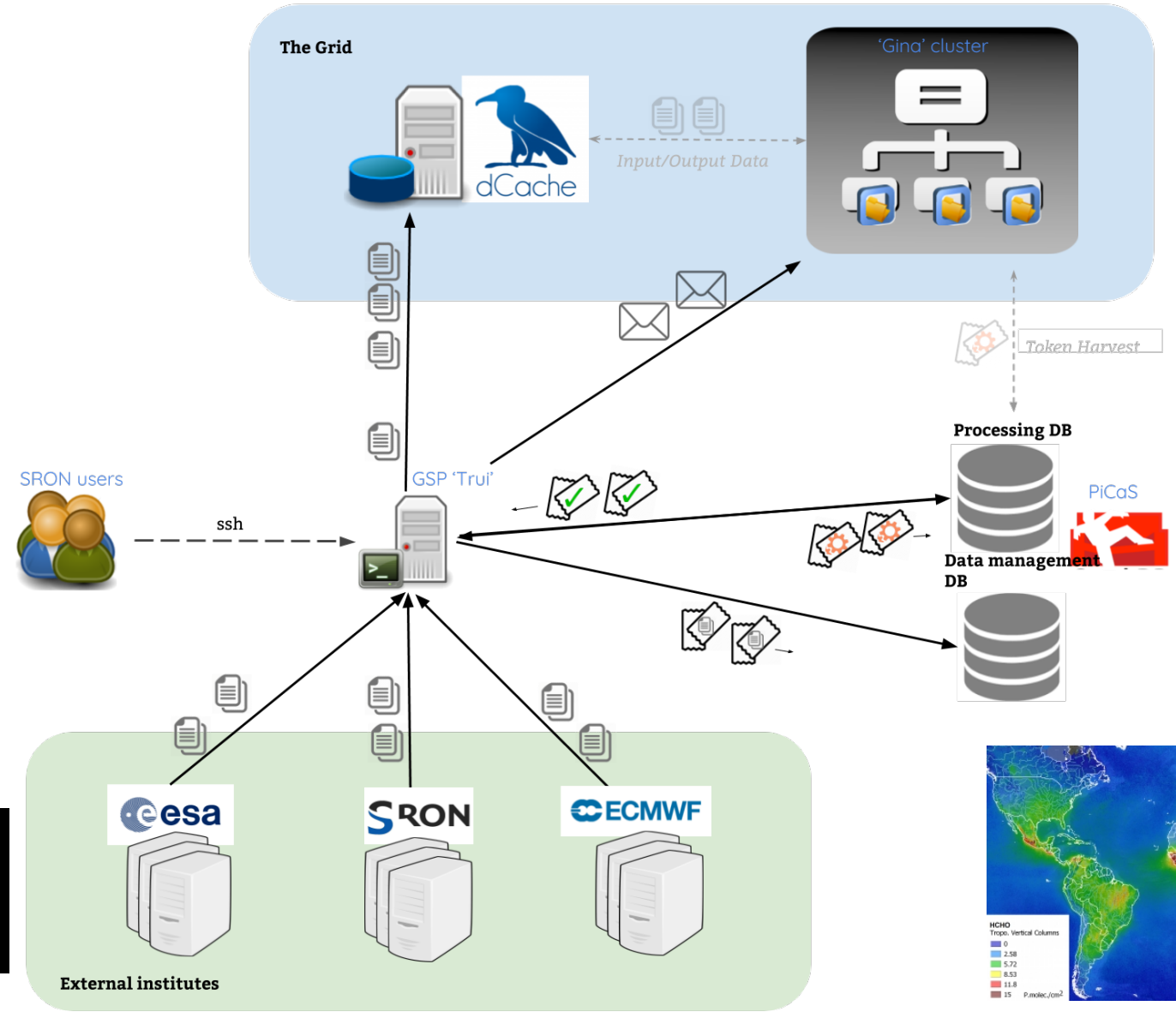
* Compute: Gina cluster (6500 cores, 8GB RAM/core, 100GB scratch/core)
* Storage: dCache (disk, tape), local scratch (SSD, HDD)
* High bandwidth network: internal (1.2 Tb/s), GEANT
* Pilot Job Framework: PiCaS (CouchDB), Dirac, Rucio, PanDA
* Workflows: Apache-Airflow, CWL
* Grid-stack:
  * UI: generic, dedicated
  * WMS: gLite, Dirac (Grid middleware)
  * VOMS: VO and VO groups for authorisation
  * CreamCE: Torque, Slurm
* Authentication: X.509 certificates *[VO + X.509 = fed. AAI]*
* Software: CVMFS (softdrive.nl), Containers (Singularity), tarball with job

SURFsara data center (AMS)

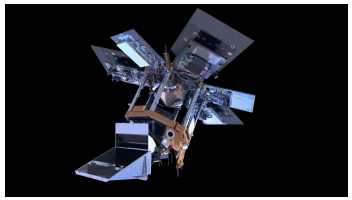
SURFsara & Grid for Climate science

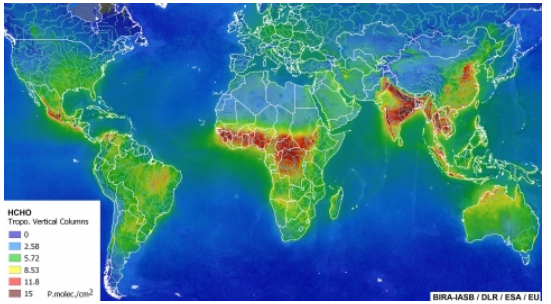
Trui: Control center
dCache: Data storage
Gina: Data processing
PiCas: Data/Job metadata

: CVMFS SW distribution
: Input token
: Processing token
: Done token
: Pilot job
: Input/output file



Tropomi (S5P)



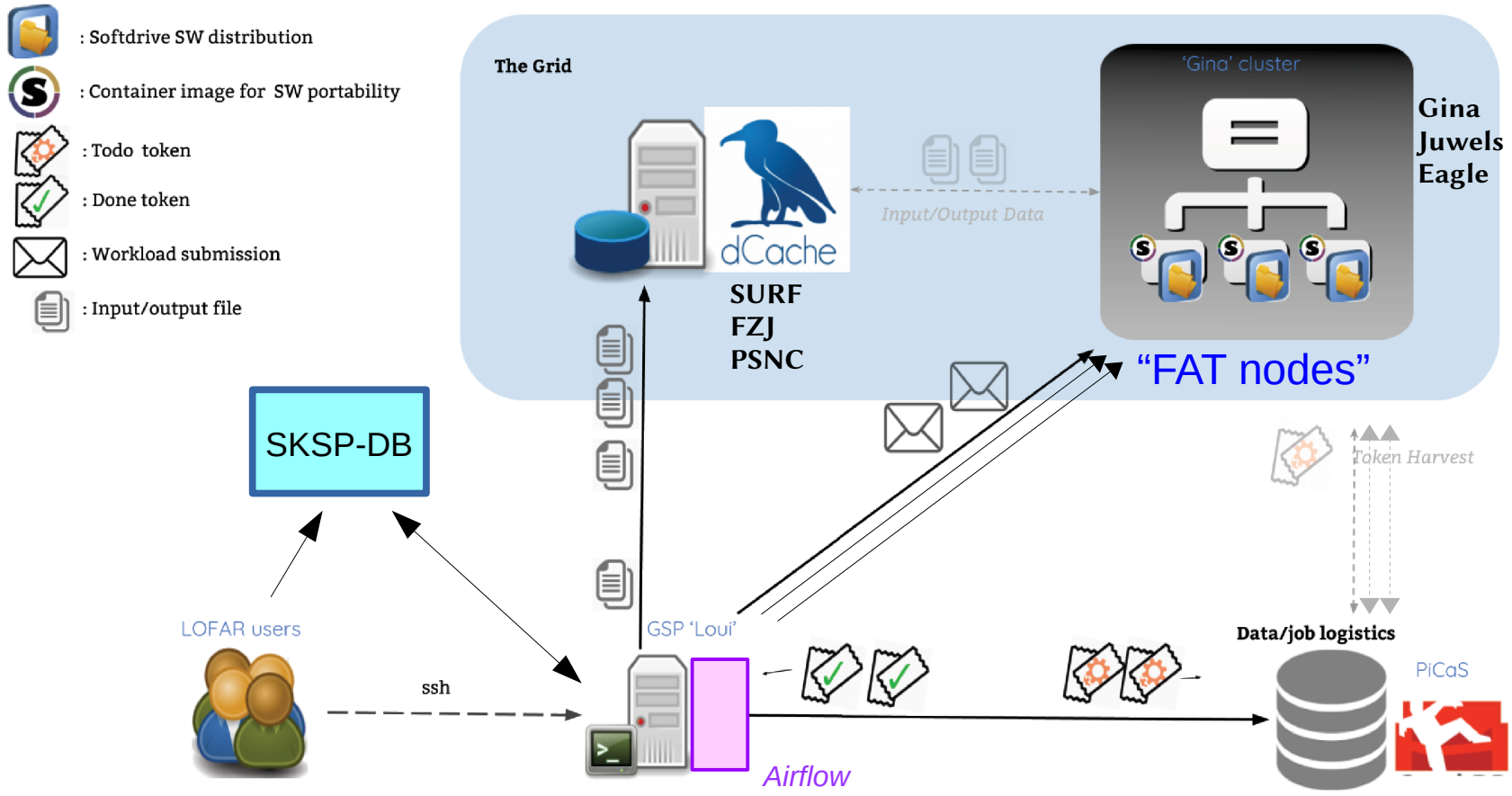


SURFsara & Grid for Radio Astronomy

LOFAR Surveys KSP

Universiteit Leiden

LOFAR

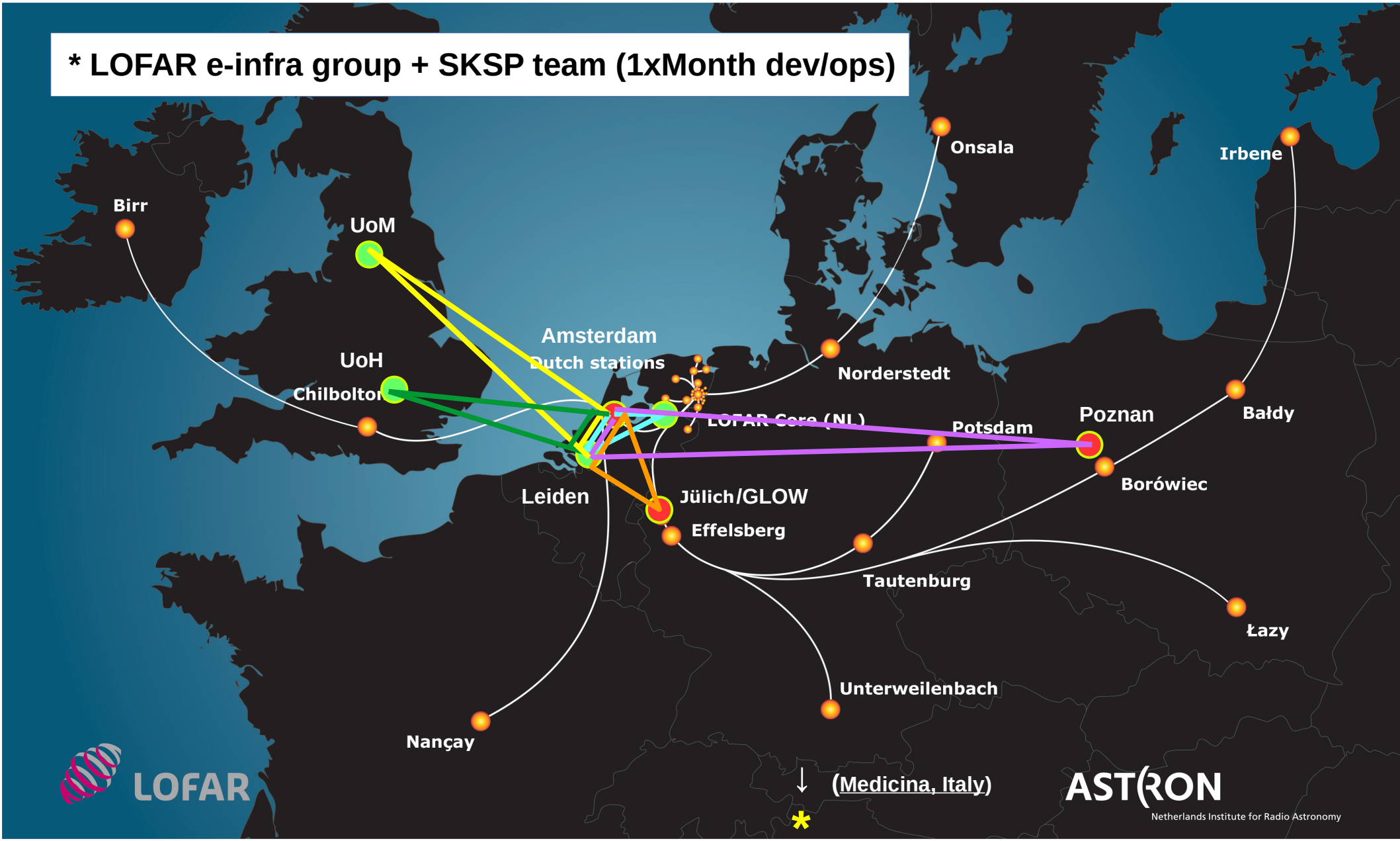


---

* **Workflow : Grid_LRT + Apache-Airflow (Mechev+2017,2018)**

* **Radio "pipelines" are a fast moving target (CI/CD → HR !)**

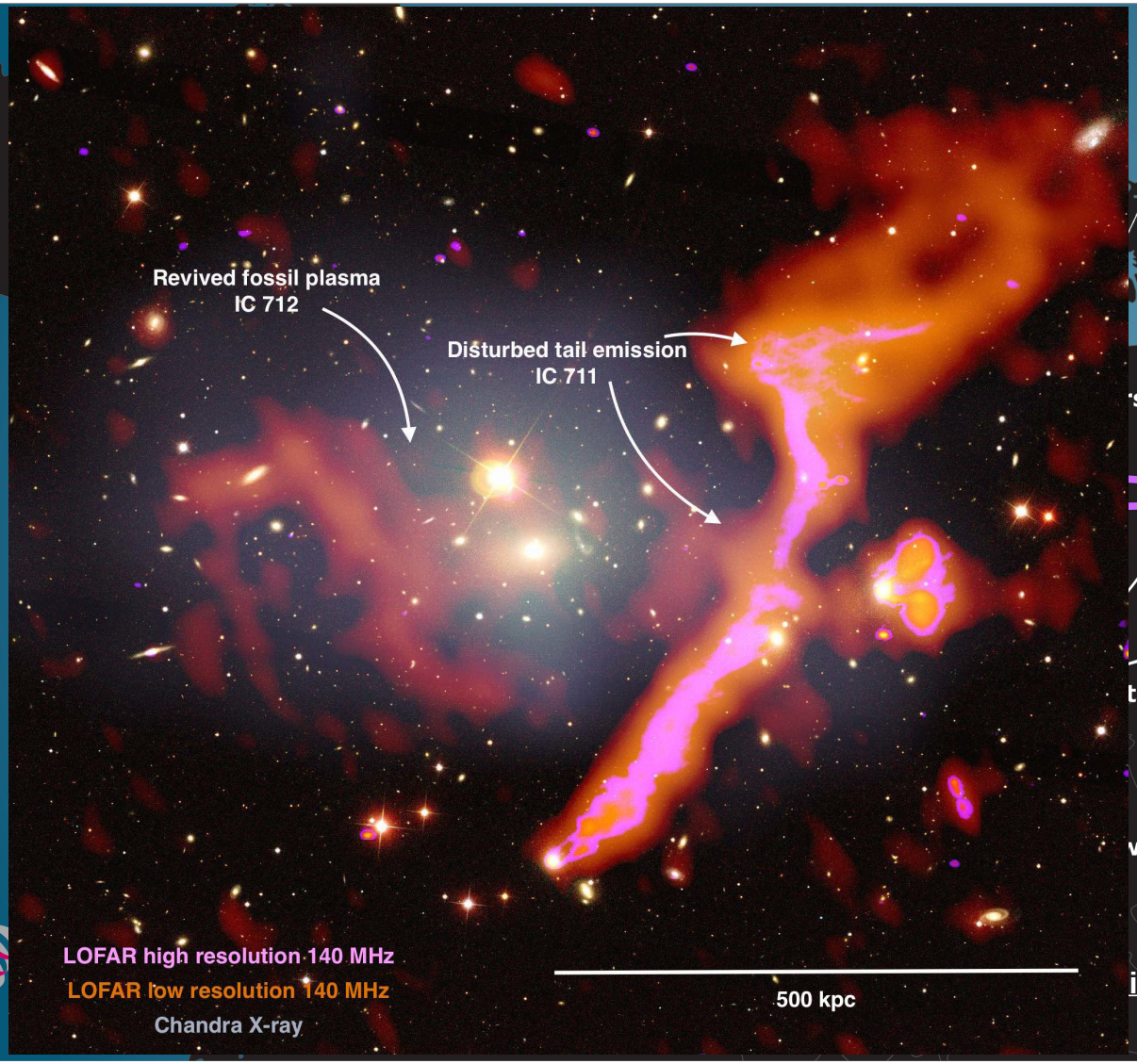
LoTSS: LOFAR “Grid++” @ Work



LoTSS: LOFAR “Grid++” @ Work

SURF SARA

* LOFAR e-infra group + SKSP team (1xMonth dev/ops)



* LoTSS Press release

“300000 new radio sources”

200 astronomers
18 countries
56 institutes

1.5 Pb data (>7Pb reduced)

CNN

Daily Mail

sky news

NOS

CBS NEWS

Lessons learned: summary

1) The Grid is 'fit for purpose'

(a) High-throughput processing power to analyse large data volumes

(b) Scalable storage capabilities to ingest and process observation data

(c) High-speed connection to the storage to optimise data transport

2018:

SURFsara Grid process: HEP, RA (LOFAR/SKA), Climate, GW, Lifescience, Geophys.

SURFsara Grid storage: 200 PB read, 100 PB processed@Gina, 25 PB written

2) RA: *complex, long and I/O, RAM & scratch intensive workflows*

* LOFAR minimal job requirements (DDF: 256 GB RAM, 3 TB scratch → 'FAT' nodes)
* SURFsara Grid: Gina cluster 6500 cores (8 GB RAM/core, 100 GB scratch/core)

3) Grid is a challenge for users & providers.

Misconceptions

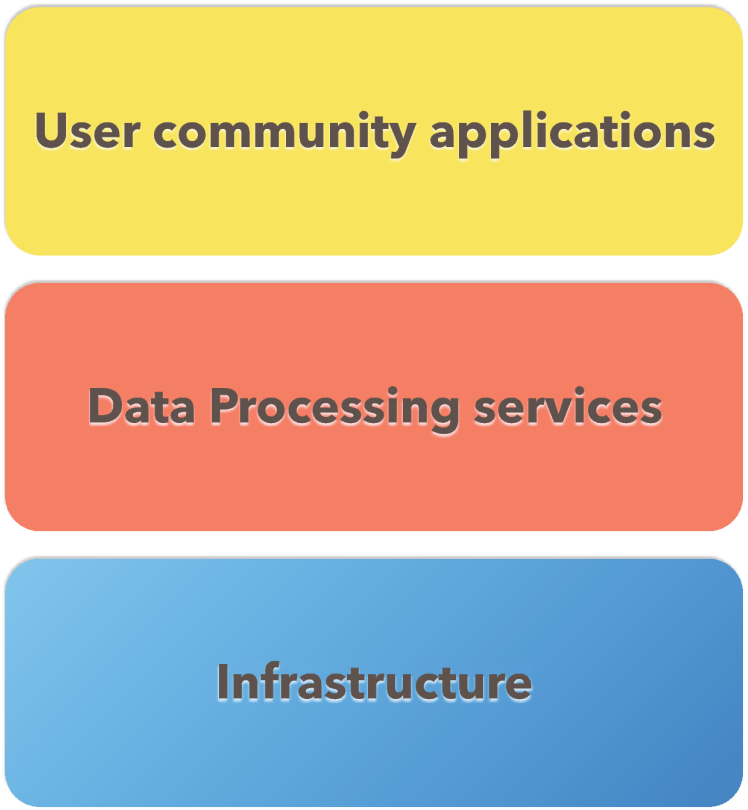
Infrastructure providers' view:

- Users are system administrators
- Users are software developers
- Users are data managers
- User requirements don't change

Users' view:

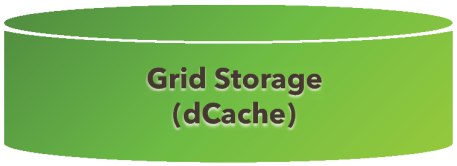
- The network is reliable
- Hardware doesn't fail
- Bandwidth is infinite
- Support scales

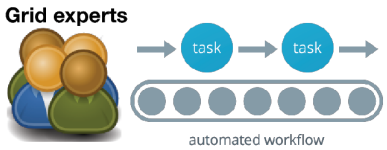
Future perspectives

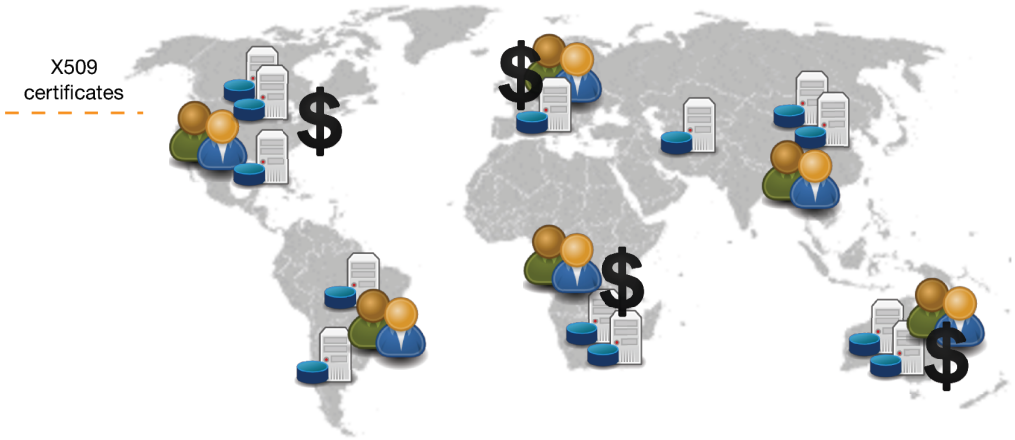


- Separation of infrastructure, services and tooling
- Low-threshold deployment of cost-efficient PaaS and SaaS services
- Co-development of services with user communities & cyclic requirement retrieval

Future perspectives







SURF SARA

Future perspectives

CERN

LIGO-VIRGO

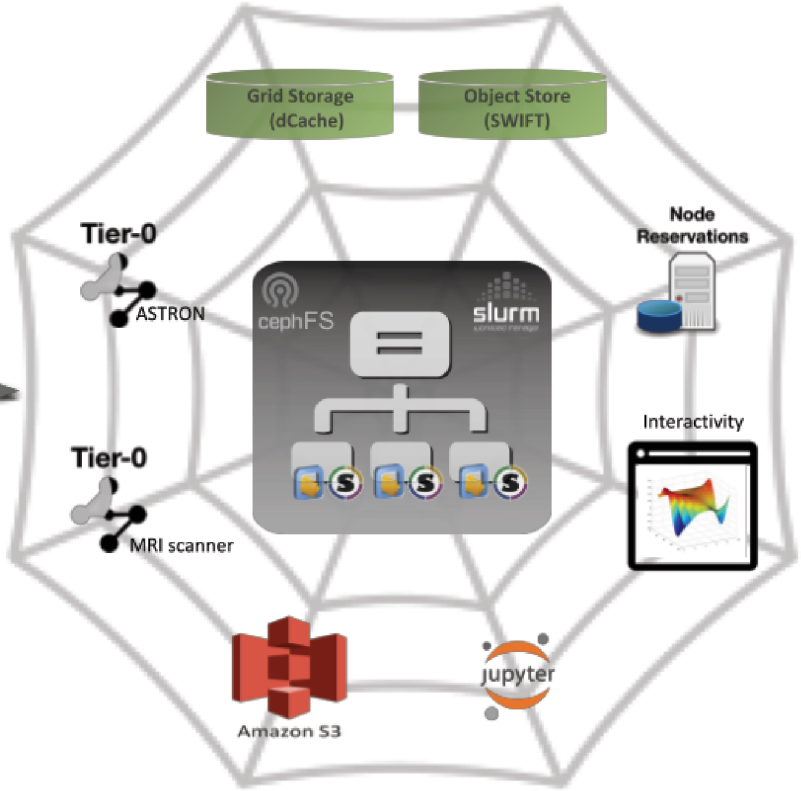
Pierre Auger

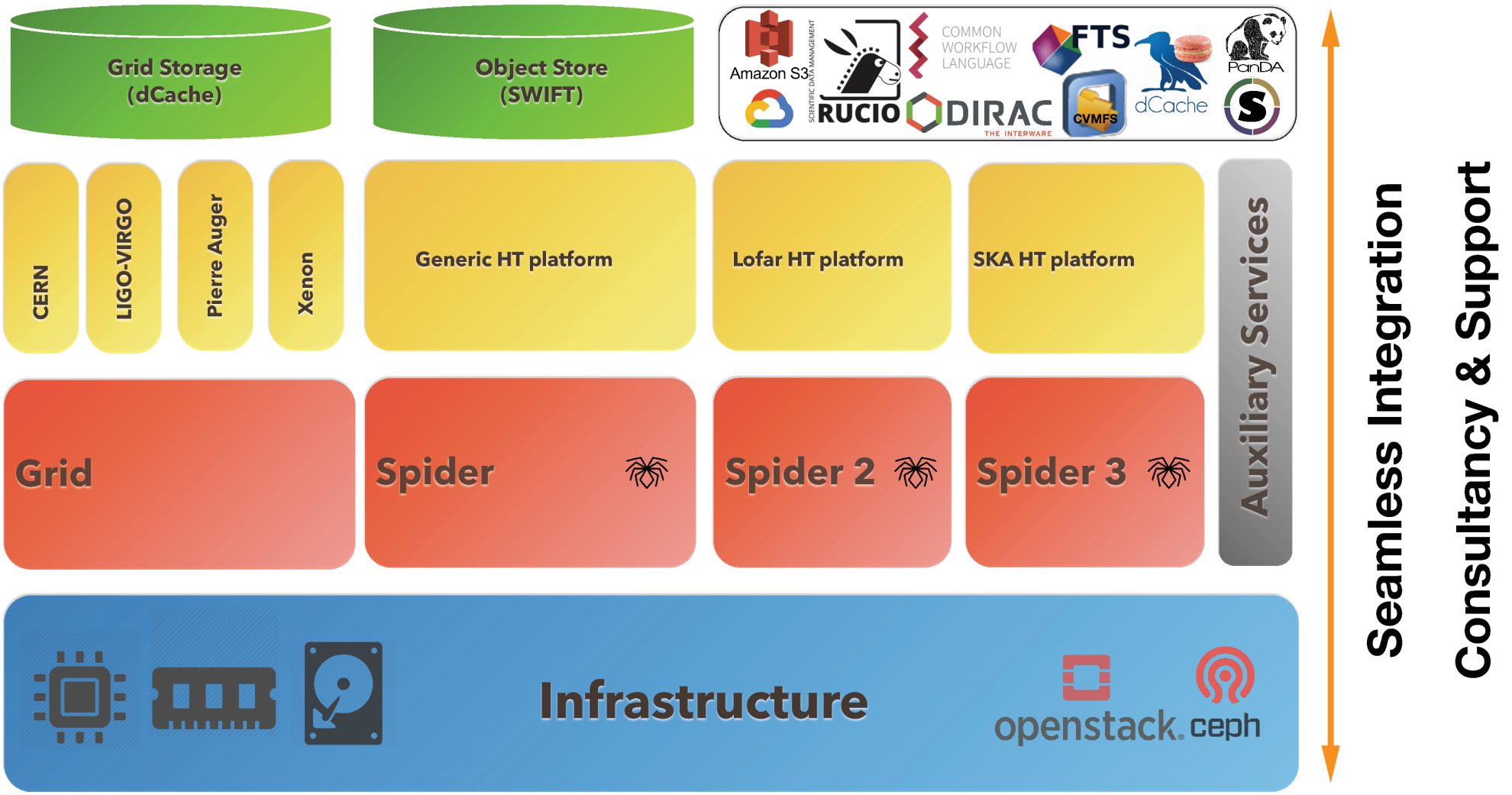
Xenon

HT platform

Grid

Spider




Grid Storage (dCache)
Object Store (SWIFT)
Amazon S3
RUCIO
COMMON WORKFLOW LANGUAGE
FTS
DIRAC
THE INTERWARE
CVMFS
dCache
PanDA
CERN
LIGO-VIRGO
Pierre Auger
Xenon
Generic HT platform
Lofar HT platform
SKA HT platform
Auxiliary Services
Grid
Spider
Spider 2
Spider 3
Infrastructure
openstack
ceph
Seamless Integration
Consultancy & Support

Extra: SURFsara & NL e-Infrastructure

Hub for e-Infrastructure and eScience: 150+ people and growing

Data center	:	Compute (HPC/HTC), Storage, Network, Market/Business
Innovation	:	SOIL, Qutech, Intel PCC-DL, OpenStack cloud, Kubernetes
Knowledge [RA]	:	EOSC-pilot/Hub, EOSC Escape, LOFAR SKSP/SDF/PWG, Aeneas
E-Science	:	Pilot Job frameworks, Containerization, DIRAC, Triple-A2
Solutions IT4RE	:	Consultancy, Collaboration & Customization

Extra: Lessons learned from the Grid

1) The Grid is ‘fit for purpose’

(a) High-throughput processing power to analyse large data volumes
(b) Scalable storage capabilities to ingest and process observation data
(c) High-speed connection to the storage to optimise data transport

* SURFsara Grid: HEP, RA (LOFAR+SKA), Climate, GW, Lifescience, Geo-physics

2) RA: *complex, long and I/O, RAM & scratch intensive workflows*

* LOFAR minimal job requirements (DDF: 256 GB RAM, 3 TB scratch → ‘FAT’ nodes)
* SURFsara Grid: Gina cluster 6500 cores (8GB RAM/core, 100GB scratch/core)

3) Grid is a challenge for users:

- complex authentication system for compute and storage
- no shared file system
- no root privileges to install software
- non-interactive environment
- bookkeeping is not easy
- steep and lengthy learning curve

4) Grid is a challenge for providers:

- coordination, monitoring, accounting, maintenance, etc.

Extra: Distributed Workflows & Users

SURF SARA

* Usage models → (Different) Design requirements !

Per observation HW requirements are similar, but usage pattern is not

1) Large organization (e.g. CERN, SKA, LOFAR, ESO, ESA)
   - HR for IT: long-term & well organized
   - Bulk data (EB) with predictable data flows
   - Non-interactive, automated processing

2) Large projects (e.g. LOFAR Surveys KSP processing)
   - HR for IT: very limited (funding → science)
   - Bulky data (PB), data flows depend on (1)
   - Some interactivity / test facilities needed

3) Individual astronomer (e.g. LOFAR DDF calibration)
   - HR for IT: none
   - Single dataset(s) (TB)
   - Highly interactive & discrete (1-2 yr/proj)

EFFICIENCY vs FLEXIBILITY