EOSC-Hub LOFAR/RACC F2F Introduction
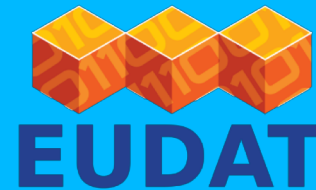## Hanno Holties

12 November 2018

EOSC Hub –
LOFAR as a Competence Center

- **EOSC Hub Objectives**

  - Simplify access through open and integrated service catalogue

  - Remove fragmentation of service provisioning

  - Consolidate e-Infrastructures

  - Widen access

  - Provide a knowledge hub

  - Increase innovation capacity

- **LOFAR - Radio Astronomy Competence Center**

  - ASTRON, SURFsara, FZJ, PSNC

  - 1/1/2018 – 31/12/2020

  - Predominantly service integration (Pilot)

11/30/18

2

Radio Astronomy Competence Center
Objectives

- **Support researchers to find, access, manage, and process data** produced by the International LOFAR Telescope.

- Provide **user access to large-scale workspace storage** facilities within the EOSC-hub to store and share temporary data and scientific data products.

- **Empower science groups to deploy their own processing workflows** within the EOSC-hub infrastructure.

- **Lessons learned** will serve as **input for** the design and construction of a **European Science Data Center** for the Square Kilometre Array (SKA)

- Build on/integrate/enhance existing efforts

11/30/18

3

Members, roles, effort

Member acronym	Role in the CC	Effort
ASTRON	Lead, science domain expert, data & service provider	29 (+3 in WP11)
JUELICH	Infra & service provider	10
PSNC	Infra & service provider	10
SURFsara	Infra & service provider	10

Radio Astronomy Competence Center
Original planning overall delay ~6 Months

- **Proof Of Concept Activities**
  - Port processing workflows to EOSC (M1 – M12)
  - Integration & Verification compute & storage (M4 – M12)
- **Pilot Activities**
  - AAI Integration (M7 – M15)
  - Metadata harvesting & PID registration (M7 – M15)
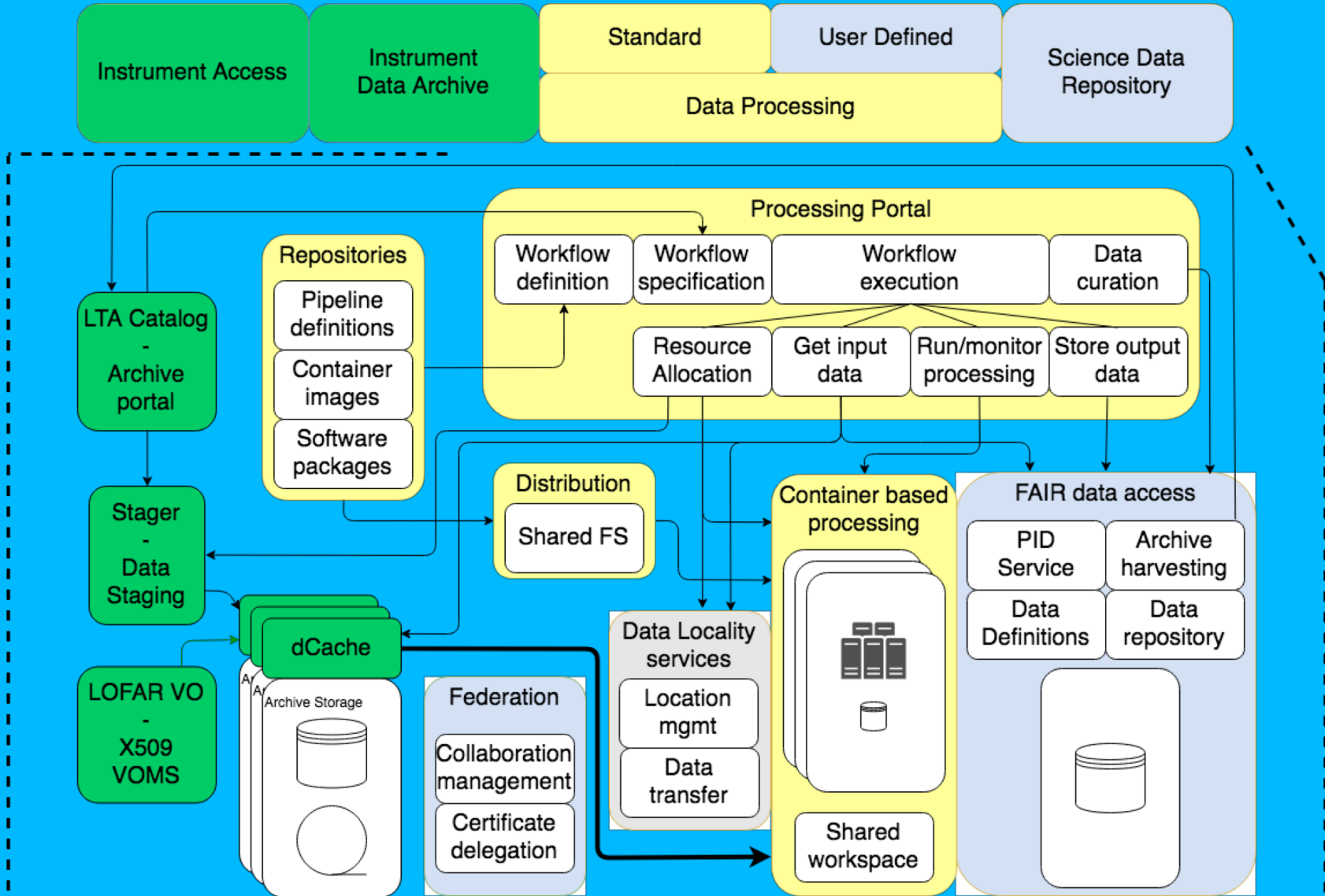 - User Workspace (M10 – M15)
- **Preparing the Production Environment** (M10 – M18)
- **Service Exploitation** (M13 – M36)

- **First WP8 Milestone**: "Technology assessment, architecture integration & validation plan for CCs" (1/3/2019)

Target Architecture

Relevant related projects

- **Asterics**

  - DADI

    - Virtual Observatory extensions

    - LOFAR 'ObsCore' datamodel

  - Obelics 'sub-projects'

    - ASTRON Data Portal (pilot ALTA VO GUI)

    - CWL Imaging pipeline (Currently focus on Apertif)

Relevant related projects

- **EOSC Pilot**
  - CWL workflows
  - Linked Data FAIR service
- **Process** (NLeSC)
  - Processing LOFAR data
- **Escape** (EU awarded, collaboration with a.o. CERN)
  - Broad program
  - Focus on SKA Science Analysis Platform

Today - Define initial tasks for

- Compute - Storage integration

- User workspace service(s)

- Open data access (Virtual Observatory, science data repository, PID service)

- Federation of services

- Workflow migration options

  - (if requested/time allows; under evaluation within ASTRON)

Compute – Storage integration

- Prefer local processing
    - Need something like B2Stage?
    - CVMFS for sharing code & containers
- Storage
    - dCache
    - Other?
- Compute
    - SURFsara: Grid, HTDP
    - FZJ: …
    - PSNC: …
    - ASTRON: SDC OpenStack cluster
- Tasks to define…
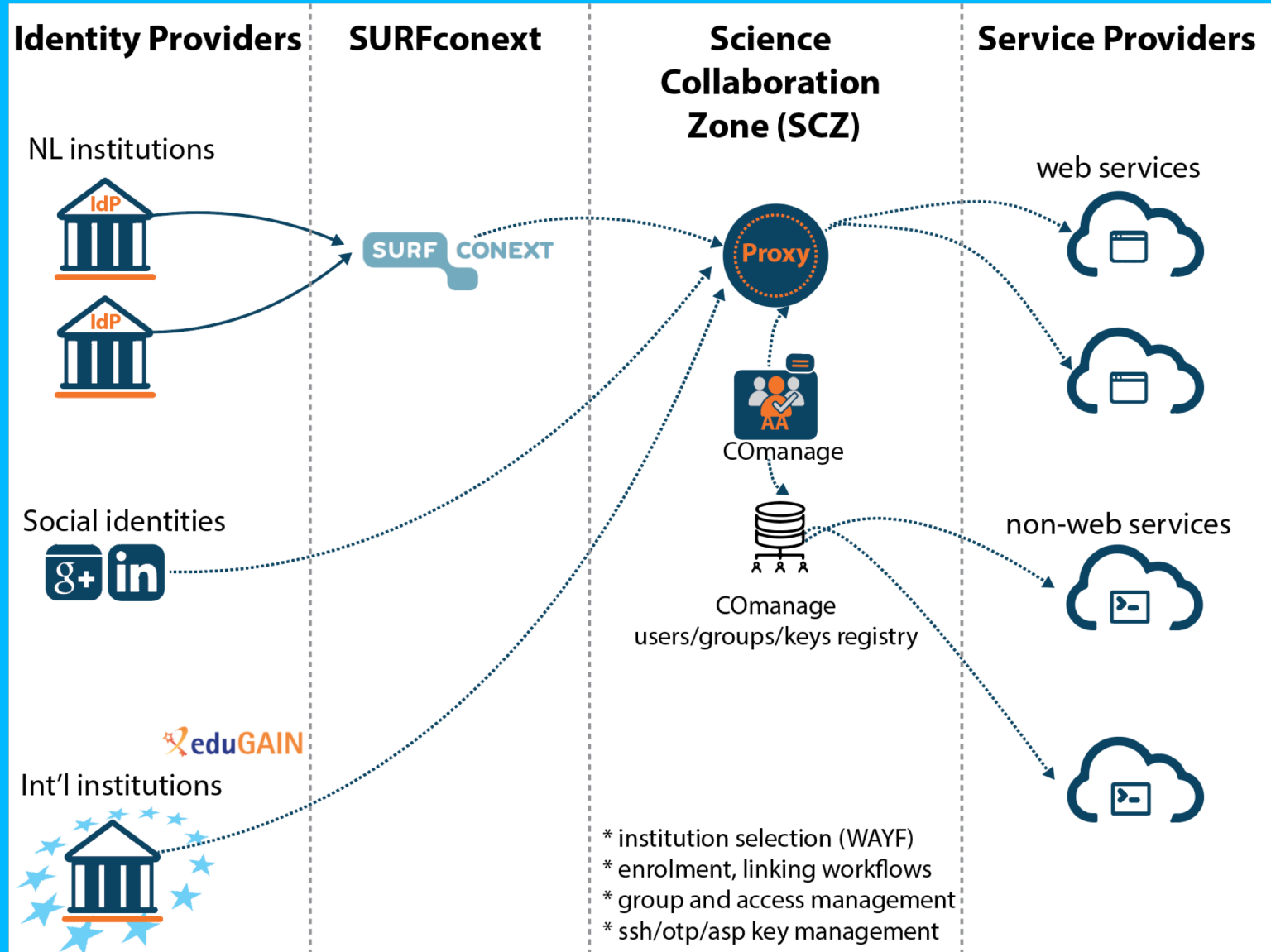
11/30/18

User Workspace

- Allow users to (temporarily) store/share data from processing pipelines

    - dCache user workspace?

    - Data transfers: (Web)FTS?

    - Other?

- Tasks to define:…

11/30/18

Open Data Access
FAIR Data Services

- **F**indable, **A**ccessible, **I**nteroperable, **R**eusable

- Virtual Observatory provides much of this (ASTERICS)

- **EOSC** Services to be selected & integrated

  - **ASTRON scoped Persistent ID's**
    Unique, persistent ID for referencing data, software, ...
    (e.g. DOI, Handle)

  - **Science Data Repository** for generated data-products

    - Associate metadata with data-products

- **Use case/requirements workshop TBD**

Open data access

- FAIR data registry

  - Virtual Observatory (TBD if part of EOSC Hub)

    - Register data

    - VO Ontology

  - Science data repository

    - Zenodo/B2Share?

    - LOFAR/Radio Astronomy metadata model

- PID service

  - SURFsara EPIC service (?)

- Tasks to define:…

11/30/18

Federation of Services
SCZ SURFnet pilot

Identity Providers | **SURFconext** | **Science Collaboration Zone (SCZ)** | **Service Providers**

NL institutions

IdP

IdP

SURF CONEXT

Proxy

COmanage

COmanage
users/groups/keys registry

web services

non-web services

Social identities

eduGAIN

Int'l institutions

* institution selection (WAYF)
* enrolment, linking workflows
* group and access management
* ssh/otp/asp key management

Federation of Services

- EOSCpilot extension:
  - Set up Comanage instance in EGI Check-In
  - Integrate <TBD> web services
  - Set up Macaroons for dCache data access
- Relevant LOFAR web services
  - AstroWise (Oracle DB), staging, download, MoM/NorthStar
- X509
  - Do we need to support X509 user certificates for Compute?
- Ssh key distribution for server/UI node access(?)
- Other services not covered above?
- Tasks to define:…

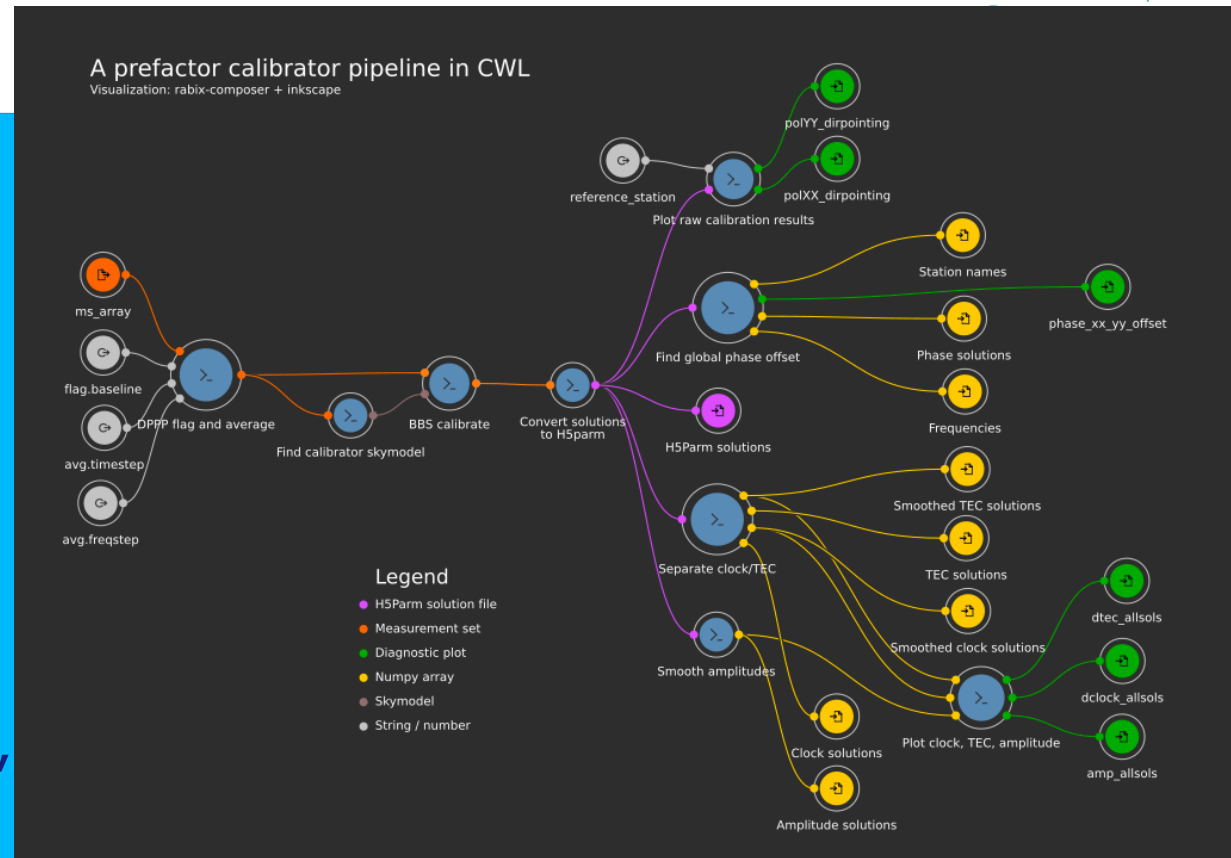# Container based pipeline processing with CWL

A prefactor calibrator pipeline in CWL
Visualization: rabix-composer + inkscape

- **Portability**
  - Containers
    - Singularity
    - Docker
    - uDocker

- **Deployed on**
  - Laptop, HTC cloud, Cartesius

- **Common Workflow Language based implementation**
  - Open standard
  - Supported by several workflow engines (Toil, Airflow, Arvados)

- **ASTRON Daily Image 16-01-2018 – Blogpost Gijs Molenaar**

Workflow integration

- Under discussion (28/11 internal meeting)
  - CWL likely for workflow definition
  - Airflow/Toil/PiCaS for running

11/30/18