NRC·CNRC

NRC.CANADA.CA

Role of an Execution Planner in an SKA Regional Centre Network Demonstrator

Séverin Gaudet

With SRC Canada, SRCNet Cyan and Coral Teams

SKA Regional Centre Network

SRCNet is an international network of interoperating science centres

SRCNet mission: To ensure that scientists can access SKA data products and use them to make discoveries

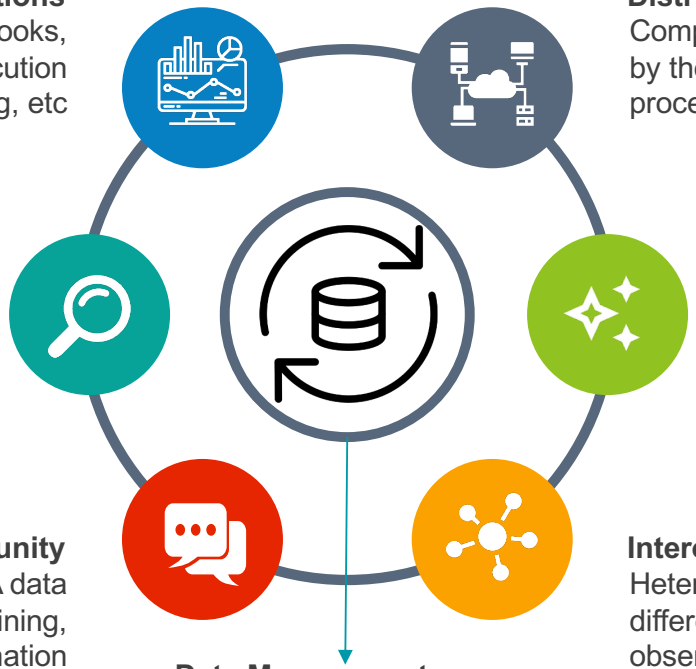
Capacity to:

- Archive ~600 PB/year of basic data products (x2)
- Provide ~21.6 PFLOPS of processing infrastructure
- Generate and archive ~100 PB/year of advanced data products (x2)
- Provide 100 FTE/year of development and operations support

SRCNet Capabilities

Science Enabling Applications
Analysis Tools, Notebooks, Workflows execution Machine Learning, etc

Distributed Data Processing
Computing capabilities provided by the SRCNet to allow data processing



Data Discovery
Discovery of SKA data from the SRCNet, local or remote, transparently to the user

Visualization
Advanced visualizers for SKA data and data from other observatories

Support to Science Community
Support community on SKA data use, SRC services use, Training, Project Impact Dissemination

Interoperability
Heterogeneous SKA data from different SRCs and other observatories

Data Management
Dissemination of Data to SRCs and Distributed Data Storage

Mini-SRCNet Demonstrator

Federating is done at the SRCNet software layer and not in the compute infrastructure layers.

- Clients, such as Execution Planning, would query various services and choose the most appropriate service based on some algorithm.

A&A is actually two different sets of services:

- Authentication and Group membership service
- Authorization as a separate service as authorization is determined differently for different resources (proprietary data, storage allocation, processing allocation, database allocation, ...)

There are no user accounts in the DRI resources. Like federating, users are known at the SRCNet software layer and do not leak into the infrastructure layer. A small set of operations accounts are required.

Mini-SRCNet Demonstrator

The key piece of software to prototype is something like the Execution Planning that determines the intersection between available compute resources, data location and size, and type of processing.

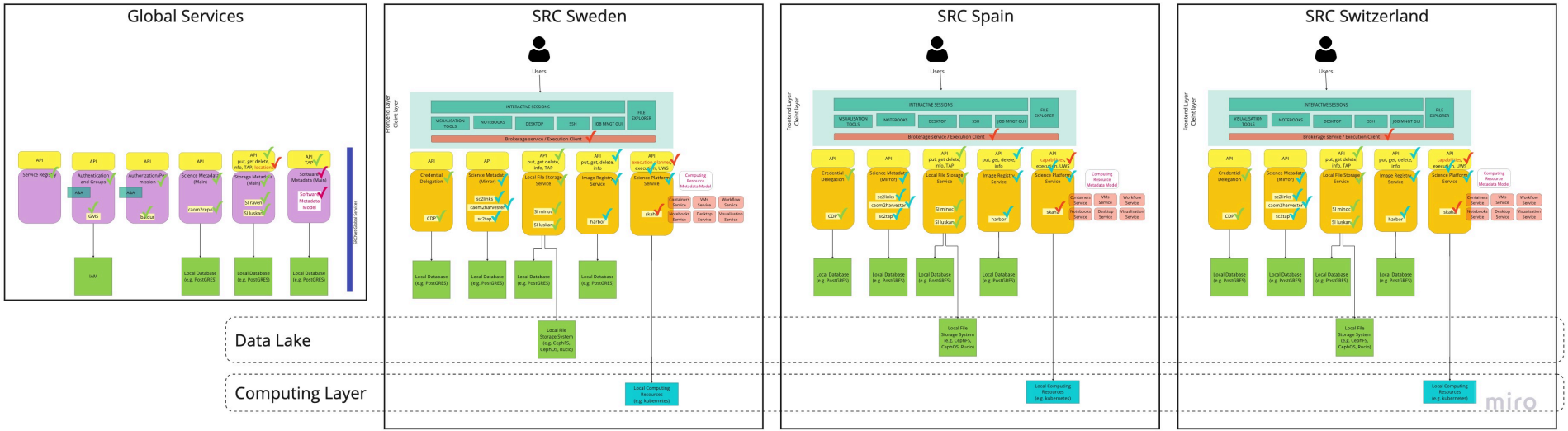
Inform:

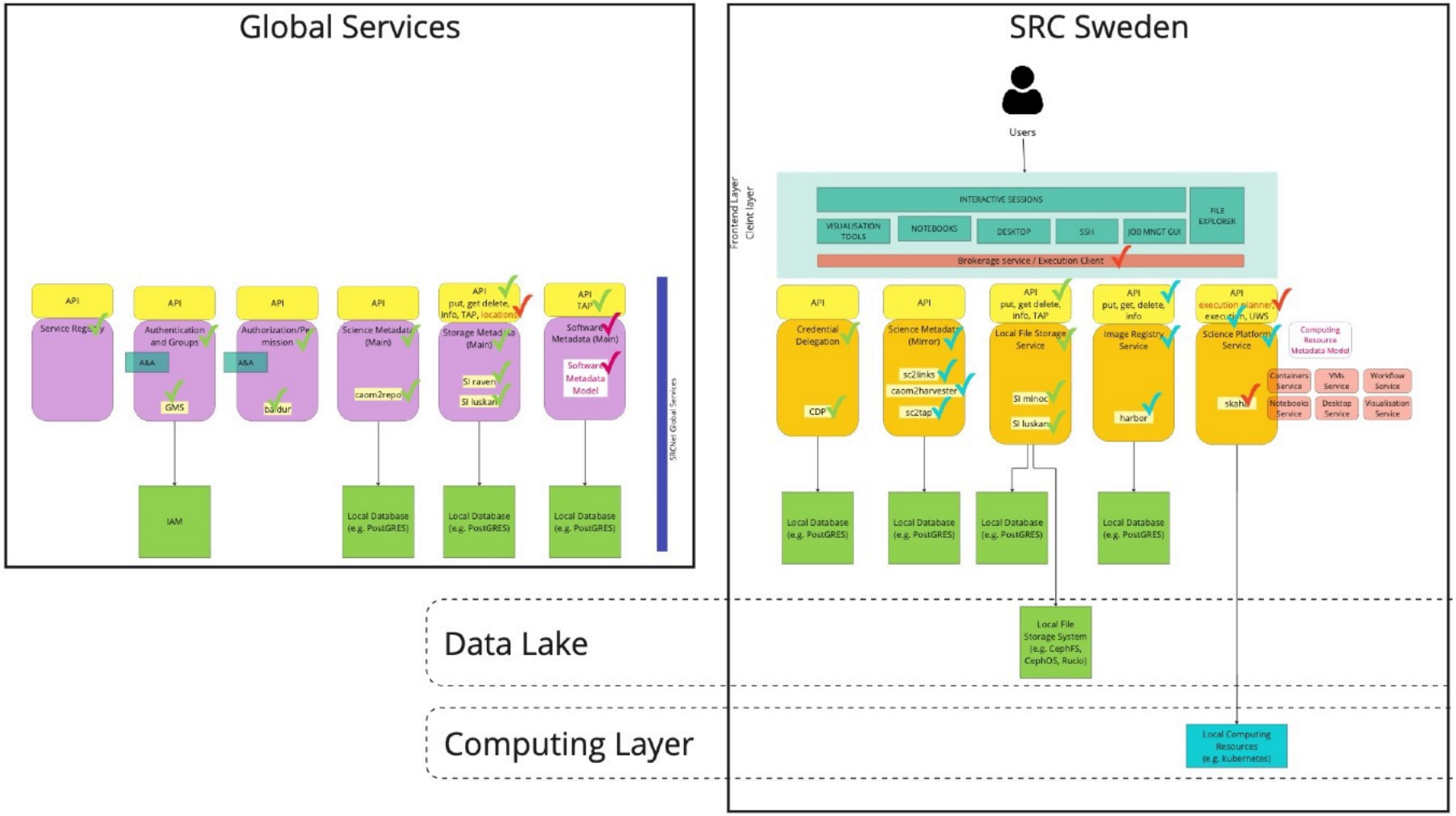
- Deployment and integration
- APIs, data models (software requiremen, compute resources, …) and IVOA standards
- Supporting multi-mission infrastructure

Two simple uses cases to start:

- Visualize a given data file.
- Discover and process data.

Mini-SRCNet Demonstrator




Global Services
API
Service Registry
API
Authentication and Groups
A&A
GMS
API
Authorization/Permission
A&A
baldur
API
Science Metadata (Main)
caom2repo
API
put, get delete, info, TAP, locations
Storage Metadata (Main)
SI raven
SI luskan
API
TAP
Software Metadata (Main)
Software Metadata Model
SRCNet Global Services
IAM
Local Database (e.g. PostGRES)
Local Database (e.g. PostGRES)
Local Database (e.g. PostGRES)
SRC Sweden
Users
Frontend Layer
Clelnt layer
INTERACTIVE SESSIONS
VISUALISATION TOOLS
NOTEBOOKS
DESKTOP
SSH
JOB MNGT GUI
FILE EXPLORER
Brokerage service / Execution Client
API
Credential Delegation
CDP
API
Science Metadata (Mirror)
sc2links
caom2harvester
sc2tap
API
put, get delete, info, TAP
Local File Storage Service
SI minoc
SI luskan
API
put, get, delete, info
Image Registry Service
harbor
API
execution planner, execution, UWS
Science Platform Service
skaha
Computing Resource Metadata Model
Containers Service
VMs Service
Workflow Service
Notebooks Service
Desktop Service
Visualisation Service
Local Database (e.g. PostGRES)
Local Database (e.g. PostGRES)
Local Database (e.g. PostGRES)
Local Database (e.g. PostGRES)
Local File Storage System (e.g. CephFS, CephOS, Ruclo)
Local Computing Resources (e.g. kubernetes)
Data Lake
Computing Layer