C-SCALE

D3.1 Initial Design of the Compute Federation

Status:	Final	**Planned due date:**	30/06/2021
Version:	V1.0	**Submission date:**	14/07/2021
Lead Participant:	EGI Foundation	**Lead Author:**	Enol Fernández
Related WP:	WP3	**Document Ref:**	D3.1
Dissemination Level:	Public (PU)		
Document Link:	https://doi.org/10.5281/zenodo.5084884		

Deliverable Abstract
The C-SCALE Compute Federation delivers access to cloud and HPC/HTC resources to support EO-related research activities. This document describes the technical architecture of this compute federation and its components.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017529.

COPYRIGHT NOTICE

C-SCALE receives funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 101017529.

DELIVERY SLIP

Date	*Name*	*Partner/Activity*
Lead Author:	Enol Fernández (EF)	EGI Foundation
Contributors:	Raymond Oonk (RO)	SURF B.V.
	Gerben Venekamp (GV)	SURF B.V.
	Giacinto Donvito (GD)	INFN
Moderated by:	Enol Fernández (EF)	EGI Foundation
Reviewed by:	Zdeněk Šustr (ZS)	CESNET
	Björn Backeberg (BB)	Deltares
Approved by:	C-SCALE Activity Management Board (AMB): Christian Briese (CB), Diego Scardaci (DS), Charis Chatzikyriakou (CC), Zdeněk Šustr (ZS), Enol Fernández (EF), Björn Backeberg (BB), Matti Heikkurinen (MH)	EODC, CESNET, EGI Foundation, Deltares

DOCUMENT LOG

Issue	*Date*	*Comment*	*Author(s)*
V0.1	2021-05-30	TOC	EF
V0.2	2021-06-15	Review ready version	EF, RO, GV, GD
V0.3	2021-06-17	Added reviewed changes from Sebastián Luna-Valero (SLV)	EF, SLV
V0.4	2021-07-03	Handled reviewer comments	EF, ZS, BB
V0.5	2021-07-07	Second iteration with reviewers	EF, ZS, BB
V1.0	2021-07-13	Version approved by AMB	C-SCALE AMB

Table of Contents

1 Introduction 7

2 C-SCALE Compute Federation 8

2.1 AAI and Use Case Onboarding 8

2.2 Software Distribution 10

2.3 Interoperability with the C-SCALE Data Federation 11

2.4 Interaction with EOSC 12

3 Cloud federation 14

3.1 Cloud Compute Providers 15

3.1.1 VM-based IaaS 16

3.1.2 Container Orchestration Platforms 16

3.2 Orchestration 17

3.3 C-SCALE AAI for Cloud 17

3.4 Provider Onboarding 18

3.5 Use Cases Onboarding 18

4 HTC/HPC Federation 20

4.1 HPC and HTC Providers 21

4.2 C-SCALE AAI for HTC/HPC 22

4.2.1 SRAM in a Nutshell 22

4.2.2 Roles 23

4.2.3 Attribute Retrieval 24

4.3 Provider Onboarding 25

4.4 Use Cases Onboarding 26

5 Conclusions 27

List of Figures

Figure 1. Cloud Federation Architecture 15
Figure 2. High-level architecture of the HPC and HTC federation. The elements highlighted in blue are the federation components delivered by the C-SCALE project. 20
Figure 3. Overview of Organisation, CO and Group elements in SRAM and the admin roles. Note: organisation admin and CO admin are different roles. 23
Figure 4. Overview of how CO information becomes available to SPs. 25

List of Acronyms

Acronym	Description
AAI	Authentication and Authorisation Infrastructure
AARC	Authentication and Authorisation for Research and Collaboration
AMB	Activity Management Board
API	Application Programming Interface
CO	Collaborative Organisation
C-SCALE	Copernicus – eoSC AnaLytics Engine
DIAS	Data and Information Access Services
EO	Earth Observation
EOSC	European Open Science Cloud
GDPR	General Data Protection Regulation
HPC	High Performance Computing
HTC	High Throughput Computing
IaaS	Infrastructure as a Service
IdP	Identity Provider
IM	Infrastructure Manager
LDAP	Lightweight Directory Access Protocol
MPI	Message Passing Interface
OIDC	OpenID Connect
PaaS	Platform as a Service
REST	Representational state transfer
SAML	Security Assertion Markup Language
SRAM	SURF Research Access Management
SP	Service Provider
SSH	Secure Shell Protocol
TOSCA	Topology and Orchestration Specification for Cloud Applications
TRL	Technology Readiness Level
VM	Virtual Machine
WP	Work Package
YAML	YAML Ain't Markup Language

Executive Summary

The C-SCALE Compute Federation brings together European commercial (e.g. Copernicus DIAS) and public (from e-infrastructure) computing providers to deliver a federated infrastructure to support Copernicus and Earth Observation (EO) use cases that deal with data- and compute-intensive workloads. The federation makes computing e-Infrastructure, potentially consisting of multiple services provided collaboratively by several service providers, accessible through homogeneous interfaces using a common AAI framework. Hence, C-SCALE facilitates the deployment of research applications that make use of Copernicus and other EO datasets, e.g. Sentinel-series satellite data, that are distributed across multiple providers.

C-SCALE Compute Federation offers interfaces to access to two types of resources:

- Cloud interface: access to Cloud resources of the federation as IaaS and container platforms (Kubernetes) with federated orchestration for the deployment of applications and platforms across providers in a seamless way;
- HPC and HTC interface: lightweight, federated and uniform access-integration layer to HPC and HTC systems.

The C-SCALE use cases and services will be deployed on top of these resources to process selected data identified through a common C-SCALE Copernicus and EO resources catalogue. Their applications will rely on the wide range of technologies supported by the federated providers (VMs, containers, batch jobs) and will be seamlessly deployed on those compute platforms thanks to federation tools.

The federation focuses on a set of basic features: (i) common authentication and authorisation compliant with the EOSC AAI that allows users to access providers according to their membership to certain communities, (ii) common software catalogue so users can easily find and distribute applications across the infrastructure, and (iii) uniform access to the data sets offered by the C-SCALE Data Federation. The cloud and HPC/HTC providers are managed independently and have specific implementations for supporting those federation features that are tailored to the needs and characteristics of each kind of resources. The implementation relies as much as possible on existing TRL 8 and higher developments from other e-Infrastructures and EOSC. C-SCALE Compute federation will become part of EOSC via the integration with EOSC Core Services (e.g. AAI) and the publication as a service in the EOSC Portal. Those interfaces developed by C-SCALE for EO data access and exploitation will be also proposed as contribution to the EOSC interoperability framework.

1 Introduction

This document describes the overall architecture of the C-SCALE Compute Federation that provides users with the baseline computing resources to perform Earth Observation related analytics on a distributed e-Infrastructure. The document first describes in Section 2 the general architecture of the federation and the common features in both cloud and HTC/HPC providers. The different federated resource types are described in Section 3 for the Cloud and in Section 4 for the HTC/HPC federation.

2 C-SCALE Compute Federation

C-SCALE Compute Federation will support EO-related compute- and data- intensive research activities within the European Open Science Cloud. The federation brings together two different types of providers to cover the different needs of the use cases to run in the infrastructure:

- HPC/HTC providers support the access to large, shared computing systems for generic batch processing via a jobs queue that is managed by a workload scheduler (e.g., Slurm). These systems are efficient and highly optimized IT solutions for data- and/or compute-intensive workloads that do not require privileged access or on-demand resources. These solutions empower users to distribute their jobs over many hardware nodes, whilst their shared nature enables high utilization of these compute resources.
- Cloud providers deliver access to IaaS resources and enable container orchestration (e.g. Kubernetes). These providers deliver a very flexible and customisable computing platform where users have complete control over the software and the supporting compute capacity. This flexibility of the computing platform enables the support of a variety of workloads: user gateways or portals, interactive computing platforms and data- and/or compute-intensive workloads using tools and frameworks that may not fit into the HPC/HTC providers. To reduce the complexity for the users to manage the complete lifecycle of the resources, C-SCALE includes orchestration tools to simplify and aid the deployment of user workloads in the infrastructure.

While these two kinds of providers form independent sub-federations, there is a set of common federation areas that are described in the sections below.

2.1 AAI and Use Case Onboarding

Access to C-SCALE compute resources uses a federated Authentication and Authorisation infrastructure (AAI) compliant with the EOSC AAI developments. This AAI builds on the previous work done by both AARC and AARC2[1] projects and the ongoing AEGIS group[2] that have contributed to establish the architecture for the EOSC AAI with the definition the AARC Blueprint Architecture (BPA)[3], which enables easy access to services and shared resources in order to collaborate.

C-SCALE relies on two technical complementary solutions that implement the AARC Blueprint Architecture for supporting the Authentication and Authorisation of users within the federation: EGI Check-in[4], for the access to cloud-based resources, and SRAM[5], for the access to HPC/HTC systems. These two solutions cater for the different existing access mechanisms for each type of providers.

On the one hand, cloud-based providers use HTTP/REST-based APIs that can be protected with OpenID Connect (OIDC). Check-in is a mature solution (TRL 9) that has been successfully used and

[1] https://aarc-project.eu/
[2] https://aarc-project.eu/about/aegis/
[3] https://aarc-project.eu/architecture/
[4] https://www.egi.eu/services/check-in/
[5] https://wiki.surfnet.nl/display/SRAM/SURF+Research+Access+Management+-+SRAM+Home

tuned to the needs of federating providers using OIDC. Check-in combines user attributes originating from various authoritative sources and delivers them to the connected service providers in a transparent way. Using Check-in simplifies the integration into C-SCALE for those providers that already are part of the EGI cloud federation as there is no need to change their setup. For new providers, the extensive use of Check-in in e-Infrastructures (more than 200 existing providers) and its tuned setup for OIDC in cloud providers make it easy to join the federation.

On the other hand, HPC/HTC systems rely on LDAP based identity management that synchronise user information at the nodes of the system, including the ssh public keys that allow users to connect to the system without an account password or passphrase. At the moment, this kind of propagation of user information based on LDAP is only supported with TRL ≥ 8 by SRAM and not by any of the other AAI solutions meeting the EOSC AAI specifications. Hence, SRAM is currently the only mature AAI choice for LDAP-based integration of HPC and HTC services.

In both cases, access to C-SCALE services is granted based on membership of users to specific communities. Each community has a manager or Principal Investigator (PI) that is responsible for (i) managing the members of the community, and (ii) the usage of the resources made available via C-SCALE services. Communities will be created after requests for C-SCALE users that are expected to be received either from the EOSC Portal, from the upcoming C-SCALE call for use cases or from the internal use cases of the project. A community creation procedure will be specified leveraging the existing procedures in Check-in and SRAM that will also include formally agreeing on and awarding the proposed use-cases with resources. Upon awarding the resources, C-SCALE, together with the relevant community, will create machine readable information, including an up-to-date list regarding these resources. The list will at a minimum include: (a) project name, (b) PI name & email, (c) project start & end data, (d) resources awarded per installation and (e) the IP addresses from where access is required (this is especially relevant for HPC nodes where the access is limited to a fixed number of IP addresses that need to be known in order to let users connect). This information will trigger the creation of the technical support resources for granting access to the providers as needed. The procedure and information needed will be defined in cooperation with the C-SCALE Data Federation to ensure it meets the overall needs of the project.

While the C-SCALE AAI supports users to define the membership to their communities, providers keep the control on the final authorisation decision (whether to grant or not access to the resources) and on the mapping of the federated users to local accounts in their systems. This mapping is performed on-the-fly for cloud resources, i.e. the local account supporting the federated user is created automatically upon the first access to the system, and via a provider-initiated provisioning mechanism supported by LDAP synchronisation for the HPC/HTC systems. The details of each AAI system for cloud and HPC/HTC are further detailed in the below chapters.

C-SCALE AAI will closely follow the developments in the EOSC AAI Federation and will seek to establish collaborations with EOSC Future to investigate further integration within the SRAM and Check-in solutions and other EOSC AAI implementations from a technical, procedural and legal/GDPR perspective.

2.2 Software Distribution

In a distributed and federated computing environment, users need solutions to efficiently manage and distribute their software across multiple providers. C-SCALE does not prescribe a single mechanism or solution to distribute software but provides users with the features of the EGI Applications Database (AppDB)[6] as an existing, production-grade, central service to store and provide information about software solutions. There are three main kind of software artifacts considered in C-SCALE:

- VM Images, provide the starting root volume of the disk for a Virtual Machine to run. Virtual Machine Images need to be packaged and distributed to the providers beforehand so users can start their VMs directly. There are several image formats (e.g. qcow2 or OVA) that can be used for packaging these images. Custom VM images can be crafted either from scratch, by creating a virtual machine and installing an operating system and the software on top of it; or from dumping an existing disk from a running VM or physical server and modify it, if needed, to run on a virtualisation platform. EGI provides guidelines on how to create these images and a set of best practices to ensure interoperability and security of the created images.
- Container images, as VM images, provide a filesystem that contains all the software dependencies needed for an application to run. There are several formats, OCI[7] supported by docker and most of the container runtimes and SIF[8], supported by Singularity. These container images can be fetched from container repositories such as Docker Hub[9] or stored in a distributed software repository as CVMFS[10]. For HPC/HTC systems, software containers must be run in the user space, thus specific container runtimes (e.g. Singularity and udocker) are available instead of docker. The creation of the container images requires root privileges. The cloud resources of C-SCALE offer users a platform that allows them to create VMs that mimic the target HPC/HTC system with root privileges to build such container images.
- Native software binaries that can be installed directly on the system. In contrast to VM images and container images, these may be non-portable and may require re-compilation for running the software on different platforms. HPC/HTC systems allow only the installation of software within user space. The use of virtual environments (e.g. virtualenv and conda) provide software installation completely in user-space without root privileges on HPC/HTC. These systems may also have pre-configured environment modules[11] that provide software applications for users, but this is not generally available in all the C-SCALE providers and requires the system administrators to manage them.

AppDB offers support for native software products and virtual machine images (support for container images is currently under development and available as a beta feature[12]) with an open

[6] AppDB https://appdb.egi.eu/
[7] Open Container Intiative (OCI) https://opencontainers.org/
[8] Singularity Image Format (SIF): https://github.com/hpcng/sif
[9] Docker Hub https://hub.docker.com/
[10] CVMFS: https://cernvm.cern.ch/fs/
[11] Environment modules: http://modules.sourceforge.net/
[12] AppDB container catalogue: https://appdb-dev.marie.hellasgrid.gr/browse/containers/cloud

browsable catalogue with information on the software metadata that facilitates reusing and sharing software across user communities. AppDB is open to any registered user for publishing software, while downloading software can be done by both registered and anonymous users using HTTP. Communities can create collections of VM/docker images that cloud providers can subscribe to. The subscription enables the periodic download, conversion, and storage of those images in the local image repository of the provider. The Compute Federation will work with C-SCALE communities to create a collection of software in AppDB that will include those applications required by the use cases.

HPC systems in general have stricter software policies and re-compilation of the applications with the libraries and configuration of each specific system is highly recommended to ensure optimal performance from the hardware. Partners providing HPC systems in C-SCALE will support re-compilation of the applications when needed. This is especially relevant for parallel MPI applications, where the correct tuning can have a considerable impact on the network performance of the applications.

Besides AppDB, other external software distribution solutions, outside of the C-SCALE project, exist already and are publicly available to users. These solutions are often interoperable with C-SCALE Compute Federation infrastructure and can also be used by end-users for obtaining software solutions, although they may not have direct support by the providers within the federation. A few publicly available, external solutions that we can mention here are: B2SHARE[13], Zenodo[14], GitHub[15], SingularityHub[16] and Docker Hub.

2.3 Interoperability with the C-SCALE Data Federation

C-SCALE Data Federation supports the redistribution of Copernicus and Earth Observation products for processing in the compute federation. Whenever possible, the user workloads should run on compute resources where the data is available locally to avoid any pre-staging operations to make the data available for the user analysis. The data federation allows to:

1. Perform metadata queries to discover the relevant products and provide a list of locations where to fetch the data. The query services in C-SCALE will use the STAC-API[17] as a query interface and are publicly accessible, i.e. do not require any authentication. Users (or tools like the PaaS Orchestrator) may dynamically chose computing providers to minimise data transfers by performing the metadata queries before submitting the actual workload.
2. Stage the data so it can be accessed by the analytics processes running in the compute platform. Data will be available via HTTP-based endpoints and rely on OpenID Connect (OIDC) federated authentication with EGI Check-in (as the C-SCALE Cloud Compute providers), i.e.

[13] B2SHARE: https://b2share.eudat.eu/
[14] Zenodo: https://zenodo.org/
[15] GitHub: https://github.com/
[16] SingularityHub: https://singularityhub.com
[17] STAC-API: https://stacspec.org/STAC-api.html

staging products requires a valid authenticated user. Tools like rclone[18] deliver a good user experience to interact with the data providers.

The different compute systems of the C-SCALE federation can have very different internal and external network setups and policies and hence a wide variety in connectivity. On the one hand, cloud and HTC systems, designed for data-intensive processing, typically have high bandwidth connectivity to the outside world. On the other hand, HPC systems focus on compute-intensive processing and aim to deliver a high speed, internal network between the worker nodes, often do not allow external connectivity from the processing nodes and sometimes also have limited bandwidth from the login nodes to the outside world.

These differences for HPC and HTC systems have consequences for data logistics that may impact how these systems interact with the Data Federation components, which may lead to further developments to support the most appropriate data query and retrieval strategy for each case. In general, use-cases can retrieve and store their data in the home and project spaces created for each use-case on the systems to which the use-case has been granted access by C-SCALE. Depending on the network setup and storage medium of the data required by a use-case, on-the-fly data staging and transport may also be possible and in fact be efficient on certain systems in C-SCALE. Each use-case will therefore be evaluated in terms of data logistics by the co-design activities of the project with support of the Data Federation and the participating providers in each of the use cases.

User authentication for interacting with the Data Federation uses OpenID Connect access tokens. These are short-lived (e.g. 1 hour) tokens that can be generated by initiating an authentication flow against Check-in. This requires the user to interact with the browser and a valid Check-in client configured to perform the authentication. This solution is not convenient for offline workloads as the user cannot login with the application and also requires the application to act as a web server so it has limited usage on certain systems. For offline processing, refresh tokens provide a more convenient method to interact with services that require authentication. Refresh tokens are longer living credentials (e.g. months) that allow an application to obtain new access tokens (short lived) whenever they are needed. This method supports offline processing as the obtention of new access tokens does not require any user interaction. Refresh tokens must be handled with special care, as they allow for impersonation of users. Tools like oidc-agent[19] support secure obtention, management and bundling these tokens with the user workload to the VMs, containers, or jobs in a distributed infrastructure so data can be reached easily.

2.4 Interaction with EOSC

C-SCALE will make the Compute Federation available to EOSC users via the ESOC Portal as a service alongside the Data Federation and Analytics Platforms. As such, the Compute Federation will align its providers with any applicable future EOSC principles, standards and values including compliance with the Rules for Participation and Interoperability Framework. The guidelines to join the C-SCALE federation will be adapted to include all the necessary steps each provider should accomplish to

[18] rclone: https://rclone.org
[19] oidc-agent: https://indigo-dc.gitbook.io/oidc-agent/

satisfy the policy, operational and technical requirements to become part of the EOSC. The project will facilitate this integration supporting the providers during the onboarding process and representing them in EOSC. This will reduce the burden for the EOSC Portal operators that should interact with only a coordinator that will represent several service providers.

The access to resources is supported by the technical solutions used in C-SCALE AAI (EGI Check-in and SRAM). These are already compliant with the existing EOSC AAI guidelines. Support for future requirements will be evaluated within C-SCALE where possible. Integration with other EOSC Core services (e.g. accounting, monitoring, helpdesk) will be done in C-SCALE at federation level. C-SCALE already relies on the existing EGI federation whenever possible, thus these adaptations will be available automatically as EGI also adapts to meet the upcoming EOSC requirements.

C-SCALE Compute Federation follows existing EOSC-hub technical specifications[20], (such as the ones for Cloud Compute and PaaS Solutions). These should enable interoperability with other EOSC services of the same area. Future EOSC interoperability frameworks and guidelines will be followed when applicable and where possible. C-SCALE is also expected to contribute to the definition of the EOSC interoperability framework defining interoperability guidelines for EO data access and exploitation.

[20] EOSC-hub technical documentation: https://www.eosc-hub.eu/technical-documentation

3 Cloud federation

The Cloud Federation of C-SCALE provides access to cloud IaaS (e.g. OpenStack) and container orchestration services (e.g. Kubernetes) for the execution and deployment of Copernicus and EO-related workloads. As shown in Figure 1, C-SCALE cloud federation has a layered architecture. A set of infrastructure providers support the execution of workloads as VMs (IaaS) or as containers (container orchestration systems) alongside with the access to data on storage resources that can interoperate with the Data Federation. This base layer provides APIs for the management of the compute and storage resources as a Service that both upper layers of the architecture and users can interact with directly. These underlying providers integrate with federated AAI mechanisms so users can seamlessly access them.

On top of the infrastructure layer, a set of federation components provides a multi-cloud orchestration framework that facilitates the execution of user applications on the distributed and federated infrastructure. The IaaS Orchestration layer provides a unified and homogeneous API for managing resources in the multi-cloud environment using the TOSCA standard[21]. The orchestrator deploys complex and customized virtual infrastructures leveraging the Infrastructure Manager (IM)[22] as its implementation over IaaS. IM automates the deployment, configuration, software installation, monitoring and update of virtual infrastructures. It supports a wide variety of back-ends, including both public IaaS Clouds (Amazon Web Services, Microsoft Azure, etc.), on-premises Cloud Management Platforms (OpenNebula, OpenStack, etc.) thus making user applications IaaS Cloud agnostic.

A PaaS Orchestration[23] layer enables the coordinated provisioning of virtualized compute and storage resources on IaaS cloud platforms and the deployment of containerised long-running services and batch jobs on container orchestration platforms. As the IaaS Orchestrator, it receives the deployment requests, expressed through templates written in TOSCA and orchestrates them on the best available infrastructure sites. To select the best site, the Orchestrator manages several sources of information: SLAs signed by the providers with the user, the monitoring of data about the availability of the compute and storage services, and the location of the data requested by the user (if any). The PaaS Orchestration layer creates a virtualised cloud where hybrid deployments spanning multiple providers are supported. The PaaS Orchestration will be used to enable the repeatable deployment of user facing platforms (e.g. batch processing platforms like openEO[24] and interactive analysis environments like Jupyter) into the Compute federation.

All the components foreseen in the architecture have mature and production-ready TRL ≥ 8 implementations to be used in C-SCALE. The Compute Federation task will work on improving the integration among these and other C-SCALE components (Data Federation, Batch and interactive processing environments) and the integration with the use cases.

[21] https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=tosca
[22] https://www.grycap.upv.es/im/
[23] https://indigo-dc.gitbook.io/indigo-paas-orchestrator/
[24] https://openeo.org/

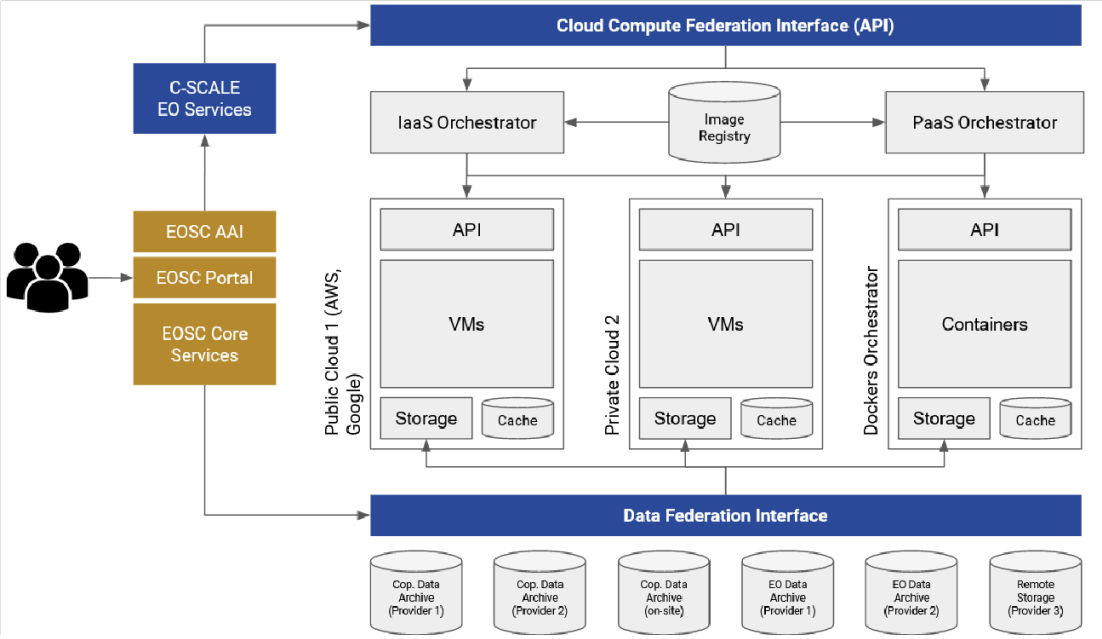

Figure 1. Cloud Federation Architecture

C-SCALE Copernicus and EO Services and user applications make use of any of the three available layers depending on the application needs: VMs & container orchestrators, IaaS orchestration or PaaS orchestration. Each of these layers offer different levels of abstraction and control. While for most cases the use of PaaS Orchestration is recommended, direct access to low-level IaaS APIs may be needed for specific features not available in the upper layers. All of them are integrated with the same AAI, thus identities of users are delegated from one layer to the next and users can manage and control their resources created from upper layers using the APIs of lower layers.

3.1 Cloud Compute Providers

All Compute Providers of the federation are discoverable via an online catalogue supported by the EGI Configuration Database (previously known as the GOCDB)[25], that contains general information about the providers participating in the EGI production infrastructure. It is a central registry to record information about the topology of the e-Infrastructure. This includes entities such as providers (Resource Centres in the EGI Configuration Database nomenclature), service endpoints and their downtimes, contact information and roles of users responsible for operations at different levels. The EGI Configuration Database allows to selectively tag endpoints and providers so that users and API queries filter for objects that define the required set of tags. A new "C-SCALE" tag will be created to easily discover and manage those providers supporting the federation. Two kinds of compute endpoints will be registered in this database: VM-based IaaS powered by OpenStack (org.openstack.nova and org.openstack.swift endpoint type in the EGI Configuration Database) and Kubernetes container orchestration endpoint (io.cncf.kubernetes endpoint type in the EGI Configuration Database). These two types of resources are described below.

[25] https://goc.egi.eu

Having the providers registered in the configuration database enables all the actors (end users, provider managers, support teams, etc.) and infrastructure tools (accounting, monitoring, orchestrators) to discover information about the infrastructure topology.

3.1.1 VM-based IaaS

IaaS providers deliver on-demand API-based access to computing resources as Virtual Machines that can run user-defined arbitrary software (including operating systems and applications). These providers also allow management of storage that can be accessed from VMs and management of network features to provide connectivity to those VMs.

C-SCALE relies on the existing EGI Cloud Federation[26] to build the IaaS layer. The federation supports IaaS providers that can integrate with the federated EOSC AAI as implemented by Check-in. APIs of the providers must be supported by the upper Orchestration layers - although there are some standards proposed for this kind of systems (most relevant are OCCI[27] and CIMI[28]), they have no support from most vendors and/or providers and therefore limited benefits.

Consumption of resources is tracked via a federated accounting system[29] that provides an integrated view about resource/service usage. The accounting system pulls together usage information from the federated providers, integrates the data and presents them in such a way that both individual users as well as whole communities can monitor their own resource/service usage across the whole federation.

3.1.2 Container Orchestration Platforms

Container orchestration platforms provide on-demand API-based management of container-based applications. They support the (automated) management of the complete lifecycle of the containers that compose an application into a set of computing resources. Several competing container orchestration platforms exist, each of them with their own and not interoperable API. Kubernetes has gained a lot of momentum in recent years and can be considered the de-facto standard in this area.

C-SCALE will focus on supporting Kubernetes as container orchestration platform for the use cases. The Kubernetes clusters can be made available either as a managed service (i.e., the deployment and configuration are performed by the provider) or deployed on top of VM-based IaaS by users manually or using tools like the IaaS Orchestrator. Kubernetes supports OpenID Connect authentication and this will be the preferred mechanism for C-SCALE configured with Check-in.

Currently, there is no production-ready federated accounting available for Kubernetes that can provide usage metrics to EOSC. C-SCALE will explore existing efforts to adopt and contribute to so

[26] https://www.egi.eu/federation/egi-federated-cloud/
[27] Open Cloud Computing Interface (OCCI): https://occi-wg.org/
[28] Cloud Infrastructure Management Interface (CIMI): https://www.dmtf.org/sites/default/files/standards/documents/DSP0263_2.0.0.pdf
[29] APEL Accounting: https://apel.github.io/

that accounting mechanisms for Kubernetes compliant with future EOSC specifications are available for the federation.

3.2 Orchestration

C-SCALE orchestration will help the end user not only to deploy seamless applications across different and heterogeneous infrastructures, but also to deal easily and powerfully with the data aware deployment. This approach, thanks to the features provided by the PaaS Orchestrator, will enable the end user to leverage the compute infrastructure closest to the data in order to improve the performance of the data analysis.

By means of a unified AAI solution users can easily access all the Cloud resources already available within the EOSC ecosystem, not only IaaS based but also in the form of container orchestrators that could easily open the possibility to exploit bare-metal HPC-like resources that expose Cloud-like APIs/capabilities.

IM is the implementation of the automation at IaaS Level by means of a standard library that is able to support a wide range of IaaS Cloud providers, including commercial public clouds like AWS, Google Cloud Compute or Microsoft Azure, and Open Source based ones like OpenStack, OpenNebula or CloudStack. The INDIGO PaaS Orchestrator is a component of the PaaS layer that allows users to instantiate resources on Cloud Management Frameworks (like OpenStack and OpenNebula, AWS, Google, Microsoft Azure, etc) and Mesos and Kubernetes clusters.

It takes the deployment requests, expressed through templates written in TOSCA YAML Simple Profile, and deploys them on the best cloud site available. In order to do that it 1) gathers SLAs, monitoring information and other data from platform services, and 2) asks the cloud provider ranker for a list of the best cloud sites. The exposed REST APIs are consumed by a web dashboard portal that offers a graphical user interface to facilitate usage of the orchestrator features.

3.3 C-SCALE AAI for Cloud

The Authentication and Authorization Infrastructure (AAI) for cloud providers of C-SCALE reuses existing elements of the EOSC AAI, particularly, it uses EGI Check-in for managing access to the providers.

EGI Check-in is a proxy service that operates as a central hub to connect federated Identity Providers (IdPs) with service providers. Check-in allows service providers to manage access to their services from users holding identities (e.g. usernames and passwords) from a very broad set of academic, community or social Identity Providers (IdPs) so EOSC users can access and use services in a uniform and easy way. Check-in brings together these IdPs, the service providers (SPs) and intermediary identity management proxies into a single, trusted, secure and interoperable infrastructure. It builds on open technologies including SAML 2.0, OpenID Connect (OIDC), OAuth 2.0 and X.509v3 to offer a flexible framework for access management.

Within C-SCALE, OpenID Connect (OIDC) is used as the main federated authentication protocol. OIDC is supported in all the C-SCALE cloud layers, from the underlying resources delivered with

OpenStack (IaaS) and Kubernetes (Container orchestration) to the IaaS and PaaS Orchestrators. Check-in provides detailed documentation on how to set-up the authentication and authorisation on providers[30], with specific details on how to configure OpenStack also available[31] . As described in Section 2.1, authorisation of users is performed locally at the providers, by checking the membership of users to the supported communities by the providers. Communities are managed as Virtual Organisations (VOs) in Perun[32], which provides user registration and approval and group-based authorization if needed.

3.4 Provider Onboarding

C-SCALE relies on the existing EGI Federation to bring providers into the cloud resources of C-SCALE. EGI has well defined procedure for onboarding new providers[33]. This procedure ensures, not only that providers follow all the technical integration steps, but also ensures that they are properly registered in the EGI Configuration Database with all the needed contacts (e.g. site admin, security) and that all the federation policies are accepted, ensuring the minimum operational conditions for to join (e.g. expected monthly availability and reliability, security coordination policies, etc.).

The registration process includes enabling the automated monitoring of the providers' services exposed to the federation. Monitoring is a key service needed to gain insights into an infrastructure. It needs to be continuous and on-demand to quickly detect, correlate, and analyse data for a fast reaction to anomalous behaviour. EGI relies on the ARGO system[34] for collecting status results from one or more monitoring engine(s) and delivers status results and/or monthly availability (A) and reliability (R) results of distributed services. Both status results and A/R metrics are presented through a Web user interface, with the ability for a user to drill-down from the availability of a site to individual services, to individual test results that contributed to the computed figure. Argo is capable also to send notifications to the service admins in case of a failure/warning on one of the services monitored.

As part of joining the federation, providers will also integrate into the EGI accounting service that collects, stores, aggregates, and displays usage information about the consumption of resources. This usage data is gathered from the providers by probes and sensors according to certain data formats (e.g. usage of CPU by VMs started at the providers). The usage data is forwarded from the sensors into a central Accounting Repository where those data are processed to generate various summaries and views for display in the Accounting Portal[35].

3.5 Use Cases Onboarding

Once a use case has been granted access to cloud resources, C-SCALE needs to set up the technical components that will support the access of the users to the resources. Each use case is mapped to

[30] Check-in Service provider integration: https://docs.egi.eu/providers/check-in/sp/
[31]OpenStack Check-in integration: https://docs.egi.eu/providers/cloud-compute/openstack/#egi-aai
[32] Perun: https://perun-aai.org/
[33] PROC 09: https://wiki.egi.eu/wiki/PROC09_Resource_Centre_Registration_and_Certification
[34] ARGO: https://argo.egi.eu
[35] EGI Accounting portal: https://accounting.egi.eu/

one or multiple Virtual Organisations (VOs). VOs are created by filling in a form in the EGI Operations Portal[36] following EGI Procedure 14[37]. This procedure, upon registration of the VO information in the Operations Portal, will trigger the creation of the VO in Perun where the VO Manager (Principal Investigator of the community) will be able to administer user membership.

Once the VO is created, a new VM image list will be created in AppDB, where the community can add those images with the software that needs to be available at the providers. By default, CentOS 7 and Ubuntu 20.04 LTS images created by the EGI support team are added to the list so the community can get started right away with the usage of the resources.

Finally, the VO is also configured at the IaaS/Container orchestrator providers supporting the use cases by setting up the mapping from EGI Check-in VO membership information provided as OpenID Connect claims to the local groups.

The support to the use case is formalised in a SLA (Service Level Agreement) backed up by a set of OLAs (Operation Level Agreement), one for each of the providers, where the expected service level objectives and service level indicators are agreed upon.

[36] EGI Operations Portal: https://operations-portal.egi.eu/vo/
[37] PROC14 VO Registration https://wiki.egi.eu/wiki/PROC14_VO_Registration

4 HTC/HPC Federation

The HPC and HTC Federation of C-SCALE provides access to and integration of HPC and HTC platform solutions for the execution and deployment of EO-related workloads. The goal is to support EO use cases that deal with data- and compute-intensive workloads. In particular, the federation focuses on enabling easy and uniform access to the HPC and HTC systems offered by the C-SCALE partners and provide the baseline input for a blueprint for extending the EO compute federation beyond C-SCALE as well as integration with potential future, external partners.

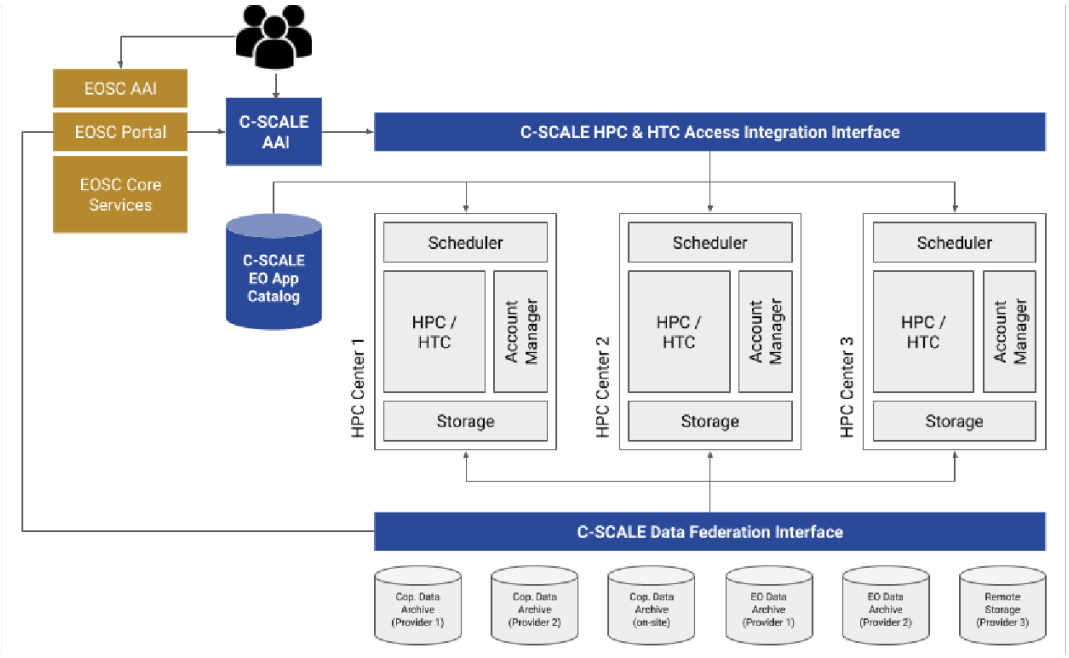

Figure 2. High-level architecture of the HPC and HTC federation. The elements highlighted in blue are the federation components delivered by the C-SCALE project.

The high-level architectural design for the HPC and HTC federation, shown in Figure 2, is aimed at solving the difficulty that users experience in accessing different HPC and HTC systems by using the C-SCALE AAI. The selected solution is based on the SURF Research Access Management (SRAM) system (see Section 4.2). Figure 2 also shows the interfaces towards the data federation, the application catalogue (see Section 2.2) and EOSC itself.

Access to the HPC and HTC federation provides the user with access via SSH to the user interface (UI) nodes at the HPC/HTC systems and the compute & storage resources that were granted to the use-case by C-SCALE. The underlying scheduling and batch processing infrastructure at each of the systems provided through the federation will then allow the data- and compute intensive EO workloads to be efficiently scheduled and processed at scale.

Upon granting a request for access to the HPC and HTC federation, C-SCALE will create the collaborative organisation (CO) in SRAM that is required to provide access for the use-case. C-SCALE will invite the PI to become an admin member of the CO. This member role will empower the PI to

conduct and lead the user management for his/her use-case team regarding access to the HPC and HTC resources delivered by the federation (e.g., the PI can invite new members and remove existing members from the CO). Once the PI, and other CO members, have accepted the invitation, created their account in SRAM and uploaded their public SSH key in their user profile, they will be able to access, over SSH, all HPC and HTC systems for which access has been granted by C-SCALE using the same SSH key pair.

To further improve the usability of the HPC and HTC federation for Copernicus and EO data workflows, just delivering compute and storage resources is not enough to ensure a smooth user experience. The C-SCALE Compute Federation will work closely with the data federation and provide support for deploying the data query, access and transfer tools developed for that purpose. Similarly, the data federation design and tools will take the constraints provided by the design of the HPC and HTC federation into account in their development plans.

Equally important for Copernicus and EO users will be how to invoke requisite software. On this topic, the HPC/HTC federation members will work closely with the use cases and provide support for deploying the needed software by use cases. The use-case onboarding process of the project must consult with the HPC and HTC federation (regarding a.o., software requirements) before granting a use case to ensure they can be supported on the federated resources.

4.1 HPC and HTC Providers

HPC and HTC service providers provide access to large, shared computing systems for batch processing via a jobs queue that is managed by a workload scheduler (e.g., Slurm). These systems are aimed at providing highly efficient IT solutions for data- and/or compute-intensive workloads that do not require privileged access or on-demand resources. These solutions empower users to distribute their jobs over many CPU (and GPU) cores and nodes, whilst their shared nature enables high utilization of these compute resources.

To become an HPC/HTC provider in the C-SCALE HPC and HTC federation, an organisation must support transnational access (at minimum within Europe) to a compute system that is governed by a workload scheduler and can be accessed via SSH keys. The provider must support SRAM as an authorization and authentication infrastructure. In particular, the provider must support the LDAP protocol provided and supported by SRAM. The providers must support the usage/installation of the C-SCALE use-cases related software targeting the HPC and HTC federation.

The HPC/HTC federation does not currently have a common monitoring and accounting mechanism that can be used to have a global view of the Availability and Reliability of the providers and the consumption of resources from the different providers. The providers are expected to have their own procedures and processes in place to provide those and report back to the project whenever requested (e.g. when reporting VA usage metrics from the supported use cases). C-SCALE will study the feasibility of further integrating these types of providers into existing monitoring and accounting frameworks, such as ARGO monitoring and APEL accounting used in the EGI federation, or similar systems that EOSC will provide to support these features. The introduction of federated monitoring would allow to impose minimum reliability and availability requirements on the providers for joining C-SCALE (e.g. at least 90 percent).

4.2 C-SCALE AAI for HTC/HPC

An Authentication and Authorization Infrastructure (AAI) is aimed at providing seamless access to services and their resources based on federated identities, i.e. reuse of a single identity. For C-SCALE's HTC/HPC federation we have selected SRAM to fulfil this requirement. SRAM is built upon open source components and adheres to the AARC blueprint architecture. This blueprint architecture has been created by the international research community to enable easy access to services and shared resources in order to collaborate. One of the components SRAM uses is eduTEAMS[38] by GÉANT. eduTEAMS is built on top of eduGAIN[39]. Hence, eduTEAMS makes registered IdPs/SPs within eduGAIN findable to SRAM.

4.2.1 SRAM in a Nutshell

In its simplest form, SRAM offers group management, identities validated by their home organisation and attributes. These attributes can be obtained through either the SAML, OIDC or LDAP protocols. In order to achieve this, SRAM has the following three concepts: i) Organisation, ii) Collaborative Organisation and iii) Groups. At the top there is the Organisation and it can create Collaborative Organisations (COs). Groups can be created within a CO. This way a hierarchy is created, where a single organisation can hold any number of COs, which in turn can have any number of groups. Depending on your role within SRAM, you can create and manage COs and groups, invite users and add or remove those to and from groups (See Figure 3).

[38] https://www.geant.org/Services/Trust_identity_and_security/Pages/eduTEAMS.aspx
[39] https://www.geant.org/Services/Trust_identity_and_security/Pages/eduGAIN.aspx

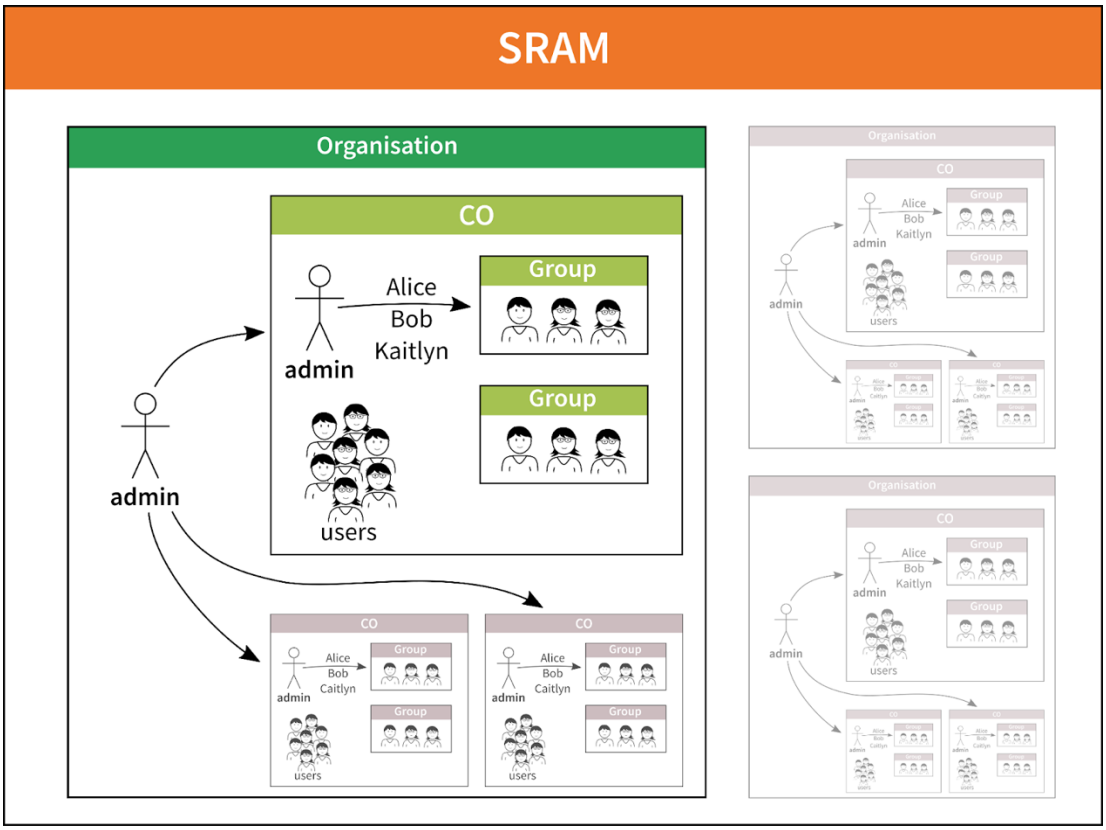

Figure 3. Overview of Organisation, CO and Group elements in SRAM and the admin roles. Note: organisation admin and CO admin are different roles.

4.2.2 Roles

At the highest level, i.e. the organisation, you have the organisation admin and organisation manager roles. Both roles are able to create new COs. The admin role has more privileges in managing the organisation itself, whereas a manager does not have these privileges. For more information about these roles see: https://wiki.surfnet.nl/display/SRAM/Organisation+admins+and+managers

When a CO is created, either the organisation admin or organisation manager is able to select or invite a CO admin. CO admins can be added or removed during the lifetime of the CO. Both the organisation admin and organisation manager have the same abilities as a CO admin. These are: i) editing CO information, ii) inviting CO admins and members, iii) managing groups and services, iv) removing members from groups or from the CO itself. More information can be found at: https://wiki.surfnet.nl/display/SRAM/CO+admins

Lastly, there are the members. As a member, which includes of course also the organisation admin or manager and the CO admin, you have a profile. One of the things you can do with that profile is to upload and manage a public SSH keys. You can also see the group that you are a member of and see who all other members within the CO are. For more information, see: https://wiki.surfnet.nl/display/SRAM/COllaboration+member

4.2.3 Attribute Retrieval

Being able to do group management is one side of the equation. The other is retrieving useful information. Any member in SRAM needs to provide an authenticated identity. This is done through either SAML or OIDC. Although the vocabulary between SAML attributes and OIDC claims differ, SRAM does not make a distinction between the two and uses SAML vocabulary only. For example, in SAML we speak about Identity Provider (IdP) and Service Provider (SP), where OIDC uses OpenID Provider (OP) and respectively Resource Provider (RP). For details on the attribute mapping to and from claims, see the White Paper for implementation of mappings between SAML 2.0 and OpenID Connect in Research and Education[40]. When members log into SRAM they login through their (selected) IdP and in doing so, that IdP releases a number of attributes. At a minimum there is a unique identifier. Other attributes for example are: first and last name or e-mail address. These can be used by SRAM to create a profile. In addition to these profiles, SRAM adds the Organisation, CO and group information of members. This information can be picked up by SPs for further processing. As stated before, using either SAML, OIDC or LDAP should yield the same information.

Attribute retrieval with LDAP

SRAM offers an LDAP for attribute retrieval at ldaps://ldap.sram.surf.nl. This is a read-only LDAP for external access. The LDAP is subdivided into subtrees. Each SP is given its own subtree. This ensures that SPs cannot obtain any information that is not destined for that SP. Only when a SP is connected to a CO, will the CO information become visible within the subtree of the SP. In case the same CO information needs to be available for multiple SPs, then each SP will get its own copy in its subtree.

Figure 4 depicts how CO information ends up in LDAP: all information is stored centrally in a database. In Figure 4, there are two sample services (Geminia and Delena) each having their own LDAP subtree (dc=Geminia, dc=services, dc=sram, dc=surf, dc=nl and dc=Delena, dc=services, dc=sram, dc=surf, dc=nl). Collaborations 'Hawking Radiation' and 'String Theory' make use of Gemina and thus appear in the subtree for Geminia. Collaborations 'String Theory', 'Dark Energy' and 'Event Horizon' make use of the Delena service and hence appear in that sub tree. As can be observed from the coloured asterisks, 'String Theory' is available to both 'Geminia' and 'Delena'.

[40] https://docs.google.com/document/d/1b-Mlet3Lq7qKLEf1BnHJ4nL1fq-vMe7fzpXyrq2wp08/

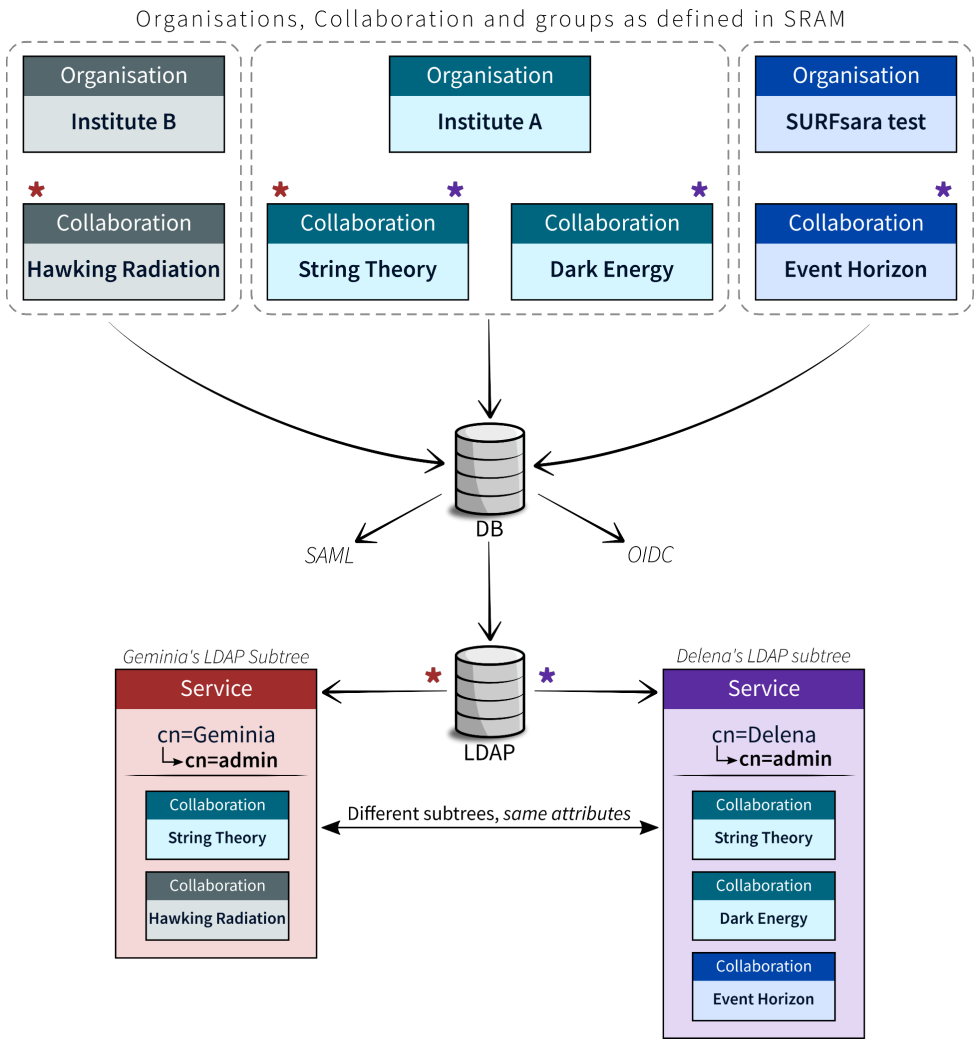

Figure 4. Overview of how CO information becomes available to SPs.

4.3 Provider Onboarding

From a technical perspective, Service Providers (SPs) wishing to become part of the C-SCALE HPC and HTC federation, need to be: i) discoverable within SRAM, and ii) coupled with a CO by a CO admin. First, an agreement is required between a CO, SRAM and an SP so a service delivered by a SP can be used by a CO via SRAM. Once this agreement has been established the CO admin will be able to find the SP, can add the service to the CO and by doing so make it available to the members of that CO. Upon coupling the service to a CO, an LDAP subtree for that CO in SRAM becomes available for the SP. This LDAP subtree is uniquely prepared for the SP, see Figure 4, and can be used by the SP for user management and access provisioning.

The current version of SRAM does not yet support an SP landing page. Connecting a service starts by the SP mailing SRAM support at: sram-support@surf.nl. This email needs to contain information about the organisation, service details, contact information, service provider policies, SRAM collaboration connections and technical information. This procedure is explained at

https://wiki.surfnet.nl/display/SRAM/Connecting+a+service+to+SRAM and describes in detail what information needs to be provided in the email. After processing the connection request for the service, the SP will receive a reply with all necessary information on how to connect.

In addition to the technical connection, it is important for a SP to know that it can specify how to treat organisations requesting to make use of its service offered via SRAM: abandon request automatically, create the connection automatically (and get informed by email), or get informed by email of the request so the SP can provide SRAM with a 'go or no go'. Further information can be found at: https://wiki.surfnet.nl/display/SRAM/SRAM+for+providers+of+resources+for+researchers.

4.4 Use Cases Onboarding

The differences for the HPC and HTC systems with regard to software and data access (see Sections 2.2 and 2.3) also have consequences for the selection and matching of use-cases. Each new use-case requires an in-depth review in order to understand e.g., the technical aspects of the proposed workflow (including data logistics) as well as the expectations of the use-case team. This is needed in order for C-SCALE to provide the best possible mapping of a use-case to the C-SCALE enabled infrastructure. This process will be led by the use cases, in close collaboration with the data and compute federation.

From a technical point of view the HPC and HTC federation will support use-cases with: (i) access and account provisioning & deprovisioning and (ii) deployment of C-SCALE use-cases software for the HPC and HTC federation. In terms of support for running the use-case workflows on the HPC and HTC infrastructure the first line of support will be handled by the user support activities included in the project. Where needed, the first line support will involve the HPC and HTC federation in technical issues as they pertain to the supported infrastructure and are reported by the use-cases. Specific support requests regarding e.g., software required by the use cases that is not easily deployed on the HTC/HPC federation, or systems that are external to C-SCALE, are outside the scope of this activity and will only be addressed by the HPC and HTC federation on a best-effort basis.

5 Conclusions

The C-SCALE Compute Federation provides access to a wide range of computing providers (IaaS VMs, container orchestration platforms, HPC, and HTC systems) to enable the analysis of Copernicus and Earth observation data under EOSC. The design of the federation allows users to deploy their applications using federated authentication mechanisms, find their software under a common catalogue, and have access to data using C-SCALE Data Federation tools. The federation relies on existing (TRL ≥ 8) tools and services already compliant with EOSC, thus facilitating the integration into the larger EOSC ecosystem of providers. Currently, those C-SCALE providers not yet federated, are going through the process of joining, following the processes described in this document. Additionally, the use-cases already identified are piloting existing onboarding procedures to access C-SCALE cloud, HPC and HTC resources. These processes will be further refined and adapted to make it possible for other external providers to join the federation in the future.