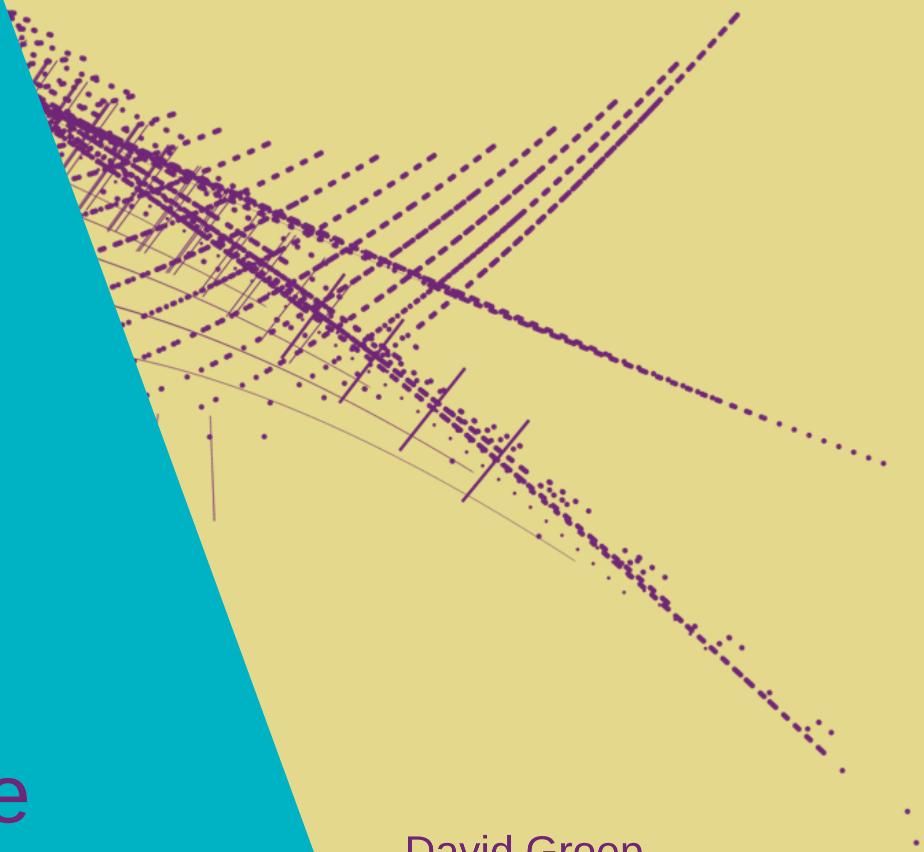
Nikhef

Maastricht University

RWTH colloquium

Infrastructure:
for the small and the large

build, trust, collaborate, govern

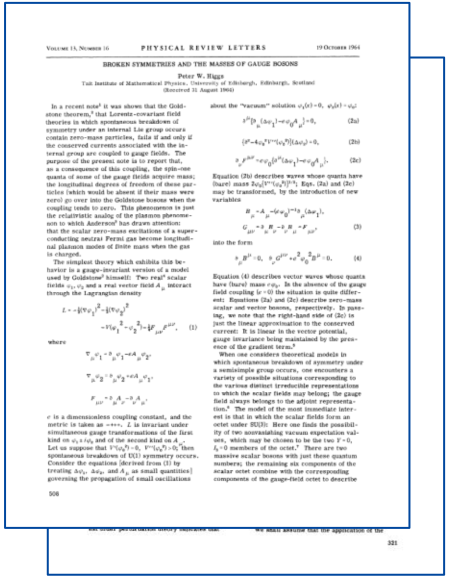


David Groep,
February 2026

Peter Higgs and Francois Englert at the 2013 Nobel prize press conference, Stockholm. Photo: Bengt Nyman, https://www.flickr.com/photos/97469566@N00

Exploding data? the Large Hadron Collider at CERN

1964



VOLUME 13, NUMBER 16 PHYSICAL REVIEW LETTERS 19 OCTOBER 1964

BROKEN SYMMETRIES AND THE MASSES OF GAUGE BOSONS

Peter W. Higgs

Tait Institute of Mathematical Physics, University of Edinburgh, Edinburgh, Scotland

(Received 31 August 1964)

In a recent note[1] it was shown that the Goldstone theorem,[2] that Lorentz-covariant field theories in which spontaneous breakdown of symmetry under an internal Lie group occurs contain zero-mass particles, fails if and only if the conserved currents associated with the internal group are coupled to gauge fields. The purpose of the present note is to report that, as a consequence of this coupling, the spin-one quanta of some of the gauge fields acquire mass; the longitudinal degrees of freedom of these particles (which would be absent if their mass were zero) go over into the Goldstone bosons when the coupling tends to zero. This phenomenon is just the relativistic analog of the plasmon phenomenon to which Anderson[3] has drawn attention: that the scalar zero-mass excitations of a superconducting neutral Fermi gas become longitudinal plasmon modes of finite mass when the gas is charged.

The simplest theory which exhibits this behavior is a gauge-invariant version of a model used by Goldstone[2] himself: Two real[4] scalar fields φ_1, φ_2 and a real vector field A_μ interact through the Lagrangian density

$$L=-\tfrac{1}{2}(\nabla\varphi_1)^2-\tfrac{1}{2}(\nabla\varphi_2)^2-V(\varphi_1^2+\varphi_2^2)-\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}, \qquad (1)$$

where

$$\nabla_\mu\varphi_1=\partial_\mu\varphi_1-eA_\mu\varphi_2,$$

$$\nabla_\mu\varphi_2=\partial_\mu\varphi_2+eA_\mu\varphi_1,$$

$$F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu,$$

e is a dimensionless coupling constant, and the metric is taken as $-+++$. L is invariant under simultaneous gauge transformations of the first kind on $\varphi_1 \pm i\varphi_2$ and of the second kind on A_μ. Let us suppose that $V'(\varphi_0^2)=0$, $V''(\varphi_0^2)>0$; then spontaneous breakdown of U(1) symmetry occurs. Consider the equations [derived from (1) by treating $\Delta\varphi_1$, $\Delta\varphi_2$, and A_μ as small quantities] governing the propagation of small oscillations about the "vacuum" solution $\varphi_1(x)=0$, $\varphi_2(x)=\varphi_0$:

$$\partial^\mu\{\partial_\mu(\Delta\varphi_1)-e\varphi_0A_\mu\}=0, \qquad (2a)$$

$$\{\partial^2-4\varphi_0^2V''(\varphi_0^2)\}(\Delta\varphi_2)=0, \qquad (2b)$$

$$\partial_\nu F^{\mu\nu}=e\varphi_0\{\partial^\mu(\Delta\varphi_1)-e\varphi_0A_\mu\}. \qquad (2c)$$

Equation (2b) describes waves whose quanta have (bare) mass $2\varphi_0\{V''(\varphi_0^2)\}^{1/2}$; Eqs. (2a) and (2c) may be transformed, by the introduction of new variables

$$B_\mu=A_\mu-(e\varphi_0)^{-1}\partial_\mu(\Delta\varphi_1),$$

$$G_{\mu\nu}=\partial_\mu B_\nu-\partial_\nu B_\mu=F_{\mu\nu}, \qquad (3)$$

into the form

$$\partial_\mu B^\mu=0, \quad \partial_\nu G^{\mu\nu}+e^2\varphi_0^2B^\mu=0. \qquad (4)$$

Equation (4) describes vector waves whose quanta have (bare) mass $e\varphi_0$. In the absence of the gauge field coupling ($e=0$) the situation is quite different: Equations (2a) and (2c) describe zero-mass scalar and vector bosons, respectively. In passing, we note that the right-hand side of (2c) is just the linear approximation to the conserved current: It is linear in the vector potential, gauge invariance being maintained by the presence of the gradient term.[5]

When one considers theoretical models in which spontaneous breakdown of symmetry under a semisimple group occurs, one encounters a variety of possible situations corresponding to the various distinct irreducible representations to which the scalar fields may belong; the gauge field always belongs to the adjoint representation.[6] The model of the most immediate interest is that in which the scalar fields form an octet under SU(3): Here one finds the possibility of two nonvanishing vacuum expectation values, which may be chosen to be the two $Y=0$, $I_3=0$ members of the octet.[7] There are two massive scalar bosons with just these quantum numbers; the remaining six components of the scalar octet combine with the corresponding components of the gauge-field octet to describe

508

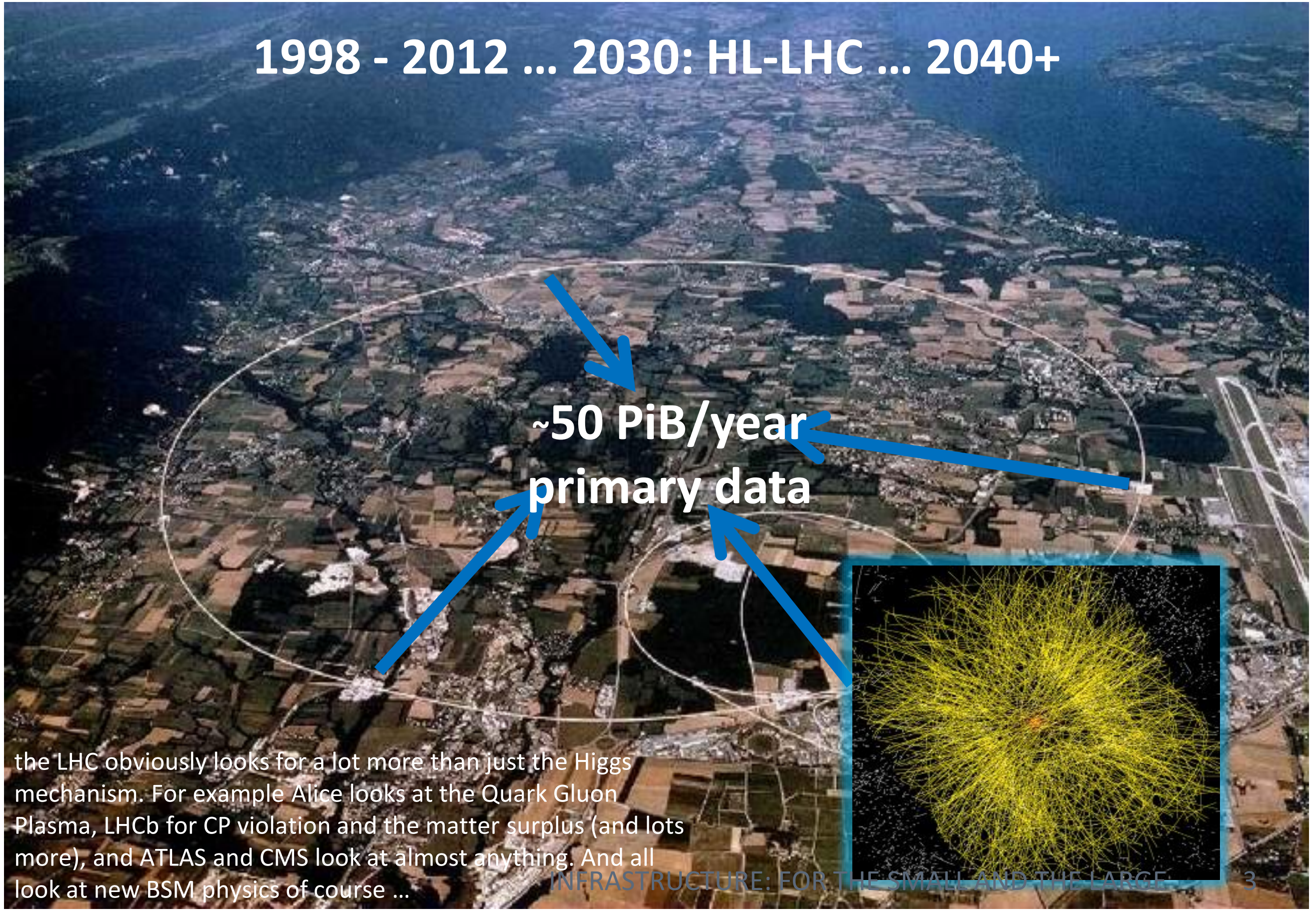
est order perturbation theory indicates that

we shall assume that the application of the

321

P. Higgs, Phys. Rev. Lett. 13, 508:

16823 characters, 165 kByte PDF

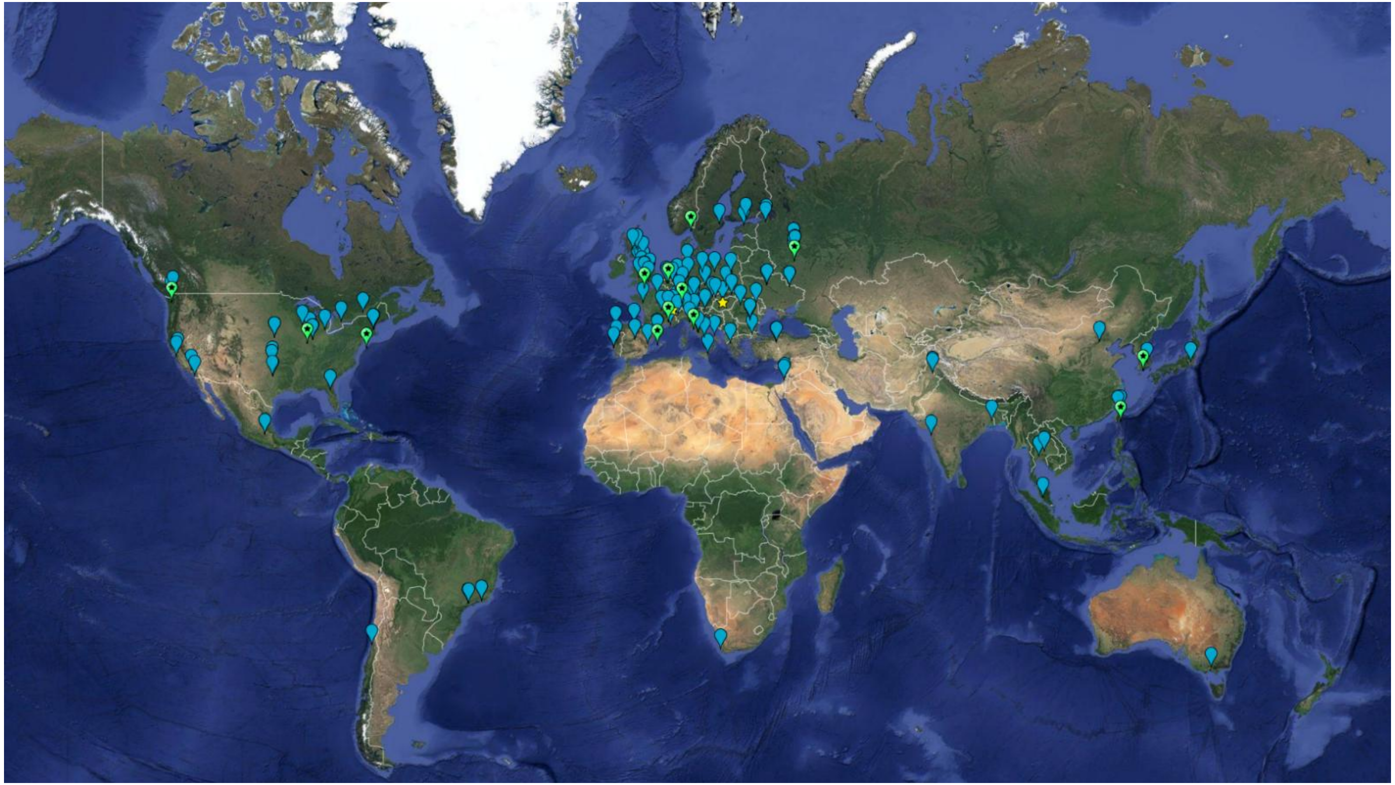


1998 - 2012 … 2030: HL-LHC … 2040+

~50 PiB/year primary data

the LHC obviously looks for a lot more than just the Higgs mechanism. For example Alice looks at the Quark Gluon Plasma, LHCb for CP violation and the matter surplus (and lots more), and ATLAS and CMS look at almost anything. And all look at new BSM physics of course …

Networked complexity: the worldwide LHC Computing



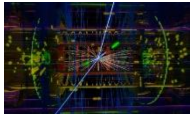
~ 1.6 million CPU cores
~ 2000 Petabyte
disk + archival

160+ institutes
40+ countries
13 ‘Tier-1 sites’
NL-T1:
SURF & Nikhef

largely based on generic e-Infrastructures
EGI
EuroHPC
NEIC
OpenScienceGrid
ACCESS-CI

Earth background: Google Earth; Data and compute animation: STFC RAL for WLCG and EGI.eu; Data: https://home.cern/science/computing/grid
For the LHC Computing Grid: wlcg.web.cern.ch, for EGI: www.egi.eu; NSF ACCESS-CI: https://access-ci.org/, for the NL-T1 and FuSE: fuse-infra.nl, https://www.surf.nl/en/research-it

Volume and computational complexity



ATLAS RAW single event
ROD File
1.60 MB

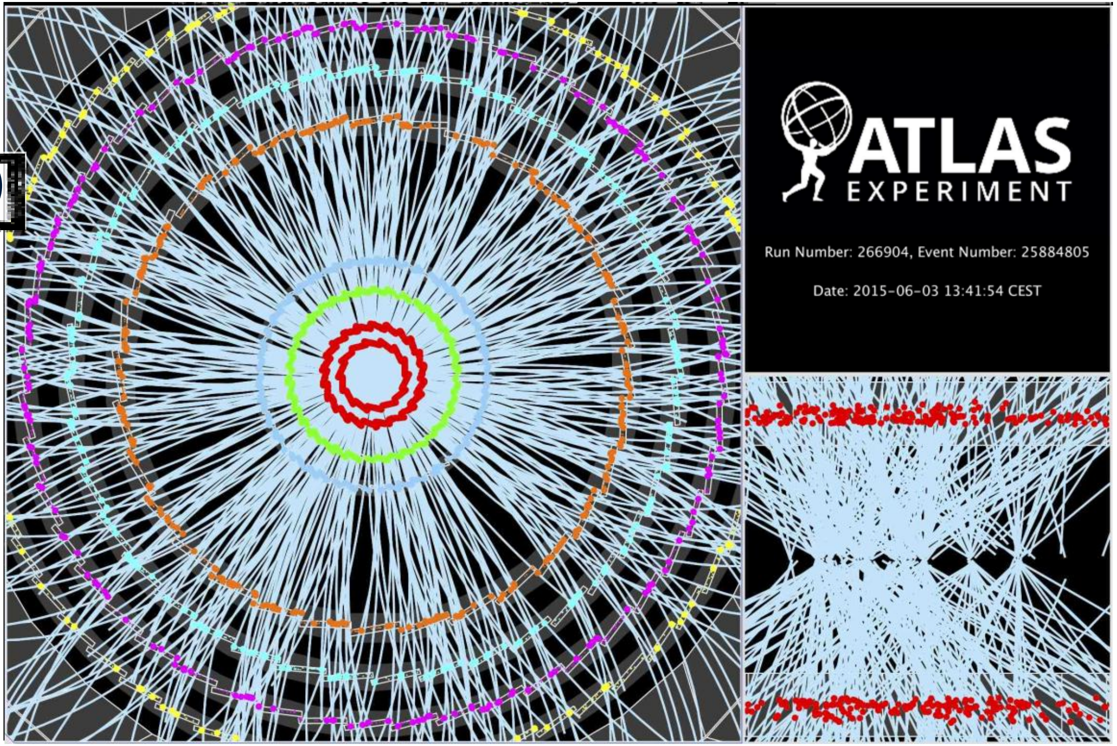
~60 TByte/s (compressed)

Trigger system selects 600 Hz ~ 1 GB/s data

~ 10 seconds compute for a single event at ATLAS with ‘jets’ containing ~30 collisions

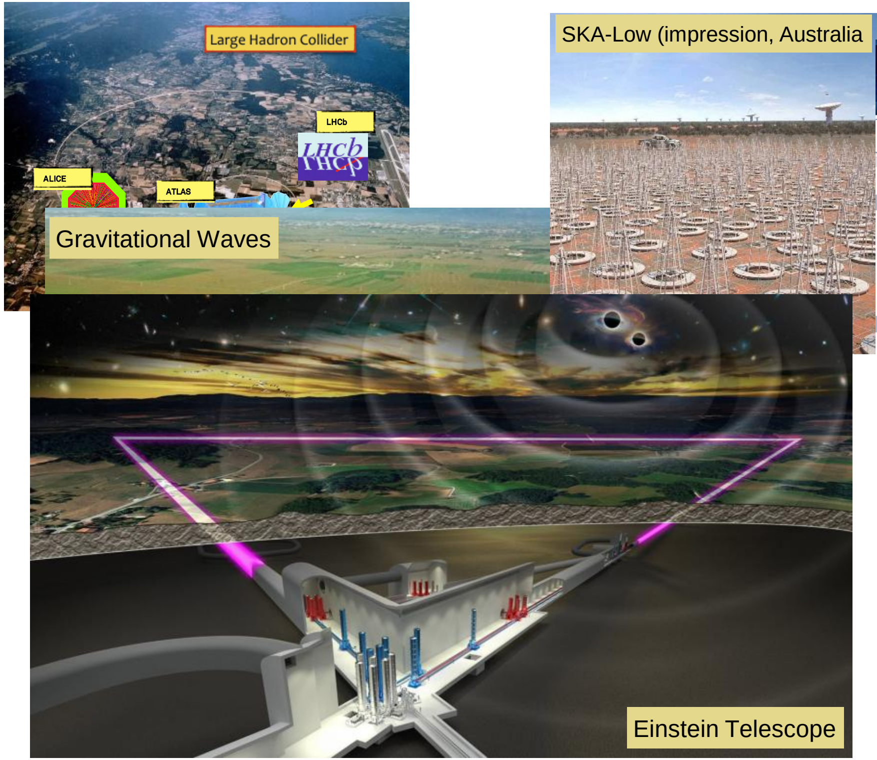
~10k researchers

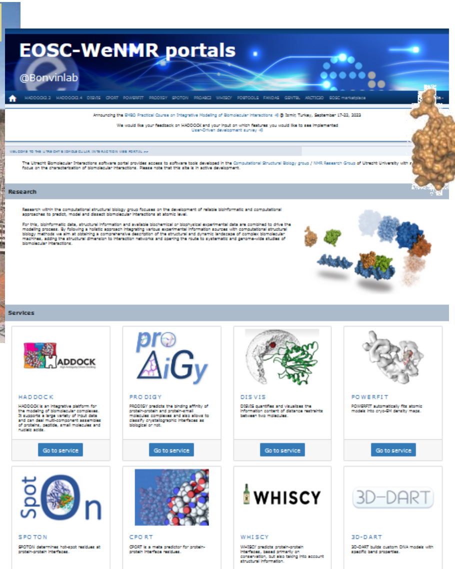
CERN and ~170 institutes


ATLAS
EXPERIMENT
Run Number: 266904, Event Number: 25884805
Date: 2015-06-03 13:41:54 CEST

Display of a proton-proton collision event recorded by ATLAS on 3 June 2015, with the first LHC stable beams at a collision energy of 13 TeV;
Event processing time: v19.0.1.1 as per Jovan Mitrevski and 2015 J. Phys.: Conf. Ser. 664 072034 (CHEP2015)

Scaling computing infrastructure – a common need




EOSC-WeNMR portals

@Bonvinlab

HADDOCK2.2 HADDOCK2.4 DISVIS CPORT POWERFIT PRODIGY SPOTON PROABC2 WHISCY PDBTOOLS FANDAS GENTBL ARCTIC3D EOSC marketplace

Announcing the EMBO Practical Course on Integrative Modelling of Biomolecular Interactions @ Izmir, Turkey, September 17-22, 2023

We would like your feedback on HADDOCK and your input on which features you would like to see implemented
User-Driven development survey

WELCOME TO THE UTRECHT BIOMOLECULAR INTERACTION WEB PORTAL »»

The Utrecht Biomolecular Interactions software portal provides access to software tools developed in the Computational Structural Biology group / NMR Research Group of Utrecht University with a focus on the characterization of biomolecular interactions. Please note that this site is in active development.

Research

Research within the computational structural biology group focuses on the development of reliable bioinformatic and computational approaches to predict, model and dissect biomolecular interactions at atomic level.

For this, bioinformatic data, structural information and available biochemical or biophysical experimental data are combined to drive the modelling process. By following a holistic approach integrating various experimental information sources with computational structural biology methods we aim at obtaining a comprehensive description of the structural and dynamic landscape of complex biomolecular machines, adding the structural dimension to interaction networks and opening the route to systematic and genome-wide studies of biomolecular interactions.

Services

HADDOCK
HADDOCK is an integrative platform for the modeling of biomolecular complexes. It supports a large variety of input data and can deal multi-component assemblies of proteins, peptide, small molecules and nucleic acids.
Go to service

PRODIGY
PRODIGY predicts the binding affinity of protein-protein and protein-small molecules complexes and also allows to classify crystallographic interfaces as biological or not.
Go to service

DISVIS
DISVIS quantifies and visualizes the information content of distance restraints between two molecules.
Go to service

POWERFIT
POWERFIT automatically fits atomic models into cryo-EM density maps.
Go to service

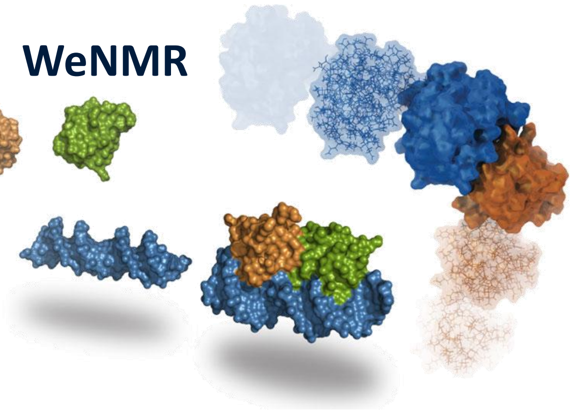
SPOTON
SPOTON determines hot-spot residues at protein-protein interfaces.
Go to service

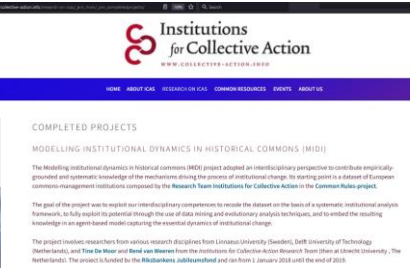
CPORT
CPORT is a meta predictor for protein-protein interface residues.
Go to service

WHISCY
WHISCY predicts protein-protein interfaces, based primarily on conservation, but also taking into account structural information.

3D-DART
3D-DART builds custom DNA models with specific bend properties.

proABC 2




Institutions for Collective Action
WWW.COLLECTIVE-ACTION.INFO

HOME ABOUT ICAS RESEARCH ON ICAS COMMON RESOURCES EVENTS ABOUT US

COMPLETED PROJECTS

MODELLING INSTITUTIONAL DYNAMICS IN HISTORICAL COMMONS (MIDI)

The Modelling institutional dynamics in historical commons (MIDI) project adopted an interdisciplinary perspective to contribute empirically-grounded and systematic knowledge of the mechanisms driving the process of institutional change. Its starting point is a dataset of European commons-management institutions composed by the Research Team Institutions for Collective Action in the Common Rules-project.

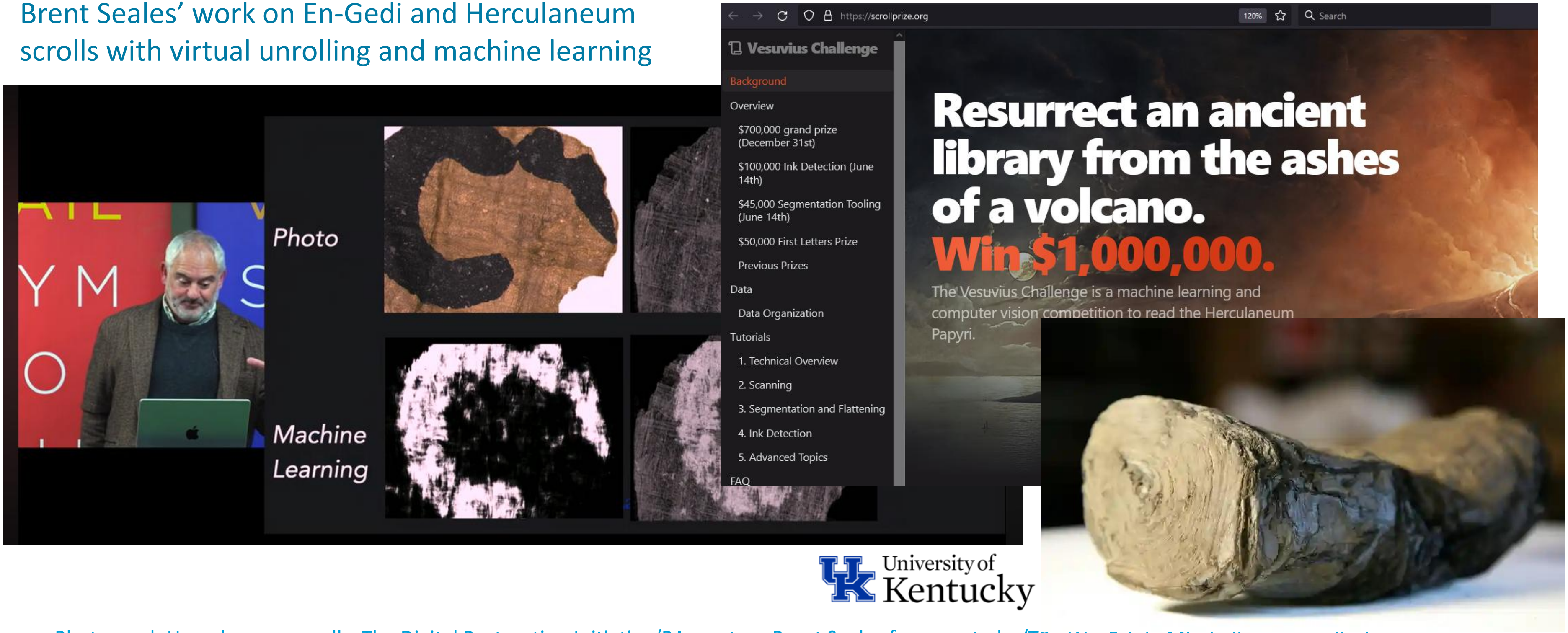
The goal of the project was to exploit our interdisciplinary competences to recode the dataset on the basis of a systematic institutional analysis framework, to fully exploit its potential through the use of data mining and evolutionary analysis techniques, and to embed the resulting knowledge in an agent-based model capturing the essential dynamics of institutional change.

The project involves researchers from various research disciplines from Linnaeus University (Sweden), Delft University of Technology (Netherlands), and Tine De Moor and René van Weeren from the *Institutions for Collective Action Research Team* (then at Utrecht University , The Netherlands). The project is funded by the Riksbankens Jubileumsfond and ran from 1 January 2018 until the end of 2019.

Sources: Einstein Telescope: https://et-emr.eu/; CERN https://wlcg.web.cern.ch/; HADDOCK, WeNMR, @Bonvinlab https://wenmr.science.uu.nl/; Virgo, Pisa, IT; SKAO: the SKA-Low observatory, Australia https://www.skatelescope.org/ - OpenMOLE simulation on EGI - https://cdn.egi.eu/app/uploads/2022/04/EGI_Use_Cases.pdf; agent-based modelling of ICAs: https://collective-action.info/research-on-icas/ Molood Dehkordi (TUDelft), Tine de Moor (EUR RSM)

Collaborative computing changing fields you may not expect

Brent Seales' work on En-Gedi and Herculaneum scrolls with virtual unrolling and machine learning



Resurrect an ancient library from the ashes of a volcano. Win $1,000,000.

The Vesuvius Challenge is a machine learning and computer vision competition to read the Herculaneum Papyri.

University of Kentucky

Photograph Herculaneum scrolls: The Digital Restoration Initiative/PA; capture Brent Seales from youtu.be/T0mWqsFrJpk; ML challenge: scrollprize.org

This is a tour of {a,one} large-scale IT landscape

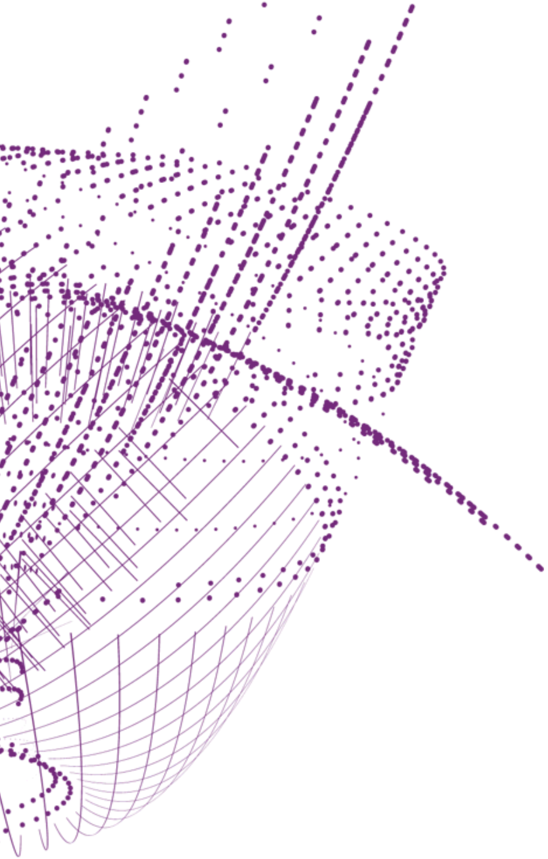
‘Exploring the e-infrastructure with use cases from data intensive research’

- **building** a compute, storage, data, and network facility
  for high-throughput computing at the LHC scale

- a global **collaborative** infrastructure
  with trust and identity, in a secure way

- **sustaining the ecosystem** we have built
  the Research Infrastructure Commons, the GORC, and how research principles can guide digitalisation

… to make ICT a research instrument rather than ‘just a tool’



Building the facilities

Single CPU scaling stopped around 2004

- limitation is power, not circuit size
  - and clock frequency is most ‘power-hungry’
  - still some packages now @ TDP of 400W

- multiple cores on the same die helps:
  - AMD EPYC Genoa (Zen 4) has 96 cores/die
  - Intel Granite Rapids, Nvidia GraceHopper, …
  - but e.g. Intel Cascade Lake AP was less useful

- CPU design-level performance gains left
  - predictive and out-of-order execution
  - on-die parallelism (multi-core)
  - pre-fetching and multi-tier caching
  - execution unit sharing (‘SMT’)

  but at increased risk for security/integrity

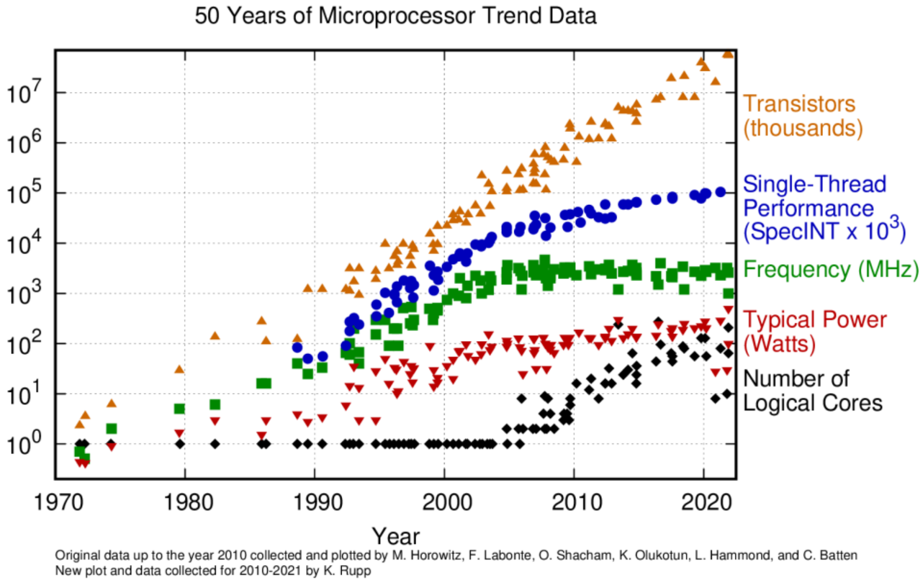
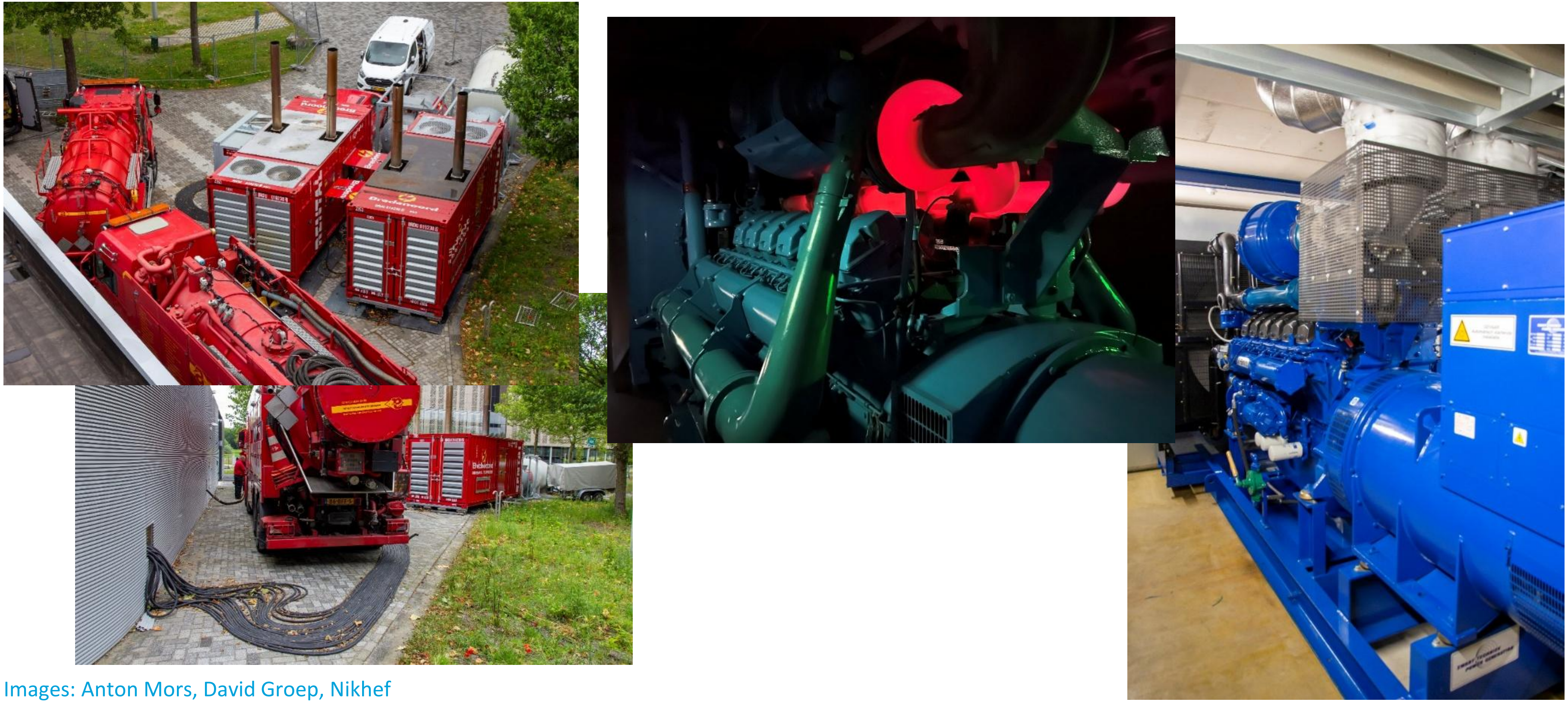


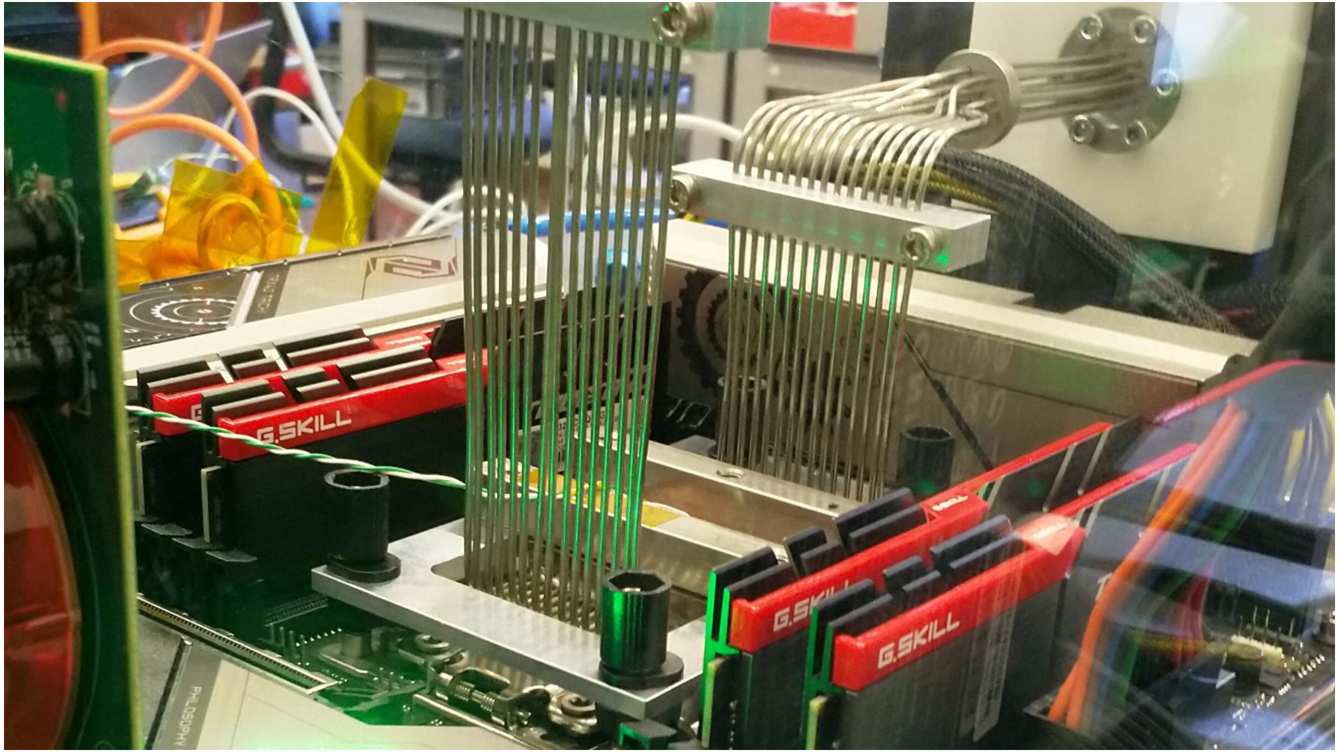
Image: K Rupp, https://github.com/karlrupp/microprocessor-trend-data

‘I got the power …’



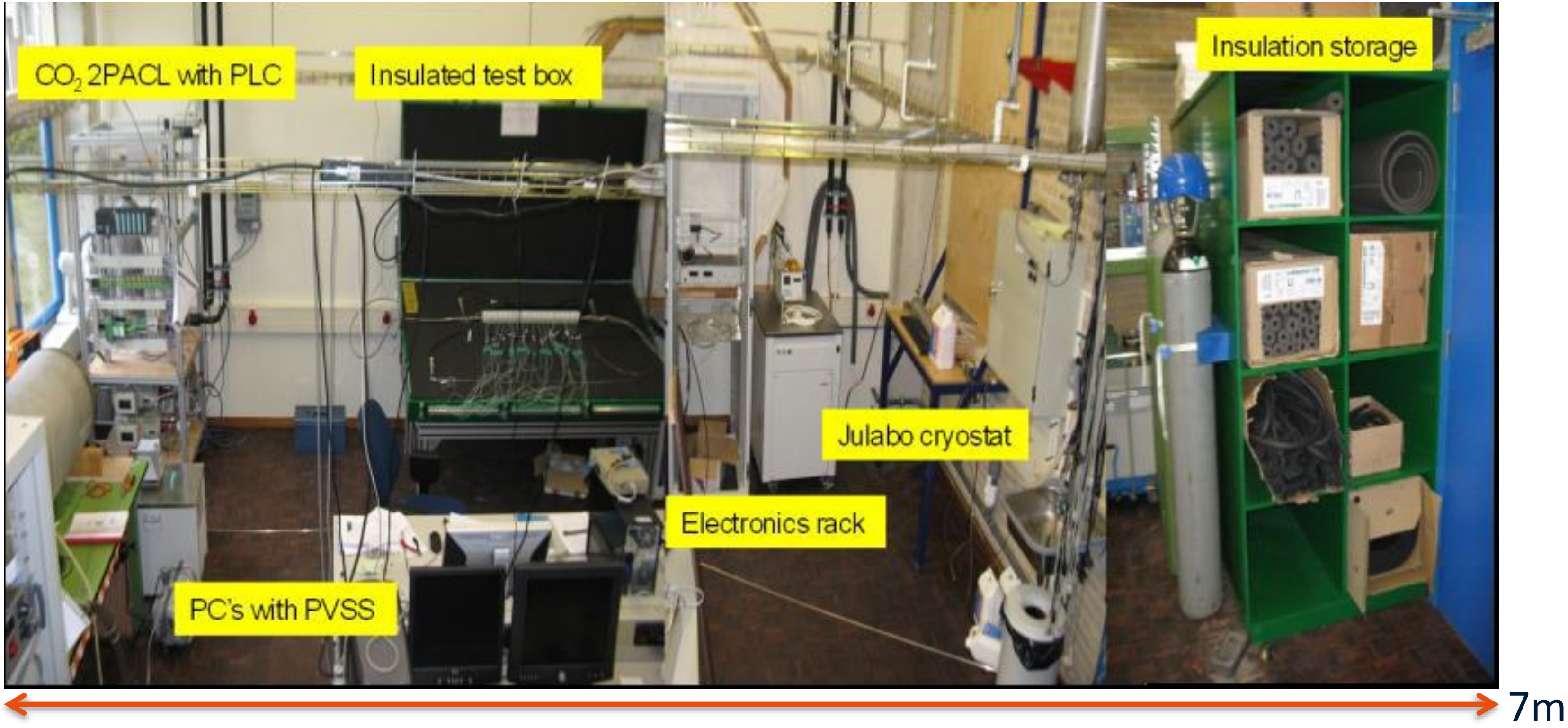
Images: Anton Mors, David Groep, Nikhef

Fix the thing that didn't scale well, CPU frequency??



LCO2 cooling of an AMD Ryzen Threadripper 3970X [56.38 °C] at 4600.1MHz processor (~1.25x nominal speed) sustained over all cores simultaneously, using the Nikhef LCO2 test bench system (https://hwbot.org/submission/4539341) - (Krista de Roo en Tristan Suerink)

… since you then need this around it …



Nikhef 2PA LCO2 cooling setup. Image from Bart Verlaat, Auke-Pieter Colijn *CO2 Cooling Developments for HEP Detectors* https://doi.org/10.22323/1.095.0031

So we scaled up *inside* one system

Multiple cores and SMT on a single die

- ‘trivial’ step-up is to do multiple sockets in one system
  2-socket, sometimes 4 socket on a motherboard
- appears as a single shared memory system, but requires
  cache coherency between CPU cores and sockets …
  which is useful for tightly coupled parallel applications
  but not needed for ‘trivially parallel’ high throughput needs
- depending on architecture
  cache coherency may limit single-thread performance
  (although AMD did better here than Intel *lakes)

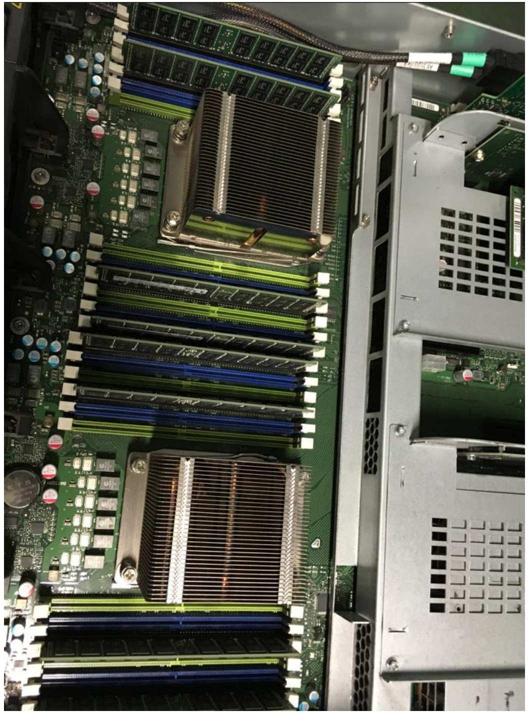


Image: dual-socket Fujitsu system at the Xenon experiment site, 2019. source: Tristan Suerink, Nikhef

The advantages of LHC-like data

- ‘tightly coupled’ HPC and (computational) cluster computing:
  - modelling for weather/climate, fluid dynamics, but also e.g. QC-simulation

- HTC and data-intensive processing for horizontal scaling:
  - lots of data, as in High Energy Physics (HEP), *omics and protein docking, …
  - conveniently parallel,
    but (intensive) local I/O requirements on memory and scratch storage

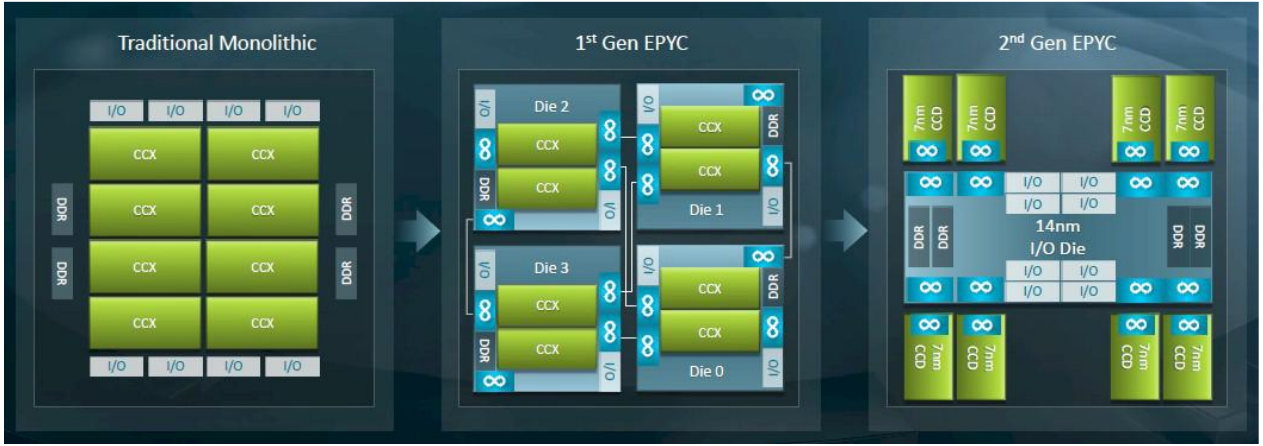
- portals and many web applications: ‘horizontal’ scaling,
  but for RI use cases often backed by HPC or HTC resources …
  - science gateways and portals (like https://wenmr.science.uu.nl/)
  - interactive notebooks and analysis environments

HPC: High Performance Computing; HTC: High Throughput Computing

CPU design changes may fit application, or not

Example: AMD EPYC effective for LHC-like workloads

- Naples → Rome added shared memory die
- links all cores directly to memory



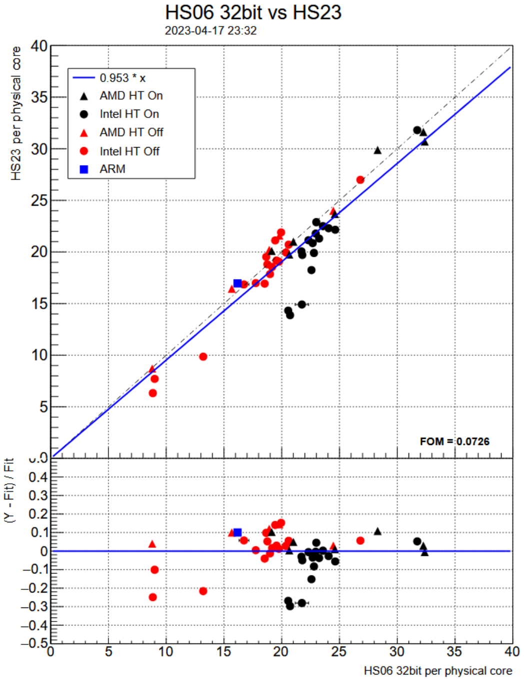


Image source: AMD, retrieved from https://m.hexus.net/tech/news/cpu/135479-amd-shares-details-zen-3-zen-4-architectures/

HEPscore: https://doi.org/10.48550/arXiv.2306.08118 ; AMD-EPYC architecture benefits memory-intensive HEPscore23 over HS06 (‘memory subset of SPECint06’)

CPU design changes may fit application, or not

But the
Rome-Milan improvement …?

- shared L3 cache
  benefits tightly coupled HPC,
  but not HTC, limited by
  ‘off-die memory’

Which is also why single-socket systems outclass dual-socket (also on TCO)

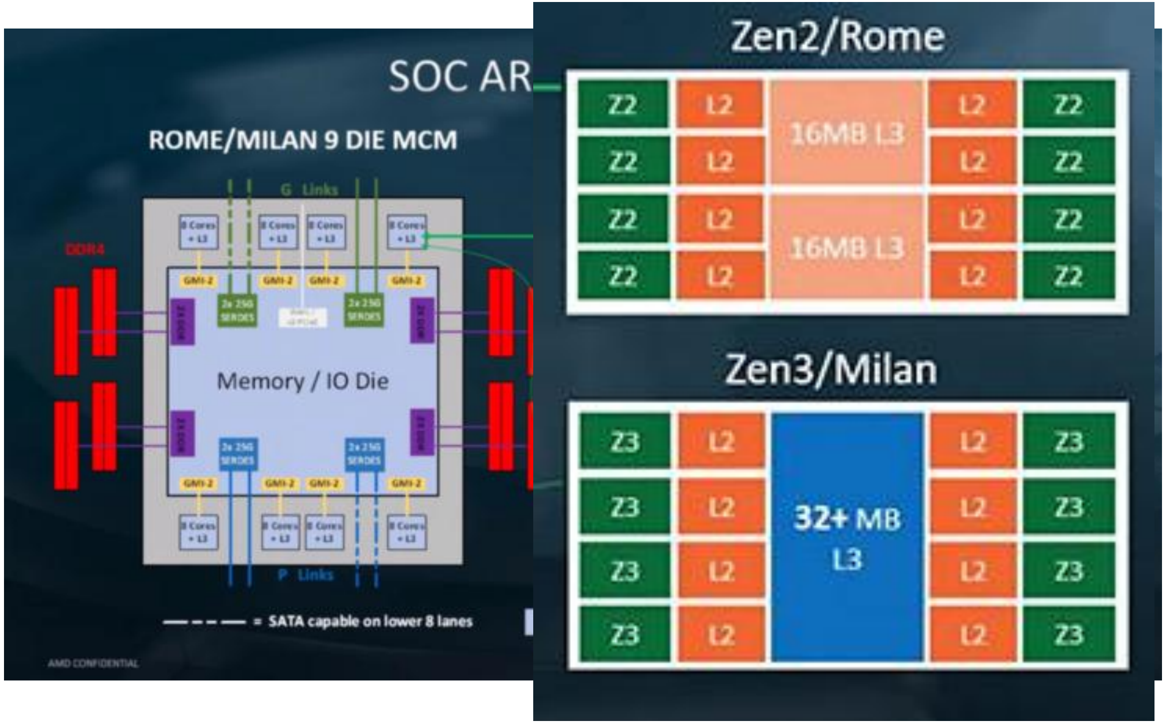
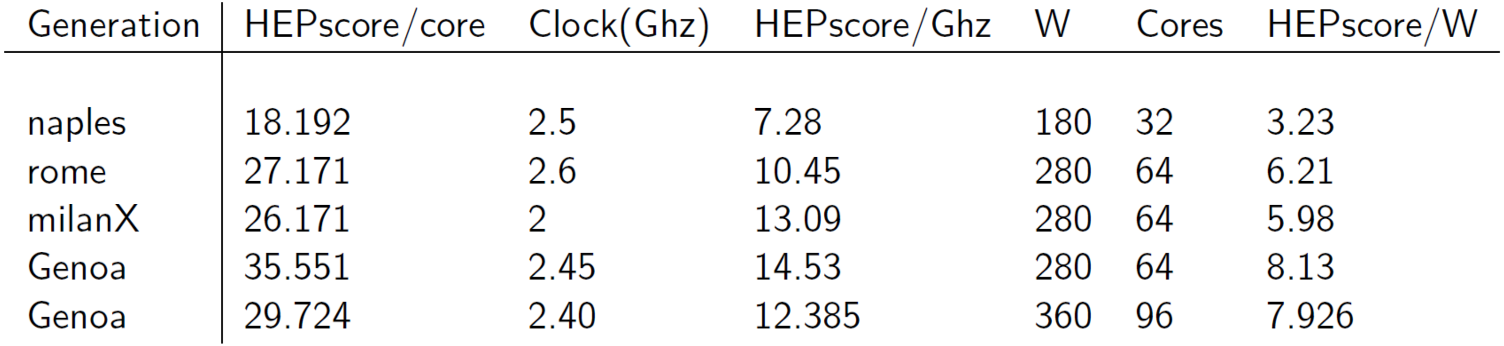


Image source: AMD, retrieved from https://m.hexus.net/tech/news/cpu/135479-amd-shares-details-zen-3-zen-4-architectures/

… and indeed we see it in the HEPscore benchmarks

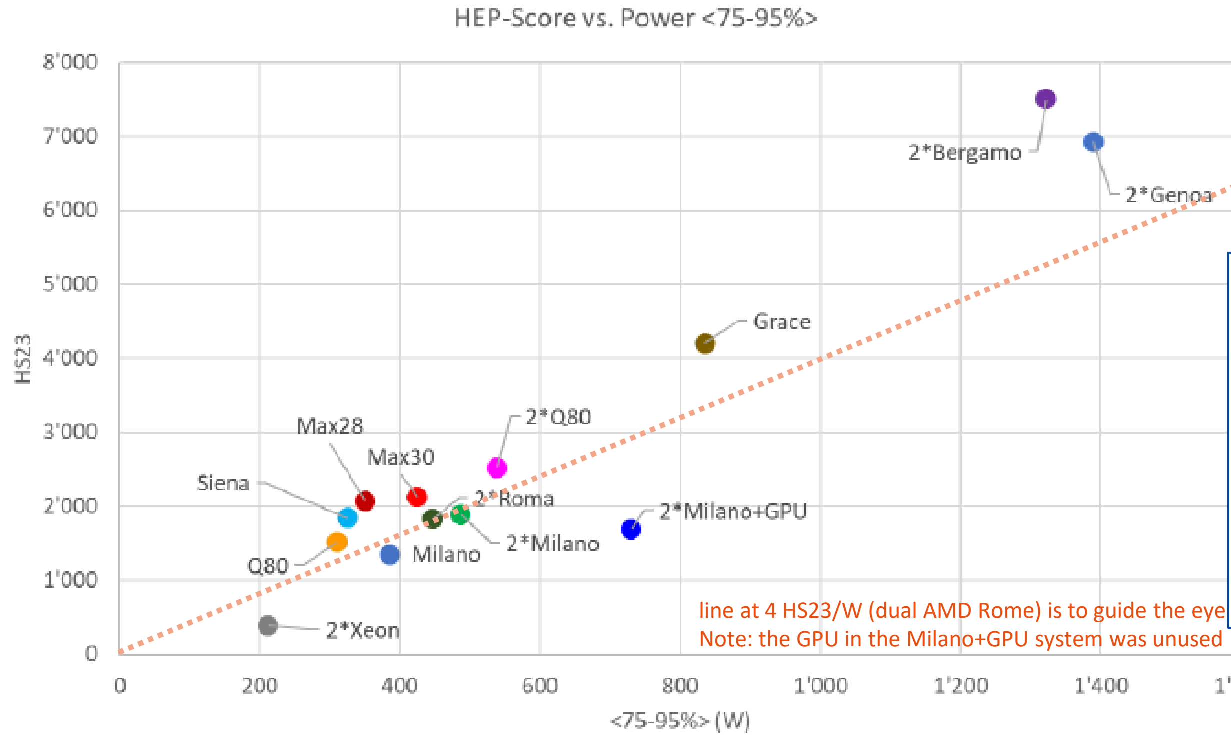


Generation	HEPscore/core	Clock(Ghz)	HEPscore/Ghz	W	Cores	HEPscore/W
naples	18.192	2.5	7.28	180	32	3.23
rome	27.171	2.6	10.45	280	64	6.21
milanX	26.171	2	13.09	280	64	5.98
Genoa	35.551	2.45	14.53	280	64	8.13
Genoa	29.724	2.40	12.385	360	96	7.926

The HEPscore/W is the most relevant number for an ‘always full’ system
and for TCO due to energy price, at least until memory prices exploded in October ’25

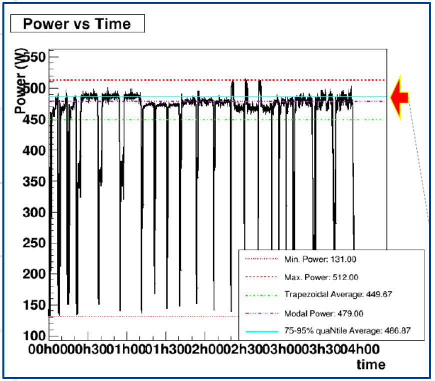
Performance data from Emily Kooistra *A Tour of Italian CPUs and the fastest road South*, Nikhef Jamboree 2023.
HEPscore: Giordano *et al.* EPJ Web of Conferences 295, 07024 (2024) https://doi.org/10.1051/epjconf/202429507024

The energy bottleneck: architecture ‘figure of merit’



line at 4 HS23/W (dual AMD Rome) is to guide the eye

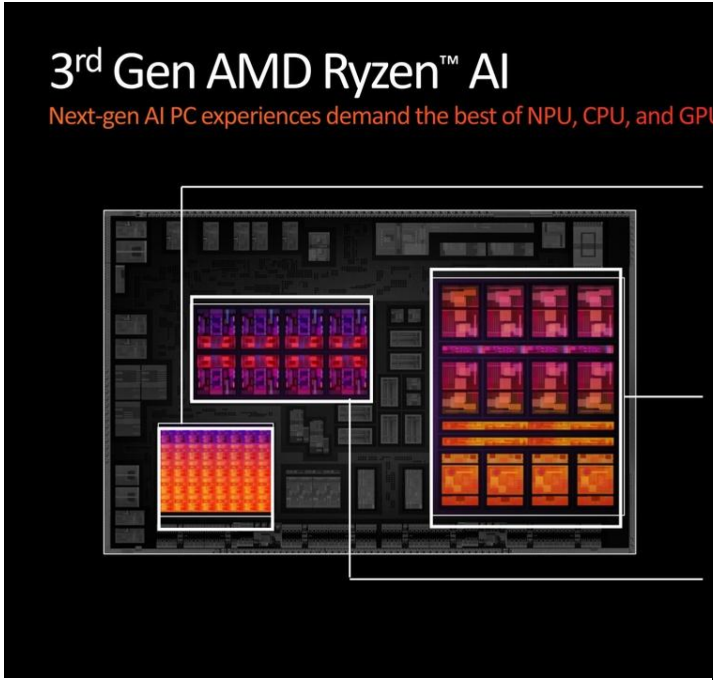
Note: the GPU in the Milano+GPU system was unused



Data and graphs: Emanuele Simili, Glasgow University, at CHEP2024 (https://indico.cern.ch/event/1338689/contributions/6011562/)

HEPSPEC23 benchmark: https://gitlab.cern.ch/hep-benchmarks/hep-benchmark-suite (‘memory-intensive’ high throughput processing application benchmark)

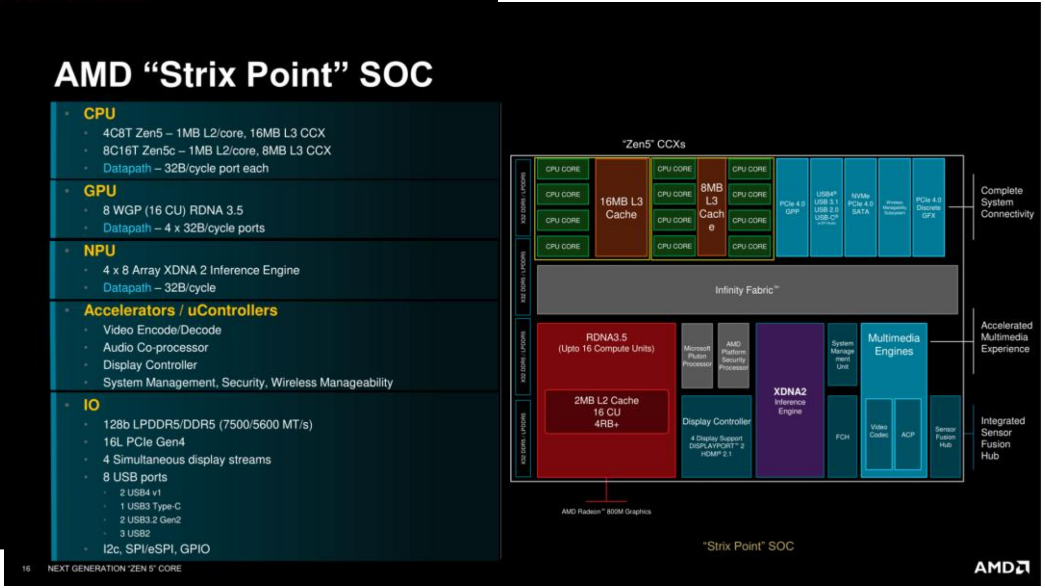
Hybrid SOCs and heterogeneous architectures



3rd Gen AMD Ryzen™ AI

Next-gen AI PC experiences demand the best of NPU, CPU, and GPU architectures

NPUs, GPUs, APUs …



AMD "Strix Point" SOC

- **CPU**
  - 4C8T Zen5 – 1MB L2/core, 16MB L3 CCX
  - 8C16T Zen5c – 1MB L2/core, 8MB L3 CCX
  - Datapath – 32B/cycle port each
- **GPU**
  - 8 WGP (16 CU) RDNA 3.5
  - Datapath – 4 x 32B/cycle ports
- **NPU**
  - 4 x 8 Array XDNA 2 Inference Engine
  - Datapath – 32B/cycle
- **Accelerators / uControllers**
  - Video Encode/Decode
  - Audio Co-processor
  - Display Controller
  - System Management, Security, Wireless Manageability
- **IO**
  - 128b LPDDR5/DDR5 (7500/5600 MT/s)
  - 16L PCIe Gen4
  - 4 Simultaneous display streams
  - 8 USB ports
    - 2 USB4 v1
    - 1 USB3 Type-C
    - 2 USB3.2 Gen2
    - 3 USB2
  - I2c, SPI/eSPI, GPIO

"Zen5" CCXs
CPU CORE
16MB L3 Cache
8MB L3 Cache
PCIe 4.0 GPP
USB4® USB 3.1 USB 2.0 USB-C®
NVMe PCIe 4.0 SATA
Wireless Manageability Subsystem
PCIe 4.0 Discrete GFX
Complete System Connectivity
X32 DDR5 / LPDDR5
Infinity Fabric™
RDNA3.5 (Upto 16 Compute Units)
2MB L2 Cache 16 CU 4RB+
Microsoft Pluton Processor
AMD Platform Security Processor
Display Controller
4 Display Support DISPLAYPORT™ 2 HDMI® 2.1
XDNA2 Inference Engine
System Management Unit
FCH
Multimedia Engines
Video Codec
ACP
Sensor Fusion Hub
Accelerated Multimedia Experience
Integrated Sensor Fusion Hub
AMD Radeon™ 800M Graphics
"Strix Point" SOC

16 NEXT GENERATION "ZEN 5" CORE

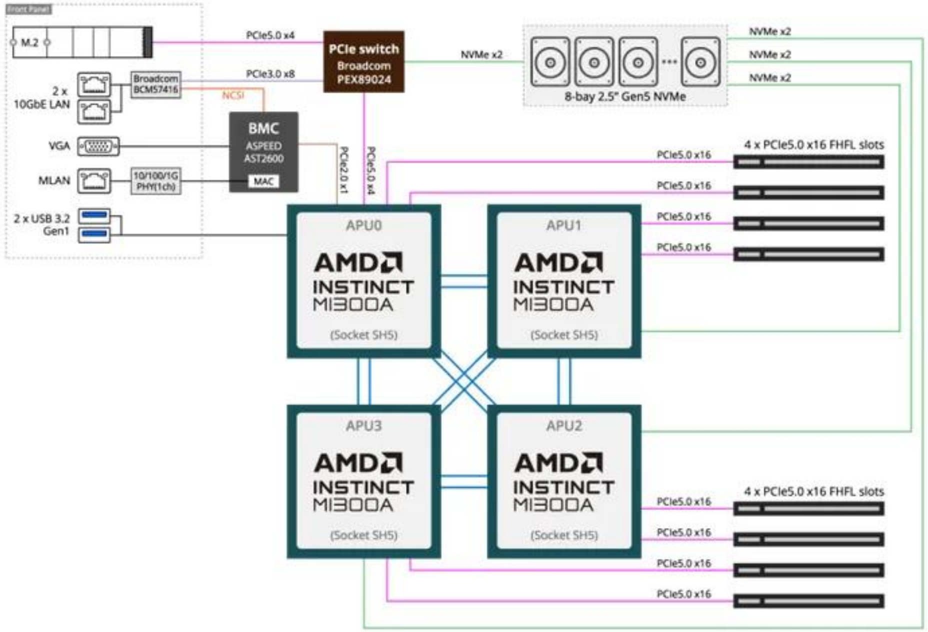
(note these are laptop/desktop SKUs, not servers)

Images: AMD Ryzen 9 HX 370 AI, Strix SOC – compare also Intel Lunar Lake architecture

ML ‘big physics models’ are changing that, but at a cost …

Current models tend to be very large, training barely fitting in an H100, and inference also needs 48-96 GB

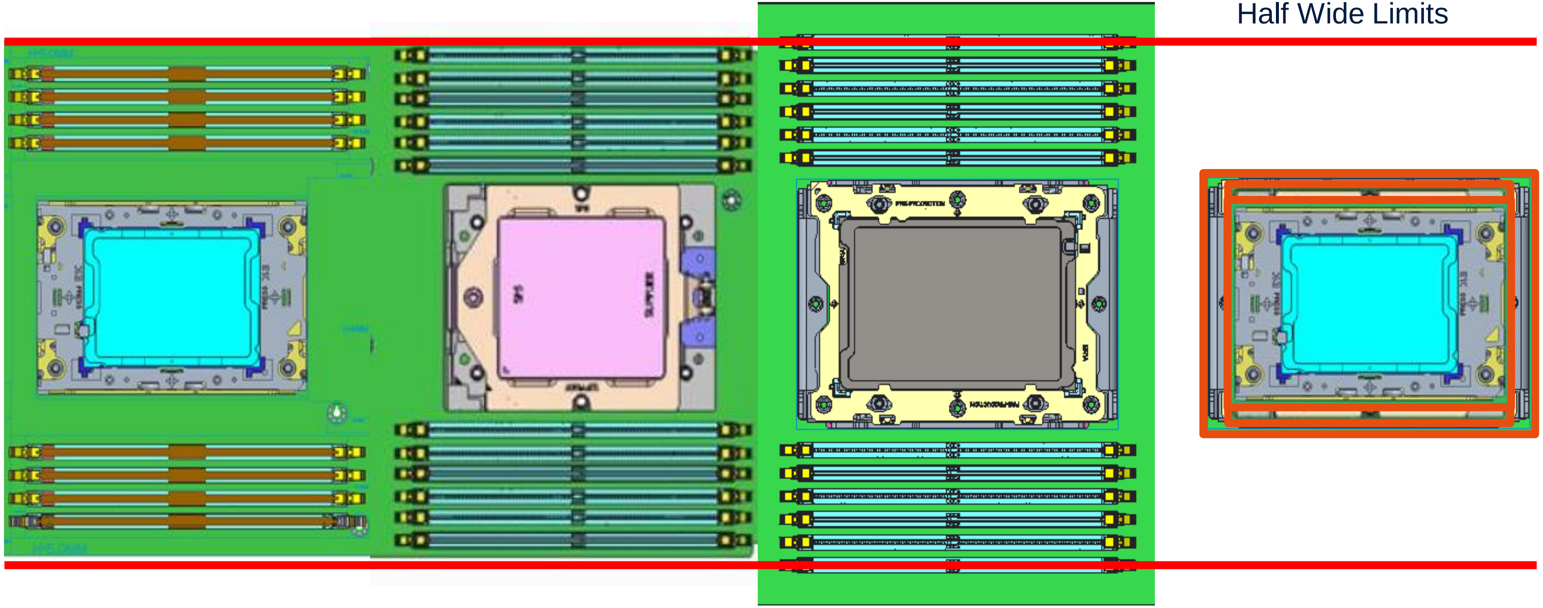
- conventional GPUs for training are outgrowing budgets very fast
- validate hybrid/APU architectures?



Challenge of course is software porting if doing more than just pyTorch – but that is more of a worry when writing kernels for real-time applications like in the HLT

Systems block diagram: GigaBYTE G383-R80-AAP1, Nikhef SIF “Bordercollie” ML training system

but there is also a serious issue with sockets …



EagleStream

Genoa

BirchStream-AP (LGA7529)

Image thanks go to Rick Koopman – Lenovo at the HTCondor Workshop 2024 https://indico.cern.ch/event/1386170/

So if large-scale IT does not quite fit … ahum …

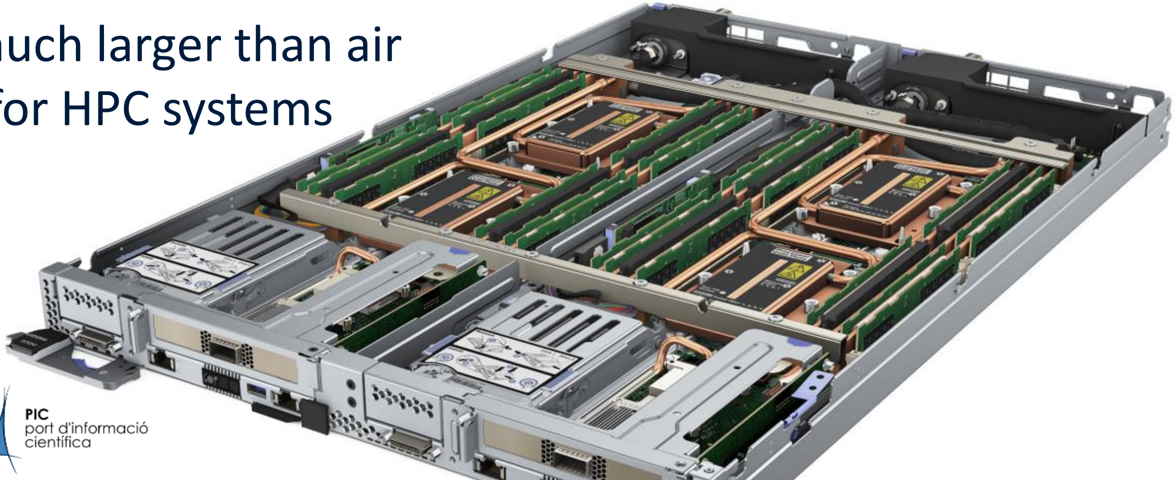
Image source: https://lambdalabs.com/products/blade

SuperMicro (branded as ‘Lambda Blade’)
4U chassis, supporting 10 consumer-grade GPUs …
… with a bump

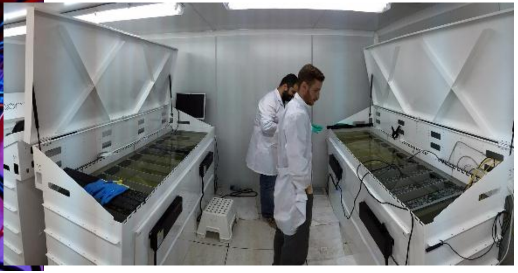
And it’s hot in there …

- Heat capacity of liquid is much larger than air
- by now (almost) standard for HPC systems

- immersive systems look cool, but are ‘a bit hard’ on maintenance



PIC port d'informació científica



Strongly depends on systems engineering: when water inlet temperature can be >40 degC, you have almost always free cooling

Image source dual-board system: Lenovo, ThinkSystem SD650

immersive cooling image https://hypertec.com/blog/sustainable-emerging-tech-liquid-immersion-cooling/, PIC T1 centre, Barcelona, ES

Scaling up – beyond one lone system

Typical compute farm @Nikhef for 'milking' computer clusters

Continuous design challenge

- **balanced features** for node throughput CPU, storage, memory bandwidth & latency, NIC & network speed

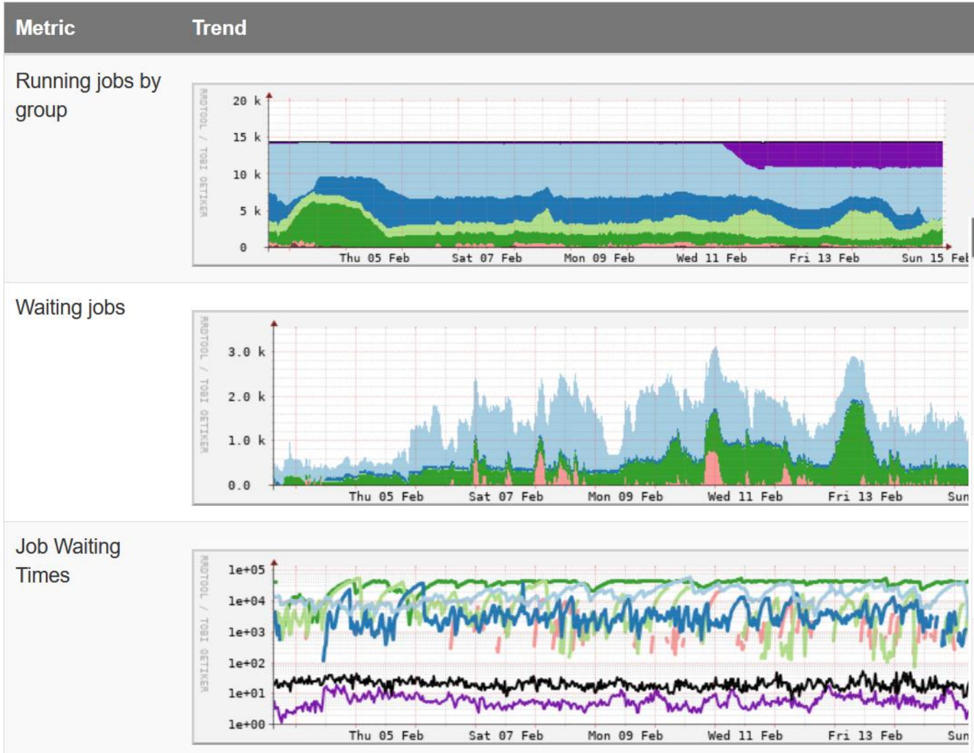
For example for WLCG:

- **single-socket** multicore systems are fine, today typically 64-128 cores per system
- **network**: 2x25/2x100Gbps (matching #cores)
- **memory**: say ~ 8 GiB/core
- **local disk**: 8-16 TB NVME (~100GB/core)

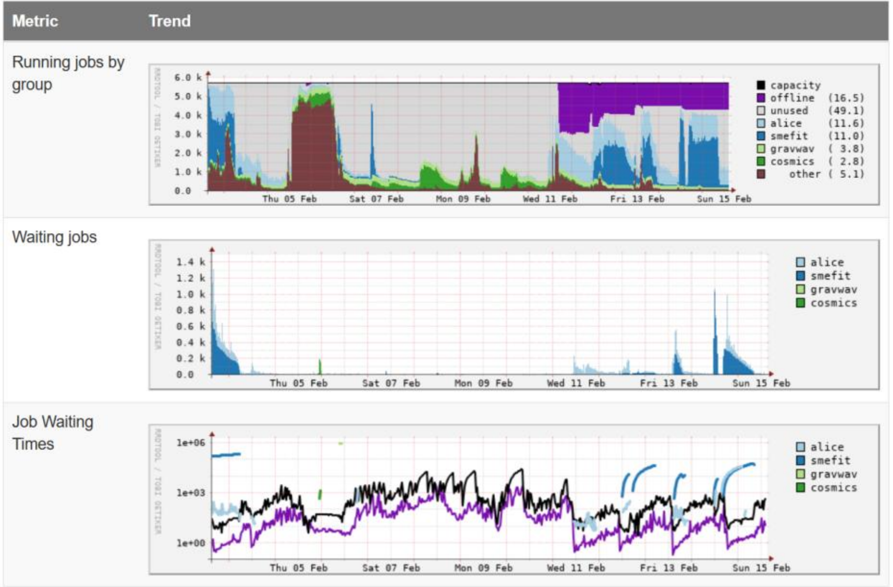
+ space (physical + power) to add **GPUs**

Image: Cluster 'Lotenfeest' at the Nikhef NDPF, acquired March 2020. Lenovo SR655 with AMD EPYC 7702P 64-Core single-socket. Some with 4 L40s Nvidia GPUs

To fill or not to fill, that’s the question …



Metric	Trend
Running jobs by group	
Waiting jobs	
Job Waiting Times	



Metric	Trend
Running jobs by group	
Waiting jobs	
Job Waiting Times	

https://www.nikhef.nl/pdp/doc/stats/ndpf-prd-grisview-week - retrieved 15 February 2026

Occupancy: balance efficient use of resources and happy users

For organized 'production' computing (planned months in advance in WLCG)

- ***predictable* scheduling** is more important (steady flow of results)
- **maximizing efficiency**: resource cost is the limiting factor in (physics) results
- co-scheduling with data (pre-placement) is required
- community-authorization based access to data sources only

For 'local' users, e.g. students whose progress tomorrow depends on results *today*

- ***response time*** is more important than efficiency
- fast turn-around/short waiting times by heterogeneous ('competing') user base
- data access must be parallelism-ready, but is 'always' local on-site
- with local credentials and sharing with desktop and Jupyter environments

so offering two distinct classes of services is (in this case) intentional

Standard interfaces for compute and data?

hourglass model ‘kind-of’ worked for IP and web with http as common standard

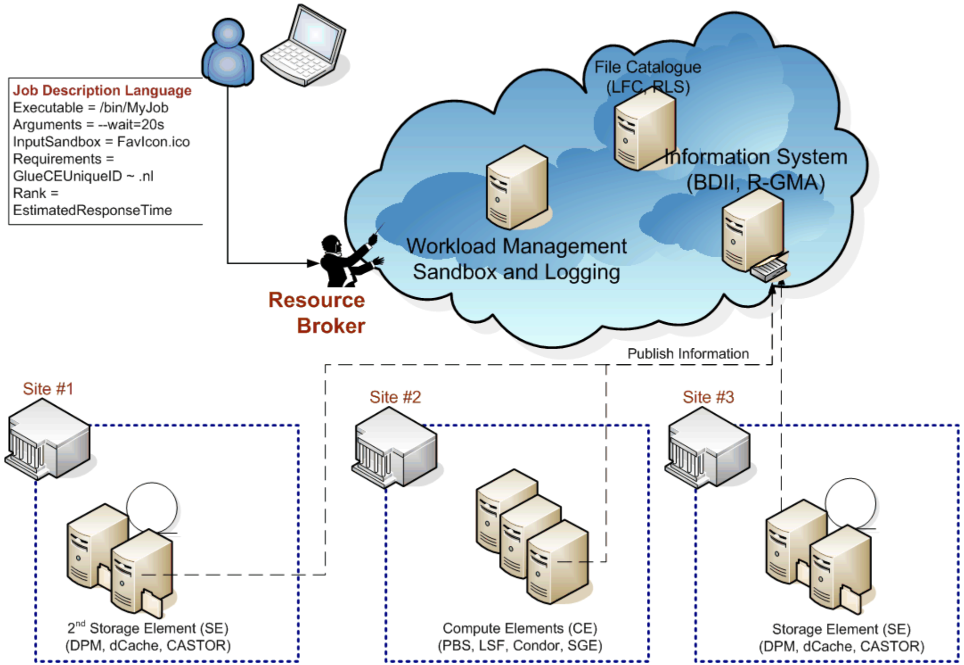
- a very simple stateless interface

protocols for higher-level services never quite reached this level of global interop

- requirements too complex and stateful
- use cases were usually scoped

slowly changing now but only for similarly simple things, like on-line object storage

Is distributed computing too bespoke …?



```
Job Description Language
Executable = /bin/MyJob
Arguments = --wait=20s
InputSandbox = FavIcon.ico
Requirements =
GlueCEUniqueID ~ .nl
Rank =
EstimatedResponseTime
```

Resource Broker
File Catalogue (LFC, RLS)
Information System (BDII, R-GMA)
Workload Management Sandbox and Logging
Publish Information
Site #1
Site #2
Site #3
2nd Storage Element (SE) (DPM, dCache, CASTOR)
Compute Elements (CE) (PBS, LSF, Condor, SGE)
Storage Element (SE) (DPM, dCache, CASTOR)

Interoperable cloud? Compare OGF’s OCCI WG GFD.221 (https://www.ogf.org/documents/GFD.221.pdf) with e.g. Amazon S3 API or the OwnCloud CS3 interfaces

DIRAC: spanning heterogeneous resource models

Add a scheduling layer!

‘any (IT) problem can be solved by adding an extra level of indirection’*

DIRAC is just one example

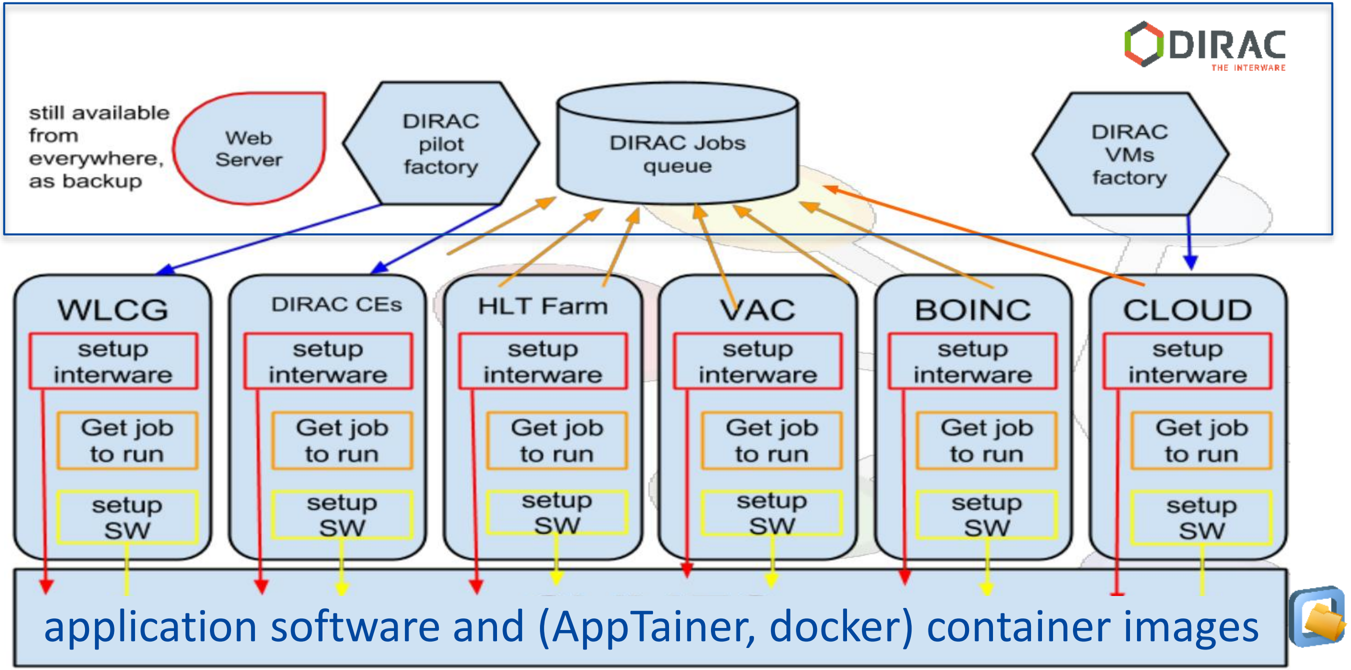
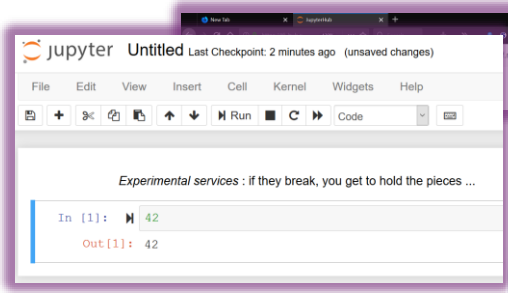


Image: DIRAC project, A. Tsaregorodtsev *et al.* CPPM Marseille, from https://dirac.readthedocs.io/ ; CVMFS (CERN VM File System) is a common software distribution platform using distributed signed data objects in a cached hierarchy using CDN techniques, see https://cernvm.cern.ch/fs/ * thanks to Miron Livny

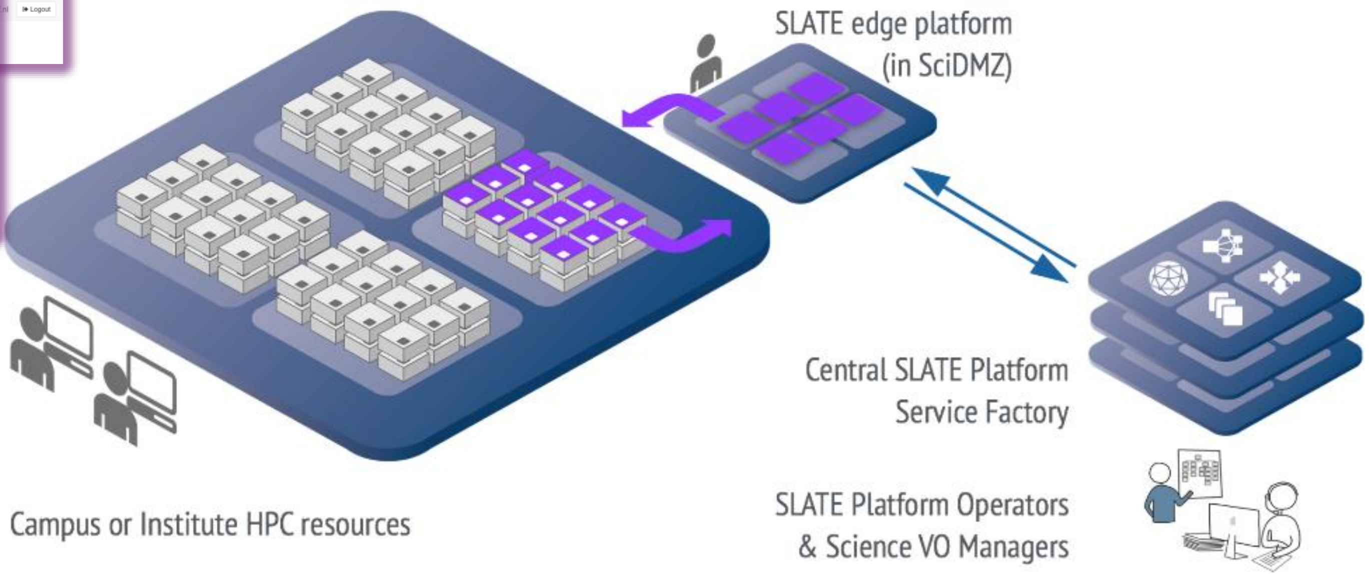
An overlay network of containers

Nobody wants a cloud per-se … what folk want is a solution …



Experimental services : if they break, you get to hold the pieces ...

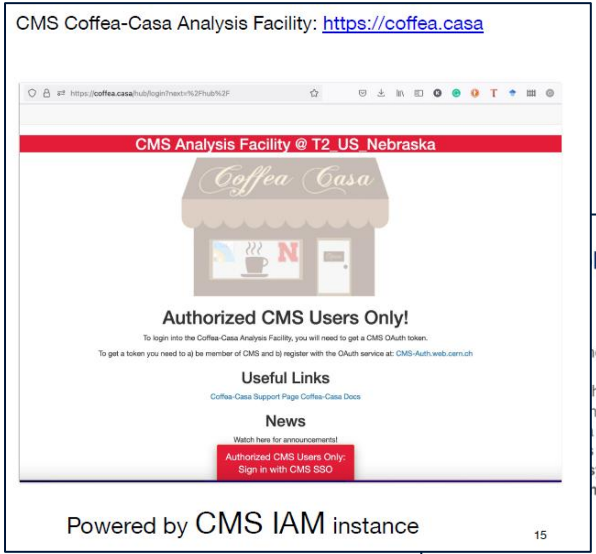
```
In [1]:  42
Out[1]: 42
```



'alien containers' HPC integration - container computing, using curated application images

Image sources: NDPF JupyterHub service "Callysto"; SLATE: Service Layer At The Edge – Rob Gartner (UChicago), Shawn KcMee (UMich) *et al.* – slateci.io

Containerised workloads: between ‘PaaS’ and ‘SaaS’



CMS Coffea-Casa Analysis Facility: https://coffea.casa

https://coffea.casa/hub/login?next=%2Fhub%2F

CMS Analysis Facility @ T2_US_Nebraska

Authorized CMS Users Only!

To login into the Coffea-Casa Analysis Facility, you will need to get a CMS OAuth token.

To get a token you need to a) be member of CMS and b) register with the OAuth service at: CMS-Auth.web.cern.ch

Useful Links

Coffea-Casa Support Page Coffea-Casa Docs

News

Watch here for announcements!

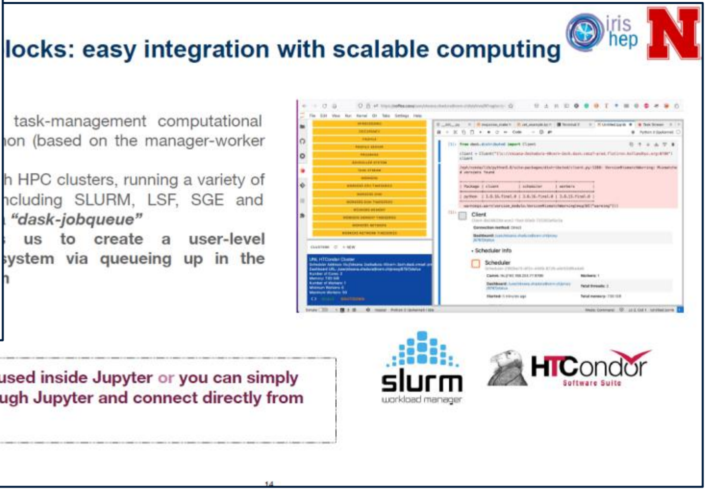
Authorized CMS Users Only: Sign in with CMS SSO

Powered by CMS IAM instance

15

CernVM File system

APPTAINER



locks: easy integration with scalable computing

task-management computational
ion (based on the manager-worker

h HPC clusters, running a variety of
ncluding SLURM, LSF, SGE and
“dask-jobqueue”
us to create a user-level
ystem via queueing up in the
n

Dask can be used inside Jupyter or you can simply launch it through Jupyter and connect directly from your laptop

slurm workload manager

HTCondor Software Suite

14

See also EESSI: the European Environment for Scientific Software Installations ...

EESSI EUROPEAN ENVIRONMENT FOR SCIENTIFIC SOFTWARE INSTALLATIONS

CernVM File system

ENVIRONMENT MODULES

Images: Oksana Shadura et al (UNebraska Lincoln), Brian Bockelman (Morgridge Institute) at CHEP2023 https://indico.jlab.org/event/459/contributions/11610/
EESSI software distribution (https://www.eessi.io/) is CVMFS + Modules

Storage for high-throughput processing

Basic storage properties are well known

- throughput
- IOPS – I/O Operations per Second
- seek-time (latency)

but not many **file systems** support *concurrent parallel access* by many clients

- both data **and** (file system or index) meta-data must be scalably distributed
- typically sacrifice either instant consistency, or (POSIX) semantics, (or scalability) in a distributed storage system

Common commercial solutions: GPFS, … but also NetApp, HDS, Dell-EMC, have their own
Common open source: BeeGFS, gluster, dCache, CephFS, Lustre, …

… likely do not use a file system if object storage does the job, *but then you need a catalogue/database*

‘Interesting’ distribution: client-side-managed GlusterFS

- scalable through independence of both clients and servers

- design is stateless: file system meta-data kept in each server’s file system

- data itself can be replicated and protected, but … inconsistencies in metadata linger around the corner in case of client failures (e.g. batch system worker nodes)

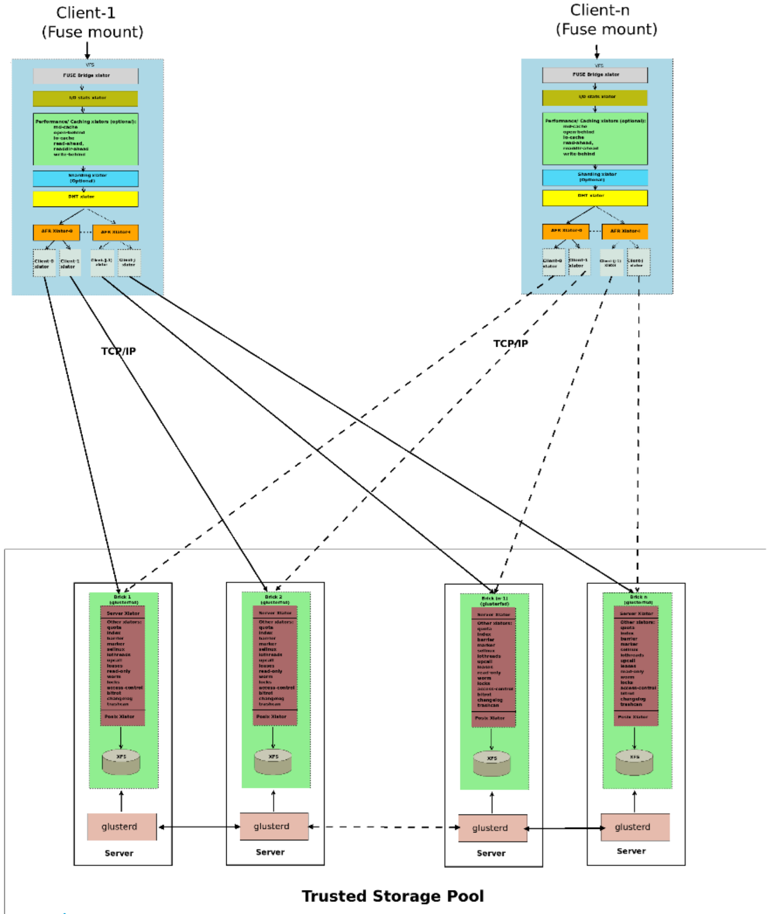
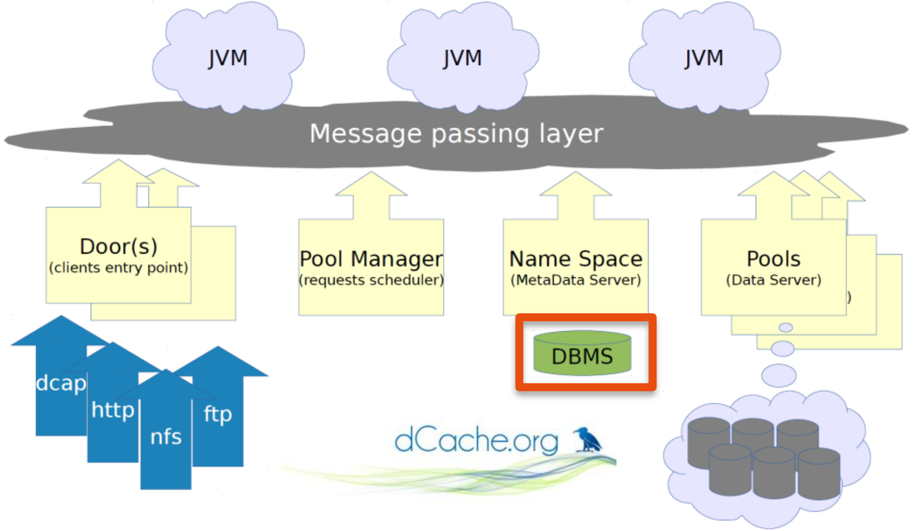
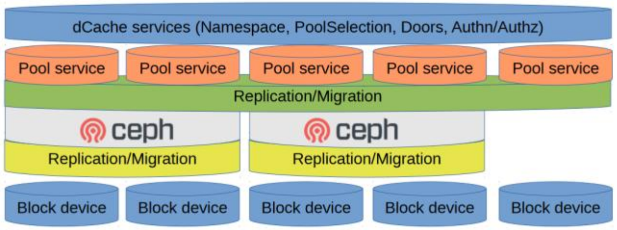


Image source Gluster community: https://docs.gluster.org/en/main/Quick-Start-Guide/Architecture/

Example: server-coherent distribution – dCache

- separate client entry points, storage access scheduling, filesystem meta-data (namespaces), and storage
- message layer for eventual consistency
- redirect-based access
  - doors and pools usually on all nodes
  - now also feature of standard NFSv4.1

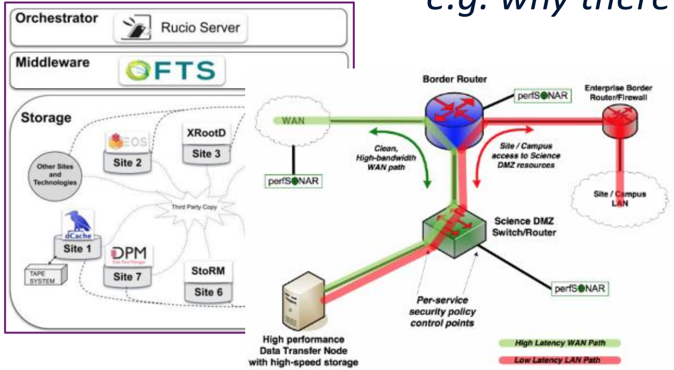




Images: Tigran Mkrtchyan (DESY, dCache.org), *dCache on steroids - delegated storage solutions*, ISGC 2016, https://dcache.org/manuals/publications.shtml

Structure of application data placement impacts storage (hardware) systems design

pre-staging all data locally allows for **latency hiding**,
posix-style access with lseek(2), and a fast, local, ‘$TMPDIR’
e.g. why there are Data Transfer Nodes (DTNs) in the ‘Science DMZ’ concept



Orchestrator
Rucio Server
Middleware
FTS
Storage
EOS
Site 2
XRootD
Site 3
Other Sites and Technologies
Third Party Copy
dCache
Site 1
TAPE SYSTEM
DPM
Site 7
StoRM
Site 6
Border Router
perfSONAR
Enterprise Border Router/Firewall
WAN
Clean, High-bandwidth WAN path
Site / Campus access to Science DMZ resources
perfSONAR
Site / Campus LAN
Science DMZ Switch/Router
perfSONAR
Per-service security policy control points
High performance Data Transfer Node with high-speed storage
High Latency WAN Path
Low Latency LAN Path

but, nowadays, pre-staging started coming at a cost, when using **SSDs**
as local ‘scratch’ area … because of their hardware characteristic ‘endurance’

Photo HGST nVME from: Dmitry Nosachev on Wikimedia Commons CC-BY-SA; Image Science DMZ and Data Transfer Nodes: ESnet fasterdata.es.net

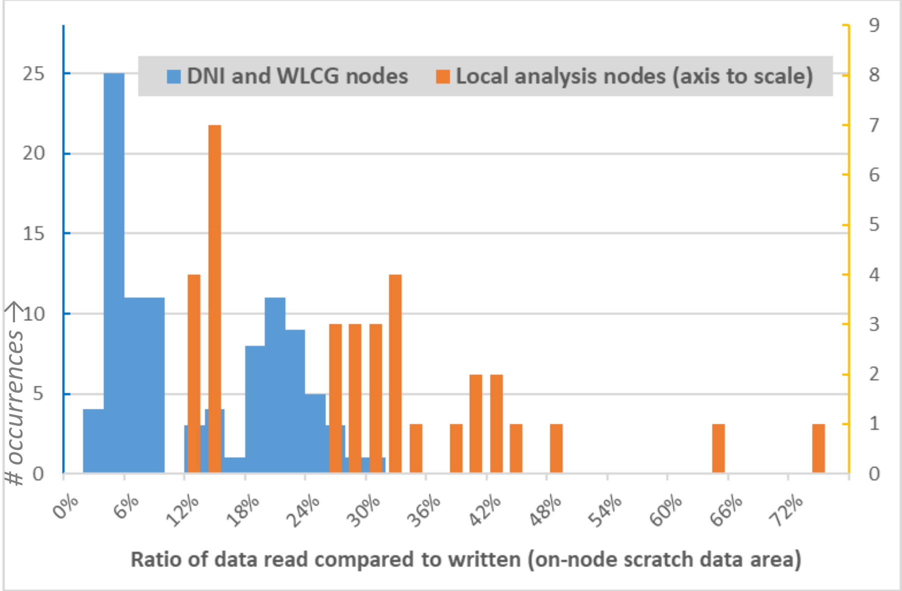
Especially with *WORN* storage: Write Once Read Never

Frequency distribution of **read-back vs. write** volume, observed on local scratch for NDPF execution nodes for ***outside ('grid') access (blue) vs local access (orange)***

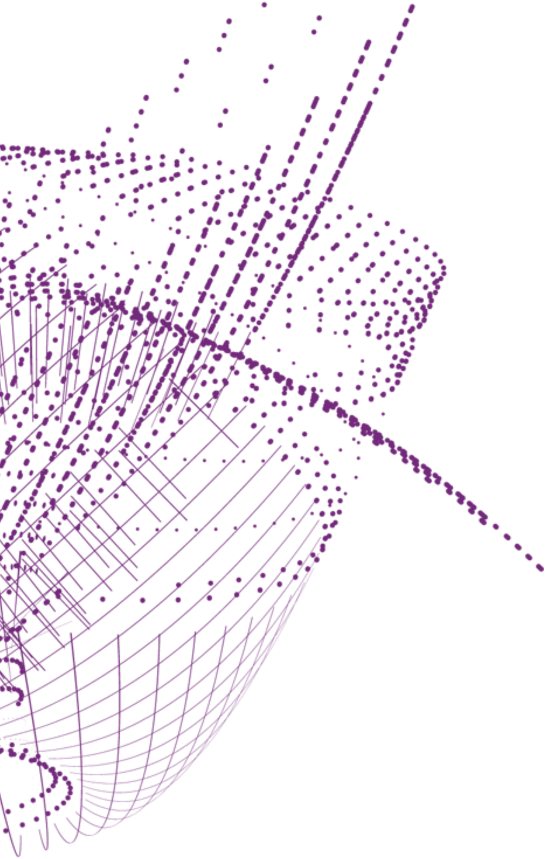
Access pattern is rather different. But why?

- external users pre-stage, because it is built into data management frameworks (like DIRAC, Athena),
- 'local' users stream output data (dCache with NFSv4) and use $TMPDIR mainly for merging partial results

Different types of workload (here analysis vs processing) determine the choice of systems hardware

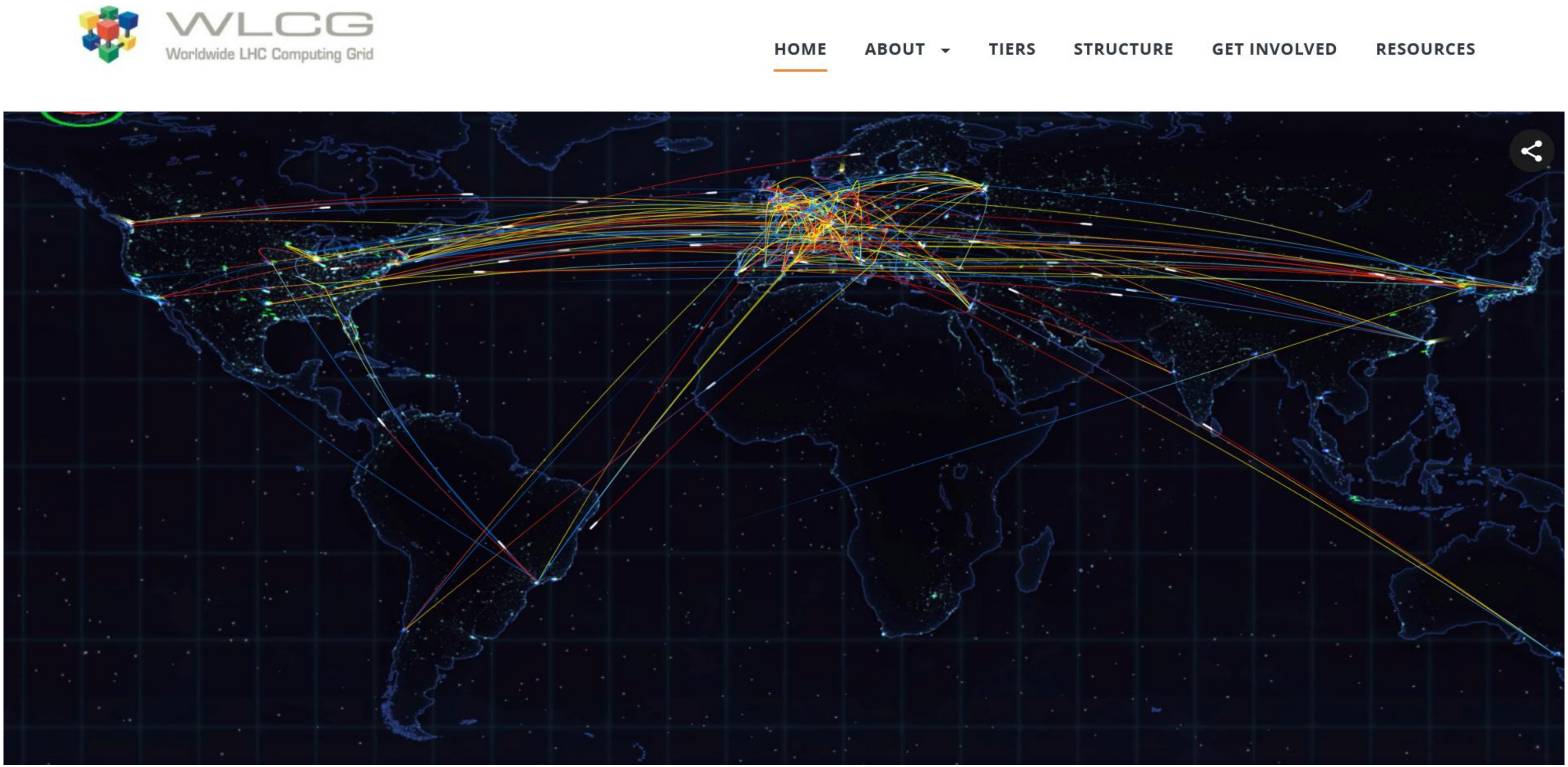


Data: NDPF execution nodes, based on SSD SMART data, integrated over total device lifetime; plot shows number of local analysis nodes scaled to DNI-WLCG count; collected using smartctl on 2020-10-28 – in total 97 'DNI' and 34 'STBC' SSDs were used in the analysis



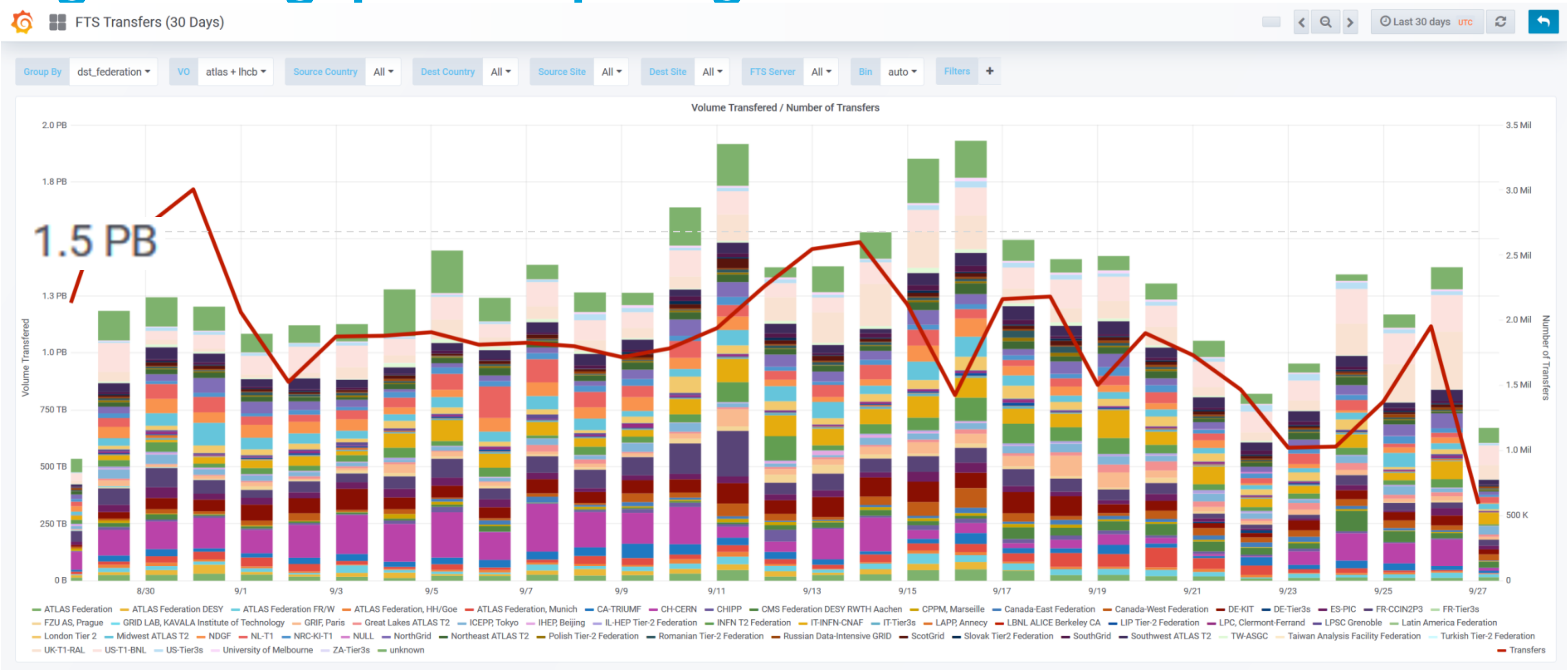
Beyond the single site

It’s all about data
… globally interconnected



https://wlcg.web.cern.ch/

High throughput computing includes data and networks



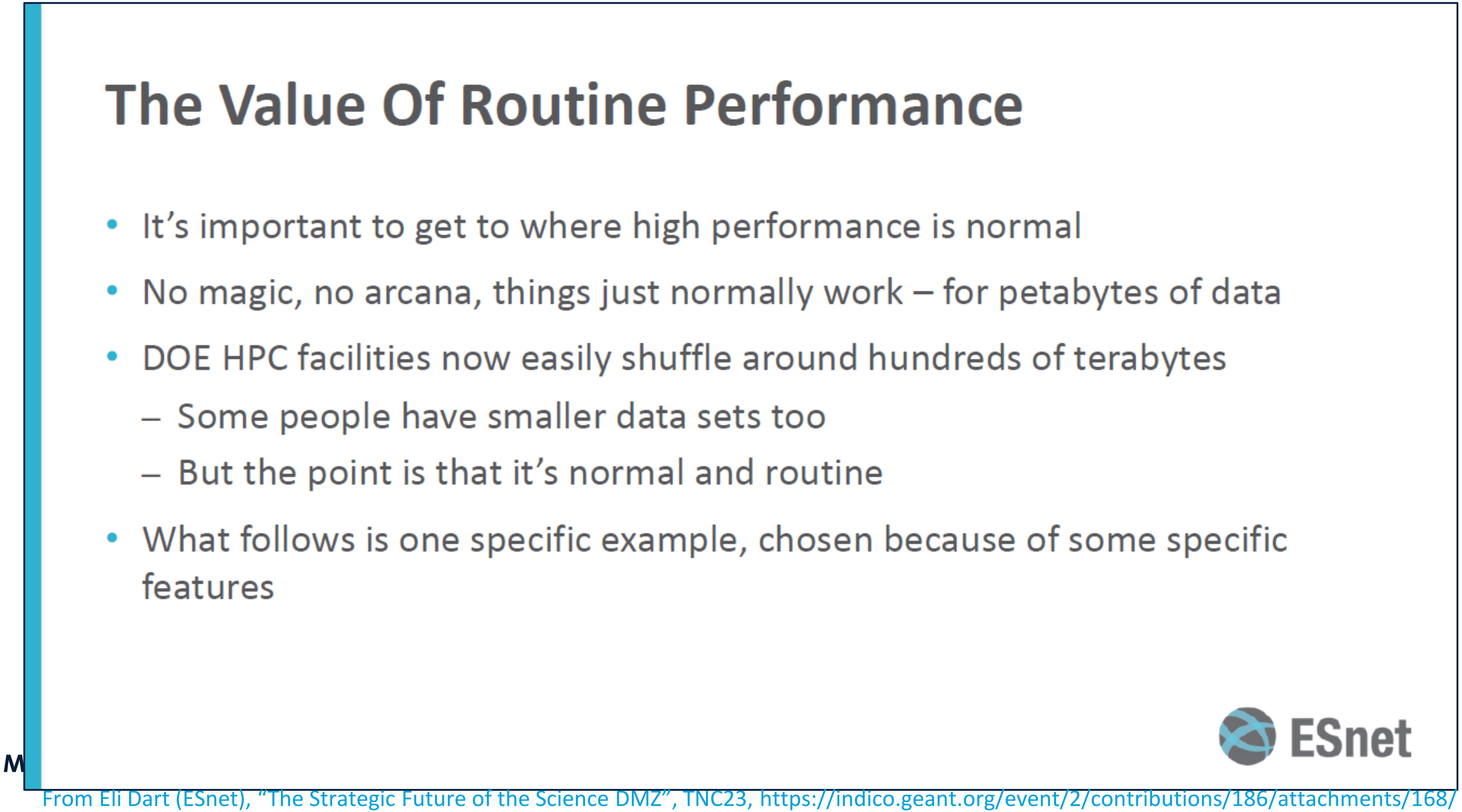
FTS Transfers (30 Days)

Last 30 days UTC

Group By: dst_federation
VO: atlas + lhcb
Source Country: All
Dest Country: All
Source Site: All
Dest Site: All
FTS Server: All
Bin: auto
Filters

source: https://monit-grafana.cern.ch/d/000000420/fts-transfers-30-day ; data: November 2020 ; CERN FTS instance WLCG: daily transfer volume ATLAS+LHCb

Can hardly be said better than Eli Dart did at TNC23



The Value Of Routine Performance

- It’s important to get to where high performance is normal
- No magic, no arcana, things just normally work – for petabytes of data
- DOE HPC facilities now easily shuffle around hundreds of terabytes
  - Some people have smaller data sets too
  - But the point is that it’s normal and routine
- What follows is one specific example, chosen because of some specific features

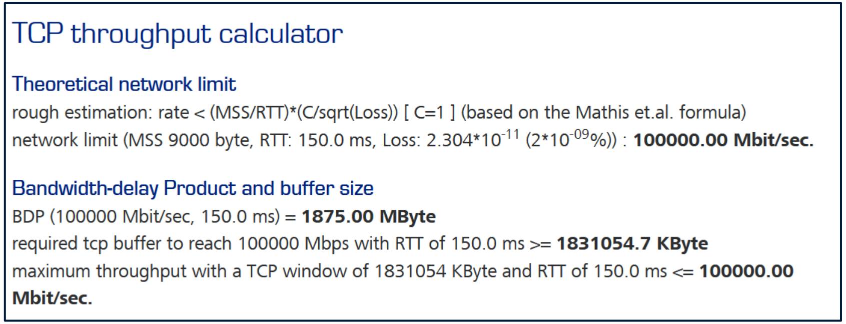
From Eli Dart (ESnet), “The Strategic Future of the Science DMZ”, TNC23, https://indico.geant.org/event/2/contributions/186/attachments/168/

Network is more than just what it says on the tin

More network bandwidth does not mean your *data* gets there faster

- memory requirements (since TCP needs a capability to re-transmit)

- tcp ‘slow start’
- congestion control algorithms



TCP throughput calculator

Theoretical network limit

rough estimation: rate < (MSS/RTT)*(C/sqrt(Loss)) [C=1] (based on the Mathis et.al. formula)

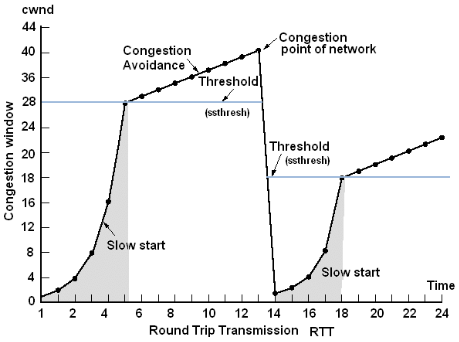
network limit (MSS 9000 byte, RTT: 150.0 ms, Loss: $2.304*10^{-11}$ ($2*10^{-09}$%)) : **100000.00 Mbit/sec.**

Bandwidth-delay Product and buffer size

BDP (100000 Mbit/sec, 150.0 ms) = **1875.00 MByte**

required tcp buffer to reach 100000 Mbps with RTT of 150.0 ms >= **1831054.7 KByte**

maximum throughput with a TCP window of 1831054 KByte and RTT of 150.0 ms <= **100000.00 Mbit/sec.**



Useful sources: https://www.switch.ch/network/tools/tcp_throughput/, https://fasterdata.es.net/

tcp slow-start graphic from Abed et al, *Improvement of TCP Congestion Window over LTE- Advanced Networks* IJoARiC&CE 2012

The cat video that destroyed it all …

latency AMS-GVA 17 ms
congestion event @20ms:
2 ms of UDP traffic to GVA

- TCP protocol sensitive to packet loss
  - 3 lost packets is enough to trigger this
- different congestion avoidance algorithms exists (~20 by now)
- loss severely impacts links w/large ‘bandwidth-delay-product’ (BDP)
- NL: ~3 ms, US East: 150ms

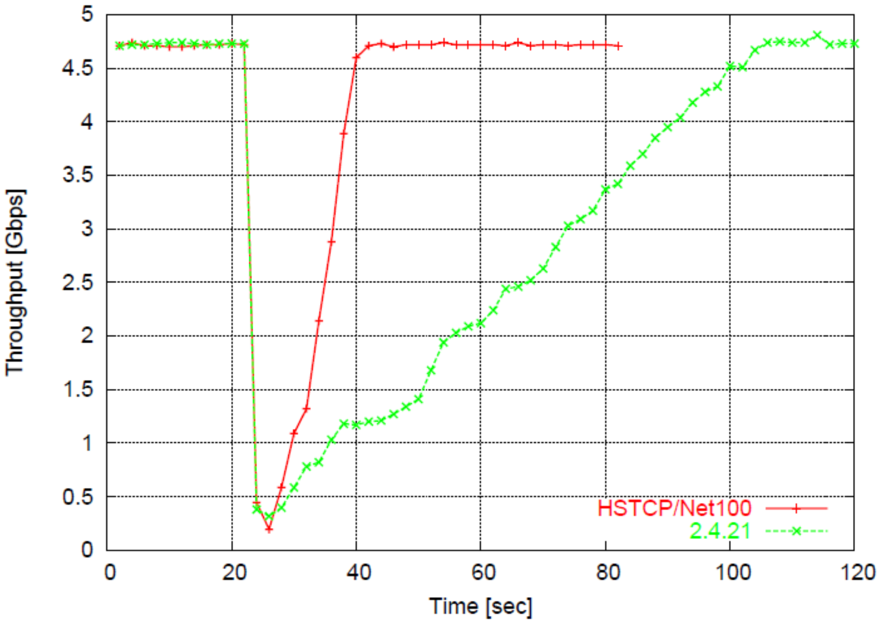
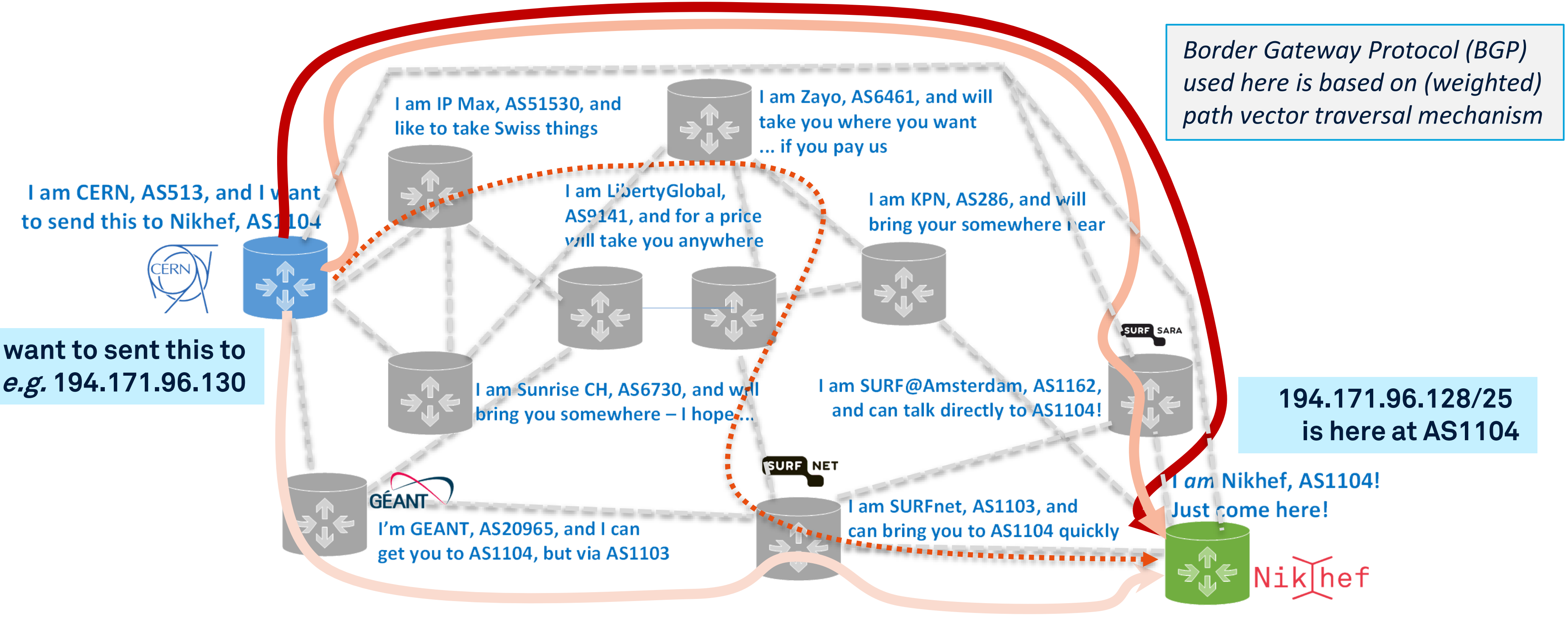


Figure 10: HSTCP versus stock TCP recovery time

source: Catalin Meirosu et al. *Native 10 Gigabit Ethernet experiments over long distances* in FGCS, doi:10.1016/j.future.2004.10.003 – aka. ATL-D-TN-0001

Fast track - and getting rid of cat videos ...



Border Gateway Protocol (BGP) used here is based on (weighted) path vector traversal mechanism

I am CERN, AS513, and I want to send this to Nikhef, AS1104

I want to sent this to *e.g.* 194.171.96.130

I am IP Max, AS51530, and like to take Swiss things

I am Zayo, AS6461, and will take you where you want ... if you pay us

I am LibertyGlobal, AS9141, and for a price will take you anywhere

I am KPN, AS286, and will bring your somewhere near

I am Sunrise CH, AS6730, and will bring you somewhere – I hope....

I am SURF@Amsterdam, AS1162, and can talk directly to AS1104!

194.171.96.128/25 is here at AS1104

I'm GEANT, AS20965, and I can get you to AS1104, but via AS1103

I am SURFnet, AS1103, and can bring you to AS1104 quickly

I am Nikhef, AS1104! Just come here!

grey-dash lines for illustration only: may not correspond to actual peerings or transit agreements; red lines: the three existing LHCOPN and R&E fall-back routes; yellow: public internet fall-back (least preferred option)

LHCOPN – distributing raw data

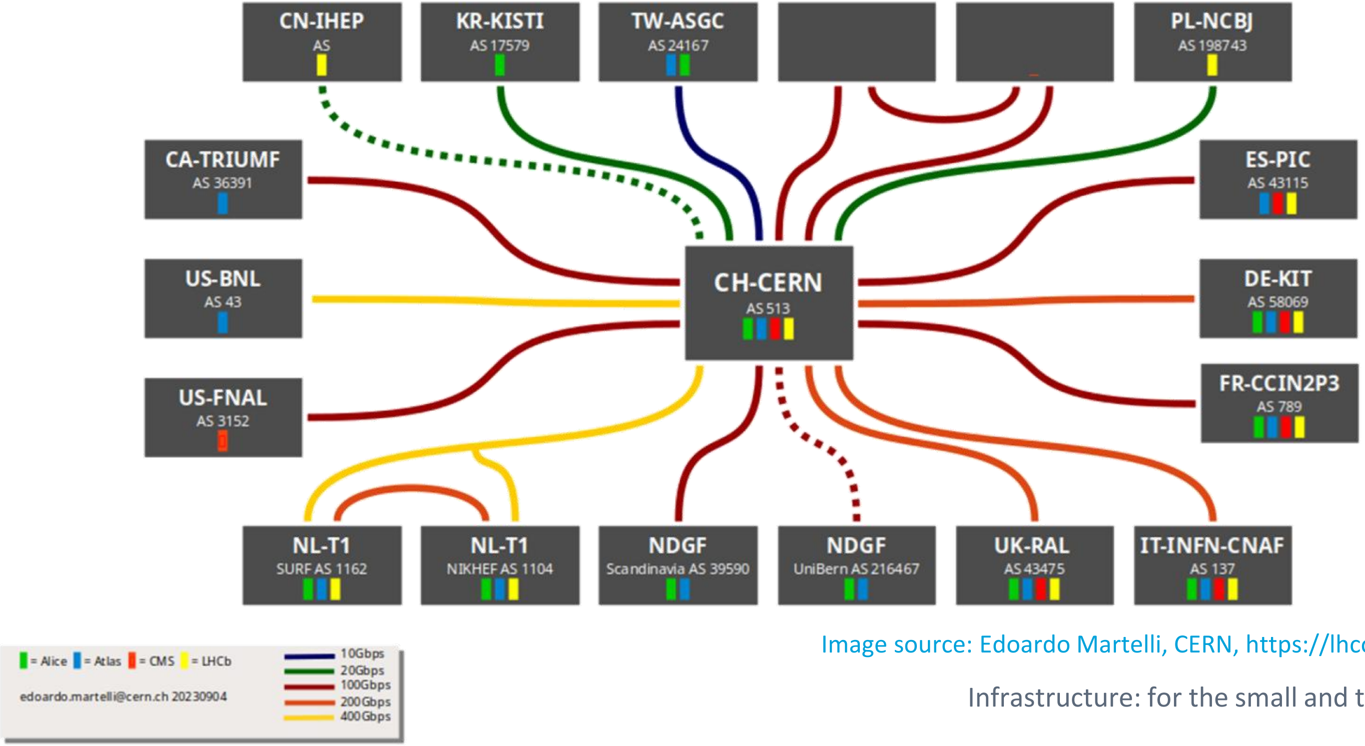
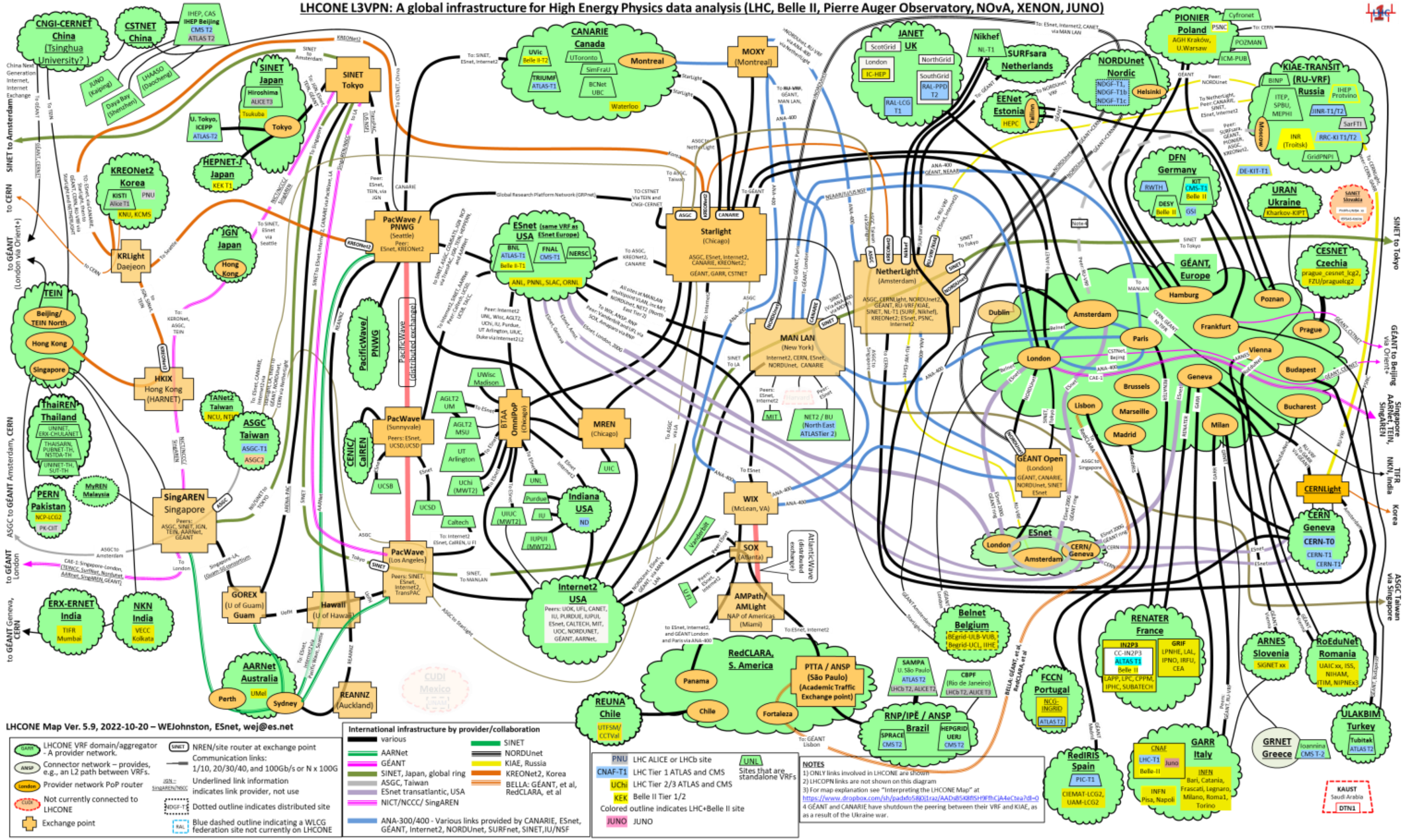


Image source: Edoardo Martelli, CERN, https://lhcopn.web.cern.ch/

LHCone


LHCONE L3VPN: A global infrastructure for High Energy Physics data analysis (LHC, Belle II, Pierre Auger Observatory, NOvA, XENON, JUNO)
CNGI-CERNET China [Tsinghua University?]
CSTNET China
IHEP, CAS IHEP Beijing CMS T2 ATLAS T2
China Next Generation Internet, Internet Exchange
JUNO (Kaiping)
Daya Bay (Shenzhen)
LHAASO (Daocheng)
SINET Japan
Hiroshima ALICE T3
Tsukuba
SINET Tokyo
Tokyo
U. Tokyo, ICEPP ATLAS T2
HEPNET-J Japan KEK T1
KREONet2 Korea
KISTI Alice T1
PNU
KNU, KCMS
KRLight Daejeon
JGN Japan
Hong Kong
TEIN
Beijing/ TEIN North
Hong Kong
Singapore
HKIX Hong Kong (HARNET)
ThaiREN Thailand
PERN Pakistan NCP-LCG2 PK-CIIT
MyREN Malaysia
TANet2 Taiwan NCU, NTU
ASGC Taiwan ASGC-T1 ASGC2
SingAREN Singapore
Peers: ASGC, SINET, JGN, TEIN, AARNet, GEANT
ERX-ERNET India TIFR Mumbai
NKN India VECC Kolkata
GOREX (U of Guam) Guam
Hawaii (U of Hawaii)
AARNet Australia UMel Perth Sydney
REANNZ (Auckland)
CUDI Mexico
PacWave / PNWG (Seattle) Peer: ESnet, KREONet2
PacWave (distributed exchange)
PacificWave/ PNWG
PacWave (Sunnyvale) Peers: ESnet, UCSD, UCSD
CENIC/ CalREN
UCSB
UCSD
PacWave (Los Angeles) Peers: SINET, ESnet, Internet2, TransPAC
CANARIE Canada
UVic Belle II-T2
UToronto
SimFraU
TRIUMF ATLAS-T1
BCNet UBC
Montreal
Waterloo
MOXY (Montreal)
ESnet (same VRF as ESnet Europe) USA
BNL ATLAS-T1 Belle II-T1
FNAL CMS-T1
NERSC
ANL, PNNL, SLAC, ORNL
Starlight (Chicago)
ASGC, ESnet, Internet2, CANARIE, KREONet2; GÉANT, GARR, CSTNET
UWisc Madison
AGLT2 UM
AGLT2 MSU
UT Arlington
UChi (MWT2)
BTAA OmniPoP (Chicago)
MREN (Chicago)
UIC
UNL
Purdue
IU
UIUC (MWT2)
IUPUI (MWT2)
Indiana USA ND
Caltech
Internet2 USA
Peers: UOK, UFL, CANET, IU, PURDUE, IUPUI, ESnet, CALTECH, MIT, UOC, NORDUNET, GÉANT, AARNet
MAN LAN (New York)
Internet2, CERN, ESnet, NORDUnet, CANARIE
MIT
NET2 / BU (North East ATLASTier 2)
WIX (McLean, VA)
SOX (Atlanta)
AtlanticWave (distributed exchange)
AMPath/ AMLight NAP of Americas (Miami)
RedCLARA, S. America
Panama
Chile
Fortaleza
PTTA / ANSP (São Paulo) (Academic Traffic Exchange point)
SAMPA U. São Paulo ATLAS T2 LHCb T2, ALICE T2
CBPF (Rio de Janeiro) LHCb T2, ALICE T3
RNP/IPÊ / ANSP Brazil
SPRACE CMS T2
HEPGRID UERJ CMS T2
REUNA Chile UTFSM/ CCTVal
JANET UK
ScotGrid
London IC-HEP
NorthGrid
SouthGrid RAL-PPD T2
RAL-LCG T1
Nikhef NL-T1
SURFsara Netherlands
NetherLight (Amsterdam)
ASGC, CERNLight, NORDUnet2, GÉANT, RU-VRF/KIAE, SINET, NL-T1 (SURF, Nikhef), KREONet2; ESnet, PSNC, Internet2
EENet Estonia HEPC
Tallinn
NORDUnet Nordic
NDGF-T1, NDGF-T1b, NDGF-T1c
Helsinki
PIONIER Poland
AGH Kraków, U.Warsaw
PSNC
Cyfronet
POZMAN
ICM-PUB
KIAE-TRANSIT (RU-VRF) Russia
BINP
ITEP, SPBU, MEPHI
IHEP Protvino
JINR-T1/T2
SarFTI
INR (Troitsk)
RRC-KI T1/T2
GridPNPI
Moscow
DFN Germany
RWTH
KIT CMS-T1 Belle II
DESY Belle II
GSI
DE-KIT-T1
URAN Ukraine Kharkov-KIPT
SANET Slovakia
CESNET Czechia prague_cesnet_lcg2 FZU/praguelcg2
GÉANT Europe
Hamburg
Poznan
Amsterdam
Dublin
Paris
Frankfurt
Prague
London
Vienna
Budapest
Geneva
Brussels
Lisbon
Marseille
Bucharest
Madrid
Milan
GÉANT Open (London)
GÉANT, CANARIE, NORDUnet, SINET ESnet
ESnet
London
Amsterdam
CERN/ Geneva
CERNLight
CERN Geneva CERN-T0 CERN-T1 CERN-T1
Belnet Belgium
BEgrid-ULB-VUB, Begrid-UCL, IIHE
RENATER France
IN2P3 CC-IN2P3 ATLAS T1 Belle II
GRIF LPNHE, LAL, IPNO, IRFU, CEA
LAPP, LPC, CPPM, IPHC, SUBATECH
FCCN Portugal NCG-INGRID ATLAS T2
RedIRIS Spain PIC-T1
CIEMAT-LCG2, UAM-LCG2
GARR Italy
CNAF LHC-T1 Juno Belle-II
INFN Bari, Catania, Frascati, Legnaro, Milano, Roma1, Torino
INFN Pisa, Napoli
ARNES Slovenia SiGNET xx
RoEduNet Romania UAIC xx, ISS, NIHAM, ITIM, NIPNEx3
ULAKBIM Turkey Tubitak ATLAS T2
GRNET Greece Ioannina CMS T-2
KAUST Saudi Arabia DTN1
SINET to Amsterdam
to CERN
to GÉANT (London via Orient+)
ASGC to GÉANT Amsterdam, CERN
to GÉANT London
to GÉANT Geneva, CERN
SINET to Tokyo
GÉANT to Beijing via Orient+
Singapore AARNet, TEIN, SingAREN
TIFR, India
Korea
ASGC Taiwan via Singapore
LHCONE Map Ver. 5.9, 2022-10-20 – WEJohnston, ESnet, wej@es.net
GARR LHCONE VRF domain/aggregator - A provider network.
ANSP Connector network – provides, e.g., an L2 path between VRFs.
London Provider network PoP router
CUDI Not currently connected to LHCONE
Exchange point
SINET NREN/site router at exchange point
Communication links: 1/10, 20/30/40, and 100Gb/s or N x 100G
JGN – SingAREN/NICC Underlined link information indicates link provider, not use
NDGF-T1 Dotted outline indicates distributed site
RAL Blue dashed outline indicating a WLCG federation site not currently on LHCONE
International infrastructure by provider/collaboration
various
AARNet
GÉANT
SINET, Japan, global ring
ASGC, Taiwan
ESnet transatlantic, USA
NICT/NCCC/ SingAREN
SINET
NORDUnet
KIAE, Russia
KREONet2, Korea
BELLA: GÉANT, et al, RedCLARA, et al
ANA-300/400 - Various links provided by CANARIE, ESnet, GÉANT, Internet2, NORDUnet, SURFnet, SINET,IU/NSF
PNU LHC ALICE or LHCb site
CNAF-T1 LHC Tier 1 ATLAS and CMS
UChi LHC Tier 2/3 ATLAS and CMS
KEK Belle II Tier 1/2
Colored outline indicates LHC+Belle II site
JUNO JUNO
UNL Sites that are standalone VRFs
NOTES
1) ONLY links involved in LHCONE are shown
2) LHCOPN links are not shown on this diagram
3) For map explanation see "Interpreting the LHCONE Map" at https://www.dropbox.com/sh/padxfo58j0j1raz/AADsB5K8fISH9FfhCjA4eCtea?dl=0
4 GÉANT and CANARIE have shutdown the peering between their VRF and KIAE, as as a result of the Ukraine war.

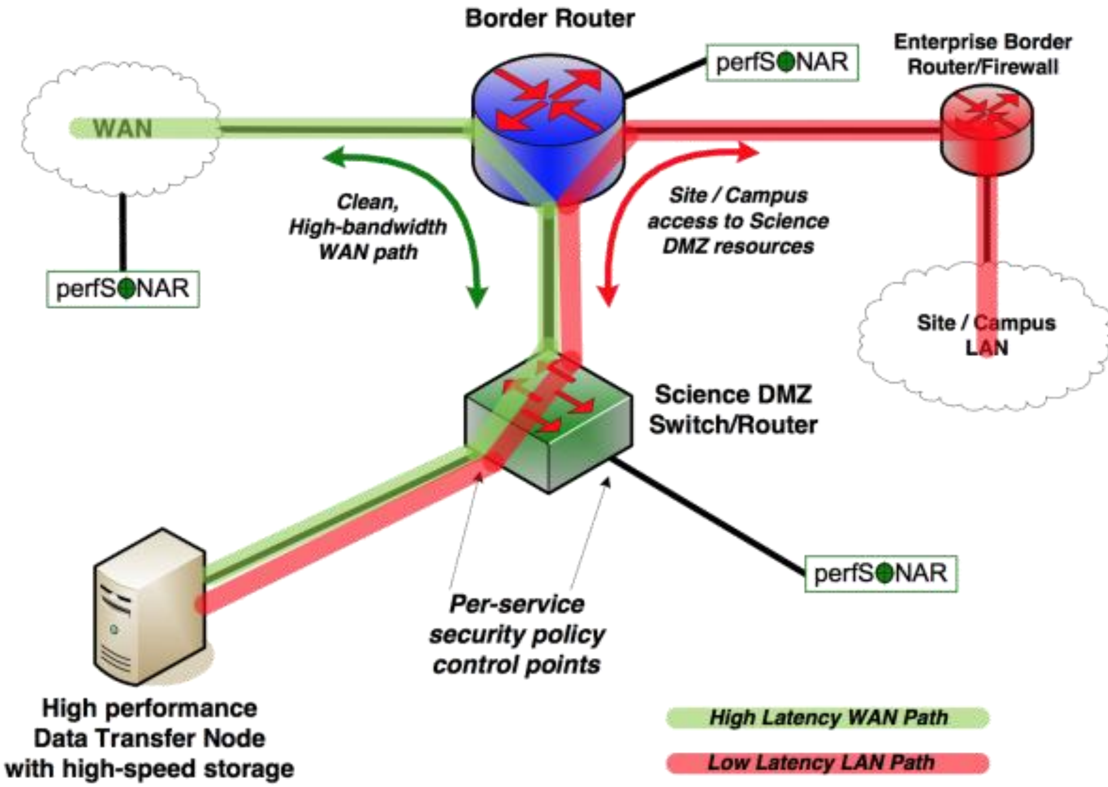
LHCone ("LHC Open Network Environment") – visualization by Bill Johnston, ESnet version: October 2022 – updated with new AS1104 links

‘ScienceDMZ’

Predicable performance and data access for research

‘where research services, data, and researchers meet’

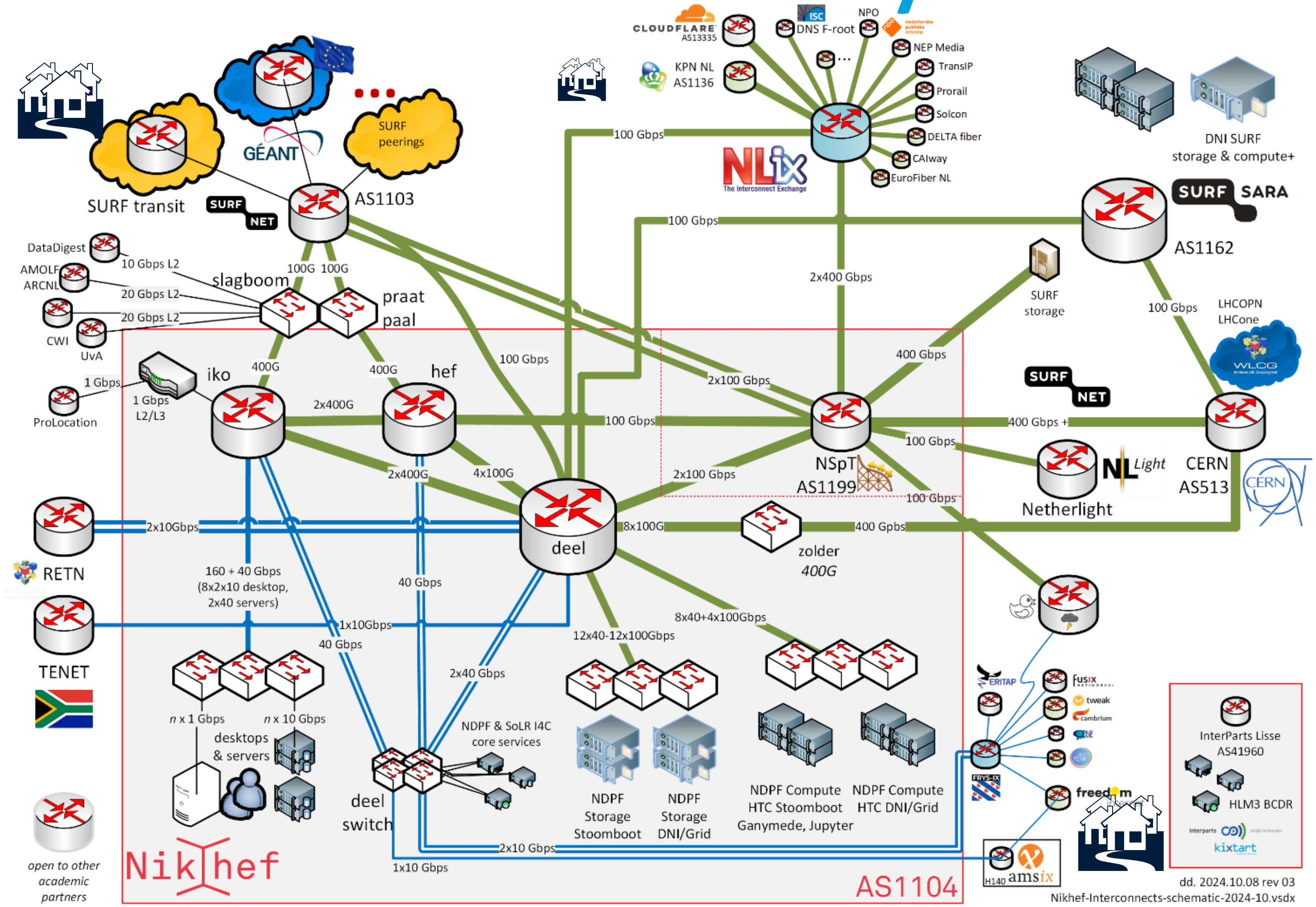
- latency hiding through caching
- **security zoning/segmentation** protects specific data sets
- **outside any enterprise perimeter**



Border Router
perfSONAR
Enterprise Border Router/Firewall
WAN
perfSONAR
Clean, High-bandwidth WAN path
Site / Campus access to Science DMZ resources
Site / Campus LAN
Science DMZ Switch/Router
perfSONAR
Per-service security policy control points
High performance Data Transfer Node with high-speed storage
High Latency WAN Path
Low Latency LAN Path

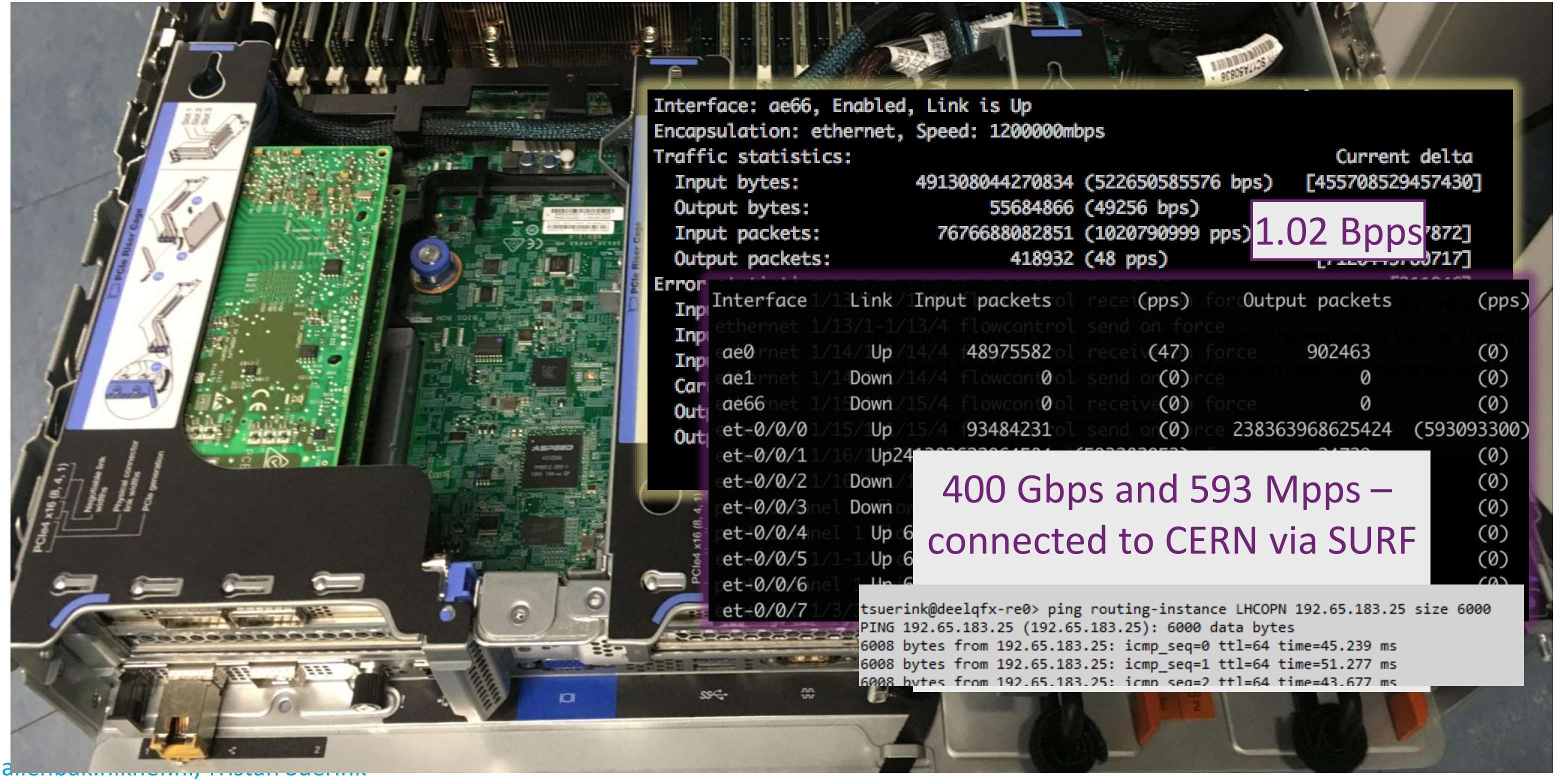
Image and ‘ScienceDMZ’ concept promulgated by ESnet (see fasterdata.es.net)

Just one random autonomous system: AS1104



AS1104
state as of Oct 2024

Exercising the network – sensor data and events



```
Interface: ae66, Enabled, Link is Up
Encapsulation: ethernet, Speed: 1200000mbps
Traffic statistics:                                    Current delta
  Input bytes:         491308044270834 (522650585576 bps)  [455708529457430]
  Output bytes:               55684866 (49256 bps)
  Input packets:         7676688082851 (1020790999 pps)            872]
  Output packets:               418932 (48 pps)                    717]
Error
  Inp
  Inp
  Inp
  Car
  Out
  Out
```

1.02 Bpps

```
Interface    Link   Input packets     (pps)   Output packets       (pps)
 ae0          Up     48975582         (47)         902463          (0)
 ae1          Down          0          (0)              0          (0)
 ae66         Down          0          (0)              0          (0)
 et-0/0/0     Up     93484231          (0)   238363968625424  (593093300)
 et-0/0/1     Up24                                                 (0)
 et-0/0/2     Down                                                 (0)
 et-0/0/3     Down                                                 (0)
 et-0/0/4     Up 6                                                 (0)
 et-0/0/5     Up 6                                                 (0)
 et-0/0/6
 et-0/0/7
```

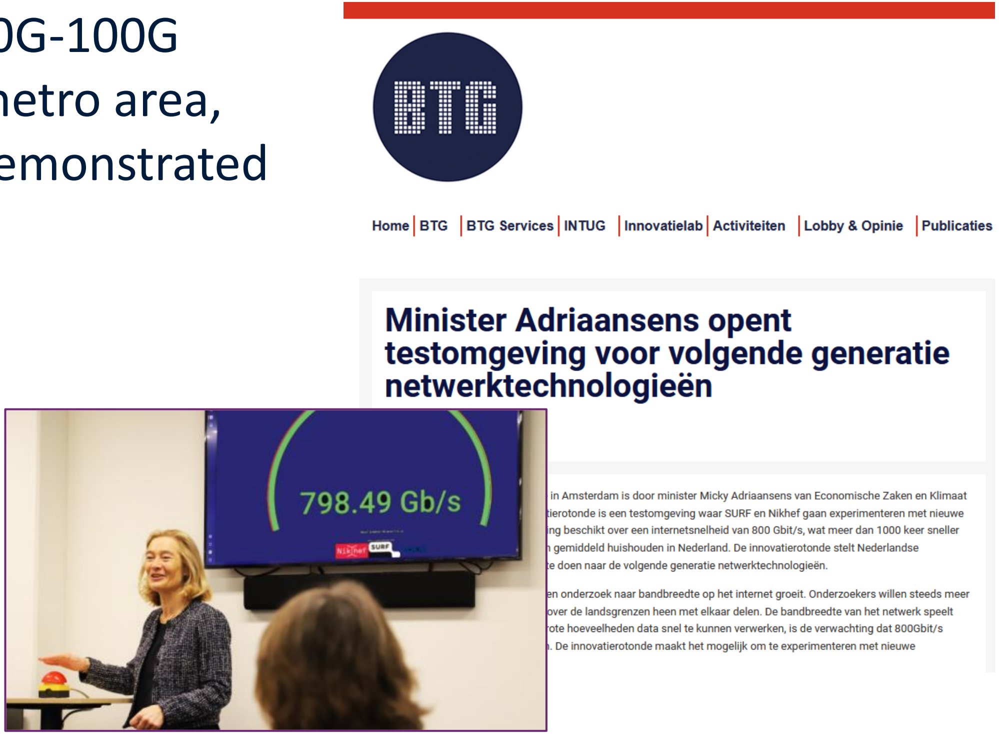
400 Gbps and 593 Mpps – connected to CERN via SURF

```
tsuerink@deelqfx-re0> ping routing-instance LHCOPN 192.65.183.25 size 6000
PING 192.65.183.25 (192.65.183.25): 6000 data bytes
6008 bytes from 192.65.183.25: icmp_seq=0 ttl=64 time=45.239 ms
6008 bytes from 192.65.183.25: icmp_seq=1 ttl=64 time=51.277 ms
6008 bytes from 192.65.183.25: icmp_seq=2 ttl=64 time=43.677 ms
```

Image: ballenbak.nikhef.nl, Tristan Suerink

For example for HL-LHC, or SKA, more is needed > 2028 …

- ‘Typical’ network is now mixed 400G-100G
- Push experiments to 800Gbps in metro area, and a local (AMS) loop has been demonstrated
- next: 800 → 1600G AMS-GVA ☺


BTG
Home | BTG | BTG Services | INTUG | Innovatielab | Activiteiten | Lobby & Opinie | Publicaties
798.49 Gb/s
Nikhef
SURF

Minister Adriaansens opent testomgeving voor volgende generatie netwerktechnologieën

in Amsterdam is door minister Micky Adriaansens van Economische Zaken en Klimaat
tierotonde is een testomgeving waar SURF en Nikhef gaan experimenteren met nieuwe
ing beschikt over een internetsnelheid van 800 Gbit/s, wat meer dan 1000 keer sneller
n gemiddeld huishouden in Nederland. De innovatierotonde stelt Nederlandse
te doen naar de volgende generatie netwerktechnologieën.

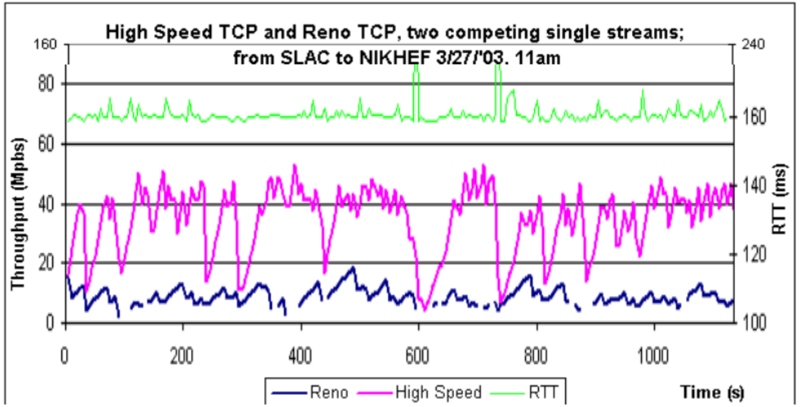
en onderzoek naar bandbreedte op het internet groeit. Onderzoekers willen steeds meer
over de landsgrenzen heen met elkaar delen. De bandbreedte van het netwerk speelt
rote hoeveelheden data snel te kunnen verwerken, is de verwachting dat 800Gbit/s
n. De innovatierotonde maakt het mogelijk om te experimenteren met nieuwe

Web screenshot: btg.org,
Images Nokia 7750-SR1x in Nikhef AMS H234b: Tristan Suerink

Scaling data access: ‘system-aware design’ at application layer

Reading data ‘scattered’ in a file - simply using POSIX-like IO - when done over the network severely exposes latency

and TCP slow-start makes that even worse



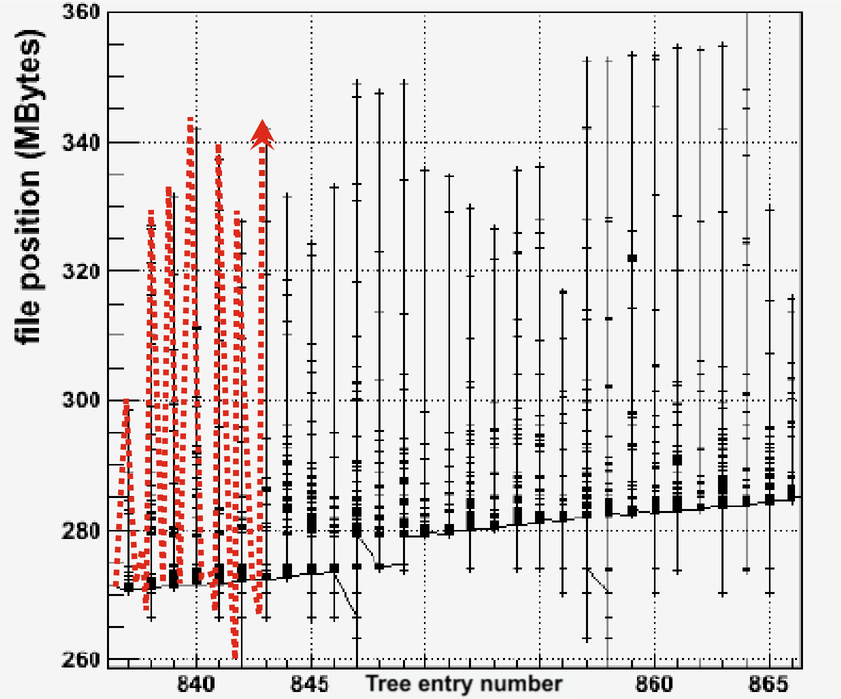
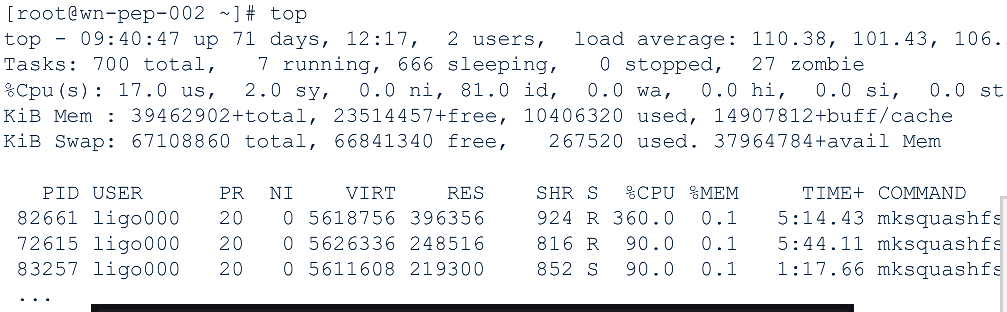


Image of TCP slow-start and packet loss impact (in Mpps): Antony Antony et al., Nikhef, for DataTAG, 2003(!)
Right: base graphic: Philippe Canal “Root I/O: the fast and the furious”, CHEP2010 Access pattern reflects Root versions < 5.28, before Ttree caching and ‘baskets’

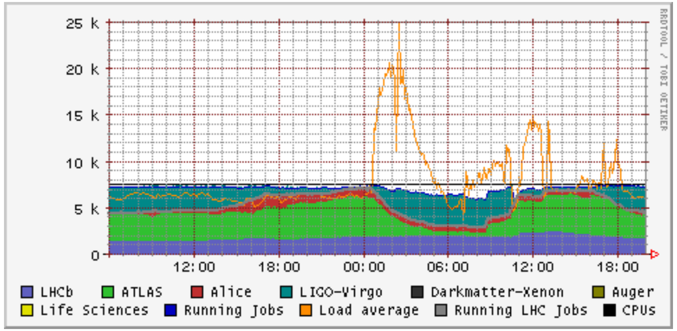
And some traffic is triggered by researchers scaling up ‘accidentally’ from a laptop to a cluster without too much thought

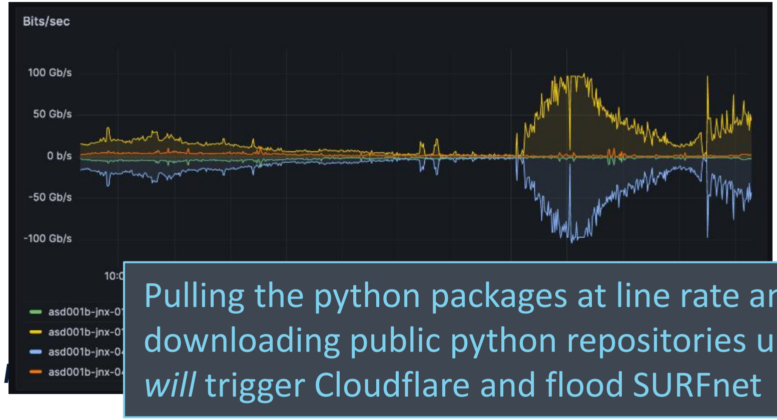
A researcher doing mass creation of containers, rebuilding their python ‘virtual env’ for each job, running on >> 4000 cores



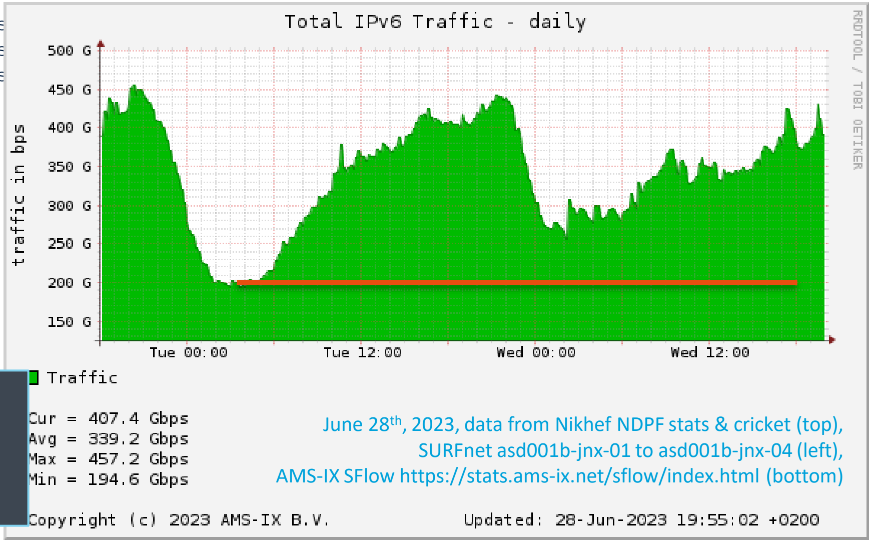
```
[root@wn-pep-002 ~]# top
top - 09:40:47 up 71 days, 12:17,  2 users,  load average: 110.38, 101.43, 106.3
Tasks: 700 total,   7 running, 666 sleeping,   0 stopped,  27 zombie
%Cpu(s): 17.0 us,  2.0 sy,  0.0 ni, 81.0 id,  0.0 wa,  0.0 hi,  0.0 si,  0.0 st
KiB Mem : 39462902+total, 23514457+free, 10406320 used, 14907812+buff/cache
KiB Swap: 67108860 total, 66841340 free,   267520 used. 37964784+avail Mem

  PID USER      PR  NI    VIRT    RES    SHR S  %CPU %MEM     TIME+ COMMAND
 82661 ligo000   20   0 5618756 396356    924 R 360.0  0.1   5:14.43 mksquashfs
 72615 ligo000   20   0 5626336 248516    816 R  90.0  0.1   5:44.11 mksquashfs
 83257 ligo000   20   0 5611608 219300    852 S  90.0  0.1   1:17.66 mksquashfs
 ...
```



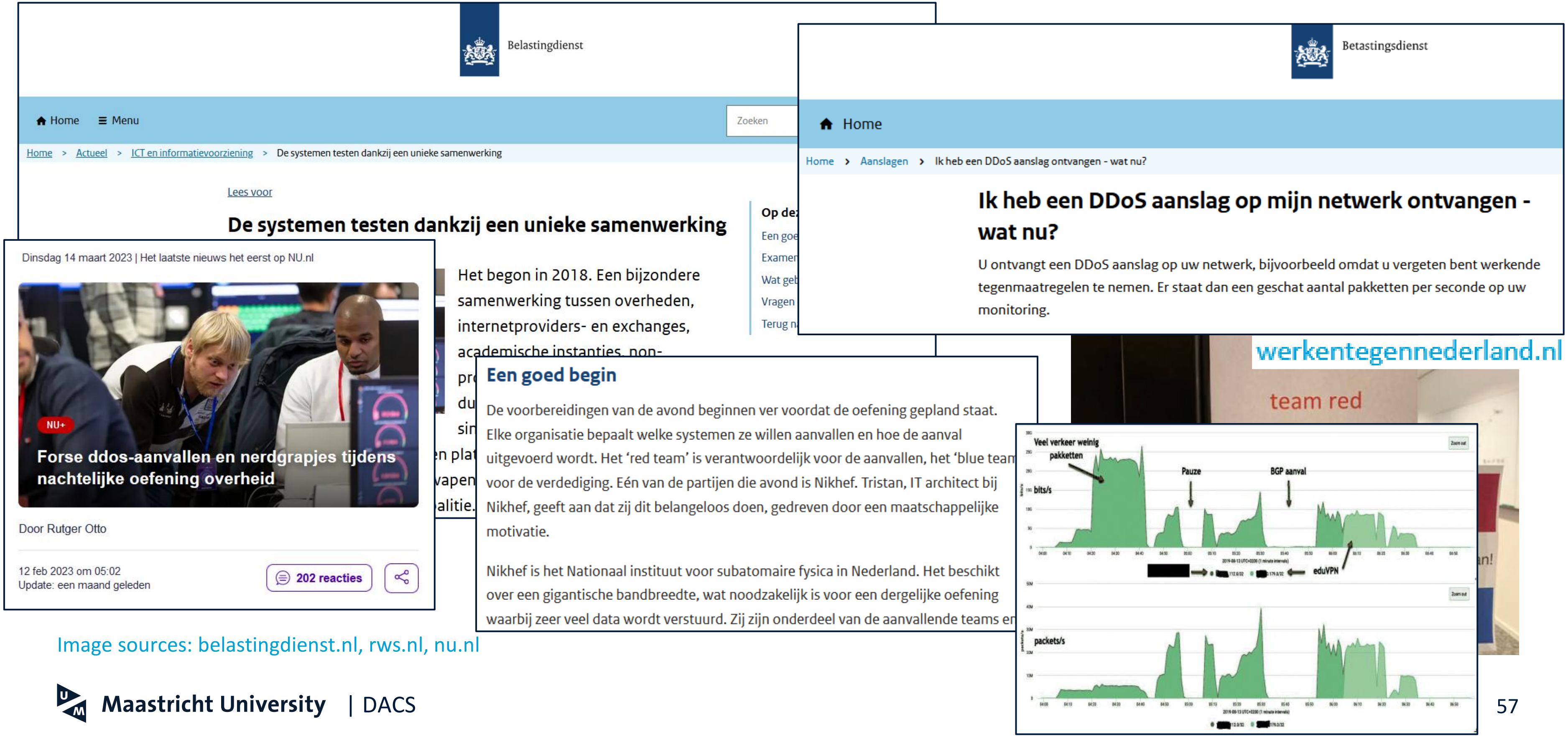


Pulling the python packages at line rate and downloading public python repositories ultimately *will* trigger Cloudflare and flood SURFnet



June 28th, 2023, data from Nikhef NDPF stats & cricket (top), SURFnet asd001b-jnx-01 to asd001b-jnx-04 (left), AMS-IX SFlow https://stats.ams-ix.net/sflow/index.html (bottom)

Research data traffic looks like … a DDoS to others ☺



Belastingdienst

Home ≡ Menu

Zoeken

Home > Actueel > ICT en informatievoorziening > De systemen testen dankzij een unieke samenwerking

Lees voor

De systemen testen dankzij een unieke samenwerking

Het begon in 2018. Een bijzondere samenwerking tussen overheden, internetproviders- en exchanges, academische instanties, non- pr du sin en plat wapen alitie.

Op de

Een goe

Examen

Wat geb

Vragen

Terug n

Betastingsdienst

Home

Home › Aanslagen › Ik heb een DDoS aanslag ontvangen - wat nu?

Ik heb een DDoS aanslag op mijn netwerk ontvangen - wat nu?

U ontvangt een DDoS aanslag op uw netwerk, bijvoorbeeld omdat u vergeten bent werkende tegenmaatregelen te nemen. Er staat dan een geschat aantal pakketten per seconde op uw monitoring.

Dinsdag 14 maart 2023 | Het laatste nieuws het eerst op NU.nl

NU+

Forse ddos-aanvallen en nerdgrapjes tijdens nachtelijke oefening overheid

Door Rutger Otto

12 feb 2023 om 05:02
Update: een maand geleden

202 reacties

Een goed begin

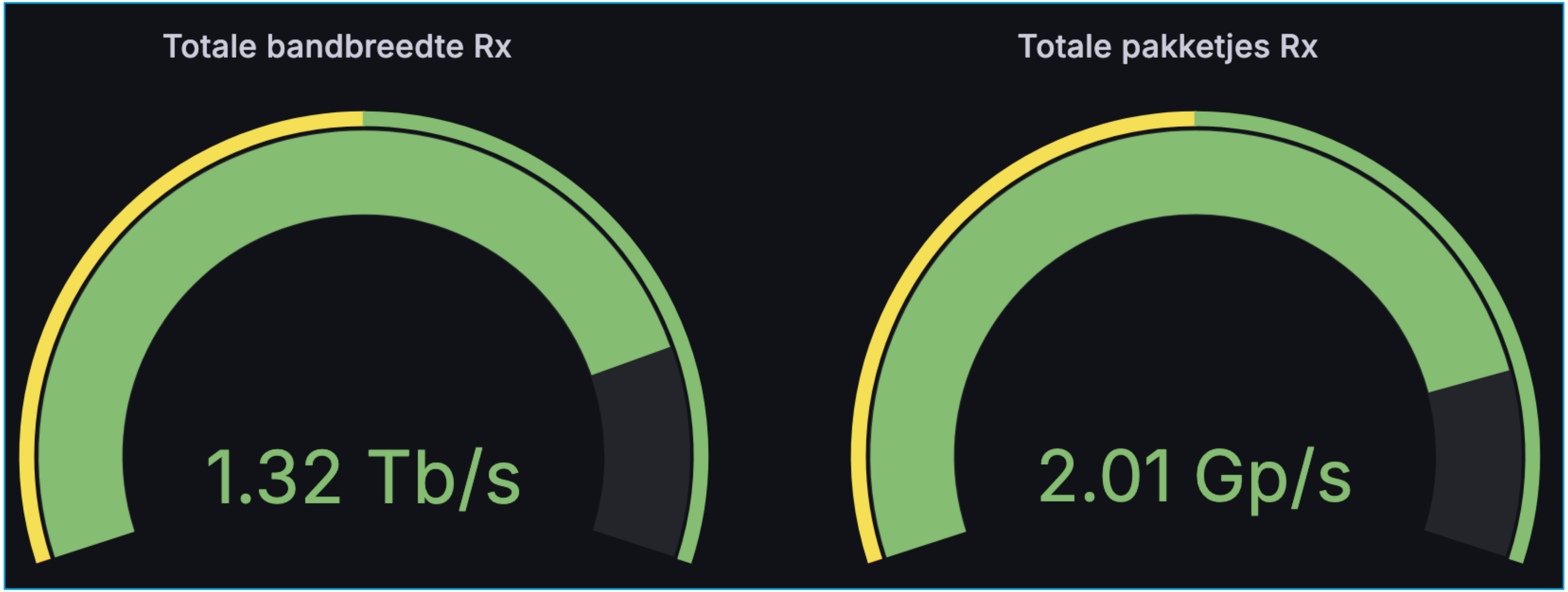
De voorbereidingen van de avond beginnen ver voordat de oefening gepland staat. Elke organisatie bepaalt welke systemen ze willen aanvallen en hoe de aanval uitgevoerd wordt. Het 'red team' is verantwoordelijk voor de aanvallen, het 'blue team voor de verdediging. Eén van de partijen die avond is Nikhef. Tristan, IT architect bij Nikhef, geeft aan dat zij dit belangeloos doen, gedreven door een maatschappelijke motivatie.

Nikhef is het Nationaal instituut voor subatomaire fysica in Nederland. Het beschikt over een gigantische bandbreedte, wat noodzakelijk is voor een dergelijke oefening waarbij zeer veel data wordt verstuurd. Zij zijn onderdeel van de aanvallende teams en

werkentegennederland.nl

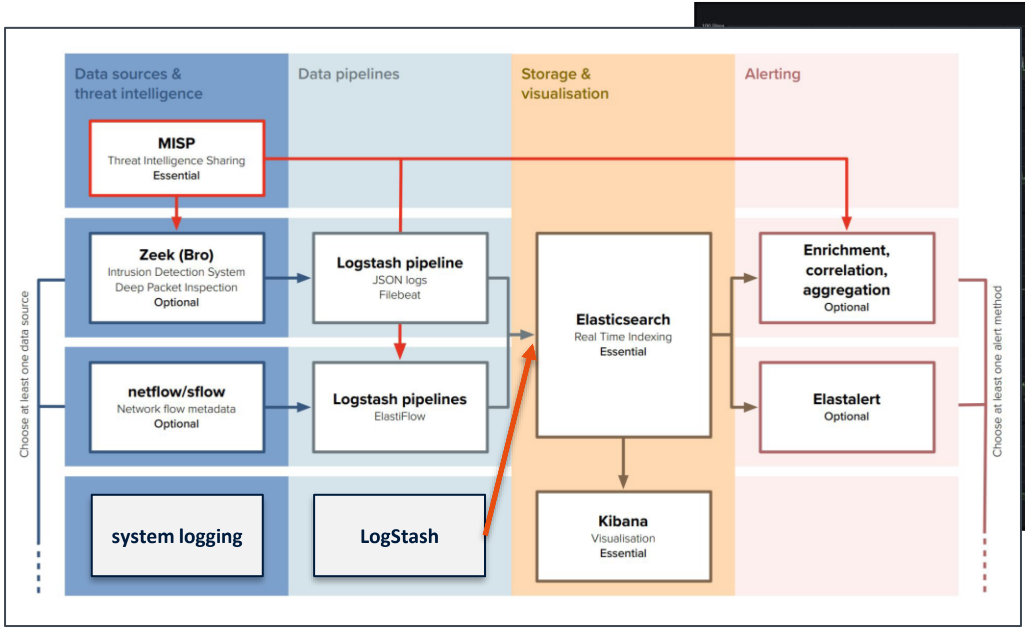
Image sources: belastingdienst.nl, rws.nl, nu.nl

with packets being more destructive than bandwidth …



https://wiki.nikhef.nl/grid/2Bpps_Machine - in preparation for the 2025 Resilience Exercise -

And ‘open’ does not mean ‘insecure’ – the WLCG SOC model



Data sources & threat intelligence

Data pipelines

Storage & visualisation

Alerting

MISP
Threat Intelligence Sharing
Essential

Zeek (Bro)
Intrusion Detection System
Deep Packet Inspection
Optional

netflow/sflow
Network flow metadata
Optional

system logging

Logstash pipeline
JSON logs
Filebeat

Logstash pipelines
ElastiFlow

LogStash

Elasticsearch
Real Time Indexing
Essential

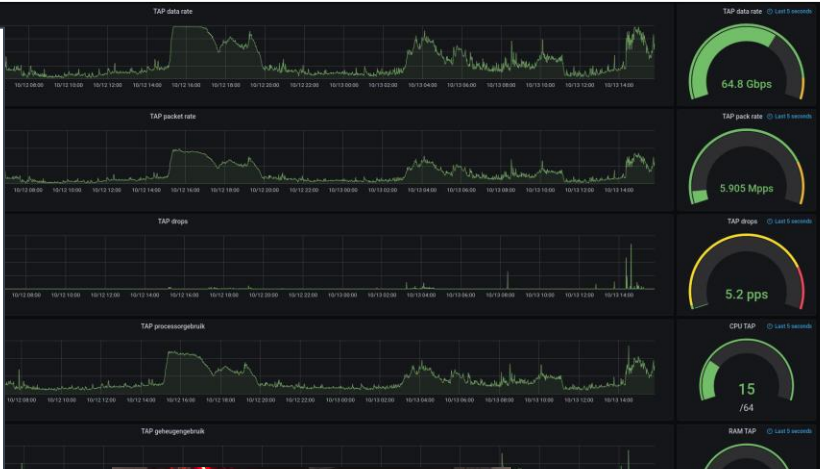
Kibana
Visualisation
Essential

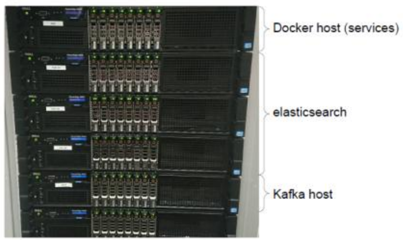
Enrichment, correlation, aggregation
Optional

Elastalert
Optional

Choose at least one data source

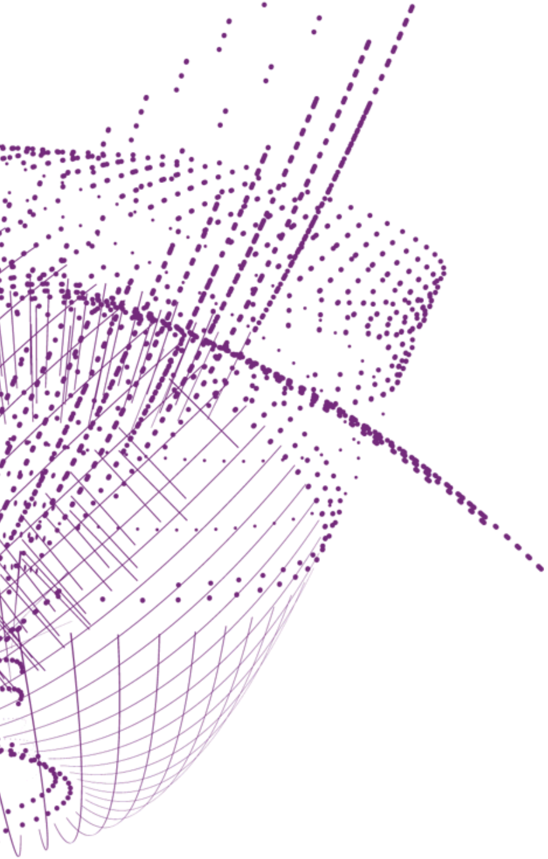
Choose at least one alert method





650 GByte/day ingest; 100Gbps+ monitoring through optical taps and mirroring; MISP intel from CERN, SURF, and private intel sources

Nikhef SOC design/management by Jouke Roorda. WLCG SOC WG model: Liviu Valsan (CERN) and David Crooks (STFC RAL)



In Infrastructure We Trust

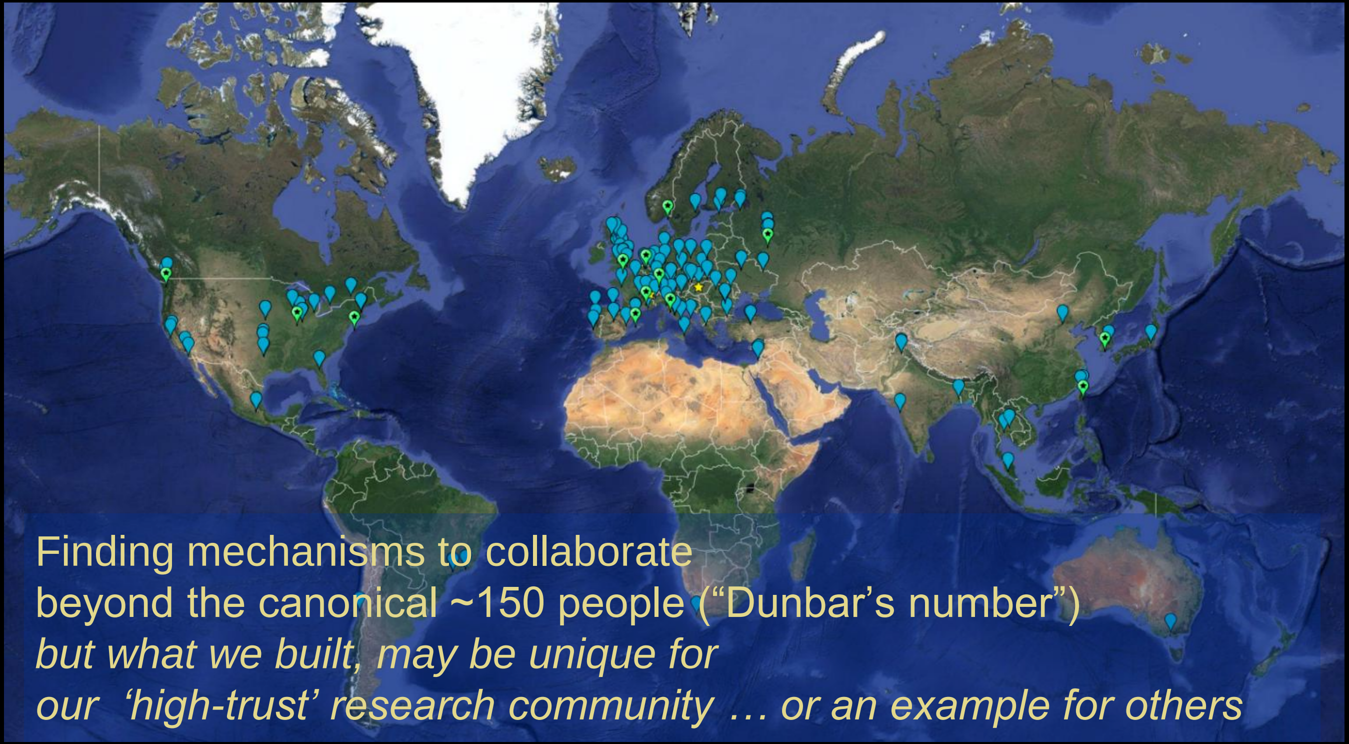
Collaborations: from small …

Nikhef user room H1.37 – terminal stations in the early 1990’s – image source: Nikhef

… to large collaborations (and shown here is a subset …)

a small part of the CMS collaboration in 2017, photo credit CERN on behalf the CMS collaboration, CMS-PHO-PUBLIC-2017-004-3

How many interactions? And just how many logins?



Finding mechanisms to collaborate
beyond the canonical ~150 people (“Dunbar’s number”)
but what we built, may be unique for
our ‘high-trust’ research community … or an example for others

Worldwide LHC Computing Grid (~ 2024)
~ 1.4 million CPU cores
~ 1500 Petabyte
disk + archival

170+ institutes
42+ countries
13 ‘Tier-1 sites’
some multi-community:
NL-T1 @ SURF & Nikhef

Earth background: Google Earth; Data and compute animation: STFC RAL for WLCG and EGI.eu; Data: https://home.cern/science/computing/grid ;
LHC Computing Grid: wlcg.web.cern.ch, EGI: www.egi.eu; ACCESS CI: https://access-ci.org/, NL-T1 and FuSE: fuse-infra.nl, https://www.surf.nl/en/research-it

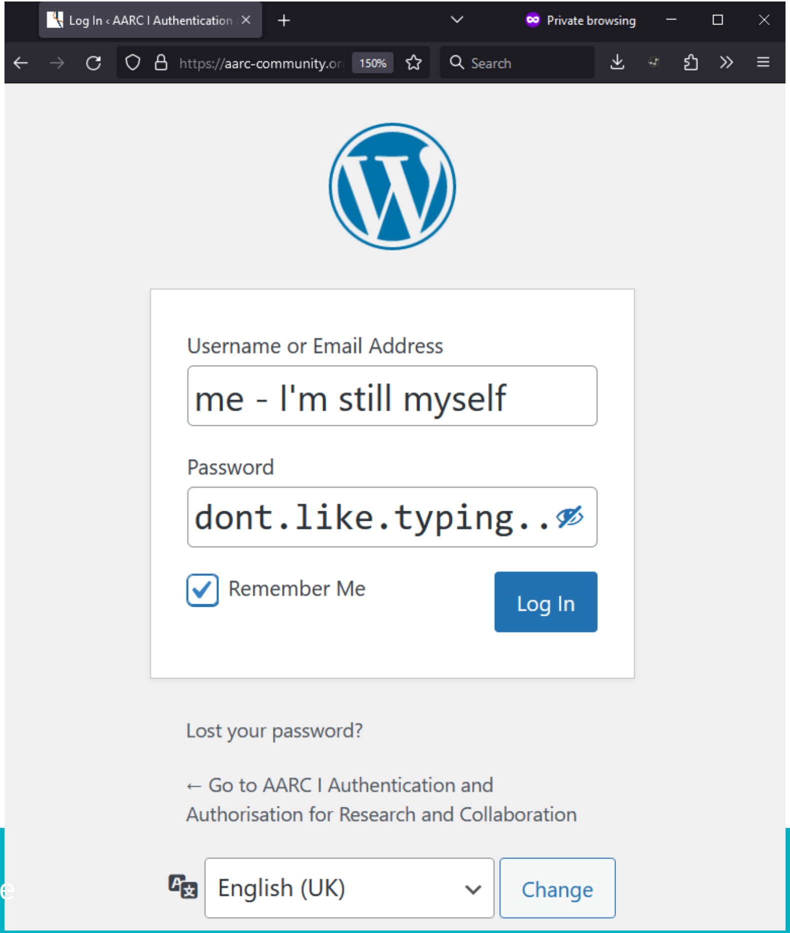
When you are asked to login again … 12 000 x 170+ times?

Authentication

demonstrating ‘you are you’

- ***authenticator***
  ‘you’ remains same ‘you’

- **vetted *identity***
  ‘you’ can be pseudonymous
  ‘you’ can be a vetted person



Username or Email Address

me - I'm still myself

Password

dont.like.typing..

☑ Remember Me

Log In

Lost your password?

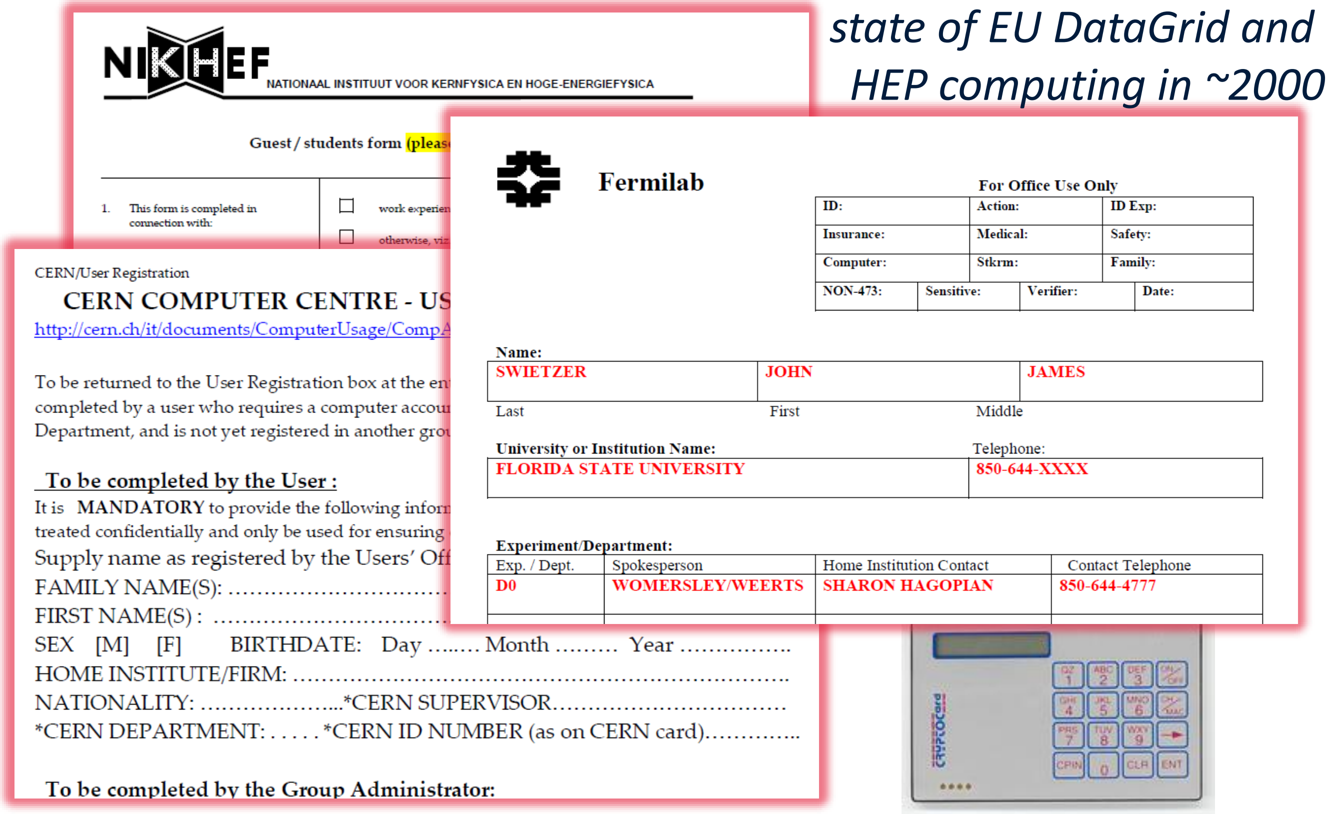
← Go to AARC I Authentication and Authorisation for Research and Collaboration

English (UK)

Change

Self-asserted or ‘pseudonymous’ often not enough

state of EU DataGrid and HEP computing in ~2000



NIKHEF NATIONAAL INSTITUUT VOOR KERNFYSICA EN HOGE-ENERGIEFYSICA

Guest / students form (pleas

1. This form is completed in connection with: □ work experien
□ otherwise, viz

CERN/User Registration

CERN COMPUTER CENTRE - US

http://cern.ch/it/documents/ComputerUsage/CompA

To be returned to the User Registration box at the en completed by a user who requires a computer accou Department, and is not yet registered in another grou

To be completed by the User :

It is MANDATORY to provide the following infor treated confidentially and only be used for ensuring

Supply name as registered by the Users’ Off

FAMILY NAME(S): ....

FIRST NAME(S) : ....

SEX [M] [F] BIRTHDATE: Day .... Month .... Year ....

HOME INSTITUTE/FIRM: ....

NATIONALITY: ....*CERN SUPERVISOR...............................

*CERN DEPARTMENT: . . . . . *CERN ID NUMBER (as on CERN card)............

To be completed by the Group Administrator:

Fermilab

For Office Use Only

ID:	**Action:**	**ID Exp:**	
Insurance:	**Medical:**	**Safety:**	
Computer:	**Stkrm:**	**Family:**	
NON-473:	**Sensitive:**	**Verifier:**	**Date:**

Name:

SWIETZER	JOHN	JAMES
Last	First	Middle

University or Institution Name:	Telephone:
FLORIDA STATE UNIVERSITY	850-644-XXXX

Experiment/Department:

Exp. / Dept.	Spokesperson	Home Institution Contact	Contact Telephone
D0	WOMERSLEY/WEERTS	SHARON HAGOPIAN	850-644-4777

Scaling credentials: per service per user

Many start with *credentials* dedicated
to each service where you need access

- In a multi-organizational system becomes

$$\mathcal{O}(n_{services}) * \mathcal{O}(n_{users})$$

- usually creates a strong link to authorization:

  different accounts for different roles,
  multiplying the number of credentials per user

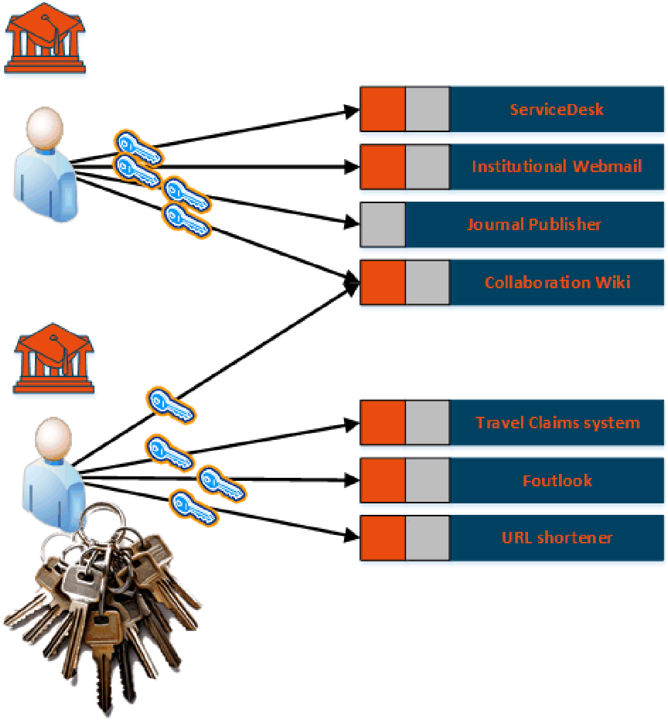


Image imspired by AARC NA2 training module “Authentication and Authorisation 101” – keychain image created by generative AI

bilateral ‘SSO’: a single service, or a single identity source

#credentials required?

from previously

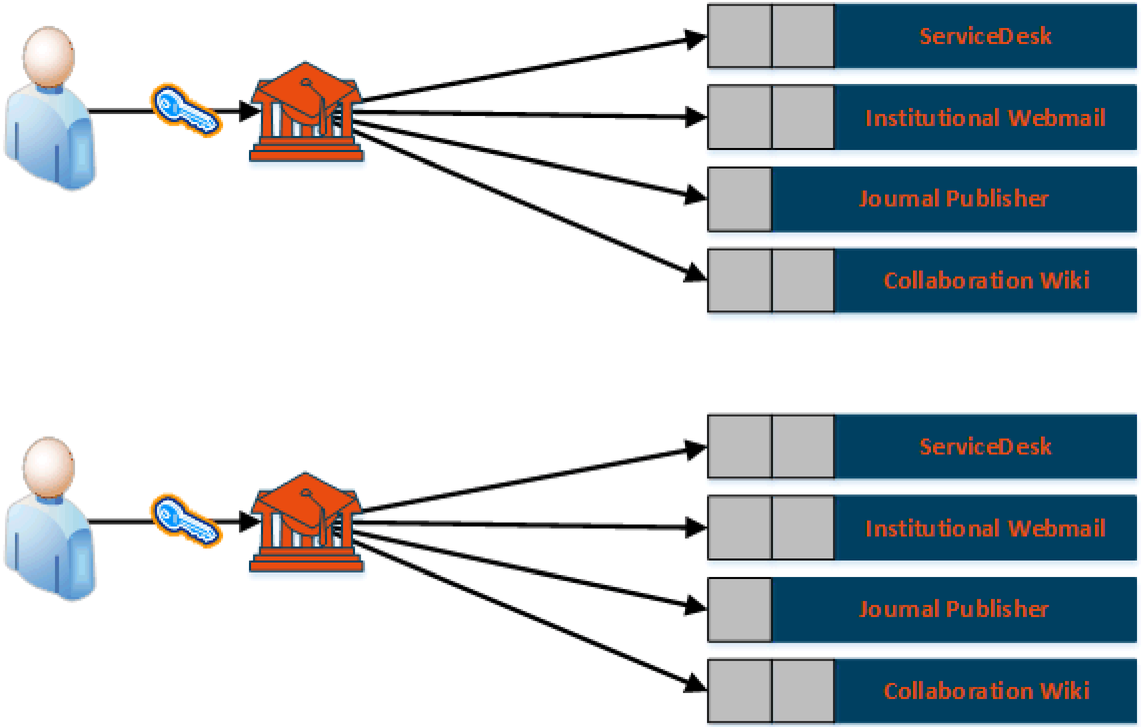
$\mathcal{O}(n_{services}) * \mathcal{O}(n_{users})$

to

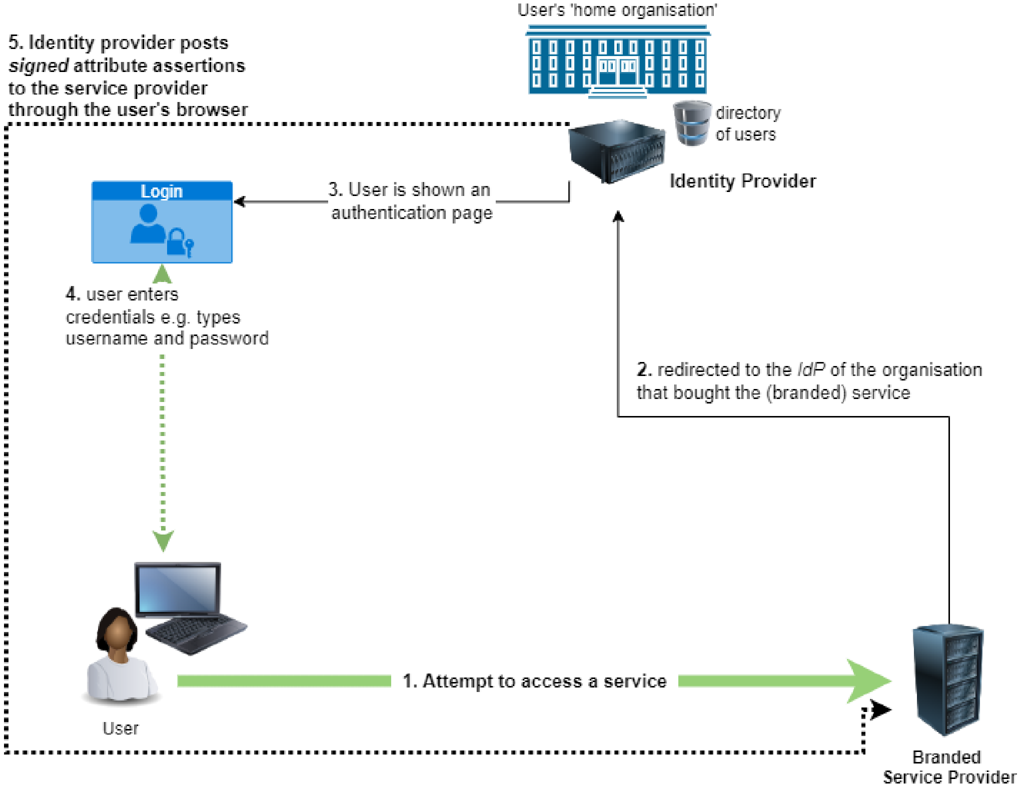
$\boldsymbol{\mathcal{O}(n_{users})}$

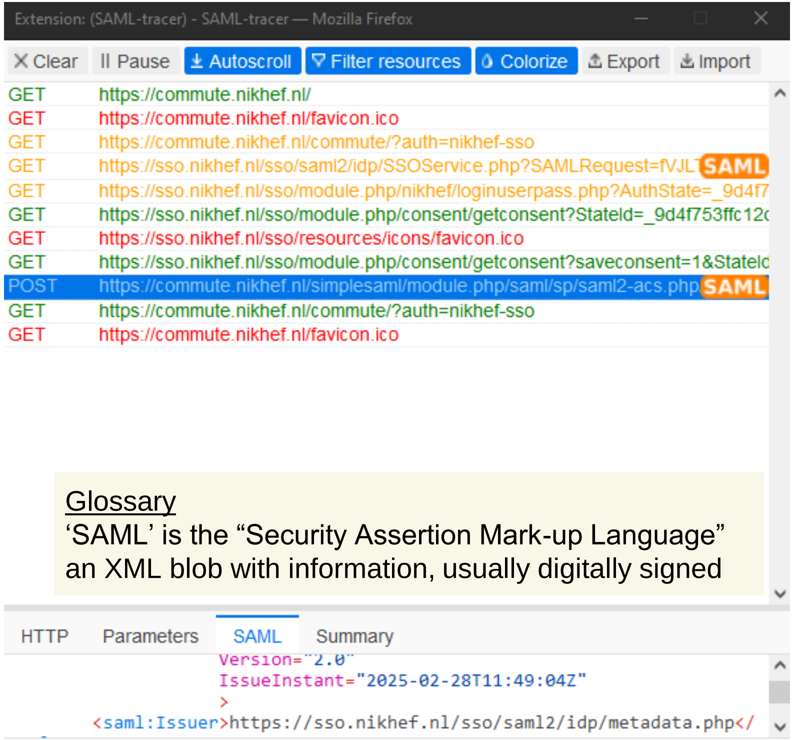
$+ \mathcal{O}(n_{services} * n_{home\text{-}orgs})$

in first order at least



Single sign-on – why your browser keeps loading things


User's 'home organisation'
5. Identity provider posts *signed* attribute assertions to the service provider through the user's browser
directory of users
Identity Provider
Login
3. User is shown an authentication page
4. user enters credentials e.g. types username and password
2. redirected to the *IdP* of the organisation that bought the (branded) service
1. Attempt to access a service
User
Branded Service Provider



```
GET     https://commute.nikhef.nl/
GET     https://commute.nikhef.nl/favicon.ico
GET     https://commute.nikhef.nl/commute/?auth=nikhef-sso
GET     https://sso.nikhef.nl/sso/saml2/idp/SSOService.php?SAMLRequest=fVJLT
GET     https://sso.nikhef.nl/sso/module.php/nikhef/loginuserpass.php?AuthState=_9d4f7
GET     https://sso.nikhef.nl/sso/module.php/consent/getconsent?StateId=_9d4f753ffc12c
GET     https://sso.nikhef.nl/sso/resources/icons/favicon.ico
GET     https://sso.nikhef.nl/sso/module.php/consent/getconsent?saveconsent=1&StateId
POST    https://commute.nikhef.nl/simplesaml/module.php/saml/sp/saml2-acs.php
GET     https://commute.nikhef.nl/commute/?auth=nikhef-sso
GET     https://commute.nikhef.nl/favicon.ico
```

```
                   Version="2.0"
                   IssueInstant="2025-02-28T11:49:04Z"
                   >
    <saml:Issuer>https://sso.nikhef.nl/sso/saml2/idp/metadata.php</
```

Glossary
'SAML' is the "Security Assertion Mark-up Language"
an XML blob with information, usually digitally signed

SAML-tracer plugin by Tim van Dijen (SSC-ICT) *et al.*
https://github.com/simplesamlphp/SAML-tracer

User-centric identity: ‘I take my passport anywhere by myself’

Your ‘home organisation’ does not have to be in the loop …

User's 'home organisation'
directory of users
Identity Provider

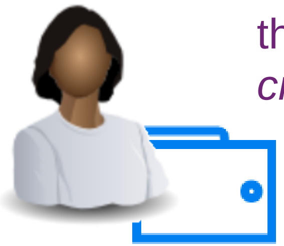
user-centric trust: you yourself hold a credential from a trusted third party and can use it *without having to ask ‘home’ each time:*

- Public Key Infrastructure client certificates (“X.509”)
- Verifiable credentials in wallets

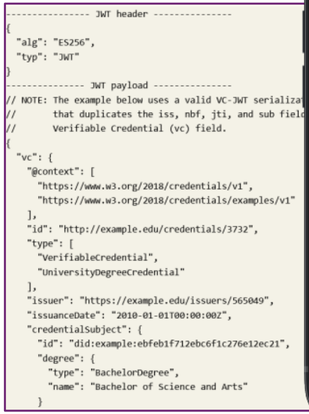
- *and who remembers CardSpace?*

Passport image: cropped from original by Jon Tyson on Unsplash https://unsplash.com/photos/Hid-yhommOg

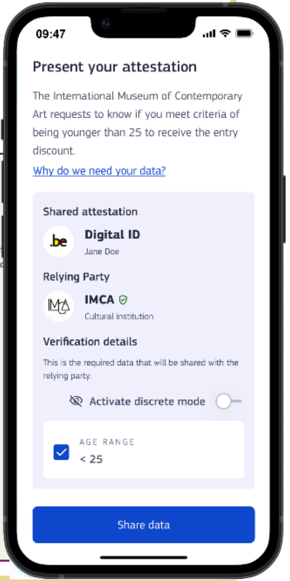
Identity wallets, held by the user, are another



the user as a
credential Holder



```
--------------- JWT header --------------
{
  "alg": "ES256",
  "typ": "JWT"
}
--------------- JWT payload --------------
// NOTE: The example below uses a valid VC-JWT serializa
//       that duplicates the iss, nbf, jti, and sub fielc
//       Verifiable Credential (vc) field.
{
  "vc": {
    "@context": [
      "https://www.w3.org/2018/credentials/v1",
      "https://www.w3.org/2018/credentials/examples/v1"
    ],
    "id": "http://example.edu/credentials/3732",
    "type": [
      "VerifiableCredential",
      "UniversityDegreeCredential"
    ],
    "issuer": "https://example.edu/issuers/565049",
    "issuanceDate": "2010-01-01T00:00:00Z",
    "credentialSubject": {
      "id": "did:example:ebfeb1f712ebc6f1c276e12ec21",
      "degree": {
        "type": "BachelorDegree",
        "name": "Bachelor of Science and Arts"
      }
```



Present your attestation

The International Museum of Contemporary Art requests to know if you meet criteria of being younger than 25 to receive the entry discount.

Why do we need your data?

Shared attestation

Digital ID
Jane Doe

Relying Party

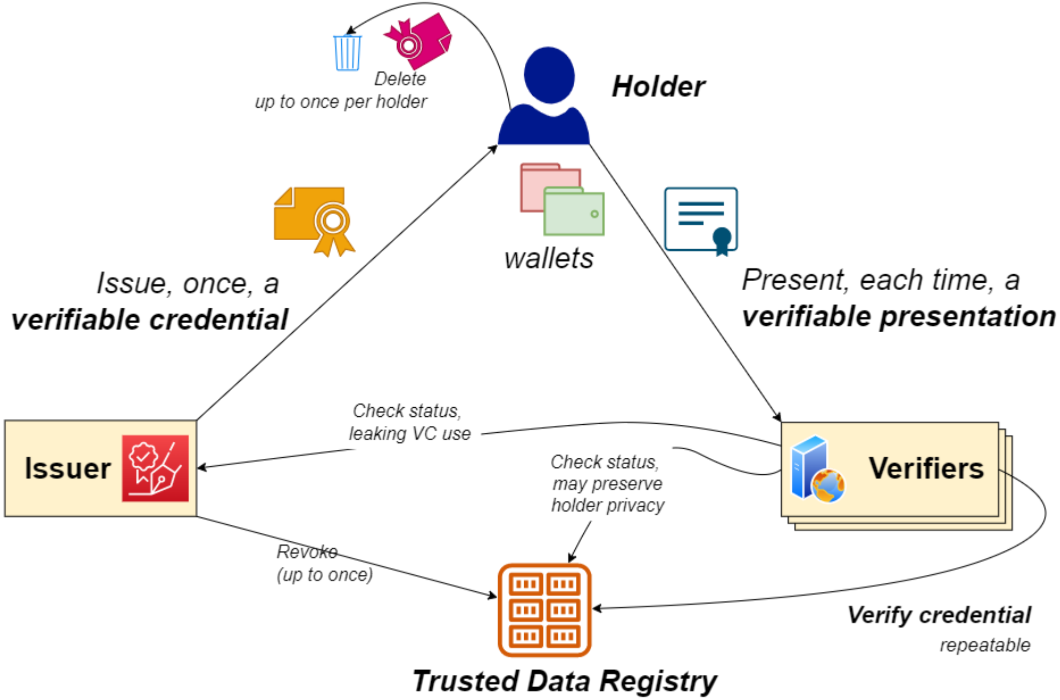
IMCA
Cultural institution

Verification details

This is the required data that will be shared with the relying party.

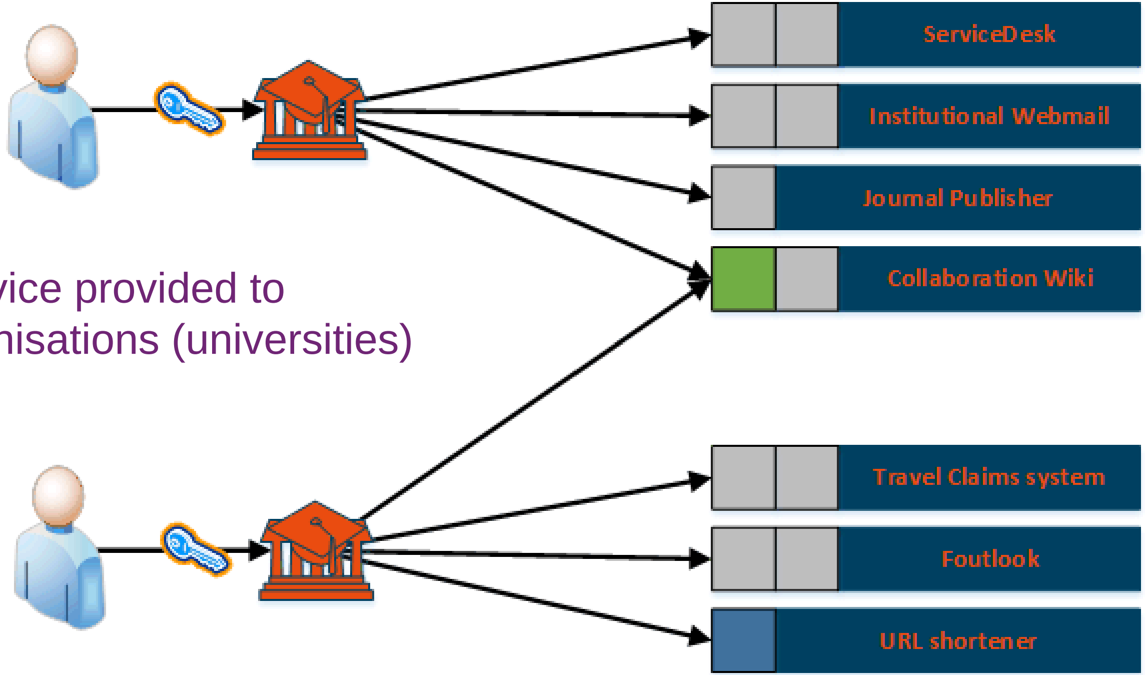
Activate discrete mode ○

☑ AGE RANGE
< 25



Flow diagram inspired by: Lifecycle Details (5.1), Verifiable Credentials Data Model v1.1, W3C Recommendation 03 March 2022, https://www.w3.org/TR/vc-data-model/
EU eID Wallet from https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-digital-identity_en
Appimage: European Commission, at https://ec.europa.eu/digital-building-blocks/sites/display/EUDIGITALIDENTITYWALLET/Security+and+Privacy

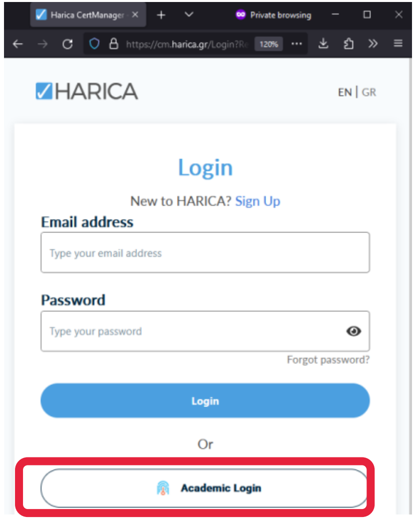
Can we scale better with an ‘federated’ Authentication and Authorisation Infrastructure (‘AAI’)



with one service provided to several organisations (universities)

we will get to authorisation in a bit …

Where are 'you' in the federated space – discovery!



Harica CertManager × | Private browsing
https://cm.harica.gr/Login?R | 120%

HARICA | EN | GR

Login

New to HARICA? Sign Up

Email address

Type your email address

Password

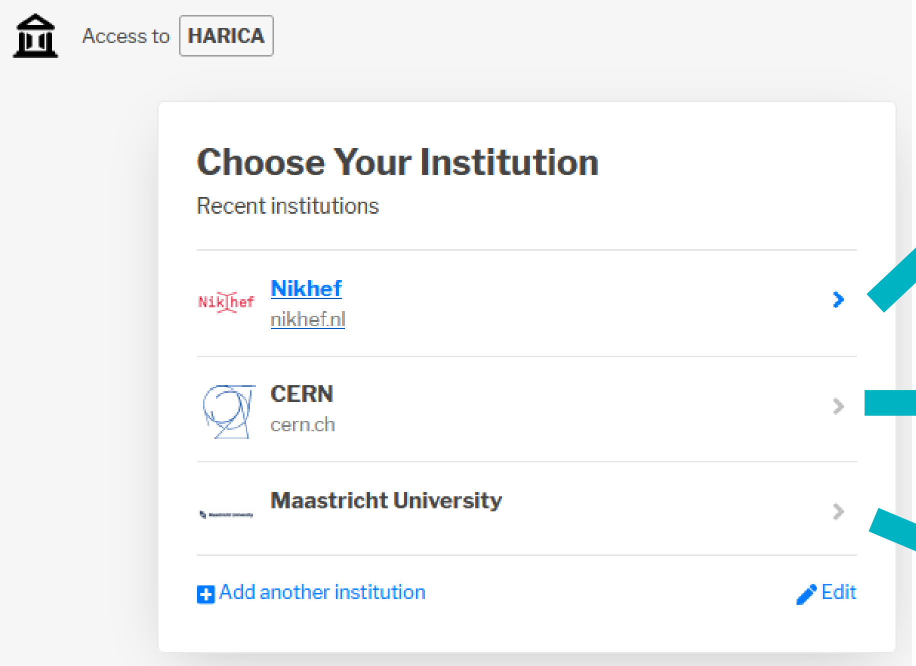
Type your password

Forgot password?

Login

Or

Academic Login



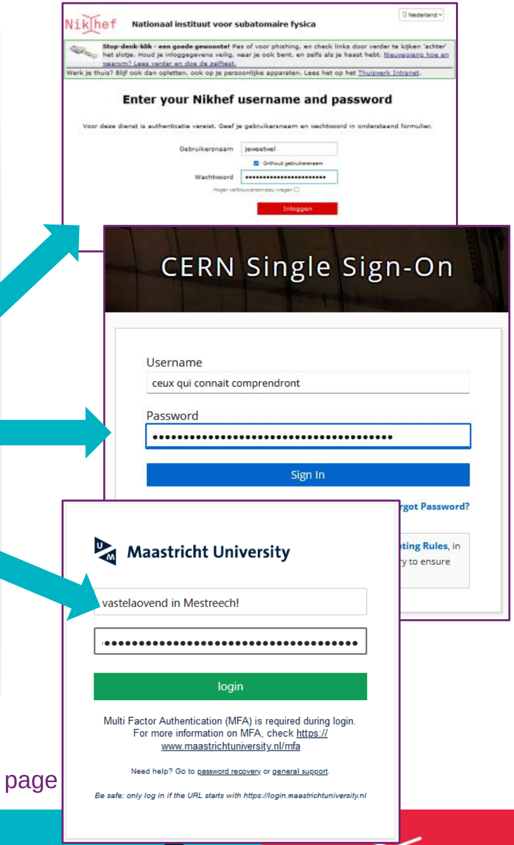
Access to **HARICA**

Choose Your Institution

Recent institutions

- **Nikhef** nikhef.nl
- **CERN** cern.ch
- **Maastricht University**

Add another institution | Edit



Nikhef | Nationaal instituut voor subatomaire fysica | Nederland

Stop-denk-klik - een goede gewoonte! Pas of voor phishing, en check links door verder te kijken 'achter' het slotje. Houd je inloggegevens veilig, waar je ook bent, en zelfs als je haast hebt. Nieuwsgierig hoe en waarom? Lees verder en doe de zelftest.

Werk je thuis? Blijf ook dan opletten, ook op je persoonlijke apparaten. Lees het op het Thuiswerk Intranet.

Enter your Nikhef username and password

Voor deze dienst is authenticatie vereist. Geef je gebruikersnaam en wachtwoord in onderstaand formulier.

Gebruikersnaam: jeweetwel

☑ Onthoud gebruikersnaam

Wachtwoord: ••••••••••••••••••••••

Hoger vertrouwensniveau vragen □

Inloggen

CERN Single Sign-On

Username: ceux qui connait comprendront

Password: ••

Sign In

rgot Password?

ting Rules, in

y to ensure

Maastricht University

vastelaovend in Mestreech!

••••••••••••••••••••••••••••••••••••

login

Multi Factor Authentication (MFA) is required during login. For more information on MFA, check https://www.maastrichtuniversity.nl/mfa

Need help? Go to password recovery or general support.

Be safe: only log in if the URL starts with https://login.maastrichtuniversity.nl

An example cross-institutional service by HARICA, the GEANT TCS G5 provider, presenting a SeamlessAccess.org discovery page

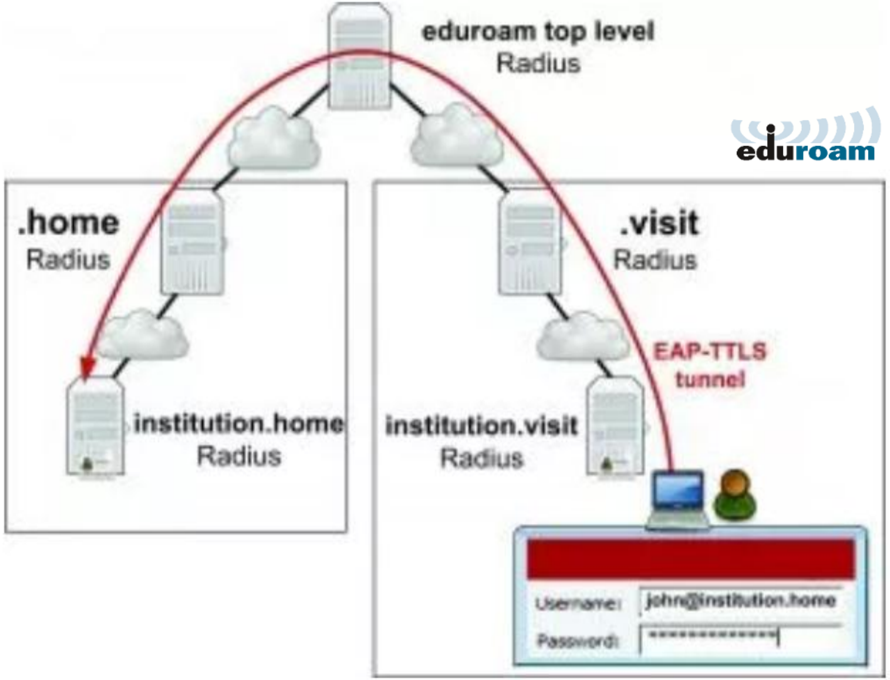
The R&E federation that was there first …

service-specific trust
between organisations

hierarchical server path, based on
a network-specific secure exchange

sending your credentials back
to *only* your home institution

found via <anon**@domain.name**>

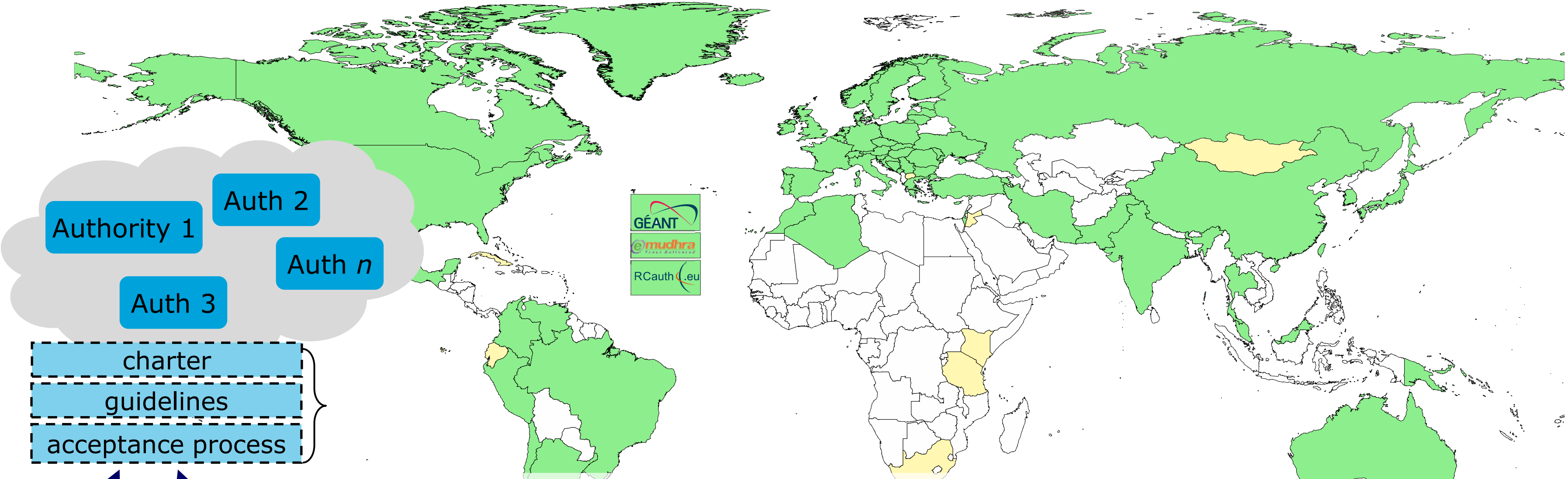


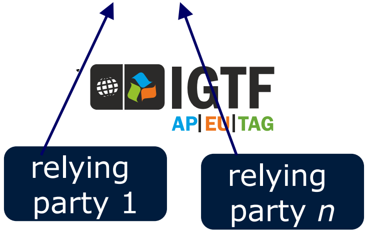
Username: john@institution.home

Password: *************

eduroam image from https://eduroam.org/how/, GEANT ; RADIUS: RC2865 https://www.rfc-editor.org/rfc/rfc2865; see also freeradius.org

IGTF: a policy-bridged global federation for research computing





A global authentication fabric & assurance standards

~ 90 Identity Providers (some leveraging a R&E federation)

~ 10 international research and e-infrastructure relying parties

> 60 countries / economic areas / international treaty orgs

> 1000 relying service provider collaborations

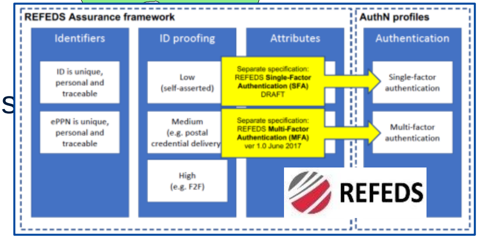
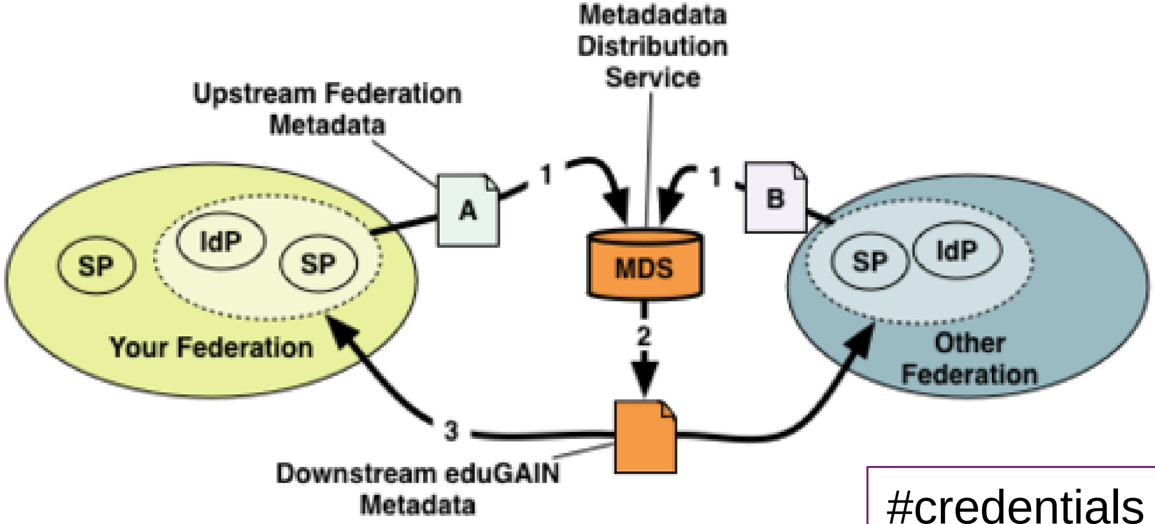

REFEDS Assurance framework
AuthN profiles
Identifiers
ID proofing
Attributes
Authentication
ID is unique, personal and traceable
ePPN is unique, personal and traceable
Low (self-asserted)
Medium (e.g. postal credential delivery)
High (e.g. F2F)
Separate specification: REFEDS Single-Factor Authentication (SFA) DRAFT
Separate specification: REFEDS Multi-Factor Authentication (MFA) ver 1.0 June 2017
Single-factor authentication
Multi-factor authentication
REFEDS

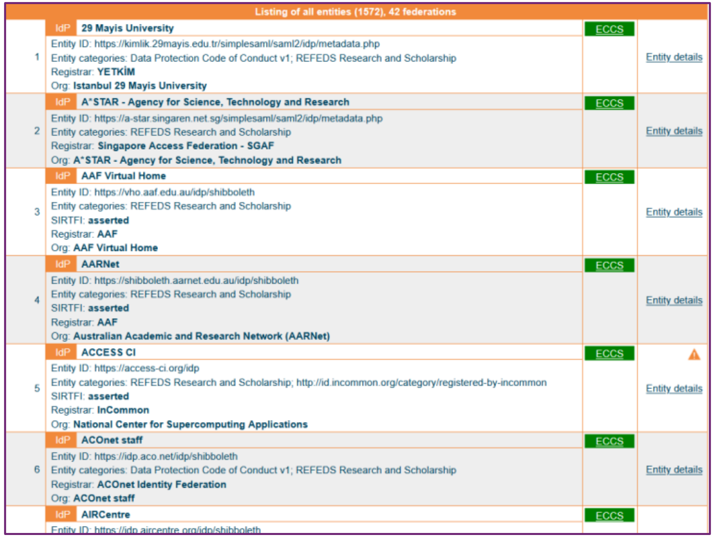
Image: Interoperable Global Trust Federation IGTF, https://igtf.net/; REFEDS Assurance Framework RAF: http://refeds.org/assurance, https://refeds.org/profile/mfa

We live in a federated world!

aconet
AUSTRALIAN ACCESS FEDERATION
eduID cz
RENATER CONNECTEUR DE SAVOIRS
GakuNin
cafe
DFN-AAI
SWITCHaai
The UK Access Management Federation FOR EDUCATION AND RESEARCH
Edugate Gaining Access to Education
SIR
haka
grnet
SURF CONEXT
RCTSaai REDE CIÊNCIA, TECNOLOGIA E SOCIEDADE
MiNGA
idem garr aai
SIFULAN
Colfire
Carsi
Belnet R&E Federation
FEIDE
AFIRE
GakuNin
InCommon
COFRe Comunidad Federada REUNA
IGTF Interoperable Global Trust Federation AP EU TAG
LIAFE
CILogon
TAAT
eduID
Tuakiri NEW ZEALAND ACCESS FEDERATION
AAI @EduHr
arnes
WAYF Where are you from

Meta-data and trust in IdP-SP ‘multi-lateral’ federations





Listing of all entities (1572), 42 federations

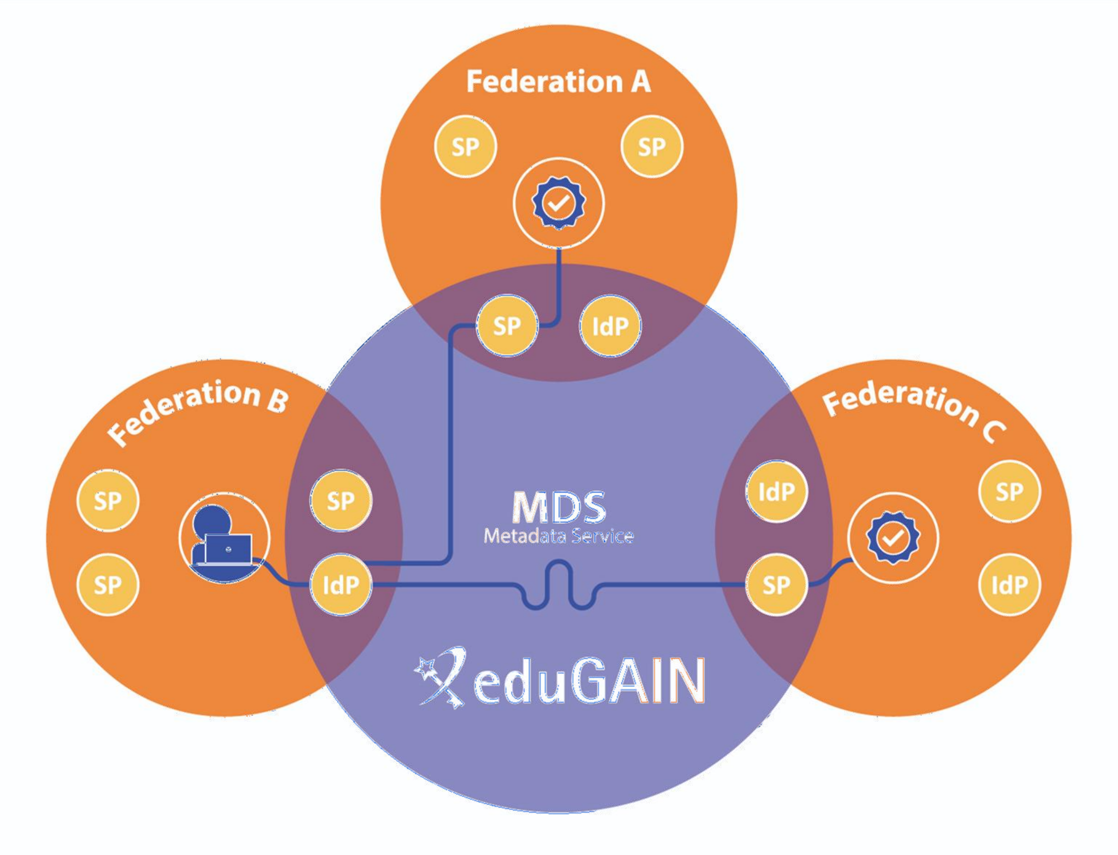
1	IdP	**29 Mayis University** Entity ID: https://kimlik.29mayis.edu.tr/simplesaml/saml2/idp/metadata.php Entity categories: Data Protection Code of Conduct v1; REFEDS Research and Scholarship Registrar: **YETKİM** Org: **Istanbul 29 Mayis University**	ECCS	Entity details
2	IdP	**A*STAR - Agency for Science, Technology and Research** Entity ID: https://a-star.singaren.net.sg/simplesaml/saml2/idp/metadata.php Entity categories: REFEDS Research and Scholarship Registrar: **Singapore Access Federation - SGAF** Org: **A*STAR - Agency for Science, Technology and Research**	ECCS	Entity details
3	IdP	**AAF Virtual Home** Entity ID: https://vho.aaf.edu.au/idp/shibboleth Entity categories: REFEDS Research and Scholarship SIRTFI: **asserted** Registrar: **AAF** Org: **AAF Virtual Home**	ECCS	Entity details
4	IdP	**AARNet** Entity ID: https://shibboleth.aarnet.edu.au/idp/shibboleth Entity categories: REFEDS Research and Scholarship SIRTFI: **asserted** Registrar: **AAF** Org: **Australian Academic and Research Network (AARNet)**	ECCS	Entity details
5	IdP	**ACCESS CI** Entity ID: https://access-ci.org/idp Entity categories: REFEDS Research and Scholarship; http://id.incommon.org/category/registered-by-incommon SIRTFI: **asserted** Registrar: **InCommon** Org: **National Center for Supercomputing Applications**	ECCS	⚠ Entity details
6	IdP	**ACOnet staff** Entity ID: https://idp.aco.net/idp/shibboleth Entity categories: Data Protection Code of Conduct v1; REFEDS Research and Scholarship Registrar: **ACOnet Identity Federation** Org: **ACOnet staff**	ECCS	
	IdP	**AIRCentre** Entity ID: https://idp.aircentre.org/idp/shibboleth	ECCS	

#credentials required?

from $\mathcal{O}(n_{users}) + \mathcal{O}(n_{services} * n_{home\text{-}orgs})$

to $\sim \boldsymbol{\mathcal{O}(n_{users}) + \mathcal{O}(n_{home\text{-}orgs}) + \mathcal{O}(n_{services})}$

MDS meta-data flow: https://wiki.geant.org/display/eduGAIN/Metadata+Flow+in+eduGAIN
eduGAIN meta-data https://mds.edugain.org/edugain-v2.xml ; table excerpt from https://technical.edugain.org/entities showing only R&S IdPs, i.e. those supporting research …



78

Identity Federations

5100+

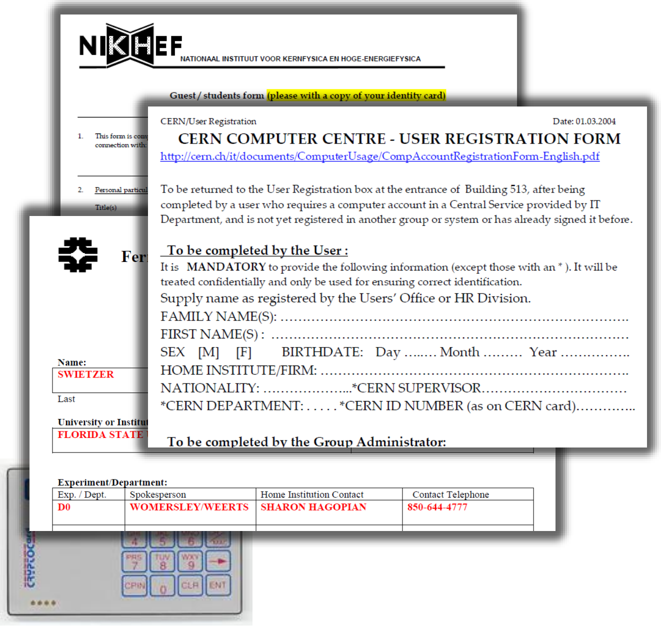
Identity Providers

3600+

Service Providers

eduGAIN image: Davide Vaghetti, GARR for GN*-*

We progressed a lot since 2003 with identity federation



NIKHEF NATIONAAL INSTITUUT VOOR KERNFYSICA EN HOGE-ENERGIEFYSICA

Guest / students form (please with a copy of your identity card)

1. This form is com
connection with:

2. Personal particul

Title(s)

Fer

Name:
SWIETZER
Last

University or Institut
FLORIDA STATE

Experiment/Department:

Exp. / Dept.	Spokesperson	Home Institution Contact	Contact Telephone
D0	WOMERSLEY/WEERTS	SHARON HAGOPIAN	850-644-4777

CERN/User Registration Date: 01.03.2004

CERN COMPUTER CENTRE - USER REGISTRATION FORM

http://cern.ch/it/documents/ComputerUsage/CompAccountRegistrationForm-English.pdf

To be returned to the User Registration box at the entrance of Building 513, after being completed by a user who requires a computer account in a Central Service provided by IT Department, and is not yet registered in another group or system or has already signed it before.

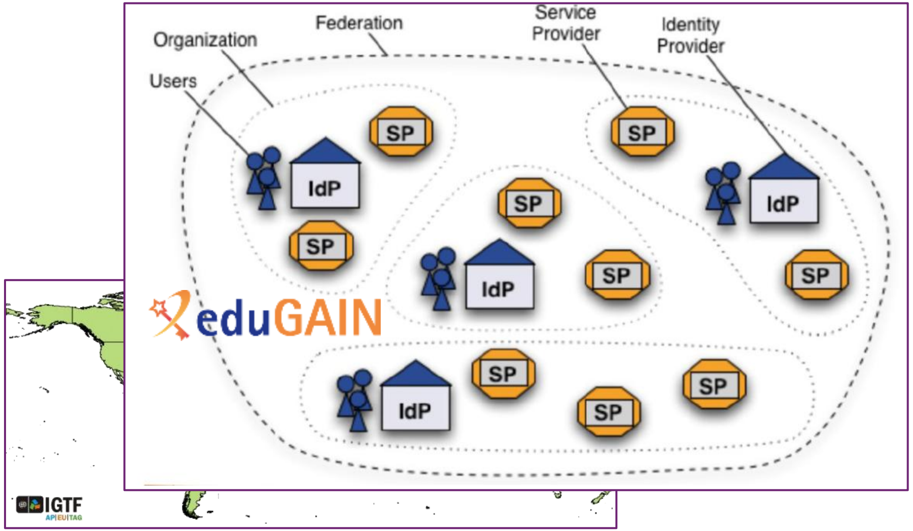
To be completed by the User :

It is MANDATORY to provide the following information (except those with an *). It will be treated confidentially and only be used for ensuring correct identification.

Supply name as registered by the Users' Office or HR Division.

FAMILY NAME(S): ………………………………………………………………………
FIRST NAME(S) : ………………………………………………………………………
SEX [M] [F] BIRTHDATE: Day …..…. Month ……… Year …………….
HOME INSTITUTE/FIRM: ………………………………………………………………
NATIONALITY: ………………...*CERN SUPERVISOR……………………………
*CERN DEPARTMENT: . . . . . *CERN ID NUMBER (as on CERN card)…………..

To be completed by the Group Administrator:



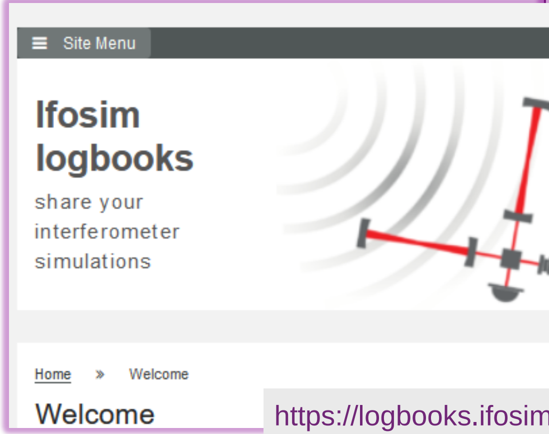
For **eduGAIN federation** the IdPs provide **authentication** from the home organisation, for the user-centric PKIX IGTF trust fabric, the CAs do.
Then **Service providers** perform **authorization**,
… maybe using attributes provided by the IdP. But do they get them??

Right-hand image: Shibboleth IdP federation, Lukas Hammerle, SWITCH (CH), user-centric PKI credentials: Interoperable Global Trust Federation, https://igtf.net/

Federated Success!

Login to GW's ifosim.org, to gitlab, or … via the service proxy

with any eduGAIN IdP for user authentication



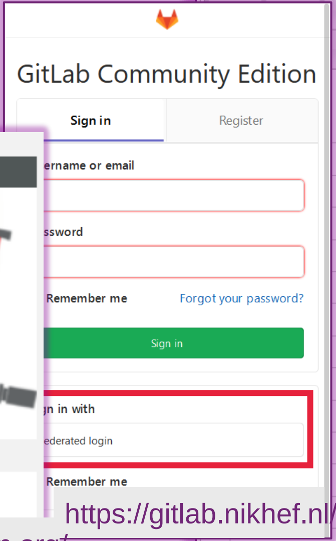
Ifosim logbooks

share your interferometer simulations

Home » Welcome

Welcome

https://logbooks.ifosim.org/



GitLab Community Edition

Sign in | Register

ername or email

ssword

Remember me Forgot your password?

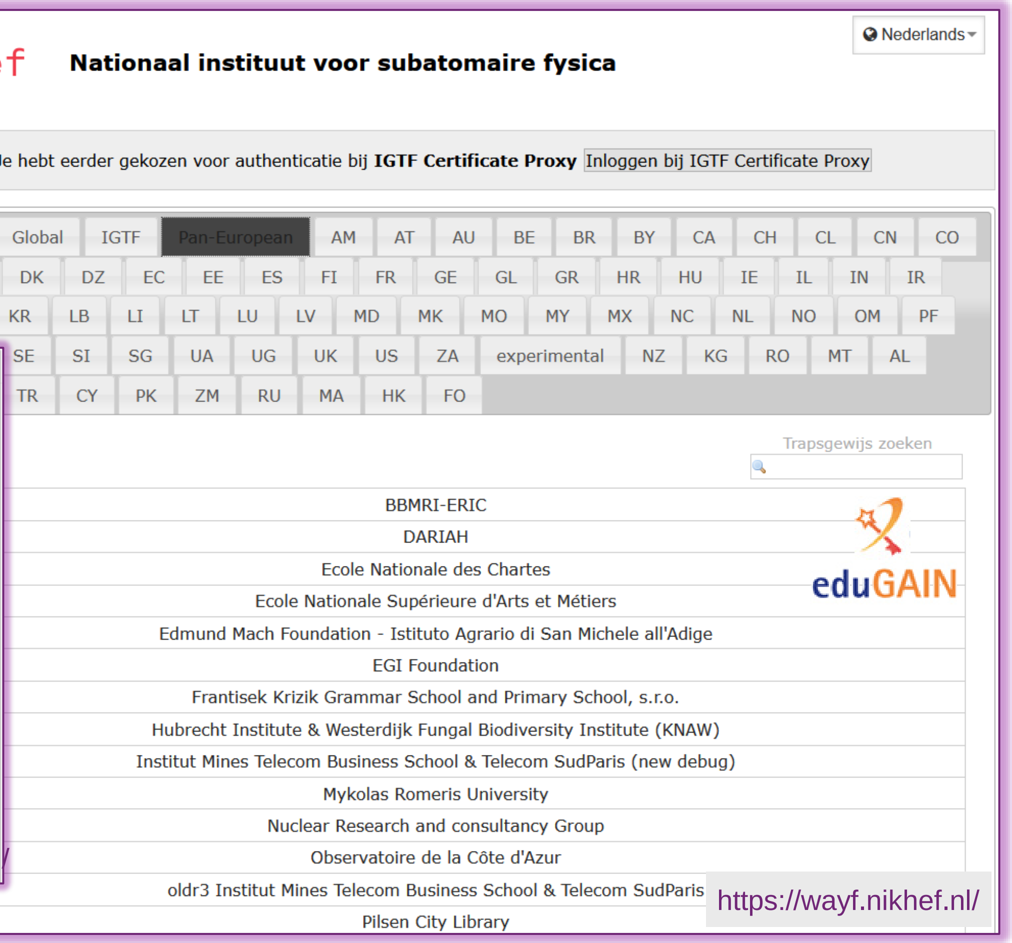
Sign in

n in with

ederated login

Remember me

https://gitlab.nikhef.nl/



Nikhef Nationaal instituut voor subatomaire fysica

Je hebt eerder gekozen voor authenticatie bij **IGTF Certificate Proxy** Inloggen bij IGTF Certificate Proxy

Preferred Global IGTF Pan-European AM AT AU BE BR BY CA CH CL CN CO
CZ DE DK DZ EC EE ES FI FR GE GL GR HR HU IE IL IN IR
IT JP KR LB LI LT LU LV MD MK MO MY MX NC NL NO OM PF
SE SI SG UA UG UK US ZA experimental NZ KG RO MT AL
TR CY PK ZM RU MA HK FO

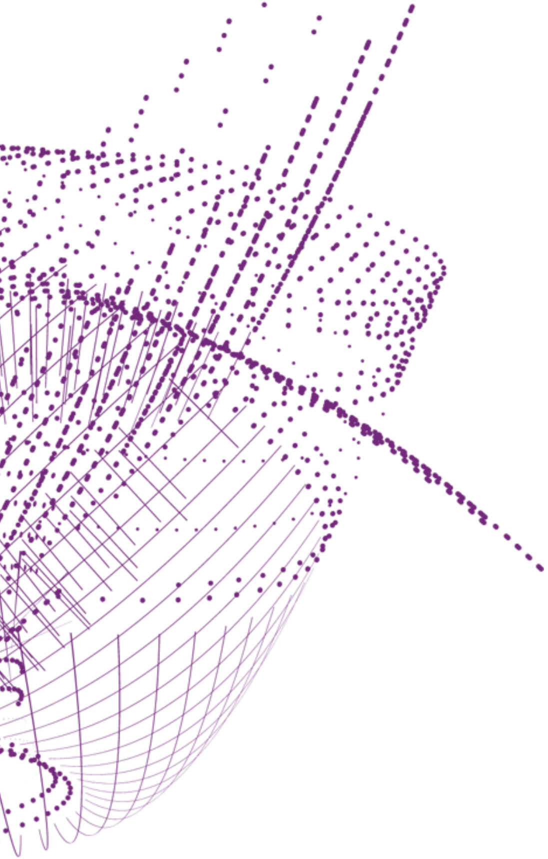
Trapsgewijs zoeken

BBMRI-ERIC
DARIAH
Ecole Nationale des Chartes
Ecole Nationale Supérieure d'Arts et Métiers
Edmund Mach Foundation - Istituto Agrario di San Michele all'Adige
EGI Foundation
Frantisek Krizik Grammar School and Primary School, s.r.o.
Hubrecht Institute & Westerdijk Fungal Biodiversity Institute (KNAW)
Institut Mines Telecom Business School & Telecom SudParis (new debug)
Mykolas Romeris University
Nuclear Research and consultancy Group
Observatoire de la Côte d'Azur
oldr3 Institut Mines Telecom Business School & Telecom SudParis
Pilsen City Library

eduGAIN

https://wayf.nikhef.nl/

ifosim federated AAI integration implementation by Mischa Sallé; per-country



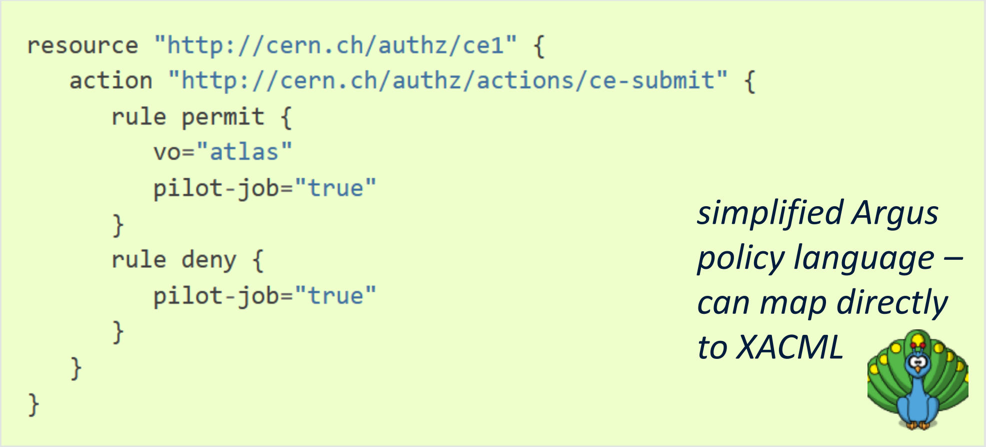
Federation works quite well for Authentication, but … not (federated) Authorisation – the important element for collaborative (research); with different complementary sources of authority, and decision power at the RP and its coordinating (e)-infrastructure

The Forgotten A in AAI

Authorization – what you are allowed to do

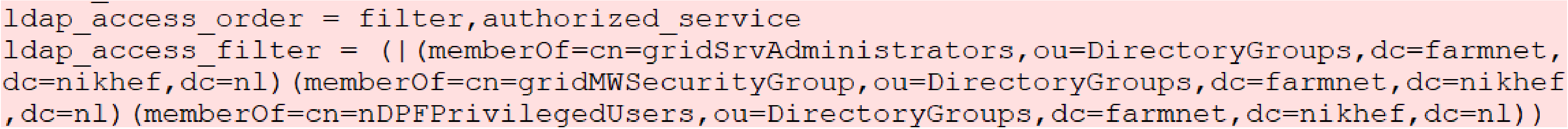
soon needs specifying **access rights** to resources, based on an access **policy**

- might be implicit or ad-hoc
- be in formal policy language
  like XACML (*example: Argus PDP)*
- or be service-specific
  example: Linux sssd config



```
resource "http://cern.ch/authz/ce1" {
   action "http://cern.ch/authz/actions/ce-submit" {
      rule permit {
         vo="atlas"
         pilot-job="true"
      }
      rule deny {
         pilot-job="true"
      }
   }
}
```

simplified Argus policy language – can map directly to XACML



```
ldap_access_order = filter,authorized_service
ldap_access_filter = (|(memberOf=cn=gridSrvAdministrators,ou=DirectoryGroups,dc=farmnet,
dc=nikhef,dc=nl)(memberOf=cn=gridMWSecurityGroup,ou=DirectoryGroups,dc=farmnet,dc=nikhef
,dc=nl)(memberOf=cn=nDPFPrivilegedUsers,ou=DirectoryGroups,dc=farmnet,dc=nikhef,dc=nl))
```

Policy example: Argus system, https://argus-documentation.readthedocs.io/en/stable/misc/examples.html; service-specific: sssd.conf ldap auth_provider

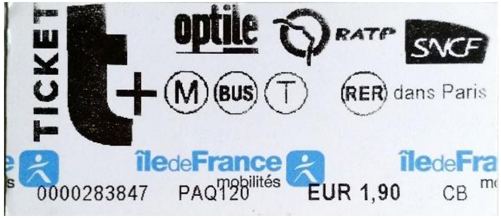
Authorization policy subjects

AuthZ policies need subject attributes (‘claims’)

- **bound to an verifiable identity** statement
  - e.g. visa are strongly linked to a specific entity, and asserted by a trusted party (by the service)
- be a **bearer token**
  - scoped to a relying party, a service, or an action
- **self-asserted**
  - quite useless unless backed by verifiable evidence, like in self-sovereign identity schemes

Transport mechanisms (see also RFC2903)

- pushed alongside the service access,
- pulled from the source as needed, or
- pushed by the attribute source as an agent

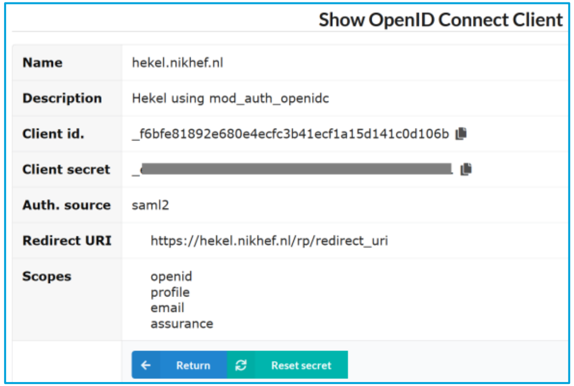


USA visa image source: https://2009-2017.state.gov/m/ds/rls/rpt/79785.htm ; RATP bearer token, issued for the Paris public transport system

OpenID Connect and OAuth2: the ‘modern’ way

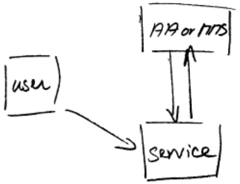
- Quite .well-known
  (used by lots modern ‘non-enterprise’ SSO)

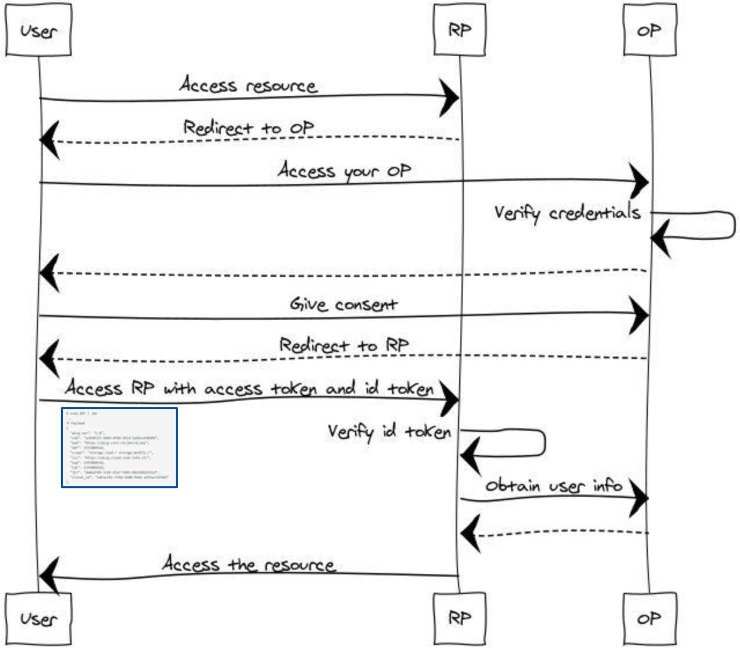
- shows signs of its initial design objective:
  one source of identity (Openid Provider, ‘OP’),
  and *many* services (Relaying Parties, ‘RP’)



Show OpenID Connect Client

Name	hekel.nikhef.nl
Description	Hekel using mod_auth_openidc
Client id.	_f6bfe81892e680e4ecfc3b41ecf1a15d141c0d106b
Client secret	_[redacted]
Auth. source	saml2
Redirect URI	https://hekel.nikhef.nl/rp/redirect_uri
Scopes	openid profile email assurance



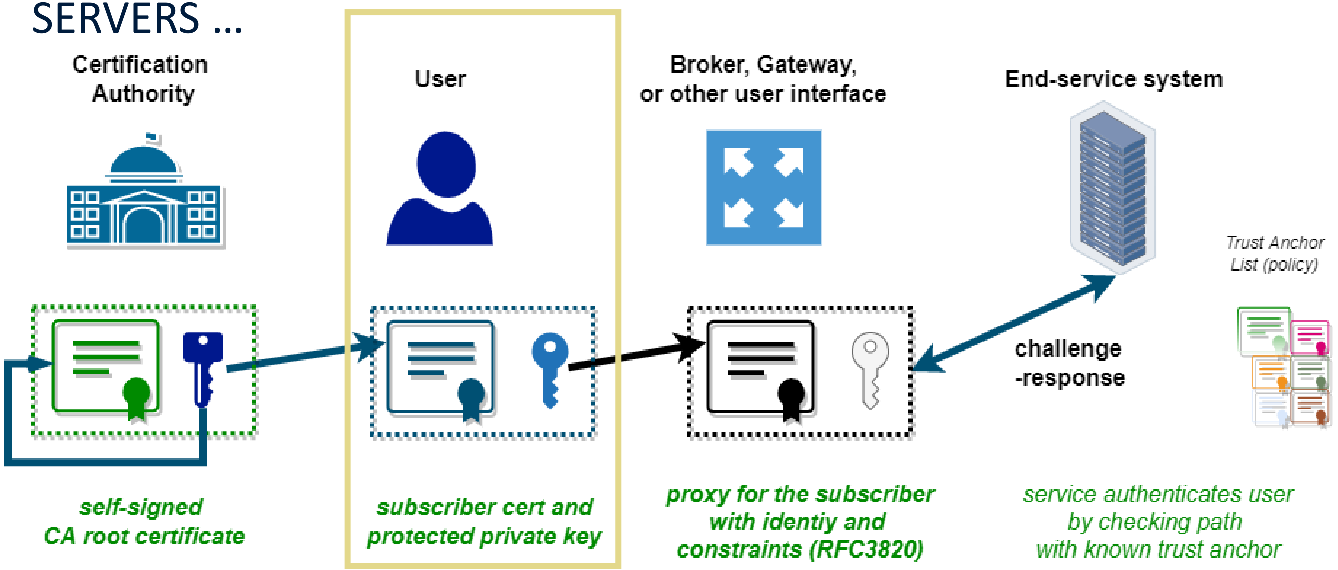


Shown is the ‘implicit flow’, other flows possible. Image source: AARC NA2 training on AAI 101
See https://openid.net/ for protocols and standardization work

>> TO FEDERATION

PKI client certificates – user* *client held* credentials

YOU HAVE SEEN *HTTPS*, BUT THE SAME PKI CERTIFICATES CAN BE USED FOR CLIENTS, NOT SERVERS …



Certification Authority

User

Broker, Gateway, or other user interface

End-service system

Trust Anchor List (policy)

challenge -response

self-signed CA root certificate

subscriber cert and protected private key

proxy for the subscriber with identiy and constraints (RFC3820)

service authenticates user by checking path with known trust anchor

```
Version: 3 (0x2)
Serial Number:
    34:f3:e3:5f:c0:53:0b:a6:ef:2b:4a:79:01:b5:50:3b
Signature Algorithm: sha384WithRSAEncryption
Issuer: C = NL, O = GEANT Vereniging, CN = GEANT eScience Personal CA 4
Validity
    Not Before: Apr  2 00:00:00 2022 GMT
    Not After : May  2 23:59:59 2023 GMT
Subject: DC = org, DC = terena, DC = tcs, C = NL, O = Nikhef, CN = David
Groep davidg@nikhef.nl
Subject Public Key Info:
    Public Key Algorithm: rsaEncryption
        RSA Public-Key: (4096 bit)
        Modulus:
            00:f0:0d:c0:ff:ee:f0:0d:f0:0d:c0:ff:ee:f0:0d:
            ...
            ff:50:6d
        Exponent: 65537 (0x10001)
X509v3 extensions:
    X509v3 Key Usage: critical
       Digital Signat
                                          Authentication
        Certificate Policies:
        Policy: 1.2.840.113612.5.2.2.5
```

RFC3820 ‘proxy’ certificates extend identity-certificate concept with (policy-restricted) delegation

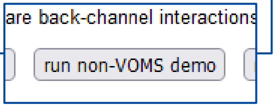
* think we found everyone in the world who could deal with certificates: about $\mathcal{O}$(100 000) people …

Maastricht University

Different tech, also an AAA push concept: X.509 and a trust PKI

```
Version: 3 (0x2)
Serial Number:
    34:f3:e3:5f:c0:53:0b:a6:ef:2b:4a:79:01:b5:50:3b
Signature Algorithm: sha384WithRSAEncryption
Issuer: C = NL, O = GEANT Vereniging, CN = GEANT eScience Personal CA 4
Validity
    Not Before: Apr  2 00:00:00 2022 GMT
    Not After : May  2 23:59:59 2023 GMT
Subject: DC = org, DC = terena, DC = tcs, C = NL, O = Nikhef, CN = David Groep davidg@nikhef.nl
Subject Public Key Info:
    Public Key Algorithm: rsaEncryption
        RSA Public-Key: (4096 bit)
        Modulus:
            00:f0:0d:c0:ff:ee:f0:0d:f0:0d:c0:ff:ee:f0:0d:
            ...
            ff:50:6d
        Exponent: 65537 (0x10001)
X509v3 extensions:
    X509v3 Key Usage: critical
        Digital Signature, Key Encipherment
    X509v3 Basic Constraints: critical
        CA:FALSE
    X509v3 Extended Key Usage:
        E-mail Protection, TLS Web Client Authentication
    X509v3 Certificate Policies:
        Policy: 1.2.840.113612.5.2.2.5
```

You should be able to get an 'IGTF-DOGWOOD' assurance certificate from RCauth.eu.
Go to https://rcdemo.nikhef.nl/ and select the 'Basic demo' and use 'run non-VOMS' to get and view your short-lived certificate



Science infrastructures using our R&E ‘federated access’

EMSO-ERIC

EMSO-ERIC, the European Multidisciplinary Seafloor and Water-Column Observatory to explore the Oceans. EGI Success Story.

EISCAT_3D

EISCAT_3D, the World's Leading Incoherent Scatter Radar. EGI Success Story.

OBSEA

OBSEA – Expandable Seafloor O… EGI Success Story

300+

Content providers

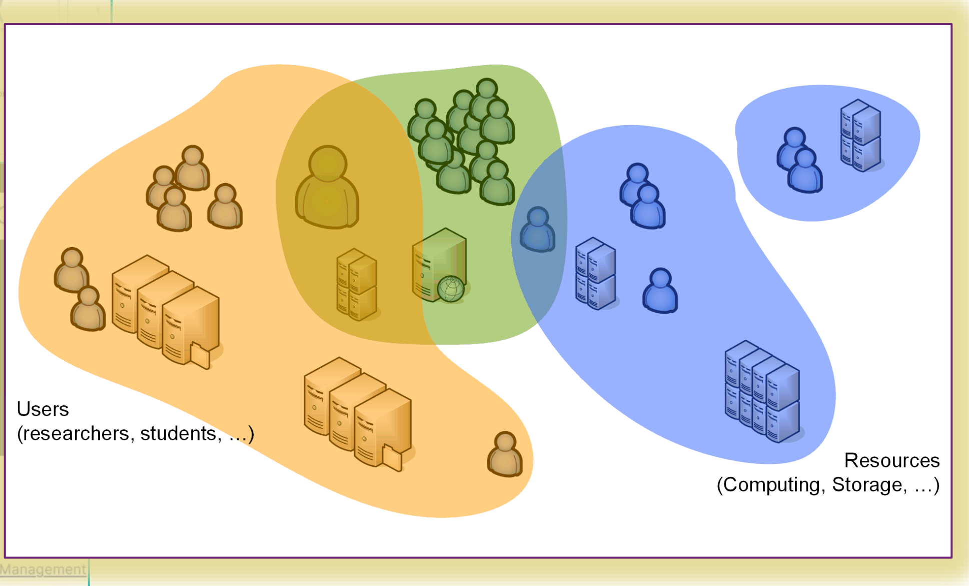
More

Storage

Research Data Management

science, cloud

Research Infrastructures


Users
(researchers, students, …)

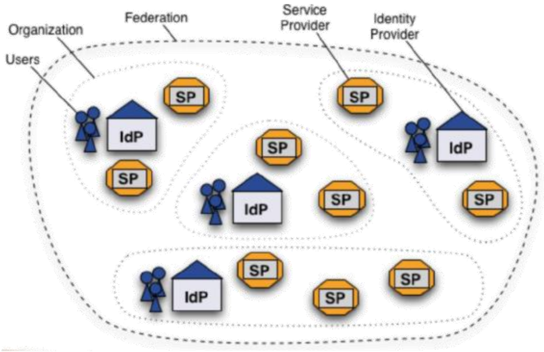
Resources
(Computing, Storage, …)

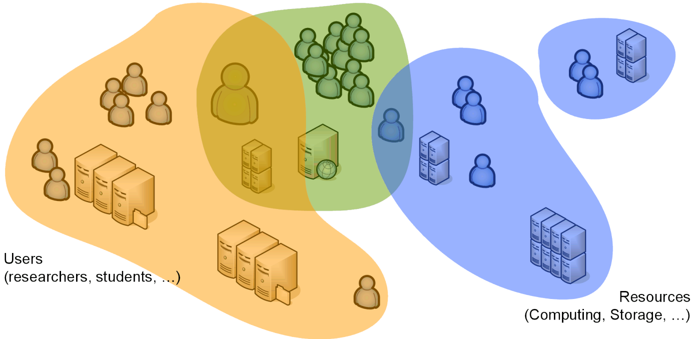
Images: CERN https://wlcg.web.cern.ch/; HADDOCK, WeNMR, @Bonvinlab https://wenmr.science.uu.nl/; Virgo, Pisa, IT; artist impression Einstein Telescope EMR region; EOSC portal in 2023, EGI catalogue https://www.egi.eu/

They look similar, yet they are not …

In the **Identity federation** picture, the source of authority is the *home organisation* via its IdP

In the **Community** picture, the source of authority is *the community itself*



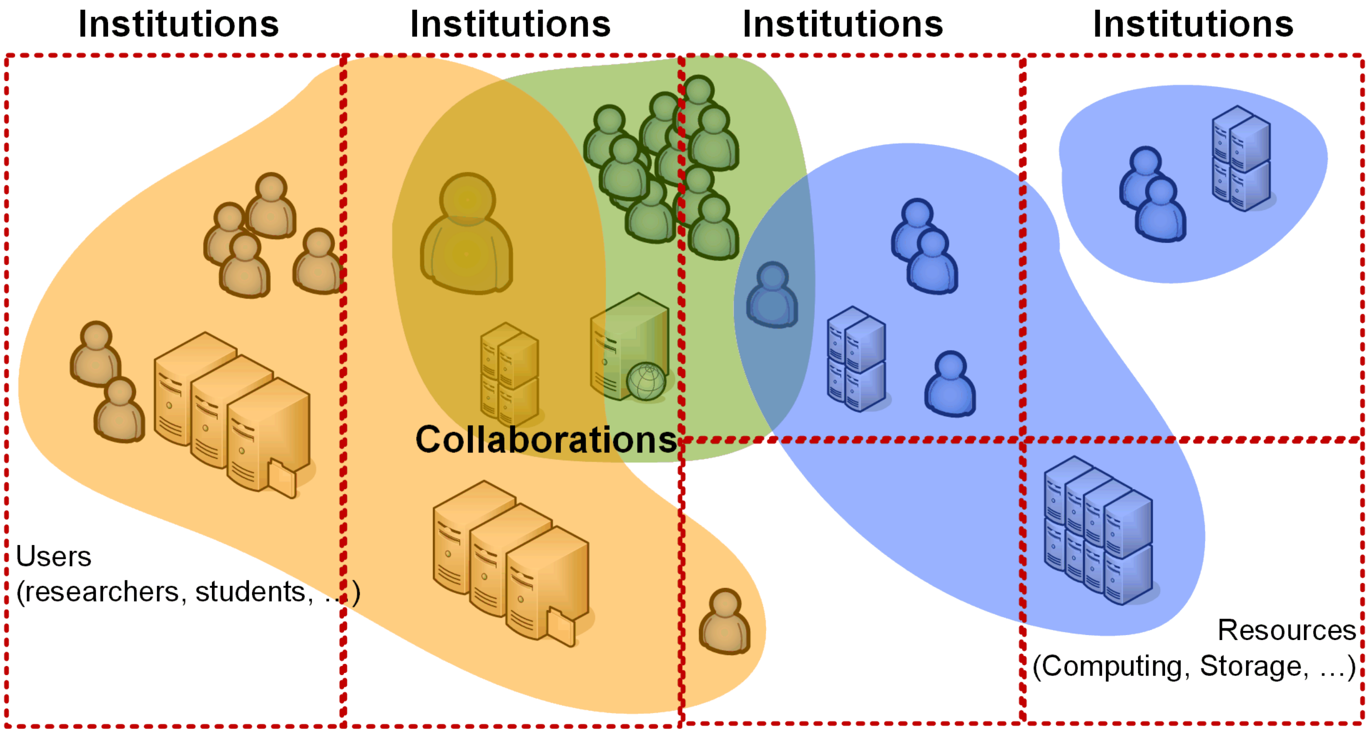


the AuthN-AuthZ separation is fundamental

to the Federated (R&E) AAI, global IGTF PKI, VOMS, 'AARC BPA' AAI architecture …

Right-hand image: Shibboleth IdP federation, Lukas Hammerle, SWITCH (CH)

Since collaborations and institutions slice in different ways



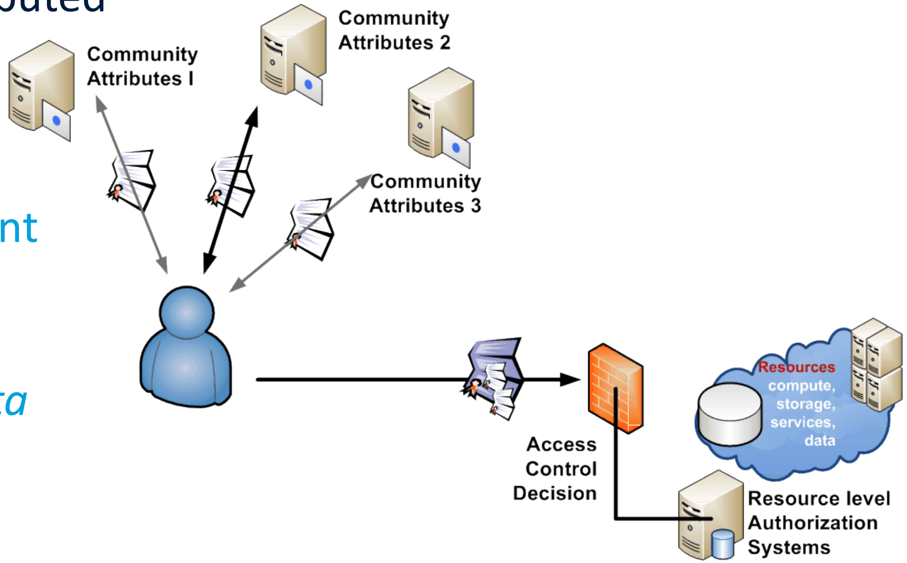
Multiple sources of authority: the community

- authorization assertion providers (attribute authorities) use the identifier(s) from authentication in their membership services

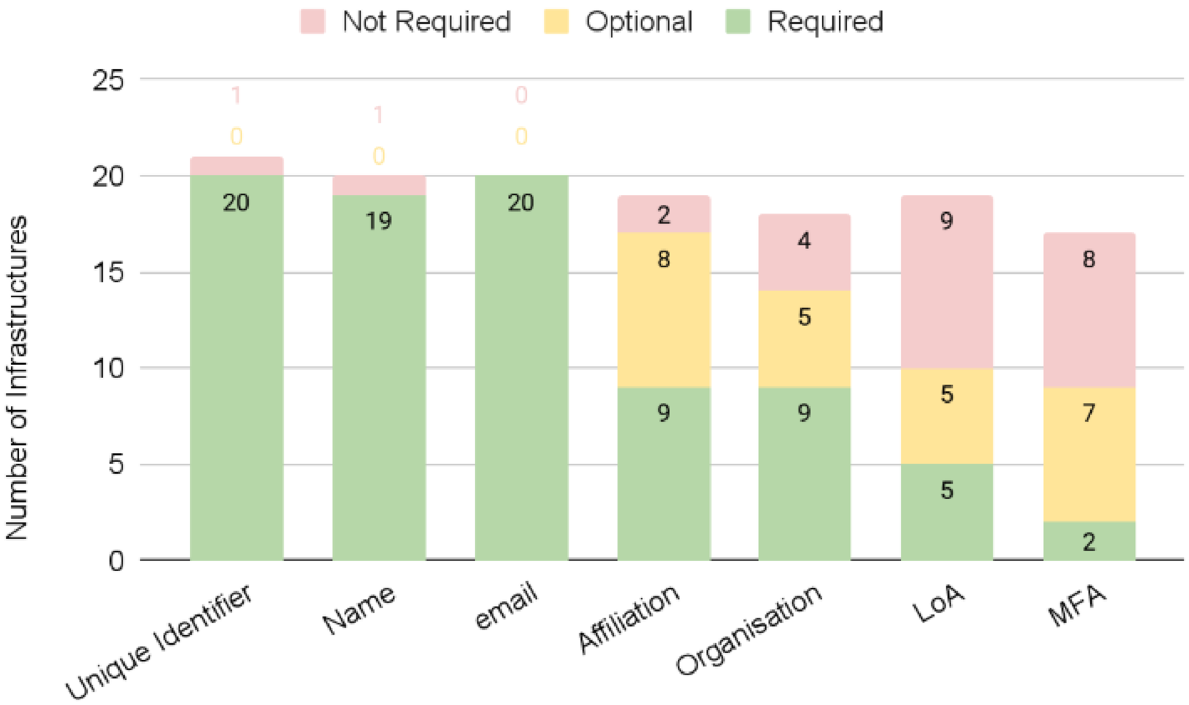
- *source of authority* for attributes is distributed

for example:
- community membership from an experiment
- affiliation status from home organisation

may be jointly needed to access sensitive data that is subject to medical-ethical clearance



Research Infrastructures: what they *actually* need from ‘home’



Glossary

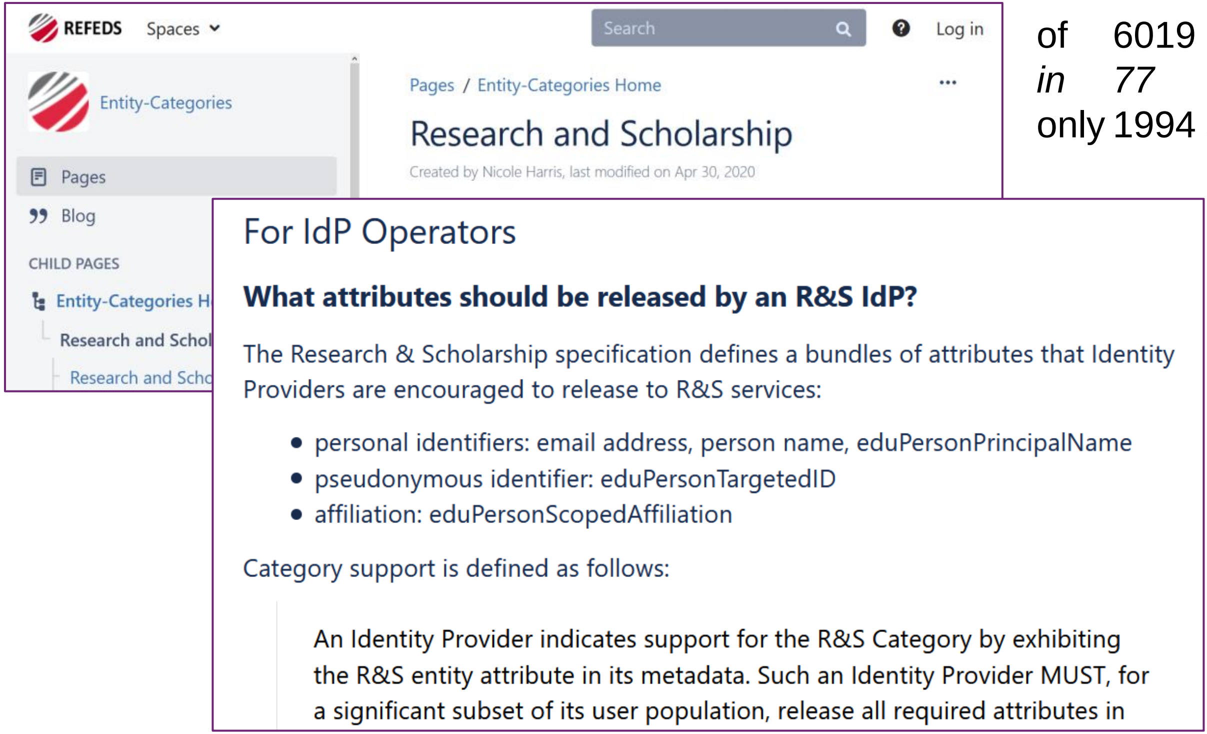
Affiliation: what *type* of entity are you (student, faculty, alumnus, …)

LoA: level of authentication assurance (like passport identity vetting and ‘freshness’ of data)

MFA: multi-factor authentication (password, 6-digit code, SMS, fingerprint)

Source: Marina Adomeit, Janos Mohasci, *et al.* AARC TREE Use-case collection and analysis (D3.2), 2025 (under review)
The one infra that did ‘not need a unique identifier’ actually stated: “<our infra> assingns own identifier upon registration” – so the unique identifier is *still* there!

For starters: sharing good user identifiers is non-trivial ☹



Research and Scholarship

Created by Nicole Harris, last modified on Apr 30, 2020

For IdP Operators

What attributes should be released by an R&S IdP?

The Research & Scholarship specification defines a bundles of attributes that Identity Providers are encouraged to release to R&S services:

- personal identifiers: email address, person name, eduPersonPrincipalName
- pseudonymous identifier: eduPersonTargetedID
- affiliation: eduPersonScopedAffiliation

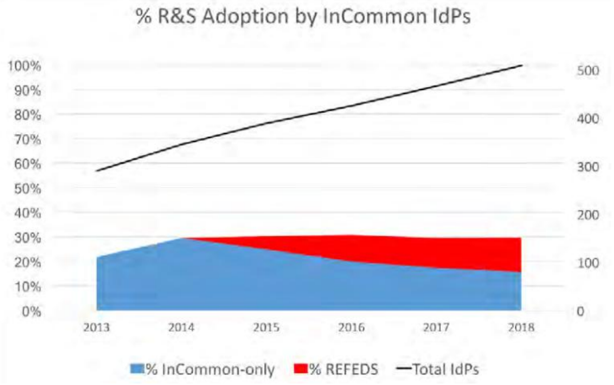
Category support is defined as follows:

> An Identity Provider indicates support for the R&S Category by exhibiting the R&S entity attribute in its metadata. Such an Identity Provider MUST, for a significant subset of its user population, release all required attributes in

of 6019 identity providers
in 77 federations,
only 1994 support R&S or Personalised access

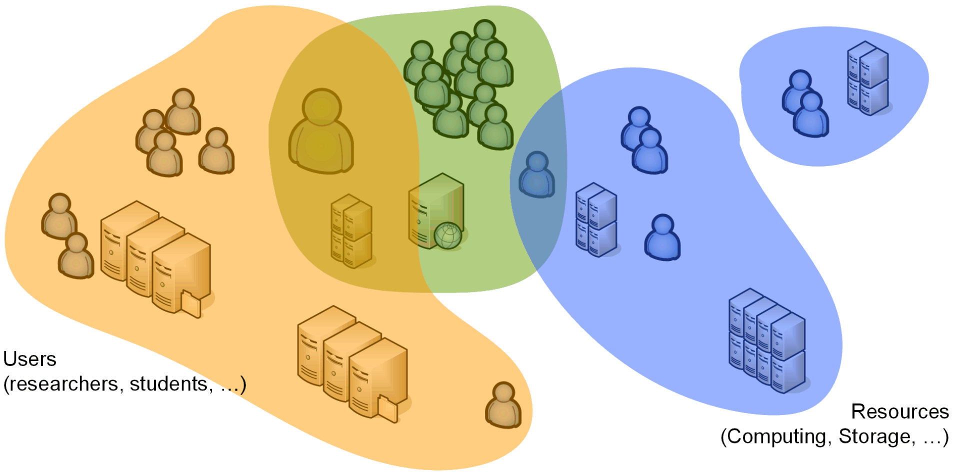
33%

~ constant since 2018 ☹

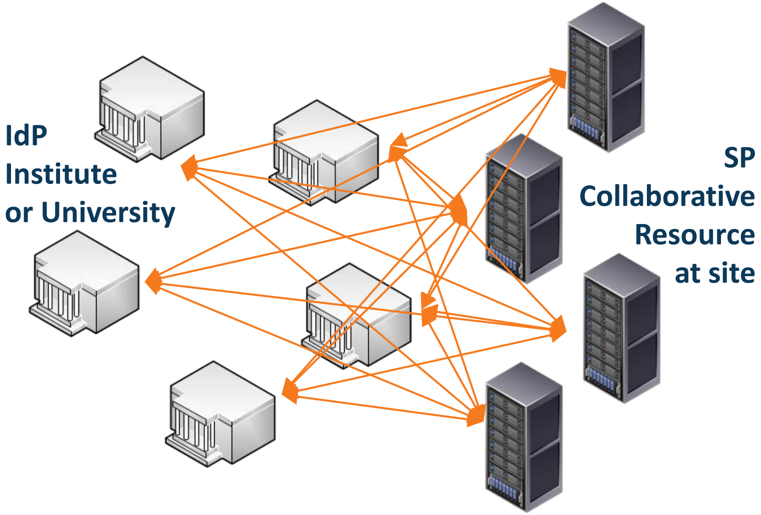


Graph: InCommon: Attributes-WG-Recommendations-May2018.pdf; Entity Category stats as per 2025-03-03, from https://technical.edugain.org/entities

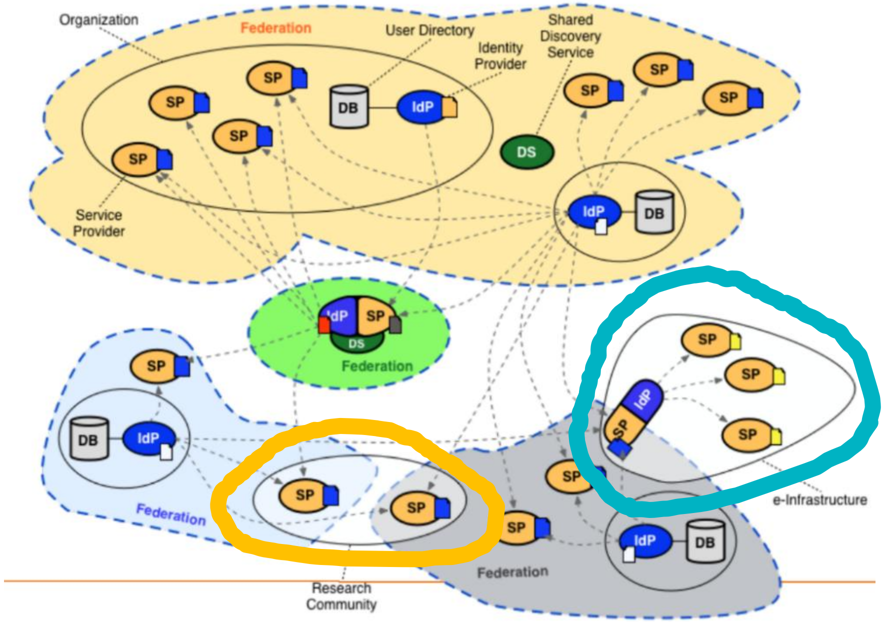
A fundamental scaling issue remained unique to research



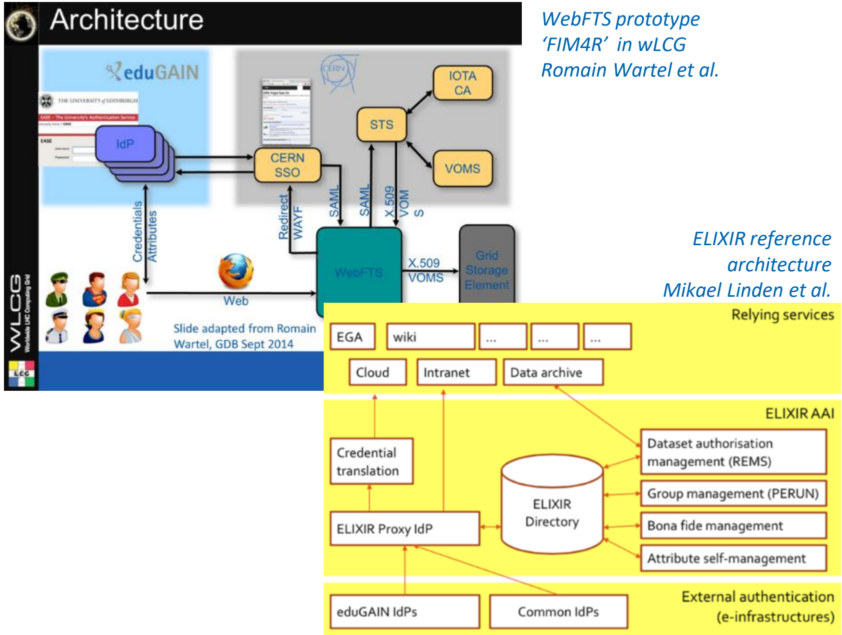
for identity and user data
‘n x m’ agreements remain(ed)



Managing complexity: distributed diverse identity sources



they were composed of many services
each of which had to manage federation complexity



Slide adapted from Romain Wartel, GDB Sept 2014

WebFTS prototype
'FIM4R' in wLCG
Romain Wartel et al.

ELIXIR reference architecture
Mikael Linden et al.

but most communities had started to invent
their own 'proxy' model to abstract complexity

Community images: Romain Wartel, CERN; Mikael Linden, CSC; Federation image (R): Lukas Hammerle, SWITCH

The IdP-SP bridge

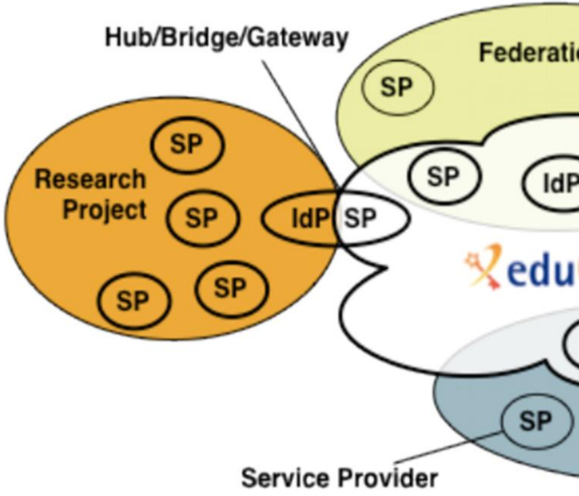
- Access services using **identities from users' Home Organizations**,
- but **hide complexity** of multiple IdPs, federations,

  and different technologies for authentication and authorisation
- **One persistent identity**

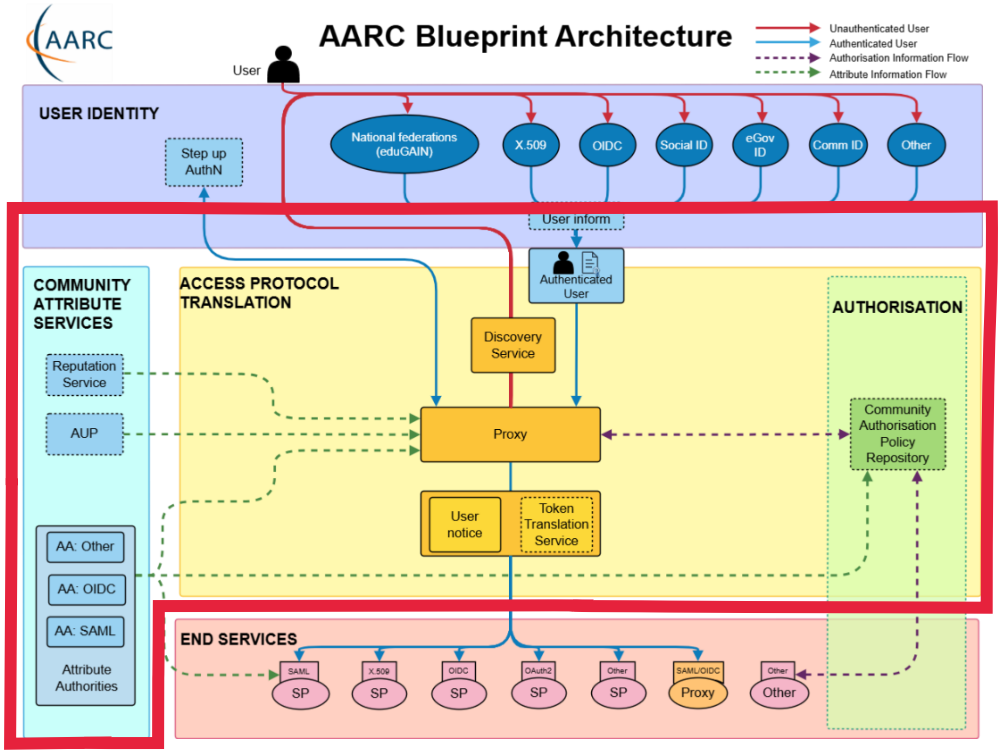
  across all the community's services through **account linking**
- **Access** services

  based on role(s) users have **in the collaboration**.
- For both **web** and **non-web** resources
- Integration of **guest identity solutions**
- **Support for stronger authentication assurance** mechanisms

often known as proxy!



AARC Blueprint – making the bridge a first-class citizen



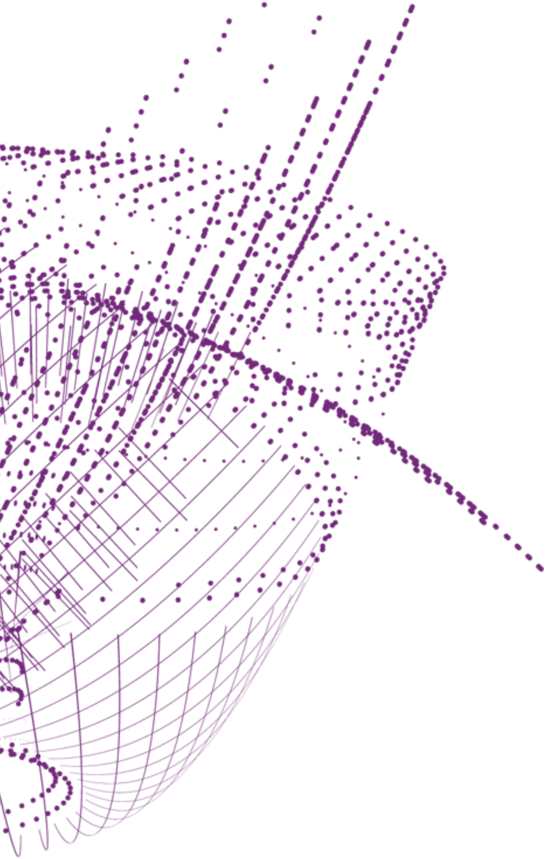
AARC BPA 2019 'Community First'
https://aarc-community.org/guidelines/aarc-g045/

Manage users and access rights

with interoperable **building blocks for 'AAI infrastructure' architects**

that are

- technology-agnostic
- have multiple implementations
- come with policy templates & good practice guides

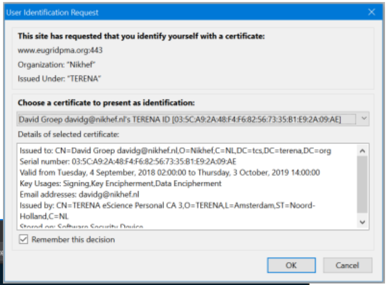


TCS and RCauth

Token translation

Bridges and Token Translation Services

TCS - for users that manage to grasp the idea



User Identification Request

This site has requested that you identify yourself with a certificate:

www.eugridpma.org:443

Organization: "Nikhef"

Issued Under: "TERENA"

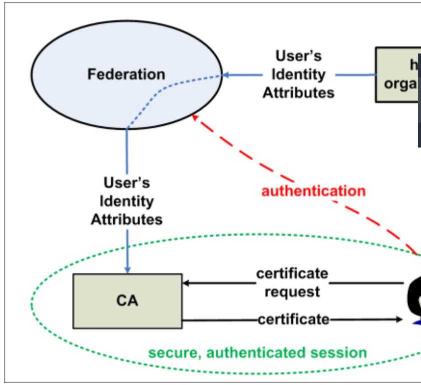
Choose a certificate to present as identification:

David Groep davidg@nikhef.nl's TERENA ID [03:5C:A9:2A:48:F4:F6:82:56:73:35:B1:E9:2A:09:AE]

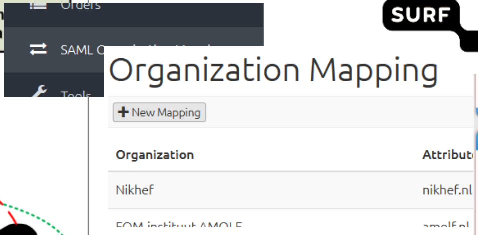
Details of selected certificate:

```
Issued to: CN=David Groep davidg@nikhef.nl,O=Nikhef,C=NL,DC=tcs,DC=terena,DC=org
Serial number: 03:5C:A9:2A:48:F4:F6:82:56:73:35:B1:E9:2A:09:AE
Valid from Tuesday, 4 September, 2018 02:00:00 to Thursday, 3 October, 2019 14:00:00
Key Usages: Signing,Key Encipherment,Data Encipherment
Email addresses: davidg@nikhef.nl
Issued by: CN=TERENA eScience Personal CA 3,O=TERENA,L=Amsterdam,ST=Noord-Holland,C=NL
Stored on: Software Security Device
```

☑ Remember this decision

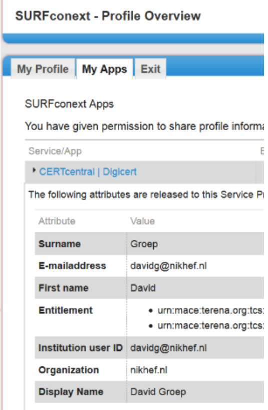


Federation
User's Identity Attributes
home organisation
User's Identity Attributes
authentication
CA
certificate request
certificate
secure, authenticated session



Organization Mapping

Organization	Attribute
Nikhef	nikhef.nl
FOM instituut AMOLF	amolf.nl



SURFconext - Profile Overview

SURFconext Apps

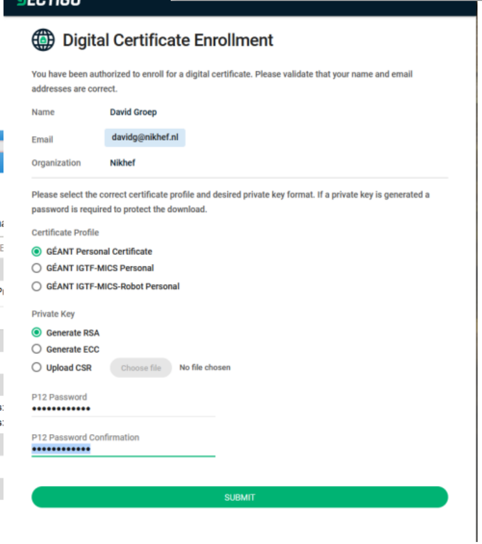
You have given permission to share profile informa

Service/App

CERTcentral | Digicert

The following attributes are released to this Service Pr

Attribute	Value
Surname	Groep
E-mailaddress	davidg@nikhef.nl
First name	David
Entitlement	• urn:mace:terena.org:tcs: • urn:mace:terena.org:tcs:
Institution user ID	davidg@nikhef.nl
Organization	nikhef.nl
Display Name	David Groep



SECTIGO

Digital Certificate Enrollment

You have been authorized to enroll for a digital certificate. Please validate that your name and email addresses are correct.

Name: David Groep

Email: davidg@nikhef.nl

Organization: Nikhef

Please select the correct certificate profile and desired private key format. If a private key is generated a password is required to protect the download.

Certificate Profile

● GÉANT Personal Certificate
○ GÉANT IGTF-MICS Personal
○ GÉANT IGTF-MICS-Robot Personal

Private Key

● Generate RSA
○ Generate ECC
○ Upload CSR Choose file No file chosen

P12 Password: ••••••••••••

P12 Password Confirmation: ••••••••••••

TCS is a SAML Service Provider (today by Sectigo) to eduGAIN: where eligible authenticated users obtain client certificates for access to many research services

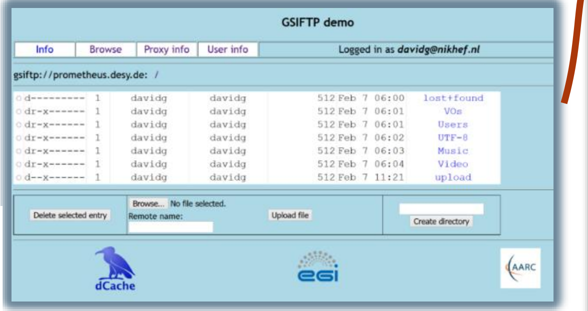
A globally recognized identity for all employees & students (they are automatically eligible!).

GEANT Trusted Certificate Service - https://ca.dutchgrid.nl/tcs/, https://cert-manager.com/customer/surfnet/idp/clientgeant, https://www.geant.org/Services/Trust_identity_and_security/Pages/TCS.aspx

Seamless in-line token translation services from ‘SAML’ to PKIX

user facing | hidden back-end

Community Science Portal

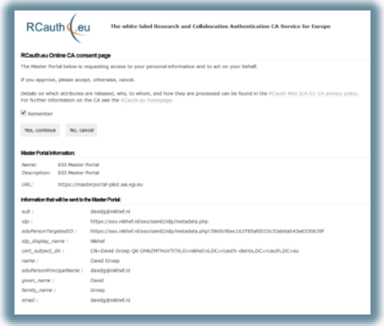


GSIFTP demo

Info | Browse | Proxy info | User info | Logged in as ***davidg@nikhef.nl***

gsiftp://prometheus.desy.de: /

```
○ d---------  1  davidg  davidg  512 Feb 7 06:00  lost+found
○ dr-x------  1  davidg  davidg  512 Feb 7 06:01  VOs
○ dr-x------  1  davidg  davidg  512 Feb 7 06:01  Users
○ dr-x------  1  davidg  davidg  512 Feb 7 06:02  UTF-8
○ dr-x------  1  davidg  davidg  512 Feb 7 06:03  Music
○ dr-x------  1  davidg  davidg  512 Feb 7 06:04  Video
○ d--x------  1  davidg  davidg  512 Feb 7 11:21  upload
```



RCauth.eu The white-label Research and Collaboration Authentication CA Service for Europe

RCauth.eu Online CA consent page

The Master Portal below is requesting access to your personal information and to act on your behalf.

If you approve, please accept, otherwise, cancel.

Details on which attributes are released, why, to whom, and how they are processed can be found in the RCauth Pilot ICA G1 CA privacy policy. For further information on the CA see the RCauth.eu homepage.

☑ Remember

Yes, continue | No, cancel

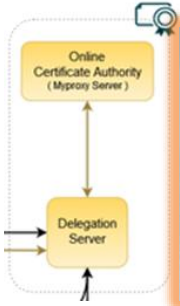
Master Portal information:

Name:	EGI Master Portal
Description:	EGI Master Portal
URL:	https://masterportal-pilot.aai.egi.eu

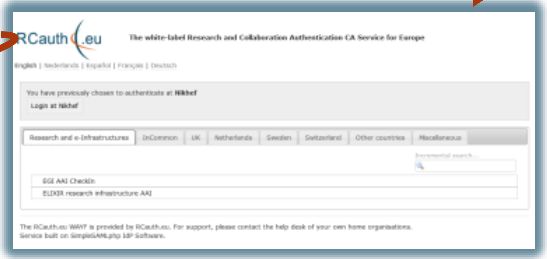
Information that will be sent to the Master Portal:

sub :	davidg@nikhef.nl
idp :	https://sso.nikhef.nl/sso/saml2/idp/metadata.php
eduPersonTargetedID :	https://sso.nikhef.nl/sso/saml2/idp/metadata.php!3960c9bec163785afd515c33ab0a543a6330639f
idp_display_name :	Nikhef
cert_subject_dn :	CN=David Groep QK-OHkZMTHoVTtT6,O=nikhef.nl,DC=rcauth-clients,DC=rcauth,DC=eu
name :	David Groep
eduPersonPrincipalName :	davidg@nikhef.nl
given_name :	David
family_name :	Groep
email :	davidg@nikhef.nl

Infrastructure Master Portal Credential Store



IGTF accredited PKIX Authority



RCauth.eu The white-label Research and Collaboration Authentication CA Service for Europe

English | Nederlands | Español | Français | Deutsch

You have previously chosen to authenticate at **Nikhef**

Login at Nikhef

Research and e-Infrastructures | InCommon | UK | Netherlands | Sweden | Switzerland | Other countries | Miscellaneous

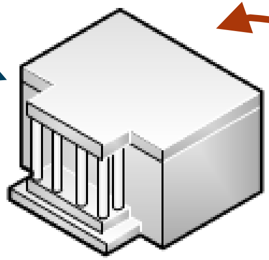
Incremental search...

- EGI AAI CheckIn
- ELIXIR research infrastructure AAI

The RCauth.eu WAYF is provided by RCauth.eu. For support, please contact the help desk of your own home organisations. Service built on SimpleSAMLphp IdP Software.

Policy Filtering WAYF to eduGAIN

REFEDS R&S
Sirtfi Trust



User Home Org
or Infrastructure IdP

see also https://rcdemo.nikhef.nl/

Built on CILogon and MyProxy, see www.cilogon.org CILogon

Unique certificated from FIM via eduPerson and REFEDS R&S

Sources of naming and uniqueness, that work *today*

- **eduPersonPrincipalName** – scoped point-in-time unique identifier, which could be, but usually is not, privacy preserving: “davidg@nikhef.nl”, “P70081609@maastrichtuniversity.nl”
- **eduPersonTargetedID** – scoped transient non-reassigned identifier, like urn:geant:nikhef.nl:nikidm:idp:sso!*27c8d63ed42c84af2875e2984*
- **subject-id** - a scoped persistent non-reassigned identifier, which should be privacy-preserving: 44f7751265a6e8b228f9@nikhef.nl

Plus the (domain-name based) schacHomeOrganisation and a ‘**representation of the real name**’

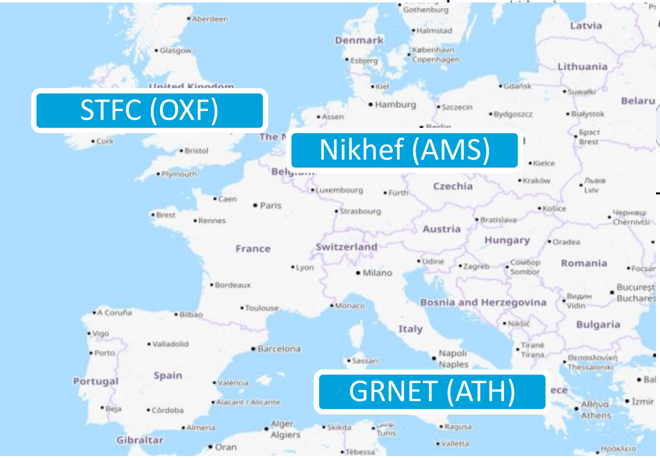
/DC=eu/DC=rcauth/DC=rcauth-clients/O=*orgdisplayname*/CN=*commonName +uniqeness*

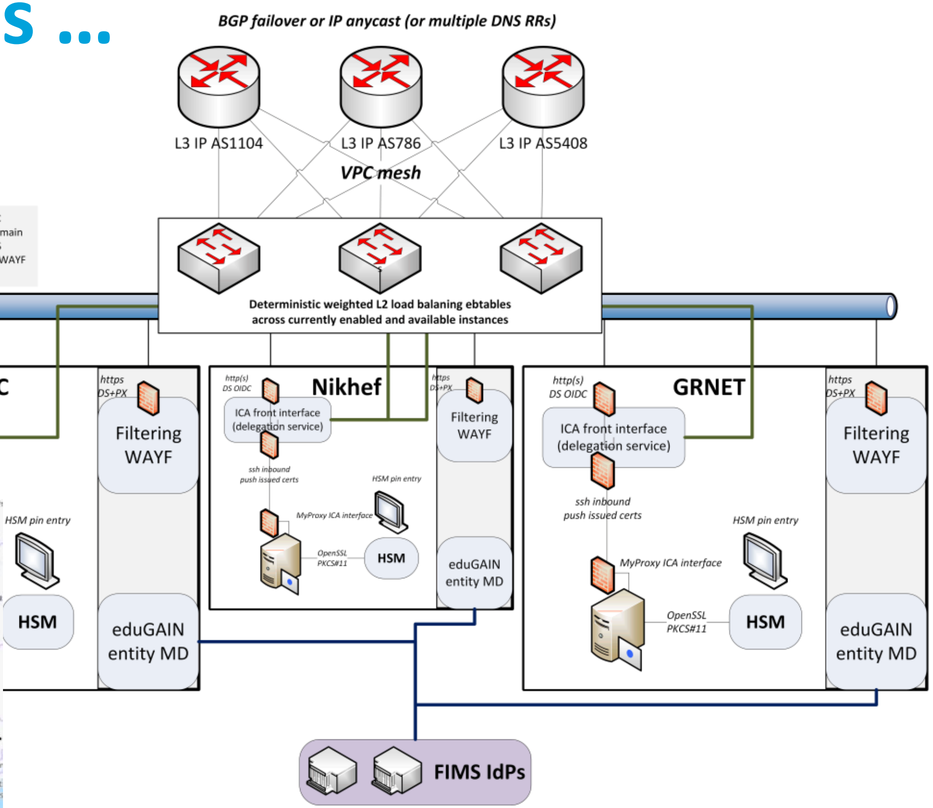
uniqueness will added to commonName via hashing of *ePPN, ePTID, subject-id*, so that an enquiry via the issuer allows unique identification of the vetted entity”

Since we do not like SPOFs …

Distributed High Availability setup

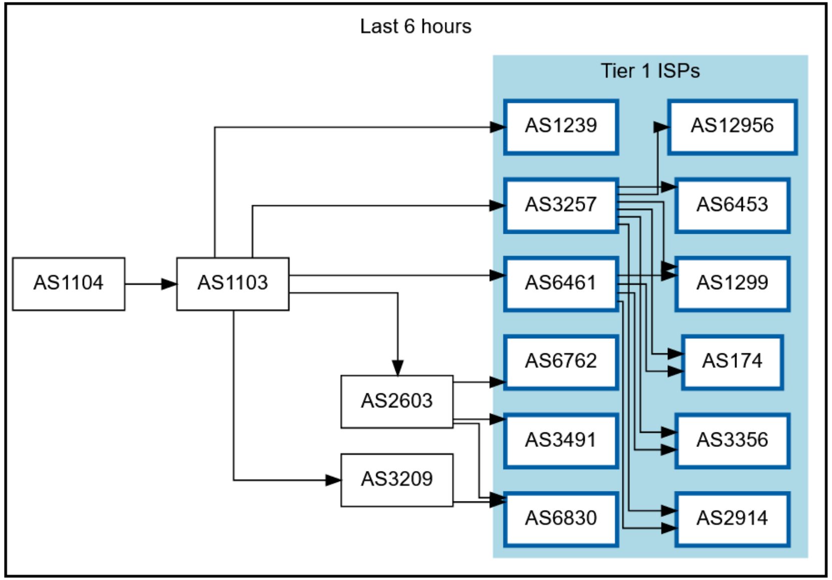
- across the 3 sites
- design for minimal effort
- readily-available techniques
  - L3 VPN (OpenVPN) or L2 VPC
  - Linux HAProxy

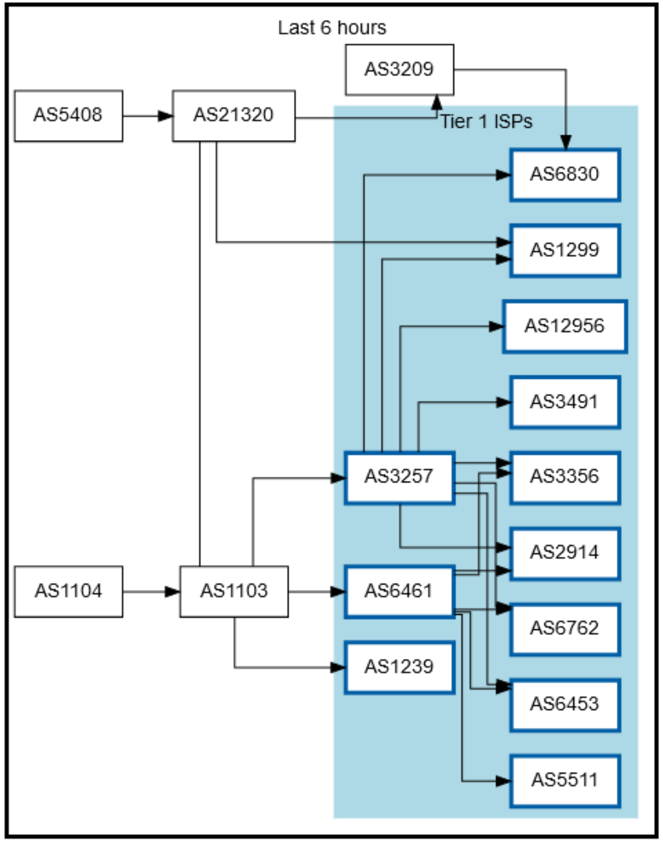



BGP failover or IP anycast (or multiple DNS RRs)
L3 IP AS1104
L3 IP AS786
L3 IP AS5408
VPC mesh
protected low-latency L2 VPC
balancing forwarding domain
connected both OIDC DS frontends as well as DS/WAYF
inter-site database sync
Deterministic weighted L2 load balaning ebtables across currently enabled and available instances
STFC
http(s) DS OIDC
ICA front interface (delegation service)
HSM pin entry
HSM
https DS+PX
Filtering WAYF
eduGAIN entity MD
Nikhef
http(s) DS OIDC
ICA front interface (delegation service)
ssh inbound push issued certs
HSM pin entry
MyProxy ICA interface
OpenSSL PKCS#11
HSM
https DS+PX
Filtering WAYF
eduGAIN entity MD
GRNET
http(s) DS OIDC
ICA front interface (delegation service)
ssh inbound push issued certs
HSM pin entry
MyProxy ICA interface
OpenSSL PKCS#11
HSM
https DS+PX
Filtering WAYF
eduGAIN entity MD
FIMS IdPs

work supported by the EOSC Hub and EOSC Future Horizon Europe projects

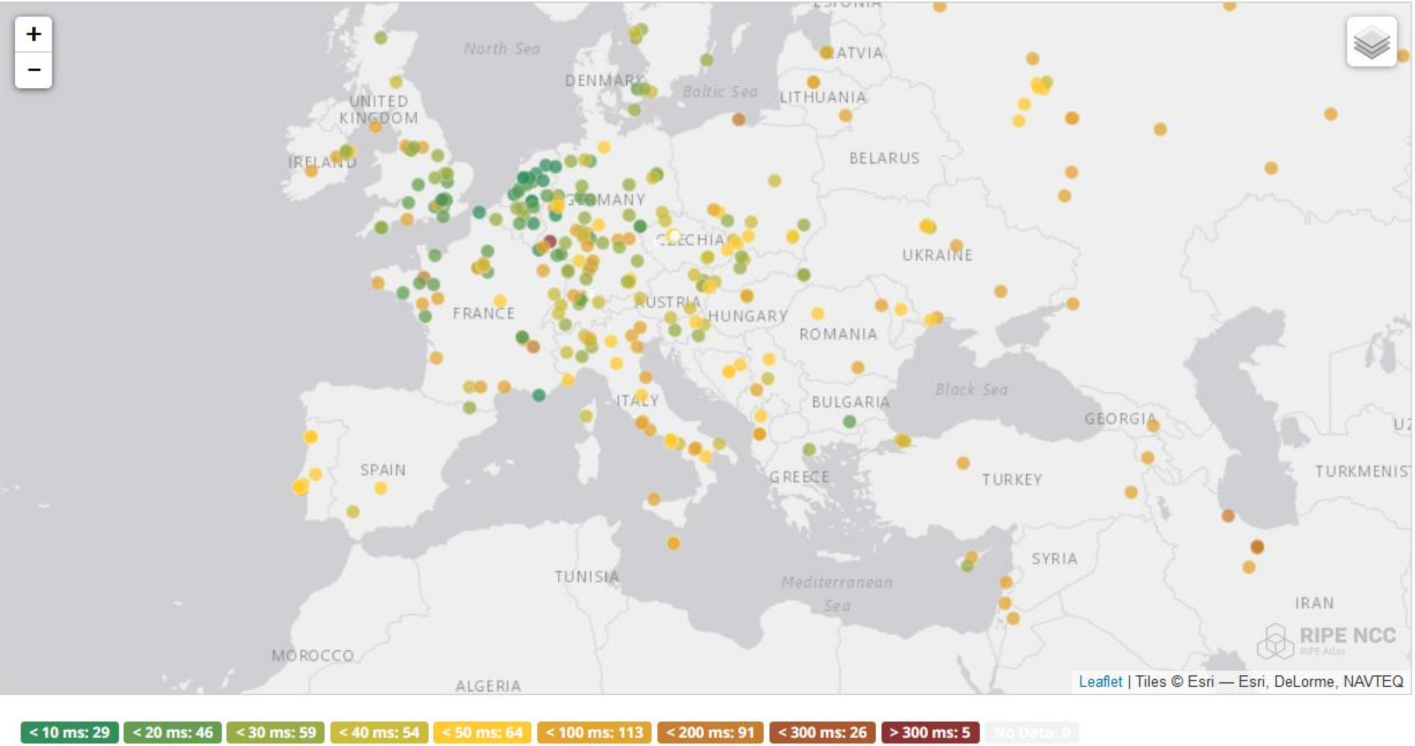
Getting 2a07:8504:1a0::/48 out there





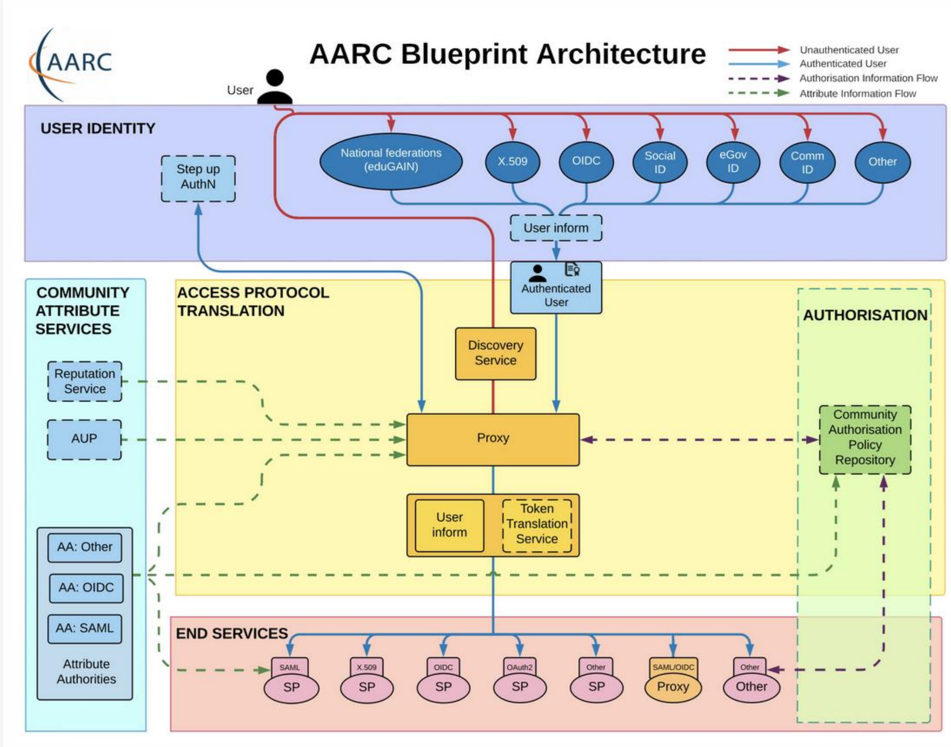
route maps: bgp.tools for 2a07:8504:1a0::/48 – IPv4 for 145.116.216.0/24 is similar – imagery from November 2022

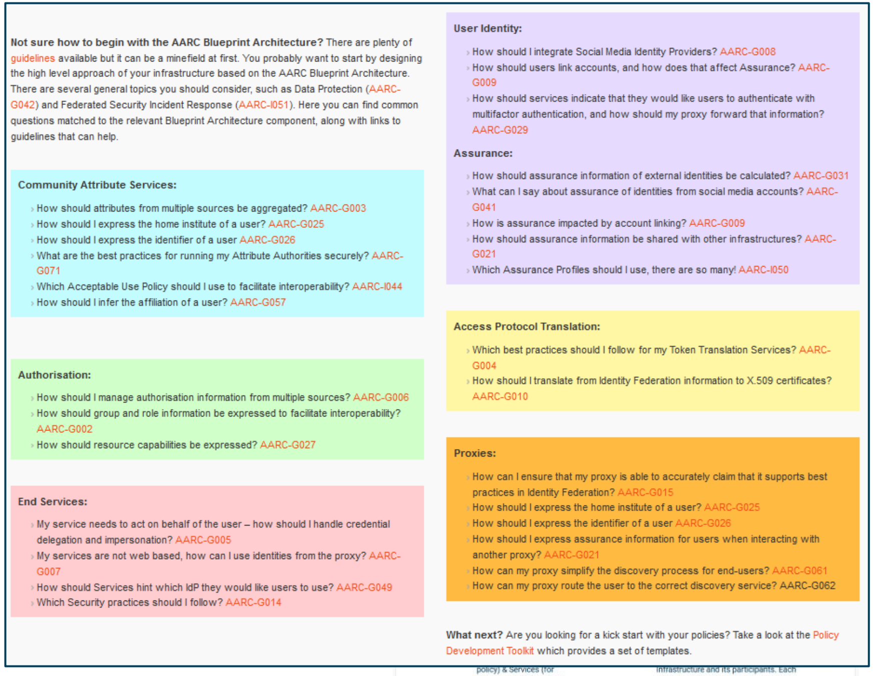
And you get reasonable load balancing in Europe for free



map: RIPE NCC RIPE Atlas - 500 probes, distributed across Europe (https://atlas.ripe.net/measurements/50949024/)

More than just nice colours





Not sure how to begin with the AARC Blueprint Architecture? There are plenty of guidelines available but it can be a minefield at first. You probably want to start by designing the high level approach of your infrastructure based on the AARC Blueprint Architecture. There are several general topics you should consider, such as Data Protection (AARC-G042) and Federated Security Incident Response (AARC-I051). Here you can find common questions matched to the relevant Blueprint Architecture component, along with links to guidelines that can help.

Community Attribute Services:

- How should attributes from multiple sources be aggregated? AARC-G003
- How should I express the home institute of a user? AARC-G025
- How should I express the identifier of a user AARC-G026
- What are the best practices for running my Attribute Authorities securely? AARC-G071
- Which Acceptable Use Policy should I use to facilitate interoperability? AARC-I044
- How should I infer the affiliation of a user? AARC-G057

Authorisation:

- How should I manage authorisation information from multiple sources? AARC-G006
- How should group and role information be expressed to facilitate interoperability? AARC-G002
- How should resource capabilities be expressed? AARC-G027

End Services:

- My service needs to act on behalf of the user – how should I handle credential delegation and impersonation? AARC-G005
- My services are not web based, how can I use identities from the proxy? AARC-G007
- How should Services hint which IdP they would like users to use? AARC-G049
- Which Security practices should I follow? AARC-G014

User Identity:

- How should I integrate Social Media Identity Providers? AARC-G008
- How should users link accounts, and how does that affect Assurance? AARC-G009
- How should services indicate that they would like users to authenticate with multifactor authentication, and how should my proxy forward that information? AARC-G029

Assurance:

- How should assurance information of external identities be calculated? AARC-G031
- What can I say about assurance of identities from social media accounts? AARC-G041
- How is assurance impacted by account linking? AARC-G009
- How should assurance information be shared with other infrastructures? AARC-G021
- Which Assurance Profiles should I use, there are so many! AARC-I050

Access Protocol Translation:

- Which best practices should I follow for my Token Translation Services? AARC-G004
- How should I translate from Identity Federation information to X.509 certificates? AARC-G010

Proxies:

- How can I ensure that my proxy is able to accurately claim that it supports best practices in Identity Federation? AARC-G015
- How should I express the home institute of a user? AARC-G025
- How should I express the identifier of a user AARC-G026
- How should I express assurance information for users when interacting with another proxy? AARC-G021
- How can my proxy simplify the discovery process for end-users? AARC-G061
- How can my proxy route the user to the correct discovery service? AARC-G062

What next? Are you looking for a kick start with your policies? Take a look at the Policy Development Toolkit which provides a set of templates.

	policy) & Services (for service specific policies)		infrastructure and its participants. Each service in the infrastructure, as well as the infrastructure itself, should complete the template.	
Service Operations Security Policy	Infrastructure Management	Services (abide by)	This policy defines requirements for running a service within the Infrastructure.	Google Doc
Acceptable Use Policy	Infrastructure Management (for baseline) & Research Communities (for community specific restrictions)	Users (abide by)	This is a template for the acceptable u policy that users must acce Research Infrastructure. It sh augmented by the Research C	oogle

Showing 1 to 9 of 9 entries

https://aarc-community.org/guidelines/

There is plenty of AARC deployments …

umbrella

FENIX RI

LUMI

Orthros

eosc

neic PUHURI

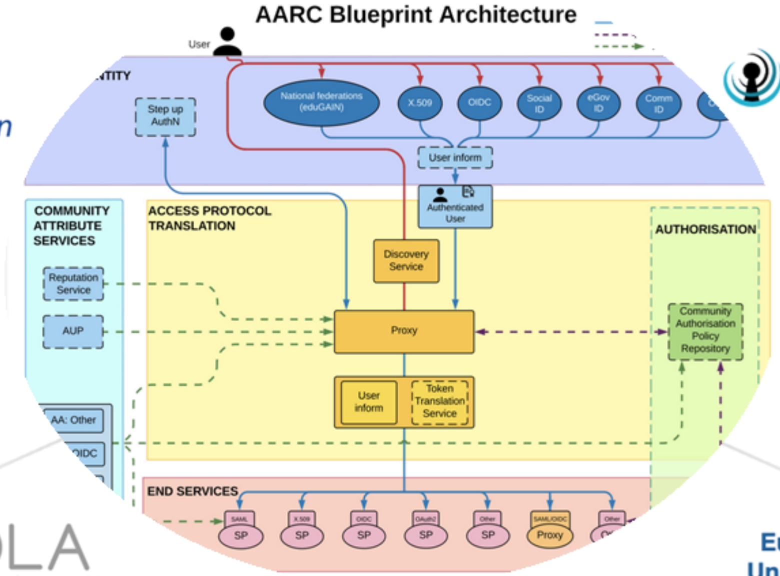
HPCI High Performance Computing Infrastructure

EUROfusion

PRACE

ACCESS

SSHOC



panosc photon and neutron open science cloud

WLCG Worldwide LHC Computing Grid

EuroHPC Joint Undertaking

GÉANT Networks • Services • People

SURF (SRAM)

EOSC-Life

LEONARDO

egi

iris

OLA online learning agreement

European Universities Alliances

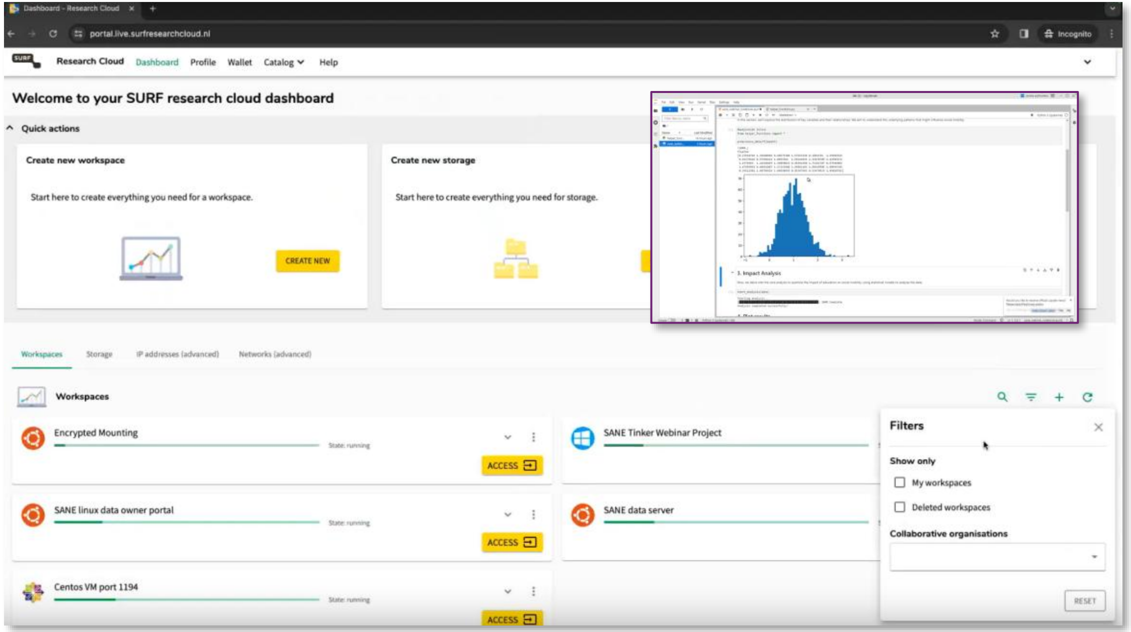
HIFIS

Erasmus+

erasmus without paper

European Commission

Example: SURF Research Cloud Secure Supercomputing



Welcome to your SURF research cloud dashboard

Create new workspace

Start here to create everything you need for a workspace.

Create new storage

Start here to create everything you need for storage.

Workspaces

Encrypted Mounting — State: running

SANE linux data owner portal — State: running

Centos VM port 1194 — State: running

SANE Tinker Webinar Project

SANE data server

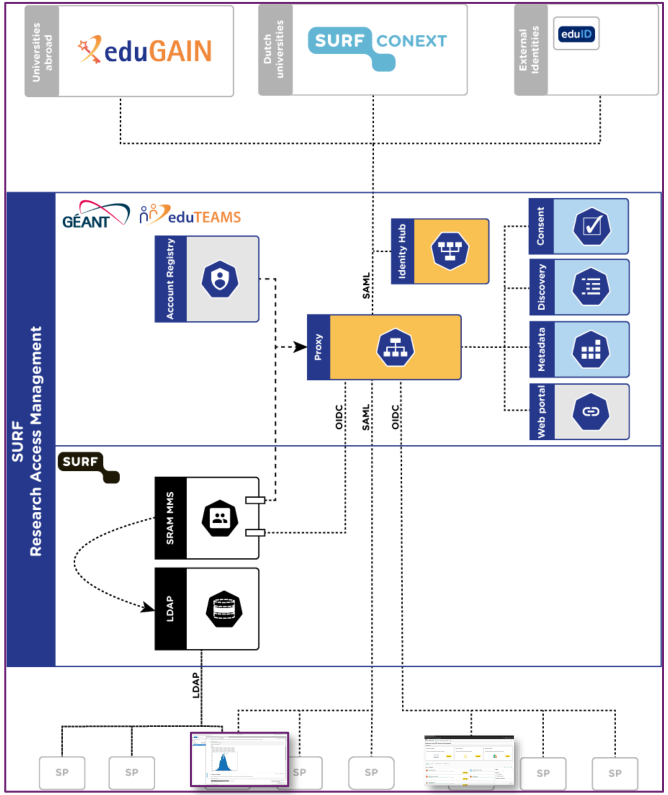
Filters

Show only

☐ My workspaces

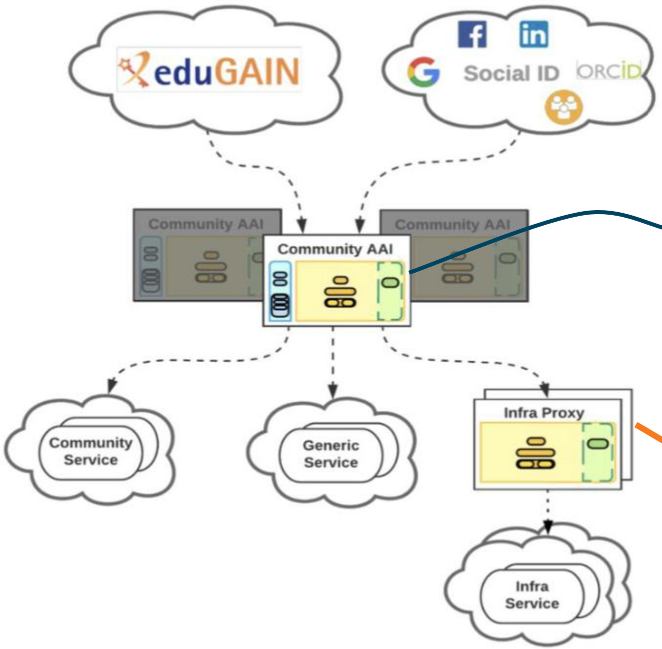
☐ Deleted workspaces

Collaborative organisations



SURF SRAM architecture, Raoul Teeuwen *et al.* from
https://servicedesk.surf.nl/wiki/display/IAM/Dienstbeschrijving+SURF+Research+Access+Management
SURF Research Cloud capture: from Introduction to SANE (Secure ANalysis Environment)
webinar February 2024, by Martin Brandt et al., SURF
https://www.surf.nl/themas/onderzoeksinfrastructuur/sane-veilige-omgeving-voor-analyse-van-gevoelige-data

… but one proxy is not enough in a research cloud



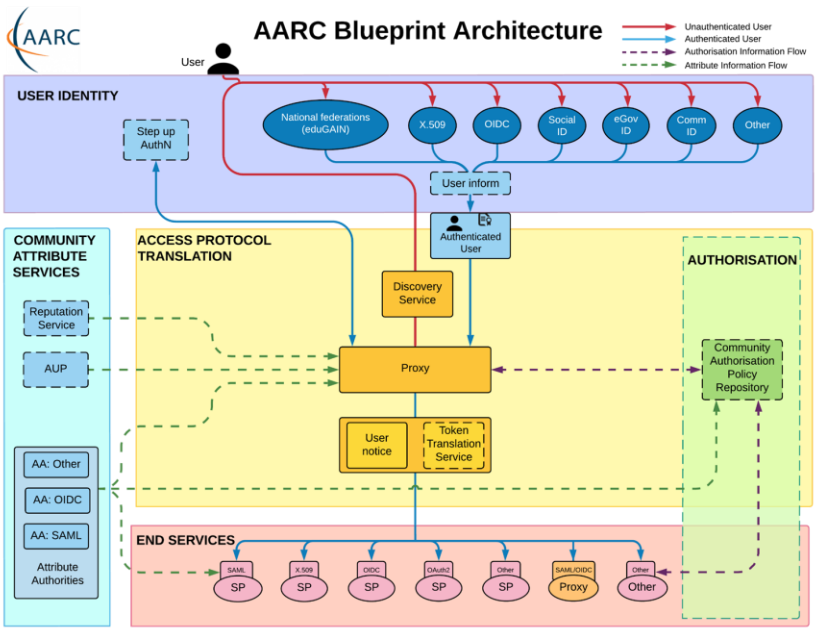
Community AAI

streamline researchers’ access to services,
both those provided by their own infrastructure
as well as the services provided
by shared infrastructures from other communities.

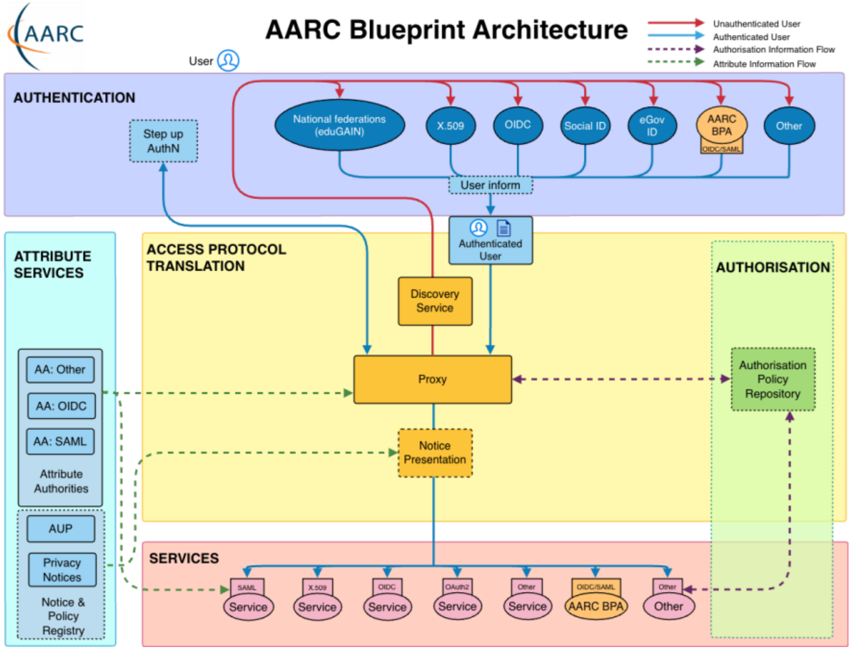
Infrastructure Proxy

enables Infrastructures with large number of resources,
to provide them through a single integration point,
where the Infrastructure can maintain centrally
all the relevant Policies and business logic
for making available resources to multiple communities

AARC Blueprint Architecture 2025: Component Layers - What has changed since AARC-BPA-2019? Can you spot the differences?

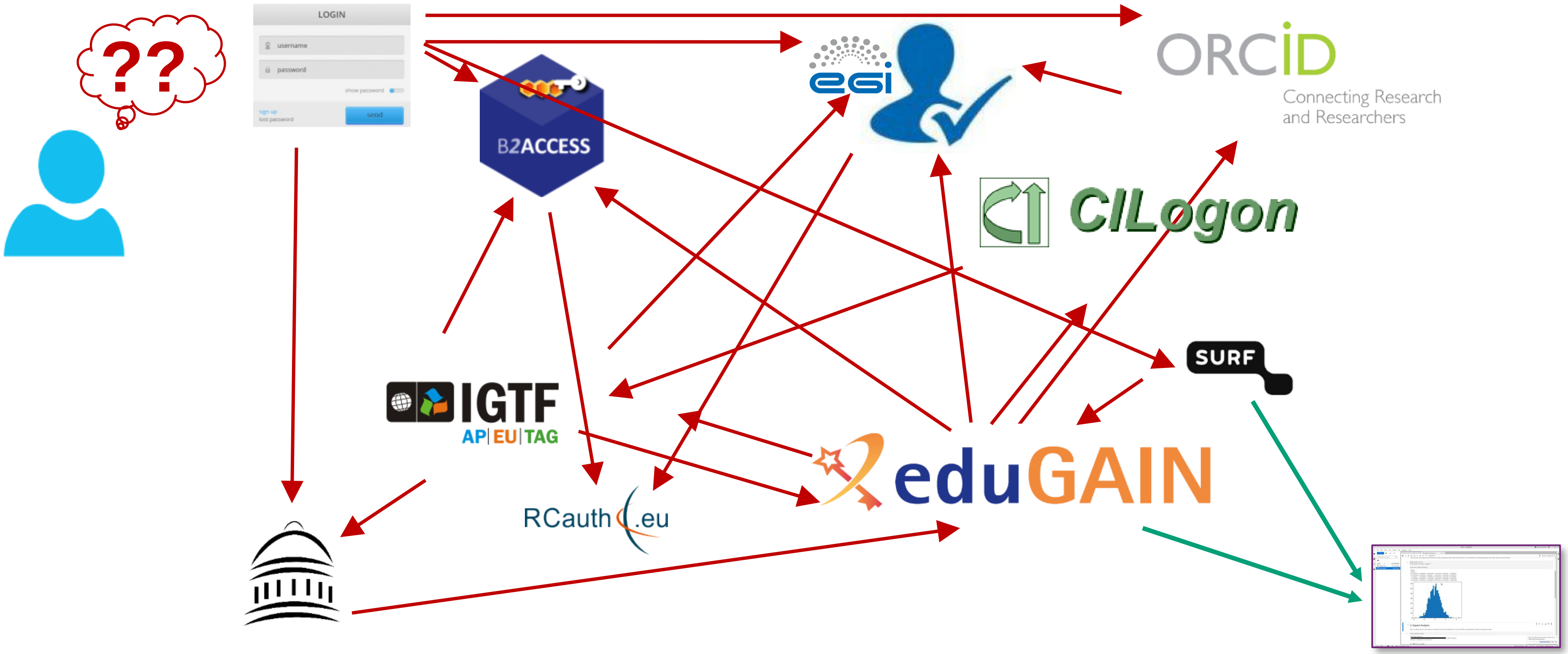


AARC-BPA-2019

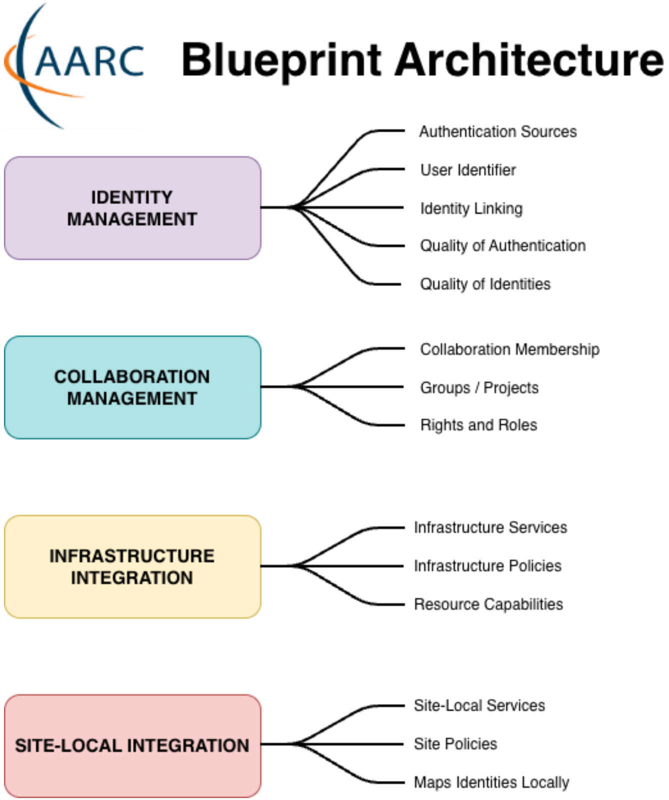


AARC-BPA-2025

Identity spaghetti: 1-loop, 2-loop and higher order diagrams



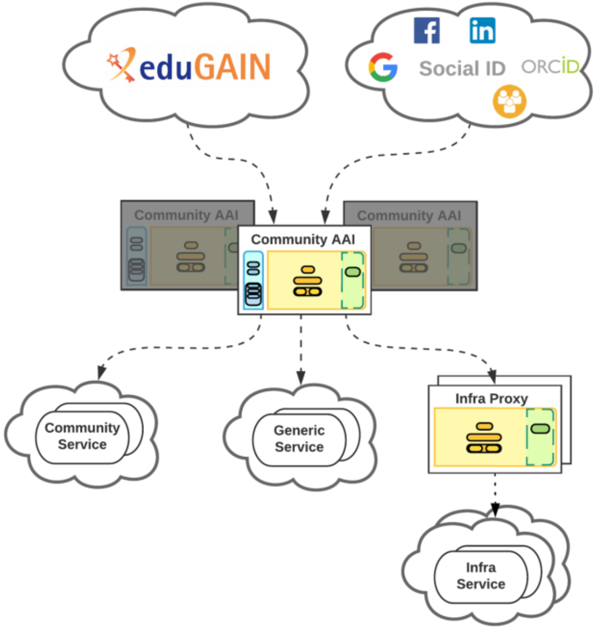
AARC Blueprint Architecture 2025: Functional Capabilities



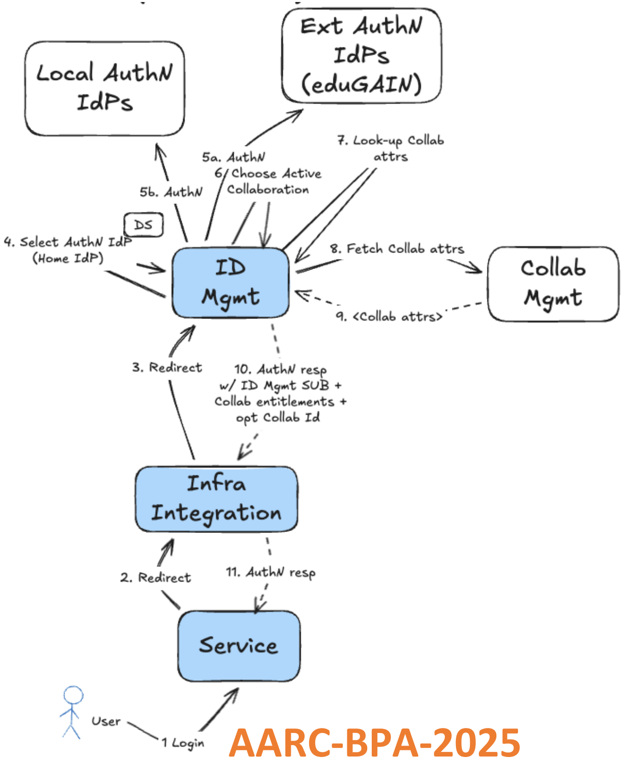
What has changed since AARC-BPA-2019?

- Added **Identity Management capability**:
  - Groups identity-related functions such as unique identifier assignment, identity assurance, authentication assurance, and identity linking
- **Community AAI → Collaboration Management**
- Added **Site-local Integration** capability:
  - Enables integration of federated users into local environments

A new way to access resources with AARC-BPA-2025

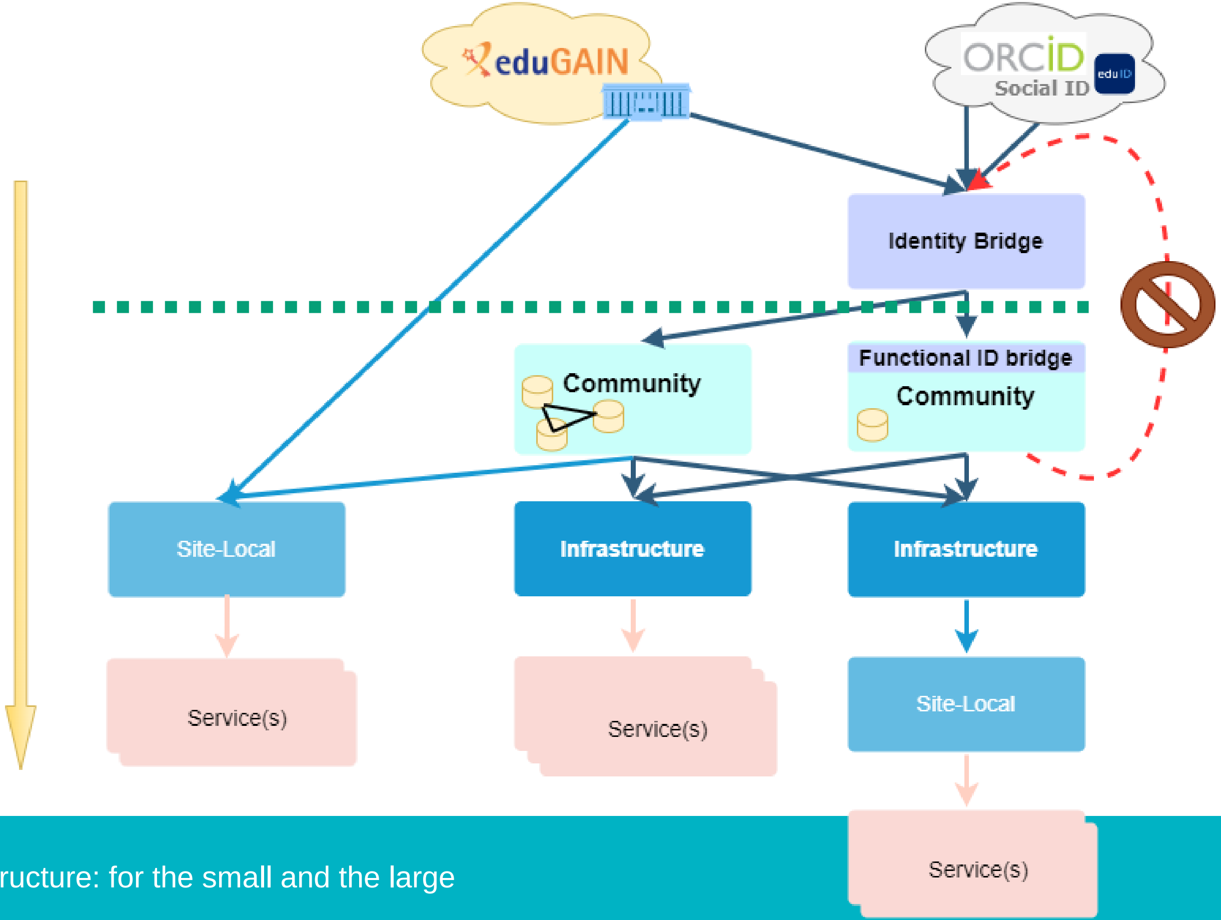


AARC-BPA-2019 aka “Community-first”

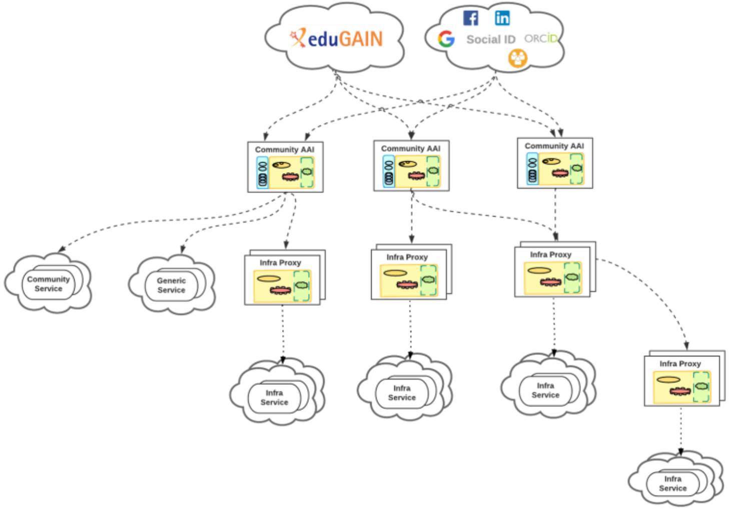


AARC-BPA-2025

Not all that is possible is allowed in the AARC BPA

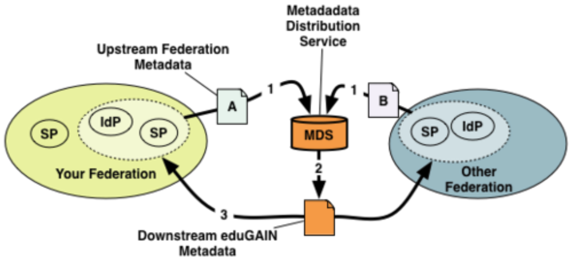


We have seen many arrows before … it needs federation!



Identity, community, infrastructure proxies and services form a ***federation of proxies***

- bilateral registration
  but then you have a scalability issue again
- meta-data distribution of trust paths
  - **OpenID Federation**
  - **SAML** meta-data


Metadadata Distribution Service
Upstream Federation Metadata
A
1
1
B
MDS
2
3
SP
IdP
SP
Your Federation
SP
IdP
Other Federation
Downstream eduGAIN Metadata

- discovery and identity provider hinting

European Open Science Cloud federation (2023 edition)

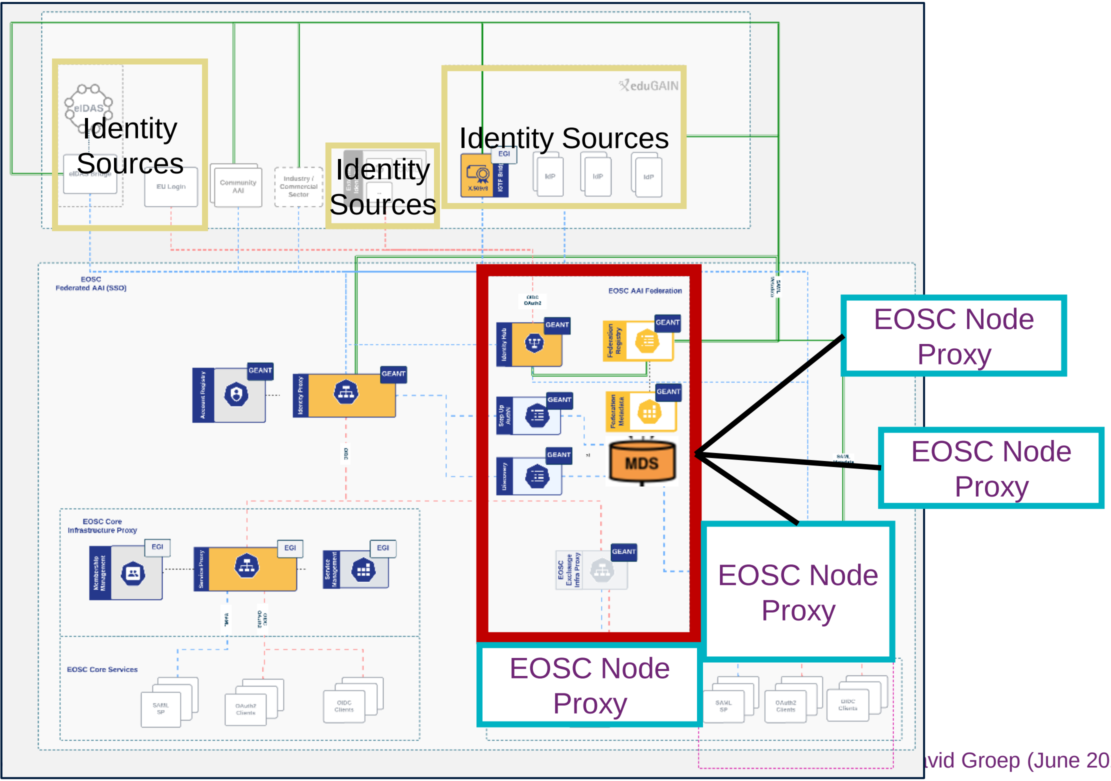


Image: EOSC AAI for the EOSC Core and Exch vid Groep (June 2023)

AARC BPA Deployment example: MyAccessID and EOSC AAI

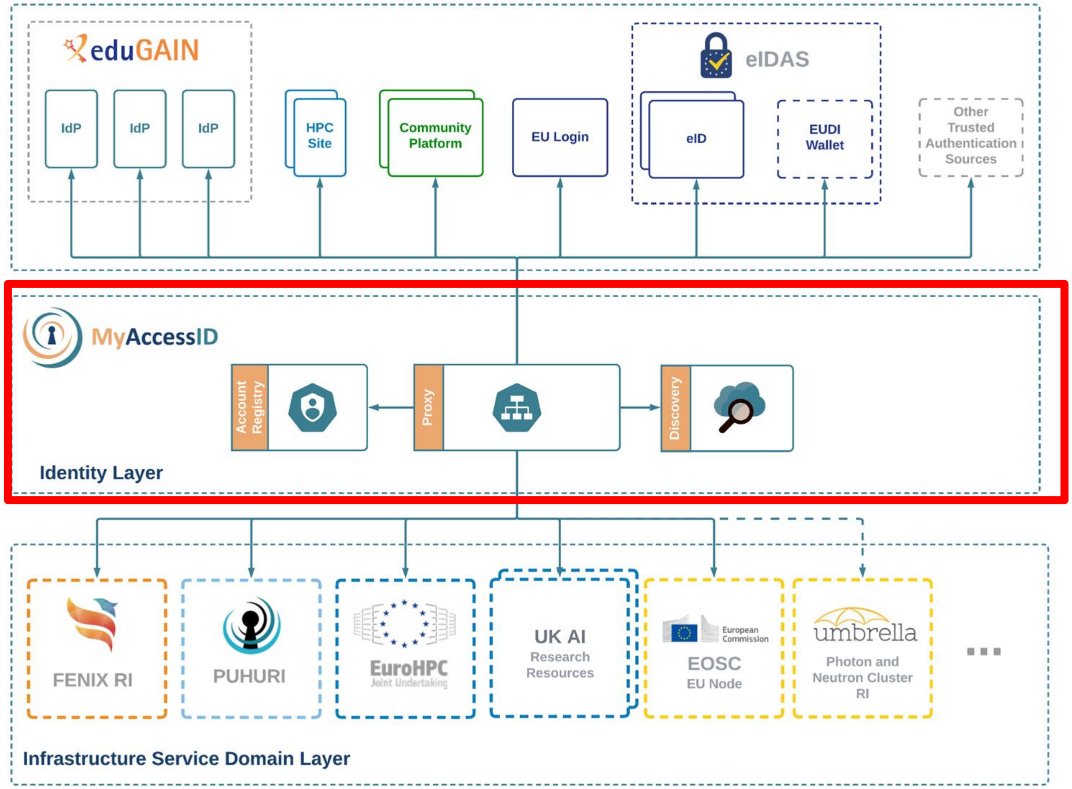
Graphics: Christos Kannelopoulos and the AARC Community

MyAccessID: A common Identity Layer for Science

- HPC Datacenters are in the process of transforming to **Infrastructure Service Providers** with **a diverse Service Portfolio**

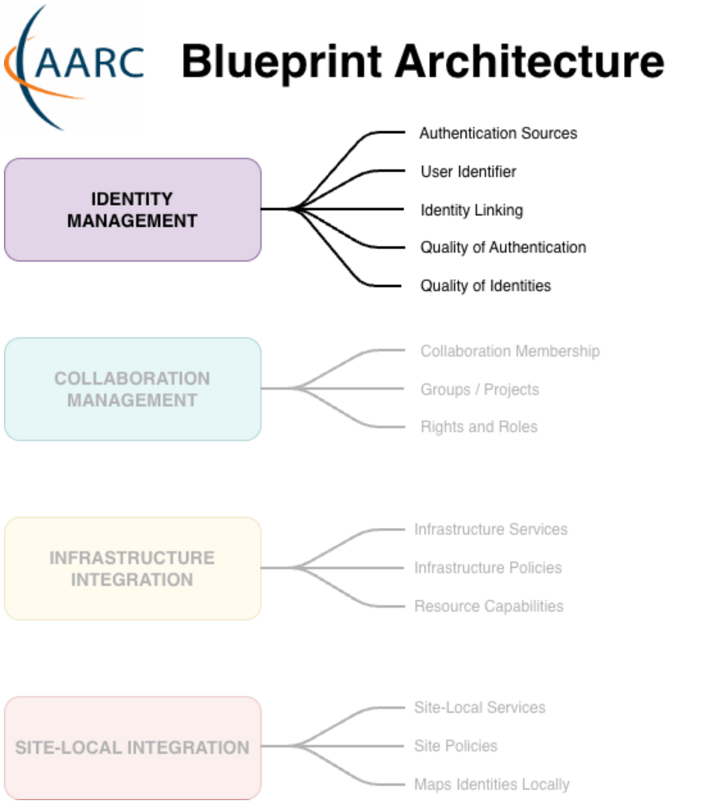
- These services become available in different administrative and policy domains, which we call **Infrastructure Service Domains**

- **A common Authentication and Authorization Infrastructure** enables uniform accessibility to scientists and engineers at European scale



Graphics: Christos Kannelopoulos and the AARC Community

MyAccessID: A common Identity Layer for Science



AARC Blueprint Architecture

IDENTITY MANAGEMENT

- Authentication Sources
- User Identifier
- Identity Linking
- Quality of Authentication
- Quality of Identities

COLLABORATION MANAGEMENT

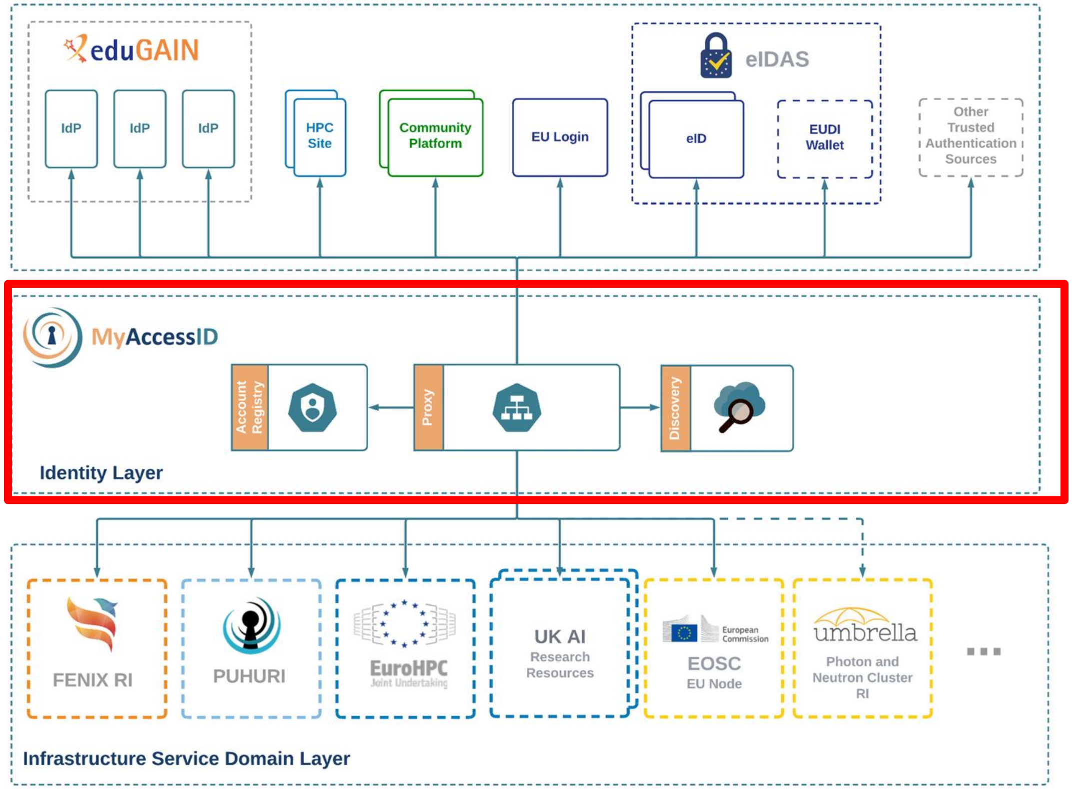
- Collaboration Membership
- Groups / Projects
- Rights and Roles

INFRASTRUCTURE INTEGRATION

- Infrastructure Services
- Infrastructure Policies
- Resource Capabilities

SITE-LOCAL INTEGRATION

- Site-Local Services
- Site Policies
- Maps Identities Locally



eduGAIN: IdP, IdP, IdP

HPC Site

Community Platform

EU Login

eIDAS: eID, EUDI Wallet

Other Trusted Authentication Sources

MyAccessID

Account Registry

Proxy

Discovery

Identity Layer

FENIX RI

PUHURI

EuroHPC Joint Undertaking

UK AI Research Resources

European Commission EOSC EU Node

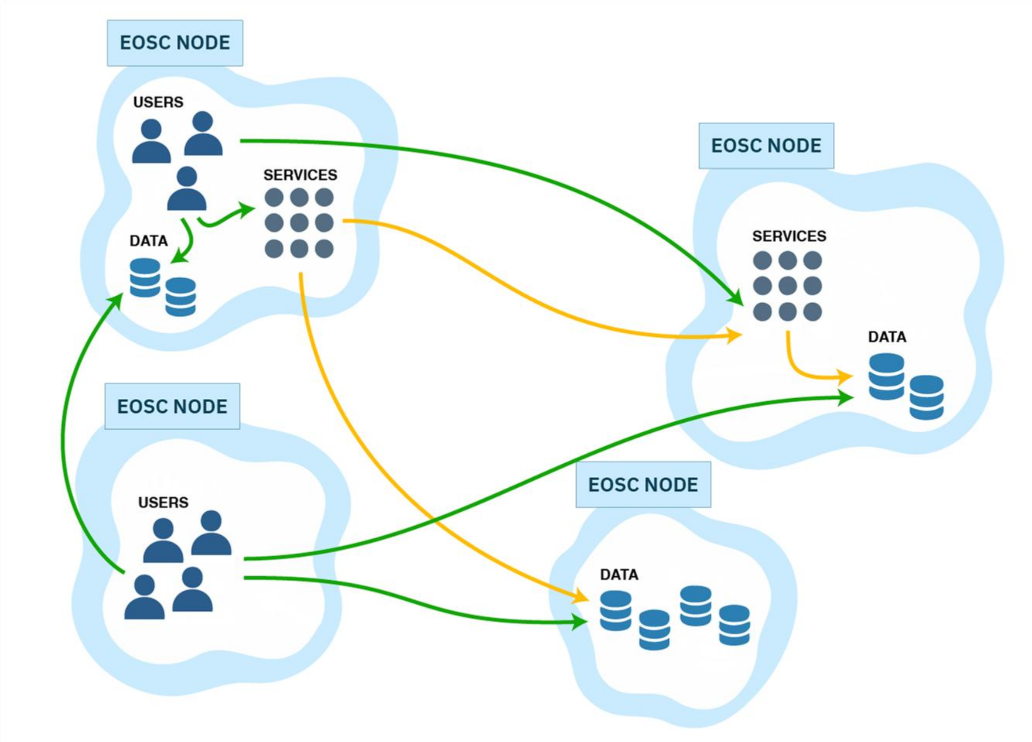
umbrella Photon and Neutron Cluster RI

...

Infrastructure Service Domain Layer

Graphics: Christos Kannelopoulos and the AARC Community

European Open Science Cloud



🟩 interactive login of users

🟨 service to service

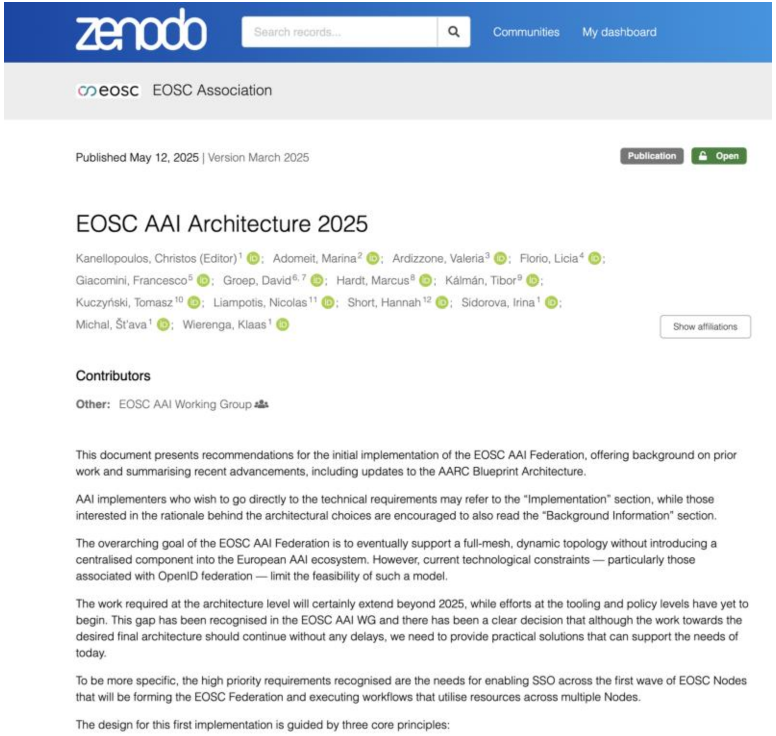
Initial EOSC AAI Use Cases

- Single Sign On Across Nodes
- Cross Node Workflows

* See Licia’s presentation from the FIM4R session

Graphics: Christos Kannelopoulos and the AARC Community

EOSC AAI Architecture - March 2025 version



Published May 12, 2025 | Version March 2025

EOSC AAI Architecture 2025

Kanellopoulos, Christos (Editor)[1]; Adomeit, Marina[2]; Ardizzone, Valeria[3]; Florio, Licia[4]; Giacomini, Francesco[5]; Groep, David[6, 7]; Hardt, Marcus[8]; Kálmán, Tibor[9]; Kuczyński, Tomasz[10]; Liampotis, Nicolas[11]; Short, Hannah[12]; Sidorova, Irina[1]; Michal, Šťava[1]; Wierenga, Klaas[1]

Contributors

Other: EOSC AAI Working Group

This document presents recommendations for the initial implementation of the EOSC AAI Federation, offering background on prior work and summarising recent advancements, including updates to the AARC Blueprint Architecture.

AAI implementers who wish to go directly to the technical requirements may refer to the “Implementation” section, while those interested in the rationale behind the architectural choices are encouraged to also read the “Background Information” section.

The overarching goal of the EOSC AAI Federation is to eventually support a full-mesh, dynamic topology without introducing a centralised component into the European AAI ecosystem. However, current technological constraints — particularly those associated with OpenID federation — limit the feasibility of such a model.

The work required at the architecture level will certainly extend beyond 2025, while efforts at the tooling and policy levels have yet to begin. This gap has been recognised in the EOSC AAI WG and there has been a clear decision that although the work towards the desired final architecture should continue without any delays, we need to provide practical solutions that can support the needs of today.

To be more specific, the high priority requirements recognised are the needs for enabling SSO across the first wave of EOSC Nodes that will be forming the EOSC Federation and executing workflows that utilise resources across multiple Nodes.

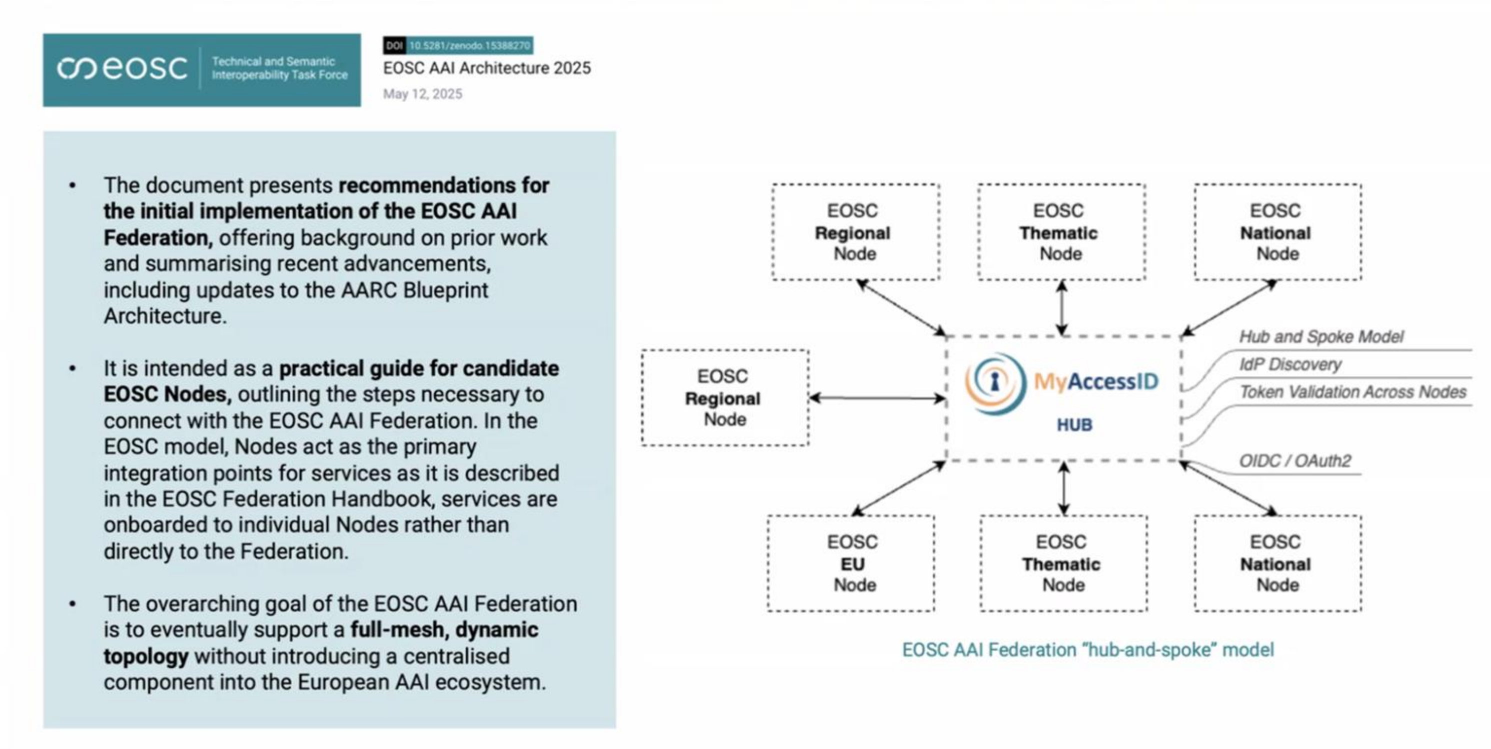
The design for this first implementation is guided by three core principles:

The EOSC AAI Architecture profiles the AARC Blueprint Architecture for EOSC

https://doi.org/10.5281/zenodo.15388270

Graphics: Christos Kannelopoulos and the AARC Community

EOSC AAI Federation



eosc | Technical and Semantic Interoperability Task Force

DOI 10.5281/zenodo.15388270

EOSC AAI Architecture 2025

May 12, 2025

- The document presents **recommendations for the initial implementation of the EOSC AAI Federation,** offering background on prior work and summarising recent advancements, including updates to the AARC Blueprint Architecture.
- It is intended as a **practical guide for candidate EOSC Nodes,** outlining the steps necessary to connect with the EOSC AAI Federation. In the EOSC model, Nodes act as the primary integration points for services as it is described in the EOSC Federation Handbook, services are onboarded to individual Nodes rather than directly to the Federation.
- The overarching goal of the EOSC AAI Federation is to eventually support a **full-mesh, dynamic topology** without introducing a centralised component into the European AAI ecosystem.

EOSC **Regional** Node

EOSC **Thematic** Node

EOSC **National** Node

EOSC **Regional** Node

MyAccessID

HUB

Hub and Spoke Model

IdP Discovery

Token Validation Across Nodes

OIDC / OAuth2

EOSC **EU** Node

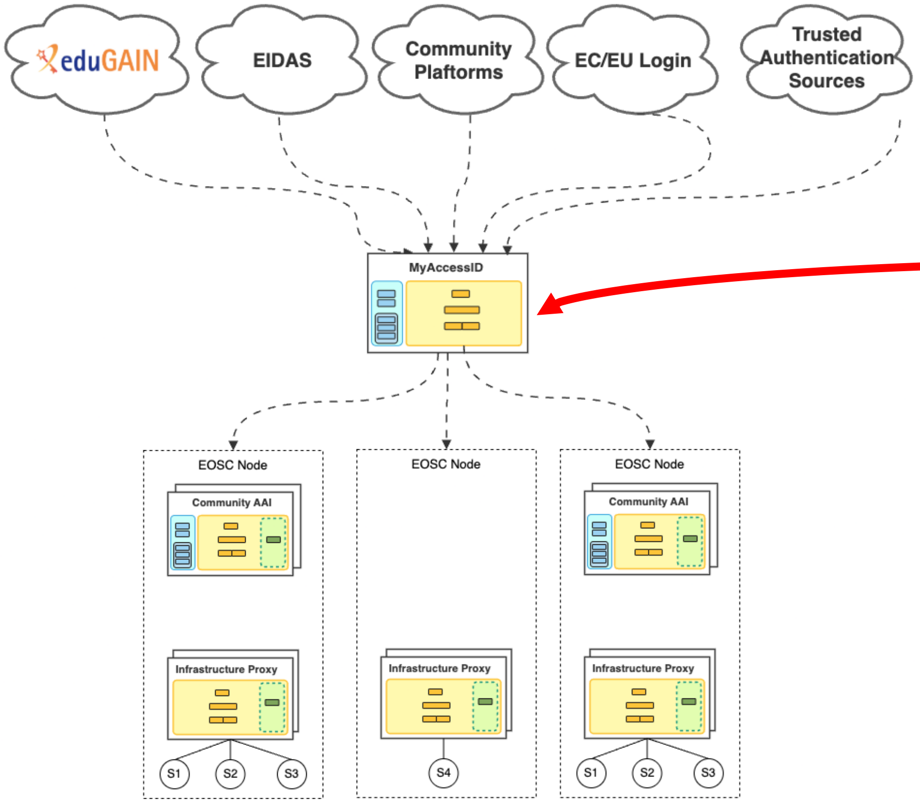
EOSC **Thematic** Node

EOSC **National** Node

EOSC AAI Federation "hub-and-spoke" model

Graphics: Christos Kannelopoulos and the AARC Community

EOSC AAI Federation - The Identity Layer



Identity Layer

Published May 12, 2025 | Version March 2025

EOSC AAI Architecture 2025

Kanellopoulos, Christos (Editor)[1]; Adamelli, Marina[2]; Ardizzone, Valeria[3]; Florio, Licia[4]; Giacomini, Francesco[5]; Groep, David[6, 7]; Hardt, Marcus[8]; Kálmán, Tibor[9]; Kuczyński, Tomasz[10]; Liampotis, Nicolas[11]; Short, Hannah[12]; Sidorova, Irina[1]; Michal, Šrāva[1]; Wierenga, Klaas[1]

Contributors

Other: EOSC AAI Working Group

This document presents recommendations for the initial implementation of the EOSC AAI Federation, offering background on prior work and summarising recent advancements, including updates to the AARC Blueprint Architecture.

AAI implementers who wish to go directly to the technical requirements may refer to the "Implementation" section, while those interested in the rationale behind the architectural choices are encouraged to also read the "Background Information" section.

The overarching goal of the EOSC AAI Federation is to eventually support a full-mesh, dynamic topology without introducing a centralised component into the European AAI ecosystem. However, current technological constraints — particularly those associated with OpenID federation — limit the feasibility of such a model.

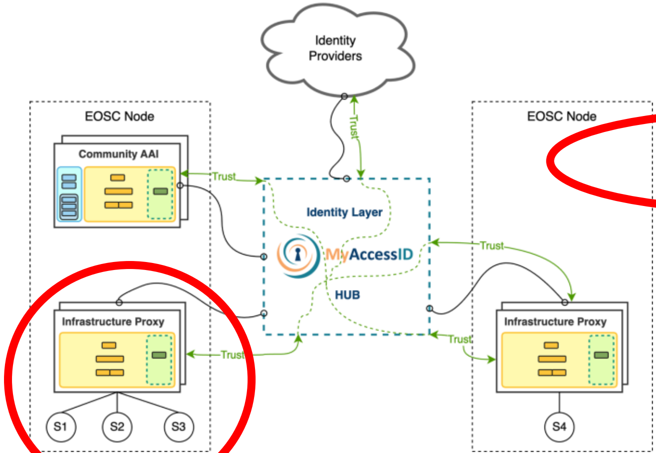
The work required at the architecture level will certainly extend beyond 2025, while efforts at the tooling and policy levels have yet to begin. This gap has been recognised in the EOSC AAI WG and there has been a clear decision that although the work towards the desired final architecture should continue without any delays, we need to provide practical solutions that can support the needs of today.

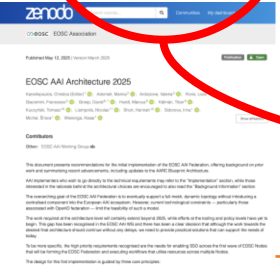
To be more specific, the high priority requirements recognised are the needs for enabling SSO across the first wave of EOSC Nodes that will be forming the EOSC Federation and executing workflows that utilise resources across multiple Nodes.

The design for this first implementation is guided by three core principles:

Graphics: Christos Kannelopoulos and the AARC Community

Infrastructure Proxy





Published May 12, 2025 | Version March 2025

EOSC AAI Architecture 2025

Kanellopoulos, Christos (Editor)[1]; Adamek, Marina[2]; Ardizzone, Valeria[3]; Florio, Lisa; Giacomini, Francesco[5]; Groep, David[6,7]; Hardt, Marcus[8]; Kálmán, Tibor[9]; Kuczyński, Tomasz[10]; Liampotis, Nicolas[11]; Short, Hannah[12]; Sidorova, Irina[1]; Michal, Šťava[1]; Wierenga, Klaas[1]

Contributors

Other: EOSC AAI Working Group

This document presents recommendations for the initial implementation of the EOSC AAI Federation, offering background on prior work and summarising recent advancements, including updates to the AARC Blueprint Architecture.

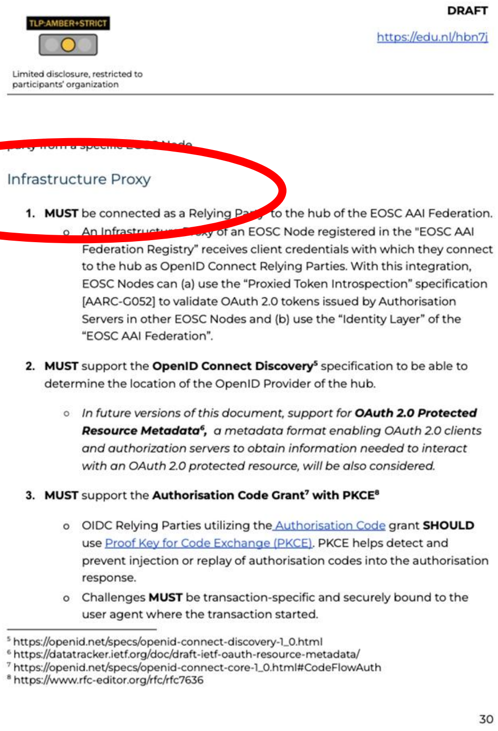
AAI implementers who wish to go directly to the technical requirements may refer to the "Implementation" section, while those interested in the rationale behind the architectural choices are encouraged to also read the "Background Information" section.

The overarching goal of the EOSC AAI Federation is to eventually support a full-mesh, dynamic topology without introducing a centralised component into the European AAI ecosystem. However, current technological constraints — particularly those associated with OpenID federation — limit the feasibility of such a model.

The work required at the architecture level will certainly extend beyond 2025, while efforts at the tooling and policy levels have yet to begin. This gap has been recognised in the EOSC AAI WG and there has been a clear decision that although the work towards the desired final architecture should continue without any delays, we need to provide practical solutions that can support the needs of today.

To be more specific, the high priority requirements recognised are the needs for enabling SSO across the first wave of EOSC Nodes that will be forming the EOSC Federation and executing workflows that utilise resources across multiple Nodes.

The design for this first implementation is guided by three core principles:



TLP:AMBER+STRICT

Limited disclosure, restricted to participants' organization

DRAFT

https://edu.nl/hbn7j

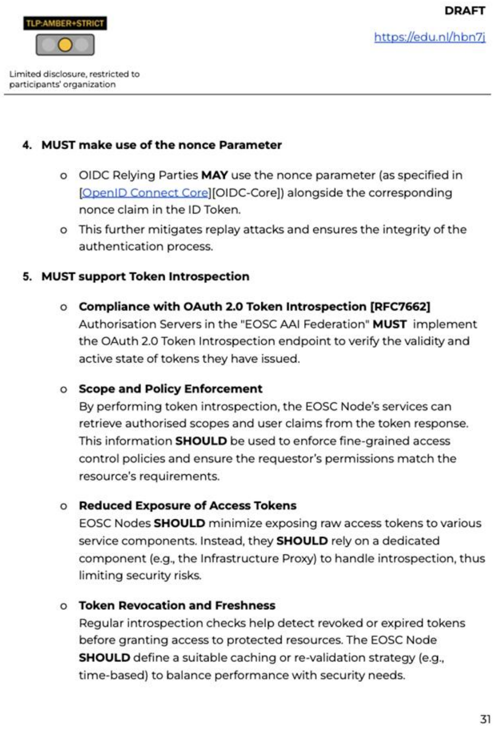
party from a specific EOSC Node.

Infrastructure Proxy

1. **MUST** be connected as a Relying Party to the hub of the EOSC AAI Federation.
   - An Infrastructure Proxy of an EOSC Node registered in the "EOSC AAI Federation Registry" receives client credentials with which they connect to the hub as OpenID Connect Relying Parties. With this integration, EOSC Nodes can (a) use the "Proxied Token Introspection" specification [AARC-G052] to validate OAuth 2.0 tokens issued by Authorisation Servers in other EOSC Nodes and (b) use the "Identity Layer" of the "EOSC AAI Federation".
2. **MUST** support the **OpenID Connect Discovery**[5] specification to be able to determine the location of the OpenID Provider of the hub.
   - *In future versions of this document, support for* ***OAuth 2.0 Protected Resource Metadata***[6]*, a metadata format enabling OAuth 2.0 clients and authorization servers to obtain information needed to interact with an OAuth 2.0 protected resource, will be also considered.*
3. **MUST** support the **Authorisation Code Grant**[7] **with PKCE**[8]
   - OIDC Relying Parties utilizing the Authorisation Code grant **SHOULD** use Proof Key for Code Exchange (PKCE). PKCE helps detect and prevent injection or replay of authorisation codes into the authorisation response.
   - Challenges **MUST** be transaction-specific and securely bound to the user agent where the transaction started.

[5] https://openid.net/specs/openid-connect-discovery-1_0.html
[6] https://datatracker.ietf.org/doc/draft-ietf-oauth-resource-metadata/
[7] https://openid.net/specs/openid-connect-core-1_0.html#CodeFlowAuth
[8] https://www.rfc-editor.org/rfc/rfc7636

30



TLP:AMBER+STRICT

Limited disclosure, restricted to participants' organization

DRAFT

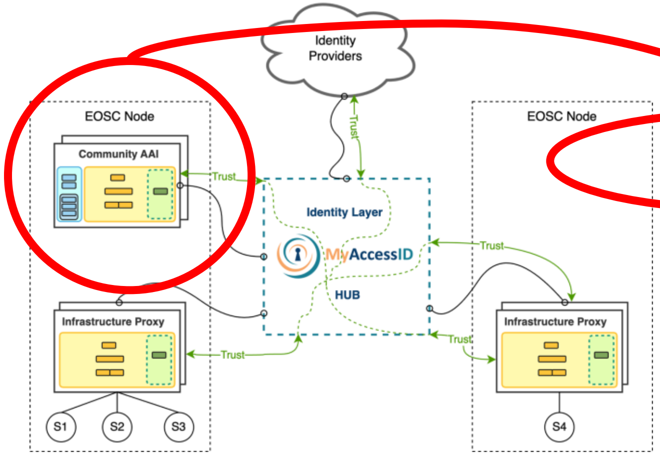
https://edu.nl/hbn7j

4. **MUST make use of the nonce Parameter**
   - OIDC Relying Parties **MAY** use the nonce parameter (as specified in [OpenID Connect Core][OIDC-Core]) alongside the corresponding nonce claim in the ID Token.
   - This further mitigates replay attacks and ensures the integrity of the authentication process.
5. **MUST support Token Introspection**
   - **Compliance with OAuth 2.0 Token Introspection [RFC7662]**
     Authorisation Servers in the "EOSC AAI Federation" **MUST** implement the OAuth 2.0 Token Introspection endpoint to verify the validity and active state of tokens they have issued.
   - **Scope and Policy Enforcement**
     By performing token introspection, the EOSC Node's services can retrieve authorised scopes and user claims from the token response. This information **SHOULD** be used to enforce fine-grained access control policies and ensure the requestor's permissions match the resource's requirements.
   - **Reduced Exposure of Access Tokens**
     EOSC Nodes **SHOULD** minimize exposing raw access tokens to various service components. Instead, they **SHOULD** rely on a dedicated component (e.g., the Infrastructure Proxy) to handle introspection, thus limiting security risks.
   - **Token Revocation and Freshness**
     Regular introspection checks help detect revoked or expired tokens before granting access to protected resources. The EOSC Node **SHOULD** define a suitable caching or re-validation strategy (e.g., time-based) to balance performance with security needs.

31

Graphics: Christos Kannelopoulos and the AARC Community

Community AAI



Published May 12, 2025 | Version March 2025

EOSC AAI Architecture 2025

Kanellopoulos, Christos (Editor)[1]; Adamali, Marina[2]; Ardizzone, Valeria[3]; Florio, Licia[4]; Giacomini, Francesco[5]; Groep, David[6, 7]; Hardt, Marcus[8]; Kálmán, Tibor[9]; Kuczyński, Tomasz[10]; Liampotis, Nicolas[11]; Short, Hannah[12]; Sidorova, Irina[1]; Michal, Štava[1]; Weeenga, Klaas[1]

Contributors

Other: EOSC AAI Working Group

This document presents recommendations for the initial implementation of the EOSC AAI Federation, offering background on prior work and summarising recent advancements, including updates to the AARC Blueprint Architecture.

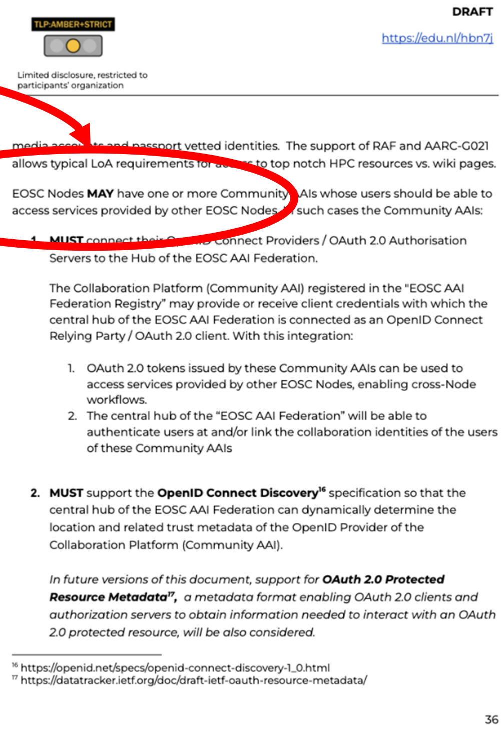
AAI implementers who wish to go directly to the technical requirements may refer to the "Implementation" section, while those interested in the rationale behind the architectural choices are encouraged to also read the "Background Information" section.

The overarching goal of the EOSC AAI Federation is to eventually support a full-mesh, dynamic topology without introducing a centralised component into the European AAI ecosystem. However, current technological constraints — particularly those associated with OpenID federation — limit the feasibility of such a model.

The work required at the architecture level will certainly extend beyond 2025, while efforts at the tooling and policy levels have yet to begin. This gap has been recognised in the EOSC AAI WG and there has been a clear decision that although the work towards the desired final architecture should continue without any delays, we need to provide practical solutions that can support the needs of today.

To be more specific, the high priority requirements recognised are the needs for enabling SSO across the first wave of EOSC Nodes that will be forming the EOSC Federation and executing workflows that utilise resources across multiple Nodes.

The design for this first implementation is guided by three core principles:



TLP:AMBER+STRICT

Limited disclosure, restricted to participants' organization

DRAFT

https://edu.nl/hbn7j

media accounts and passport vetted identities. The support of RAF and AARC-G021 allows typical LoA requirements for access to top notch HPC resources vs. wiki pages.

EOSC Nodes **MAY** have one or more Community AAIs whose users should be able to access services provided by other EOSC Nodes. In such cases the Community AAIs:

1. **MUST** connect their OpenID Connect Providers / OAuth 2.0 Authorisation Servers to the Hub of the EOSC AAI Federation.

   The Collaboration Platform (Community AAI) registered in the "EOSC AAI Federation Registry" may provide or receive client credentials with which the central hub of the EOSC AAI Federation is connected as an OpenID Connect Relying Party / OAuth 2.0 client. With this integration:

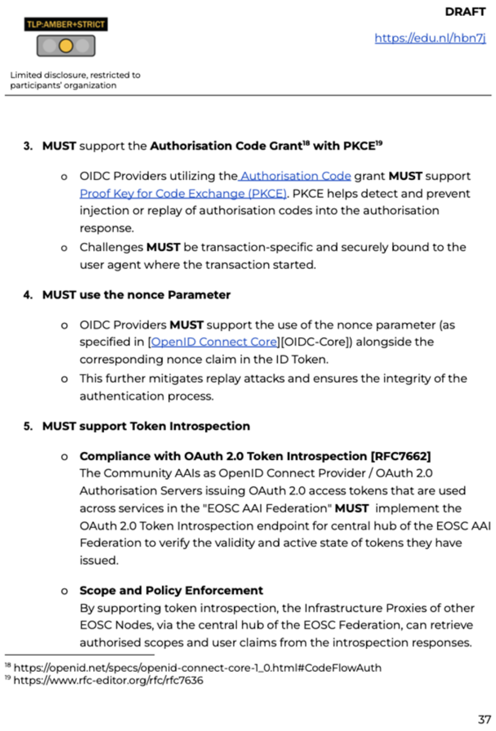
   1. OAuth 2.0 tokens issued by these Community AAIs can be used to access services provided by other EOSC Nodes, enabling cross-Node workflows.
   2. The central hub of the "EOSC AAI Federation" will be able to authenticate users at and/or link the collaboration identities of the users of these Community AAIs

2. **MUST** support the **OpenID Connect Discovery**[16] specification so that the central hub of the EOSC AAI Federation can dynamically determine the location and related trust metadata of the OpenID Provider of the Collaboration Platform (Community AAI).

   *In future versions of this document, support for **OAuth 2.0 Protected Resource Metadata**[17], a metadata format enabling OAuth 2.0 clients and authorization servers to obtain information needed to interact with an OAuth 2.0 protected resource, will be also considered.*

[16] https://openid.net/specs/openid-connect-discovery-1_0.html
[17] https://datatracker.ietf.org/doc/draft-ietf-oauth-resource-metadata/

36



TLP:AMBER+STRICT

Limited disclosure, restricted to participants' organization

DRAFT

https://edu.nl/hbn7j

3. **MUST** support the **Authorisation Code Grant**[18] **with PKCE**[19]
   - OIDC Providers utilizing the Authorisation Code grant **MUST** support Proof Key for Code Exchange (PKCE). PKCE helps detect and prevent injection or replay of authorisation codes into the authorisation response.
   - Challenges **MUST** be transaction-specific and securely bound to the user agent where the transaction started.

4. **MUST use the nonce Parameter**
   - OIDC Providers **MUST** support the use of the nonce parameter (as specified in [OpenID Connect Core][OIDC-Core]) alongside the corresponding nonce claim in the ID Token.
   - This further mitigates replay attacks and ensures the integrity of the authentication process.

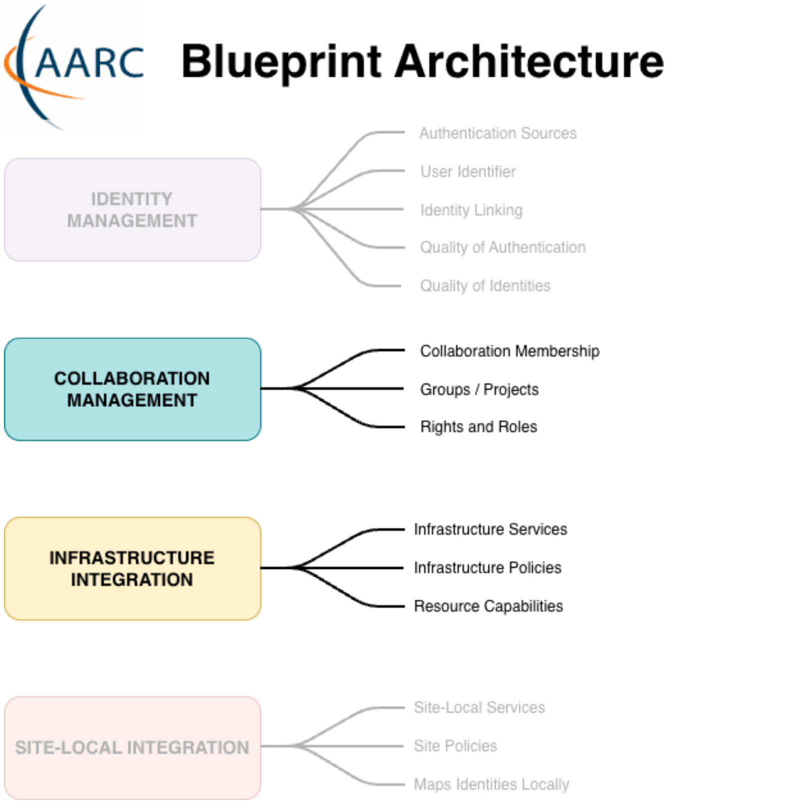
5. **MUST support Token Introspection**
   - **Compliance with OAuth 2.0 Token Introspection [RFC7662]**
     The Community AAIs as OpenID Connect Provider / OAuth 2.0 Authorisation Servers issuing OAuth 2.0 access tokens that are used across services in the "EOSC AAI Federation" **MUST** implement the OAuth 2.0 Token Introspection endpoint for central hub of the EOSC AAI Federation to verify the validity and active state of tokens they have issued.
   - **Scope and Policy Enforcement**
     By supporting token introspection, the Infrastructure Proxies of other EOSC Nodes, via the central hub of the EOSC Federation, can retrieve authorised scopes and user claims from the introspection responses.

[18] https://openid.net/specs/openid-connect-core-1_0.html#CodeFlowAuth
[19] https://www.rfc-editor.org/rfc/rfc7636

37

Graphics: Christos Kannelopoulos and the AARC Community

EOSC EU Node



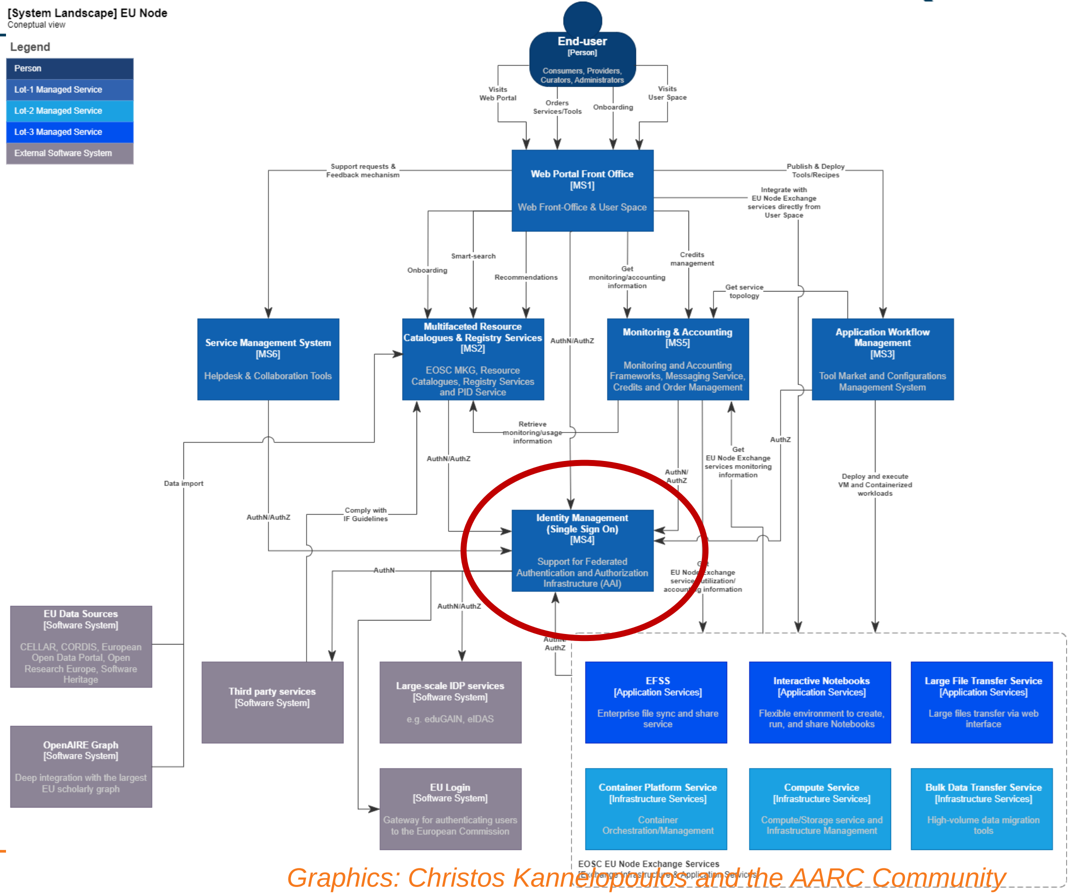
AARC **Blueprint Architecture**

IDENTITY MANAGEMENT
- Authentication Sources
- User Identifier
- Identity Linking
- Quality of Authentication
- Quality of Identities

COLLABORATION MANAGEMENT
- Collaboration Membership
- Groups / Projects
- Rights and Roles

INFRASTRUCTURE INTEGRATION
- Infrastructure Services
- Infrastructure Policies
- Resource Capabilities

SITE-LOCAL INTEGRATION
- Site-Local Services
- Site Policies
- Maps Identities Locally



[System Landscape] EU Node
Coneptual view

Legend
- Person
- Lot-1 Managed Service
- Lot-2 Managed Service
- Lot-3 Managed Service
- External Software System

End-user [Person]
Consumers, Providers, Curators, Administrators

Web Portal Front Office [MS1]
Web Front-Office & User Space

Service Management System [MS6]
Helpdesk & Collaboration Tools

Multifaceted Resource Catalogues & Registry Services [MS2]
EOSC MKG, Resource Catalogues, Registry Services and PID Service

Monitoring & Accounting [MS5]
Monitoring and Accounting Frameworks, Messaging Service, Credits and Order Management

Application Workflow Management [MS3]
Tool Market and Configurations Management System

Identity Management (Single Sign On) [MS4]
Support for Federated Authentication and Authorization Infrastructure (AAI)

EU Data Sources [Software System]
CELLAR, CORDIS, European Open Data Portal, Open Research Europe, Software Heritage

OpenAIRE Graph [Software System]
Deep integration with the largest EU scholarly graph

Third party services [Software System]

Large-scale IDP services [Software System]
e.g. eduGAIN, eIDAS

EU Login [Software System]
Gateway for authenticating users to the European Commission

EFSS [Application Services]
Enterprise file sync and share service

Interactive Notebooks [Application Services]
Flexible environment to create, run, and share Notebooks

Large File Transfer Service [Application Services]
Large files transfer via web interface

Container Platform Service [Infrastructure Services]
Container Orchestration/Management

Compute Service [Infrastructure Services]
Compute/Storage service and Infrastructure Management

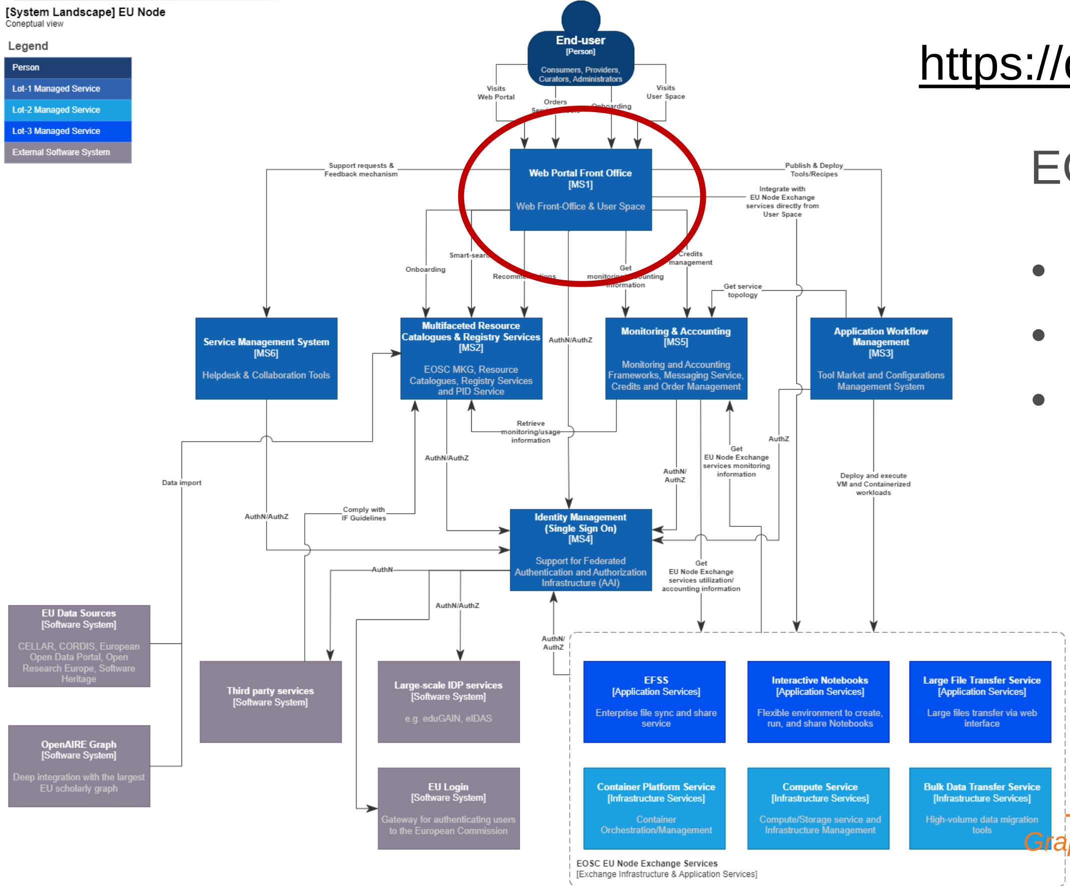
Bulk Data Transfer Service [Infrastructure Services]
High-volume data migration tools

EOSC EU Node Exchange Services
[Exchange Infrastructure & Application Services]

Visits Web Portal
Orders Services/Tools
Onboarding
Visits User Space
Support requests & Feedback mechanism
Publish & Deploy Tools/Recipes
Integrate with EU Node Exchange services directly from User Space
Onboarding
Smart-search
Recommendations
Get monitoring/accounting information
Credits management
Get service topology
AuthN/AuthZ
Retrieve monitoring/usage information
AuthN/AuthZ
AuthN/AuthZ
AuthZ
Get EU Node Exchange services monitoring information
Deploy and execute VM and Containerized workloads
Data import
AuthN/AuthZ
Comply with IF Guidelines
AuthN
AuthN/AuthZ
Get EU Node Exchange service utilization/accounting information
AuthN/AuthZ

Graphics: Christos Kanellopoulos and the AARC Community

EOSC EU Node


[System Landscape] EU Node
Coneptual view
Legend
Person
Lot-1 Managed Service
Lot-2 Managed Service
Lot-3 Managed Service
External Software System
End-user
[Person]
Consumers, Providers, Curators, Administrators
Visits Web Portal
Orders Services
Onboarding
Visits User Space
Web Portal Front Office
[MS1]
Web Front-Office & User Space
Support requests & Feedback mechanism
Publish & Deploy Tools/Recipes
Integrate with EU Node Exchange services directly from User Space
Onboarding
Smart-search
Recommendations
Get monitoring/accounting information
Credits management
Get service topology
AuthN/AuthZ
Service Management System
[MS6]
Helpdesk & Collaboration Tools
Multifaceted Resource Catalogues & Registry Services
[MS2]
EOSC MKG, Resource Catalogues, Registry Services and PID Service
Monitoring & Accounting
[MS5]
Monitoring and Accounting Frameworks, Messaging Service, Credits and Order Management
Application Workflow Management
[MS3]
Tool Market and Configurations Management System
Retrieve monitoring/usage information
Data import
AuthN/AuthZ
AuthN/AuthZ
Comply with IF Guidelines
AuthN/AuthZ
Get EU Node Exchange services monitoring information
AuthZ
Deploy and execute VM and Containerized workloads
Identity Management
(Single Sign On)
[MS4]
Support for Federated Authentication and Authorization Infrastructure (AAI)
AuthN
AuthN/AuthZ
Get EU Node Exchange services utilization/ accounting information
AuthN/ AuthZ
EU Data Sources
[Software System]
CELLAR, CORDIS, European Open Data Portal, Open Research Europe, Software Heritage
OpenAIRE Graph
[Software System]
Deep integration with the largest EU scholarly graph
Third party services
[Software System]
Large-scale IDP services
[Software System]
e.g. eduGAIN, eIDAS
EU Login
[Software System]
Gateway for authenticating users to the European Commission
EFSS
[Application Services]
Enterprise file sync and share service
Interactive Notebooks
[Application Services]
Flexible environment to create, run, and share Notebooks
Large File Transfer Service
[Application Services]
Large files transfer via web interface
Container Platform Service
[Infrastructure Services]
Container Orchestration/Management
Compute Service
[Infrastructure Services]
Compute/Storage service and Infrastructure Management
Bulk Data Transfer Service
[Infrastructure Services]
High-volume data migration tools
EOSC EU Node Exchange Services
[Exchange Infrastructure & Application Services]

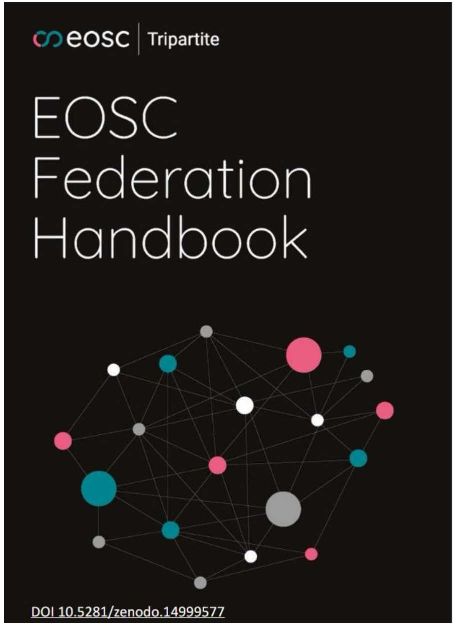
https://open-science-cloud.ec.europa.eu/

EOSC EU Node Web Portal

- Users register
- Users requests and manage projects
- Integrates with the EOSC EU Node AAI (SCIM)

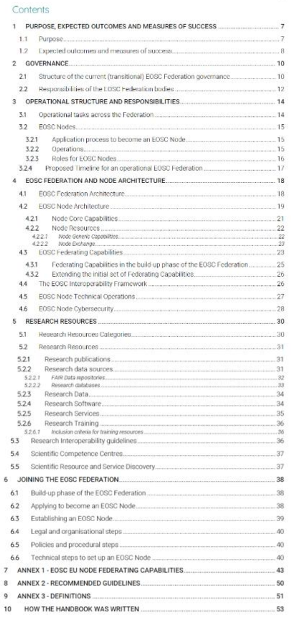
Graphics: Christos Kannelopoulos and the AARC Community

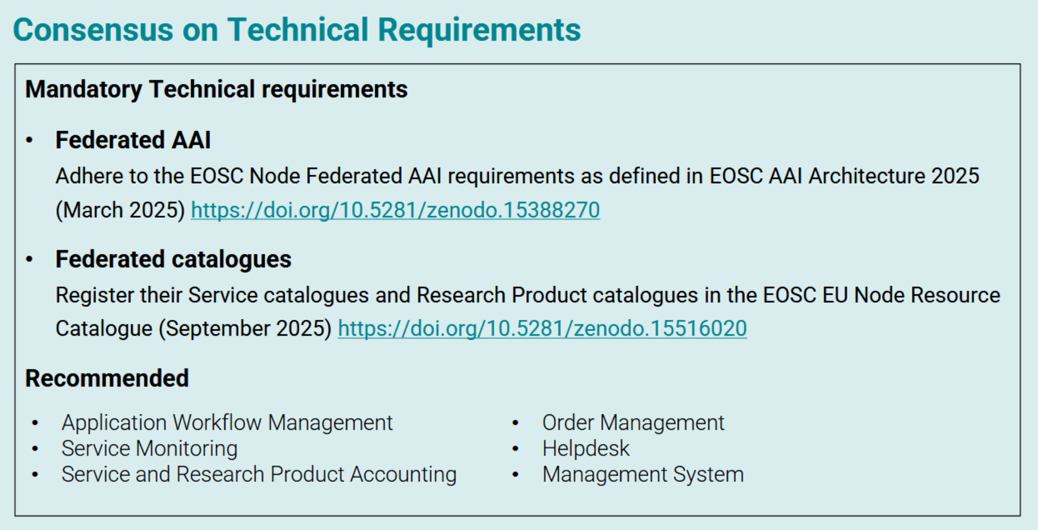
Joining the EOSC Federation


eosc | Tripartite

EOSC Federation Handbook

DOI 10.5281/zenodo.14999577


Contents


Consensus on Technical Requirements

Mandatory Technical requirements

- **Federated AAI**
  Adhere to the EOSC Node Federated AAI requirements as defined in EOSC AAI Architecture 2025 (March 2025) https://doi.org/10.5281/zenodo.15388270
- **Federated catalogues**
  Register their Service catalogues and Research Product catalogues in the EOSC EU Node Resource Catalogue (September 2025) https://doi.org/10.5281/zenodo.15516020

Recommended

- Application Workflow Management
- Service Monitoring
- Service and Research Product Accounting
- Order Management
- Helpdesk
- Management System

DOI 10.5281/zenodo.14999577

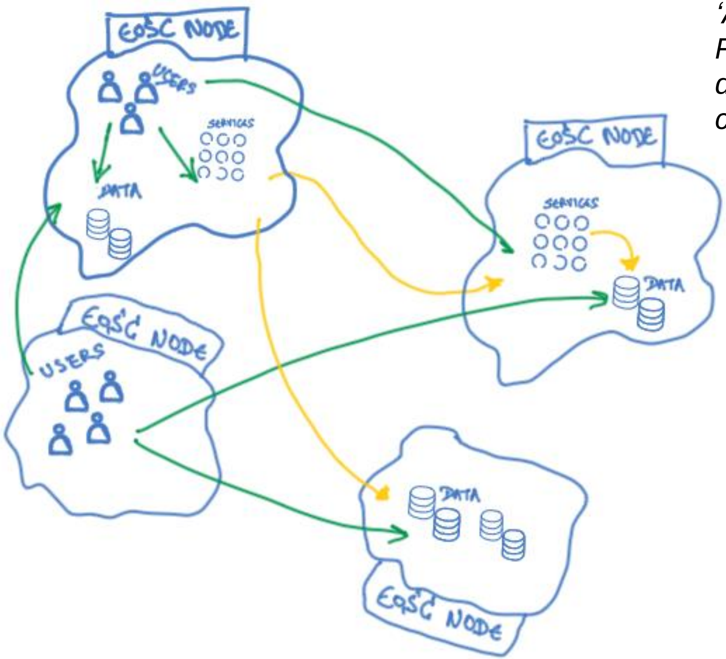
EOSC Federation Handbook

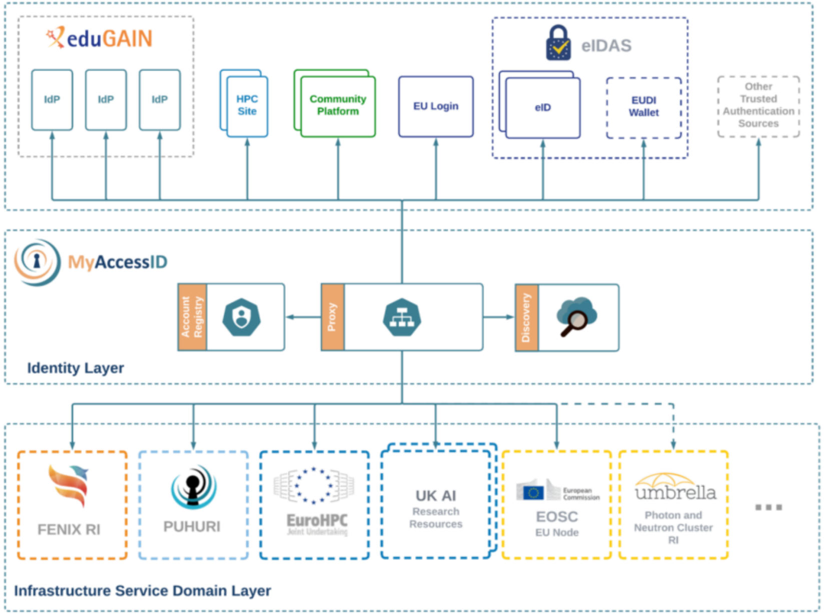
March 27, 2025

Excerpt from the slides by Bob Jones (EOSC-A) for the November 2025 EOSC Symposium

EOSC Federation structure and the evolution of its AAI

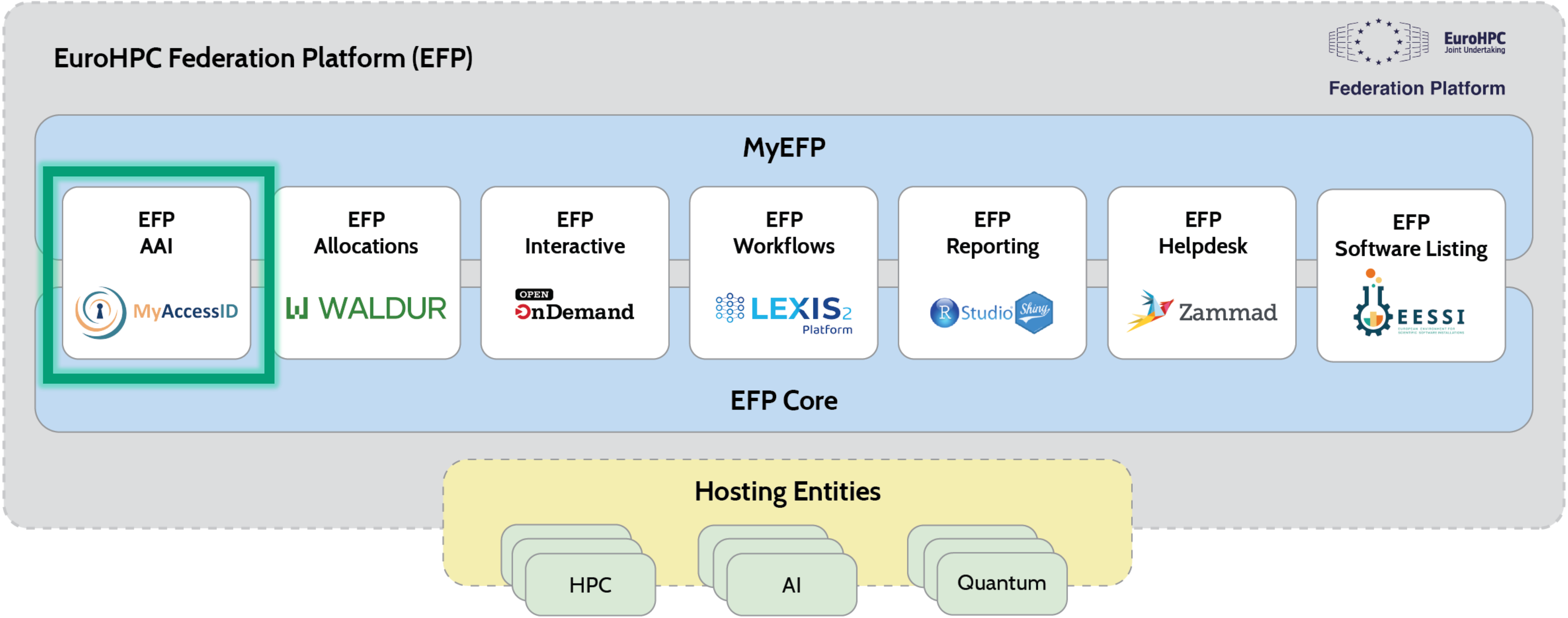
‘Adopting the “hub-and-spoke” model in the initial phase of the EOSC AAI Federation is a practical step forward, and it is implementable today, while the design and development work for the OpenID Federation and “full-mesh” topologies continues in the background in the AARC Architecture WG and the EOSC AAI WG.’



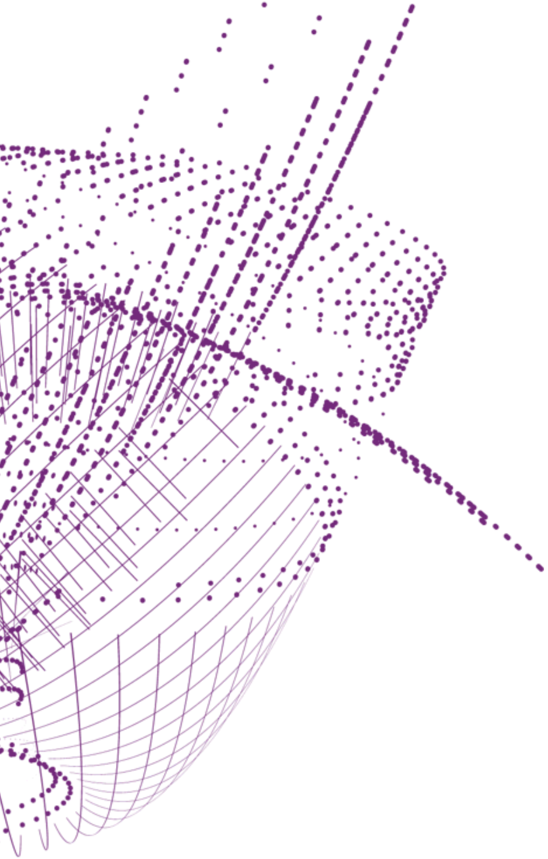


EOSC AAI Architecture 2025, https://doi.org/10.5281/zenodo.15388269
(EOSC AAI Working Group, 2025)

And the EOSC is not alone in adopting this structure

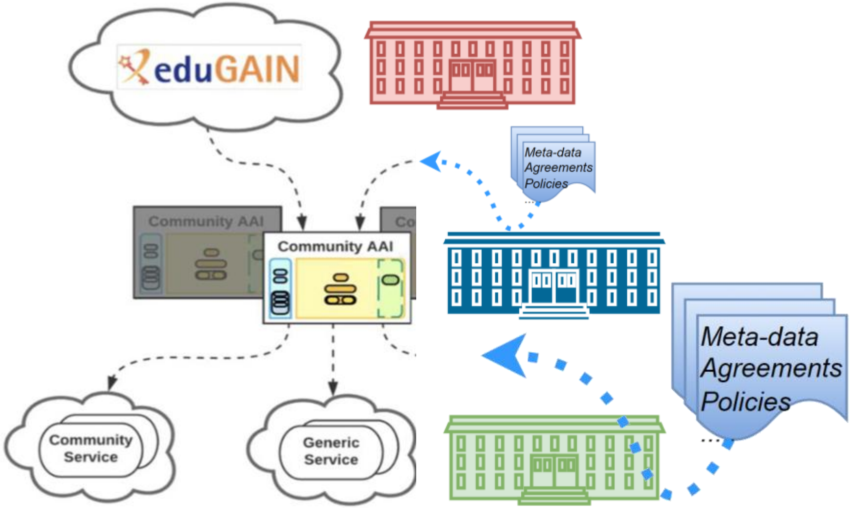


EuroHPC JU Federation Platform, see e.g. https://my-eurohpc.eu/ (image retrieved from https://my-eurohpc.eu/ February 2026)



Trust and the AARC Policy Development Kit

Now we need to ‘decorate’ the arrows with trust

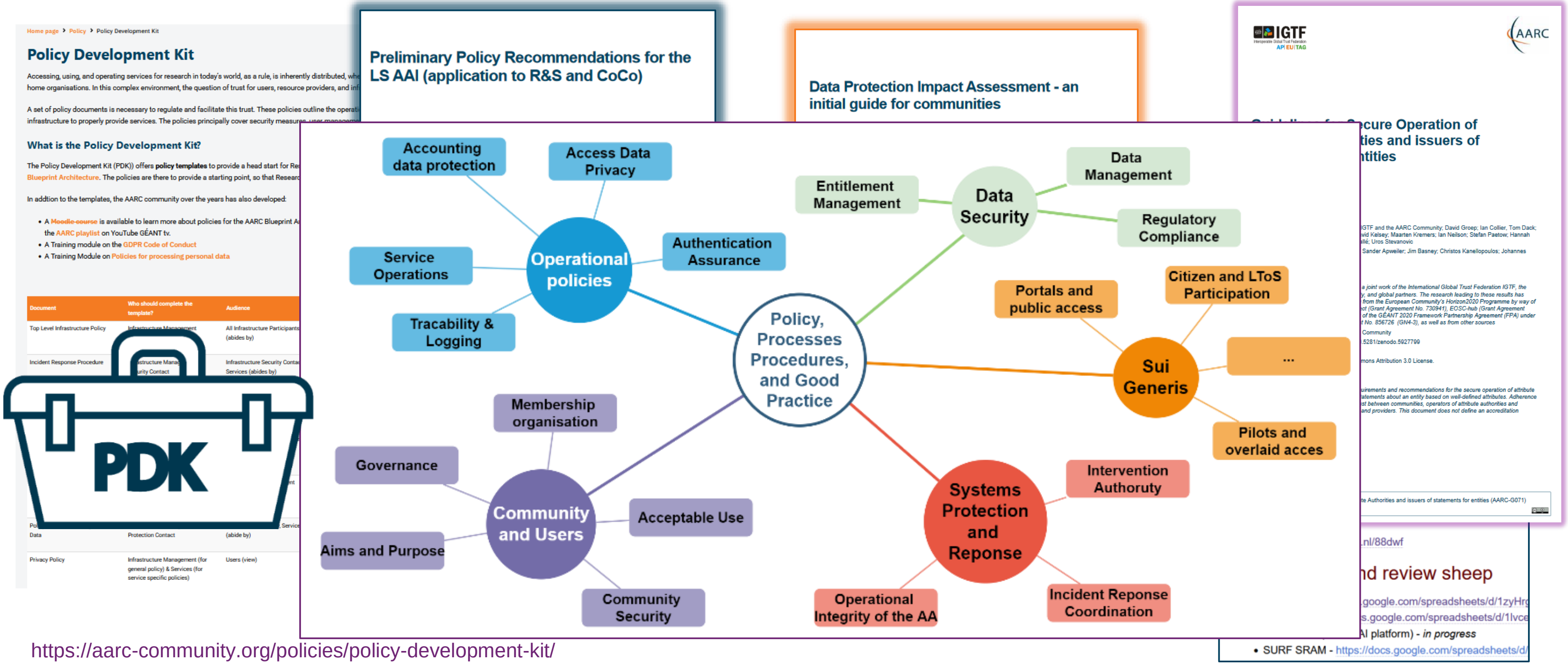


Each side of each arrow has *independent* parties

- we allow *them* to do part of the work *we* would otherwise do
- to make it easier and faster for users to perform their research
- but **we relinquish some control** beyond our organisation, our own policies, our own jurisdiction

***Why* would we trust them to do that?**

Structuring trust ‘between boxes and arrows’ is complex!



Home page › Policy › Policy Development Kit

Policy Development Kit

Accessing, using, and operating services for research in today’s world, as a rule, is inherently distributed, whe
home organisations. In this complex environment, the question of trust for users, resource providers, and inf

A set of policy documents is necessary to regulate and facilitate this trust. These policies outline the operati
infrastructure to properly provide services. The policies principally cover security measures, user manageme

What is the Policy Development Kit?

The Policy Development Kit (PDK)) offers **policy templates** to provide a head start for Re
Blueprint Architecture. The policies are there to provide a starting point, so that Researc

In addtion to the templates, the AARC community over the years has also developed:

- A ~~Moodle course~~ is available to learn more about policies for the AARC Blueprint A
  the AARC playlist on YouTube GÉANT tv.
- A Training module on the GDPR Code of Conduct
- A Training Module on Policies for processing personal data

Document	Who should complete the template?	Audience
Top Level Infrastructure Policy	Infrastructure Management	All Infrastructure Participants (abides by)
Incident Response Procedure	astructure Manag urity Contact	Infrastructure Security Contac Services (abides by)
		ent
Po Data	Protection Contact	Service (abide by)
Privacy Policy	Infrastructure Management (for general policy) & Services (for service specific policies)	Users (view)

PDK

Preliminary Policy Recommendations for the LS AAI (application to R&S and CoCo)

Data Protection Impact Assessment - an initial guide for communities

IGTF Interoperable Global Trust Federation API EU TAG

AARC

cure Operation of
ties and issuers of
ntities

IGTF and the AARC Community; David Groep; Ian Collier, Tom Dack;
vid Kelsey; Maarten Kremers; Ian Neilson; Stefan Paetow; Hannah
llé; Uros Stevanovic
Sander Apweiler; Jim Basney; Christos Kanellopoulos; Johannes

a joint work of the International Global Trust Federation IGTF, the
y, and global partners. The research leading to these results has
from the European Community’s Horizon2020 Programme by way of
ct (Grant Agreement No. 730941), EOSC-hub (Grant Agreement
of the GÉANT 2020 Framework Partnership Agreement (FPA) under
t No. 856726 (GN4-3), as well as from other sources
Community
.5281/zenodo.5927799

mons Attribution 3.0 License.

uirements and recommendations for the secure operation of attribute
tatements about an entity based on well-defined attributes. Adherence
st between communities, operators of attribute authorities and
and providers. This document does not define an accreditation

te Authorities and issuers of statements for entities (AARC-G071)

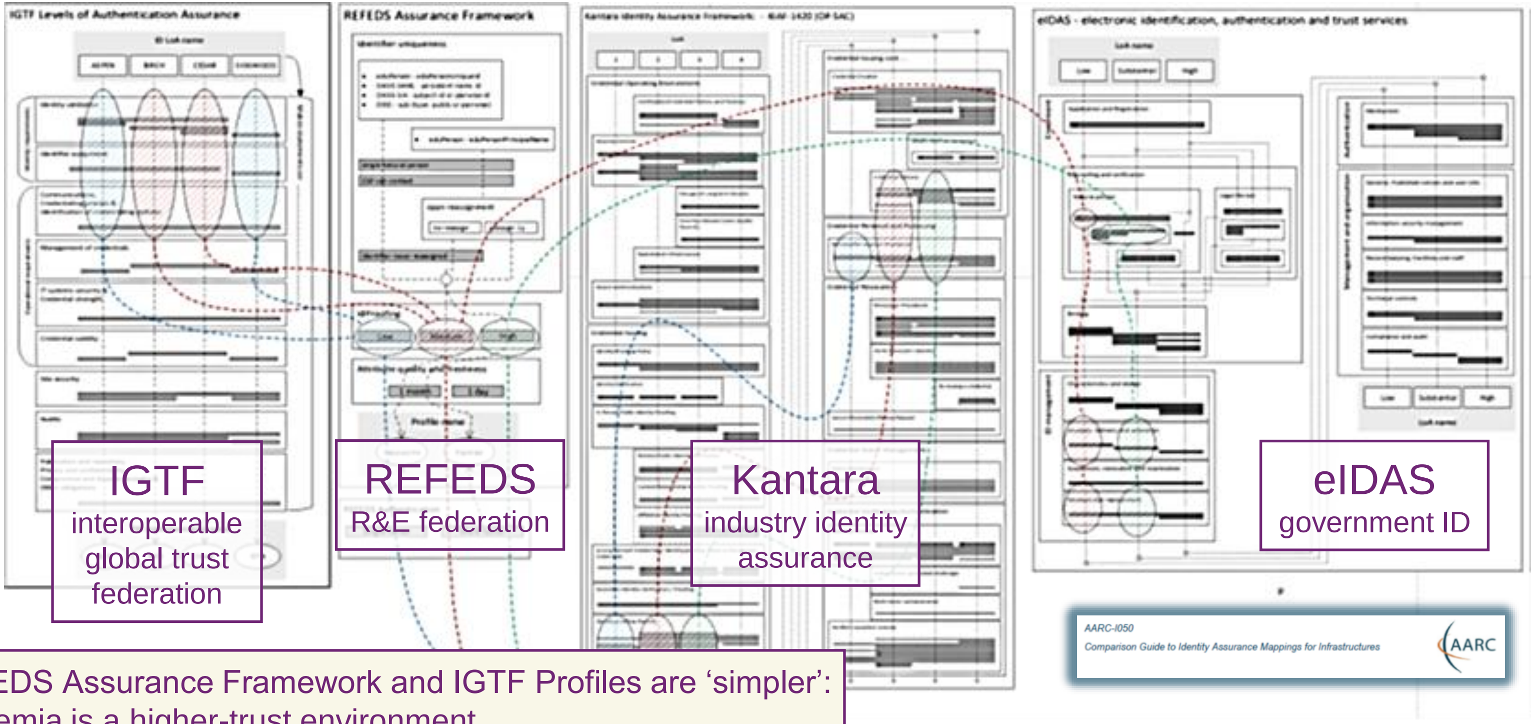
.nl/88dwf

nd review sheep

google.com/spreadsheets/d/1zyHrg
s.google.com/spreadsheets/d/1lvce
AI platform) - *in progress*
- SURF SRAM - https://docs.google.com/spreadsheets/d/

https://aarc-community.org/policies/policy-development-kit/

And even a simple ‘Who are you?’ is not always easy …


IGTF Levels of Authentication Assurance
REFEDS Assurance Framework
eIDAS - electronic identification, authentication and trust services
AARC-I050
Comparison Guide to Identity Assurance Mappings for Infrastructures
AARC

IGTF
interoperable
global trust
federation

REFEDS
R&E federation

Kantara
industry identity
assurance

eIDAS
government ID

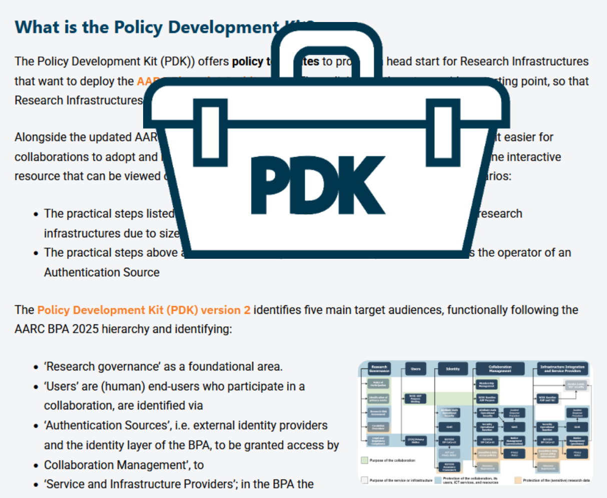
REFEDS Assurance Framework and IGTF Profiles are ‘simpler’:
academia is a higher-trust environment,
leveraging self-assessed peer review

Source: https://aarc-community.org/guidelines/aarc-i050, Ian Neilson et al.

Developing the Trust framework, guidelines and best practice for BPA proxies and interaction with research services

minimise the number of divergent policies

empower identity providers, service providers, user communities to rely on interoperable policies



What is the Policy Development Kit?

The Policy Development Kit (PDK)) offers **policy t** **tes** to pro head start for Research Infrastructures that want to deploy the AA ting point, so that Research Infrastructures

Alongside the updated AAR t easier for collaborations to adopt and ne interactive resource that can be viewed rios:

- The practical steps listed research infrastructures due to size
- The practical steps above s the operator of an Authentication Source

The Policy Development Kit (PDK) version 2 identifies five main target audiences, functionally following the AARC BPA 2025 hierarchy and identifying:

- 'Research governance' as a foundational area.
- 'Users' are (human) end-users who participate in a collaboration, are identified via
- 'Authentication Sources', i.e. external identity providers and the identity layer of the BPA, to be granted access by
- Collaboration Management', to
- 'Service and Infrastructure Providers'; in the BPA the

PDK
Research Governance
Users
Identity
Collaboration Management
Infrastructure Integration and Service Providers
Purpose of the collaboration
Purpose of the service or infrastructure
Protection of the collaboration, its users, ICT services, and resources
Protection of the (sensitive) research data

From the AARC2 infrastructure-oriented Policy Development Kit to

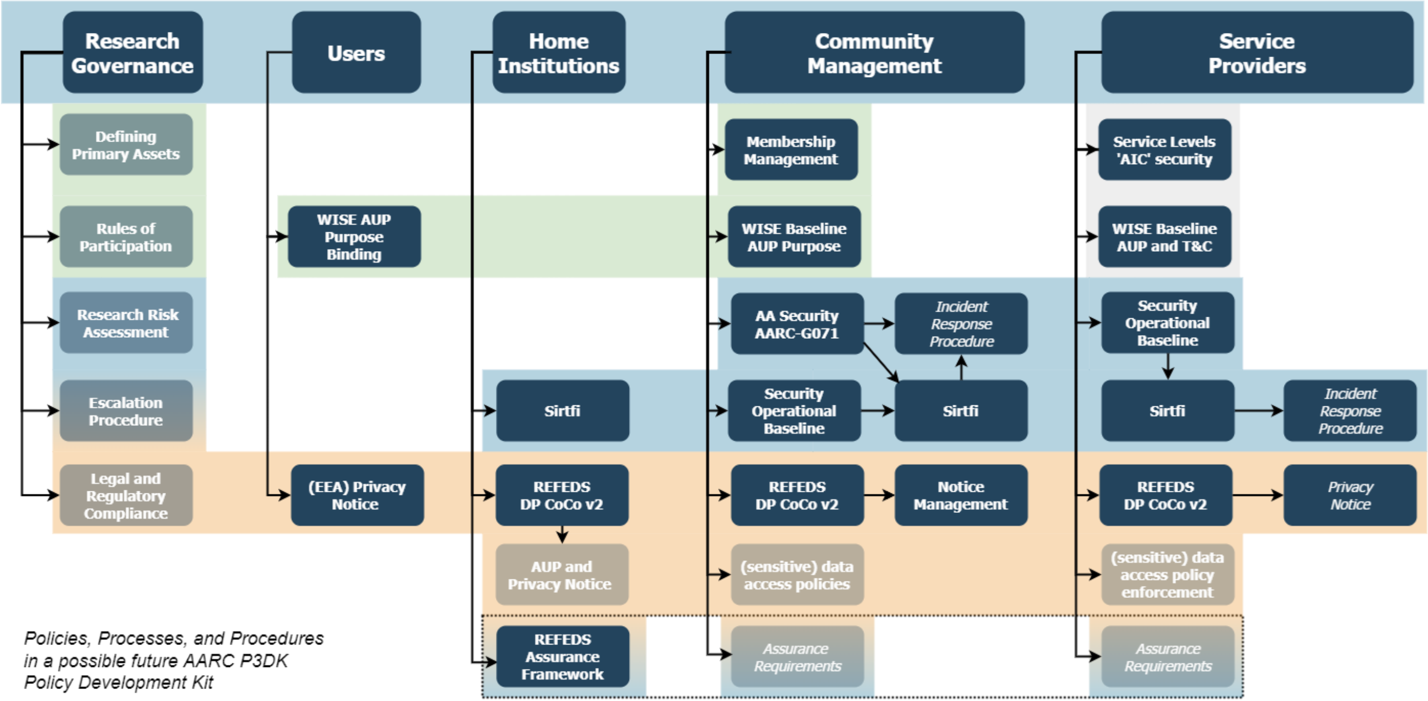
a *simpler* and *deployment-oriented*
Policy, Process, and Procedure Development Kit version 2

- comprehensive review of existing policy suite to reduce complexity
- input from national research infrastructures and EOSC nodes, but not *only* in Europe but e.g. also Australia
- leverage the works we co-created with REFEDS and EOSC

https://aarc-community.org/policies/policy-development-kit/

AARC-I082

Building the trust framework: development of the new *full* PDK structure



Policies, Processes, and Procedures in a possible future AARC P3DK Policy Development Kit

🟩 Purpose of the collaboration

🟦 Protection of the collaboration, its ICT services and resources

🟧 Protection of the collaborating users (and any sensitive research data)

AARC-I082

Today specialised AAI platform providers have established themselves

- Previous PDK policies targeted primarily at *infrastructure AAIs* and at *operators* of the few multi-community AAIs
- BPA2025 identifies platform layers, and AAI platform *operators* serving many collaborations and infrastructures with a common layer are a key player today
- A ‘trusted proxy operator’ can now be either self-hosted or used ‘as a service’

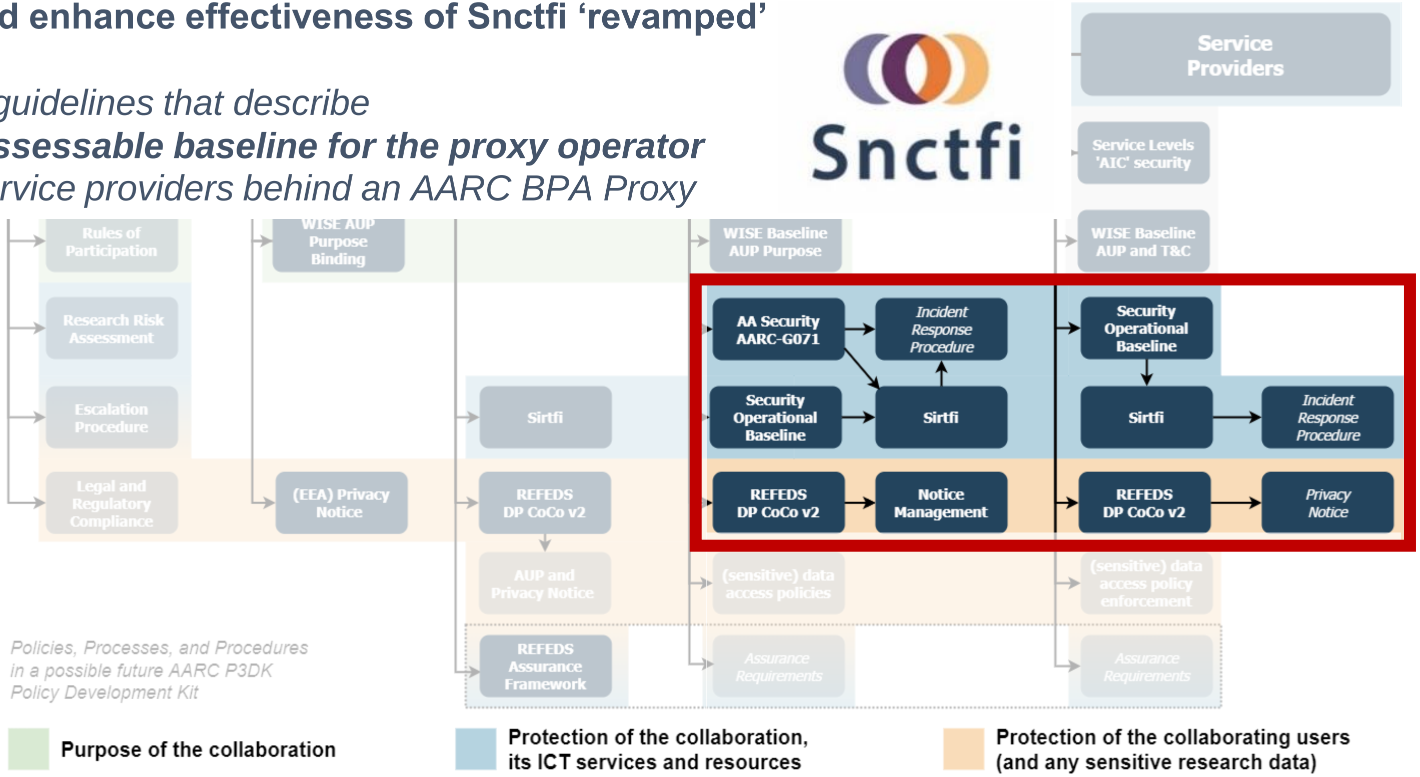
This has changed the policy landscape:
the more complex policy implementations can now be ‘sourced’ from trusted providers

AARC-I082

AAI infrastructure providers for communities: a new ‘Snctfi’ trust mark

review and enhance effectiveness of Snctfi ‘revamped’

the set of guidelines that describe
*a **(self-) assessable baseline for the proxy operator***
a set of service providers behind an AARC BPA Proxy



Snctfi
Service Providers
Service Levels 'AIC' security
Rules of Participation
WISE AUP Purpose Binding
WISE Baseline AUP Purpose
WISE Baseline AUP and T&C
Research Risk Assessment
AA Security AARC-G071
Incident Response Procedure
Security Operational Baseline
Escalation Procedure
Sirtfi
Security Operational Baseline
Sirtfi
Sirtfi
Incident Response Procedure
Legal and Regulatory Compliance
(EEA) Privacy Notice
REFEDS DP CoCo v2
REFEDS DP CoCo v2
Notice Management
REFEDS DP CoCo v2
Privacy Notice
AUP and Privacy Notice
(sensitive) data access policies
(sensitive) data access policy enforcement
REFEDS Assurance Framework
Assurance Requirements
Assurance Requirements

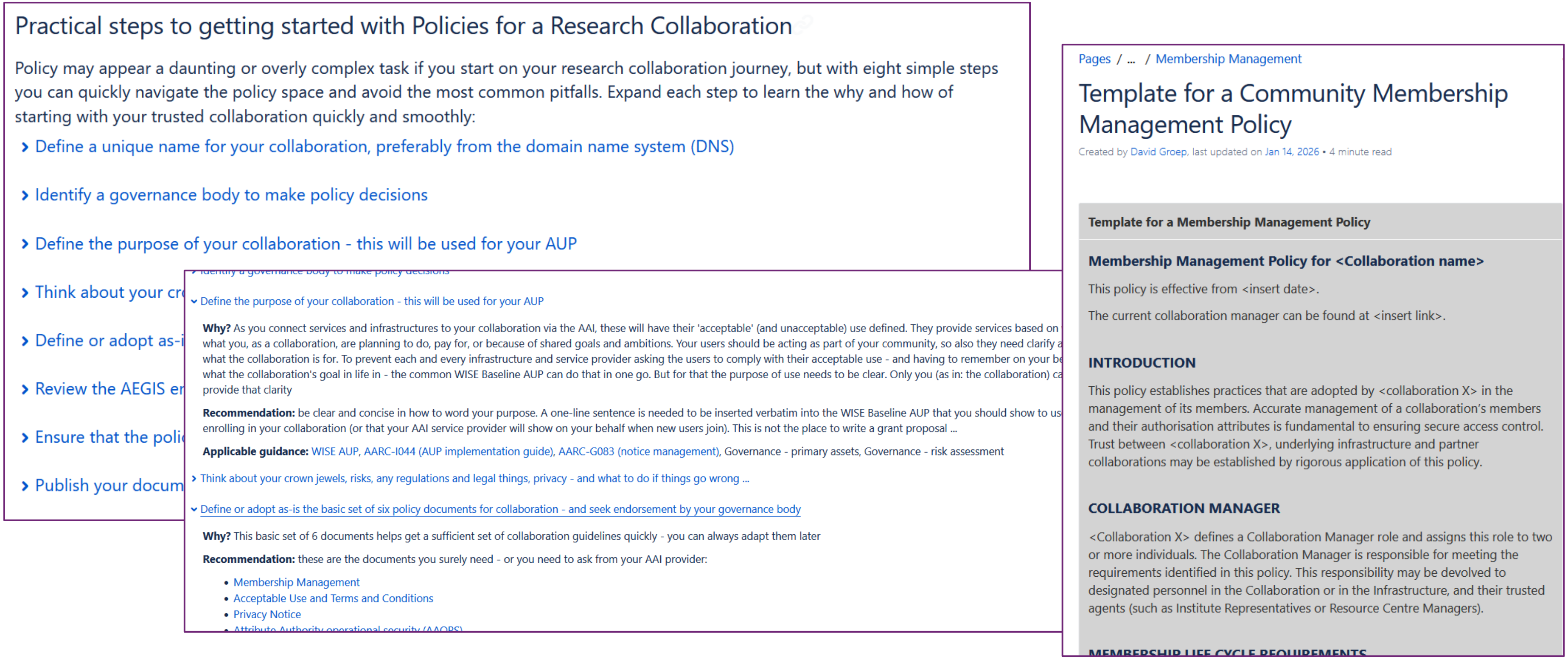
Policies, Processes, and Procedures in a possible future AARC P3DK Policy Development Kit

🟩 Purpose of the collaboration

🟦 Protection of the collaboration, its ICT services and resources

🟧 Protection of the collaborating users (and any sensitive research data)

Collaboration: foundational guidance



Practical steps to getting started with Policies for a Research Collaboration

Policy may appear a daunting or overly complex task if you start on your research collaboration journey, but with eight simple steps you can quickly navigate the policy space and avoid the most common pitfalls. Expand each step to learn the why and how of starting with your trusted collaboration quickly and smoothly:

› Define a unique name for your collaboration, preferably from the domain name system (DNS)

› Identify a governance body to make policy decisions

› Define the purpose of your collaboration - this will be used for your AUP

› Think about your cr

› Define or adopt as-i

› Review the AEGIS e

› Ensure that the poli

› Publish your docum

› Identify a governance body to make policy decisions

⌄ Define the purpose of your collaboration - this will be used for your AUP

Why? As you connect services and infrastructures to your collaboration via the AAI, these will have their 'acceptable' (and unacceptable) use defined. They provide services based on what you, as a collaboration, are planning to do, pay for, or because of shared goals and ambitions. Your users should be acting as part of your community, so also they need clarify a what the collaboration is for. To prevent each and every infrastructure and service provider asking the users to comply with their acceptable use - and having to remember on your be what the collaboration's goal in life in - the common WISE Baseline AUP can do that in one go. But for that the purpose of use needs to be clear. Only you (as in: the collaboration) ca provide that clarity

Recommendation: be clear and concise in how to word your purpose. A one-line sentence is needed to be inserted verbatim into the WISE Baseline AUP that you should show to us enrolling in your collaboration (or that your AAI service provider will show on your behalf when new users join). This is not the place to write a grant proposal ...

Applicable guidance: WISE AUP, AARC-I044 (AUP implementation guide), AARC-G083 (notice management), Governance - primary assets, Governance - risk assessment

› Think about your crown jewels, risks, any regulations and legal things, privacy - and what to do if things go wrong ...

⌄ Define or adopt as-is the basic set of six policy documents for collaboration - and seek endorsement by your governance body

Why? This basic set of 6 documents helps get a sufficient set of collaboration guidelines quickly - you can always adapt them later

Recommendation: these are the documents you surely need - or you need to ask from your AAI provider:

- Membership Management
- Acceptable Use and Terms and Conditions
- Privacy Notice
- Attribute Authority operational security (AAOPS)

Pages / ... / Membership Management

Template for a Community Membership Management Policy

Created by David Groep, last updated on Jan 14, 2026 • 4 minute read

Template for a Membership Management Policy

Membership Management Policy for <Collaboration name>

This policy is effective from <insert date>.

The current collaboration manager can be found at <insert link>.

INTRODUCTION

This policy establishes practices that are adopted by <collaboration X> in the management of its members. Accurate management of a collaboration's members and their authorisation attributes is fundamental to ensuring secure access control. Trust between <collaboration X>, underlying infrastructure and partner collaborations may be established by rigorous application of this policy.

COLLABORATION MANAGER

<Collaboration X> defines a Collaboration Manager role and assigns this role to two or more individuals. The Collaboration Manager is responsible for meeting the requirements identified in this policy. This responsibility may be devolved to designated personnel in the Collaboration or in the Infrastructure, and their trusted agents (such as Institute Representatives or Resource Centre Managers).

MEMBERSHIP LIFE CYCLE REQUIREMENTS

More importantly:
AARC Guidelines series as a pathways to policy sustainability and impact

AARC-I082 Trust framework for proxies and Snctfi research services landscape analysis and structure

AARC-G083 Guidance for Notice Management by Proxies reducing user frustration by streamlining

AARC-G084 Security Operational Baseline trusted and secure infrastructure and incident response

AARC-I085 eID Assurance Model Assessment investigates capabilities for leveraging national eID

AARC-I086 Membership Management Policy Development at light-weight and infrastructure-level

AARC-PDK Policy Development Kit an interactive resource for jumpstarting collaboration

Cross-cutting guidelines

AARC-G080 Blueprint Architecture 2025 as the conceptual foundation

AARC-G081 Recommendations for Token Lifetimes balancing usage patterns and security

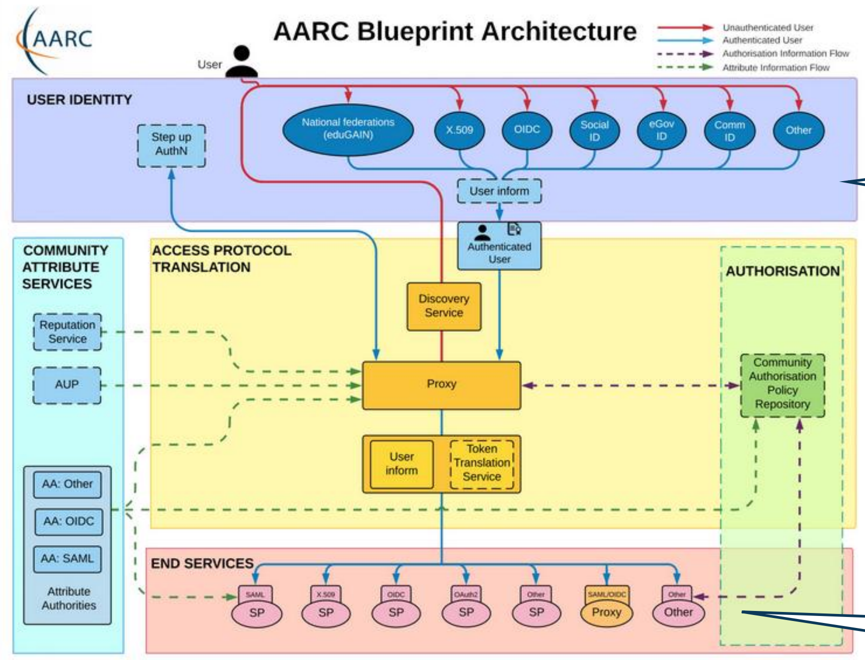
Adoption stimuli through the Policy Development Kit version 2 for

AARC-G071 'Attribute Authority and Proxy Operations', **AARC-I044** 'Baseline AUP implementation'
AARC-I051 and SIRTFI federated incident response, REFEDS DPCoCo v2, **AARC-G042** 'DPIA' for research collaborations, REFEDS Assurance Framework

https://aarc-community.org/guidelines/; PDK: https://aarc-community.org/policy/policy-development-kit/

Practices we already have, practices we need to harmonise

AARC-G071



Authentication/identity sources

NIST SP800-63

FIPS140

ISO 27001

IGTF AP Profiles

REFEDS MFA

REFEDS Assurance Framework

so … what about standards for the Community Attribute Authority (AA) or for operation of the Proxy?

Service provider operations

ISO27k

NIS2

ITSRM2

while for identity sources and for services there is extensive normalisation, our AARC BPA ‘proxy’ did not …

AARC-G071

How to establish secure operation for your (AARC BPA) proxy?

The Challenge

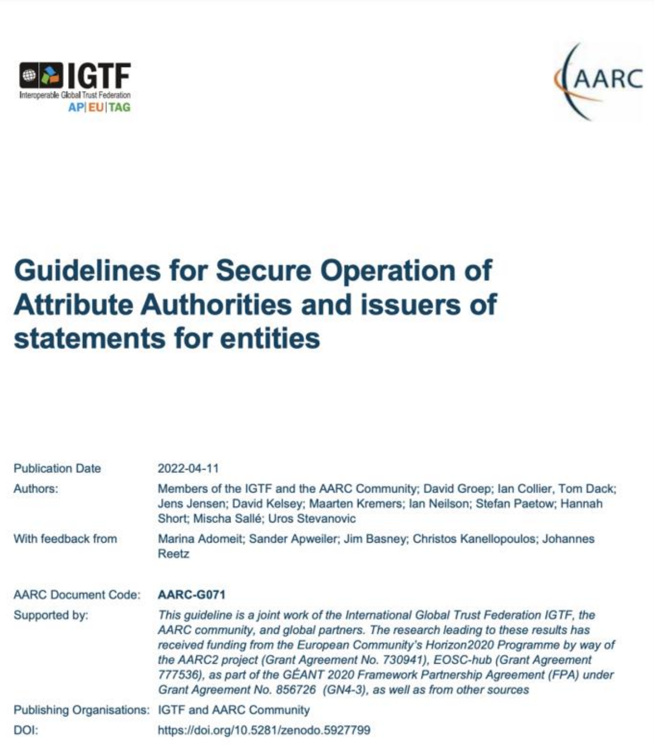
- How to securely operate proxies, attribute authorities and issuers of statements for entities?

Guideline

- AARC-G071 Guidelines for Secure Operation of Attribute Authorities

Summary

- Operational security processes and procedures
- Requirements on traceability, auditability, and logging
- Requirements on the secure operation
- Requirements on securing the interactions



IGTF Interoperable Global Trust Federation AP|EU|TAG

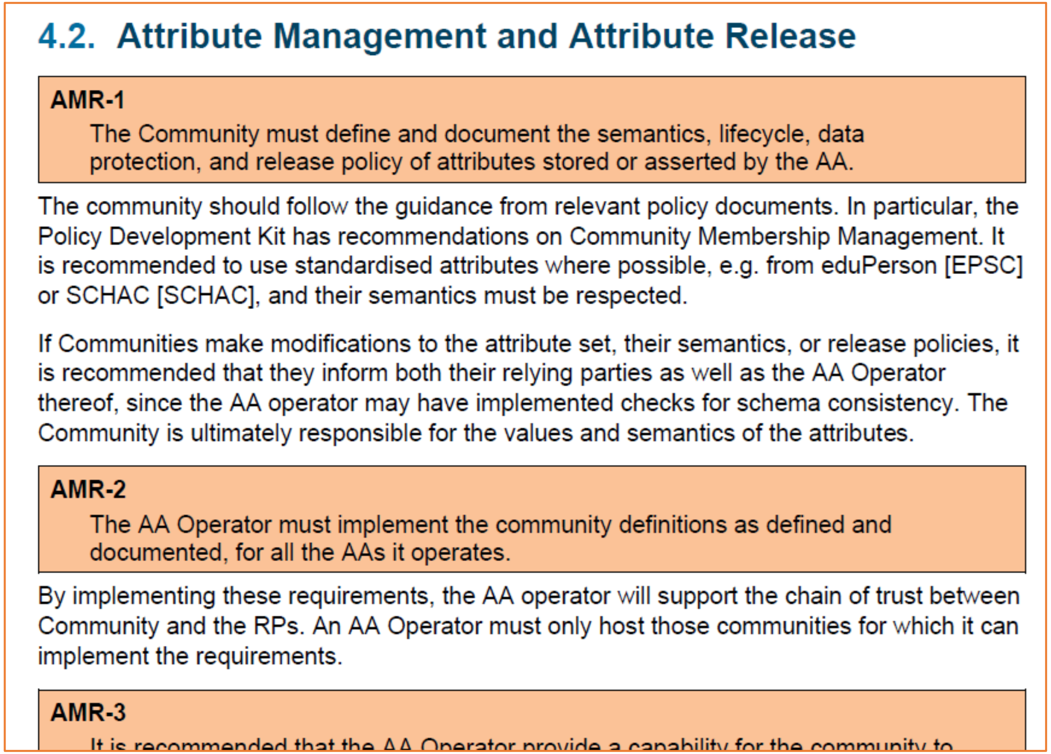
AARC

Guidelines for Secure Operation of Attribute Authorities and issuers of statements for entities

Publication Date	2022-04-11
Authors:	Members of the IGTF and the AARC Community; David Groep; Ian Collier, Tom Dack; Jens Jensen; David Kelsey; Maarten Kremers; Ian Neilson; Stefan Paetow; Hannah Short; Mischa Sallé; Uros Stevanovic
With feedback from	Marina Adomeit; Sander Apweiler; Jim Basney; Christos Kanellopoulos; Johannes Reetz
AARC Document Code:	**AARC-G071**
Supported by:	*This guideline is a joint work of the International Global Trust Federation IGTF, the AARC community, and global partners. The research leading to these results has received funding from the European Community's Horizon2020 Programme by way of the AARC2 project (Grant Agreement No. 730941), EOSC-hub (Grant Agreement 777536), as part of the GÉANT 2020 Framework Partnership Agreement (FPA) under Grant Agreement No. 856726 (GN4-3), as well as from other sources*
Publishing Organisations:	IGTF and AARC Community
DOI:	https://doi.org/10.5281/zenodo.5927799

Deployment guidance, self-assessment, and peer feed-back

AARC-G071



4.2. Attribute Management and Attribute Release

AMR-1

The Community must define and document the semantics, lifecycle, data protection, and release policy of attributes stored or asserted by the AA.

The community should follow the guidance from relevant policy documents. In particular, the Policy Development Kit has recommendations on Community Membership Management. It is recommended to use standardised attributes where possible, e.g. from eduPerson [EPSC] or SCHAC [SCHAC], and their semantics must be respected.

If Communities make modifications to the attribute set, their semantics, or release policies, it is recommended that they inform both their relying parties as well as the AA Operator thereof, since the AA operator may have implemented checks for schema consistency. The Community is ultimately responsible for the values and semantics of the attributes.

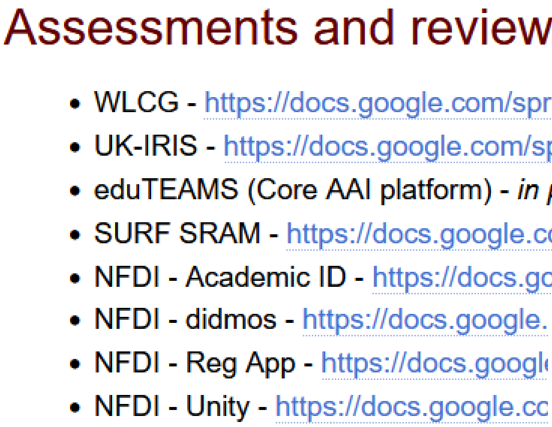
AMR-2

The AA Operator must implement the community definitions as defined and documented, for all the AAs it operates.

By implementing these requirements, the AA operator will support the chain of trust between Community and the RPs. An AA Operator must only host those communities for which it can implement the requirements.

AMR-3

It is recommended that the AA Operator provide a capability for the community to



Assessments and review shee

- WLCG - https://docs.google.com/spreadsheet
- UK-IRIS - https://docs.google.com/spreadshee
- eduTEAMS (Core AAI platform) - *in progress*
- SURF SRAM - https://docs.google.com/spread
- NFDI - Academic ID - https://docs.google.com
- NFDI - didmos - https://docs.google.
- NFDI - Reg App - https://docs.googl
- NFDI - Unity - https://docs.google.cc

http://wiki.eugridpma.org/Main/AAOperationsGuidelines

Proxy Operations: Information Security and Security Operational Baseline

AARC-G084

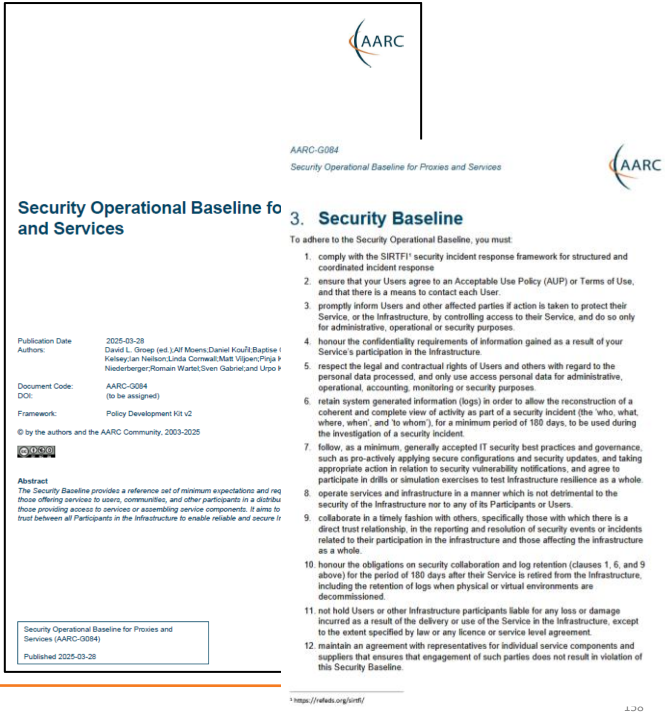
> *‘address information security for disciplines and infrastructures - some of which process sensitive data’*

Service Security Policy from AARC PDK v1
was successful but diverged in several directions:

- national implementations and specialisations
- included in EOSC Interoperability Framework as ‘Security Operational Baseline’

The new PDK in AARC TREE converges on a common **Baseline** - with guidance and FAQ

- Included in the EOSC AAI WG Federation 2025



Security Operational Baseline fo
and Services

Publication Date 2025-03-28
Authors: David L. Groep (ed.);Alf Moens;Daniel Kouřil;Baptise ‹
Kelsey;Ian Neilson;Linda Cornwall;Matt Viljoen;Pinja K
Niederberger;Romain Wartel;Sven Gabriel;and Urpo K

Document Code: AARC-G084
DOI: (to be assigned)

Framework: Policy Development Kit v2

© by the authors and the AARC Community, 2003-2025

Abstract
The Security Baseline provides a reference set of minimum expectations and req
those offering services to users, communities, and other participants in a distribu
those providing access to services or assembling service components. It aims to
trust between all Participants in the Infrastructure to enable reliable and secure Ir

Security Operational Baseline for Proxies and Services (AARC-G084)

Published 2025-03-28

AARC-G084

Security Operational Baseline for Proxies and Services

3. Security Baseline

To adhere to the Security Operational Baseline, you must:

1. comply with the SIRTFI[1] security incident response framework for structured and coordinated incident response
2. ensure that your Users agree to an Acceptable Use Policy (AUP) or Terms of Use, and that there is a means to contact each User.
3. promptly inform Users and other affected parties if action is taken to protect their Service, or the Infrastructure, by controlling access to their Service, and do so only for administrative, operational or security purposes.
4. honour the confidentiality requirements of information gained as a result of your Service's participation in the Infrastructure.
5. respect the legal and contractual rights of Users and others with regard to the personal data processed, and only use access personal data for administrative, operational, accounting, monitoring or security purposes.
6. retain system generated information (logs) in order to allow the reconstruction of a coherent and complete view of activity as part of a security incident (the 'who, what, where, when', and 'to whom'), for a minimum period of 180 days, to be used during the investigation of a security incident.
7. follow, as a minimum, generally accepted IT security best practices and governance, such as pro-actively applying secure configurations and security updates, and taking appropriate action in relation to security vulnerability notifications, and agree to participate in drills or simulation exercises to test Infrastructure resilience as a whole.
8. operate services and infrastructure in a manner which is not detrimental to the security of the Infrastructure nor to any of its Participants or Users.
9. collaborate in a timely fashion with others, specifically those with which there is a direct trust relationship, in the reporting and resolution of security events or incidents related to their participation in the infrastructure and those affecting the infrastructure as a whole.
10. honour the obligations on security collaboration and log retention (clauses 1, 6, and 9 above) for the period of 180 days after their Service is retired from the Infrastructure, including the retention of logs when physical or virtual environments are decommissioned.
11. not hold Users or other Infrastructure participants liable for any loss or damage incurred as a result of the delivery or use of the Service in the Infrastructure, except to the extent specified by law or any licence or service level agreement.
12. maintain an agreement with representatives for individual service components and suppliers that ensures that engagement of such parties does not result in violation of this Security Baseline.

[1] https://refeds.org/sirtfi/

AARC-G084 https://doi.org/10.5281/zenodo.7396724

AARC-G084

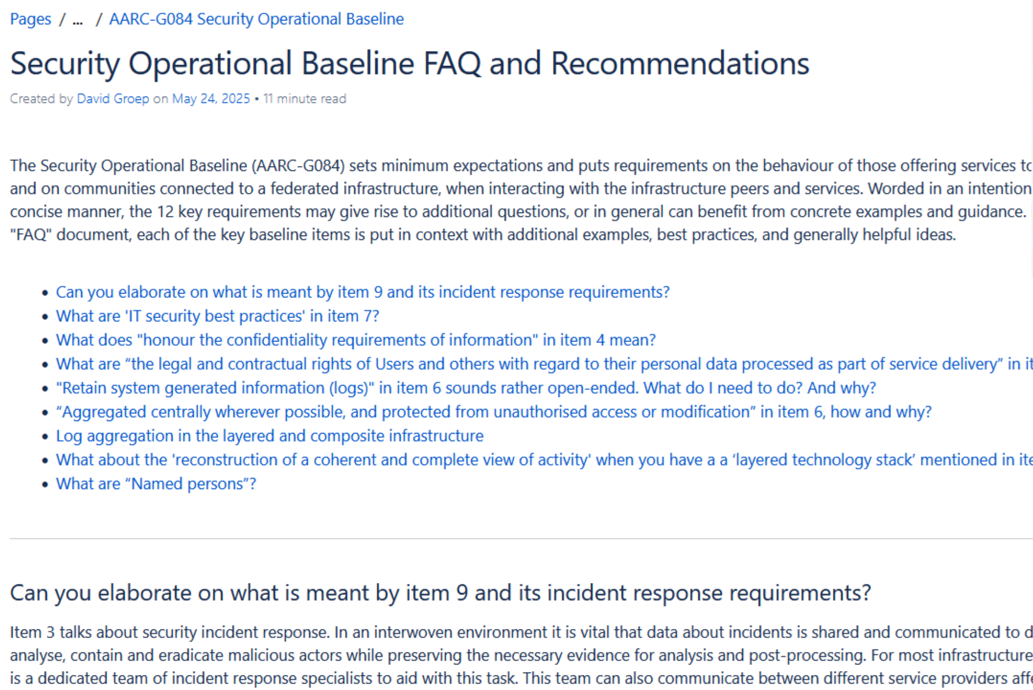
The 12 points of AARC-G084

1. comply with the SIRTFI security incident response framework for structured and coordinated incident response [ref to SIRTFI]
2. ensure that your Users agree to an Acceptable Use Policy (AUP) or Terms of Use, and that there is a means to contact each User.
3. promptly inform Users and other affected parties if action is taken to protect their Service, or the Infrastructure, by controlling access to their Service, and do so only for administrative, operational or security purposes.
4. honour the confidentiality requirements of information gained as a result of your Service's participation in the Infrastructure.
5. respect the legal and contractual rights of Users and others with regard to the personal data processed, and only use access personal data for administrative, operational, accounting, monitoring or security purposes.
6. retain system generated information (logs) in order to allow the reconstruction of a coherent and complete view of activity as part of a security incident (the 'who, what, where, when', and 'to whom'), for a minimum period of 180 days, to be used during the investigation of a security incident.
7. follow, as a minimum, generally accepted IT security best practices and governance, such as pro-actively applying secure configurations and security updates, and taking appropriate action in relation to security vulnerability notifications, and agree to participate in drills or simulation exercises to test Infrastructure resilience as a whole.
8. operate services and infrastructure in a manner which is not detrimental to the security of the Infrastructure nor to any of its Participants or Users.
9. collaborate in a timely fashion with others, specifically those with which there is a direct trust relationship, in the reporting and resolution of security events or incidents related to their participation in the infrastructure and those affecting the infrastructure as a whole.
10. honour the obligations on security collaboration and log retention (clauses 1, 6, and 9 above) for the period of 180 days after their Service is retired from the Infrastructure, including the retention of logs when physical or virtual environments are decommissioned.
11. not hold Users or other Infrastructure participants liable for any loss or damage incurred as a result of the delivery or use of the Service in the Infrastructure, except to the extent specified by law or any licence or service level agreement.
12. maintain an agreement with representatives for individual service components and suppliers that ensures that engagement of such parties does not result in violation of this Security Baseline

FAQ and implementation guidance

AARC-G084

https://wiki.geant.org/spaces/AARC/pages/1049624759/view



Pages / ... / AARC-G084 Security Operational Baseline

Security Operational Baseline FAQ and Recommendations

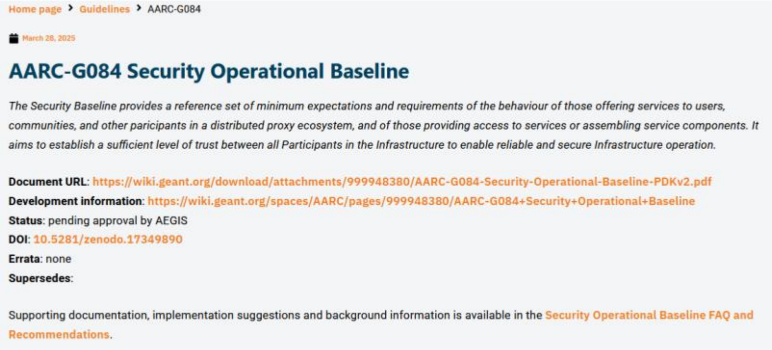
Created by David Groep on May 24, 2025 • 11 minute read

The Security Operational Baseline (AARC-G084) sets minimum expectations and puts requirements on the behaviour of those offering services tc and on communities connected to a federated infrastructure, when interacting with the infrastructure peers and services. Worded in an intention concise manner, the 12 key requirements may give rise to additional questions, or in general can benefit from concrete examples and guidance. "FAQ" document, each of the key baseline items is put in context with additional examples, best practices, and generally helpful ideas.

- Can you elaborate on what is meant by item 9 and its incident response requirements?
- What are 'IT security best practices' in item 7?
- What does "honour the confidentiality requirements of information" in item 4 mean?
- What are "the legal and contractual rights of Users and others with regard to their personal data processed as part of service delivery" in item 5?
- "Retain system generated information (logs)" in item 6 sounds rather open-ended. What do I need to do? And why?
- "Aggregated centrally wherever possible, and protected from unauthorised access or modification" in item 6, how and why?
- Log aggregation in the layered and composite infrastructure
- What about the 'reconstruction of a coherent and complete view of activity' when you have a a 'layered technology stack' mentioned in item 6?
- What are "Named persons"?

Can you elaborate on what is meant by item 9 and its incident response requirements?

Item 3 talks about security incident response. In an interwoven environment it is vital that data about incidents is shared and communicated to detect, analyse, contain and eradicate malicious actors while preserving the necessary evidence for analysis and post-processing. For most infrastructures, there is a dedicated team of incident response specialists to aid with this task. This team can also communicate between different service providers affected by



Home page › Guidelines › AARC-G084

March 28, 2025

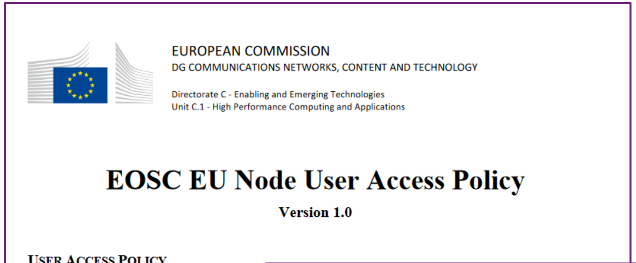
AARC-G084 Security Operational Baseline

The Security Baseline provides a reference set of minimum expectations and requirements of the behaviour of those offering services to users, communities, and other participants in a distributed proxy ecosystem, and of those providing access to services or assembling service components. It aims to establish a sufficient level of trust between all Participants in the Infrastructure to enable reliable and secure Infrastructure operation.

Document URL: https://wiki.geant.org/download/attachments/999948380/AARC-G084-Security-Operational-Baseline-PDKv2.pdf
Development information: https://wiki.geant.org/spaces/AARC/pages/999948380/AARC-G084+Security+Operational+Baseline
Status: pending approval by AEGIS
DOI: 10.5281/zenodo.17349890
Errata: none
Supersedes:

Supporting documentation, implementation suggestions and background information is available in the Security Operational Baseline FAQ and Recommendations.

Helping community and users: how much clicking through?



EUROPEAN COMMISSION
DG COMMUNICATIONS NETWORKS, CONTENT AND TECHNOLOGY

Directorate C - Enabling and Emerging Technologies
Unit C.1 - High Performance Computing and Applications

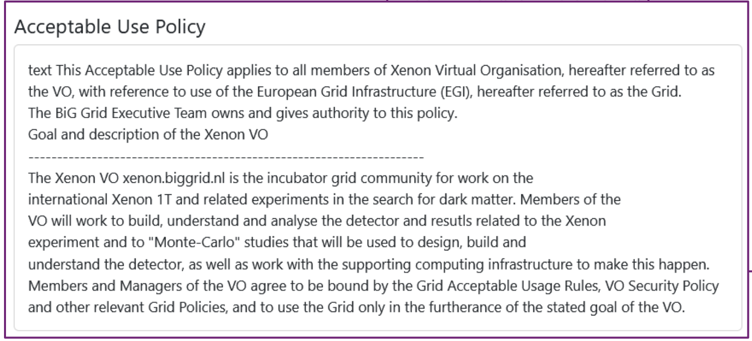
EOSC EU Node User Access Policy

Version 1.0

USER ACCESS POLICY

1. Purpose

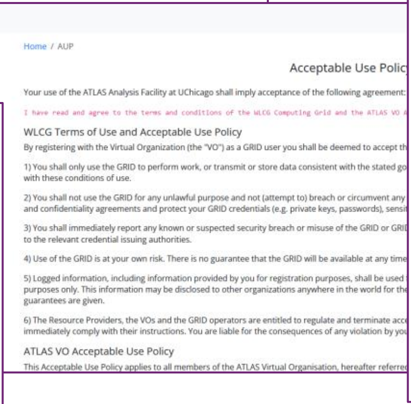
This User Access Policy ("UAP") defi



Acceptable Use Policy

text This Acceptable Use Policy applies to all members of Xenon Virtual Organisation, hereafter referred to as the VO, with reference to use of the European Grid Infrastructure (EGI), hereafter referred to as the Grid.
The BiG Grid Executive Team owns and gives authority to this policy.
Goal and description of the Xenon VO

The Xenon VO xenon.biggrid.nl is the incubator grid community for work on the
international Xenon 1T and related experiments in the search for dark matter. Members of the
VO will work to build, understand and analyse the detector and resutls related to the Xenon
experiment and to "Monte-Carlo" studies that will be used to design, build and
understand the detector, as well as work with the supporting computing infrastructure to make this happen.
Members and Managers of the VO agree to be bound by the Grid Acceptable Usage Rules, VO Security Policy and other relevant Grid Policies, and to use the Grid only in the furtherance of the stated goal of the VO.



Home / AUP

Acceptable Use Polic

Your use of the ATLAS Analysis Facility at UChicago shall imply acceptance of the following agreement:

```
I have read and agree to the terms and conditions of the WLCG Computing Grid and the ATLAS VO A
```

WLCG Terms of Use and Acceptable Use Policy

By registering with the Virtual Organization (the "VO") as a GRID user you shall be deemed to accept th

1) You shall only use the GRID to perform work, or transmit or store data consistent with the stated go
with these conditions of use.

2) You shall not use the GRID for any unlawful purpose and not (attempt to) breach or circumvent any
and confidentiality agreements and protect your GRID credentials (e.g. private keys, passwords), sensit

3) You shall immediately report any known or suspected security breach or misuse of the GRID or GRII
to the relevant credential issuing authorities.

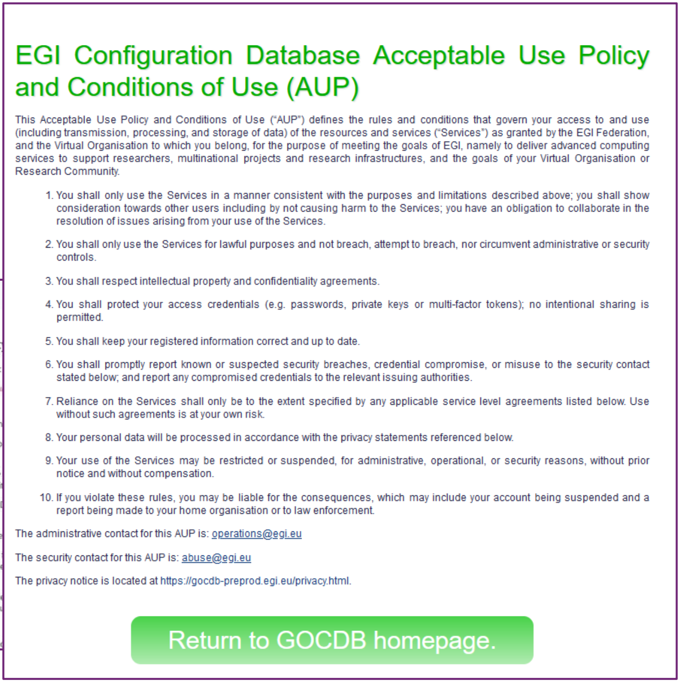
4) Use of the GRID is at your own risk. There is no guarantee that the GRID will be available at any time

5) Logged information, including information provided by you for registration purposes, shall be used
purposes only. This information may be disclosed to other organizations anywhere in the world for the
guarantees are given.

6) The Resource Providers, the VOs and the GRID operators are entitled to regulate and terminate acce
immediately comply with their instructions. You are liable for the consequences of any violation by you

ATLAS VO Acceptable Use Policy

This Acceptable Use Policy applies to all members of the ATLAS Virtual Organisation, hereafter referre



EGI Configuration Database Acceptable Use Policy and Conditions of Use (AUP)

This Acceptable Use Policy and Conditions of Use ("AUP") defines the rules and conditions that govern your access to and use (including transmission, processing, and storage of data) of the resources and services ("Services") as granted by the EGI Federation, and the Virtual Organisation to which you belong, for the purpose of meeting the goals of EGI, namely to deliver advanced computing services to support researchers, multinational projects and research infrastructures, and the goals of your Virtual Organisation or Research Community.

1. You shall only use the Services in a manner consistent with the purposes and limitations described above; you shall show consideration towards other users including by not causing harm to the Services; you have an obligation to collaborate in the resolution of issues arising from your use of the Services.
2. You shall only use the Services for lawful purposes and not breach, attempt to breach, nor circumvent administrative or security controls.
3. You shall respect intellectual property and confidentiality agreements.
4. You shall protect your access credentials (e.g. passwords, private keys or multi-factor tokens); no intentional sharing is permitted.
5. You shall keep your registered information correct and up to date.
6. You shall promptly report known or suspected security breaches, credential compromise, or misuse to the security contact stated below; and report any compromised credentials to the relevant issuing authorities.
7. Reliance on the Services shall only be to the extent specified by any applicable service level agreements listed below. Use without such agreements is at your own risk.
8. Your personal data will be processed in accordance with the privacy statements referenced below.
9. Your use of the Services may be restricted or suspended, for administrative, operational, or security reasons, without prior notice and without compensation.
10. If you violate these rules, you may be liable for the consequences, which may include your account being suspended and a report being made to your home organisation or to law enforcement.

The administrative contact for this AUP is: operations@egi.eu

The security contact for this AUP is: abuse@egi.eu

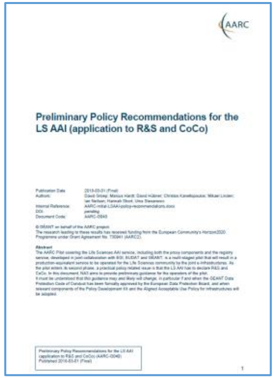
The privacy notice is located at https://gocdb-preprod.egi.eu/privacy.html.

AARC-G083

Proxies have their 'experience challenges': AUPs, T&Cs, Privacy notices, …

For **large 'multi-tenant' proxies**

- some subset users in some communities use a set of services – how to I present their Terms and Conditions, and their privacy policies, so that the users
  - only see the T&Cs and notices for services they will access
  - this does not to need to be manually configured for each community
  - is automatically updated when services join



AARC

Preliminary Policy Recommendations for the LS AAI (application to R&S and CoCo)

Publication Date: 2019-03-01 (Final)
Authors: David Groep; Marcus Hardt; David Hübner; Christos Kanellopoulos; Mikael Linden; Ian Neilson; Hannah Short; Uros Stevanovic
Internal Reference: AARC-initial-LSAAI-policy-recommendations.docx
DOI: pending
Document Code: AARC-G048

© GÉANT on behalf of the AARC project.
The research leading to these results has received funding from the European Community's Horizon2020 Programme under Grant Agreement No. 730941 (AARC2).

Abstract
The AARC Pilot covering the Life Sciences AAI service, including both the proxy components and the registry service, developed in joint collaboration with EGI, EUDAT and GÉANT, is a multi-staged pilot that will result in a production-equivalent service to be operated for the Life Sciences community by the joint e-Infrastructures. As the pilot enters its second phase, a practical policy related issue is that the LS AAI has to declare R&S and CoCo. In this document, NA3 aims to provide preliminary guidance for the operators of the pilot.
It must be understood that this guidance may and likely will change, in particular if and when the GÉANT Data Protection Code of Conduct has been formally approved by the European Data Protection Board, and when relevant components of the Policy Development Kit and the Aligned Acceptable Use Policy for infrastructures will be adopted.

Preliminary Policy Recommendations for the LS AAI (application to R&S and CoCo) (AARC-G048)
Published 2019-03-01 (Final)

1

beyond bespoke guidance

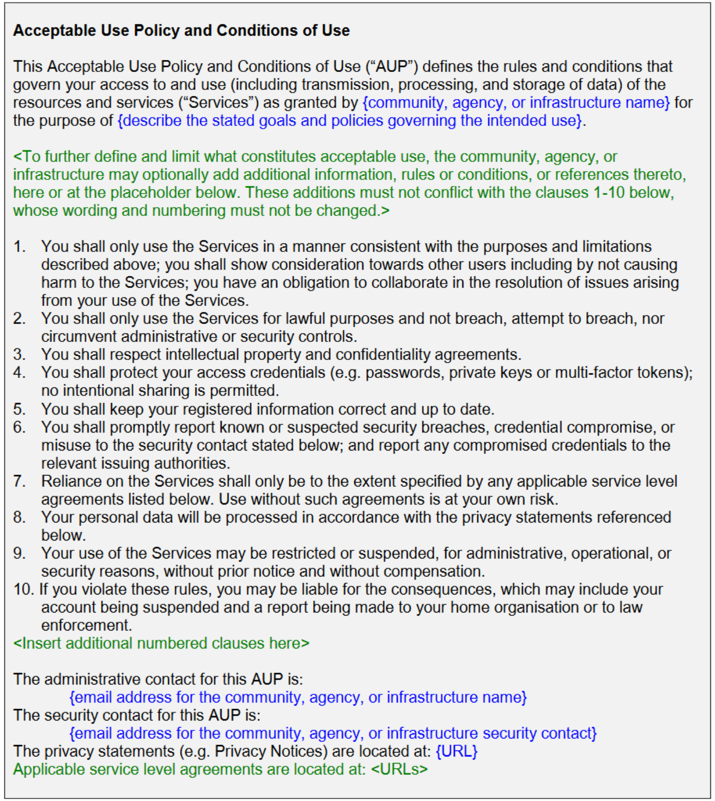
as well as for **community and dedicated proxies**

- when new (sensitive) services join, who actually needs to *see* the new T&Cs?
- can we communicate acceptance of T&Cs to services even if 'we' are small and 'they' are large?

What is an acceptable user experience in clicking through agreements?
What is most effective in exploiting the WISE Baseline AUP? What do *you* need?

With Fewer Clicks to More Resources!

Good common practice: the WISE Baseline AUP



Acceptable Use Policy and Conditions of Use

This Acceptable Use Policy and Conditions of Use (“AUP”) defines the rules and conditions that govern your access to and use (including transmission, processing, and storage of data) of the resources and services (“Services”) as granted by {community, agency, or infrastructure name} for the purpose of {describe the stated goals and policies governing the intended use}.

<To further define and limit what constitutes acceptable use, the community, agency, or infrastructure may optionally add additional information, rules or conditions, or references thereto, here or at the placeholder below. These additions must not conflict with the clauses 1-10 below, whose wording and numbering must not be changed.>

1. You shall only use the Services in a manner consistent with the purposes and limitations described above; you shall show consideration towards other users including by not causing harm to the Services; you have an obligation to collaborate in the resolution of issues arising from your use of the Services.
2. You shall only use the Services for lawful purposes and not breach, attempt to breach, nor circumvent administrative or security controls.
3. You shall respect intellectual property and confidentiality agreements.
4. You shall protect your access credentials (e.g. passwords, private keys or multi-factor tokens); no intentional sharing is permitted.
5. You shall keep your registered information correct and up to date.
6. You shall promptly report known or suspected security breaches, credential compromise, or misuse to the security contact stated below; and report any compromised credentials to the relevant issuing authorities.
7. Reliance on the Services shall only be to the extent specified by any applicable service level agreements listed below. Use without such agreements is at your own risk.
8. Your personal data will be processed in accordance with the privacy statements referenced below.
9. Your use of the Services may be restricted or suspended, for administrative, operational, or security reasons, without prior notice and without compensation.
10. If you violate these rules, you may be liable for the consequences, which may include your account being suspended and a report being made to your home organisation or to law enforcement.

<Insert additional numbered clauses here>

The administrative contact for this AUP is:
{email address for the community, agency, or infrastructure name}
The security contact for this AUP is:
{email address for the community, agency, or infrastructure security contact}
The privacy statements (e.g. Privacy Notices) are located at: {URL}
Applicable service level agreements are located at: <URLs>

Purpose binding
ensure use is as intended for access grant

Terms and Conditions
research data access conditions, permits, grant conditions

WISE Baseline AUP
common 10 commandments that allow seamless cross-sectoral user movement

Service level agreements
promises and recourse

Privacy notice references
for *access* personal data policies

Guidance for Notice Management by Proxies

https://wise-community.org/wise-baseline-aup/

AARC-G083

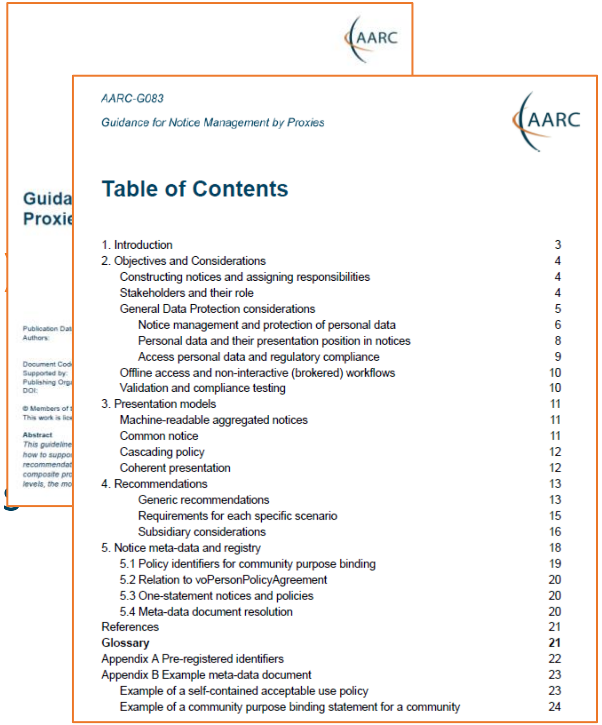
New AARC guidance on Notice Management by Proxies

Four **presentation models** In order of preference

1. **machine-readable** aggregated notice
2. common notice (single common **authority domain**)
3. cascading notices (**assume responsibility** for underlings)
4. coherent presentation: you show what you need (but not more)

Recommend WISE Baseline AUP plus model to **construct notices and communicate acceptance** based on the AARC ID-community-infra hierarchy of proxies

- sufficient to build you a comprehensive WISE Baseline AUP
- and a set of privacy notices (for those GDPR encumbered)
- plus a namespace inspired by RFC6711's LoA registry



Guida
Proxie

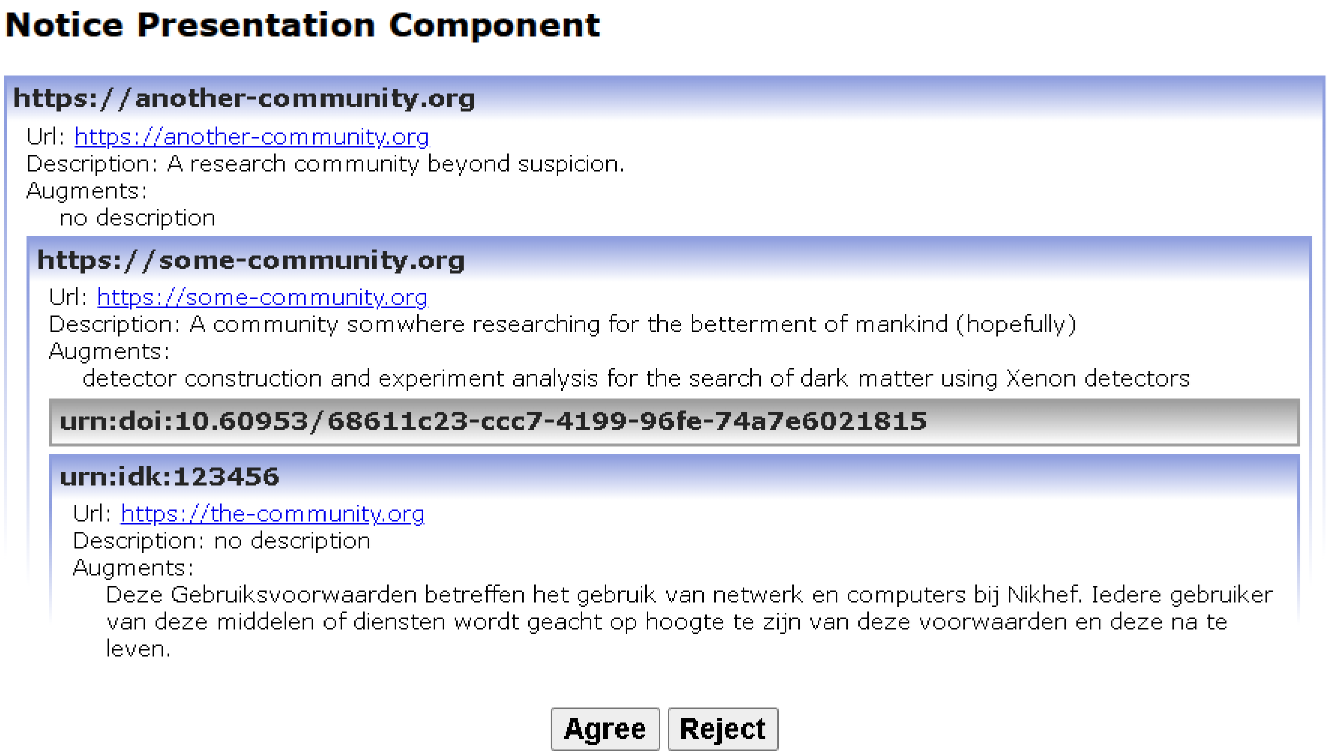
AARC-G083

Guidance for Notice Management by Proxies

Table of Contents

Notice presentation (PoC example implementation from the Validator)

AARC-G083



Notice Presentation Component

https://another-community.org

Url: https://another-community.org
Description: A research community beyond suspicion.
Augments:
no description

https://some-community.org

Url: https://some-community.org
Description: A community somwhere researching for the betterment of mankind (hopefully)
Augments:
detector construction and experiment analysis for the search of dark matter using Xenon detectors

urn:doi:10.60953/68611c23-ccc7-4199-96fe-74a7e6021815

urn:idk:123456

Url: https://the-community.org
Description: no description
Augments:
Deze Gebruiksvoorwaarden betreffen het gebruik van netwerk en computers bij Nikhef. Iedere gebruiker van deze middelen of diensten wordt geacht op hoogte te zijn van deze voorwaarden en deze na te leven.

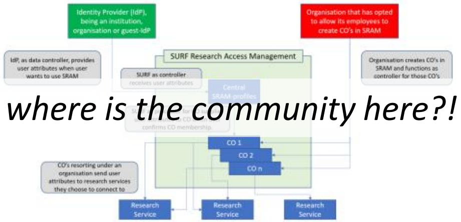
Helping out the community – a simpler policy toolkit for communities

AARC-I086

provide a revised policy development kit for mid-sized communities using the research infrastructures

Requirement from the AAI operators in FIM4R and BPA operators:

> ***“small to mid-sized communities do not have the resources to maintain a bespoke community management policy”***



Identity Provider (IdP), being an institution, organisation or guest-IdP
Organisation that has opted to allow its employees to create CO's in SRAM
IdP, as data controller, provides user attributes when user wants to use SRAM
SURF Research Access Management
SURF as controller receives user attributes
Central SRAM profiles
Organisation creates CO's in SRAM and functions as controller for those CO's
confirms CO membership.
CO 1
CO 2
CO n
CO's resorting under an organisation send user attributes to research services they choose to connect to
Research Service
Research Service
Research Service

where is the community here?!

But both communities and operators of membership management services are today unclear about trust assurance level of members: current templates in toolkit too complex and prescriptive

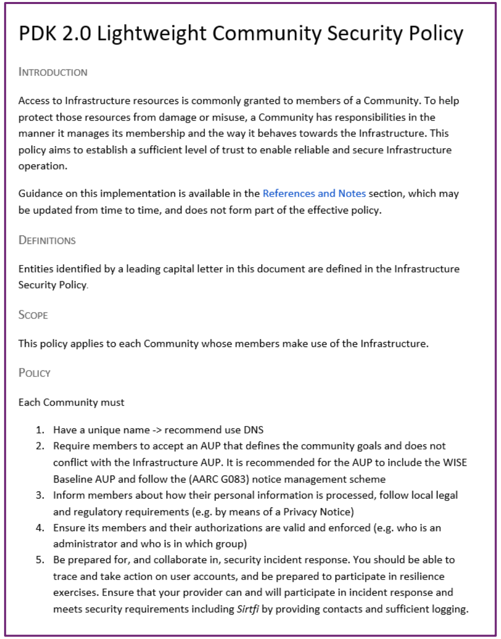
- develop ‘minimum viable community management’ for most small and mid-sized use cases
- give template and implementation guidance (FAQ) on community lifecycle management
- leverage complement of PDK practices that communities can ‘source’ from trusted providers

I086: Simplified Community Management policy – down to five items!

AARC-I086

Each Community must

- Have a **unique name** (we recommend use DNS domain names)
- Require **members to accept an AUP** that defines the community goals and does not conflict with the Infrastructure AUP. It is recommended for the AUP to include the WISE Baseline AUP and follow the (AARC G083) notice management scheme
- Inform members about how their **personal information is processed**, follow local legal and regulatory requirements (e.g. by means of a Privacy Notice)
- Ensure its **members and their authorizations are valid** and enforced (e.g. who is an administrator and who is in which group)
- Be prepared for, and collaborate in, **security incident response**. You should be able to trace and take action on user accounts, and be prepared to participate in resilience exercises. Ensure that your provider can and will participate in incident response and meets security requirements including *Sirtfi* by providing contacts and sufficient logging.



PDK 2.0 Lightweight Community Security Policy

INTRODUCTION

Access to Infrastructure resources is commonly granted to members of a Community. To help protect those resources from damage or misuse, a Community has responsibilities in the manner it manages its membership and the way it behaves towards the Infrastructure. This policy aims to establish a sufficient level of trust to enable reliable and secure Infrastructure operation.

Guidance on this implementation is available in the References and Notes section, which may be updated from time to time, and does not form part of the effective policy.

DEFINITIONS

Entities identified by a leading capital letter in this document are defined in the Infrastructure Security Policy.

SCOPE

This policy applies to each Community whose members make use of the Infrastructure.

POLICY

Each Community must

1. Have a unique name -> recommend use DNS
2. Require members to accept an AUP that defines the community goals and does not conflict with the Infrastructure AUP. It is recommended for the AUP to include the WISE Baseline AUP and follow the (AARC G083) notice management scheme
3. Inform members about how their personal information is processed, follow local legal and regulatory requirements (e.g. by means of a Privacy Notice)
4. Ensure its members and their authorizations are valid and enforced (e.g. who is an administrator and who is in which group)
5. Be prepared for, and collaborate in, security incident response. You should be able to trace and take action on user accounts, and be prepared to participate in resilience exercises. Ensure that your provider can and will participate in incident response and meets security requirements including *Sirtfi* by providing contacts and sufficient logging.

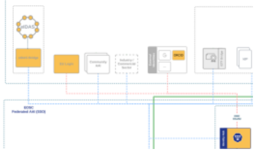
https://aarc-community.org/guidelines/aarc-i086/

More diverse sources of researcher identity & assurance with eID wallets

AARC-I085

Most reliable (and most ‘available’) source of assurance could be government identity!

- Step-up can now readily be done ‘at home’ by users through their national eID schemes
- eID wallets could solve the blockage by home IdPs to release assurance



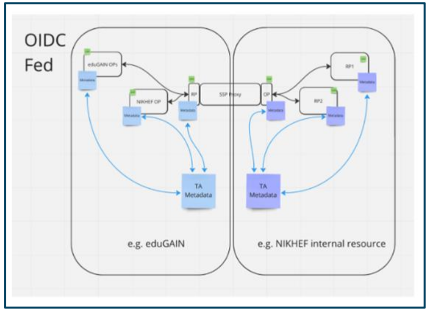
… but their applicability to research and education use cases remains limited:

- eIDAS 1.0 suffers from inconsistent national uptake, asymmetrical cross-border connectivity, and protocol incompatibilities
- eIDAS 2 *at this point in time*, has incomplete roll-out, national implementations vary widely, and support for non-governmental use cases remains immature
- non-European users in Europe and international linking are not addressed at all today

Verifiable Credentials and digital wallets offer a complementary path forward, but lack of ecosystem maturity, lack of common standards, and adoption are (too) far in the future …

Build trusted baseline expectations to increase reach of RI proxies

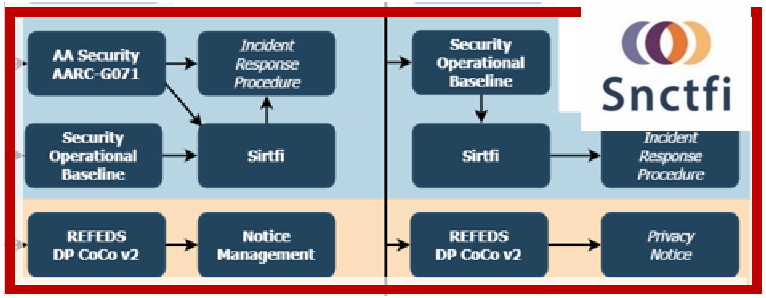
… with R&E identity providers … and with new sources of information



Service and data providers need *unique identifier and affiliation,* with name and email, and 'fresh' assurance from home IdPs, but:

- proxies have met with scepticism by IdPs:
  lack of even basic personalised and R&S attribute release
- how do these trust qualities 'traverse' proxies?
- how do operators rely on adherence to guidelines by their 'downstream' providers?

Position of the proxy makes trust bidirectional, and ***platform operators*** *are facilitating this trust today*

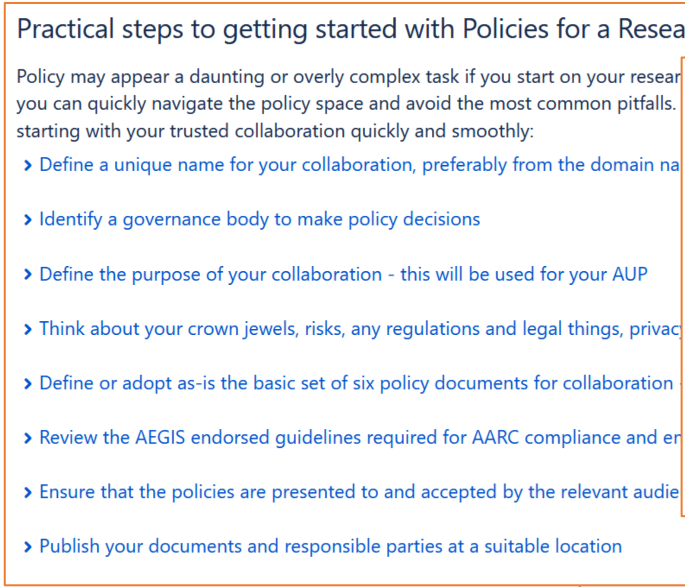

AA Security
AARC-G071
Incident Response Procedure
Security Operational Baseline
Sirtfi
REFEDS
DP CoCo v2
Notice Management
Security Operational Baseline
Snctfi
Sirtfi
Incident Response Procedure
REFEDS
DP CoCo v2
Privacy Notice

AARC TREE D2.2

AARC PDK

Bringing it together: the Policy Development Kit

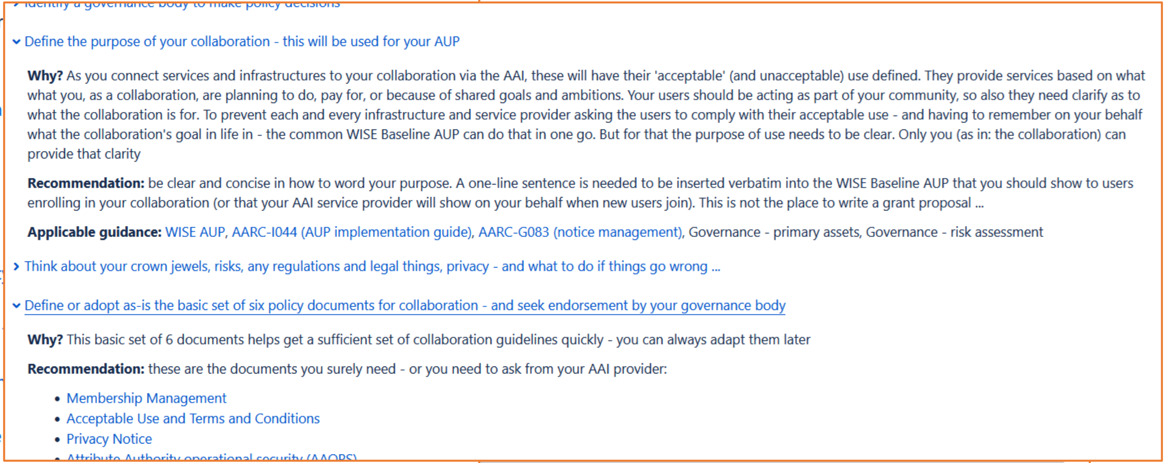
PDK v2 has guidelines *and* explanations, hints, and accessible recommendations



Practical steps to getting started with Policies for a Research Collaboration

Policy may appear a daunting or overly complex task if you start on your resear
you can quickly navigate the policy space and avoid the most common pitfalls.
starting with your trusted collaboration quickly and smoothly:

- › Define a unique name for your collaboration, preferably from the domain na
- › Identify a governance body to make policy decisions
- › Define the purpose of your collaboration - this will be used for your AUP
- › Think about your crown jewels, risks, any regulations and legal things, privac
- › Define or adopt as-is the basic set of six policy documents for collaboration
- › Review the AEGIS endorsed guidelines required for AARC compliance and en
- › Ensure that the policies are presented to and accepted by the relevant audie
- › Publish your documents and responsible parties at a suitable location



› Identify a governance body to make policy decisions

⌄ Define the purpose of your collaboration - this will be used for your AUP

Why? As you connect services and infrastructures to your collaboration via the AAI, these will have their 'acceptable' (and unacceptable) use defined. They provide services based on what what you, as a collaboration, are planning to do, pay for, or because of shared goals and ambitions. Your users should be acting as part of your community, so also they need clarify as to what the collaboration is for. To prevent each and every infrastructure and service provider asking the users to comply with their acceptable use - and having to remember on your behalf what the collaboration's goal in life in - the common WISE Baseline AUP can do that in one go. But for that the purpose of use needs to be clear. Only you (as in: the collaboration) can provide that clarity

Recommendation: be clear and concise in how to word your purpose. A one-line sentence is needed to be inserted verbatim into the WISE Baseline AUP that you should show to users enrolling in your collaboration (or that your AAI service provider will show on your behalf when new users join). This is not the place to write a grant proposal ...

Applicable guidance: WISE AUP, AARC-I044 (AUP implementation guide), AARC-G083 (notice management), Governance - primary assets, Governance - risk assessment

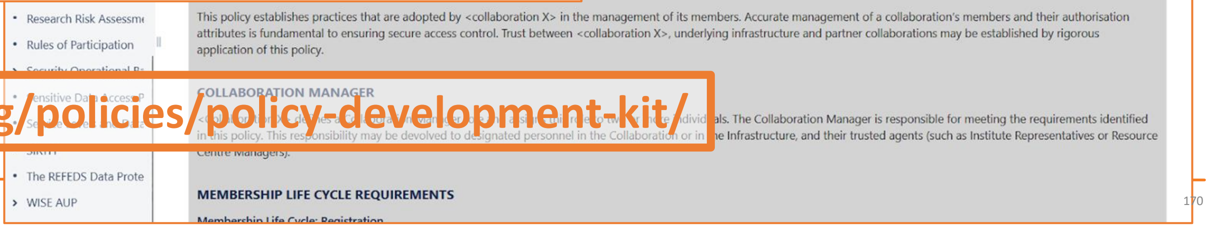
› Think about your crown jewels, risks, any regulations and legal things, privacy - and what to do if things go wrong ...

⌄ Define or adopt as-is the basic set of six policy documents for collaboration - and seek endorsement by your governance body

Why? This basic set of 6 documents helps get a sufficient set of collaboration guidelines quickly - you can always adapt them later

Recommendation: these are the documents you surely need - or you need to ask from your AAI provider:

- Membership Management
- Acceptable Use and Terms and Conditions
- Privacy Notice
- Attribute Authority operational security (AAOPS)



- Research Risk Assessme
- Rules of Participation
- Security Operational Ba
- Sensitive Data Access P
- Service Level and Data
- SIRTFI
- The REFEDS Data Prote
- › WISE AUP

This policy establishes practices that are adopted by <collaboration X> in the management of its members. Accurate management of a collaboration's members and their authorisation attributes is fundamental to ensuring secure access control. Trust between <collaboration X>, underlying infrastructure and partner collaborations may be established by rigorous application of this policy.

COLLABORATION MANAGER

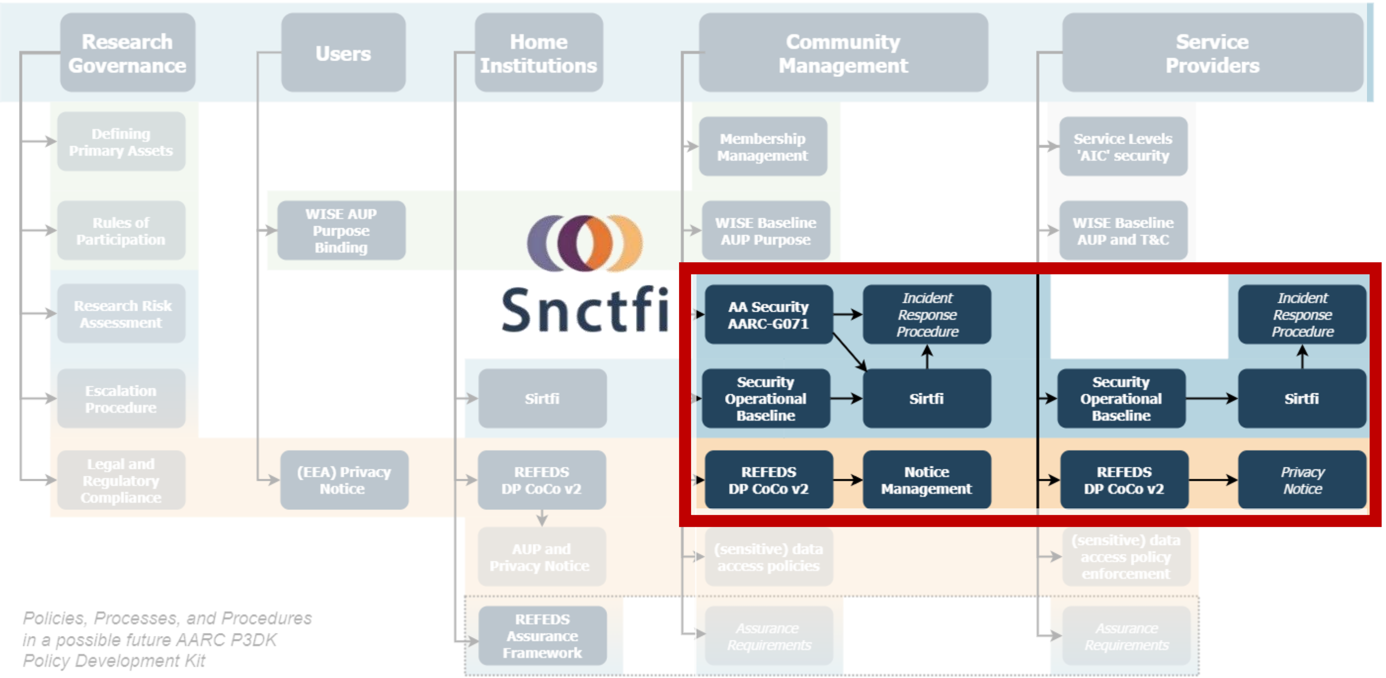
<collaboration X> defines a Collaboration Manager role and assigns this role to two or more individuals. The Collaboration Manager is responsible for meeting the requirements identified in this policy. This responsibility may be devolved to designated personnel in the Collaboration or in the Infrastructure, and their trusted agents (such as Institute Representatives or Resource Centre Managers).

MEMBERSHIP LIFE CYCLE REQUIREMENTS

Membership Life Cycle: Registration

https://aarc-community.org/policies/policy-development-kit/

Providers manage complexity for research communities



Policies, Processes, and Procedures in a possible future AARC P3DK Policy Development Kit

🟩 Purpose of the collaboration

🟦 Protection of the collaboration, its ICT services and resources

🟧 Protection of the collaborating users (and any sensitive research data)

communities sourcing 'well-operated' community platforms

egi

eduTEAMS

B2ACCESS

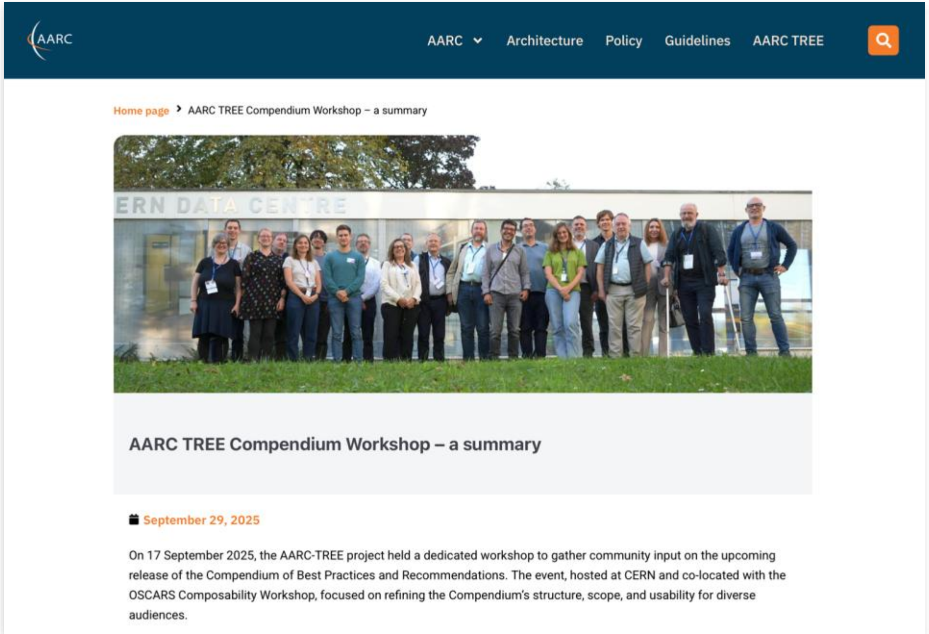
SURF (SRAM)

CILogon

and a few more …

through their scale gets federations to trust our AARC 'middle boxes'

Towards the AARC Compendium

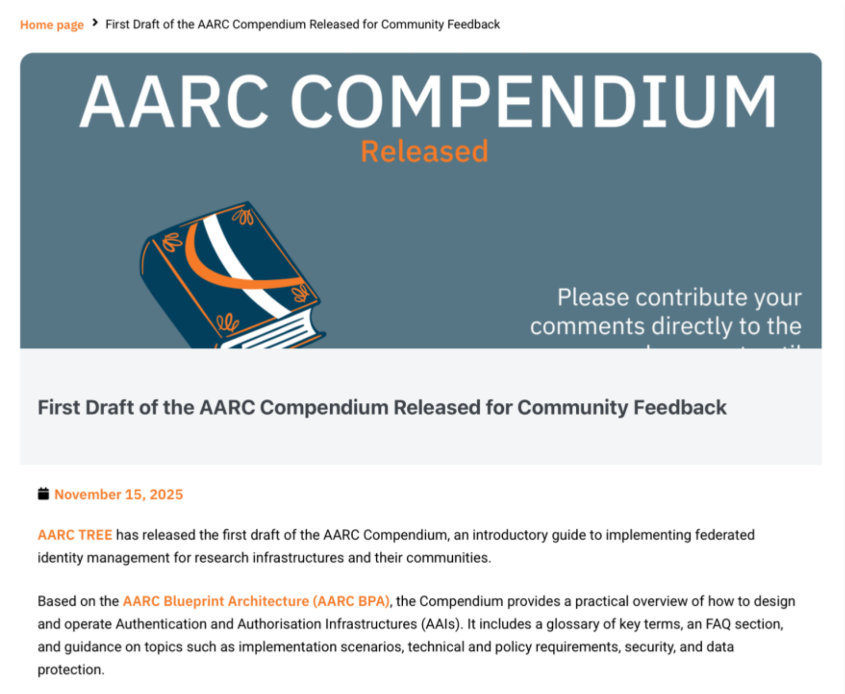


Home page › AARC TREE Compendium Workshop – a summary

AARC TREE Compendium Workshop – a summary

September 29, 2025

On 17 September 2025, the AARC-TREE project held a dedicated workshop to gather community input on the upcoming release of the Compendium of Best Practices and Recommendations. The event, hosted at CERN and co-located with the OSCARS Composability Workshop, focused on refining the Compendium's structure, scope, and usability for diverse audiences.



Home page › First Draft of the AARC Compendium Released for Community Feedback

First Draft of the AARC Compendium Released for Community Feedback

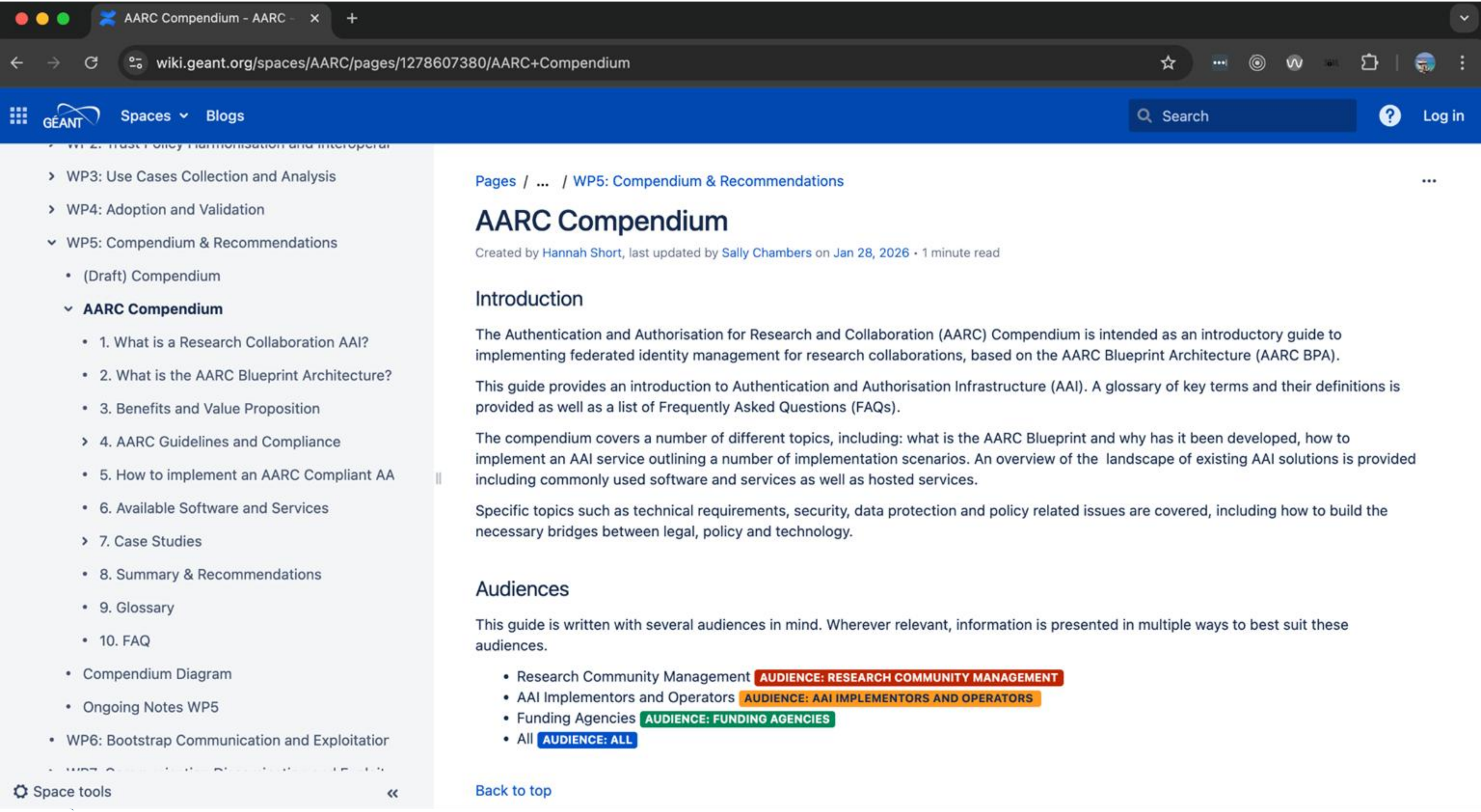
November 15, 2025

AARC TREE has released the first draft of the AARC Compendium, an introductory guide to implementing federated identity management for research infrastructures and their communities.

Based on the AARC Blueprint Architecture (AARC BPA), the Compendium provides a practical overview of how to design and operate Authentication and Authorisation Infrastructures (AAIs). It includes a glossary of key terms, an FAQ section, and guidance on topics such as implementation scenarios, technical and policy requirements, security, and data protection.

https://aarc-community.org/aarc-tree-compendium-workshop-a-summary/

https://aarc-community.org/first-draft-of-the-aarc-compendium-released-for-community-feedback/


AARC Compendium - AARC -
wiki.geant.org/spaces/AARC/pages/1278607380/AARC+Compendium
GÉANT
Spaces
Blogs
Search
Log in
WP2: Trust Policy Harmonisation and Interoperal
WP3: Use Cases Collection and Analysis
WP4: Adoption and Validation
WP5: Compendium & Recommendations
(Draft) Compendium
AARC Compendium
1. What is a Research Collaboration AAI?
2. What is the AARC Blueprint Architecture?
3. Benefits and Value Proposition
4. AARC Guidelines and Compliance
5. How to implement an AARC Compliant AA
6. Available Software and Services
7. Case Studies
8. Summary & Recommendations
9. Glossary
10. FAQ
Compendium Diagram
Ongoing Notes WP5
WP6: Bootstrap Communication and Exploitatior
Space tools
Pages / ... / WP5: Compendium & Recommendations

AARC Compendium

Created by Hannah Short, last updated by Sally Chambers on Jan 28, 2026 • 1 minute read

Introduction

The Authentication and Authorisation for Research and Collaboration (AARC) Compendium is intended as an introductory guide to implementing federated identity management for research collaborations, based on the AARC Blueprint Architecture (AARC BPA).

This guide provides an introduction to Authentication and Authorisation Infrastructure (AAI). A glossary of key terms and their definitions is provided as well as a list of Frequently Asked Questions (FAQs).

The compendium covers a number of different topics, including: what is the AARC Blueprint and why has it been developed, how to implement an AAI service outlining a number of implementation scenarios. An overview of the landscape of existing AAI solutions is provided including commonly used software and services as well as hosted services.

Specific topics such as technical requirements, security, data protection and policy related issues are covered, including how to build the necessary bridges between legal, policy and technology.

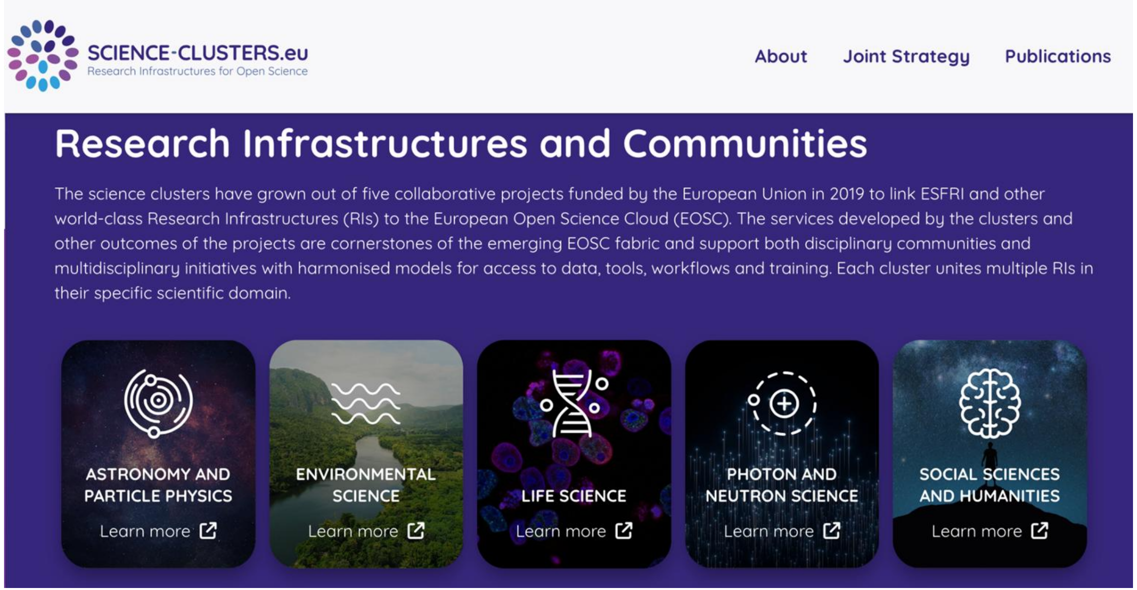
Audiences

This guide is written with several audiences in mind. Wherever relevant, information is presented in multiple ways to best suit these audiences.

- Research Community Management AUDIENCE: RESEARCH COMMUNITY MANAGEMENT
- AAI Implementors and Operators AUDIENCE: AAI IMPLEMENTORS AND OPERATORS
- Funding Agencies AUDIENCE: FUNDING AGENCIES
- All AUDIENCE: ALL

Back to top

Compendium Outreach Campaign: Engagement through Use Cases



SCIENCE-CLUSTERS.eu
Research Infrastructures for Open Science

About Joint Strategy Publications

Research Infrastructures and Communities

The science clusters have grown out of five collaborative projects funded by the European Union in 2019 to link ESFRI and other world-class Research Infrastructures (RIs) to the European Open Science Cloud (EOSC). The services developed by the clusters and other outcomes of the projects are cornerstones of the emerging EOSC fabric and support both disciplinary communities and multidisciplinary initiatives with harmonised models for access to data, tools, workflows and training. Each cluster unites multiple RIs in their specific scientific domain.

ASTRONOMY AND PARTICLE PHYSICS
Learn more
ENVIRONMENTAL SCIENCE
Learn more
LIFE SCIENCE
Learn more
PHOTON AND NEUTRON SCIENCE
Learn more
SOCIAL SCIENCES AND HUMANITIES
Learn more

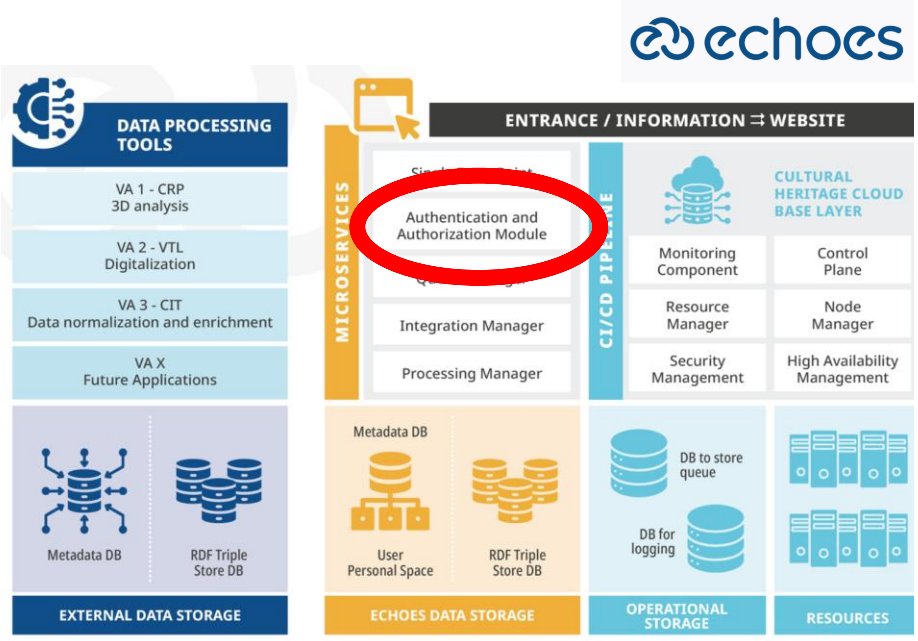
https://science-clusters.eu/

OSCARS
Open Science Clusters' Action for Research & Society

https://oscars-project.eu/

- Highlight Compendium within various communities, e.g. compendium as a resource in the SSH Open Marketplace

Compendium Outreach Campaign: Engagement through Use Cases

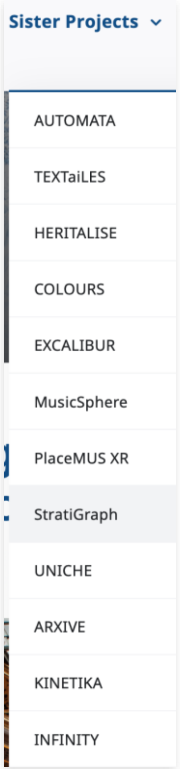


https://www.echoes-eccch.eu/infrastructure/

https://www.echoes-eccch.eu/

ECHOES Integration Task Force (EITF)

Kotzinos, D., Chambers, S., Barbot, L., Durco, M., & Dalla Torre, G. (2025). *ECHOES Integration Strategy for datasets, tools and workflows with potential for reuse in ECHOES* https://doi.org/10.5281/zenodo.17751335



Recommendations: Key messages

- There is benefit to investing in common AAI solutions, both for funding agencies and research collaborations themselves
- AAI is complex and our community generally recommends using a hosted/managed solution rather than starting from scratch. This will aid future interoperability as the landscape is highly dynamic.
- A “consulting” service where Research Collaborations can seek advice would be highly useful
- Many thanks to **all** contributors. Particular mention to the **Australian Access Federation (AAF).**

Visit

https://wiki.geant.org/spaces/AARC/pages/1278607380/AARC+Compendium

The AARC Blueprint – a very digestible architecture … so

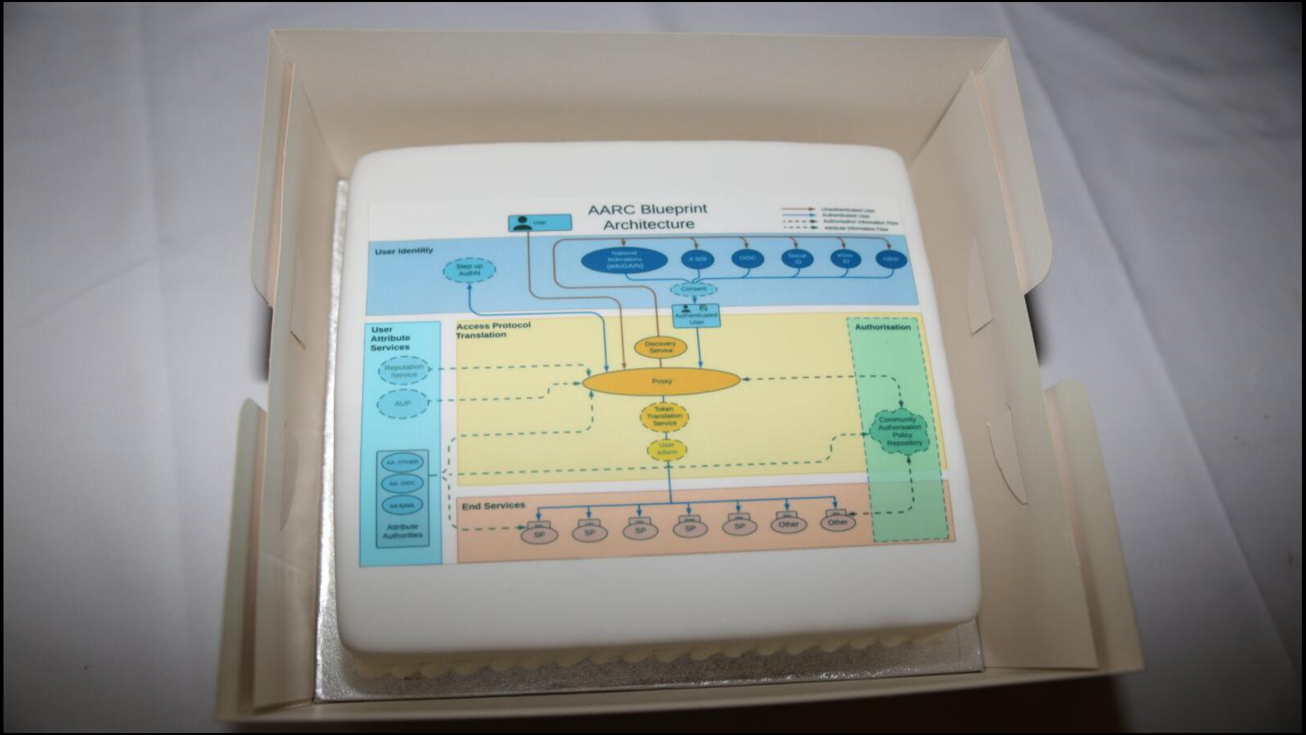
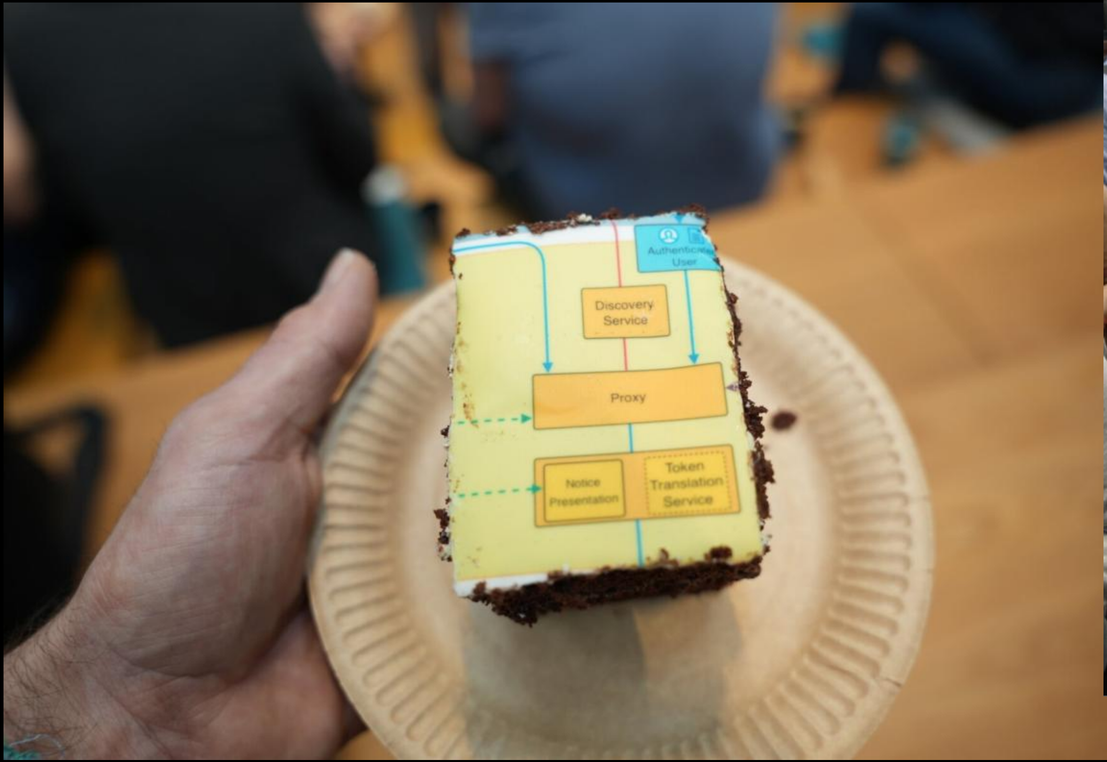


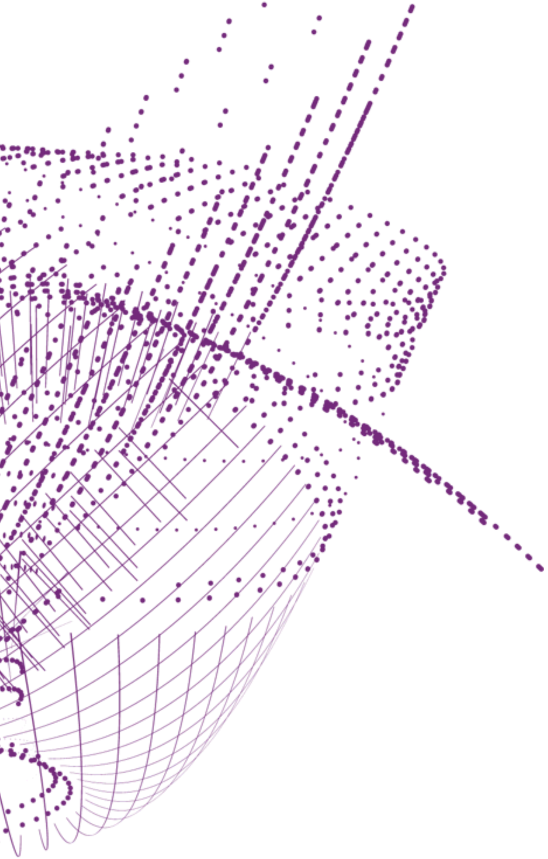
Photo credit: Marcus Hardt

The AARC Blueprint – take a piece and feed collaboration!

Photo credit: Marcus Hardt

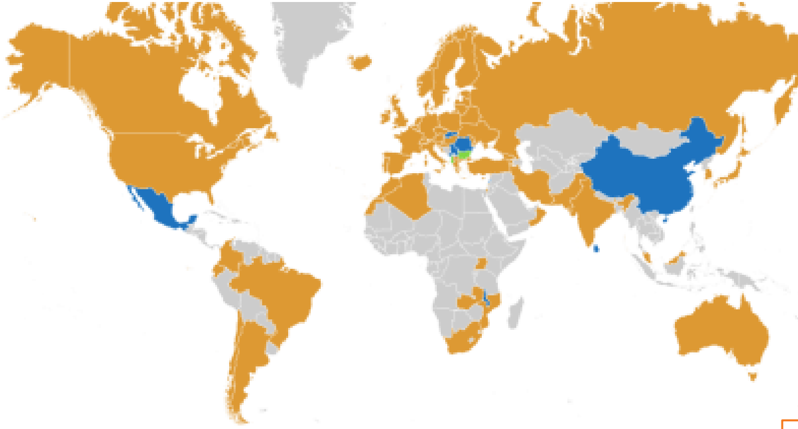
And Marcus Hardt then ate the proxy

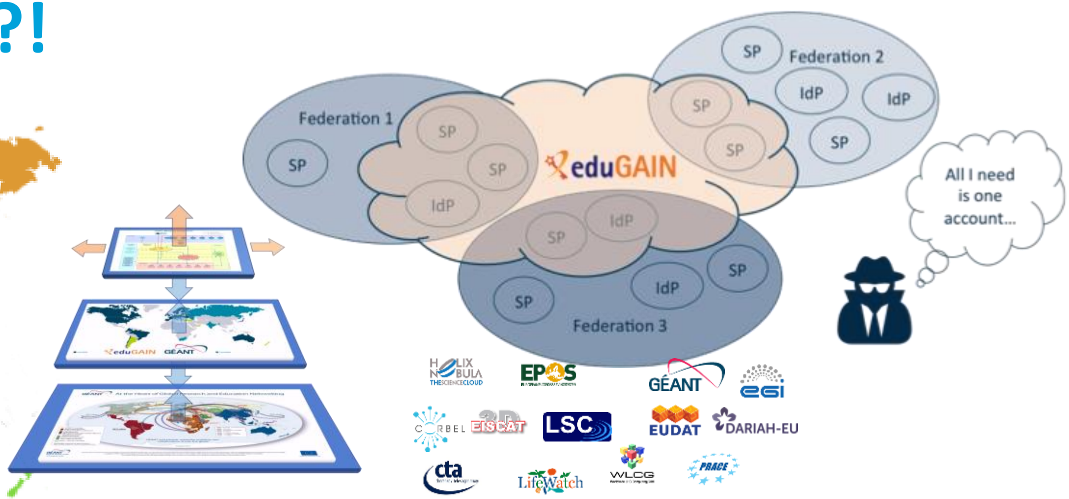




Securing our federated world

Now *what* have we built?!





full of valuable resources
(data, network, services)

We have federation and single sign-on …
… but can we share security information when needed?
… timely and confidentially, protecting everyone’s reputation?

left: eduGAIN interfederation extent in 2020; logos on the right from the European e-Infrastructures and ESFRIs; center graphic: AARC collaboration

‘Sirtfi’ – what makes federated security different?

SIRTFI
Security Incident Response Trust Framework for Federated Identity

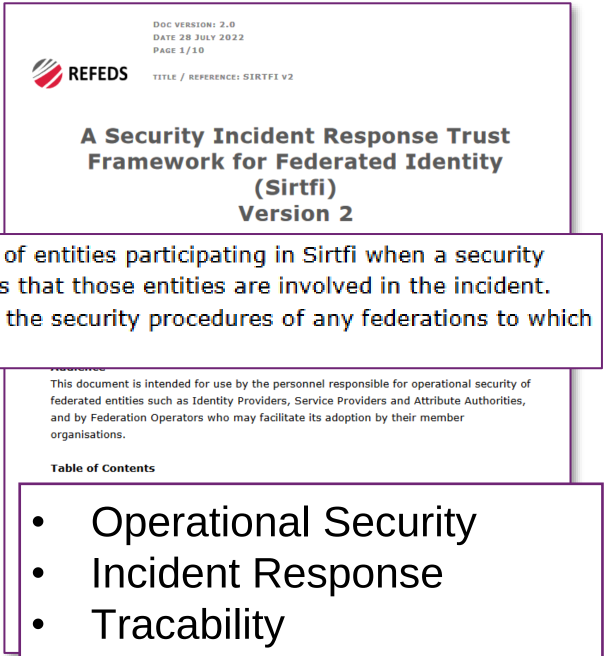
Organisations probably do ‘something reasonable’ for their own security ... but may not realise the implications for others

Sirtfi targets coordinated **response in a federated context**:

1. Enable **communication** and coordination in managing federated security incidents
2. Relevant **event data** is available to help collaborating incident responders.
3. **Security protections are applied** to federated transactions

Define capabilities for security incident response an IdP or SP **organisation can self-asserts** in federation meta-data

https://refeds.org/sirtfi



Doc version: 2.0
Date 28 July 2022
Page 1/10

REFEDS

title / reference: SIRTFI v2

A Security Incident Response Trust Framework for Federated Identity (Sirtfi) Version 2

This document is intended for use by the personnel responsible for operational security of federated entities such as Identity Providers, Service Providers and Attribute Authorities, and by Federation Operators who may facilitate its adoption by their member organisations.

Table of Contents

- [IR3] Notify security contacts of entities participating in Sirtfi when a security incident investigation suggests that those entities are involved in the incident. Notification should also follow the security procedures of any federations to which your organisation belongs.

- Operational Security
- Incident Response
- Tracability
- User Rules & Conditions

Sirtfi – Security Incident Response Trust framework for Federated Identity

SIRTFI
Security Incident Response Trust Framework for Federated Identity

A means by which to enable a **coordinated response to a security incident in a federated context** that does not depend on a centralised authority or governance structure to assign roles and responsibilities for doing so.

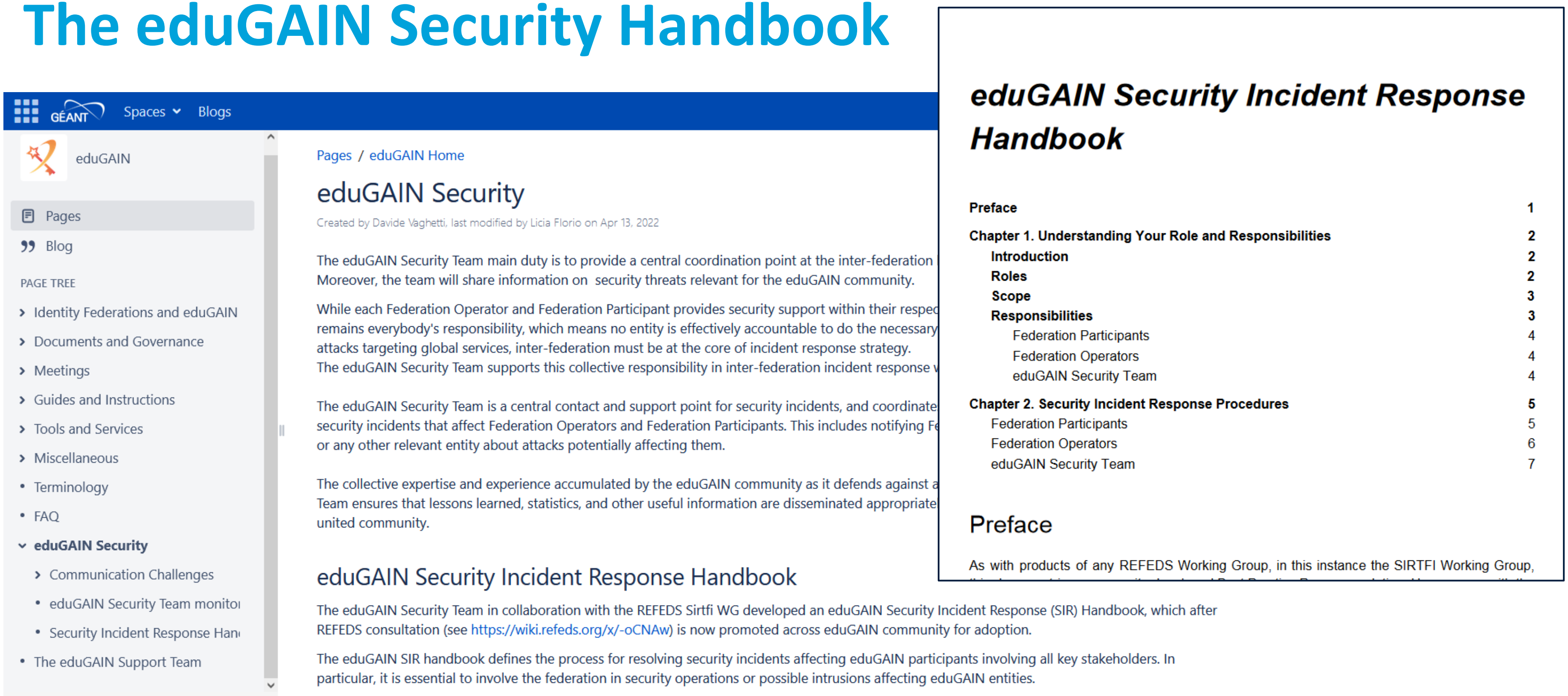
Defines a set of capabilities and roles associated with security incident response that an IdP or SP **organisation self-asserts**. The Sirtfi trust framework posits that organisations asserting conformance with these will coordinate their response to security incidents.

Derived from the first four elements of the SCI Framework:

- **Operational Security**: patch and vulnerability management; IDS and threat mitigation; service ownership management; user suspension and termination; CSIRT capability
- **Incident Response**: CSIRT contact in meta-data; timely response; collaborate in IR; defined processes; privacy respect; TLP information sharing
- **Traceability**: timestamped accurate logs are available; log retention process in place
- **Participant Responsibilities**: users agree to an AUP; awareness and acceptance of the AUP

https://refeds.org/SIRTFI

The eduGAIN Security Handbook



Pages / eduGAIN Home

eduGAIN Security

Created by Davide Vaghetti, last modified by Licia Florio on Apr 13, 2022

The eduGAIN Security Team main duty is to provide a central coordination point at the inter-federation
Moreover, the team will share information on security threats relevant for the eduGAIN community.

While each Federation Operator and Federation Participant provides security support within their respec
remains everybody's responsibility, which means no entity is effectively accountable to do the necessary
attacks targeting global services, inter-federation must be at the core of incident response strategy.
The eduGAIN Security Team supports this collective responsibility in inter-federation incident response v

The eduGAIN Security Team is a central contact and support point for security incidents, and coordinate
security incidents that affect Federation Operators and Federation Participants. This includes notifying Fe
or any other relevant entity about attacks potentially affecting them.

The collective expertise and experience accumulated by the eduGAIN community as it defends against a
Team ensures that lessons learned, statistics, and other useful information are disseminated appropriate
united community.

eduGAIN Security Incident Response Handbook

The eduGAIN Security Team in collaboration with the REFEDS Sirtfi WG developed an eduGAIN Security Incident Response (SIR) Handbook, which after REFEDS consultation (see https://wiki.refeds.org/x/-oCNAw) is now promoted across eduGAIN community for adoption.

The eduGAIN SIR handbook defines the process for resolving security incidents affecting eduGAIN participants involving all key stakeholders. In particular, it is essential to involve the federation in security operations or possible intrusions affecting eduGAIN entities.

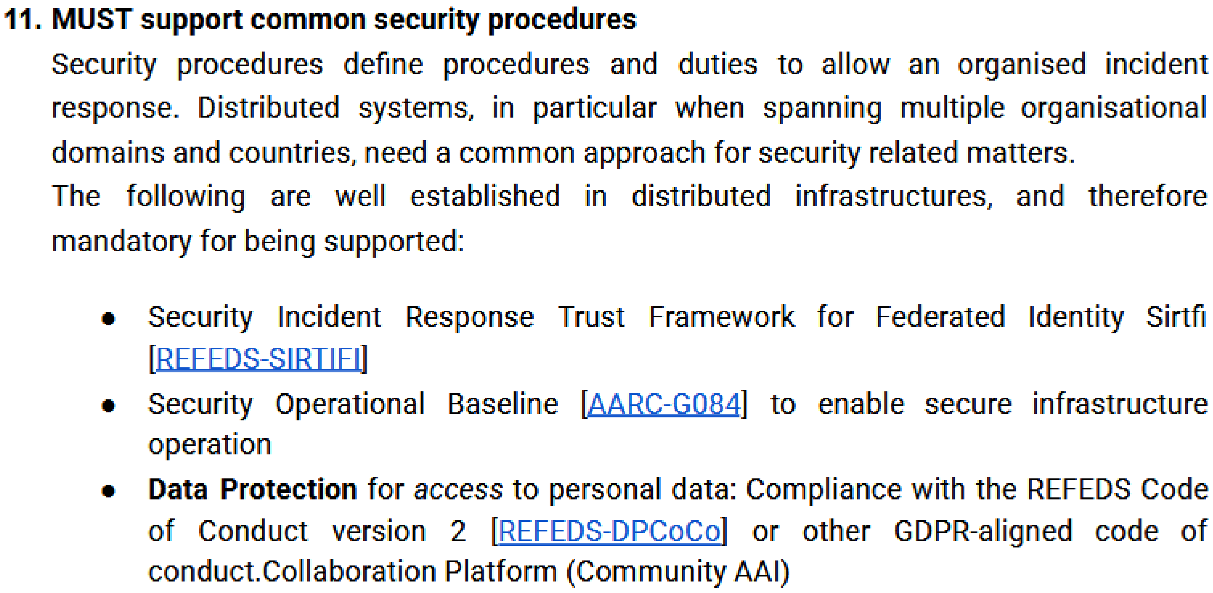
eduGAIN Security Incident Response Handbook

Preface

As with products of any REFEDS Working Group, in this instance the SIRTFI Working Group,

https://edugain.org/edugain-security/references/ eduGAIN Security activities supported by the GN4-3 and GN5-1 Trust and Identity activities

And security is a mandatory corner stone also of the EOSC

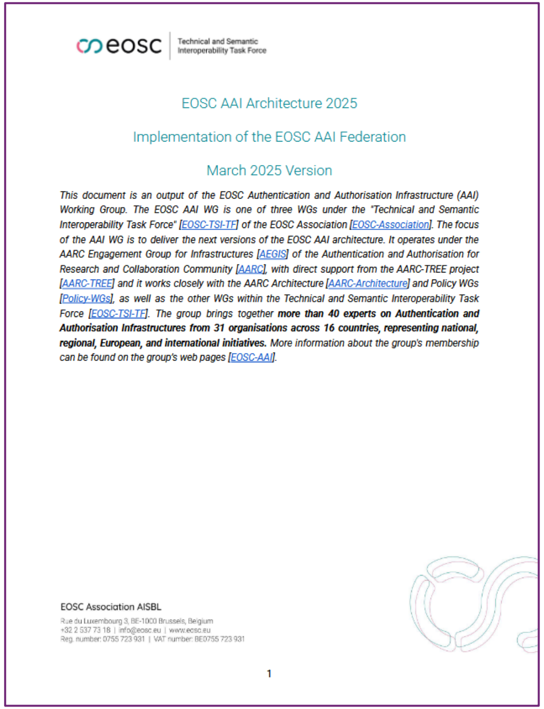


11. MUST support common security procedures

Security procedures define procedures and duties to allow an organised incident response. Distributed systems, in particular when spanning multiple organisational domains and countries, need a common approach for security related matters.

The following are well established in distributed infrastructures, and therefore mandatory for being supported:

- Security Incident Response Trust Framework for Federated Identity Sirtfi [REFEDS-SIRTIFI]
- Security Operational Baseline [AARC-G084] to enable secure infrastructure operation
- **Data Protection** for *access* to personal data: Compliance with the REFEDS Code of Conduct version 2 [REFEDS-DPCoCo] or other GDPR-aligned code of conduct.Collaboration Platform (Community AAI)



eosc | Technical and Semantic Interoperability Task Force

EOSC AAI Architecture 2025

Implementation of the EOSC AAI Federation

March 2025 Version

This document is an output of the EOSC Authentication and Authorisation Infrastructure (AAI) Working Group. The EOSC AAI WG is one of three WGs under the "Technical and Semantic Interoperability Task Force" [EOSC-TSI-TF] of the EOSC Association [EOSC-Association]. The focus of the AAI WG is to deliver the next versions of the EOSC AAI architecture. It operates under the AARC Engagement Group for Infrastructures [AEGIS] of the Authentication and Authorisation for Research and Collaboration Community [AARC], with direct support from the AARC-TREE project [AARC-TREE] and it works closely with the AARC Architecture [AARC-Architecture] and Policy WGs [Policy-WGs], as well as the other WGs within the Technical and Semantic Interoperability Task Force [EOSC-TSI-TF]. The group brings together ***more than 40 experts on Authentication and Authorisation Infrastructures from 31 organisations across 16 countries, representing national, regional, European, and international initiatives.*** *More information about the group's membership can be found on the group's web pages [EOSC-AAI].*

EOSC Association AISBL
Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

1

EOSC AAI Architecture 2025 - Implementation of the EOSC AAI Federation; https://doi.org/10.5281/zenodo.15388270

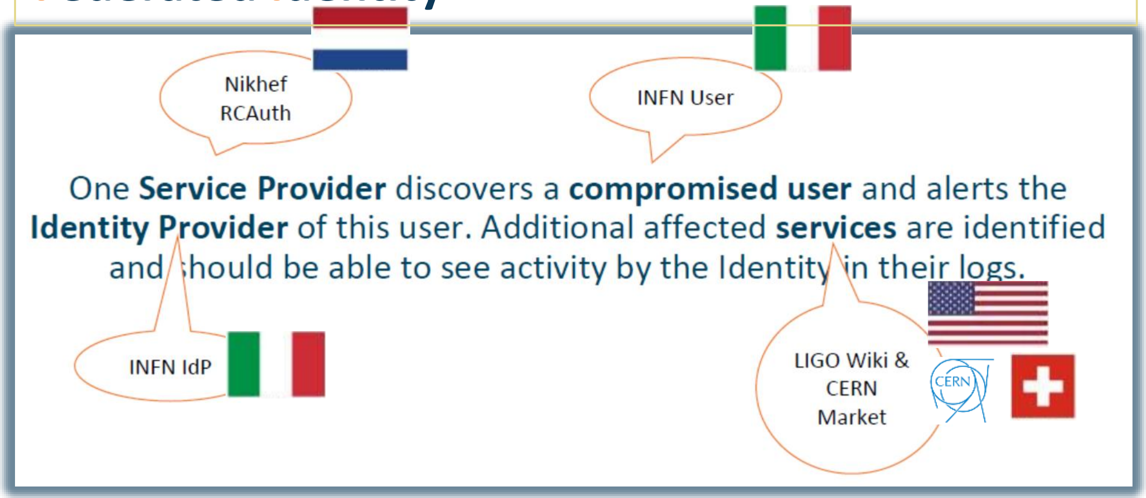
A federated community security challenge

SIRTFI
Security Incident Response Trust
Framework for Federated Identity

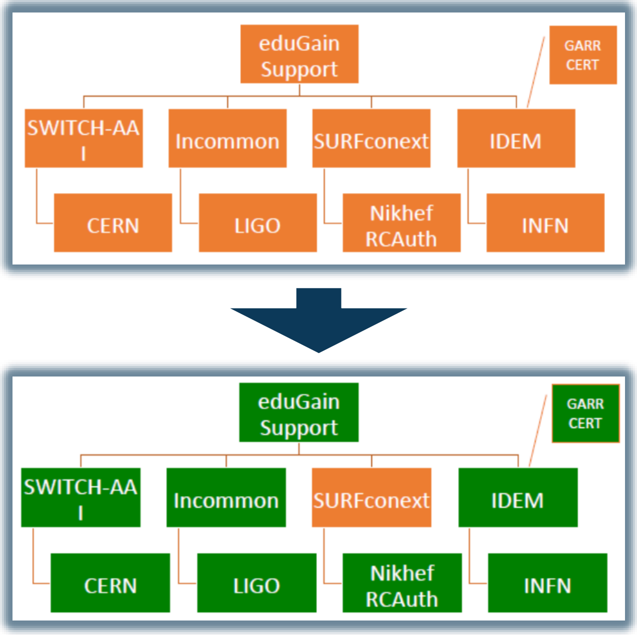
Can we coordinate our collective R&E response?

‘challenges’ based on the *Sirtfi* contact model

Security Incident Response Trust Framework for Federated Identity



One **Service Provider** discovers a **compromised user** and alerts the **Identity Provider** of this user. Additional affected **services** are identified and should be able to see activity by the Identity in their logs.

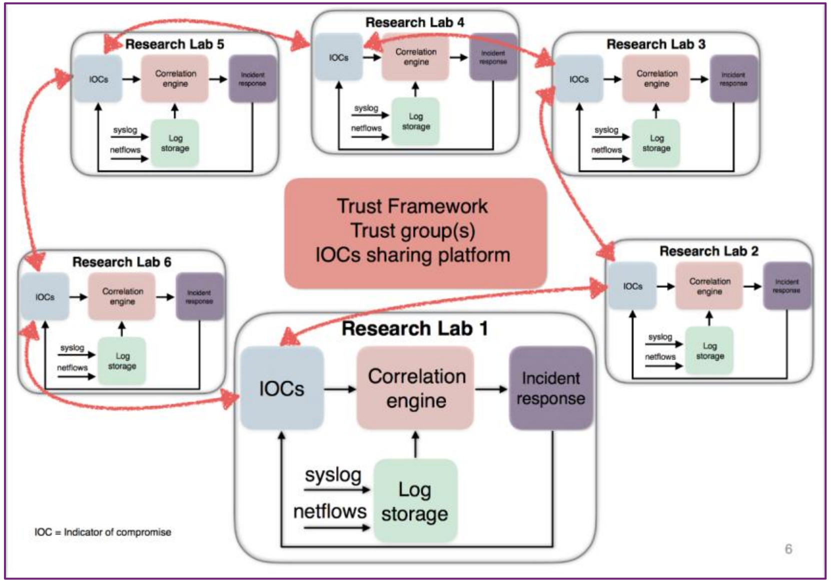


parties involved in response challenge

Report-outs see **https://wiki.geant.org/display/AARC/Sirtfi+Communications+Challenges%2C+AARC2-TNA3.1**

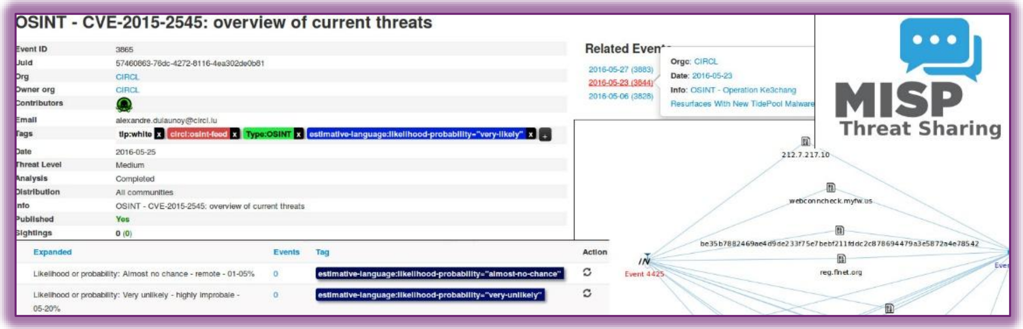
Sharing threat intel – working with our community

MITRE | ATT&CK®



IOC = Indicator of compromise

AARC I-051 Guide to federated incident response
https://aarc-community.org/guidelines/aarc-i051/



OSINT - CVE-2015-2545: overview of current threats

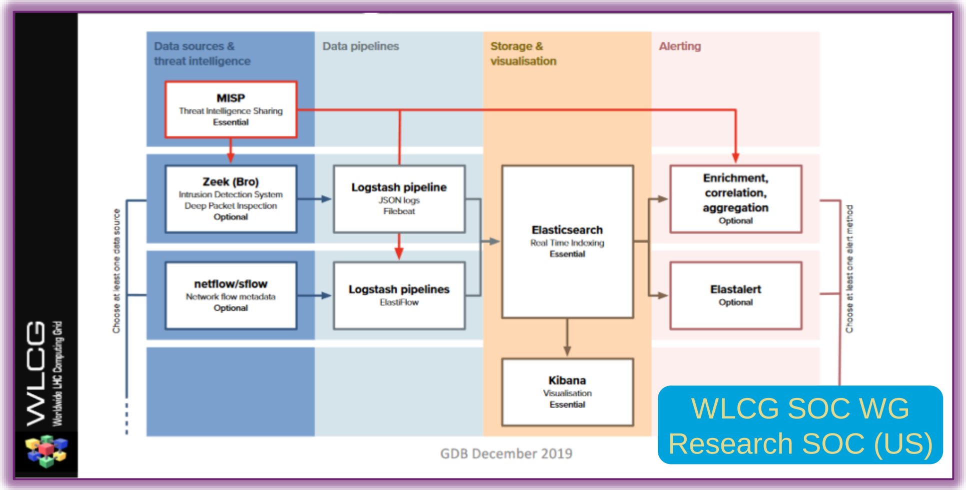
Event ID	3865
Uuid	57460863-76dc-4272-8116-4ea302de0b81
Org	CIRCL
Owner org	CIRCL
Contributors	
Email	alexandre.dulaunoy@circl.lu
Tags	tlp:white x circl:osint-feed x Type:OSINT x estimative-language:likelihood-probability="very-likely" x +
Date	2016-05-25
Threat Level	Medium
Analysis	Completed
Distribution	All communities
Info	OSINT - CVE-2015-2545: overview of current threats
Published	Yes
Sightings	0 (0)

Related Even…

2016-05-27 (3883)
2016-05-23 (3844)
2016-05-06 (3828)

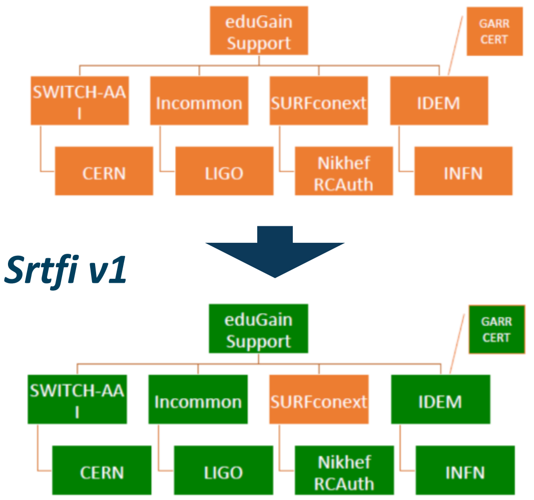
Orgc: CIRCL
Date: 2016-05-23
Info: OSINT - Operation Ke3chang Resurfaces With New TidePool Malware

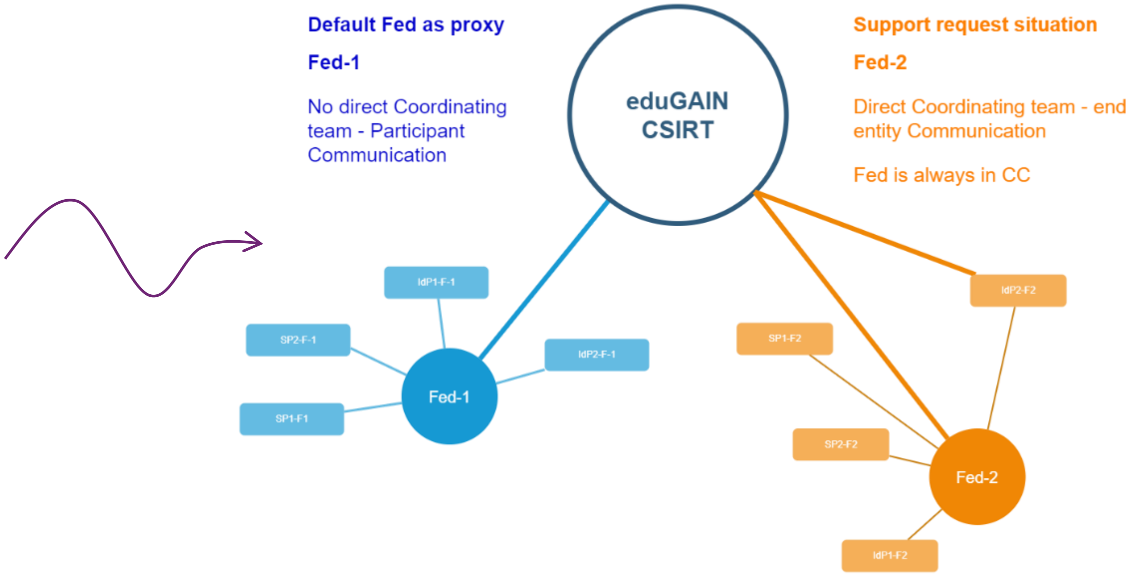
Expanded	Events	Tag	Action
Likelihood or probability: Almost no chance - remote - 01-05%	0	estimative-language:likelihood-probability="almost-no-chance"	
Likelihood or probability: Very unlikely - highly improbale - 05-20%	0	estimative-language:likelihood-probability="very-unlikely"	



WLCG SOC WG
Research SOC (US)

Response across IdP-SP Proxies: the limits of Sirtfi version 1





Default Fed as proxy

Fed-1

No direct Coordinating team - Participant Communication

Support request situation

Fed-2

Direct Coordinating team - end entity Communication

Fed is always in CC

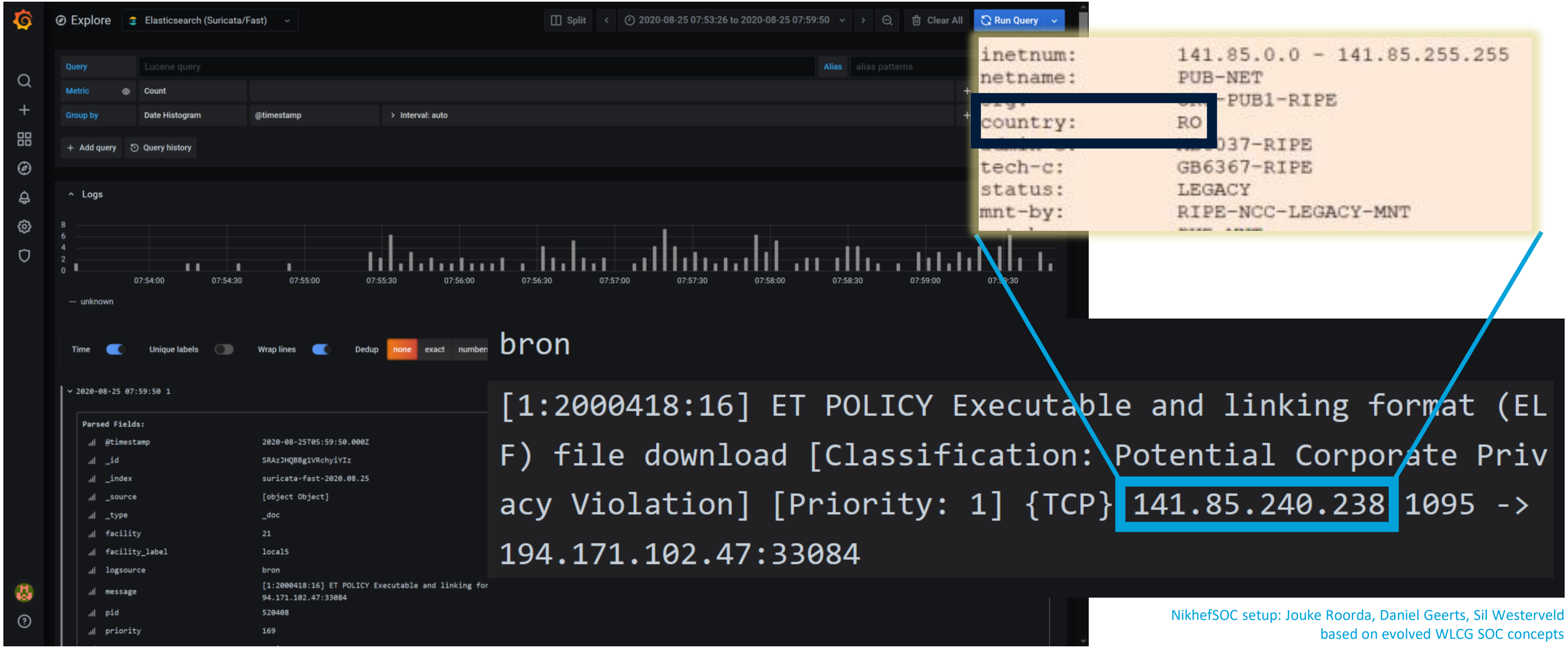
GÉANT | eduGAIN | CSIRT

joint work with GN5 EnCo and eduGAIN CSIRT

eduGAIN Security Incident Response Handbook

A single site sees only so much …

many ‘false warnings’ when industry-standard (e.g. Suricata) rules are used. You need R&E specific ones!



```
2020-08-25 07:59:50 1

Parsed Fields:
  @timestamp        2020-08-25T05:59:50.000Z
  _id               SRAzJHQBBg1VRchyiYIz
  _index            suricata-fast-2020.08.25
  _source           [object Object]
  _type             _doc
  facility          21
  facility_label    local5
  logsource         bron
  message           [1:2000418:16] ET POLICY Executable and linking for
                    94.171.102.47:33084
  pid               520408
  priority          169
```

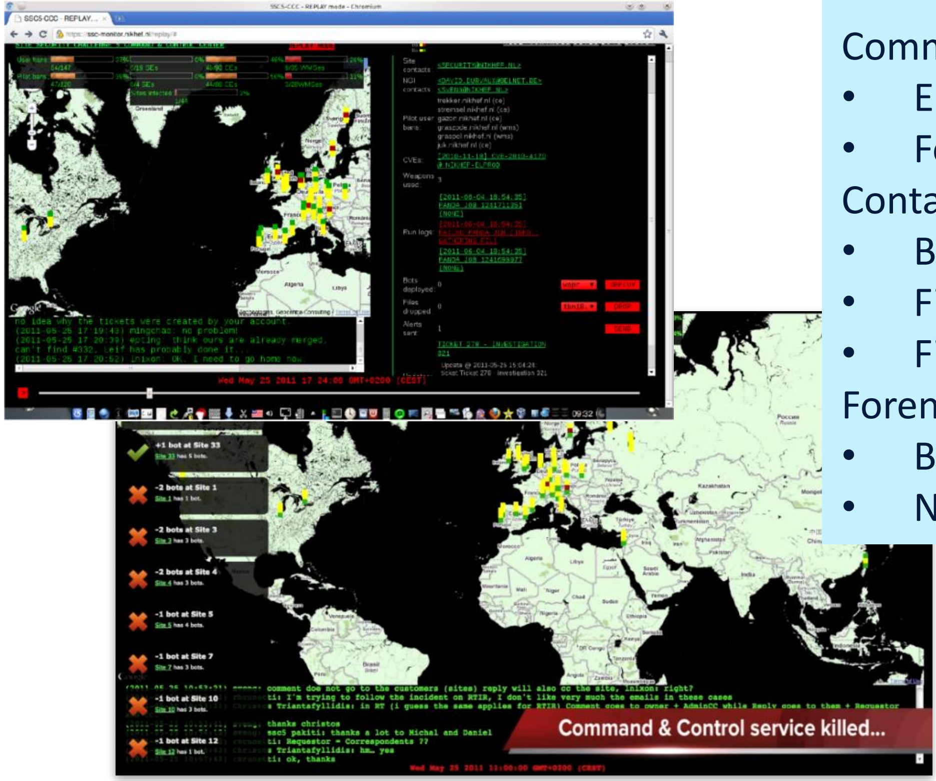
```
inetnum:        141.85.0.0 - 141.85.255.255
netname:        PUB-NET
                -PUB1-RIPE
country:        RO
                037-RIPE
tech-c:         GB6367-RIPE
status:         LEGACY
mnt-by:         RIPE-NCC-LEGACY-MNT
```

```
bron

[1:2000418:16] ET POLICY Executable and linking format (EL
F) file download [Classification: Potential Corporate Priv
acy Violation] [Priority: 1] {TCP} 141.85.240.238:1095 ->
194.171.102.47:33084
```

NikhefSOC setup: Jouke Roorda, Daniel Geerts, Sil Westerveld
based on evolved WLCG SOC concepts

A question of *when*, not *if* – hence we run security challenges



<SECURITY@NIKHEF.NL>
<DAVID.DURVAUX@BELNET.BE>
<SVENG@NIKHEF.NL>
trekker.nikhef.nl (ce)
stremsel.nikhef.nl (ce)
gazon.nikhef.nl (ce)
graszode.nikhef.nl (wms)
graspol.nikhef.nl (wms)
juk.nikhef.nl (ce)
[2010-11-18] CVE-2010-4170 @ NIKHEF-ELPROD
3
[2011-06-04 18:54:35] PANDA_JOB_1241711351 (NONE)
[2011-06-04 18:54:35] FAILED PANDA JOB (INFO - GATHERING FIL)
[2011-06-04 18:54:35] PANDA_JOB_1241699977 (NONE)
0
0
1
TICKET 278 - INVESTIGATION 321
Update @ 2011-05-25 15:04:28: ticket Ticket 278 - investigation 321

```
no idea why the tickets were created by your account.
(2011-05-25 17:19:43) mingchao: no problem!
(2011-05-25 17:20:39) epting: think ours are already merged.
can't find #332. Leif has probably done it...
(2011-05-25 17:20:52) inixon: OK. I need to go home now.
```

Wed May 25 2011 17:24:00 GMT+0200 (CEST)

+1 bot at Site 33
Site 33 has 5 bots.

-2 bots at Site 1
Site 1 has 1 bot.

-2 bots at Site 3
Site 3 has 3 bots.

-2 bots at Site 4
Site 4 has 3 bots.

-1 bot at Site 5
Site 5 has 4 bots.

-1 bot at Site 7
Site 7 has 3 bots.

-1 bot at Site 10
Site 10 has 3 bots.

-1 bot at Site 12
Site 12 has 1 bot.

```
(2011-05-25 10:52:21) svenag: comment doe not go to the customers (sites) reply will also cc the site, inixon: right?
christos ti: I'm trying to follow the incident on RTIR, I don't like very much the emails in these cases
Christos Triantafyllidis: in RT (i guess the same applies for RTIR) Comment goes to owner + AdminCC while Reply goes to them + Requestor
svenag: thanks christos
svenag: ssc5 pakiti: thanks a lot to Michal and Daniel
christos ti: Requestor = Correspondents ??
Christos Triantafyllidis: hm… yes
christos ti: ok, thanks
```

Command & Control service killed...

Wed May 25 2011 11:00:00 GMT+0200 (CEST)

Communication:

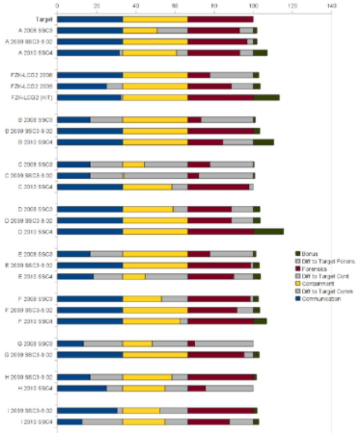
- Endpoints valid?
- Form/Content OK ?

Containment

- Ban ”malicious” users
- Find/Stop malicious processes
- Find submission IP

Forensics

- Basic Forensics on binary
- Network traffic



Nikhef

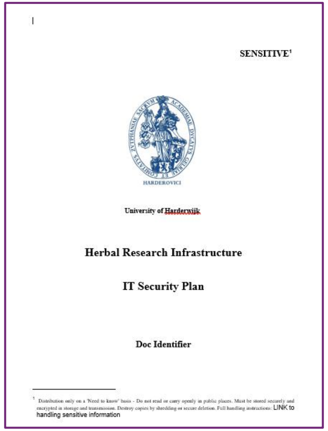
Nikhef CSIRT Traceability Challenge

Introduction

Deze Traceability Challenge bestaat uit drie onderdelen, in (naar verwachting) oplopende moeilijkheidsgraad. Iedere challenge begint met een externe 'trigger' – aan het eind van dit document staan de hints en de goede (of in ieder geval: de 'gewenste') oplossing.

Veel plezier!

Federation security table-top exercises



SENSITIVE[1]

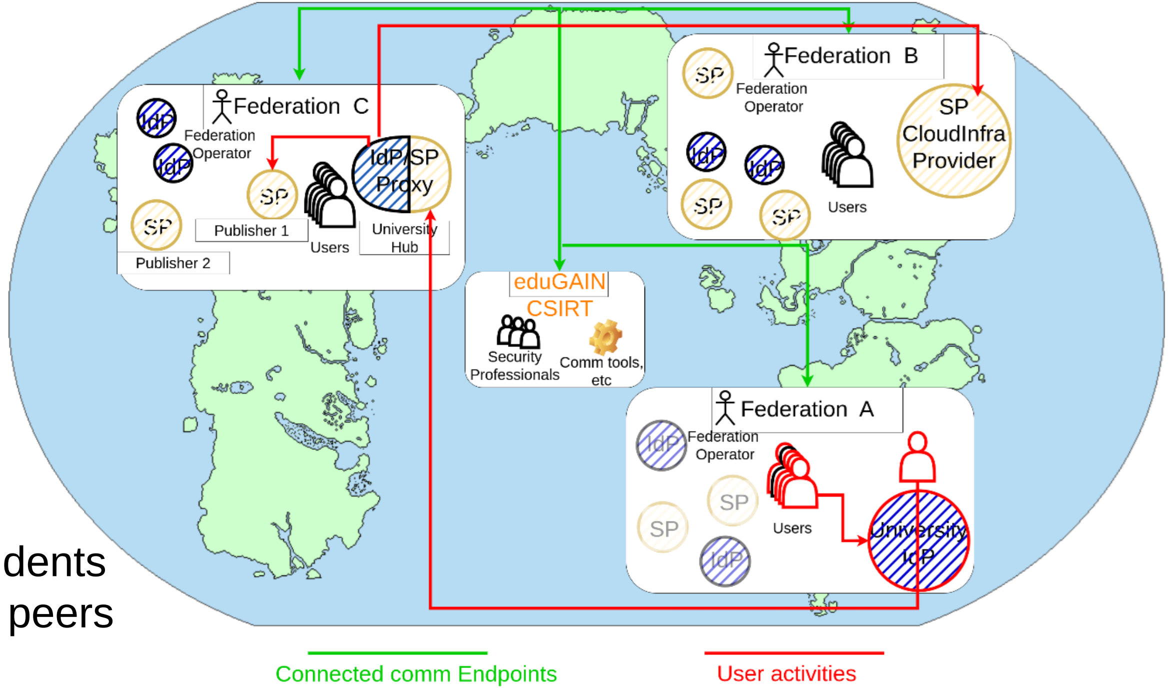
University of Harderwijk

Herbal Research Infrastructure

IT Security Plan

Doc Identifier

[1] Distribution only on a 'Need to know' basis - Do not read or carry openly in public places. Must be stored securely and encrypted in storage and transmission. Destroy copies by shredding or secure deletion. Full handling instructions: LINK to handling sensitive information

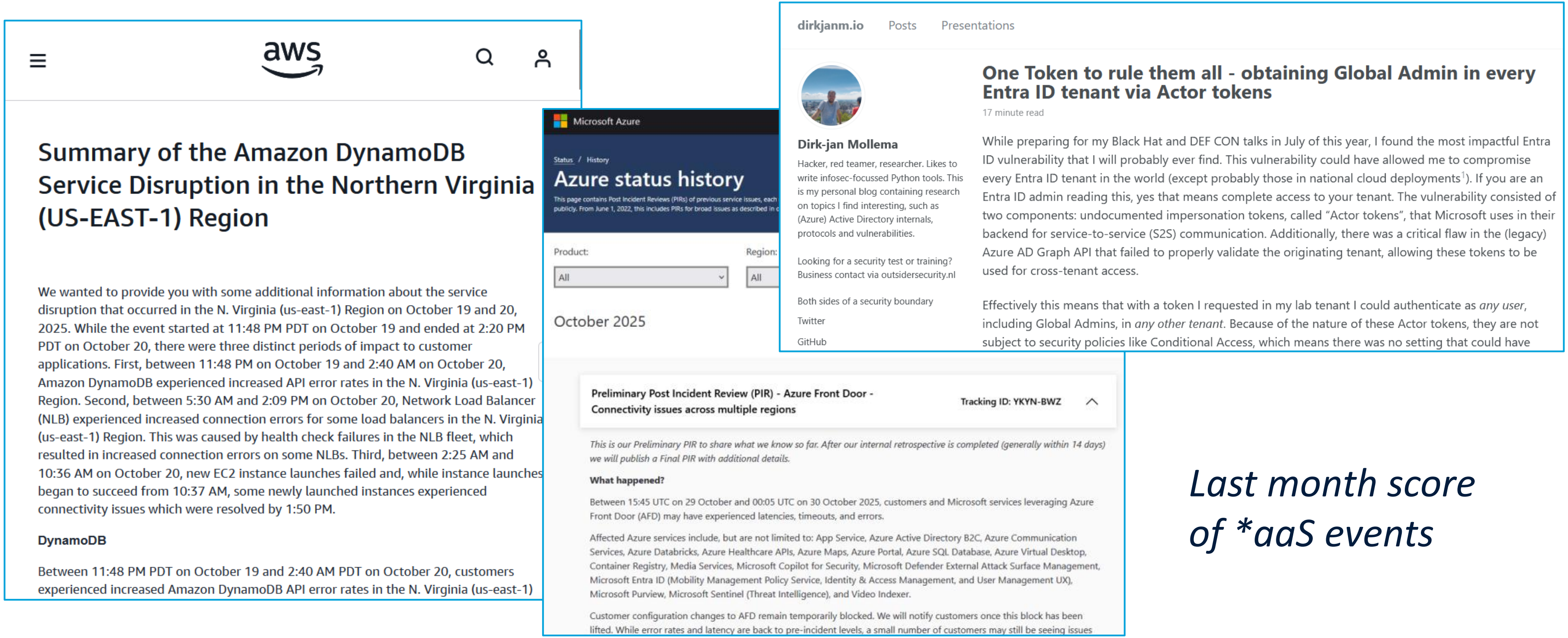


What your role play brings you ☺

- real time pressure to contain incidents
- true gratitude for protecting your peers
- collective recovery
- exploring some gruelling conflicts of interest!

eduGAIN TTX – role play scenario from the ISGC Security Workshop 2024, 2025

On leaky abstractions and circular dependencies



Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region

We wanted to provide you with some additional information about the service disruption that occurred in the N. Virginia (us-east-1) Region on October 19 and 20, 2025. While the event started at 11:48 PM PDT on October 19 and ended at 2:20 PM PDT on October 20, there were three distinct periods of impact to customer applications. First, between 11:48 PM on October 19 and 2:40 AM on October 20, Amazon DynamoDB experienced increased API error rates in the N. Virginia (us-east-1) Region. Second, between 5:30 AM and 2:09 PM on October 20, Network Load Balancer (NLB) experienced increased connection errors for some load balancers in the N. Virginia (us-east-1) Region. This was caused by health check failures in the NLB fleet, which resulted in increased connection errors on some NLBs. Third, between 2:25 AM and 10:36 AM on October 20, new EC2 instance launches failed and, while instance launches began to succeed from 10:37 AM, some newly launched instances experienced connectivity issues which were resolved by 1:50 PM.

DynamoDB

Between 11:48 PM PDT on October 19 and 2:40 AM PDT on October 20, customers experienced increased Amazon DynamoDB API error rates in the N. Virginia (us-east-1)

Azure status history

This page contains Post Incident Reviews (PIRs) of previous service issues, each publicly. From June 1, 2022, this includes PIRs for broad issues as described in c

Product: All

Region: All

October 2025

Preliminary Post Incident Review (PIR) - Azure Front Door - Connectivity issues across multiple regions Tracking ID: YKYN-BWZ

This is our Preliminary PIR to share what we know so far. After our internal retrospective is completed (generally within 14 days) we will publish a Final PIR with additional details.

What happened?

Between 15:45 UTC on 29 October and 00:05 UTC on 30 October 2025, customers and Microsoft services leveraging Azure Front Door (AFD) may have experienced latencies, timeouts, and errors.

Affected Azure services include, but are not limited to: App Service, Azure Active Directory B2C, Azure Communication Services, Azure Databricks, Azure Healthcare APIs, Azure Maps, Azure Portal, Azure SQL Database, Azure Virtual Desktop, Container Registry, Media Services, Microsoft Copilot for Security, Microsoft Defender External Attack Surface Management, Microsoft Entra ID (Mobility Management Policy Service, Identity & Access Management, and User Management UX), Microsoft Purview, Microsoft Sentinel (Threat Intelligence), and Video Indexer.

Customer configuration changes to AFD remain temporarily blocked. We will notify customers once this block has been lifted. While error rates and latency are back to pre-incident levels, a small number of customers may still be seeing issues

Dirk-jan Mollema

Hacker, red teamer, researcher. Likes to write infosec-focussed Python tools. This is my personal blog containing research on topics I find interesting, such as (Azure) Active Directory internals, protocols and vulnerabilities.

Looking for a security test or training? Business contact via outsidersecurity.nl

Both sides of a security boundary

Twitter

GitHub

One Token to rule them all - obtaining Global Admin in every Entra ID tenant via Actor tokens

17 minute read

While preparing for my Black Hat and DEF CON talks in July of this year, I found the most impactful Entra ID vulnerability that I will probably ever find. This vulnerability could have allowed me to compromise every Entra ID tenant in the world (except probably those in national cloud deployments[1]). If you are an Entra ID admin reading this, yes that means complete access to your tenant. The vulnerability consisted of two components: undocumented impersonation tokens, called “Actor tokens”, that Microsoft uses in their backend for service-to-service (S2S) communication. Additionally, there was a critical flaw in the (legacy) Azure AD Graph API that failed to properly validate the originating tenant, allowing these tokens to be used for cross-tenant access.

Effectively this means that with a token I requested in my lab tenant I could authenticate as *any user*, including Global Admins, in *any other tenant*. Because of the nature of these Actor tokens, they are not subject to security policies like Conditional Access, which means there was no setting that could have

*Last month score of *aaS events*

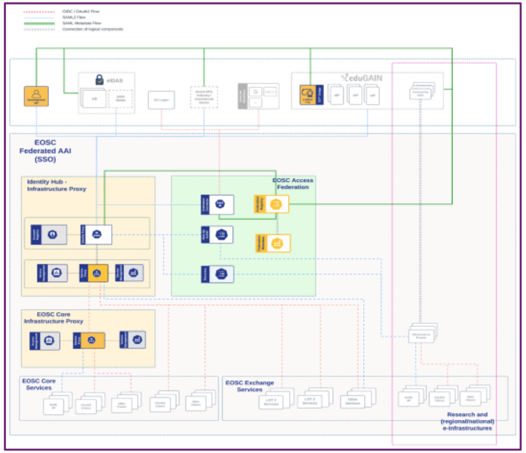
https://aws.amazon.com/message/101925/, https://azure.status.microsoft/en-us/status/history/, https://dirkjanm.io/obtaining-global-admin-in-every-entra-id-tenant-with-actor-tokens/

Collaboration for research

Enabling research: using the 'EOSC' with federated login

AARC compliant federation of 'national' and 'thematic' nodes in the European Open Science Cloud

linked with other 'data spaces' and infrastructures



eosc | Federation

ESRF
panosc
Italian National Research Council
eosc | Blue-Cloud2026
CERN
NATIONAL SCIENCE CENTRE POLAND
ICSC
Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing
nfdi
Nationale Forschungsdateninfrastruktur
cnrs
DATA TERRA
EUDAT
CSC
SURF
BBMRI-ERIC
elixir
EMBL
EURO BIOIMAGING
instruct ERIC

The organisations invited to join the March 2025 kick-off workshop for the build-up phase of the EOSC Federation. All of the organisations are among the membership of the EOSC Association.

https://eosc.eu/eosc-about/building-the-eosc-federation/contributing-to-the-build-up-phase-of-the-eosc-federation/; See also https://wiki.geant.org/display/AARC/EOSC+AAI

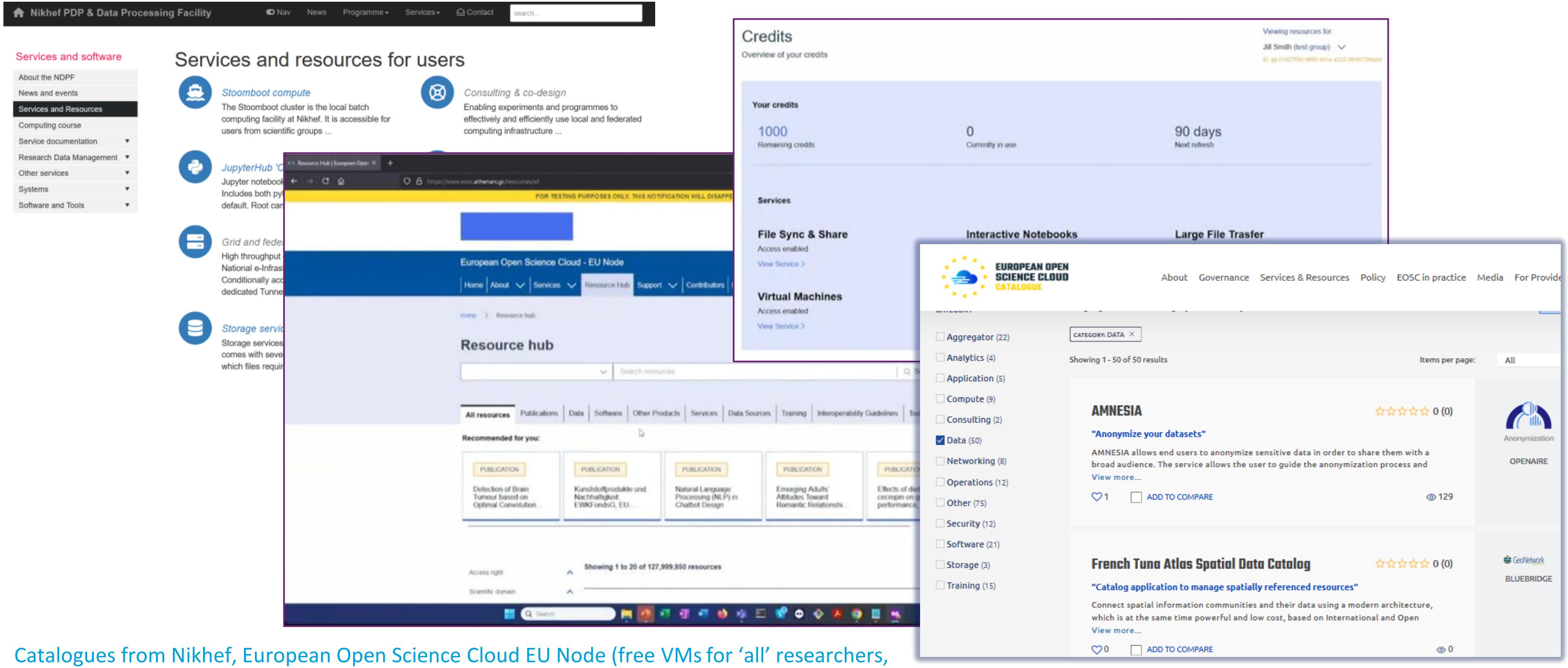
Where do researchers find services & collaboration …

FIM4L FIM4R

SeaDataNet
BBMRI-ERIC®
WLCG
Worldwide LHC Computing Grid
DARIAH-EU
LIFE SCIENCE RI
EISCAT SCIENTIFIC ASSOCIATION
KM3NeT
SSHOC
social sciences & humanities open cloud
elixir
cta
cherenkov telescope array
ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures
SKA
SQUARE KILOMETRE ARRAY
CLARIN
neic
PUHURI
panosc
photon and neutron open science cloud
LUMI
FENIX
RESEARCH INFRASTRUCTURE
eosc
EUDAT
Collaborative Data Infrastructure
egi
GÉANT
OpenAIRE
iris
SURF
nfdi
HIFIS
HELMHOLTZ FEDERATED IT SERVICES
Open Science Grid
HPCI
High Performance Computing Infrastructure of Japan
RIST
Open to the Global HPC Communities
ACCESS

represented by logos: some of the (AARC BPA) Research Communities (top) providing federated access using the AAI proxy architecture.
At the ~ bottom: (global) e-Infrastructures, which all use the AARC BPA collaborative model

Service portfolios – what do you offer, and to whom



Nikhef PDP & Data Processing Facility

Nav | News | Programme | Services | Contact | search...

Services and software

- About the NDPF
- News and events
- Services and Resources
- Computing course
- Service documentation
- Research Data Management
- Other services
- Systems
- Software and Tools

Services and resources for users

Stoomboot compute
The Stoomboot cluster is the local batch computing facility at Nikhef. It is accessible for users from scientific groups ...

Consulting & co-design
Enabling experiments and programmes to effectively and efficiently use local and federated computing infrastructure ...

JupyterHub 'C
Jupyter noteboo
Includes both py
default. Root can

Grid and fede
High throughput
National e-Infras
Conditionally acc
dedicated Tunne

Storage servic
Storage services
comes with seve
which files requi

Resource Hub | European Open

FOR TESTING PURPOSES ONLY. THIS NOTIFICATION WILL DISAPPE

European Open Science Cloud - EU Node

Home | About | Services | Resource Hub | Support | Contributors

Home > Resource hub

Resource hub

Search resources

All resources | Publications | Data | Software | Other Products | Services | Data Sources | Training | Interoperability Guidelines | Too

Recommended for you:

PUBLICATION Detection of Brain Tumour based on Optimal Convolution...	PUBLICATION Kunststoffprodukte und Nachhaltigkeit: EWKFondsG, EU-...	PUBLICATION Natural Language Processing (NLP) in Chatbot Design	PUBLICATION Emerging Adults' Attitudes Toward Romantic Relationshi...	PUBLICATIO Effects of diet cecropin on g performance,

Access right

Scientific domain

Showing 1 to 20 of 127,999,850 resources

Credits

Overview of your credits

Viewing resources for:
Jill Smith (test group)

Your credits

1000	0	90 days
Remaining credits	Currently in use	Next refresh

Services

File Sync & Share
Access enabled
View Service >

Interactive Notebooks

Large File Trasfer

Virtual Machines
Access enabled
View Service >

EUROPEAN OPEN SCIENCE CLOUD CATALOGUE

About | Governance | Services & Resources | Policy | EOSC in practice | Media | For Provide

- □ Aggregator (22)
- □ Analytics (4)
- □ Application (5)
- □ Compute (9)
- □ Consulting (2)
- ☑ Data (50)
- □ Networking (8)
- □ Operations (12)
- □ Other (75)
- □ Security (12)
- □ Software (21)
- □ Storage (3)
- □ Training (15)

CATEGORY: DATA ✕

Showing 1 - 50 of 50 results

Items per page: All

AMNESIA ☆☆☆☆☆ 0 (0)

"Anonymize your datasets"

AMNESIA allows end users to anonymize sensitive data in order to share them with a broad audience. The service allows the user to guide the anonymization process and
View more...

♡1 □ ADD TO COMPARE 129

Anonymization
OPENAIRE

French Tuna Atlas Spatial Data Catalog ☆☆☆☆☆ 0 (0)

"Catalog application to manage spatially referenced resources"

Connect spatial information communities and their data using a modern architecture, which is at the same time powerful and low cost, based on International and Open
View more...

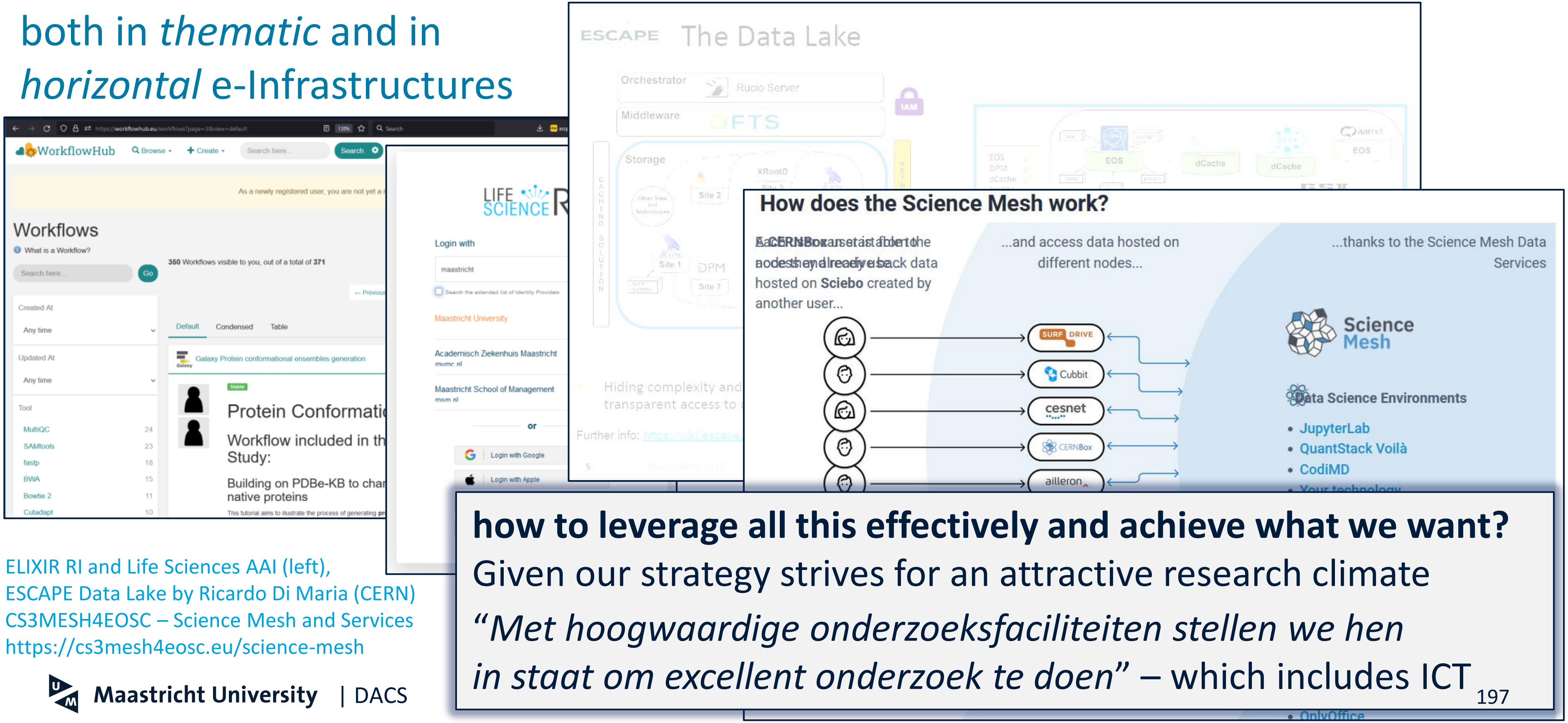
♡0 □ ADD TO COMPARE 0

GeoNetwork
BLUEBRIDGE

Catalogues from Nikhef, European Open Science Cloud EU Node (free VMs for 'all' researchers, subject to https://open-science-cloud.ec.europa.eu/system/files?file=2024-10/EOSC-EU-Node-User-Access-Policy-v1.0.pdf)

‘Services await us’ in global research & e-infrastructures

both in *thematic* and in *horizontal* e-Infrastructures



WorkflowHub · Browse · Create · Search here... · Search

As a newly registered user, you are not yet a

Workflows

What is a Workflow?

350 Workflows visible to you, out of a total of 371

Tool	
MultiQC	24
SAMtools	23
fastp	18
BWA	15
Bowtie 2	11
Cutadapt	10

Created At: Any time

Updated At: Any time

Default · Condensed · Table

Galaxy Protein conformational ensembles generation

Stable

Protein Conformati

Workflow included in th Study:

Building on PDBe-KB to char native proteins

This tutorial aims to illustrate the process of generating pr

LIFE SCIENCE R

Login with

maastricht

☐ Search the extended list of Identity Providers

Maastricht University

Academisch Ziekenhuis Maastricht mumc.nl

Maastricht School of Management msm.nl

or

Login with Google

Login with Apple

ESCAPE The Data Lake

Orchestrator: Rucio Server

Middleware: FTS

IAM

Storage: XRootD, Site 1, Site 2, Site 3, Site 7, DPM, Other Sites and Technologies

EOS, DPM, dCache, EOS, dCache, dCache, EOS

Hiding complexity and transparent access to

Further info: https://wiki.escape

How does the Science Mesh work?

EaCERNBoxarsetatstabtetothe aodesthaydreadyusback data hosted on **Sciebo** created by another user...

...and access data hosted on different nodes...

...thanks to the Science Mesh Data Services

Data Science Environments

- JupyterLab
- QuantStack Voilà
- CodiMD
- Your technology
- OnlyOffice

how to leverage all this effectively and achieve what we want?

Given our strategy strives for an attractive research climate

“Met hoogwaardige onderzoeksfaciliteiten stellen we hen in staat om excellent onderzoek te doen” – which includes ICT

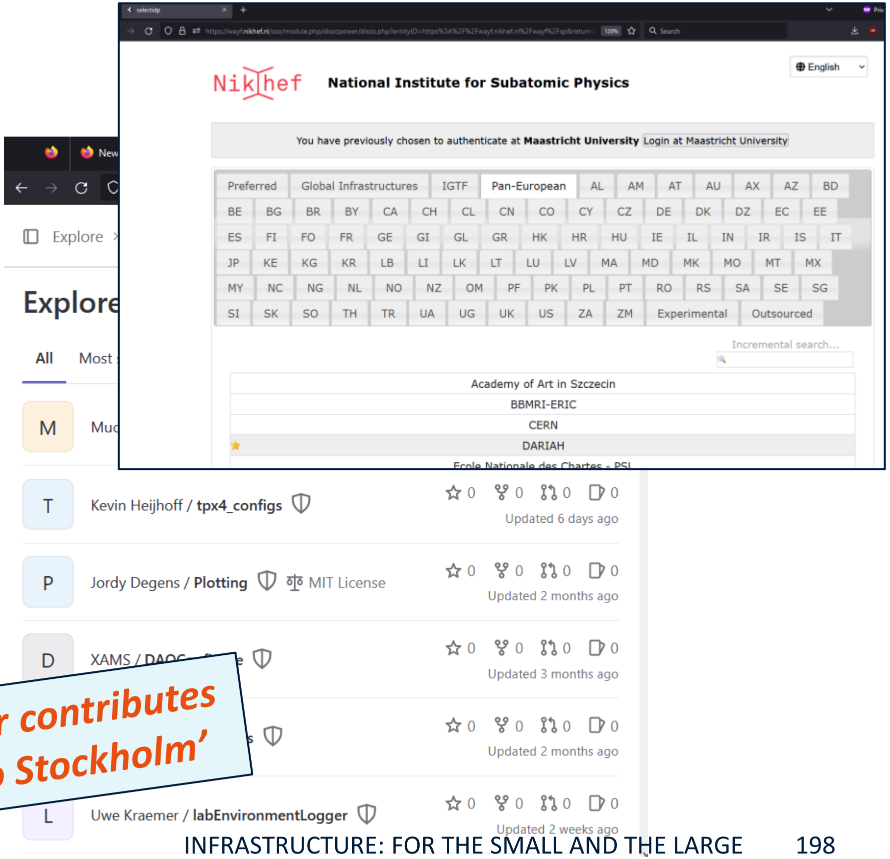
ELIXIR RI and Life Sciences AAI (left), ESCAPE Data Lake by Ricardo Di Maria (CERN) CS3MESH4EOSC – Science Mesh and Services https://cs3mesh4eosc.eu/science-mesh

Collaborative services are distributed and federated

Collaborative services are
spread across the research community

- logbooks with federated login from LIGO and IGWN for the ET pathfinder
- analysis notebooks and control software in open to the collaboration via eduGAIN
- our aforementioned RCauth.eu

need mix of local expertise and resources, national systems, research infra services, and European (global) resources



selectidp
English
Nikhef
National Institute for Subatomic Physics
Preferred
Global Infrastructures
IGTF
Pan-European
AL AM AT AU AX AZ BD
BE BG BR BY CA CH CL CN CO CY CZ DE DK DZ EC EE
ES FI FO FR GE GI GL GR HK HR HU IE IL IN IR IS IT
JP KE KG KR LB LI LK LT LU LV MA MD MK MO MT MX
MY NC NG NL NO NZ OM PF PK PL PT RO RS SA SE SG
SI SK SO TH TR UA UG UK US ZA ZM
Experimental
Outsourced
Incremental search...
New
Explore
Explore
All
Most

You have previously chosen to authenticate at **Maastricht University** Login at Maastricht University

Academy of Art in Szczecin
BBMRI-ERIC
CERN
DARIAH
Ecole Nationale des Chartes - PSL

M Mud

T Kevin Heijhoff / **tpx4_configs** ☆ 0 0 0 0 Updated 6 days ago

P Jordy Degens / **Plotting** MIT License ☆ 0 0 0 0 Updated 2 months ago

D XAMS / DAQC... e ☆ 0 0 0 0 Updated 3 months ago

s ☆ 0 0 0 0 Updated 2 months ago

L Uwe Kraemer / **labEnvironmentLogger** ☆ 0 0 0 0 Updated 2 weeks ago

'every partner contributes to the trip to Stockholm'

Infrastructure for research is an ecosystem: hardware, software, services, and … people

High-performance data processing

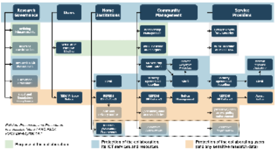
Do you want to process and store large volumes of data? Our team of experts can support you in using our high-throughput data processing systems and storage solutions.

For processing large, struc

Our high-throughput data processi
structured dat
escopes, and s
services for:

Sunday, Oct 6, 2019
● Bytes: **463 383 065 878 740 200**

Images: ATLAS Rucio volume, (from rucio.cern.ch); optical network: NDPF ‘deel’; User meeting Stoomboot Office Hours (both Nikhef)); Snellius opening visit; HPDC service page (both SURF)



Research Governance
Users
Home Institutions
Community Management
Service Providers
Purpose of the collaboration
Protection of the collaboration, its ICT services and resources
Protection of the collaborating users (and any sensitive research data)

… but what about that first pillar
of the Policy Development Kit?

Collaboration & governance

Make and treat computing as the research instrument it is today – institutionally and globally

Institutional:
Nikhef “Stoomboot”
Analysis Facility

National Infrastructure
SURF Snellius HPC

EuroHPC
Joint Undertaking

PRACE

as well as JP’s HPCI,
US’s AccessCI, &c of course!

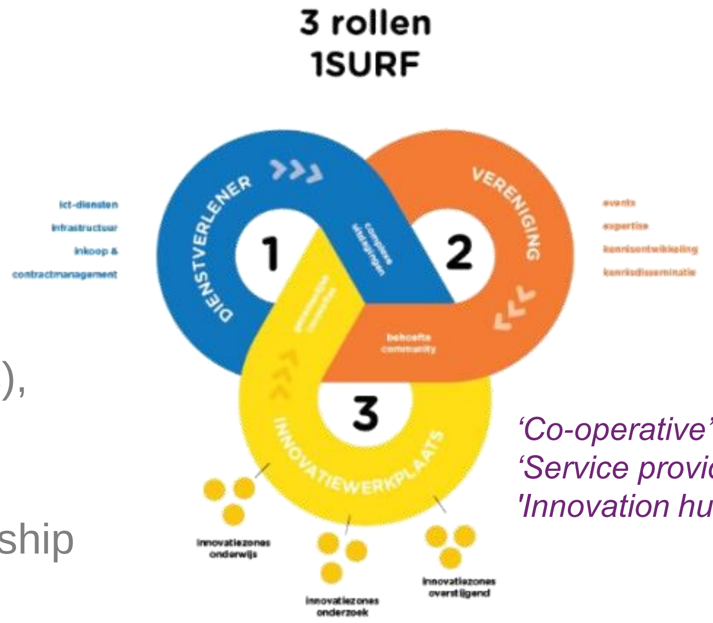
There are today as much part of science as detectors are to physics
and: users should move seamlessly between tiers

Photos: Nikhef NDPF, DelftBlue/TUDelft, SURF Data Repository, Snellius, SURF @ DigitalRealty; EuroHPC images: EuroHPC, LUMI Consortium, Jules Verne consortium

SURF, our ‘collaborative organisation for IT in edu and research’

Besides ‘commodity’ services (network, software licensing, joint procurement, &c) as a service provider, SURF adds

- **Research IT** facilities
  HPC ‘Snellius’, HPC Cloud, Grid, Data Archive, collaborative analysis platform, ResearchCloud, …
- **research support and open science**
  national coordination research data management (LCRDM), digital competence centre initiatives (DCCs), European Open Science Cloud, expertise networks
- **Federated access** and **collaboration**
  SURFconext, projects, international liaison & membership
- **Innovation** of the IT knowledge basis
  at SURF and at co-creating member organisations



‘Co-operative’
‘Service provider’
'Innovation hub’

Integral Approach to ICT Digital Competences for Research

- need for a federated networked scheme for data, computing (and expertise) remains as relevant today as it was in 2017
- ‘local’ digital competence centres in their role as “*node in a federated network for data, computing en expertise*” did not get attention for *infrastructure* that was intended
- expertise bundling and development of “Tier-2” facilities in national landscape set as institutional responsibility, (‘strengthening research support’) with some central funding
- but using national funding also means: be open to national collaboration, and ensure facilities (expertise, but also datasets, computing, storage, networks) are actually accessible in a FAIR and federated way, open to researchers from outside – based on e.g. federated SRAM, MyAccessID, or IGTF

Integrale aanpak voor digitalisering in de wetenschap

Uitvoeringsplan investeringen digitale onderzoeksinfrastructuur

A comprehensive approach to research digitalization

Specific execution plans at SURF for research support,
funded by the NWO ‘Apers’ means

Integrale aanpak voor digitalisering in de wetenschap

Uitvoeringsplan investeringen digitale onderzoeksinfrastructuur

NWO

A. Rekenfaciliteiten: Aanschaf **nieuwe supercomputer**

B. Rekenfaciliteiten: Toekomstige **investeringen in HPC**
   & jaarlijkse gebruikersbijdrage reserveringen HPC

C. Rekenfaciliteiten: Investeringen in **overige rekenfaciliteiten en opslag** hardware

D. Rekenfaciliteiten: **Vernieuwing van de kennisbasis**

E. Rekenfaciliteiten: **Expertise en ondersteuning** van rekenfaciliteiten en datacentra

F. Digitalisering: Impulsfinanciering **lokale DCC’s**

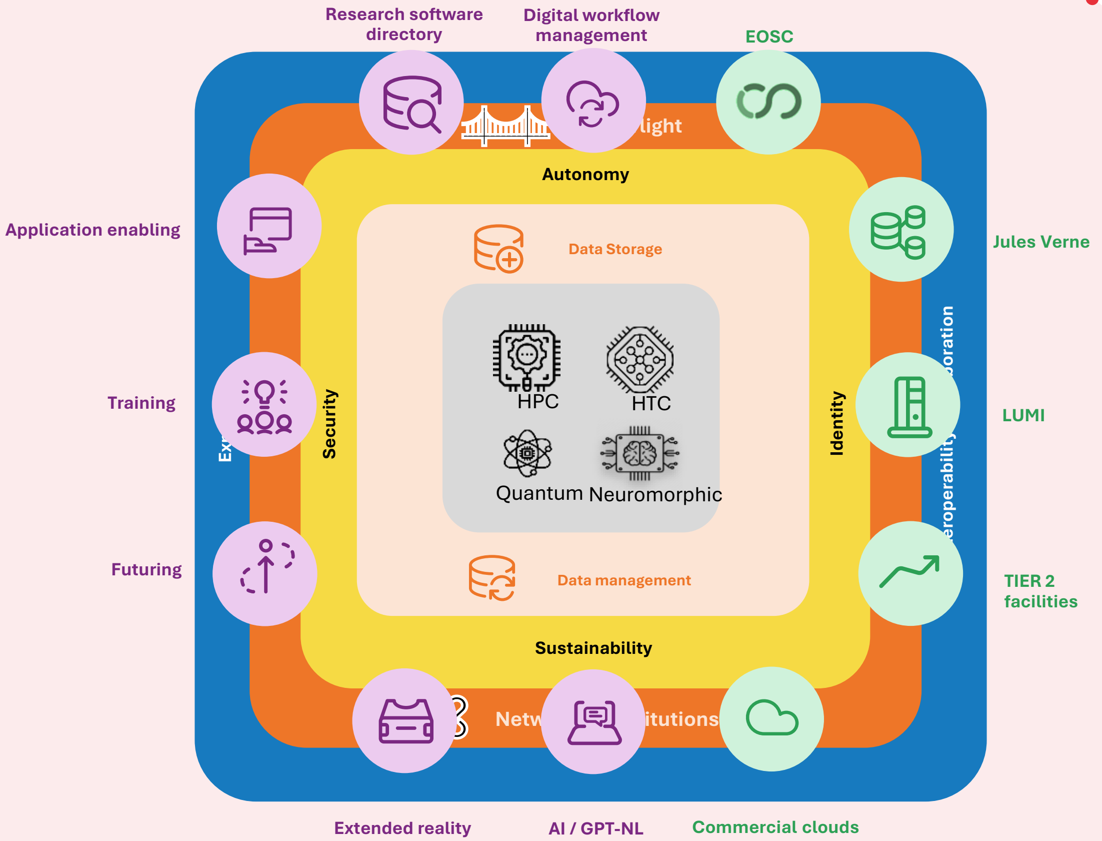
G. Digitalisering: Stimulering **thematische DCC’s**

H. Digitalisering: Ondersteuning DCC’s door SURF

I. Digitalisering: Investeringen in **eScience**

See: https://www.nwo.nl/onderzoeksprogrammas/uitvoeringsplan-ict-infrastructuur

SURF Research Infrastructure Ecosystem



Research software directory

Digital workflow management

EOSC

light

Autonomy

Application enabling

Jules Verne

Data Storage

HPC

HTC

Training

Ex

Security

Quantum

Neuromorphic

Identity

oration

eroperability

LUMI

Futuring

Data management

TIER 2 facilities

Sustainability

Netw

itutions

Extended reality

AI / GPT-NL

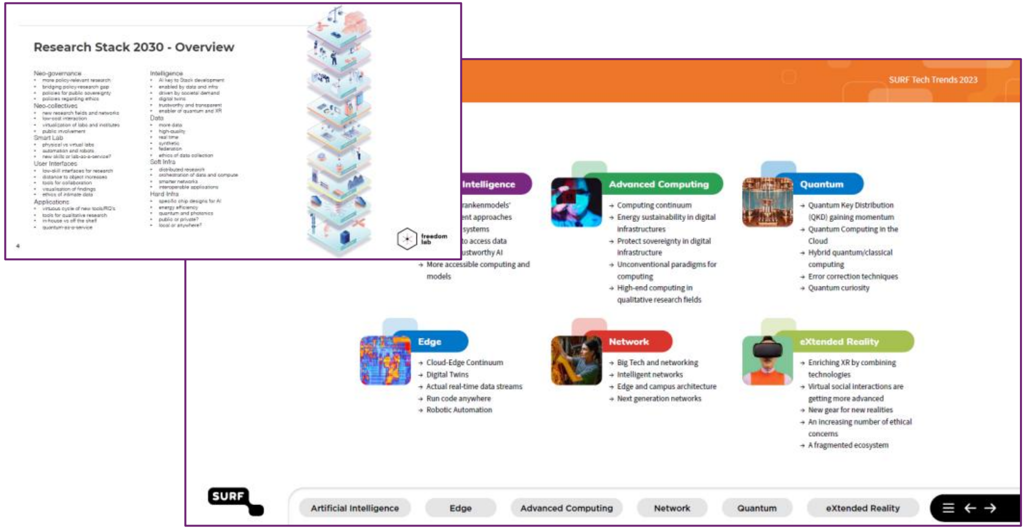
Commercial clouds

SURF

Acknowledgements: SURF, Jet de Ranitz, Magchiel Bijsterbosch

Joint innovation and the Dutch HPC Tier2 landscape

Representative expert group on knowledge basis innovation drives **joint innovation with institutes** on tech trends:
AI, Advanced Computing, Quantum, Edge, Network, XR



Research Stack 2030 - Overview

Neo-governance
- more policy-relevant research
- bridging policy-research gap
- policies for public sovereignty
- policies regarding ethics

Neo-collectives
- new research fields and networks
- low-cost interaction
- virtualization of labs and institutes
- public involvement

Smart Lab
- physical vs virtual labs
- automation and robots
- new skills or lab-as-a-service?

User Interfaces
- low-skill interfaces for research
- distance to object increases
- tools for collaboration
- visualisation of findings
- ethics of intimate data

Applications
- virtuous cycle of new tools/RQ's
- tools for qualitative research
- in-house vs off the shelf
- quantum-as-a-service

Intelligence
- AI key to Stack development
- enabled by data and infra
- driven by societal demand
- digital twins
- trustworthy and transparent
- enabler of quantum and XR

Data
- more data
- high-quality
- real time
- synthetic
- federation
- ethics of data collection

Soft Infra
- distributed research
- orchestration of data and compute
- smarter networks
- interoperable applications

Hard Infra
- specific chip designs for AI
- energy efficiency
- quantum and photonics
- public or private?
- local or anywhere?

4

freedom lab

SURF Tech Trends 2023

Intelligence
- rankenmodels'
- ent approaches
- systems
- to access data
- ustworthy AI
- More accessible computing and models

Advanced Computing
- Computing continuum
- Energy sustainability in digital infrastructures
- Protect sovereignty in digital infrastructure
- Unconventional paradigms for computing
- High-end computing in qualitative research fields

Quantum
- Quantum Key Distribution (QKD) gaining momentum
- Quantum Computing in the Cloud
- Hybrid quantum/classical computing
- Error correction techniques
- Quantum curiosity

Edge
- Cloud-Edge Continuum
- Digital Twins
- Actual real-time data streams
- Run code anywhere
- Robotic Automation

Network
- Big Tech and networking
- Intelligent networks
- Edge and campus architecture
- Next generation networks

eXtended Reality
- Enriching XR by combining technologies
- Virtual social interactions are getting more advanced
- New gear for new realities
- An increasing number of ethical concerns
- A fragmented ecosystem

Integrale aanpak voor digitalisering in de wetenschap

Uitvoeringsplan investeringen digitale onderzoeksinfrastructuur

Strengthening Tier-2 infrastructure

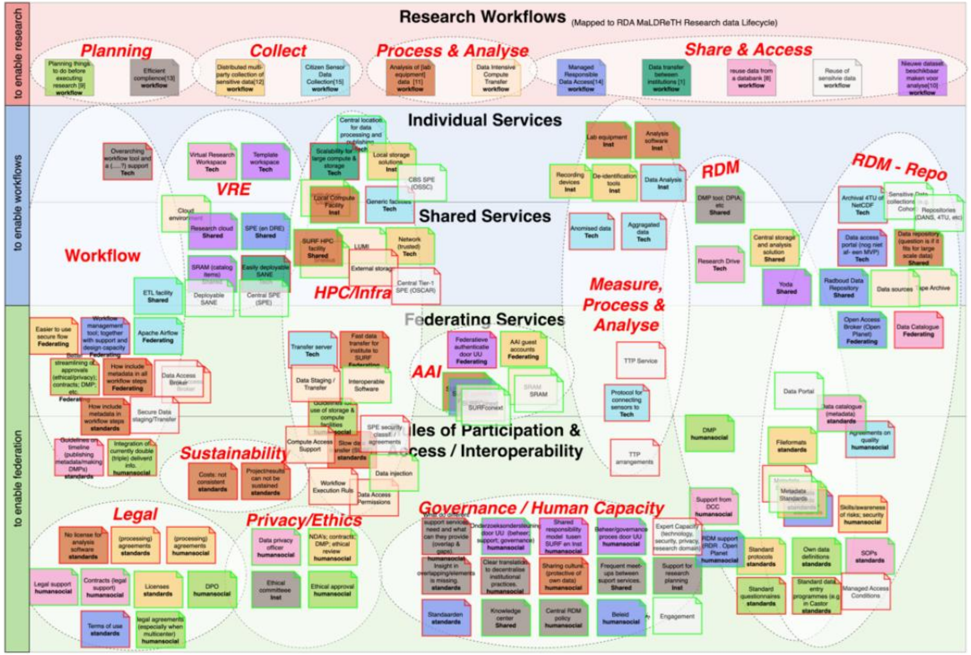
- support interaction between SURF and institutes & universities
- workshops on Federated HPC & Data
- strengthen organisational Research IT support

Sources: https://www.surf.nl/en/tech-trends, FreedomLab Innovation Expert group report-out, HPC Tier-2 contact group

Federating hitherto independent centres is non-trivial

Even with national funding available, getting a coherent national plan for Tier-2 facilities proved challenging

- Local build-up of computing expertise
- Technical alignment
  (ESSII/CVMFS, SRAM, OpenOnDemand, ssh, …)
- Scaling: how to reach researchers?
- Divergent business models between institutions *and* between institutions and national/EU resources
- ‘Competition’ from ‘free’ national (NWO) compute grants for ‘pilot/small requests’



Research Workflows (Mapped to RDA MaLDReTH Research data Lifecycle)
to enable research
to enable workflows
to enable federation
Planning
Collect
Process & Analyse
Share & Access
Planning things to do before executing research [9] workflow
Efficient complience[13] workflow
Distributed multi-party collection of sensitive data[12] workflow
Citizen Sensor Data Collection[15] workflow
Analysis of [lab equipment] data [11] workflow
Data Intensive Compute Transfer workflow
Managed Responsible Data Access[14] workflow
Data transfer between institutions [1] workflow
reuse data from a databank [8] workflow
Reuse of sensitvie data workflow
Nieuwe dataset beschikbaar maken voor analyse[10] workflow
Individual Services
Shared Services
Federating Services
Workflow
VRE
HPC/Infra
Measure, Process & Analyse
RDM
RDM - Repo
AAI
Sustainability
Rules of Participation & Access / Interoperability
Legal
Privacy/Ethics
Governance / Human Capacity
Overarching workflow tool and a (.....?) support Tech
Virtual Research Workspace Tech
Template workspace Tech
Central location for data processing and publishing Tech
Scalability for large compute & storage Tech
Local storage solutions Inst
CBS SPE (OSSC)
Local Compute Facility Inst
Generic facilities Tech
Cloud environment
Research cloud Shared
SPE (en DRE) Shared
SURF HPC facility Shared
LUMI
Network (trusted) Tech
SRAM (catalog items)
Easily deployable SANE
External storage
ETL facility Shared
Deployable SANE
Central SPE (SPE)
Central Tier-1 SPE (OSCAR)
Lab equipment Inst
Analysis software Inst
Recording devices Inst
De-identification tools Inst
Data Analysis Inst
Anomised data Tech
Aggregated data Tech
DMP tool; DPIA; etc Shared
Research Drive Tech
Central storage and analysis solution Shared
Yoda Shared
Archival 4TU of NetCDF Tech
Sensitive Data collections (e.g. Cohort)
Repositories (DANS, 4TU, etc)
Data access portal (nog niet af- een MVP) Tech
Data repository (question is if it fits for large scale data) Shared
Radboud Data Repository Shared
Data sources
Tape Archive
Open Access Broker (Open Planet) Federating
Data Catalogue Federating
Easier to use secure flow Federating
Workflow management tool; together with support and design capacity Federating
Apache Airflow Federating
Better streamlining of approvals (ethical/privacy); contracts; DMP; etc. Federating
How include metadata in all workflow steps Federating
Data Access Broker
How include metadata in workflow steps standards
Secure Data staging/Transfer
Guidelines on timeline (publishing metadata/making DMPs) standards
Integration of currently double (triple) deliverd info. humansocial
Transfer server Tech
Fast data transfer for institute to SURF Federating
Data Staging / Transfer
Interoperable Software
Guidelines for use of storage & compute facilities humansocial
Compute Access Support
Slow data transfer (SURF) standards
SPE security classif. / agreements
Data injection
Workflow Execution Rules
Data Access Permissions
Costs: not consistent standards
Project/results can not be sustained standards
Federatieve authenticatie door UU Federating
AAI guest accounts Federating
SURFconext
SRAM
TTP Service
Protocol for connecting sensors to Tech
TTP arrangements
DMP humansocial
Data Portal
Data catalogue (metadata) standards
Agreements on quality humansocial
Fileformats standards
Metadata Standards
Support from DCC humansocial
Skills/awareness of risks; security humansocial
No license for analysis software standards
(processing) agreements standards
(processing) agreements humansocial
Legal support humansocial
Contracts (legal support) humansocial
Licenses standards
DPO humansocial
Terms of use standards
legal agreements (especially when multicenter) humansocial
Data privacy officer humansocial
NDA's; contracts; DMP; ethical review humansocial
Ethical committee Inst
Ethical approval humansocial
What do different support services need and what can they provide (overlap & gaps) humansocial
Onderzoeksondersteuning door UU (beheer, support, governance) humansocial
Shared responsibility model tusen SURF en Inst humansocial
Beheer/governance proces door UU humansocial
Expert Capacity (technology, security, privacy, research domain)
RDM support (RDR . Open Planet) humansocial
Insight in overlapping/elements is missing standards
Clear translation to decentralise institutional practices. humansocial
Sharing culture (protective of own data) humansocial
Frequent meet-ups between suport services. Shared
Support for research planning Inst
Standaarden standards
Knowledge center Shared
Central RDM policy humansocial
Beleid humansocial
Engagement
Standard protocols standards
Own data definitions standards
SOPs standards
Standard questionnaires standards
Standard data entry programmes (e.g in Castor) standards
Managed Access Conditions

Figure 6: Output of workshops, for illustration purposes

SURF and NWO DCC-T2 proposal 1.0

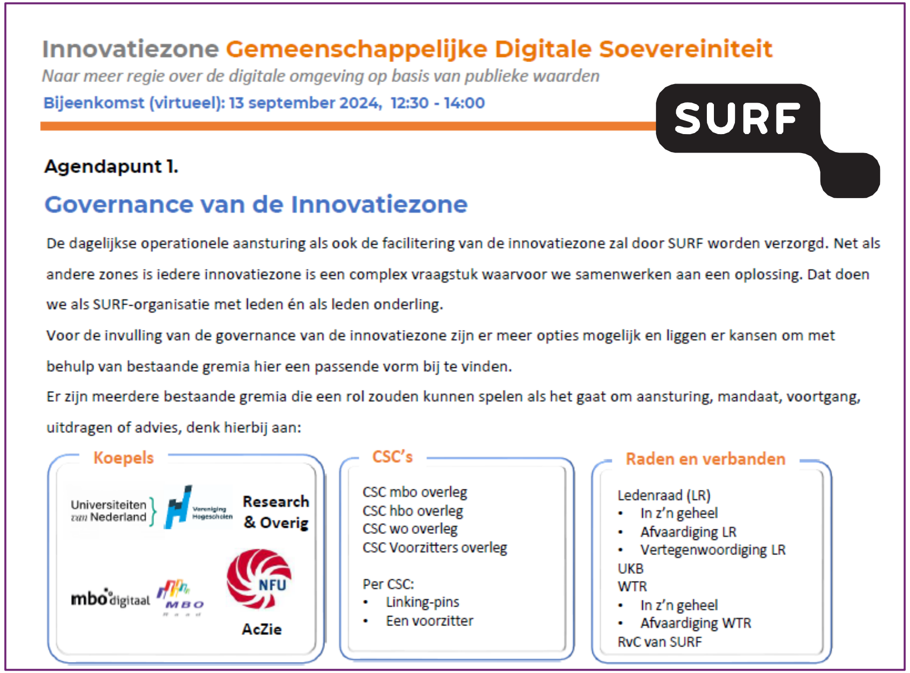
Limitations to the membership model ?

Coherence between funding, research mission, and *representation* is challenging

- *Representation in the co-operative* is
  from mostly enterprise-IT focussed organisational units
  – that then drive key decision processes on the whole co-operative
- *Influence* on services by researchers, research funding, and funders
  is hard to anchor in the governance (no representation = limited influence)
- *Divergent directions* even within the co-operative
  example: innovation action on digital sovereignty has no effect on
  enterprise-member-enforced discontinuation of existing autonomous services
  – regardless of strategy and impact ☹

Collaborating on ICT infrastructure collaboration
Innovatiezone: *Gemeenschappelijke Digitale Soevereiniteit*

- Bewustwording en kennis over alternatieven, creëren van draagvlak en bereidheid om dit serieus te ontwikkelen;
- Sturing op aanpassing van de Sourcing strategieën (van alle leden), acceptatie van Open Source;
- Investering in AI op basis van alternatieve ontwikkelingen;
- Mandaat om hier binnen SURF al pilots op te ondernemen; en
- Breed communiceren over beschikbare alternatieven en geleerde lessen.



Innovatiezone Gemeenschappelijke Digitale Soevereiniteit

Naar meer regie over de digitale omgeving op basis van publieke waarden

Bijeenkomst (virtueel): 13 september 2024, 12:30 - 14:00

SURF

Agendapunt 1.

Governance van de Innovatiezone

De dagelijkse operationele aansturing als ook de facilitering van de innovatiezone zal door SURF worden verzorgd. Net als andere zones is iedere innovatiezone is een complex vraagstuk waarvoor we samenwerken aan een oplossing. Dat doen we als SURF-organisatie met leden én als leden onderling.

Voor de invulling van de governance van de innovatiezone zijn er meer opties mogelijk en liggen er kansen om met behulp van bestaande gremia hier een passende vorm bij te vinden.

Er zijn meerdere bestaande gremia die een rol zouden kunnen spelen als het gaat om aansturing, mandaat, voortgang, uitdragen of advies, denk hierbij aan:

Koepels

Universiteiten van Nederland
Vereniging Hogescholen
Research & Overig
mbo digitaal
MBO Raad
NFU
AcZie

CSC's

CSC mbo overleg
CSC hbo overleg
CSC wo overleg
CSC Voorzitters overleg

Per CSC:
- Linking-pins
- Een voorzitter

Raden en verbanden

Ledenraad (LR)
- In z'n geheel
- Afvaardiging LR
- Vertegenwoordiging LR

UKB
WTR
- In z'n geheel
- Afvaardiging WTR

RvC van SURF

Even in NL - traditionally a big-tech-lax country …

‘Digital autonomy is the starting point for government.
We choose a European digital infrastructure […]
so that more Dutch and European SMEs can participate.

[…] based on security-by-design, zero-trust, sovereignty, open source, and supply-chain security.
Government authorities will use their purchasing power to enforce safe standard and will put in place minimum requirements for security for central government.’

‘Digitale autonomie moet het uitgangspunt zijn voor de overheid. We kiezen voor een Europese digitale infrastructuur, bouwen strategische afhankelijkheden in cloud, data en cruciale systemen doelgericht af, en we splitsen grote projecten op zodat meer Nederlandse en Europese mkb’ers kunnen meedoen. Digitale inkoop en aanbestedingen worden gestandaardiseerd en gecentraliseerd, gestuurd op security-by-design, zero-trust, soevereiniteit, open source en ketenveiligheid. De overheid benut haar marktmacht om veilige standaarden af te dwingen en stelt rijksbrede minimumeisen op voor security.’

Aan de slag
Bouwen aan een beter Nederland

Coalitieakkoord 2026-2030
D66, VVD en CDA

Maastricht University
UNU MERIT

Working Paper Series

nd Digital Sovereignty for the
tribution to a Coalition Agreement
and Nicholas Rounding

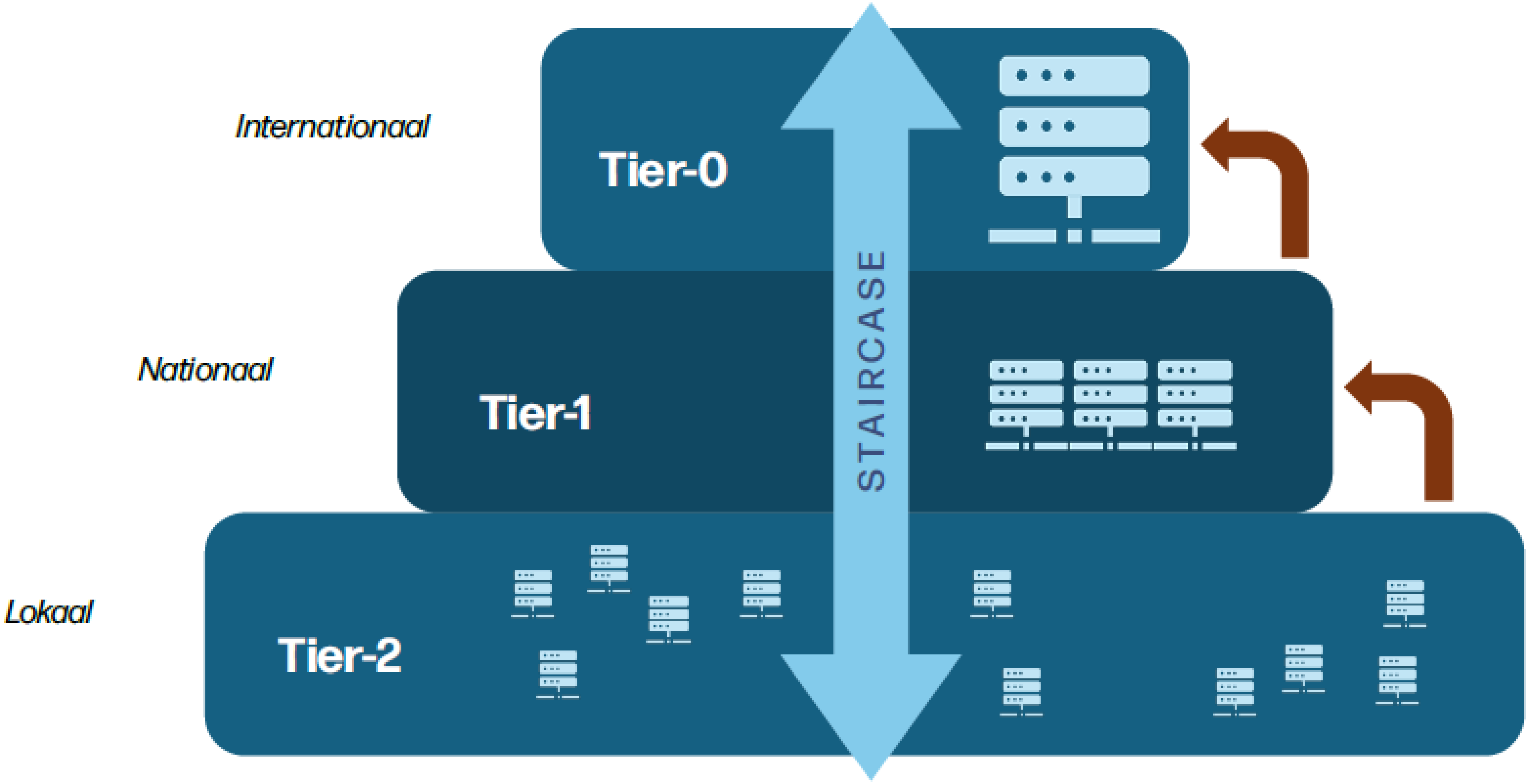
n and Technology (UNU-MERIT)

DE ROUTE NAAR TOEKOMSTIGE WELVAART

Een sterk Nederland
in een relevant Europa

RAPPORT WENNINK

https://www.kabinetsformatie2025.nl/documenten/2026/01/30/aan-de-slag---coalitieakkoord-2026-2030;
Akhan, Groep, Ritzen and Rounding: https://doi.org/10.53330/ZMFF2486, Rapport Wennink https://www.rapportwennink.nl/



BALANCED TIER MODEL

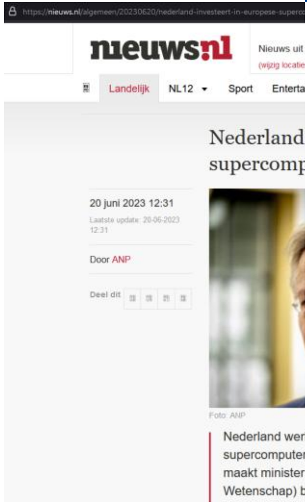
EuroHPC targets large-scale *compute* (and some data)

Dutch direct investments: 2M€ LUMI, 8M€ JV
+ access through 'Europe' and the JU

But: it's not the 'one single solution' …

e.g. EuroHPC has overly many controls, it being subject to more export controls

- harder to use for research (like for DestinE portals) that need to run services or use service accounts
- tension with open and citizen science



https://nieuws.nl/algemeen/20230620/nederland-investeert-in-europese-superco

nieuws.nl

Nieuws uit
(wijzig locatie

Nederland
supercomp

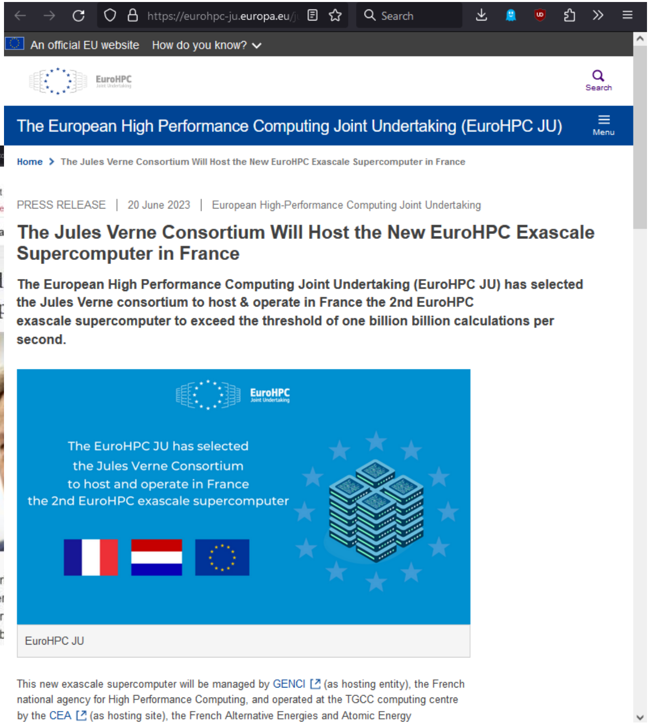
20 juni 2023 12:31
Laatste update: 20-06-2023 12:31

Door ANP

Deel dit

Foto: ANP

Nederland wer
supercompute
maakt minister
Wetenschap) b



https://eurohpc-ju.europa.eu/

An official EU website How do you know?

EuroHPC

The European High Performance Computing Joint Undertaking (EuroHPC JU)

Home > The Jules Verne Consortium Will Host the New EuroHPC Exascale Supercomputer in France

PRESS RELEASE | 20 June 2023 | European High-Performance Computing Joint Undertaking

The Jules Verne Consortium Will Host the New EuroHPC Exascale Supercomputer in France

The European High Performance Computing Joint Undertaking (EuroHPC JU) has selected the Jules Verne consortium to host & operate in France the 2nd EuroHPC exascale supercomputer to exceed the threshold of one billion billion calculations per second.

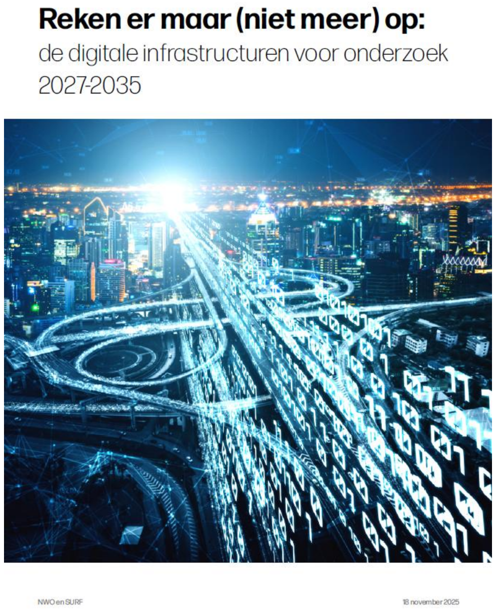
EuroHPC
The EuroHPC JU has selected
the Jules Verne Consortium
to host and operate in France
the 2nd EuroHPC exascale supercomputer

EuroHPC JU

This new exascale supercomputer will be managed by GENCI (as hosting entity), the French national agency for High Performance Computing, and operated at the TGCC computing centre by the CEA (as hosting site), the French Alternative Energies and Atomic Energy

Images: https://nieuws.nl/algemeen/20230620/nederland-investeert-in-europese-supercomputer/, https://eurohpc-ju.europa.eu/jules-verne-consortium-will-host-new-eurohpc-exascale-supercomputer-france-2023-06-20_en. EuroHPC comments, see also Thomas Geenen, ECMWF & DestinE (at EGI2023)

Dealing with the digitisation ‘explosion’



Reken er maar (niet meer) op:

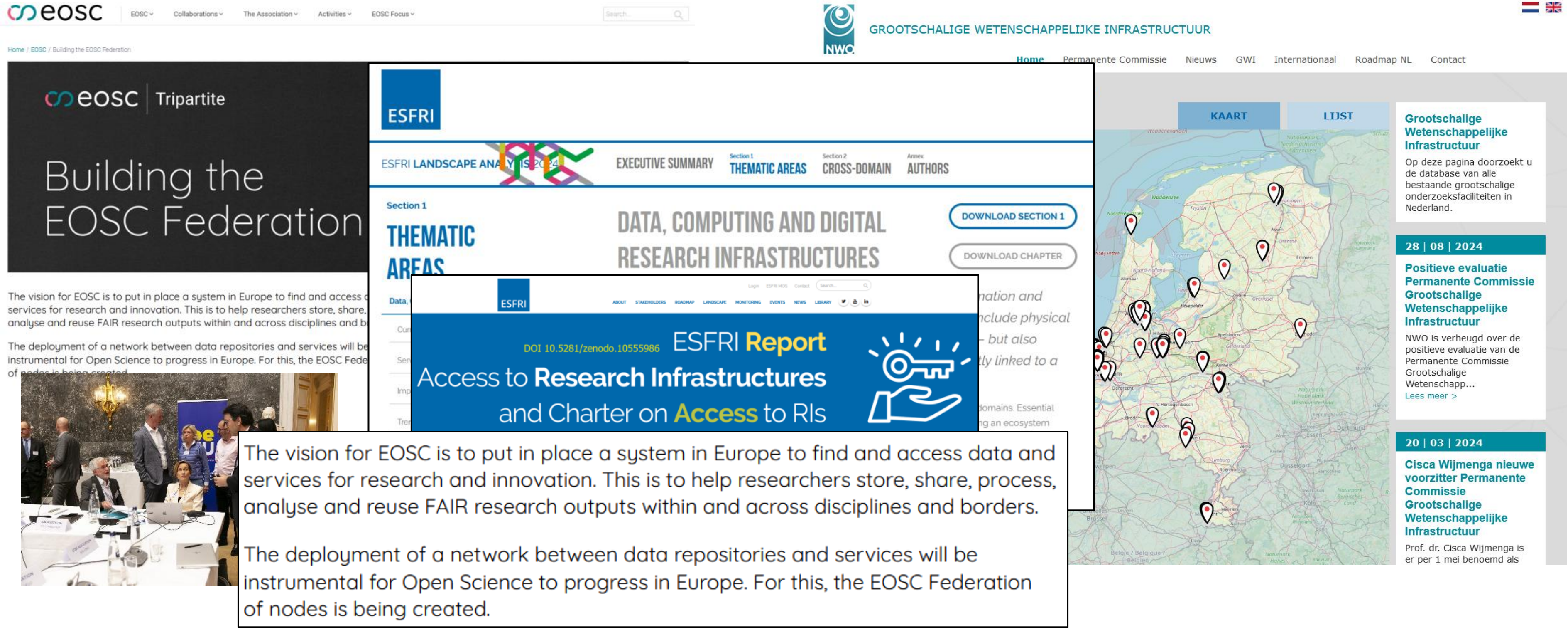
de digitale infrastructuren voor onderzoek 2027-2035

NWO en SURF

18 november 2025

https://www.rijksoverheid.nl/documenten/rapporten/2025/11/18/reken-er-maar-niet-meer-op-de-digitale-infrastructuren-voor-onderzoek-2027-2035

On ESFRIs, GWIs, the European Open Science Cloud



Home / EOSC / Building the EOSC Federation

Building the EOSC Federation

Section 1

THEMATIC AREAS

DATA, COMPUTING AND DIGITAL RESEARCH INFRASTRUCTURES

DOI 10.5281/zenodo.10555986

ESFRI Report

Access to Research Infrastructures and Charter on Access to RIs

Grootschalige Wetenschappelijke Infrastructuur

Op deze pagina doorzoekt u de database van alle bestaande grootschalige onderzoeksfaciliteiten in Nederland.

28 | 08 | 2024

Positieve evaluatie Permanente Commissie Grootschalige Wetenschappelijke Infrastructuur

NWO is verheugd over de positieve evaluatie van de Permanente Commissie Grootschalige Wetenschapp...

Lees meer >

20 | 03 | 2024

Cisca Wijmenga nieuwe voorzitter Permanente Commissie Grootschalige Wetenschappelijke Infrastructuur

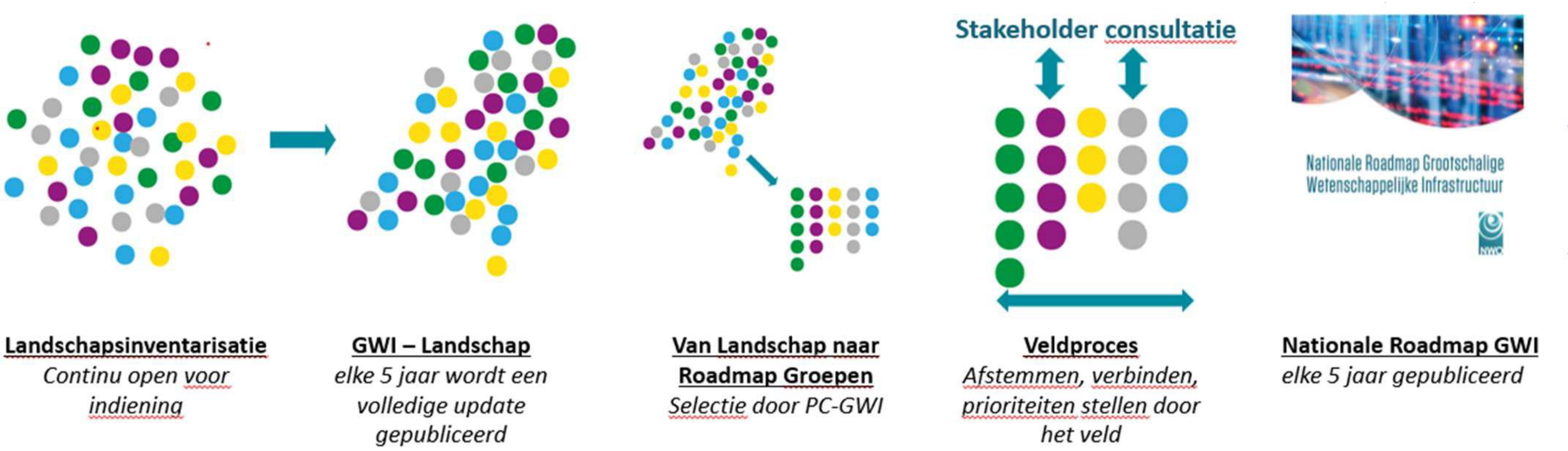
Prof. dr. Cisca Wijmenga is er per 1 mei benoemd als

The vision for EOSC is to put in place a system in Europe to find and access data and services for research and innovation. This is to help researchers store, share, process, analyse and reuse FAIR research outputs within and across disciplines and borders.

The deployment of a network between data repositories and services will be instrumental for Open Science to progress in Europe. For this, the EOSC Federation of nodes is being created.

See e.g. https://www.onderzoeksfaciliteiten.nl/; https://www.esfri.eu/; https://landscape2024.esfri.eu/; https://eosc.eu/building-the-eosc-federation/

The Dutch Research Infrastructure Landscape Process


Stakeholder consultatie
Nationale Roadmap Grootschalige Wetenschappelijke Infrastructuur
NWO

Landschapsinventarisatie
Continu open voor indiening

GWI – Landschap
elke 5 jaar wordt een volledige update gepubliceerd

Van Landschap naar Roadmap Groepen
Selectie door PC-GWI

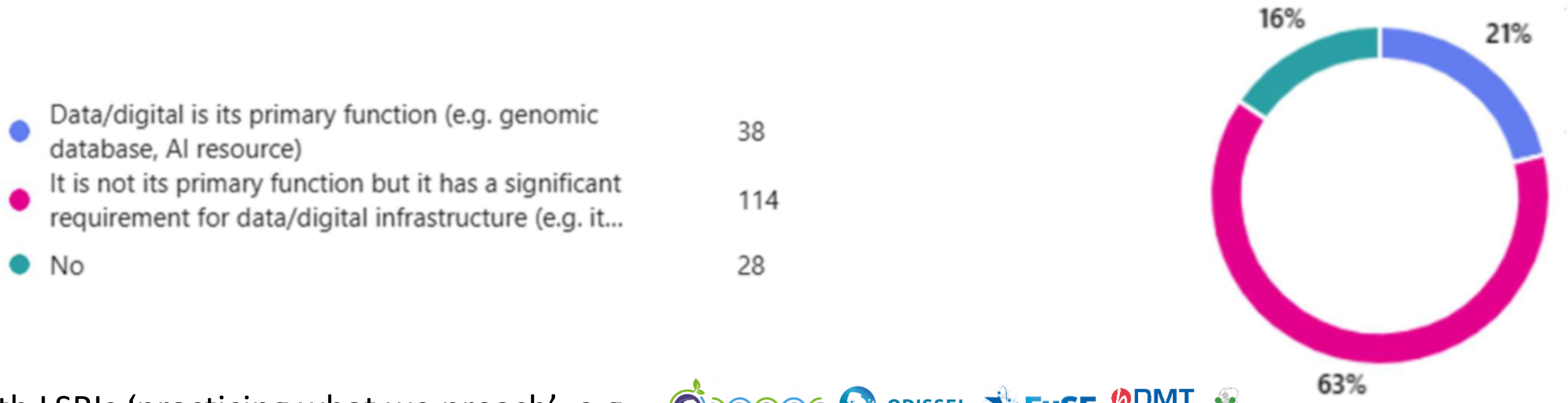
Veldproces
Afstemmen, verbinden, prioriteiten stellen door het veld

Nationale Roadmap GWI
elke 5 jaar gepubliceerd

https://www.nwo.nl/en/large-scale-research-infrastructure

Digitisation is everywhere

80%+ of all LSRIs in the Netherlands have significant data or digital requirements



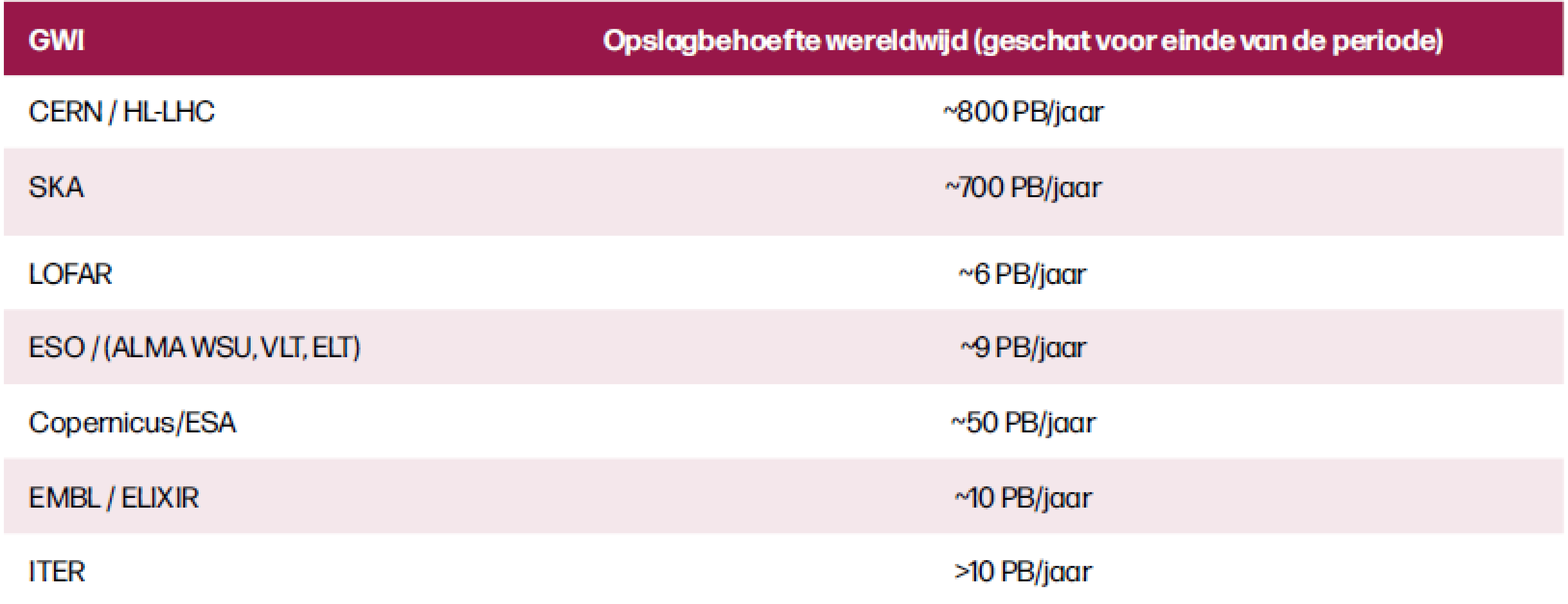
With LSRIs 'practicing what we preach', e.g.:

ARISE (biodiversity), ODISSEI (social sciences), FuSE (HEP physics & astronomy), hDMT (human Disease Model Technologies), NL-BioImaging (distributed multi-model microscopy)

(… and apologies to those forgotten)

Data from NWO GWI Landscape Analysis, 2025 (Katrien Uytterhoeven *et al*., NWO); logos from the respective GWIs

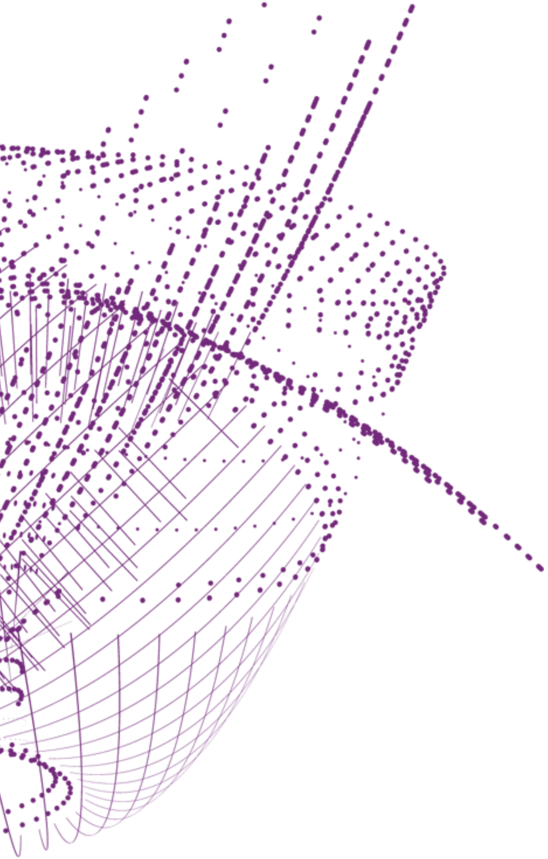
And more LSRIs relying on *much more data* than before



GWI	Opslagbehoefte wereldwijd (geschat voor einde van de periode)
CERN / HL-LHC	~800 PB/jaar
SKA	~700 PB/jaar
LOFAR	~6 PB/jaar
ESO / (ALMA WSU, VLT, ELT)	~9 PB/jaar
Copernicus/ESA	~50 PB/jaar
EMBL / ELIXIR	~10 PB/jaar
ITER	>10 PB/jaar

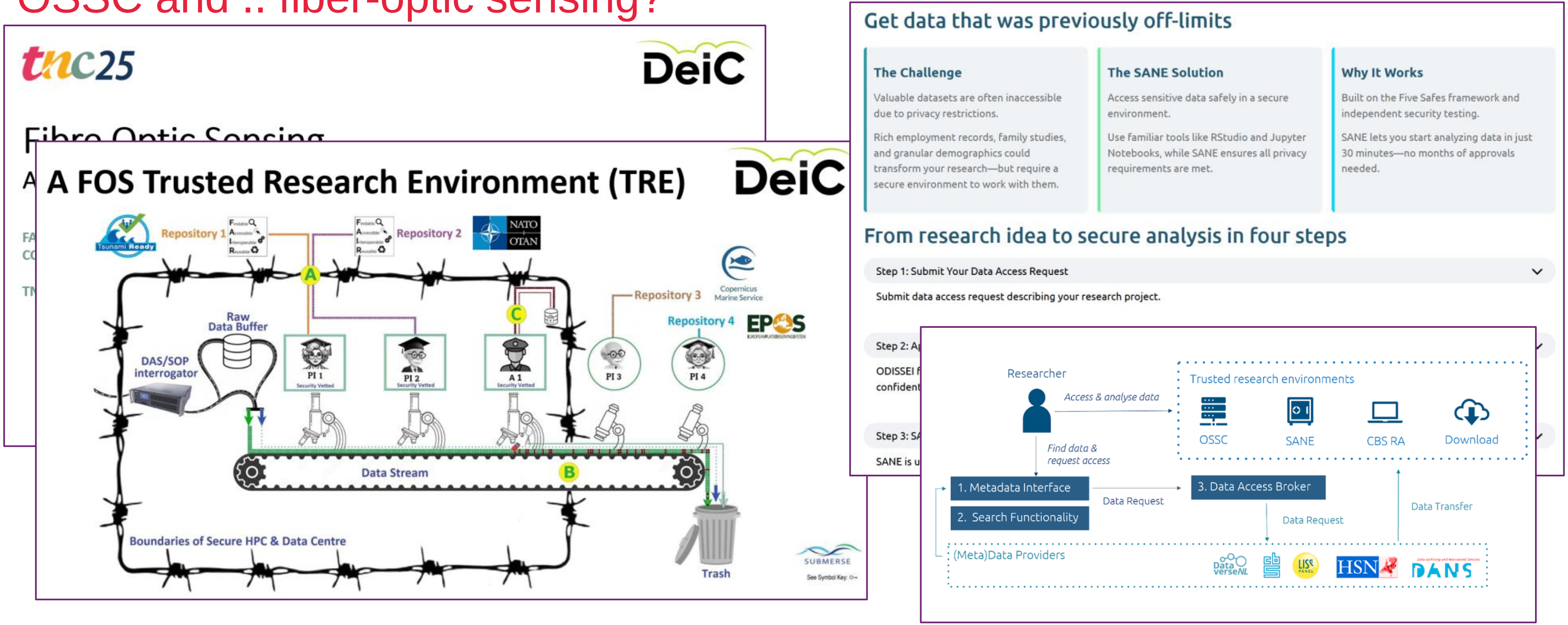
Tabel 5.1.2: Geschatte opslagbehoefte per jaar voor een aantal GWI's waaraan Nederland deelneemt.

From the report 'Reken er maar (niet meer) op', November 2025, https://open.overheid.nl/documenten/be82689f-08e7-40de-8898-181c119e080a/file



Research Infrastructure Commons

Example of common challenges and foundations: OSSC and .. fiber-optic sensing?



tnc25 DeiC

Fibre Optic Sensing

A

FA
CC

TN

A FOS Trusted Research Environment (TRE)

DeiC

Get data that was previously off-limits

The Challenge

Valuable datasets are often inaccessible due to privacy restrictions.

Rich employment records, family studies, and granular demographics could transform your research—but require a secure environment to work with them.

The SANE Solution

Access sensitive data safely in a secure environment.

Use familiar tools like RStudio and Jupyter Notebooks, while SANE ensures all privacy requirements are met.

Why It Works

Built on the Five Safes framework and independent security testing.

SANE lets you start analyzing data in just 30 minutes—no months of approvals needed.

From research idea to secure analysis in four steps

Step 1: Submit Your Data Access Request

Submit data access request describing your research project.

Step 2: Ap

ODISSEI f
confident

Step 3: SA

SANE is u

FOS slides from DeIC, **Rene Belsø**, at TNC25 (https://indico.geant.org/event/5/contributions/420/), and
ODISSEI Secure Supercomputing (OSSC) by SURF & ODISSEI GWI consortium (https://odissei-data.nl/facility/secure-analysis-environment-sane/), https://odissei-data.nl/facility/odissei-portal/

Identifying elements of the Commons in LSRIs

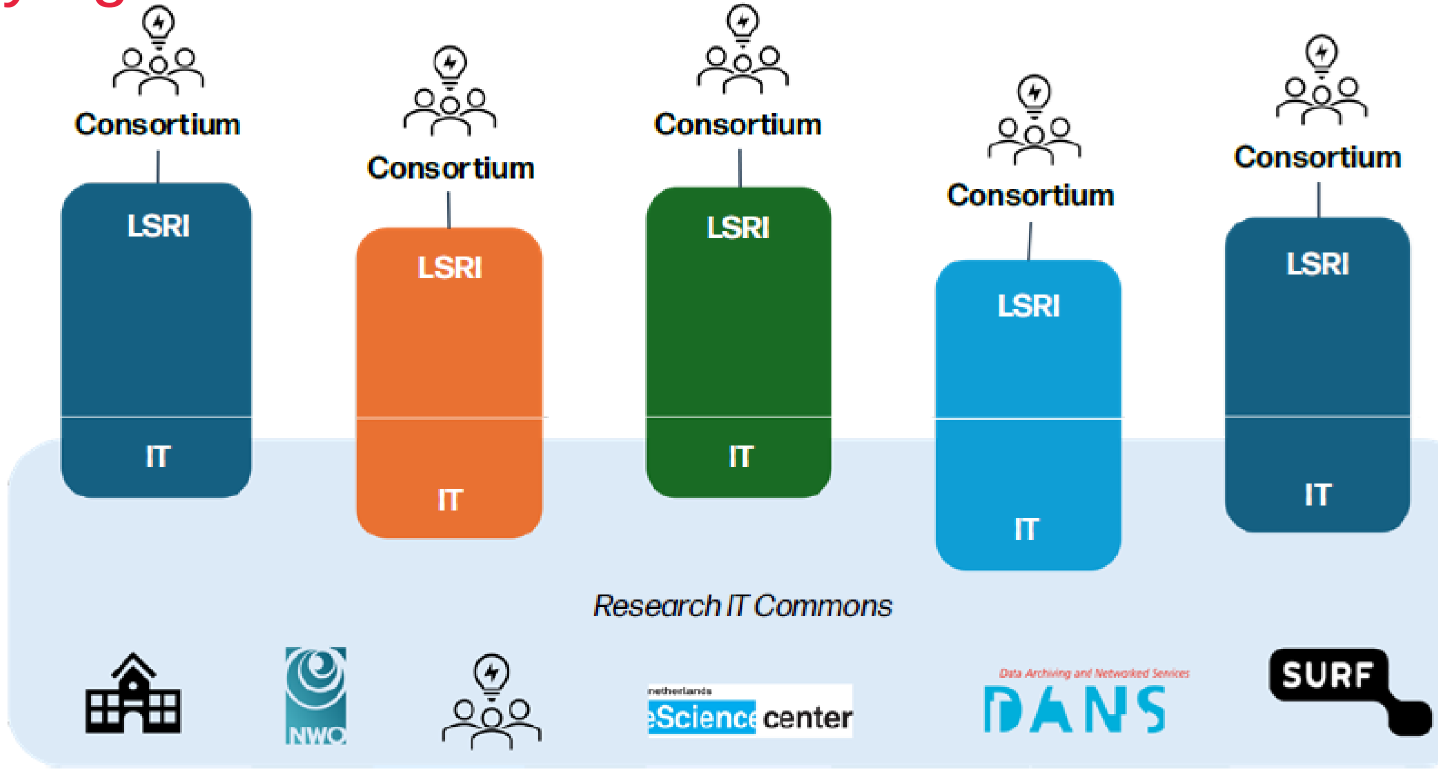
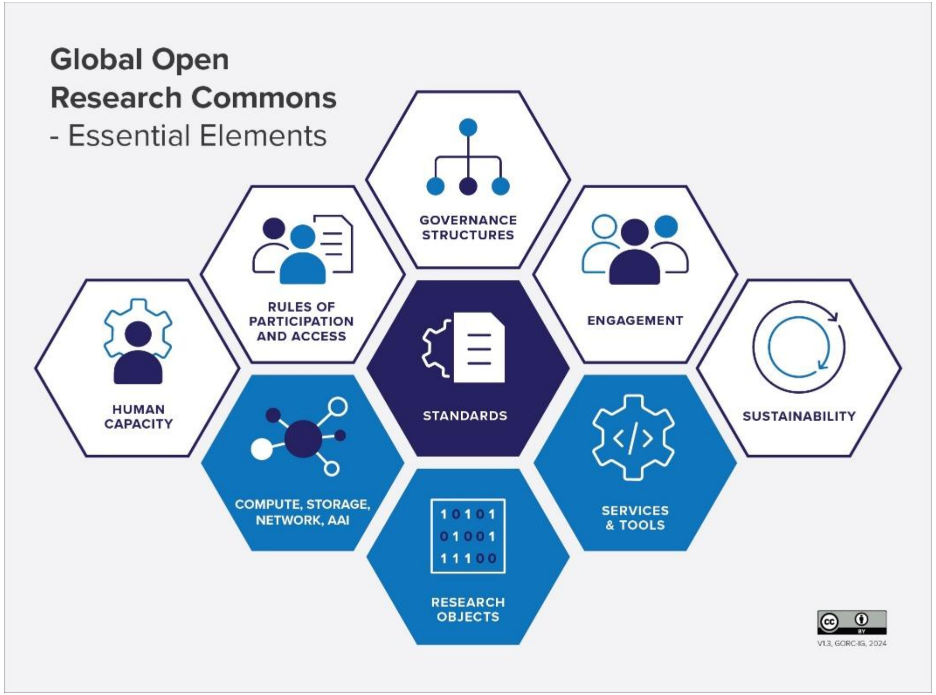


Image source: Machgiel Bijsterbos, SURF

‘Managing the commons’



GORC IG: Typology and Definitions, https://doi.org/10.15497/RDA00087 - for the original book ‘Governing the Commons’ see of course Elanor Ostrom ☺

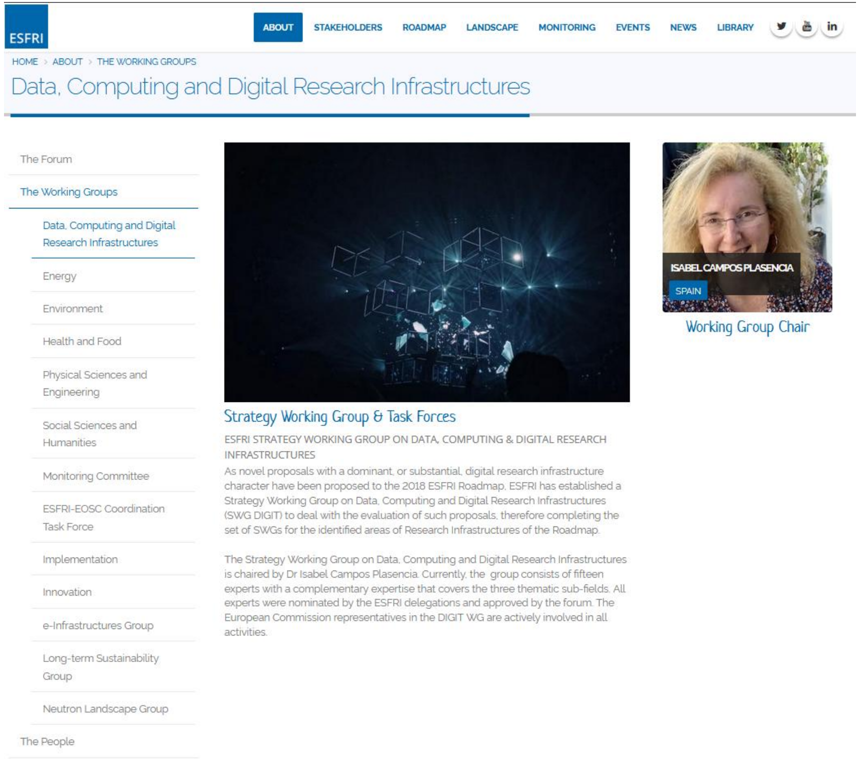
Horizontal platforms – GWI-DIGIT … and ESFRI-DIGIT

Ontdek de GWI Faciliteiten:

Overzicht 0 Op de kaart 0

NWO
Grootschalige
Wetenschappelijke
Infrastructuur

Digit-platform



HOME › ABOUT › THE WORKING GROUPS

Data, Computing and Digital Research Infrastructures

- The Forum
- The Working Groups
  - Data, Computing and Digital Research Infrastructures
  - Energy
  - Environment
  - Health and Food
  - Physical Sciences and Engineering
  - Social Sciences and Humanities
  - Monitoring Committee
  - ESFRI-EOSC Coordination Task Force
  - Implementation
  - Innovation
  - e-Infrastructures Group
  - Long-term Sustainability Group
  - Neutron Landscape Group
- The People

ISABEL CAMPOS PLASENCIA
SPAIN

Working Group Chair

Strategy Working Group & Task Forces

ESFRI STRATEGY WORKING GROUP ON DATA, COMPUTING & DIGITAL RESEARCH INFRASTRUCTURES

As novel proposals with a dominant, or substantial, digital research infrastructure character have been proposed to the 2018 ESFRI Roadmap, ESFRI has established a Strategy Working Group on Data, Computing and Digital Research Infrastructures (SWG DIGIT) to deal with the evaluation of such proposals, therefore completing the set of SWGs for the identified areas of Research Infrastructures of the Roadmap.

The Strategy Working Group on Data, Computing and Digital Research Infrastructures is chaired by Dr Isabel Campos Plasencia. Currently, the group consists of fifteen experts with a complementary expertise that covers the three thematic sub-fields. All experts were nominated by the ESFRI delegations and approved by the forum. The European Commission representatives in the DIGIT WG are actively involved in all activities.

https://www.esfri.eu/working-groups/data-computing-and-digital-research-infrastructures

How to make ICT infrastructure into our ‘research instrument’ ?

All these use cases seem diverse, but still result in **common infrastructure capabilities**

- Interactive analysis, collaboration and ‘research service bursting’ platform
  - DSRI is there now to fill this space –can evolve to the ‘interactive gateway’ for all users
- HTC/HPC computing facilities at reasonable ‘T2’ scale, based on application co-design
  - solves short-turnaround needs at limited scale, is the place for growing expertise for scale out to national (SURF) and international (EuroHPC, EGI, EOSC, …) level
- High-throughput data storage and sharing services
  - targeting data processing compute integration and effective fast access to FAIR data
- Open network for collaborative & data intensive sciences
  - ‘ye shall not have stateful devices in thy data path’ – ScienceDMZ or better
  - is *essential* prerequisite for open science, EOSC, and collaborative (& citizen science) services
- Tools for digital research collaboration
  - sustainable research software, collaborative spaces with *global* partners, SRAM, eduGAIN & EOSC federated access, ubiquitous access to *external R&S* services

Infrastructure is more than just the tools or technology

The ‘*Uitvoeringsplan*’ (‘*commissie Apers’, 2019*) deliberately identified digital competences to be broad and include not only data, but also software **and a federated expertise network** at the ‘local’ digital competence centres (LDCCs):

- “Knooppunt in een gefedereerd netwerk voor data, computing en expertise”
- “Belangrijk is dat de aangesloten lokale infrastructuren middels het gefedereerde systeem geïntegreerd moeten kunnen worden in de European Open Science Cloud (EOSC), die in ontwikkeling is.”

This means we require expertise and alignment, also for governance and policy, with the goals for federated Open Science which our nationally initiatives are funding

https://zoek.officielebekendmakingen.nl/kst-29338-189

Redefining the concept of service management

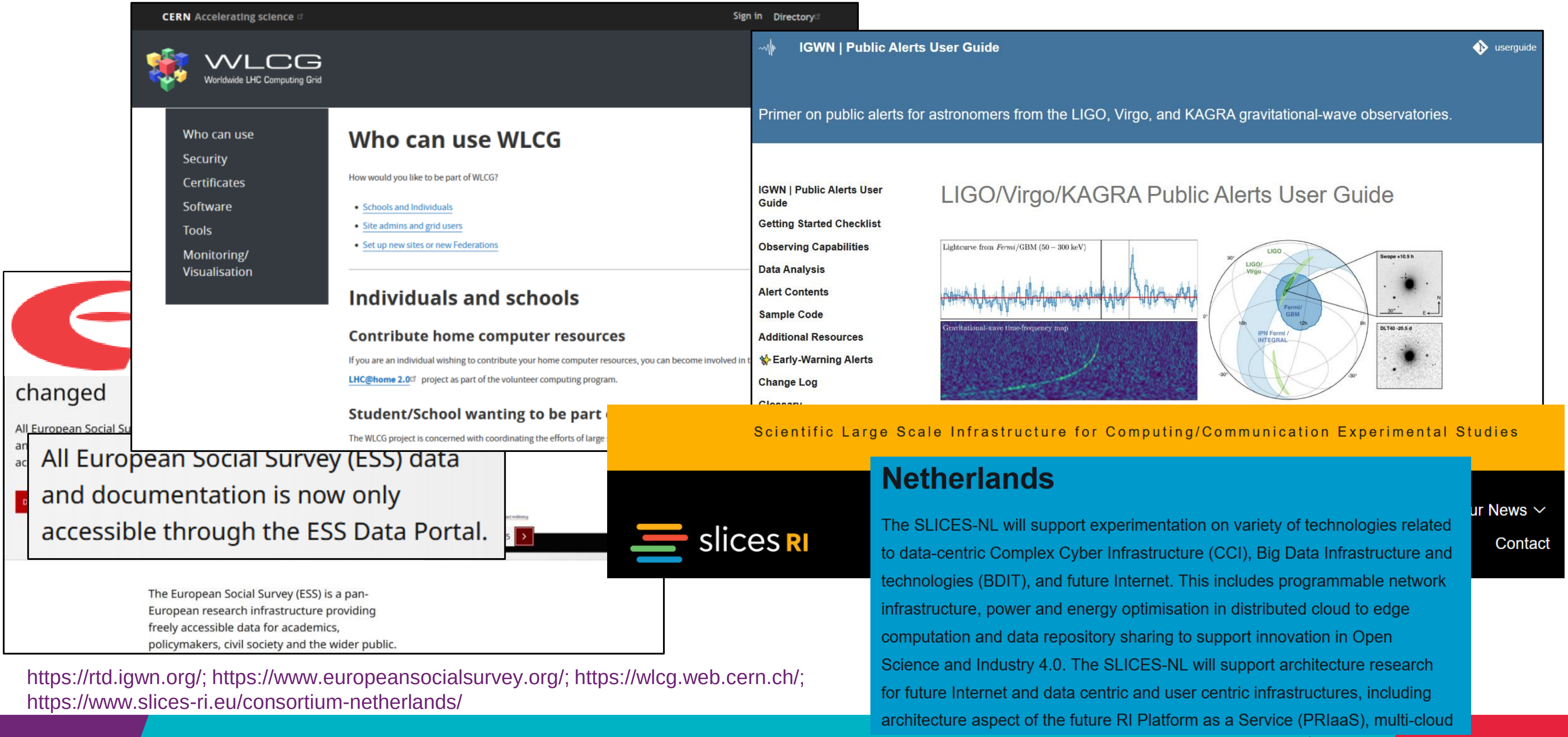
IT Service Management (ISO20000, ITIL, FitSM) promote a *specific definition* of service portfolio management:

> “**Product/service portfolio** The product/service portfolio is the complete set of products and/or services *that are managed by the organization*, and it represents the organization’s commitments and investments across all its customers and market spaces. …”
> [ITIL v4, chapter 5]

Concept of ‘services’ in collaborative research is entirely different:

- coming from existing collaborations and infrastructures, many or most services *already exists and used extensively* by research and collaborative administration;
- they are an *essential part of collaborative research*: they should be embraced;
- whether a ‘service’ is operated by a third party, our outside (local) ITSM control is *immaterial to the value of the service*

Research Infrastructures both users and providers … and ‘we’ are as well!



Who can use WLCG

How would you like to be part of WLCG?

- Schools and Individuals
- Site admins and grid users
- Set up new sites or new Federations

Individuals and schools

Contribute home computer resources

If you are an individual wishing to contribute your home computer resources, you can become involved in t

LHC@home 2.0 project as part of the volunteer computing program.

Student/School wanting to be part

The WLCG project is concerned with coordinating the efforts of large

Primer on public alerts for astronomers from the LIGO, Virgo, and KAGRA gravitational-wave observatories.

LIGO/Virgo/KAGRA Public Alerts User Guide

changed

All European Social Su

an

ac

All European Social Survey (ESS) data and documentation is now only accessible through the ESS Data Portal.

The European Social Survey (ESS) is a pan-European research infrastructure providing freely accessible data for academics, policymakers, civil society and the wider public.

Scientific Large Scale Infrastructure for Computing/Communication Experimental Studies

Netherlands

The SLICES-NL will support experimentation on variety of technologies related to data-centric Complex Cyber Infrastructure (CCI), Big Data Infrastructure and technologies (BDIT), and future Internet. This includes programmable network infrastructure, power and energy optimisation in distributed cloud to edge computation and data repository sharing to support innovation in Open Science and Industry 4.0. The SLICES-NL will support architecture research for future Internet and data centric and user centric infrastructures, including architecture aspect of the future RI Platform as a Service (PRIaaS), multi-cloud

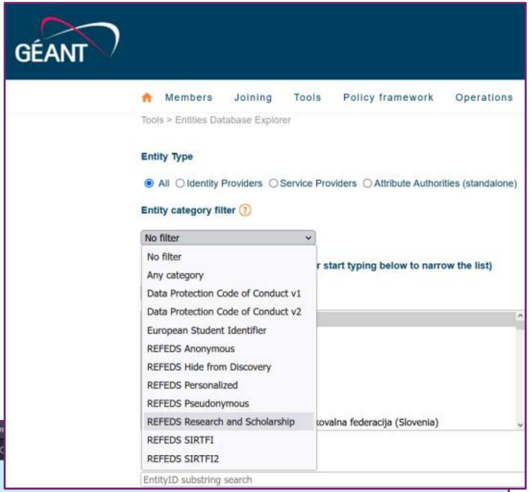
https://rtd.igwn.org/; https://www.europeansocialsurvey.org/; https://wlcg.web.cern.ch/; https://www.slices-ri.eu/consortium-netherlands/

And it needs everyone to work together

To scale trust in research infrastructures,
we need to keep challenging ourselves …

- *for eduGAIN: do we choose more trustworthiness and target baseline assurance, or more inclusiveness, but maybe less trust?*
- *for your university IT department: prioritize the primary mission of education and research, as both are now globally connected*
  - ‘we can *use* existing services from outside’
  - ‘we can *contribute* in collaborations in education and research’
  - ‘we teach our students to understand, study, and work with interconnected services and systems that are globally connected’

  … rather than get stuck in an enterprise egg-shell approach?
- *do our networks support a perimeter ‘fit for collaboration’?*



Tools > Entities Database Explorer

Entity Type

● All ○ Identity Providers ○ Service Providers ○ Attribute Authorities (standalone)

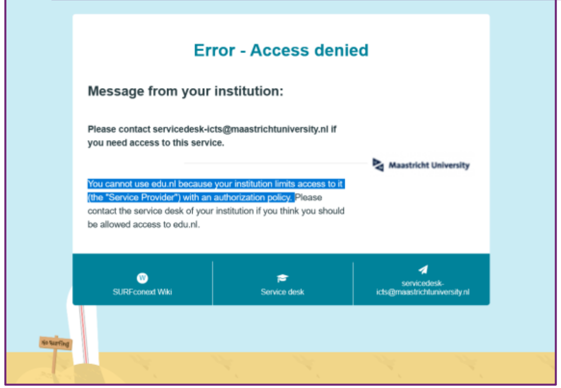
Entity category filter

No filter
- No filter
- Any category
- Data Protection Code of Conduct v1
- Data Protection Code of Conduct v2
- European Student Identifier
- REFEDS Anonymous
- REFEDS Hide from Discovery
- REFEDS Personalized
- REFEDS Pseudonymous
- REFEDS Research and Scholarship
- REFEDS SIRTFI
- REFEDS SIRTFI2

r start typing below to narrow the list)

ovalna federacija (Slovenia)

EntityID substring search



Error - Access denied

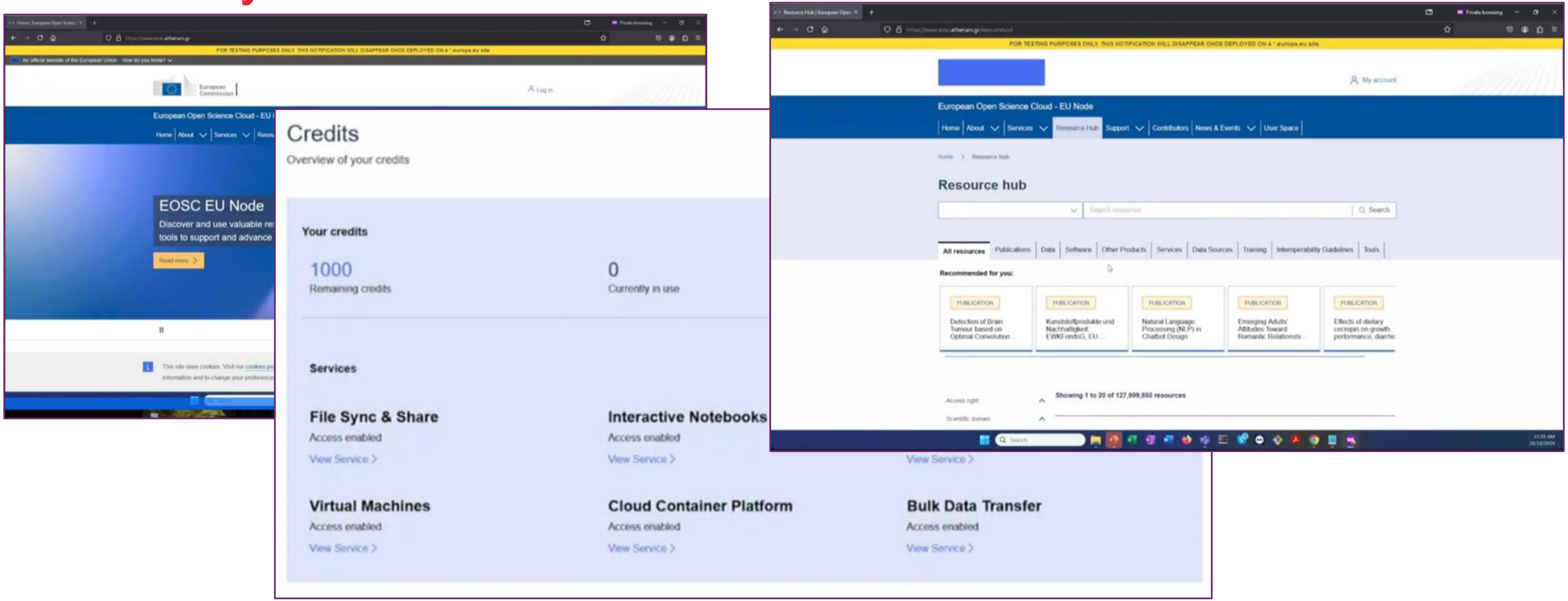
Message from your institution:

Please contact servicedesk-icts@maastrichtuniversity.nl if you need access to this service.

You cannot use edu.nl because your institution limits access to it (the "Service Provider") with an authorization policy. Please contact the service desk of your institution if you think you should be allowed access to edu.nl.

Images: https://technical.edugain.org/entities, Maastricht University blocking access to … a privacy-friendly URI shortener ☹,

And service use is coming from the EOSC ‘whether you want it or not’ …



FOR TESTING PURPOSES ONLY. THIS NOTIFICATION WILL DISAPPEAR ONCE DEPLOYED ON A *.europa.eu site

An official website of the European Union How do you know?

European Commission

Log in

European Open Science Cloud - EU

Home | About | Services | Resou

EOSC EU Node

Discover and use valuable re
tools to support and advance

Read more >

This site uses cookies. Visit our cookies po
information and to change your preference

Credits

Overview of your credits

Your credits

1000
Remaining credits

0
Currently in use

Services

File Sync & Share
Access enabled
View Service >

Interactive Notebooks
Access enabled
View Service >

View Service >

Virtual Machines
Access enabled
View Service >

Cloud Container Platform
Access enabled
View Service >

Bulk Data Transfer
Access enabled
View Service >

FOR TESTING PURPOSES ONLY. THIS NOTIFICATION WILL DISAPPEAR ONCE DEPLOYED ON A *.europa.eu site

My account

European Open Science Cloud - EU Node

Home | About | Services | Resource Hub | Support | Contributors | News & Events | User Space

Home > Resource hub

Resource hub

Search resources | Search

All resources | Publications | Data | Software | Other Products | Services | Data Sources | Training | Interoperability Guidelines | Tools

Recommended for you:

PUBLICATION
Detection of Brain Tumour based on Optimal Convolution...

PUBLICATION
Kunststoffprodukte und Nachhaltigkeit: EWKFondsG, EU-...

PUBLICATION
Natural Language Processing (NLP) in Chatbot Design

PUBLICATION
Emerging Adults' Attitudes Toward Romantic Relationshi...

PUBLICATION
Effects of dietary cecropin on growth performance, diarrhe

Access right

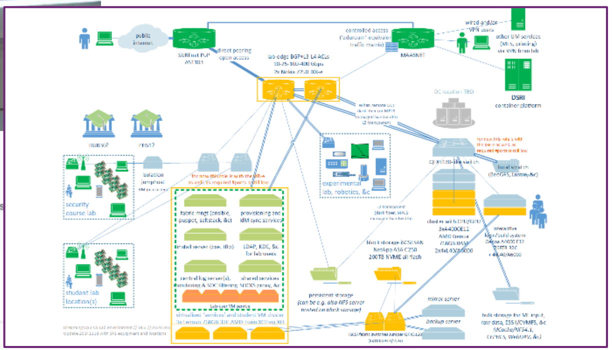
Scientific domain

Showing 1 to 20 of 127,999,850 resources

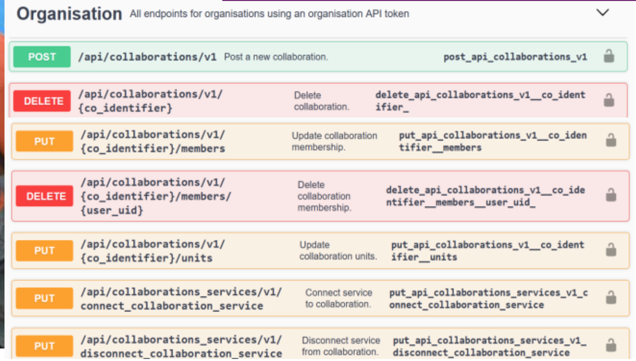
https://webcast.ec.europa.eu/eu-node-technical-launch-event-24-10-10

And education labs are much like ad-hoc research collaboration

just slightly more organised than research … ?



public internet
SURFnet PoP
direct peering open access
controlled access ('eduroam'-equivalent traffic matrix)
MAASNET
wired and/or VPN users
other UM services (MFS, printing) via VPN from lab
lab edge BGP+L3 L4 ACLs
10-25-100-400 Gbps
DC location TBD
DSRI
container platform
security course lab
isolation jumphost
student lab location(s)
fabric mngt (ansible, puppet, saltstack, &c)
provisioning and IdM sync services
install server (pxe, tftp)
LDAP, KDC, &c for lab users
central log server(s), monitoring & SDC filtering
shared services SOCKS proxy, &c
lab-user VM service
virtualised 'services' and student VM cluster
experimental lab, robotics, &c
local scratch (BeeGFS, Lustre, &c)
cluster with CPU/GPU
interactive login/build system
block storage iSCSI SAN NetApp ASA C250 200TB NVME all flash
persistent storage (can be e.g. also NFS server hosted on block storage)
mirror server
backup server
bulk storage for ML input, raw data, ESSI/CVMFS, &c (dCache/NFS4.1, CIFS, WebDAV, &c)



SRAM API 1.0

/swagger_api/apispec.json

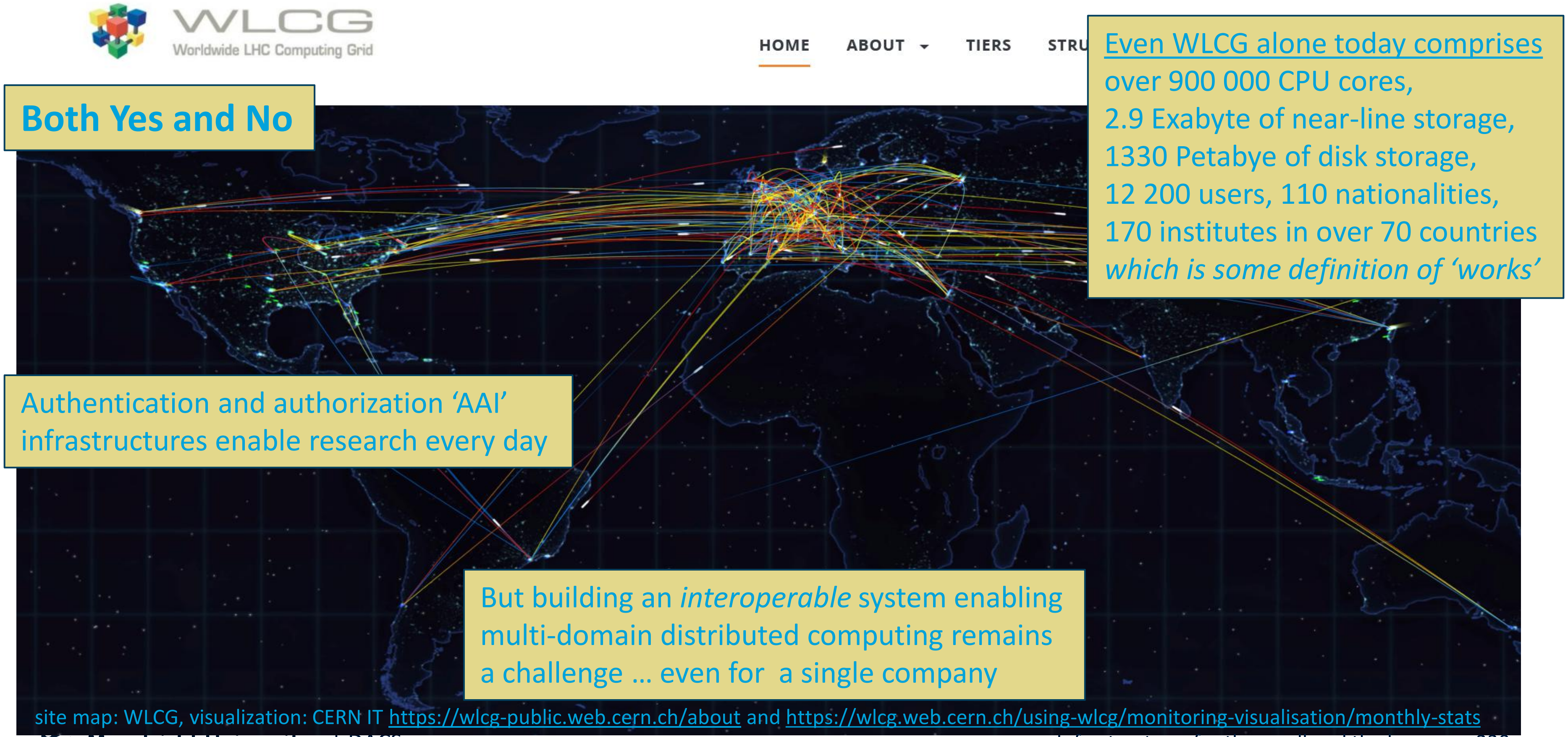
Documentation for the public APIs of SURF Res

Organisation All endpoints for organisations using an organisation API token

POST	`/api/collaborations/v1`	Post a new collaboration.	`post_api_collaborations_v1`
DELETE	`/api/collaborations/v1/{co_identifier}`	Delete collaboration.	`delete_api_collaborations_v1__co_identifier_`
PUT	`/api/collaborations/v1/{co_identifier}/members`	Update collaboration membership.	`put_api_collaborations_v1__co_identifier__members`
DELETE	`/api/collaborations/v1/{co_identifier}/members/{user_uid}`	Delete collaboration membership.	`delete_api_collaborations_v1__co_identifier__members__user_uid_`
PUT	`/api/collaborations/v1/{co_identifier}/units`	Update collaboration units.	`put_api_collaborations_v1__co_identifier__units`
PUT	`/api/collaborations_services/v1/connect_collaboration_service`	Connect service to collaboration.	`put_api_collaborations_services_v1_connect_collaboration_service`
PUT	`/api/collaborations_services/v1/disconnect_collaboration_service`	Disconnect service from collaboration.	`put_api_collaborations_services_v1_disconnect_collaboration_service`

Photo by sunrise University on Unsplash; network diagram: FSE CSLab, Maastricht University; SRAM API: https://sram.surf.nl/apidocs/

So did we solve our - research - infrastructure challenge?



Both Yes and No

<u>Even WLCG alone today comprises</u>
over 900 000 CPU cores,
2.9 Exabyte of near-line storage,
1330 Petabye of disk storage,
12 200 users, 110 nationalities,
170 institutes in over 70 countries
which is some definition of 'works'

Authentication and authorization 'AAI' infrastructures enable research every day

But building an *interoperable* system enabling multi-domain distributed computing remains a challenge … even for a single company

site map: WLCG, visualization: CERN IT https://wlcg-public.web.cern.ch/about and https://wlcg.web.cern.ch/using-wlcg/monitoring-visualisation/monthly-stats

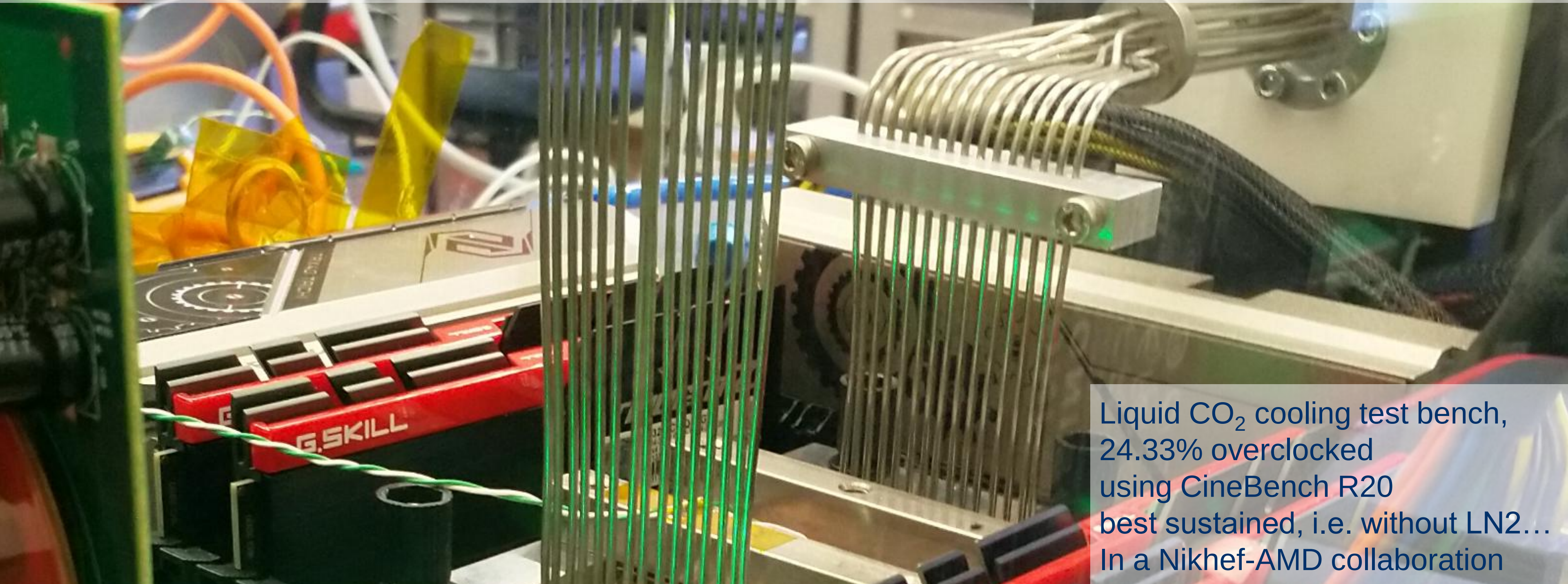
Looking for the common pattern …

- It’s all about *balance*
  - systems are like congested highways: no use solving just *one* bottleneck
  - and the bottlenecks may be inside the system
    as well as in interconnects, trust, interoperability, and governance

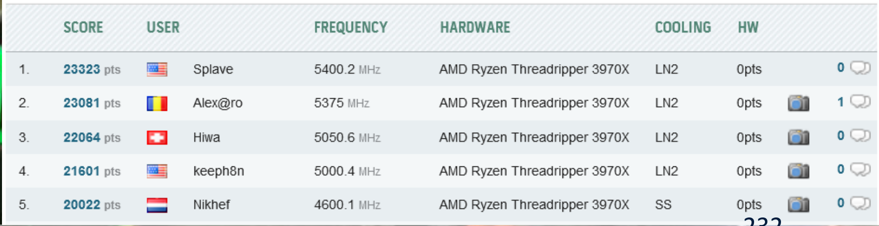
- Scaling *collaboration and trust* federation is as complex as scaling systems
  - composing services across administrative domains is ubiquitous
  - but beyond a certain size, $\mathcal{O}(100)$, you will also find need for policy and review

And you may move problems around, but it’s hard to actually *solve* them!

… since some things are fun, but not quite *that* scalable …



Liquid CO_2 cooling test bench,
24.33% overclocked
using CineBench R20
best sustained, i.e. without LN2…
In a Nikhef-AMD collaboration



	SCORE	USER		FREQUENCY	HARDWARE	COOLING	HW		
1.	23323 pts		Splave	5400.2 MHz	AMD Ryzen Threadripper 3970X	LN2	0pts		0
2.	23081 pts		Alex@ro	5375 MHz	AMD Ryzen Threadripper 3970X	LN2	0pts		1
3.	22064 pts		Hiwa	5050.6 MHz	AMD Ryzen Threadripper 3970X	LN2	0pts		0
4.	21601 pts		keeph8n	5000.4 MHz	AMD Ryzen Threadripper 3970X	LN2	0pts		0
5.	20022 pts		Nikhef	4600.1 MHz	AMD Ryzen Threadripper 3970X	SS	0pts		0

T Suerink, K de Roo: https://hwbot.org/submission/4539341_nikhef_cinebench___r20_with_benchmate_ryzen_threadripper_3970x_20022_pts

Discussion time … !

All views expressed herein are my own, and do not necessarily reflect the position of any co-funding organisations, the European Commission, SURF, NWO, or collaborators mentioned below. Yet works presented here are part of many collaborations. Thanks especially to – in random order – Mischa Sallé, Tristan Suerink, Dennis van Dok, Mary Hester, Andrew Pickford, Jeff Templon, Roel Aaij, Emily Kooistra, Jouke Roorda, Sven Gabriel, Lennie de Roo, Arjen van Rijn, Krista de Roo (Nikhef), Nicolas Liampotis, Kyriakos Glinis (GRNET), Liam Atherton, Jens Jensen, Dave Kelsey, David Crooks (STFC RAL), Hannah Short, Liviu Valsan (CERN), Uros Stevanovic, Marcus Hardt (KIT), Maarten Kremers (SURF), Licia Florio (NORDUNET), Christos Kanellopoulos, Klaas Wieringa (GEANT), Tom Barton (Internet2, UChicago), Tiziana Ferrari, Matt Viljoen, Baptise Grenier (EGI.eu), and the EGI, GEANT, PRACE, WLCG, REFEDS, WISE, IGTF, AARC & FIM4R communities!

David Groep
david.groep@maastrichtuniversity.nl
davidg@nikhef.nl Nikhef
https://www.nikhef.nl/~davidg/presentations/
iD https://orcid.org/0000-0003-1026-6606

Elements of this work have been co-supported by projects that have received funding from the European Union's Horizon research and innovation programmes under grants GN4-3, EOSC Future, AARC2, AARC-TREE, GN5-1, and GN5-2.

SURF Elements of this work have been co-supported by SURF, the collaborative organisation for IT in Dutch education and research.

This work has also been co-supported by projects that have received funding from the European Union's Horizon research and innovation programmes under AARC-TREE and GN5-2.

This work is supported by SURF under the Innovation Programme, part of the Execution Plan Digital Infrastructures for Research and other sources

time for (more) discussion ...

David Groep
davidg@nikhef.nl
https://www.nikhef.nl/~davidg/presentations/
https://orcid.org/0000-0003-1026-6606

Maastricht University

Nikhef

Nulla folia post hoc sunt

Thanks for watching!

“En daarmee, geachte luisteraars, laat ik u over aan
de verpozing die uw babbelklant u gemeenlijk pleegt te bieden.”