# Building RCauth

## *a proxy for our federated global infrastructure*

A multinational service
for federated authentication in research infrastructure
using a networked-systems integration approach

Nikhef

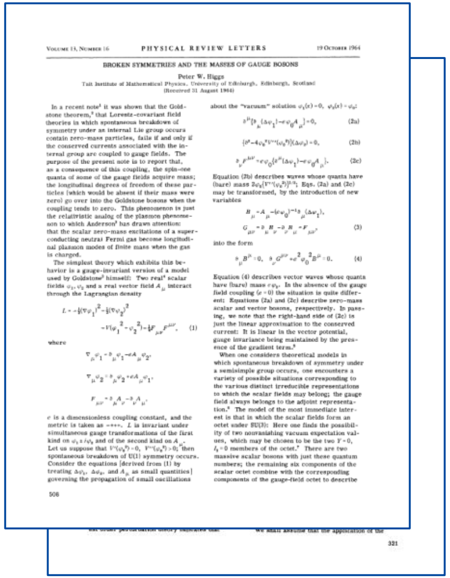
David Groep
DACS & Nikhef
March 2023

Maastricht University | Department of Advanced Computing Sciences

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

# A ‘big science’ facility: 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 $\varphi_1, \varphi_2$ and a real vector field $A_\mu$ interact through the Lagrangian density

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

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 $A_\mu$. 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 $A_\mu$ as small quantities] governing the propagation of small oscillations about the "vacuum" solution $\varphi_1(x)=0$, $\varphi_2(x)=\varphi_0$:

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

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

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

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

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

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

into the form

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

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]

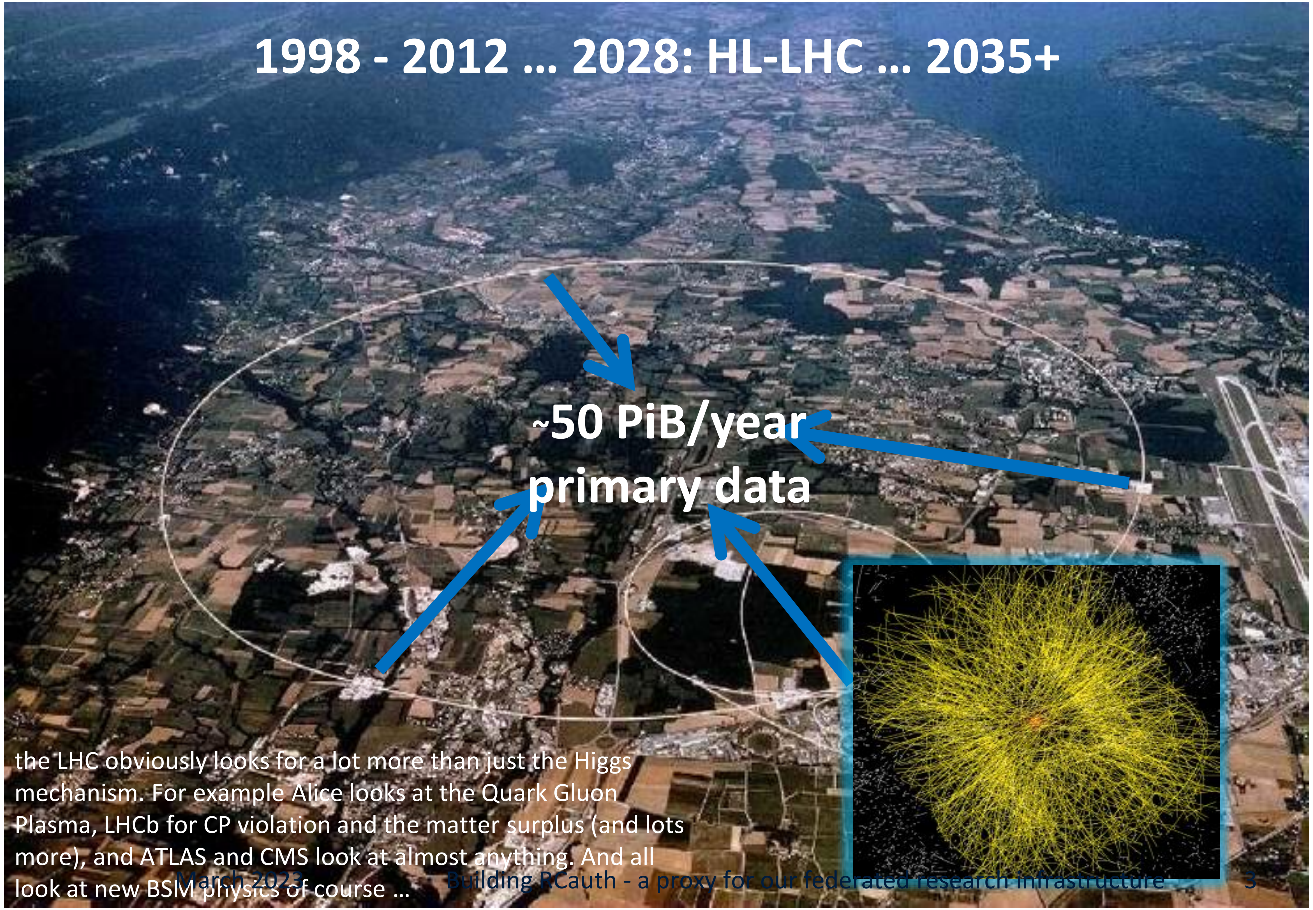
When one considers theoretical models in which spontaneous breakdown of symmetry under a semisimple group occurs, one encounters a variety of possible situations corresponding to the various distinct irreducible representations to which the scalar fields may belong; the gauge field always belongs to 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

321

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

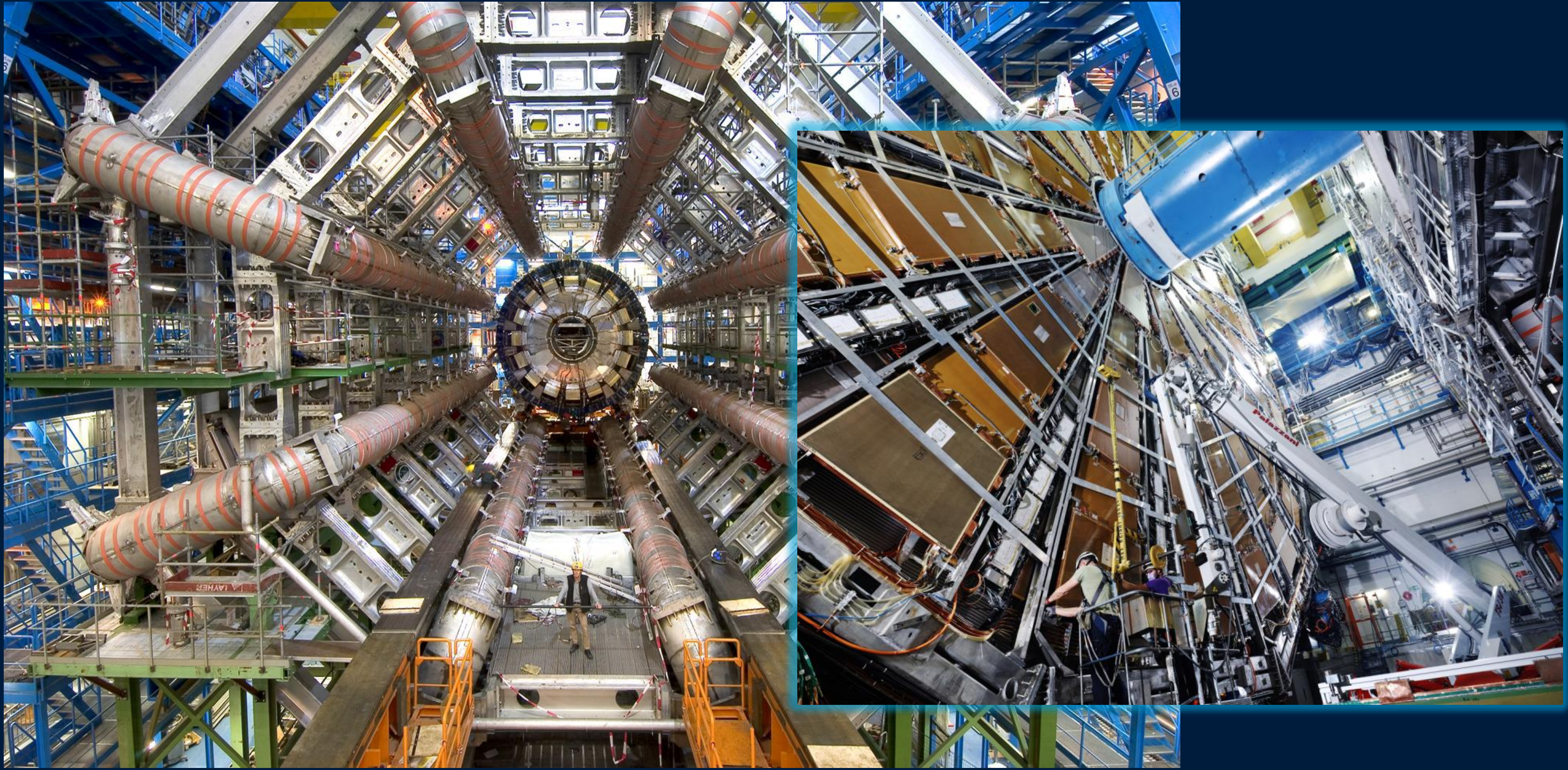
**16823 characters, 165 kByte PDF**



1998 - 2012 … 2028: HL-LHC … 2035+

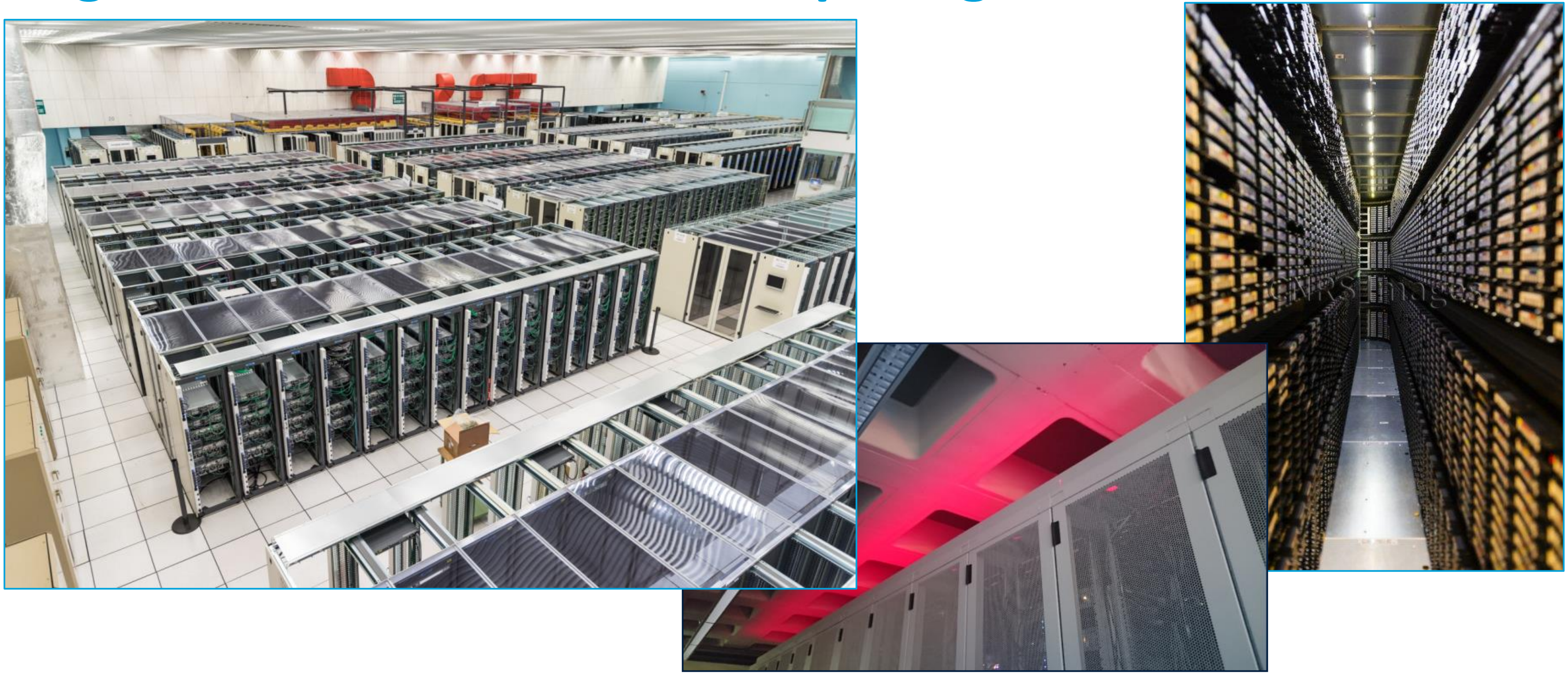
~50 PiB/year primary data

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 …



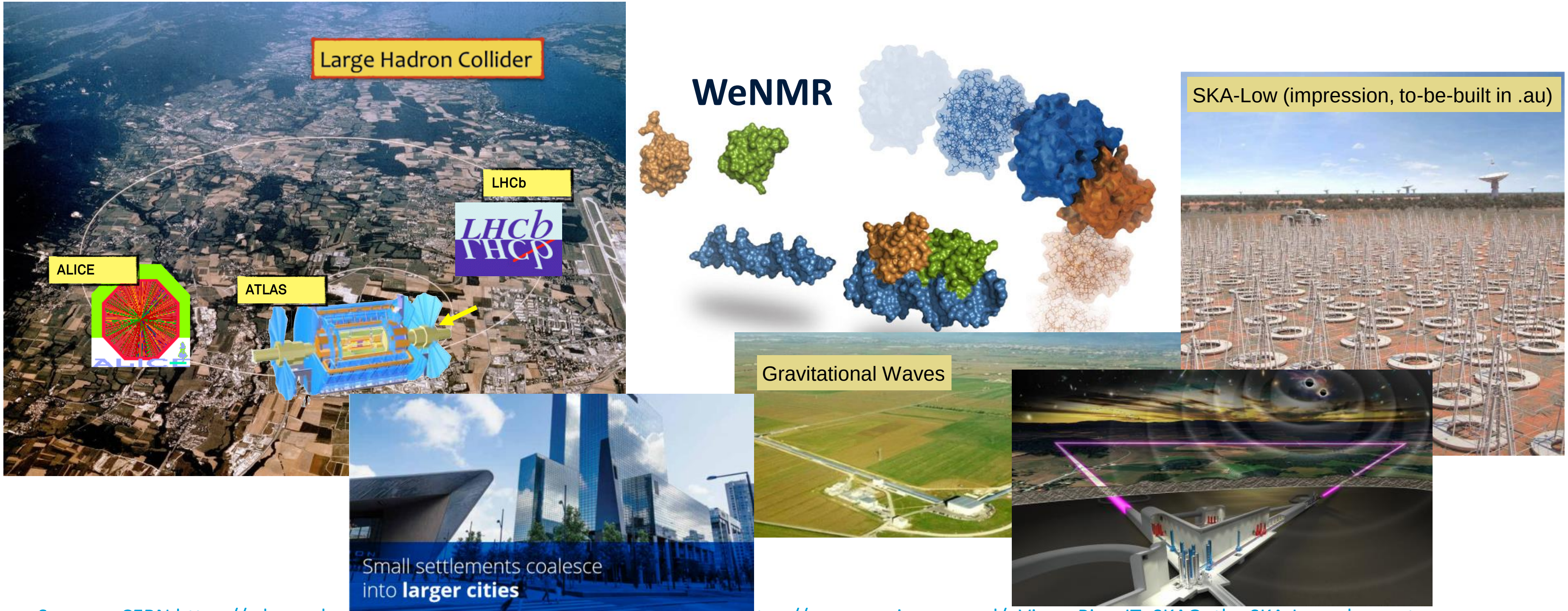
Images: ATLAS detector in the cavern at CERN. Source: CERN

# ‘Big Science’ needs some computing …



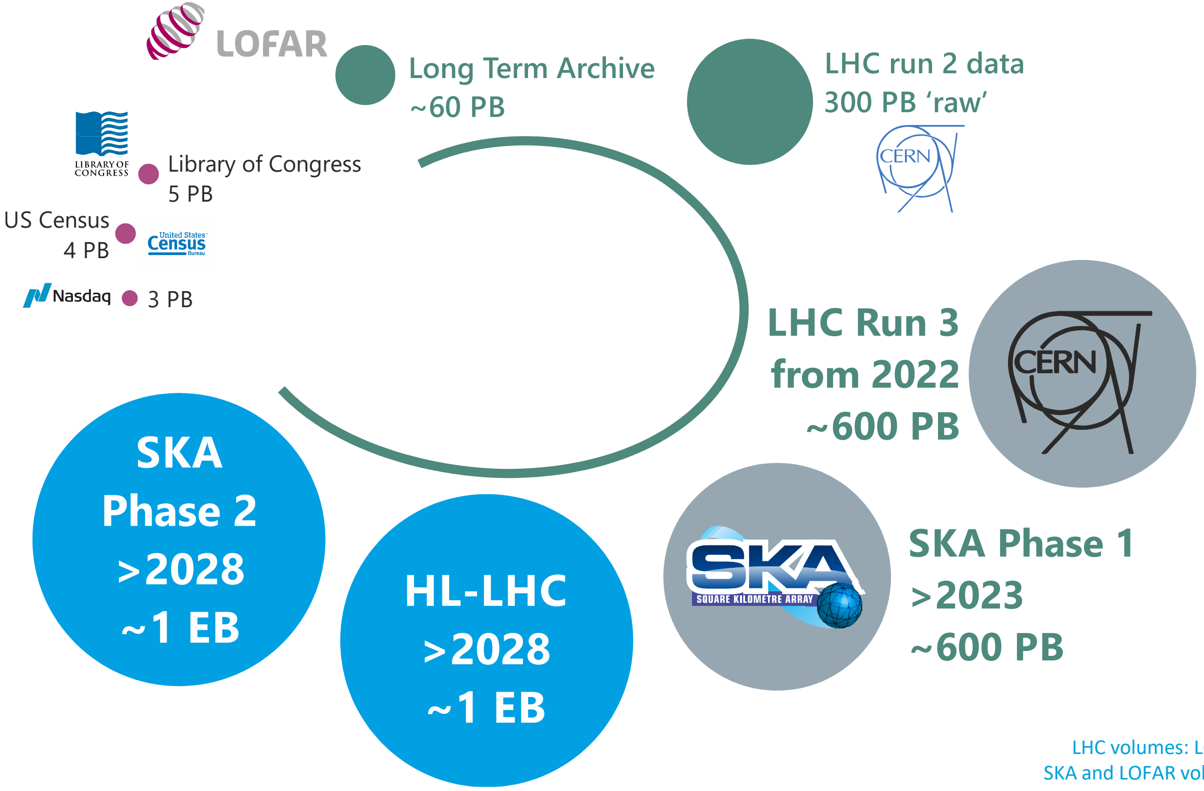
CERN Computing Centre B513, image: CERN, https://cds.cern.ch/record/2127440; tape library image CC-IN2P3 with LHC and LSST data; cabinets: Nikhef H234b

# Larger scales for both facilities and computing


Large Hadron Collider
LHCb
ALICE
ATLAS
WeNMR
SKA-Low (impression, to-be-built in .au)
Gravitational Waves
Small settlements coalesce into larger cities

Sources: 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

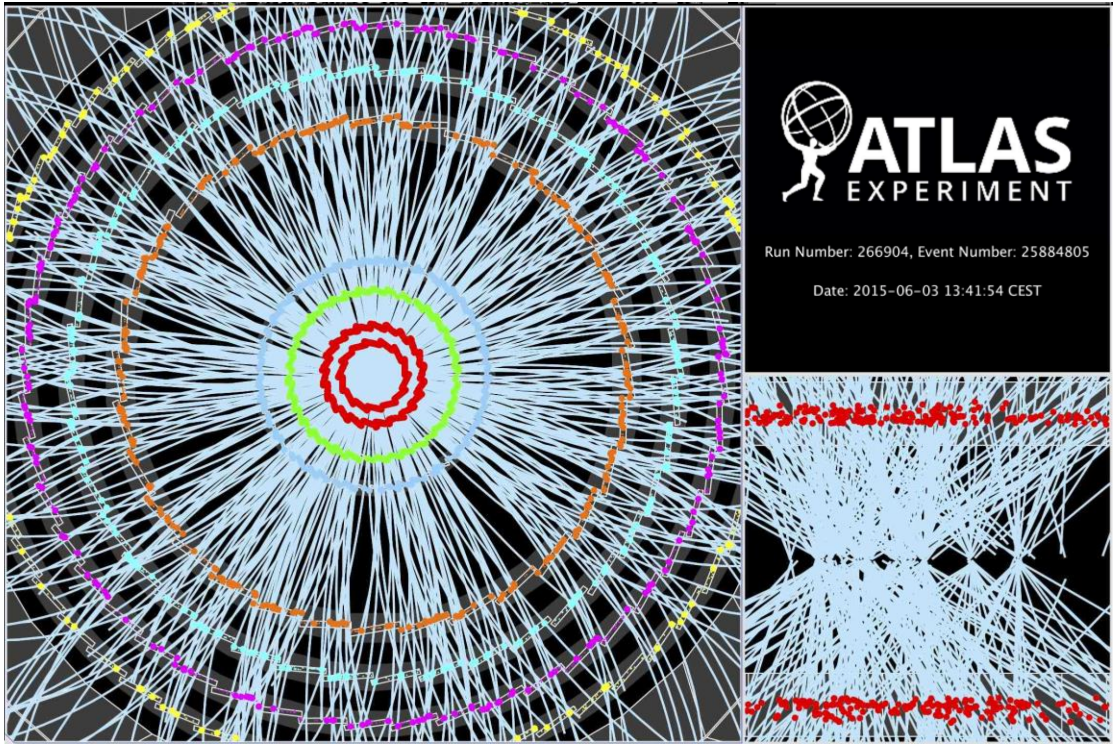
# Processing at scale for data intensive science



Data from various sources, for
public entities: data ca. 2018,
indicative, within ~ factor 2
LHC volumes: LCG Resource Scrutiny Group & CERN; 2020
SKA and LOFAR volumes: ASTRON/Michiel van Haarlem, 2020

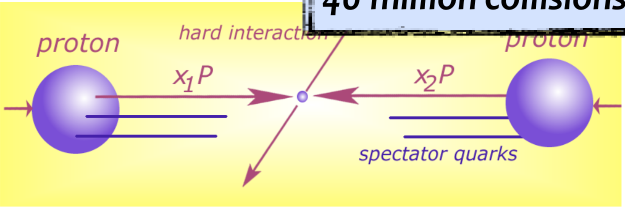
# Computing on lots of data – 40Mevents/sec

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

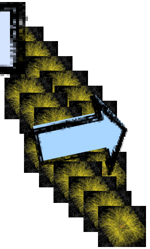

ATLAS
EXPERIMENT
Run Number: 266904, Event Number: 25884805
Date: 2015-06-03 13:41:54 CEST

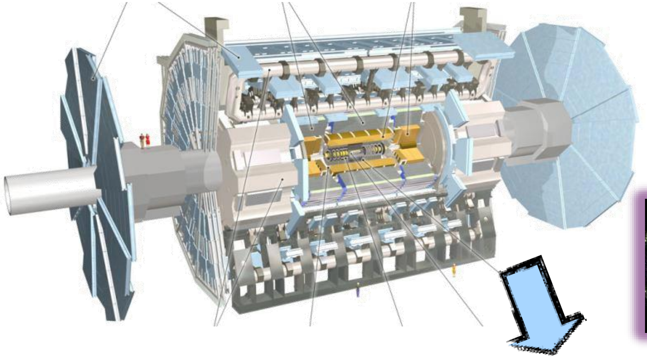
Display of a proton-proton collision event recorded by ATLAS on 3 June 2015, with the first LHC stable beams at a collision energy of 13 TeV;
Event processing time: v19.0.1.1 as per Jovan Mitrevski and 2015 J. Phys.: Conf. Ser. 664 072034 (CHEP2015)

# Detector to doctor workflow

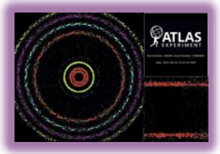


*40 million collisions / second*





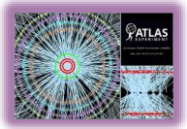
*Trigger system selects 600 Hz ~ 1 GB/s data*



**Classify particles in collision and their physics properties:**

- *electrons*
- *muons*
- *jets consisting of hadrons*

...



***Physics analysis by (PhD) students, in papers & analysis notes***

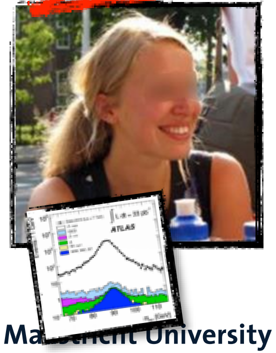
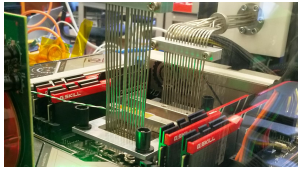


diagram adapted from Frank Linde; images: ATLAS collaboration, Nikhef. … and sorry for the GDPR-blur

# Processing … at different scales

## algorithms and systems design

- designing for accelerators and high-performance processors
- rethinking design patterns for work & data orchestration



**algorithms and systems**

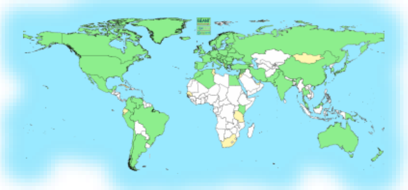
## collective compute, storage, and networks

- building ‘research IT facilities’
- peering and global networks
- stressing the network
- research ‘cloudy’ services

**systems and interactions**

## accessing services collaboratively & securely

- trust and identity services
- securing the infrastructure of an open science cloud
- managing complexity of collaboration mechanisms



**interactions and people**

# More than one system

‘HTC’ – high throughput computing
sharing workflows across multiple sites
and multiple solutions

# Physical farms: selecting the ‘worker nodes’

For HTC applications
– like WLCG, SKA, WeNMR – typically

- **balanced features for node throughput**
  (CPU, storage, memory bandwidth, network)

- **single-socket** multicore systems are fine, typical: 64-128 cores per system
- **network**: 2x25Gbps
  (+ ‘out of band’ management like IPMI)
- **memory**: 8 GiB/core
- **local disk**: 4TB NVME PCIe Gen4 x4

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

Image: Cluster ‘Lotenfeest’ at the Nikhef NDPF, acquired March 2020. Lenovo SR655 with AMD EPYC 7702P 64-Core single-socket

# Using clusters of nodes in a scalable way

Batch queuing and distributing jobs

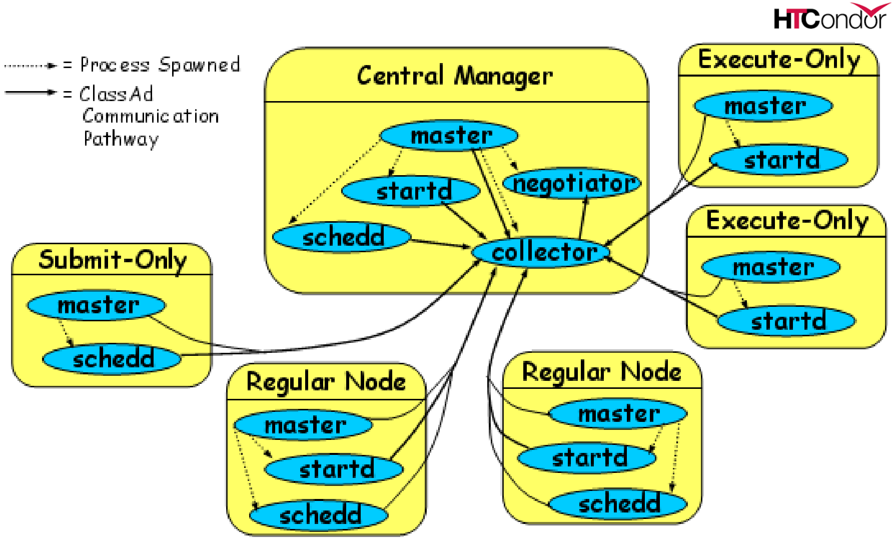
- scheduling over a known set of nodes (SLURM, Torque, …)

Or matchmaking based on 'ClassAds'

- both jobs and machines advertise their requirements and capabilities in 'classified advertisements'
- Matchmaking done by the negotiator execution nodes mostly autonomous

helps for scalability and resilience, but

- handling fair-share based allocation and accounting is much more difficult (no central component to do that …)

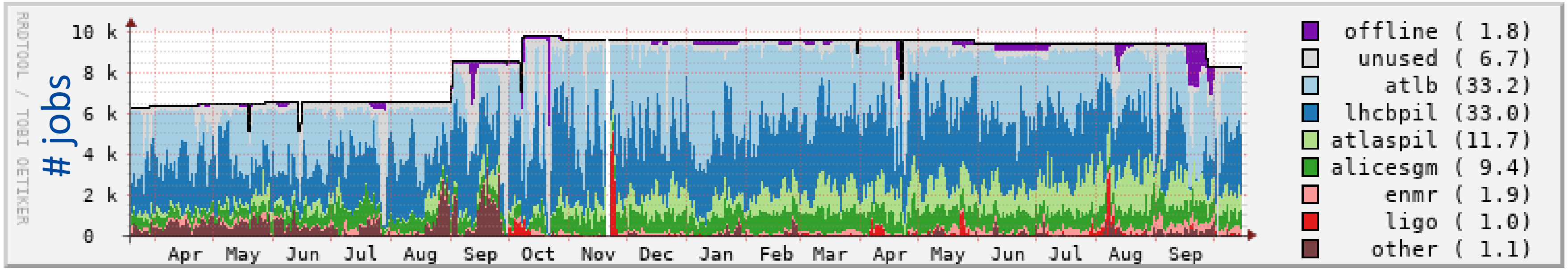


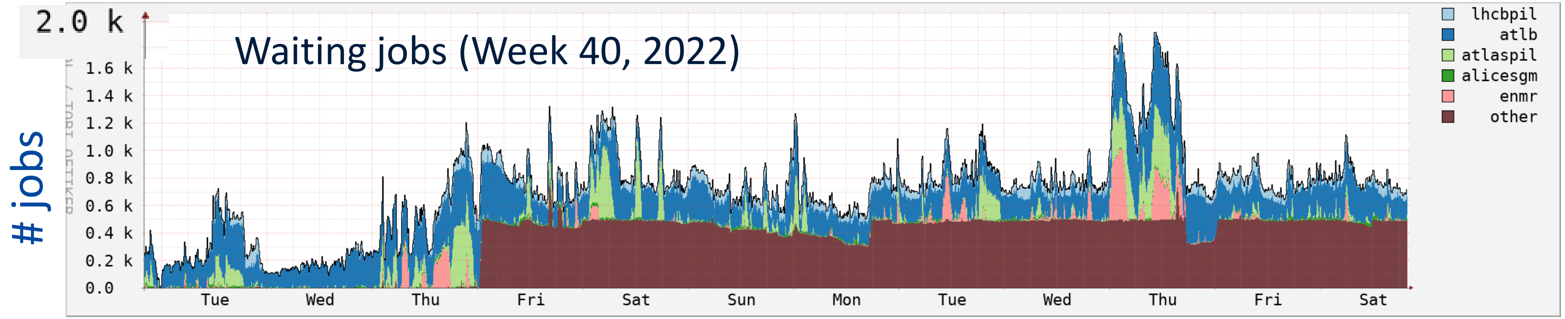
HTCondor, Miron Livny et al, UWMadison; https://research.cs.wisc.edu/htcondor/CondorWeek2008/condor_presentations/desmet_admin_tutorial/

# NDPF ‘WLCG and Dutch National Infra’ cluster

Running jobs:

period: March 2021 .. October 2022

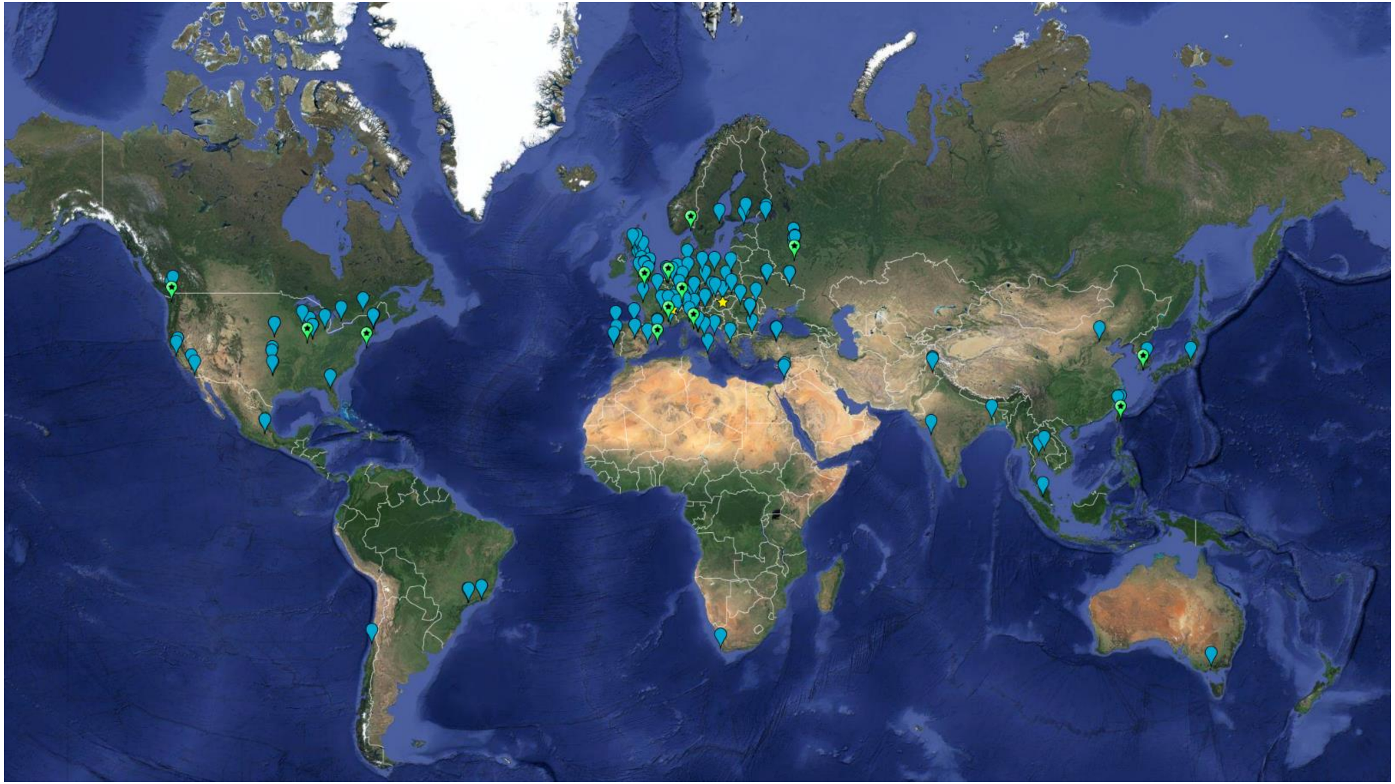




drainage event on Sept 27 are nodes being moved to the LIGO-VIRGO specific cluster; Source: NDPF Statistics overview, https://www.nikhef.nl/pdp/doc/stats/
‘other’ waiting jobs are almost all for the Auger experiment - GRISview images: Jeff Templon for NDPF and STBC

Image source: Free Software Foundation Europe - https://fsfe.org/

# Example: the worldwide LHC Computing Grid



~ 1.4 million CPU cores
~ 1500 Petabyte
disk + archival

170+ institutes
40+ countries
13 'Tier-1 sites'
**NL-T1:**
**SURF & Nikhef**

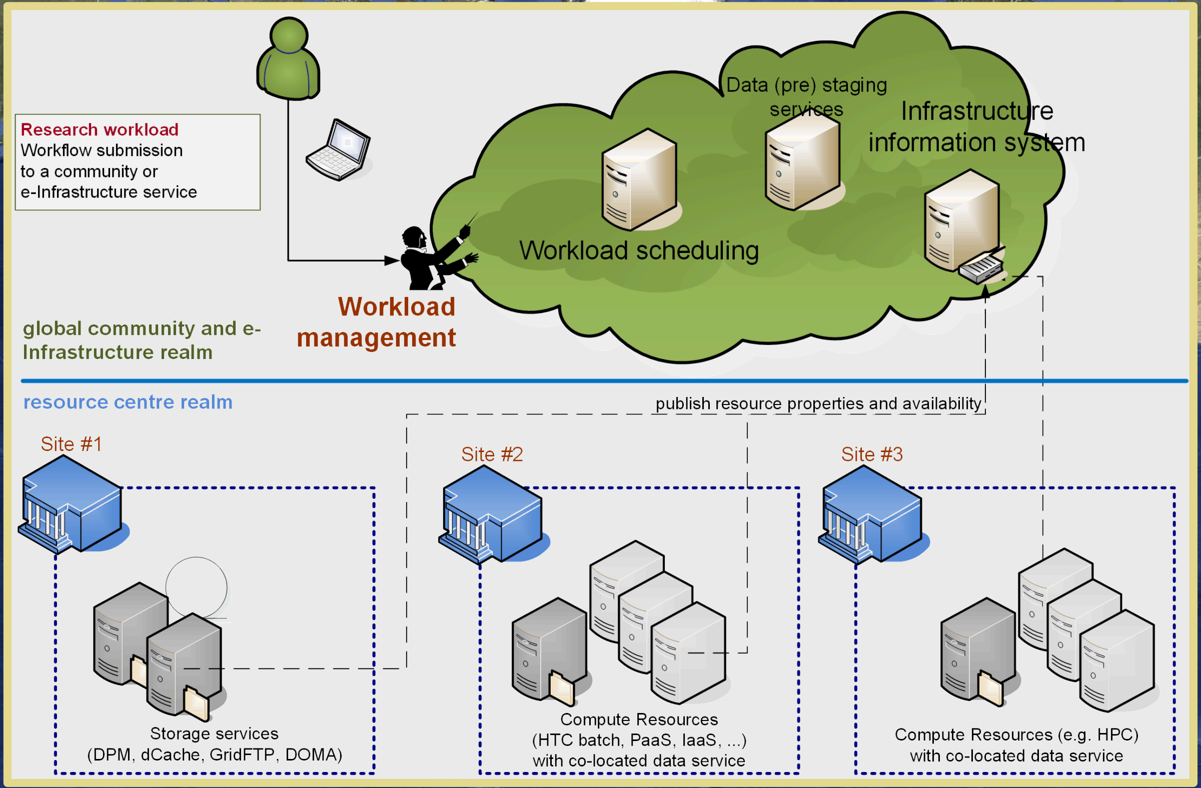
*e-Infrastructures*
EGI
PRACE-RI
EuroHPC
OpenScienceGrid
XSEDE (ACCESS)

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; ACCESS (XSEDE): https://access-ci.org/, for the NL-T1 and FuSE: fuse-infra.nl, https://www.surf.nl/en/research-it

# Global distribution of computing and data placement

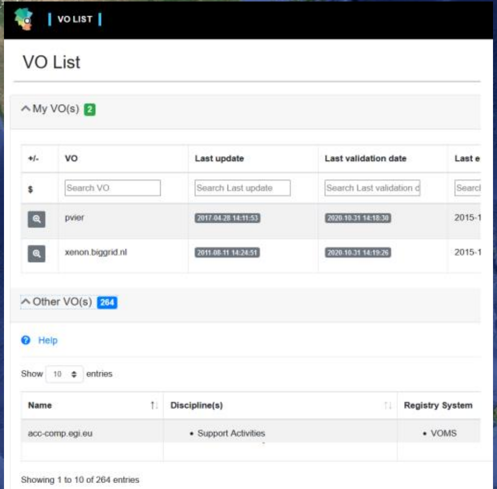
WLCG and EGI Advanced Computing for Research

# Conveniently parallel: a global infrastructure for research


Research workload
Workflow submission to a community or e-Infrastructure service
Data (pre) staging services
Infrastructure information system
Workload scheduling
Workload management
global community and e-Infrastructure realm
resource centre realm
publish resource properties and availability
Site #1
Site #2
Site #3
Storage services (DPM, dCache, GridFTP, DOMA)
Compute Resources (HTC batch, PaaS, IaaS, ...) with co-located data service
Compute Resources (e.g. HPC) with co-located data service

shared multi-community infrastructure

Already EGI e-infra has >250 communities just doing HTC


VO LIST

VO List

My VO(s) 2

| +/- | VO | Last update | Last validation date | Last e |
|---|---|---|---|---|
| $ | Search VO | Search Last update | Search Last validation d | Searc |
|  | pvier | 2017-04-28 14:11:53 | 2020-10-31 14:18:30 | 2015-1 |
|  | xenon.biggrid.nl | 2011-08-11 14:24:51 | 2020-10-31 14:19:26 | 2015-1 |

Other VO(s) 264

Help

Show 10 entries

| Name | Discipline(s) | Registry System |
|---|---|---|
| acc-comp.egi.eu | • Support Activities | • VOMS |

Showing 1 to 10 of 264 entries

Right-hand graphic: EGI operations portal, https://operations-portal.egi.eu/vo/

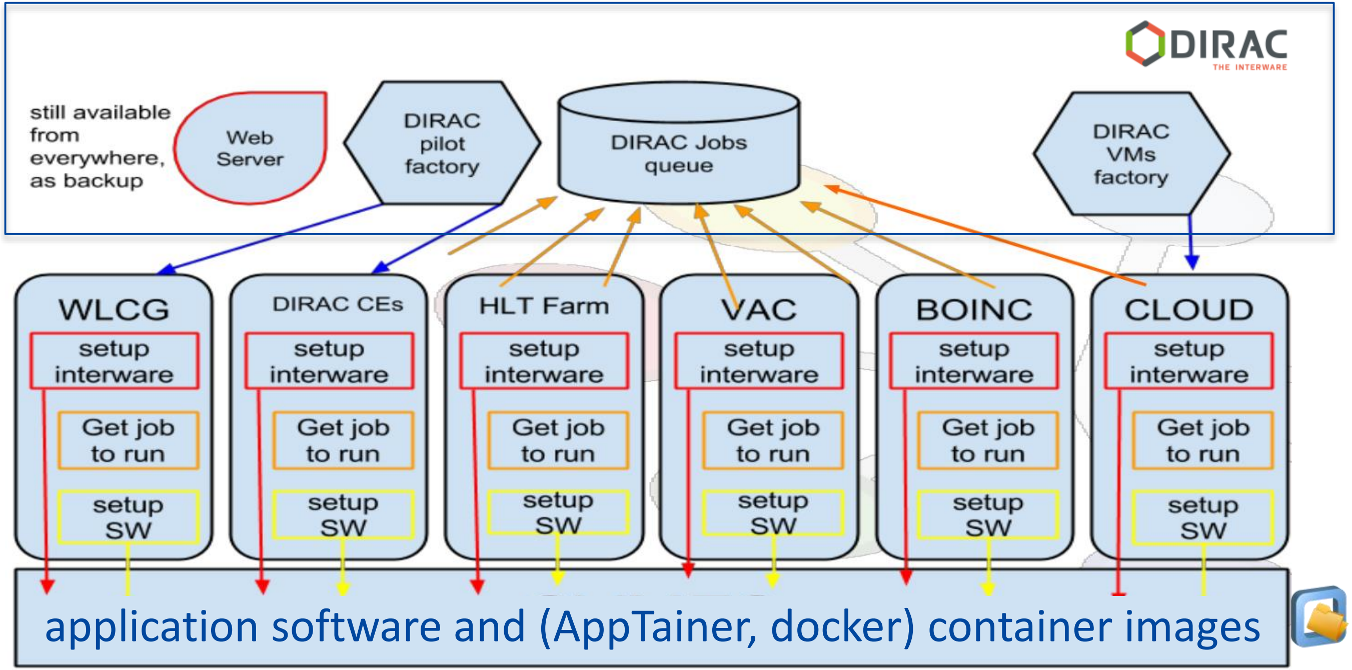
# Brokered access spanning heterogeneous resources

Adding a scheduling layer on top

since all sites are autonomous, and global standards failed

‘any (IT) problem can be solved by adding one layer of indirection’

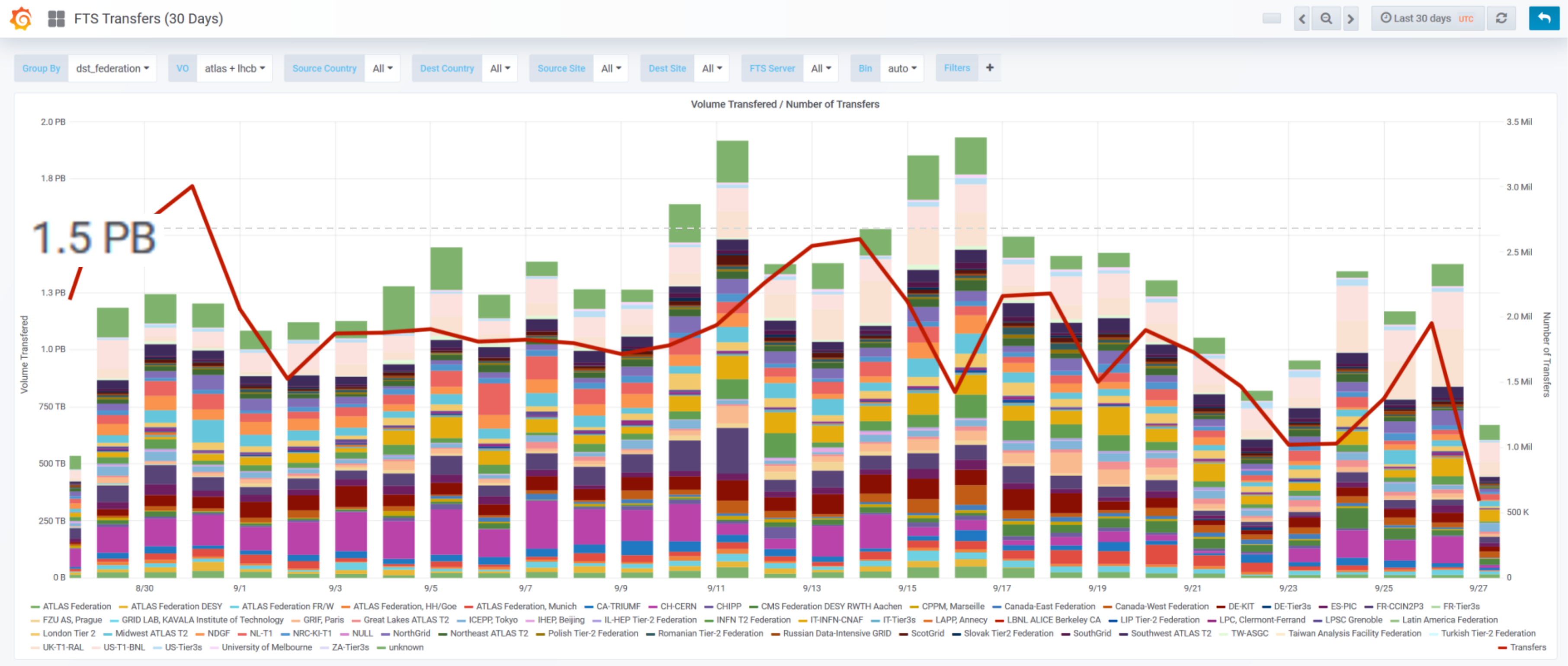
DIRAC is just one example



application software and (AppTainer, docker) container images

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/

# High throughput computing is also about data



FTS Transfers (30 Days)

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

Last 30 days UTC

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

# Connecting together

application design and the network
separating network flows for high-throughput data processing
stressing the network … beyond LHC data processing

# This does not quite cut it for the LHC

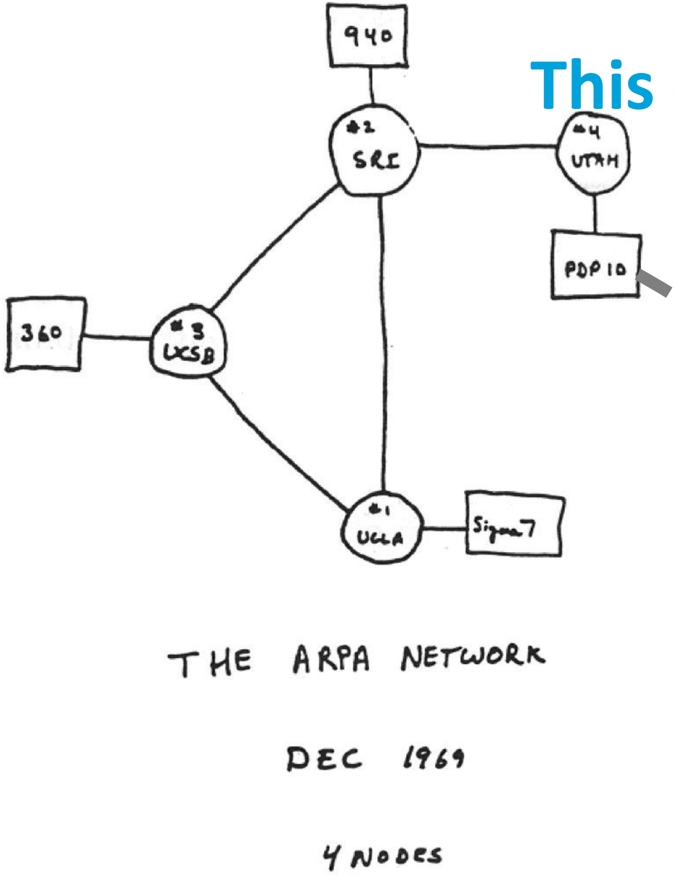
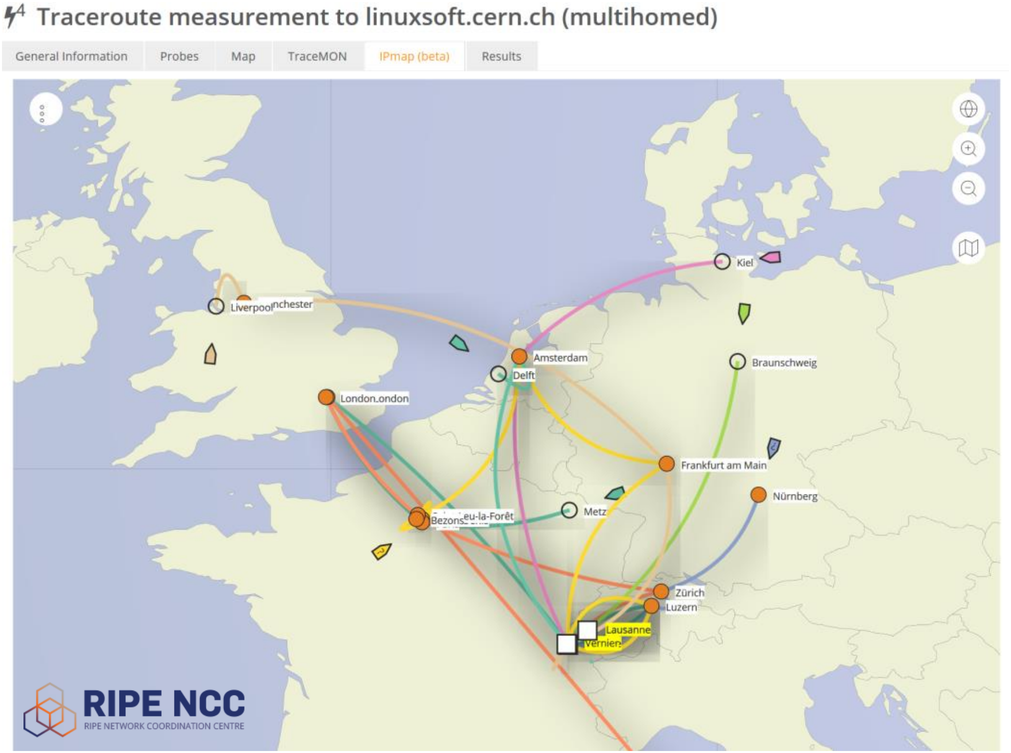


Image source: Alex McKenzie and "Casting the Net", page 56. See https://personalpages.manchester.ac.uk/staff/m.dodge/cybergeography/atlas/arpanet2.gif ; acoustocoupler: Wikimedia

# A quick look at internet routing …

network paths
from various places
in Western Europe

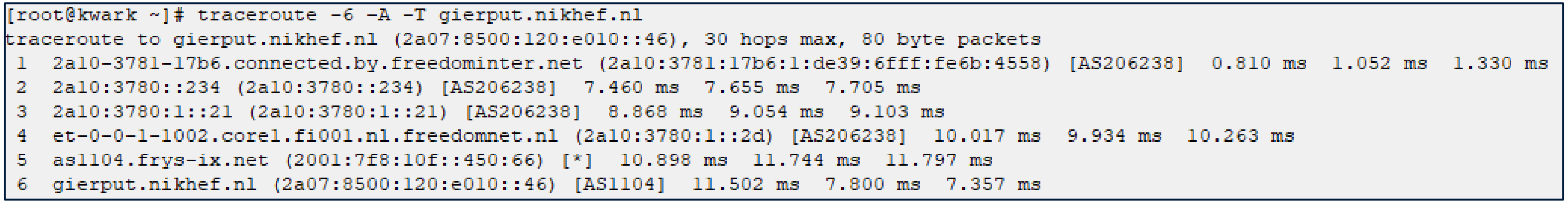
towards an IP address
at CERN



Data: RIPE NCC Atlas project, TraceMON IPmap, atlas.ripe.net, measurement 9249079

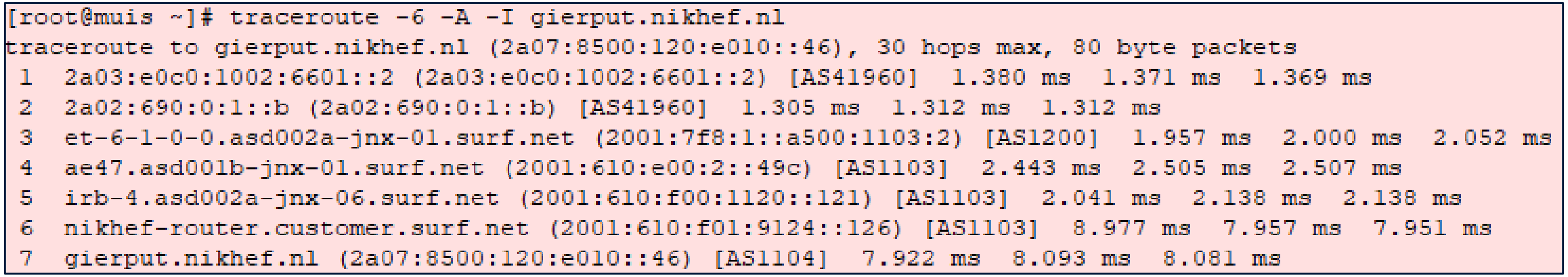
# Many paths to Rome … i.e. to your server

From a home connected to the Freedom Internet ISP to *spiegel.nikhef.nl*



```
[root@kwark ~]# traceroute -6 -A -T gierput.nikhef.nl
traceroute to gierput.nikhef.nl (2a07:8500:120:e010::46), 30 hops max, 80 byte packets
 1  2a10-3781-17b6.connected.by.freedominter.net (2a10:3781:17b6:1:de39:6fff:fe6b:4558) [AS206238]  0.810 ms  1.052 ms  1.330 ms
 2  2a10:3780::234 (2a10:3780::234) [AS206238]  7.460 ms  7.655 ms  7.705 ms
 3  2a10:3780:1::21 (2a10:3780:1::21) [AS206238]  8.868 ms  9.054 ms  9.103 ms
 4  et-0-0-1-1002.core1.fi001.nl.freedomnet.nl (2a10:3780:1::2d) [AS206238]  10.017 ms  9.934 ms  10.263 ms
 5  as1104.frys-ix.net (2001:7f8:10f::450:66) [*]  10.898 ms  11.744 ms  11.797 ms
 6  gierput.nikhef.nl (2a07:8500:120:e010::46) [AS1104]  11.502 ms  7.800 ms  7.357 ms
```

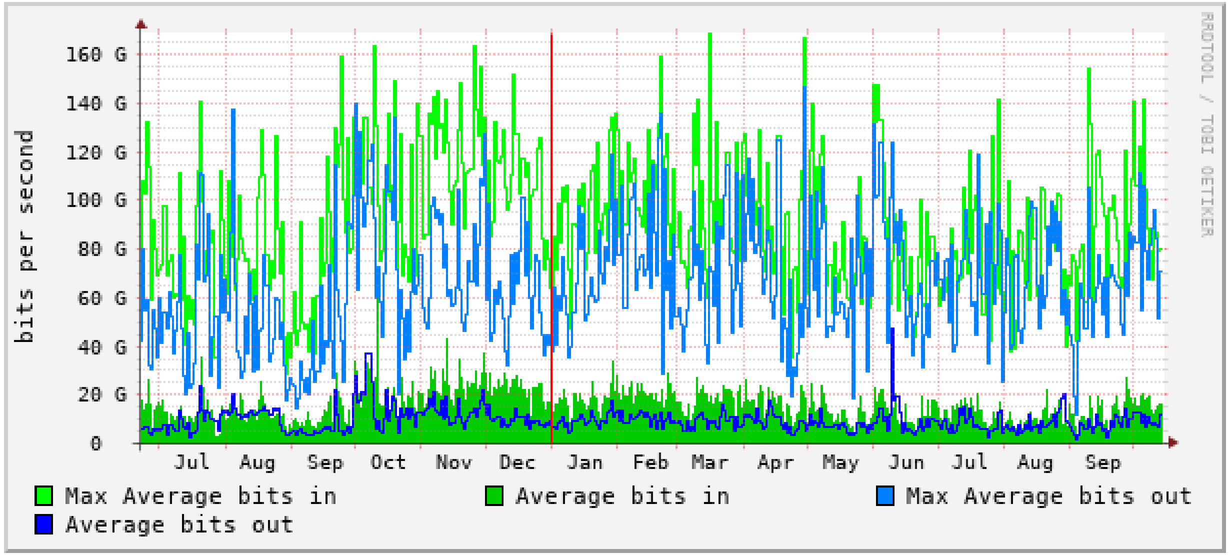
but from Interparts in Lisse, NH:



```
[root@muis ~]# traceroute -6 -A -I gierput.nikhef.nl
traceroute to gierput.nikhef.nl (2a07:8500:120:e010::46), 30 hops max, 80 byte packets
 1  2a03:e0c0:1002:6601::2 (2a03:e0c0:1002:6601::2) [AS41960]  1.380 ms  1.371 ms  1.369 ms
 2  2a02:690:0:1::b (2a02:690:0:1::b) [AS41960]  1.305 ms  1.312 ms  1.312 ms
 3  et-6-1-0-0.asd002a-jnx-01.surf.net (2001:7f8:1::a500:1103:2) [AS1200]  1.957 ms  2.000 ms  2.052 ms
 4  ae47.asd001b-jnx-01.surf.net (2001:610:e00:2::49c) [AS1103]  2.443 ms  2.505 ms  2.507 ms
 5  irb-4.asd002a-jnx-06.surf.net (2001:610:f00:1120::121) [AS1103]  2.041 ms  2.138 ms  2.138 ms
 6  nikhef-router.customer.surf.net (2001:610:f01:9124::126) [AS1103]  8.977 ms  7.957 ms  7.951 ms
 7  gierput.nikhef.nl (2a07:8500:120:e010::46) [AS1104]  7.922 ms  8.093 ms  8.081 ms
```

AS41960: Interparts; AS1200: AMS-IX route reflector; AS1103: SURFnet; AS1104: Nikhef; AS206238: Freedom Internet – on the FrysIX there is direct L2 peering

# Typical data traffic to and from the processing cluster



Source: Nikhef cricket graphs period June 2021 – October 2022 – aggregated (research) traffic to external peers from deelqfx – https://cricket.nikhef.nl/

# That viral cat video 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

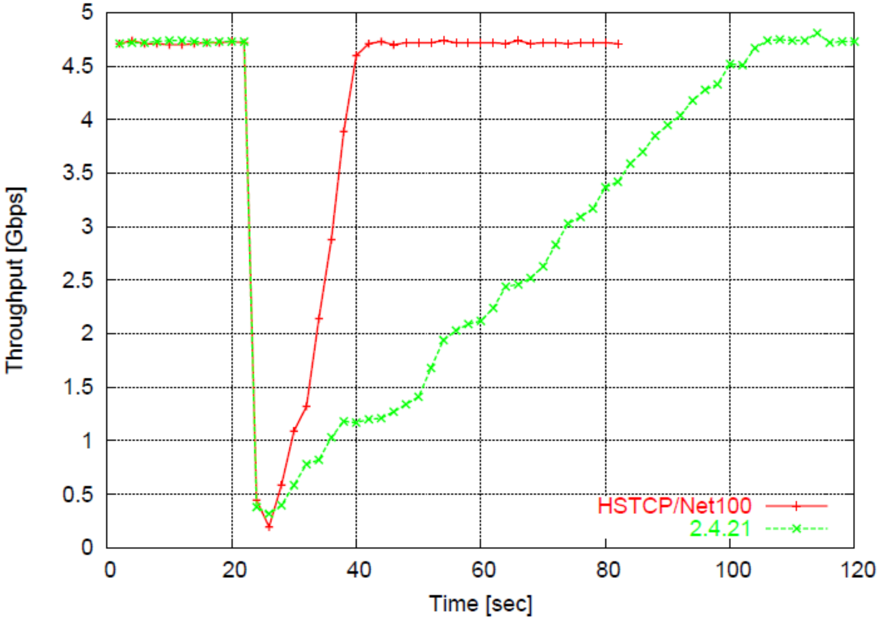


Figure 10: HSTCP versus stock TCP recovery time

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

# ‘Elephant streams in a packet-switched internet’

**Can cat videos survive in an internet dominated by big science data flows?**

Most of the internet is ‘packet switched’: each packet can go somewhere else …

so use waggons on a train, or ships,
that always go from A-to-B anyway?
A conveyer belt will do much better!

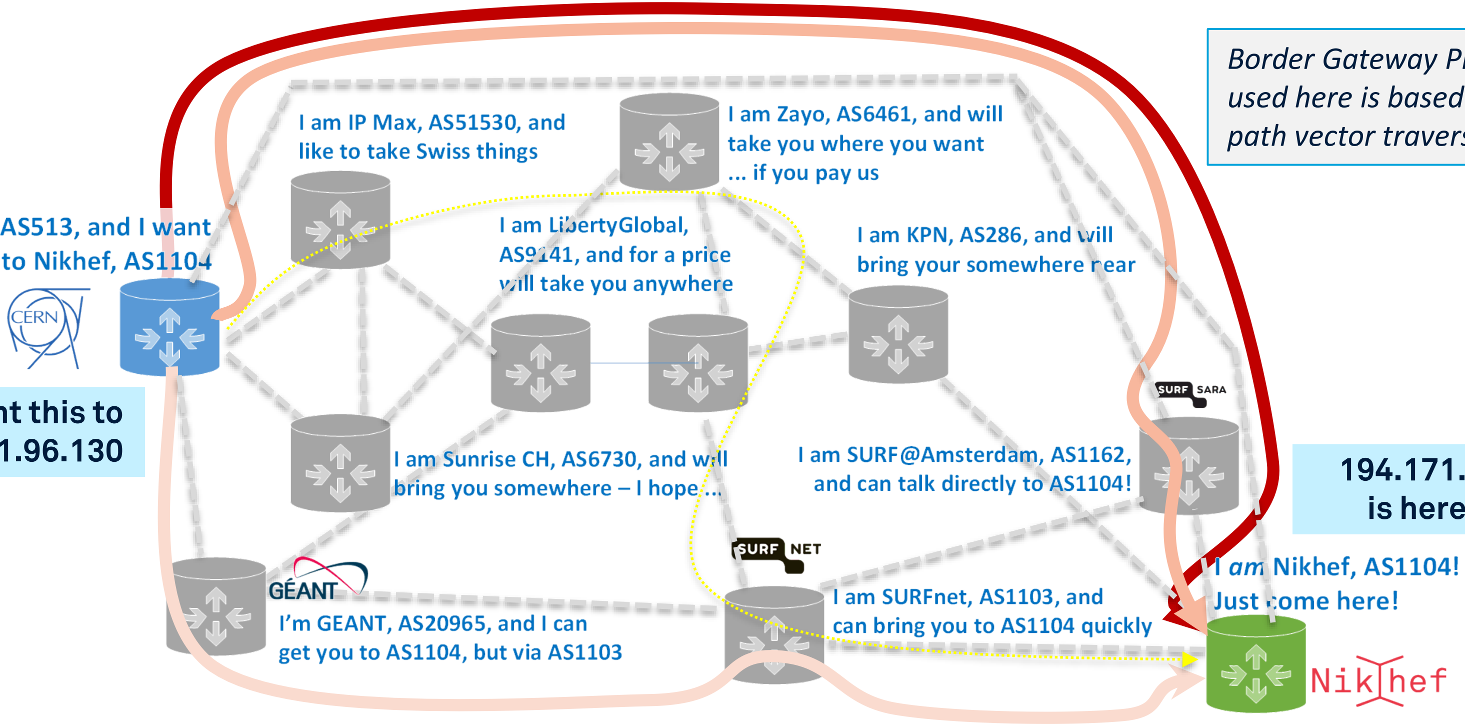
… *although you still need*
*a hole to dump it in …*

Image conveyor belt tunnel near Bluntisham, Cambridgeshire by Hugh Venables, CC-BY-SA-4.0 from https://www.geograph.org.uk/photo/4344525

# Where do internet packets go anyway?

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

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



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

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!

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

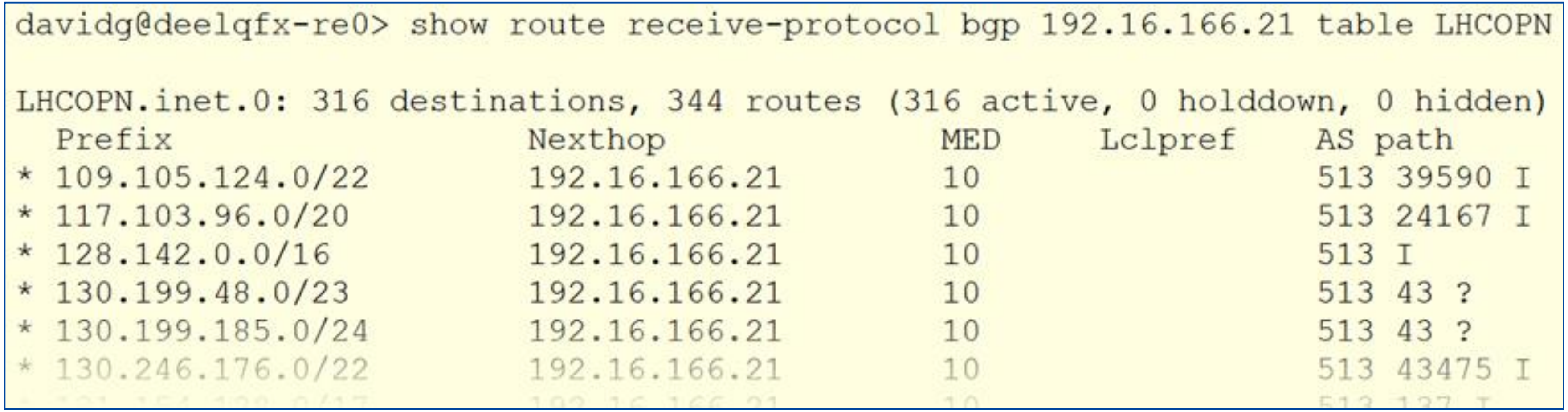
I am SURFnet, AS1103, and can bring you to AS1104 quickly

I *am* Nikhef, AS1104! Just come here!

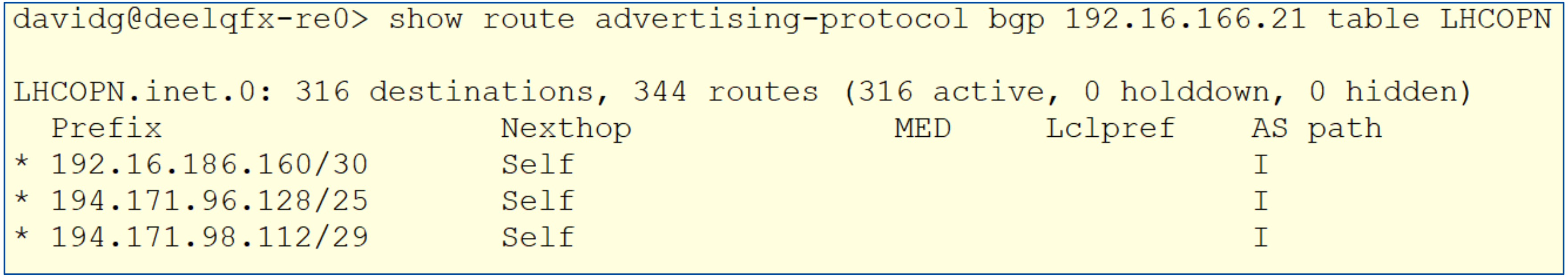
**194.171.96.128/25 is here at AS1104**

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)

# Announcing routes: the Border Gateway Protocol



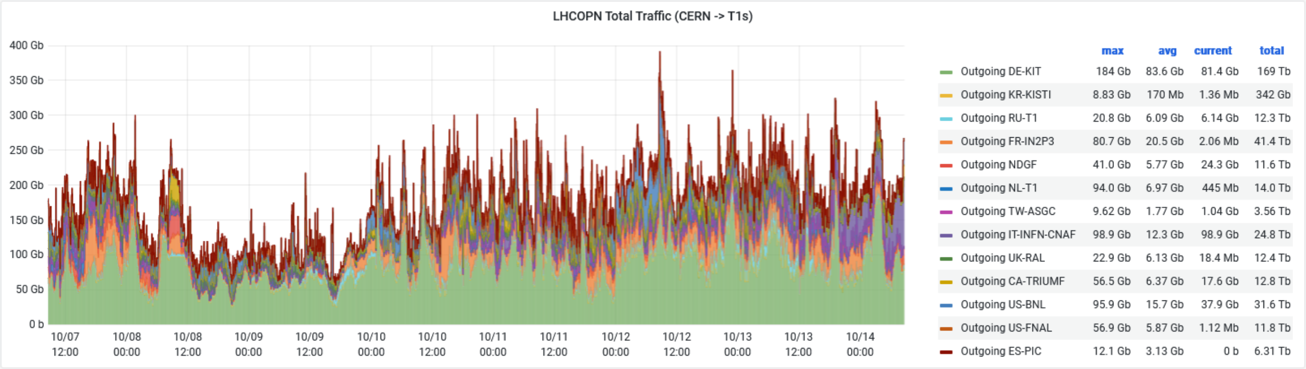
```
davidg@deelqfx-re0> show route receive-protocol bgp 192.16.166.21 table LHCOPN

LHCOPN.inet.0: 316 destinations, 344 routes (316 active, 0 holddown, 0 hidden)
  Prefix                  Nexthop              MED     Lclpref    AS path
* 109.105.124.0/22        192.16.166.21        10                 513 39590 I
* 117.103.96.0/20         192.16.166.21        10                 513 24167 I
* 128.142.0.0/16          192.16.166.21        10                 513 I
* 130.199.48.0/23         192.16.166.21        10                 513 43 ?
* 130.199.185.0/24        192.16.166.21        10                 513 43 ?
* 130.246.176.0/22        192.16.166.21        10                 513 43475 I
* 131.154.128.0/17        192.16.166.21        10                 513 137 I
```



```
davidg@deelqfx-re0> show route advertising-protocol bgp 192.16.166.21 table LHCOPN

LHCOPN.inet.0: 316 destinations, 344 routes (316 active, 0 holddown, 0 hidden)
  Prefix                  Nexthop              MED     Lclpref    AS path
* 192.16.186.160/30       Self                                    I
* 194.171.96.128/25       Self                                    I
* 194.171.98.112/29       Self                                    I
```

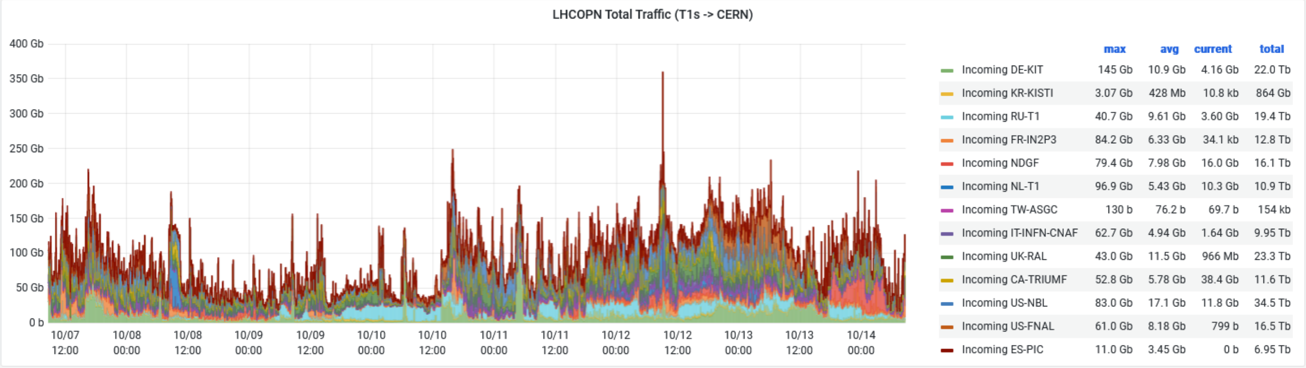
IPv4 routes advertised from AS513/CERN (for all sites on LHCOPN) to AS1104/Nikhef (top), and the routes announced by AS1104/Nikhef to CERN, on 5 Nov 2022

# LHCOPN – traffic levels for T1T1 data transfer



LHCOPN Total Traffic (CERN -> T1s)

| | max | avg | current | total |
|---|---|---|---|---|
| 🟩 Outgoing DE-KIT | 184 Gb | 83.6 Gb | 81.4 Gb | 169 Tb |
| 🟨 Outgoing KR-KISTI | 8.83 Gb | 170 Mb | 1.36 Mb | 342 Gb |
| 🟦 Outgoing RU-T1 | 20.8 Gb | 6.09 Gb | 6.14 Gb | 12.3 Tb |
| 🟧 Outgoing FR-IN2P3 | 80.7 Gb | 20.5 Gb | 2.06 Mb | 41.4 Tb |
| 🟥 Outgoing NDGF | 41.0 Gb | 5.77 Gb | 24.3 Gb | 11.6 Tb |
| 🟦 Outgoing NL-T1 | 94.0 Gb | 6.97 Gb | 445 Mb | 14.0 Tb |
| 🟪 Outgoing TW-ASGC | 9.62 Gb | 1.77 Gb | 1.04 Gb | 3.56 Tb |
| 🟪 Outgoing IT-INFN-CNAF | 98.9 Gb | 12.3 Gb | 98.9 Gb | 24.8 Tb |
| 🟩 Outgoing UK-RAL | 22.9 Gb | 6.13 Gb | 18.4 Mb | 12.4 Tb |
| 🟨 Outgoing CA-TRIUMF | 56.5 Gb | 6.37 Gb | 17.6 Gb | 12.8 Tb |
| 🟦 Outgoing US-BNL | 95.9 Gb | 15.7 Gb | 37.9 Gb | 31.6 Tb |
| 🟧 Outgoing US-FNAL | 56.9 Gb | 5.87 Gb | 1.12 Mb | 11.8 Tb |
| 🟥 Outgoing ES-PIC | 12.1 Gb | 3.13 Gb | 0 b | 6.31 Tb |

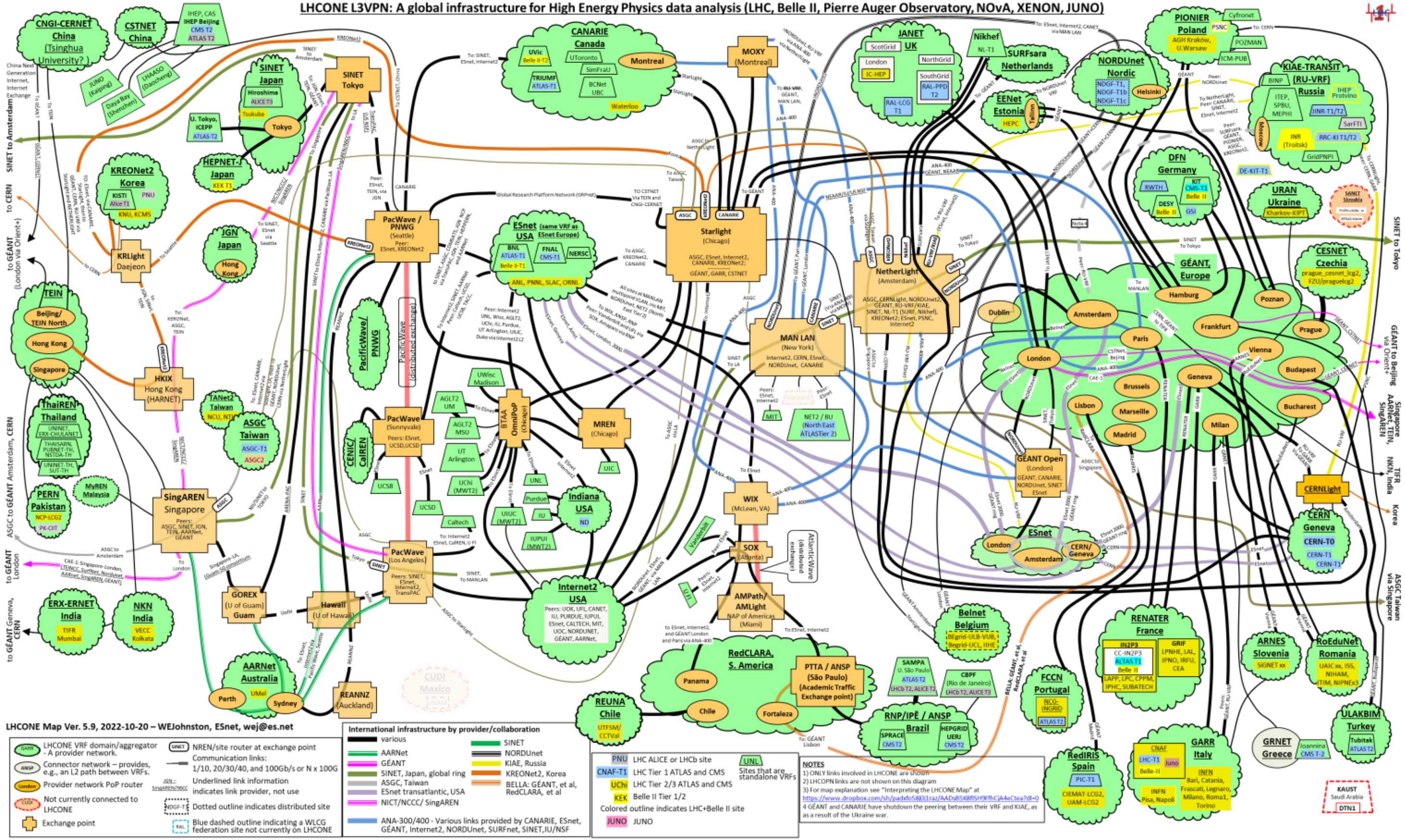


LHCOPN Total Traffic (T1s -> CERN)

| | max | avg | current | total |
|---|---|---|---|---|
| 🟩 Incoming DE-KIT | 145 Gb | 10.9 Gb | 4.16 Gb | 22.0 Tb |
| 🟨 Incoming KR-KISTI | 3.07 Gb | 428 Mb | 10.8 kb | 864 Gb |
| 🟦 Incoming RU-T1 | 40.7 Gb | 9.61 Gb | 3.60 Gb | 19.4 Tb |
| 🟧 Incoming FR-IN2P3 | 84.2 Gb | 6.33 Gb | 34.1 kb | 12.8 Tb |
| 🟥 Incoming NDGF | 79.4 Gb | 7.98 Gb | 16.0 Gb | 16.1 Tb |
| 🟦 Incoming NL-T1 | 96.9 Gb | 5.43 Gb | 10.3 Gb | 10.9 Tb |
| 🟪 Incoming TW-ASGC | 130 b | 76.2 b | 69.7 b | 154 kb |
| 🟪 Incoming IT-INFN-CNAF | 62.7 Gb | 4.94 Gb | 1.64 Gb | 9.95 Tb |
| 🟩 Incoming UK-RAL | 43.0 Gb | 11.5 Gb | 966 Mb | 23.3 Tb |
| 🟨 Incoming CA-TRIUMF | 52.8 Gb | 5.78 Gb | 38.4 Gb | 11.6 Tb |
| 🟦 Incoming US-NBL | 83.0 Gb | 17.1 Gb | 11.8 Gb | 34.5 Tb |
| 🟧 Incoming US-FNAL | 61.0 Gb | 8.18 Gb | 799 b | 16.5 Tb |
| 🟥 Incoming ES-PIC | 11.0 Gb | 3.45 Gb | 0 b | 6.95 Tb |

CERN OpenMonIT LHCOPN, period Oct 7 .. Oct 14 2022, from https://monit-grafana-open.cern.ch/d/HreVOyc7z/all-lhcopn-traffic

# LHCone



LHCONE L3VPN: A global infrastructure for High Energy Physics data analysis (LHC, Belle II, Pierre Auger Observatory, NOvA, XENON, JUNO)

LHCONE Map Ver. 5.9, 2022-10-20 – WEJohnston, ESnet, wej@es.net

- LHCONE VRF domain/aggregator - A provider network.
- Connector network – provides, e.g., an L2 path between VRFs.
- Provider network PoP router
- Not currently connected to LHCONE
- Exchange point
- NREN/site router at exchange point
- Communication links: 1/10, 20/30/40, and 100Gb/s or N x 100G
- Underlined link information indicates link provider, not use
- Dotted outline indicates distributed site
- 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
- ANA-300/400 - Various links provided by CANARIE, ESnet, GÉANT, Internet2, NORDUnet, SURFnet, SINET,IU/NSF
- SINET
- NORDUnet
- KIAE, Russia
- KREONet2, Korea
- BELLA: GÉANT, et al, RedCLARA, et al

| | |
|---|---|
| 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

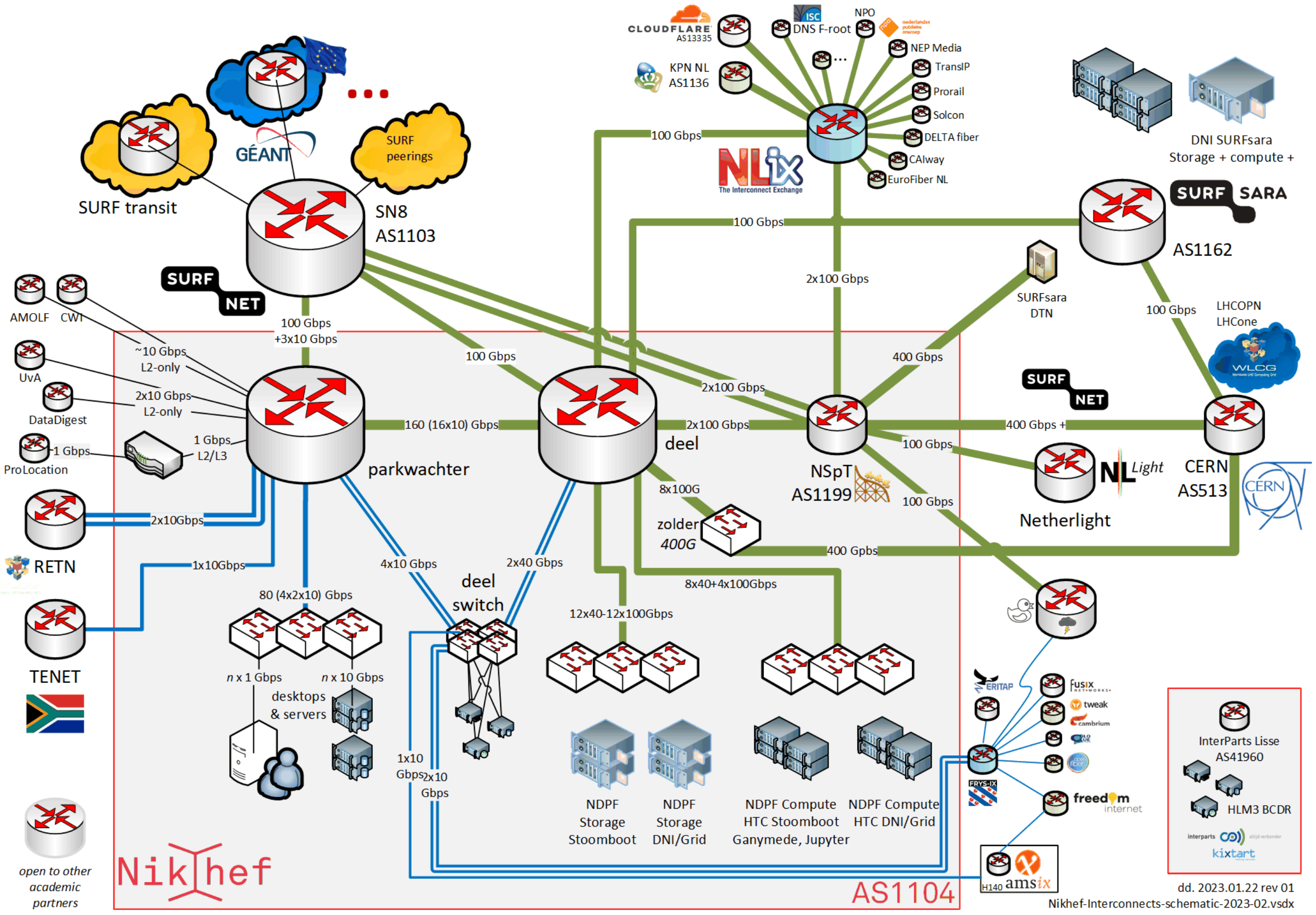
| | |
|---|---|
| JUNO | JUNO |

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/AADs85K8flSH9FfhCjA4eCtea?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

# AS1104

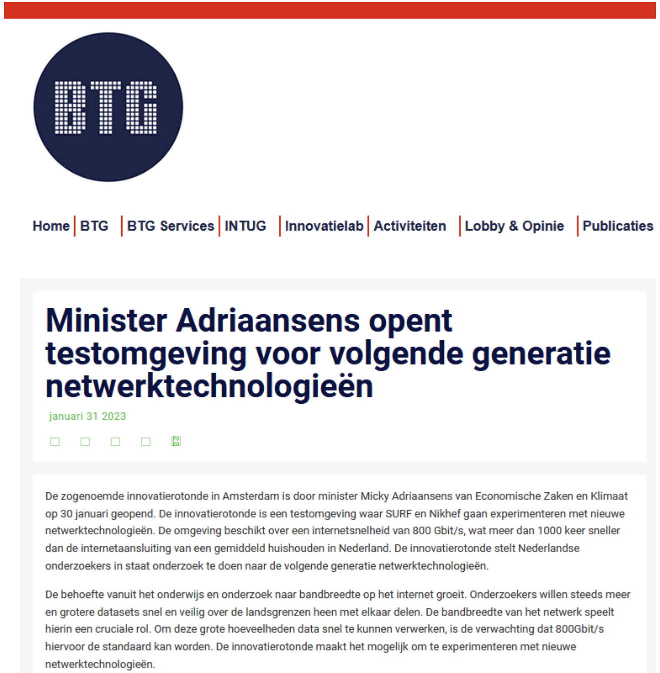


dd. 2023.01.22 rev 01
Nikhef-Interconnects-schematic-2023-02.vsdx

https://www.nikhef.nl/pdp/doc/facility

# For the HL-LHC and SKA, more is needed!

- Core network is now 400G-100G mixed
- Experiments with 800Gbps now ongoing
- local (AMS) has been demonstrated
- next: 400 → 800G AMS-GVA ☺


BTG
Home | BTG | BTG Services | INTUG | Innovatielab | Activiteiten | Lobby & Opinie | Publicaties

## Minister Adriaansens opent testomgeving voor volgende generatie netwerktechnologieën

januari 31 2023

De zogenoemde innovatierotonde in Amsterdam is door minister Micky Adriaansens van Economische Zaken en Klimaat op 30 januari geopend. De innovatierotonde is een testomgeving waar SURF en Nikhef gaan experimenteren met nieuwe netwerktechnologieën. De omgeving beschikt over een internetsnelheid van 800 Gbit/s, wat meer dan 1000 keer sneller dan de internetaansluiting van een gemiddeld huishouden in Nederland. De innovatierotonde stelt Nederlandse onderzoekers in staat onderzoek te doen naar de volgende generatie netwerktechnologieën.

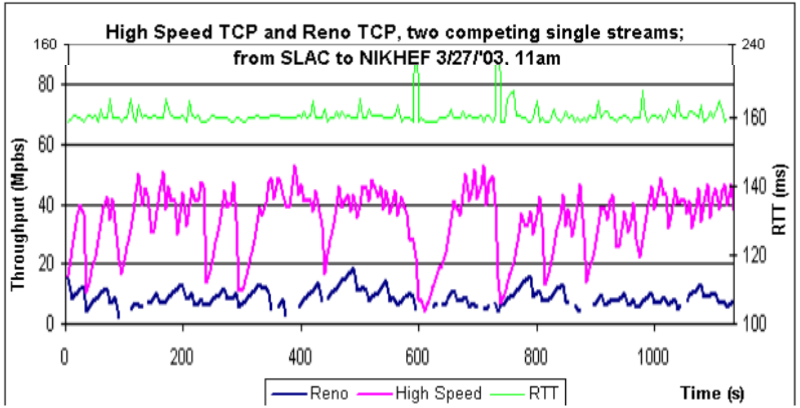
De behoefte vanuit het onderwijs en onderzoek naar bandbreedte op het internet groeit. Onderzoekers willen steeds meer en grotere datasets snel en veilig over de landsgrenzen heen met elkaar delen. De bandbreedte van het netwerk speelt hierin een cruciale rol. Om deze grote hoeveelheden data snel te kunnen verwerken, is de verwachting dat 800Gbit/s hiervoor de standaard kan worden. De innovatierotonde maakt het mogelijk om te experimenteren met nieuwe netwerktechnologieën.

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



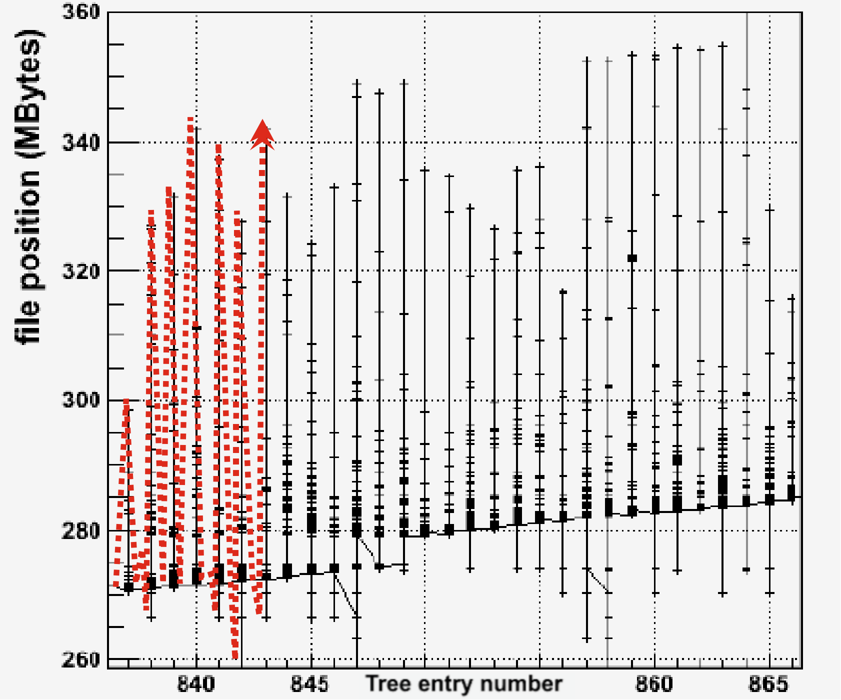
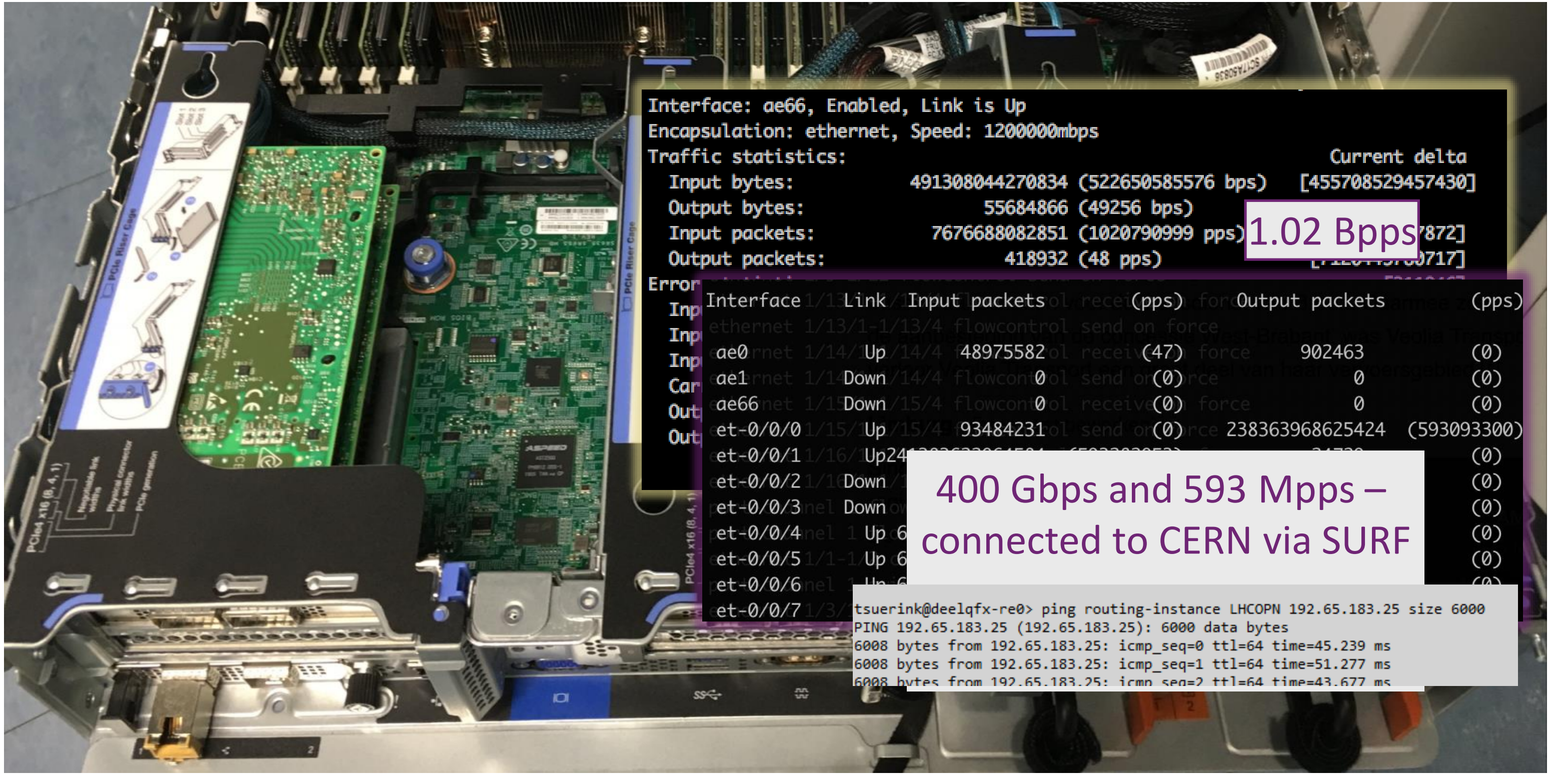


Image of TCP slow-start and packet loss impact (in Mpps): Antony Antony et al., Nikhef, for DataTAG, 2003(!)

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’

# 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

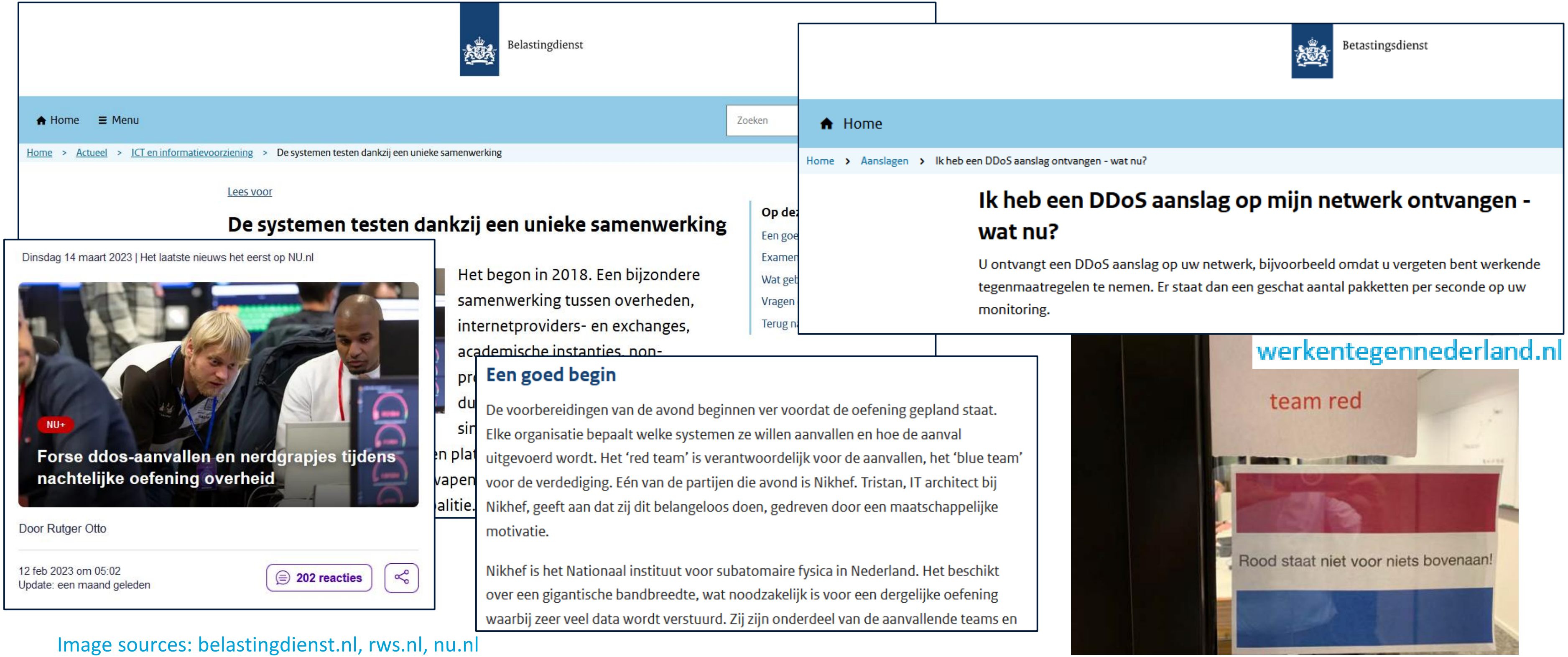
```
Interface   Link  Input packets       (pps)   Output packets        (pps)

 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

# 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
vapen
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

# ‘ScienceDMZ’

Multi-pronged role

- network isolation
- security zoning
- ‘latency hiding’
- caching

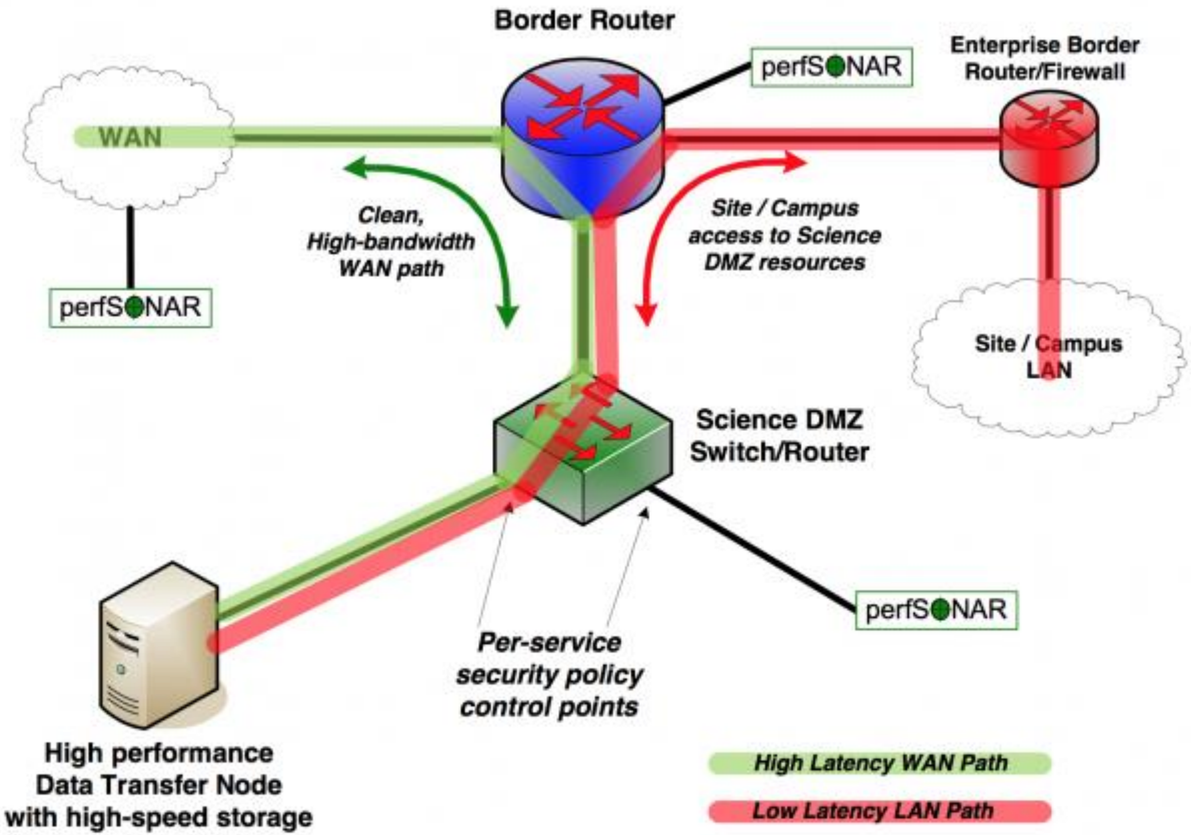
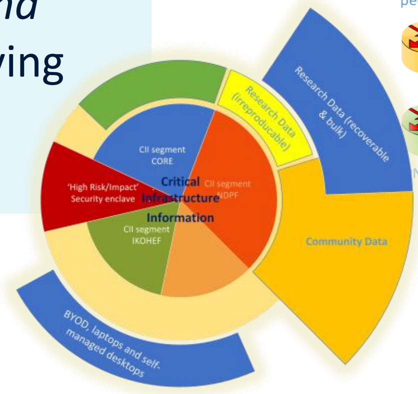

Border Router
perfSONAR
Enterprise Border Router/Firewall
WAN
perfSONAR
Clean, High-bandwidth WAN path
Site / Campus access to Science DMZ resources
Site / Campus LAN
Science DMZ Switch/Router
perfSONAR
Per-service security policy control points
High performance Data Transfer Node with high-speed storage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

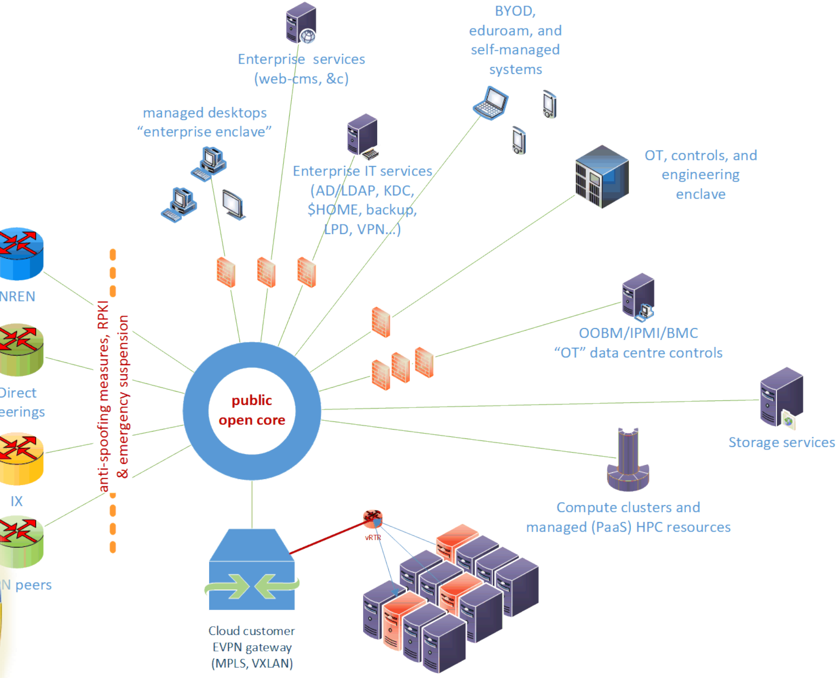
Image and ‘ScienceDMZ’ promulgated by ESnet (see fasterdata.es.net)

# Although for a research organisation

… you want a **science network**
with a ‘back-office enclave’

‘open-core’ research network model implements enclave structure *and* protects against overload by having *no stateful components in the network path*


Research Data (irreproducable)
Research Data (recoverable & bulk)
CII segment CORE
Critical Infrastructure Information
CII segment NDPF
'High Risk/Impact' Security enclave
CII segment IKOHEF
Community Data
BYOD, laptops and self-managed desktops


Enterprise services (web-cms, &c)
BYOD, eduroam, and self-managed systems
managed desktops “enterprise enclave”
Enterprise IT services (AD/LDAP, KDC, $HOME, backup, LPD, VPN…)
OT, controls, and engineering enclave
NREN
Direct peerings
IX
N peers
anti-spoofing measures, RPKI & emergency suspension
public open core
OOBM/IPMI/BMC “OT” data centre controls
Storage services
Compute clusters and managed (PaaS) HPC resources
vRTR
Cloud customer EVPN gateway (MPLS, VXLAN)

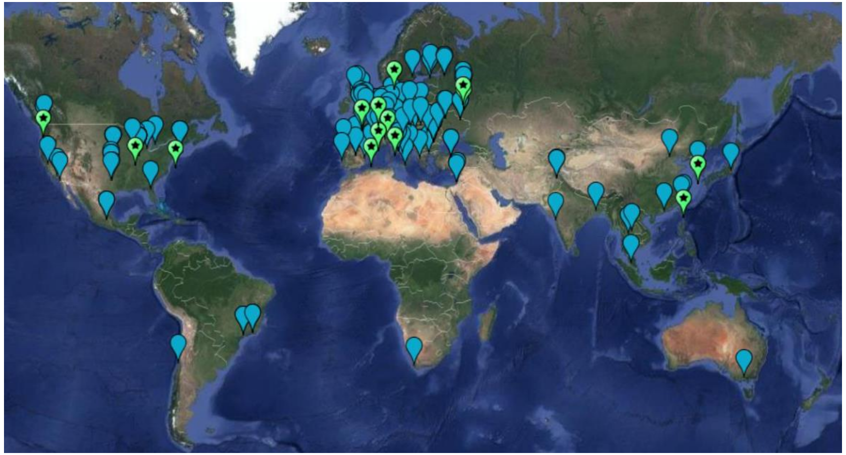
# Trust & Identity

## *Safe access for open data processing*

More than one user, *from*
more than one organizational domain, *in*
more than one country!

# WLCG: when we met a global trust scaling issue

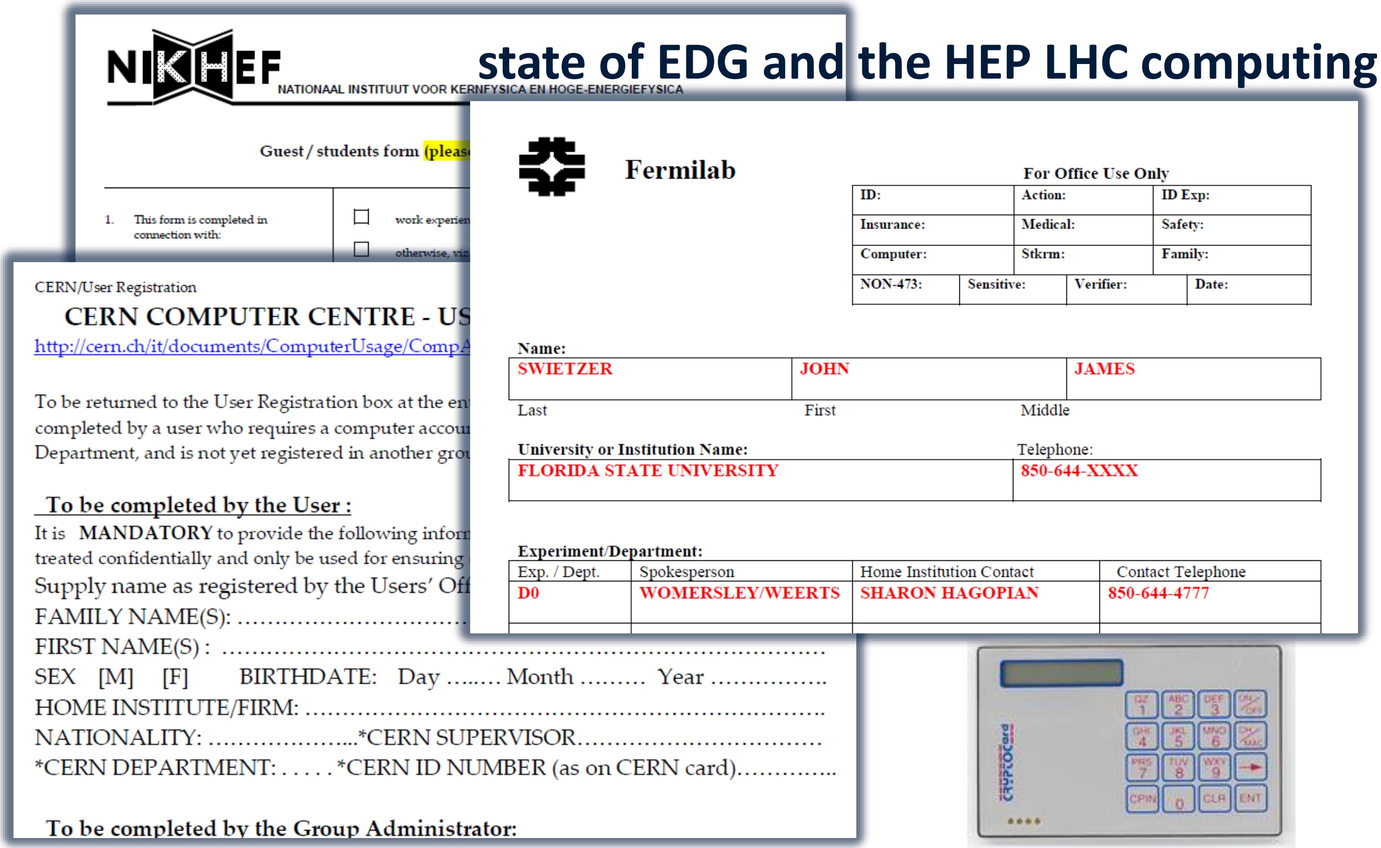
170 sites
~60 countries & regions
~20000 users
just how many interactions ??



people photo: a small part of the CMS collaboration in 2017, Credit: CMS-PHO-PUBLIC-2017-004-3; site map: WLCG sites from Maarten Litmaath (CERN) 2021

# Scaling issues – credentials at each site does not work

**state of EDG and the HEP LHC 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, vi

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 gro

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 |

# Access control in a single domain

- Dedicated to each service where you need access
- Usually strongly linked to authorization: at times even different accounts for different roles
- In a multi-organizational system becomes

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

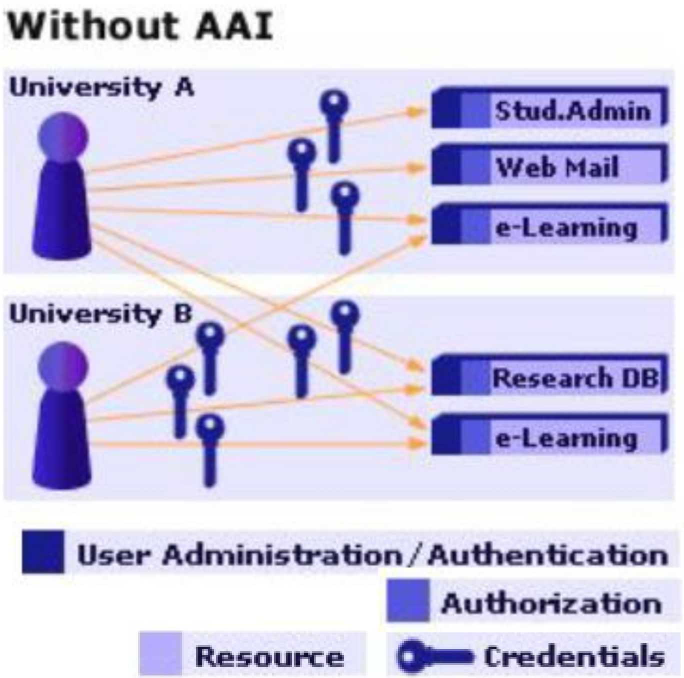


Image: AARC NA2 training module “Authentication and Authorisation 101” - https://aarc-community.org/training/aai-101/

# Authentication and Authorization Infrastructure

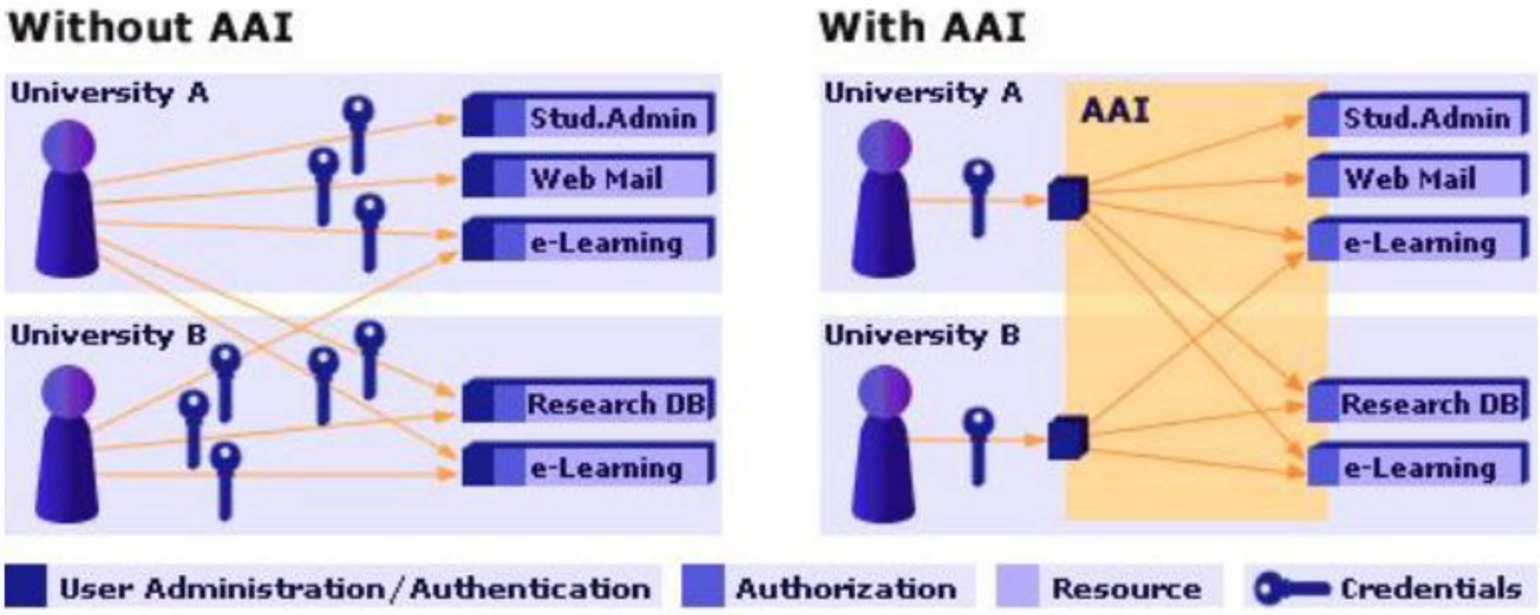
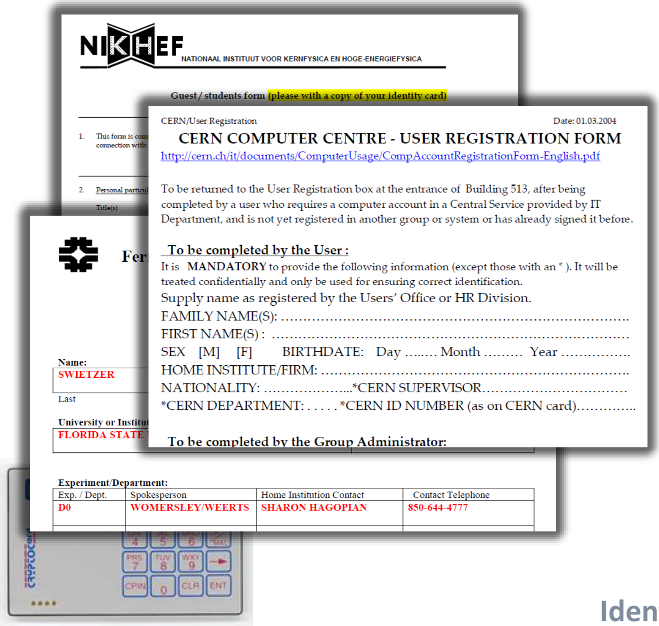


Image: AARC NA2 training module “Authentication and Authorisation 101” - https://aarc-community.org/training/aai-101/

# Whence we came: the long road to federated access

From disparate systems in ~2000



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)

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:**

Fer

**Name:**

SWIETZER

Last

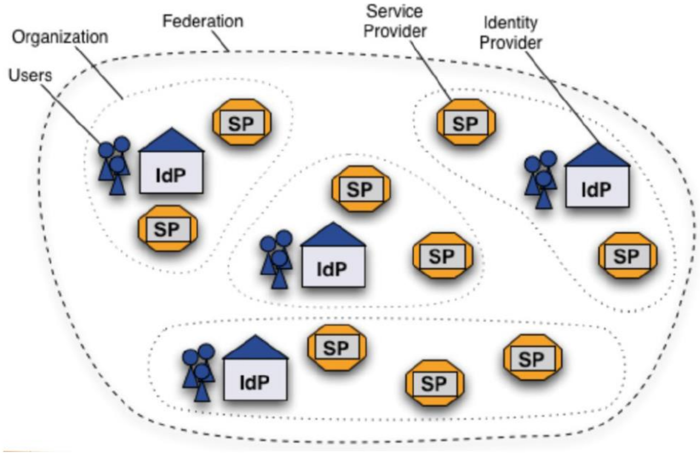
**University or Institut**

FLORIDA STATE

**Experiment/Department:**

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

**AuthN-AuthZ separation fundamental**
to the Federated (R&E) AAI, global IGTF PKI, VOMS, 'AARC BPA' AAI architecture …



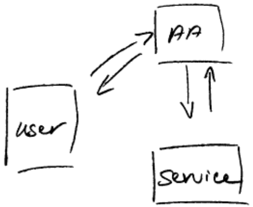
**Identity federation** provides authentication from the home organisation (IdP, "identity provider")
**Service providers** perform authorization, maybe using attributed provided by the IdP

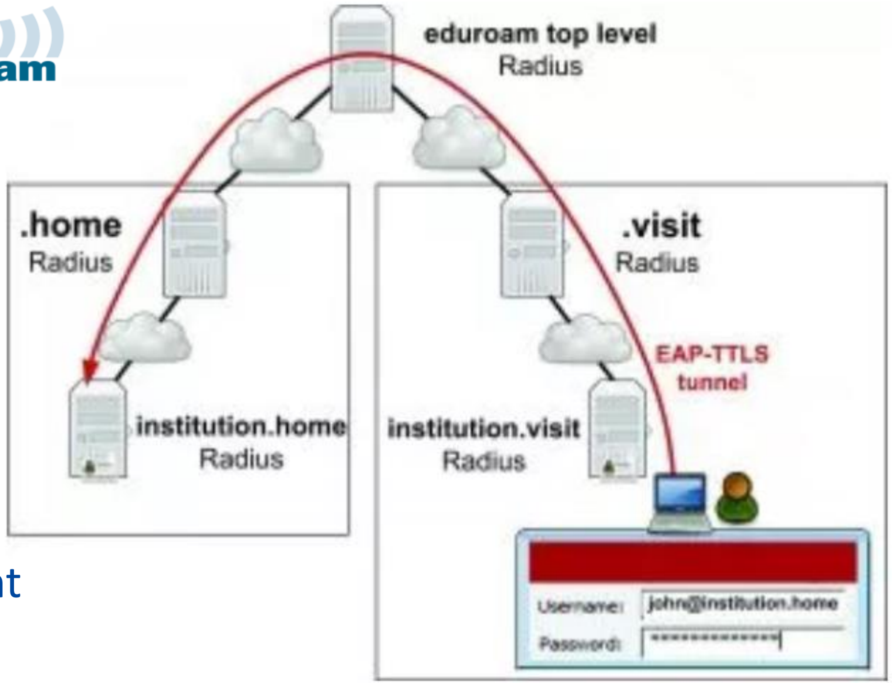
Shibboleth IdP image: SWITCH (CH)

# One simple federation you know: eduroam

*service-specific* trust
between organisations
globally

hierarchical RADIUS servers based
an 802.1x secure exchange
over TLS or EAP-TTLS
tunneling your credentials
back to your home institution





Username: john@institution.home

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

RADIUS server then instructs WiFi access point

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

# Federation: different technologies, same idea

**SAML - Security Assertion Markup Language and WebSSO ('SAML2Int')**

- XML-formatted 'attribute statements' over web transport (usually POST)
- SAML-Metadata: list of entities with description of bindings with entityAttributes

**PKI - Public Key Infrastructures**

- certification authority (CA) signing X.509 formatted certificates
  with name, issuer, serial number, and extensions
- CAs can sign end-entities as well as other CAs (hierarchically or by cross-signing)
- bridge CAs render a technical implementation of a shared policy (assurance)
- policy-bridges don't sign anything, but curate *distribution* (like browsers and operating systems based on CA/BF requirements, or the IGTF for research infras)

**OIDC Fed - OpenID Connect Federation**

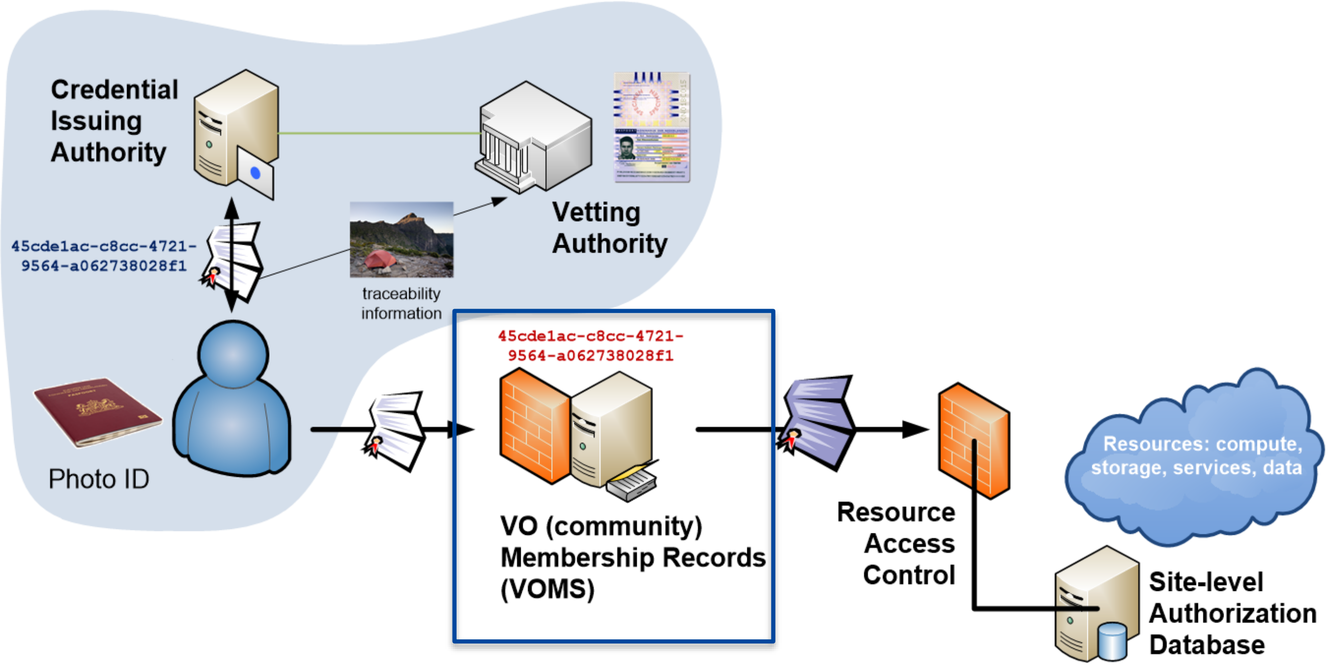
- for end-points for OIDC Providers and Relying Parties – otherwise quite similar

*federation based on 'ultimate trust' domains (e.g. cross-realm Kerberos) also exists, but …*

See www.oasis.org for SAML, RFC5280 (tech) & RFC3247 (policy) for PKIX, https://igtf.net/ and https://cabforum.org;
OpenID Connect Federation: https://openid.net/specs/openid-connect-federation-1_0.html

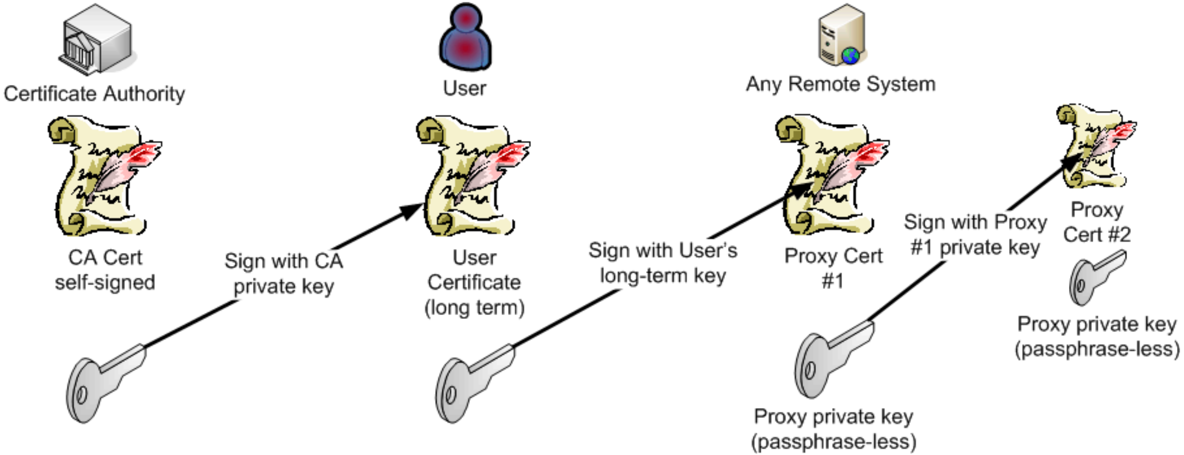
# Identity federations give … identity (“AuthN”)

Authorization (what may you do) still needs to be added to the mix


Credential Issuing Authority
Vetting Authority
45cde1ac-c8cc-4721-9564-a062738028f1
traceability information
Photo ID
45cde1ac-c8cc-4721-9564-a062738028f1
VO (community) Membership Records (VOMS)
Resource Access Control
Resources: compute, storage, services, data
Site-level Authorization Database

# PKIX certificates using proxies also for *non-web access*

- Certificates are ASN.1 structures with (issuer, subject, serial) + extensions
- The digest (hash) signed with the private key of the issuer
- Verifiable using the issuer’s public key



RFC3820 ‘proxy’ certificates extend this concept to (restricted) identity delegation

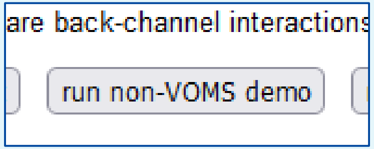
To get an RFC3820 proxy certificate using your own federated identity, use RCauth.eu – see https://rcdemo.nikhef.nl/ and use the “Basic Demo” option

# An X.509 RFC5280 Certificate (textually)

```
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
```

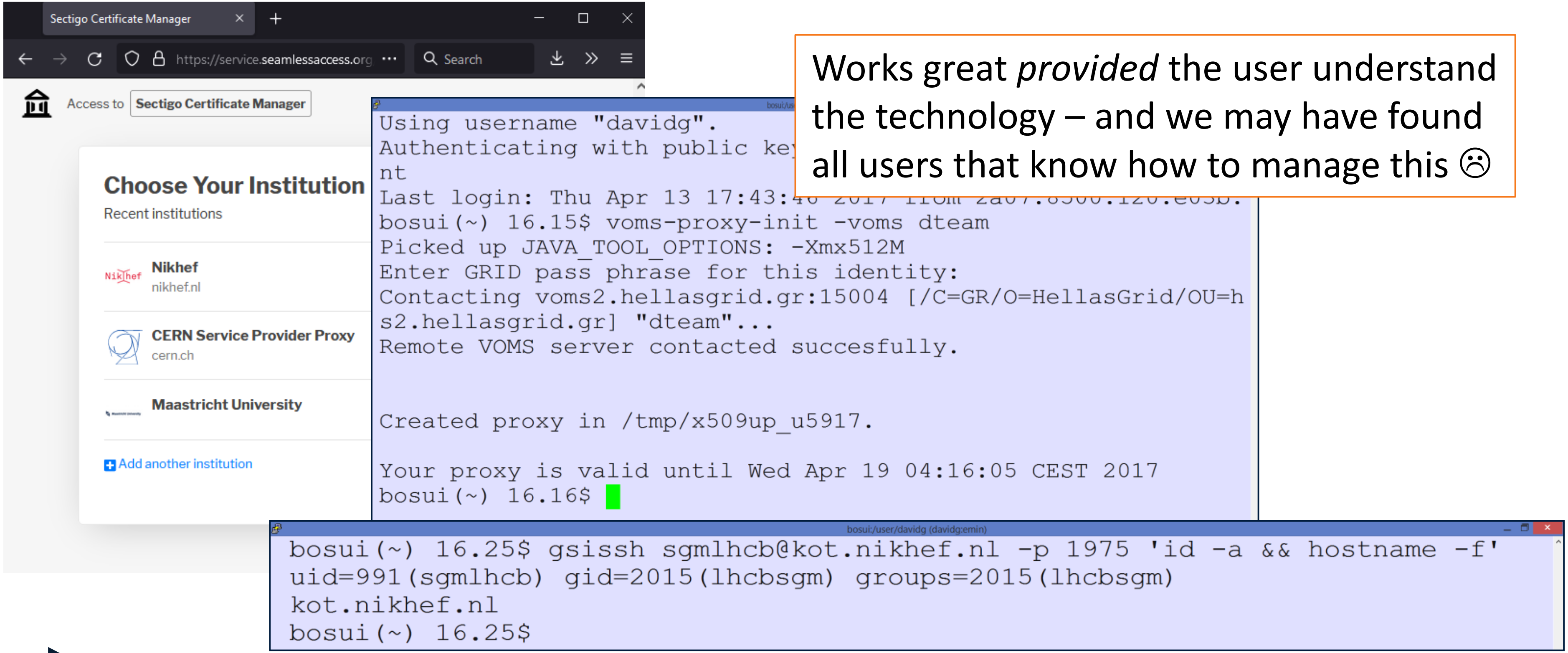
You should be able to get a 'DOGWOOD' assurance certificate from RCauth.eu right now:

- go to https://rcdemo.nikhef.nl/
- select the 'Basic demo'
- use 'run non-VOMS' to get and view your short-lived certificate



# Seamless (eduGAIN) Access to (non-Web) Resources using PKIX?

## Traditional workflow – using a client-held credential



Sectigo Certificate Manager
https://service.seamlessaccess.org
Search
Access to Sectigo Certificate Manager
Choose Your Institution
Recent institutions
Nikhef
nikhef.nl
CERN Service Provider Proxy
cern.ch
Maastricht University
Add another institution

```
Using username "davidg".
Authenticating with public key
nt
Last login: Thu Apr 13 17:43:46 2017 from 2a07:8500:120:e03b:
bosui(~) 16.15$ voms-proxy-init -voms dteam
Picked up JAVA_TOOL_OPTIONS: -Xmx512M
Enter GRID pass phrase for this identity:
Contacting voms2.hellasgrid.gr:15004 [/C=GR/O=HellasGrid/OU=h
s2.hellasgrid.gr] "dteam"...
Remote VOMS server contacted succesfully.


Created proxy in /tmp/x509up_u5917.

Your proxy is valid until Wed Apr 19 04:16:05 CEST 2017
bosui(~) 16.16$
```

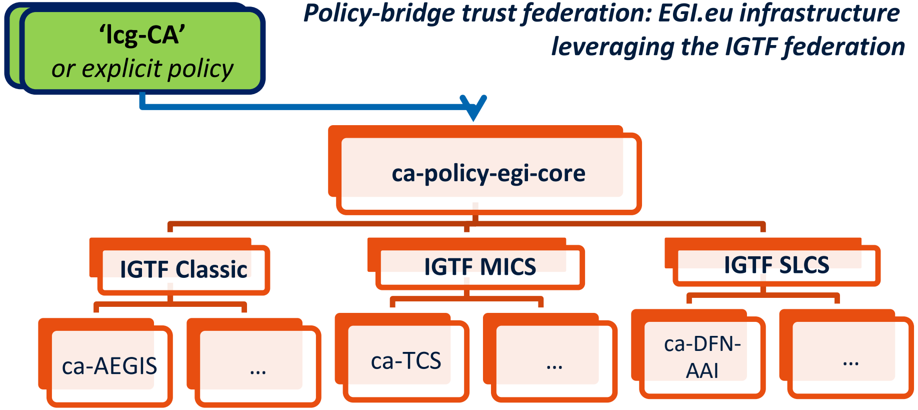
```
bosui(~) 16.25$ gsissh sgmlhcb@kot.nikhef.nl -p 1975 'id -a && hostname -f'
uid=991(sgmlhcb) gid=2015(lhcbsgm) groups=2015(lhcbsgm)
kot.nikhef.nl
bosui(~) 16.25$
```

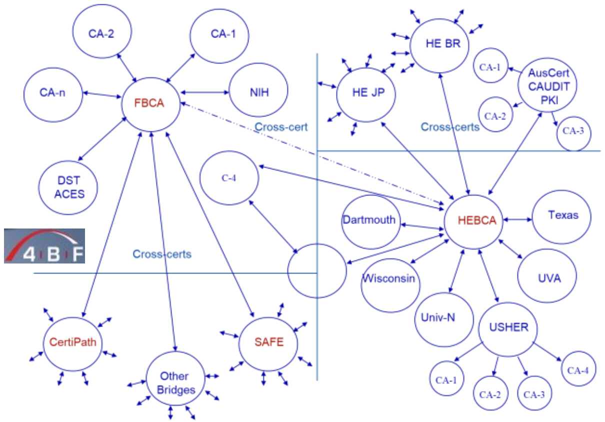
Works great *provided* the user understand the technology – and we may have found all users that know how to manage this ☹

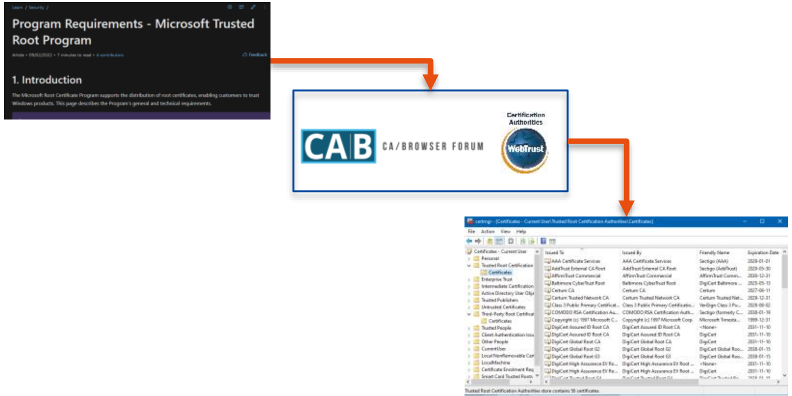
# PKIX federation

trust remains with the relying party can be *bridged* by either cross-signing (left) or by policy agreements (right)

*Policy-bridge trust federation: EGI.eu infrastructure leveraging the IGTF federation*







Learn / Security /

Program Requirements - Microsoft Trusted Root Program

Article • 09/02/2022 • 7 minutes to read • 4 contributors

Feedback

1. Introduction

The Microsoft Root Certificate Program supports the distribution of root certificates, enabling customers to trust Windows products. This page describes the Program's general and technical requirements.

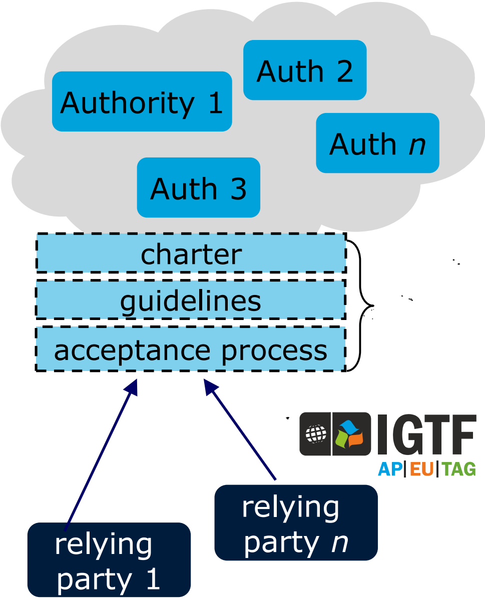
certmgr - [Certificates - Current User\Trusted Root Certification Authorities\Certificates]

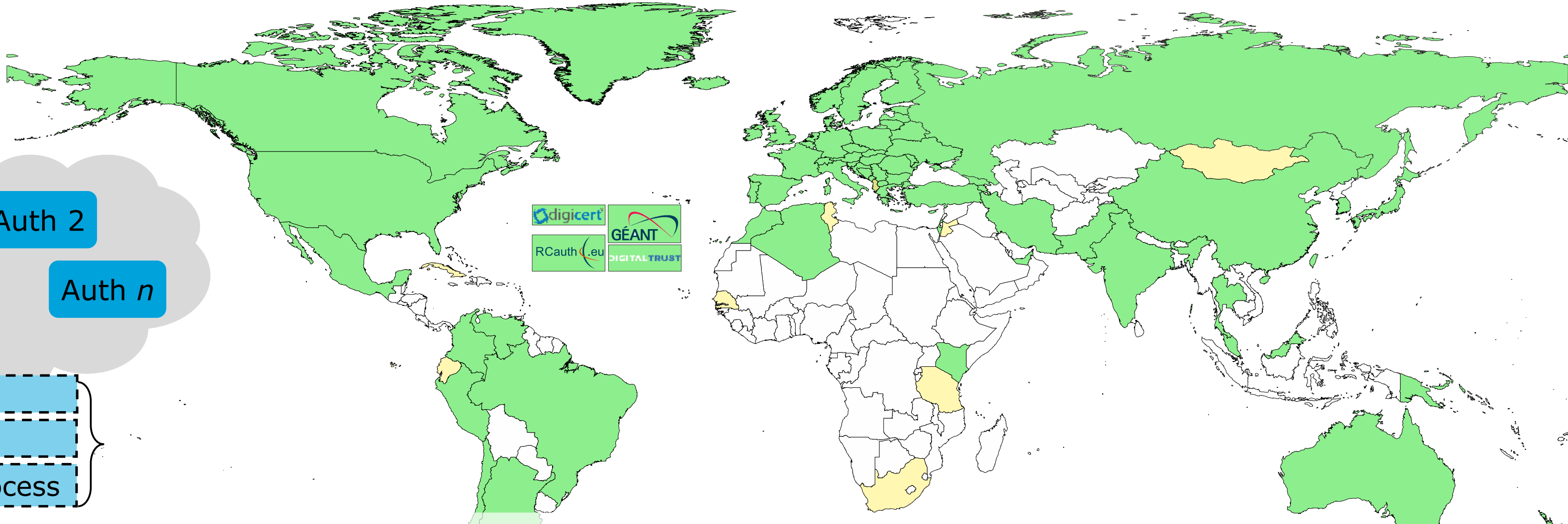
| Issued To | Issued By | Friendly Name | Expiration Date |
|---|---|---|---|
| AAA Certificate Services | AAA Certificate Services | Sectigo (AAA) | 2028-01-01 |
| AddTrust External CA Root | AddTrust External CA Root | Sectigo (AddTrust) | 2020-05-30 |
| AffirmTrust Commercial | AffirmTrust Commercial | AffirmTrust Comm... | 2030-12-31 |
| Baltimore CyberTrust Root | Baltimore CyberTrust Root | DigiCert Baltimore ... | 2025-05-13 |
| Certum CA | Certum CA | Certum | 2027-06-11 |
| Certum Trusted Network CA | Certum Trusted Network CA | Certum Trusted Net... | 2029-12-31 |
| Class 3 Public Primary Certificat... | Class 3 Public Primary Certificatio... | VeriSign Class 3 Pu... | 2028-08-02 |
| COMODO RSA Certification Au... | COMODO RSA Certification Auth... | Sectigo (formerly C... | 2038-01-19 |
| Copyright (c) 1997 Microsoft C... | Copyright (c) 1997 Microsoft Corp. | Microsoft Timesta... | 1999-12-31 |
| DigiCert Assured ID Root CA | DigiCert Assured ID Root CA | <None> | 2031-11-10 |
| DigiCert Assured ID Root CA | DigiCert Assured ID Root CA | DigiCert | 2031-11-10 |
| DigiCert Global Root CA | DigiCert Global Root CA | DigiCert | 2031-11-10 |
| DigiCert Global Root G2 | DigiCert Global Root G2 | DigiCert Global Roo... | 2038-01-15 |
| DigiCert Global Root G3 | DigiCert Global Root G3 | DigiCert Global Roo... | 2038-01-15 |
| DigiCert High Assurance EV Ro... | DigiCert High Assurance EV Root ... | <None> | 2031-11-10 |
| DigiCert High Assurance EV Ro... | DigiCert High Assurance EV Root ... | DigiCert | 2031-11-10 |

Trusted Root Certification Authorities store contains 56 certificates.

Left-hand image: 4 Bridges Forum, source: Scott Rea (then: Dartmouth)
Images: cabforum.org, WebTrust logo: from DigiCert.com; image MS root store, https://learn.microsoft.com/en-us/security/trusted-root/program-requirements

# Policy-bridged global federations for research computing





**3 regional IGTF chapters: EMEA, Americas, Asia Pacific**
~ 90 Identity Providers (some leveraging a R&E federation)
~ 10 international major relying parties
~ 60 countries / economic areas / international treaty orgs
> 1000 relying service provider collaborations

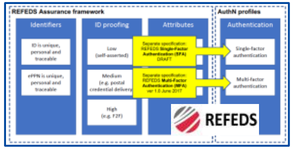
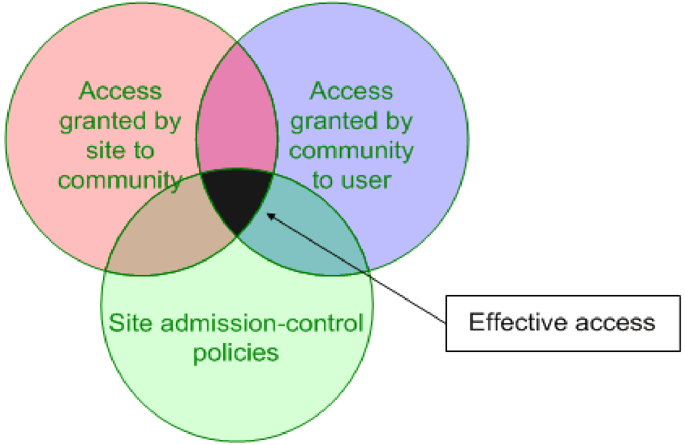

REFEDS Assurance framework
AuthN profiles
Identifiers
ID proofing
Attributes
Authentication
ID is unique, personal and traceable
ePPN is unique, personal and traceable
Low (self-asserted)
Medium (e.g. postal credential delivery)
High (e.g. F2F)
Separate specification: REFEDS Single-Factor Authentication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

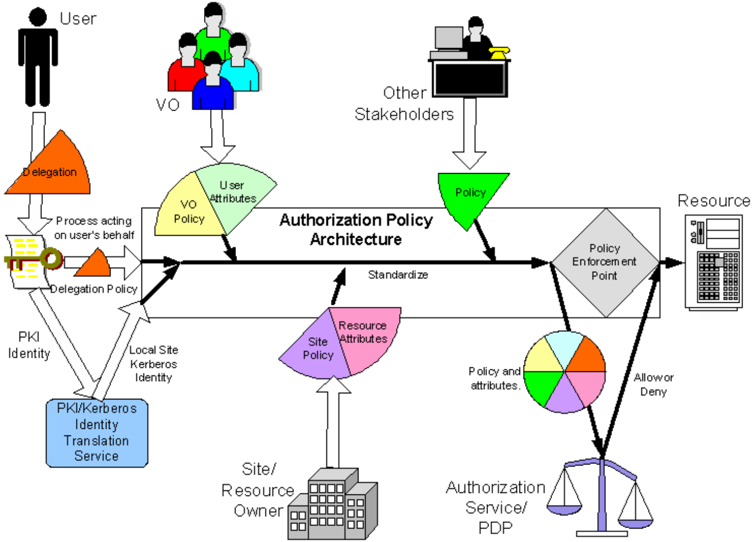
Image: Interoperable Global Trust Federation IGTF, https://igtf.net/; REFEDS Assurance Framework RAF: http://refeds.org/assurance, https://refeds.org/profile/mfa

# Separating authN and authZ for access control

Access control ultimately enforced by service provider
(unless data-level encryption is used)

but role/attribute-based access control needs… attributes

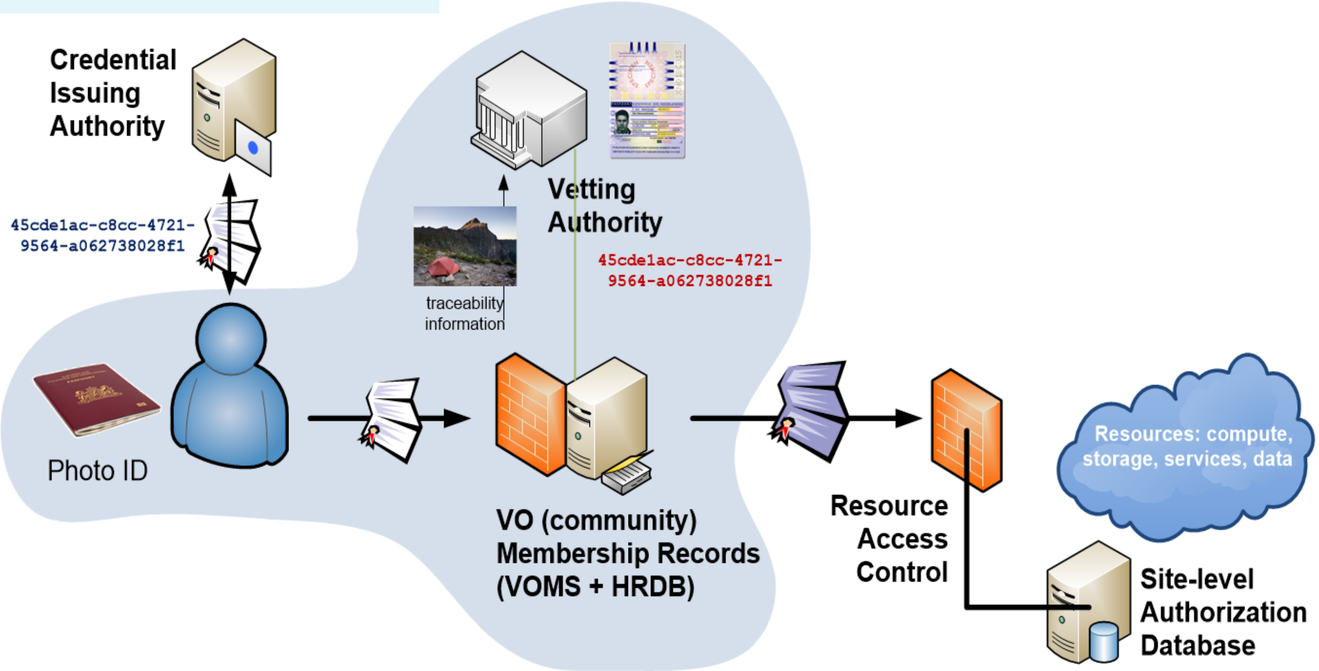

Access granted by site to community
Access granted by community to user
Site admission-control policies
Effective access



policy overlap diagram by Olle Mulmo, KTH for EGEE-I JRA3, ‘policy pie’ from: OpenGrid Forum OGSA working group and Globus Alliance

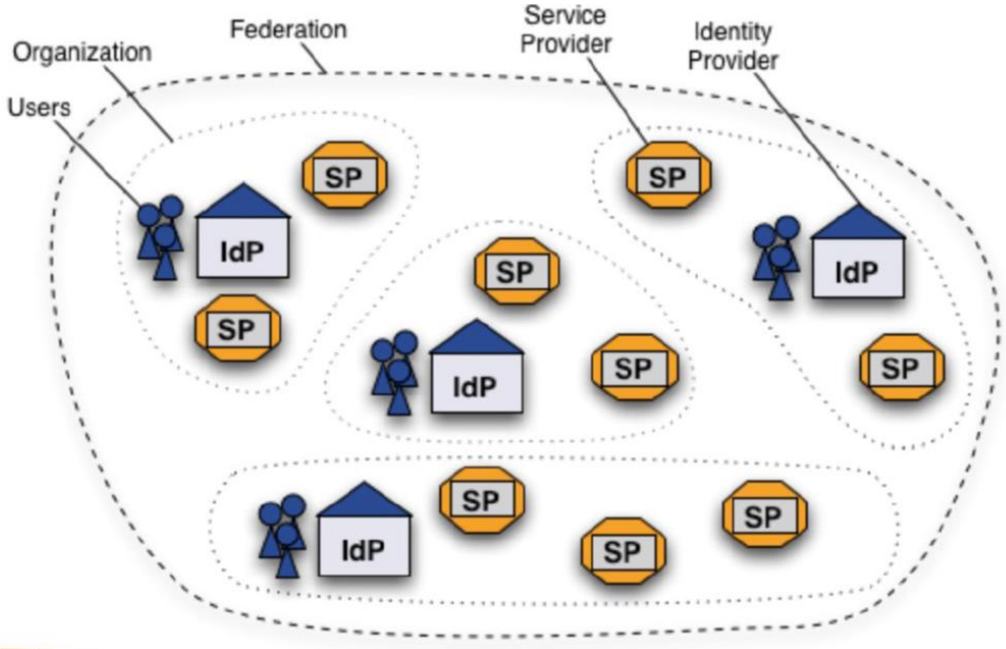
# Separating source of authenticator and identity

'Identifier Only Trust Assurance', i.e. *IOTA Certification Authorities*


Credential Issuing Authority
45cdelac-c8cc-4721-9564-a062738028f1
Vetting Authority
45cdelac-c8cc-4721-9564-a062738028f1
traceability information
Photo ID
VO (community) Membership Records (VOMS + HRDB)
Resource Access Control
Resources: compute, storage, services, data
Site-level Authorization Database

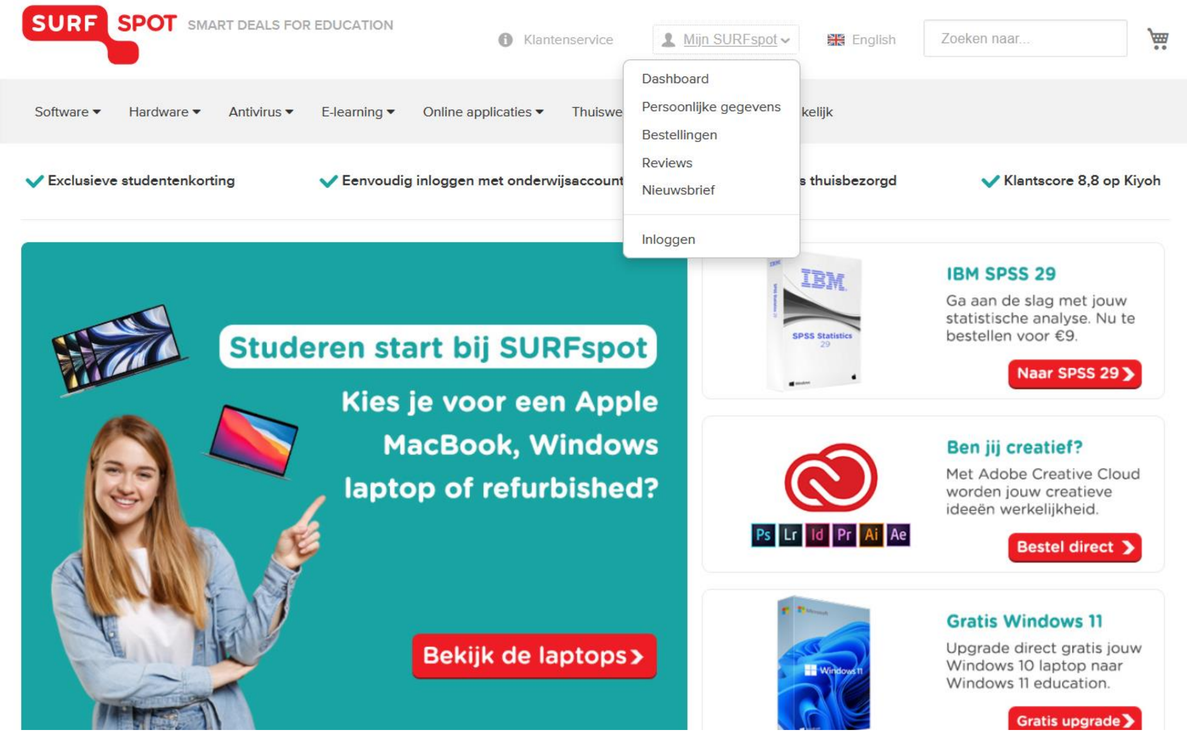
# SAML Federation

portability of identity information across otherwise autonomous administrative domains



Shibboleth IdP image and SAML2 auth flow by SWITCH (CH) – see also https://refeds.org/ on federation structure and (assurance and security) guidelines

# Your favourite federated service?



✓ Exclusieve studentenkorting ✓ Eenvoudig inloggen met onderwijsaccount s thuisbezorgd ✓ Klantscore 8,8 op Kiyoh

Studeren start bij SURFspot

Kies je voor een Apple MacBook, Windows laptop of refurbished?

**IBM SPSS 29**

Ga aan de slag met jouw statistische analyse. Nu te bestellen voor €9.

**Ben jij creatief?**

Met Adobe Creative Cloud worden jouw creatieve ideeën werkelijkheid.

**Gratis Windows 11**

Upgrade direct gratis jouw Windows 10 laptop naar Windows 11 education.

https://surfspot.nl/

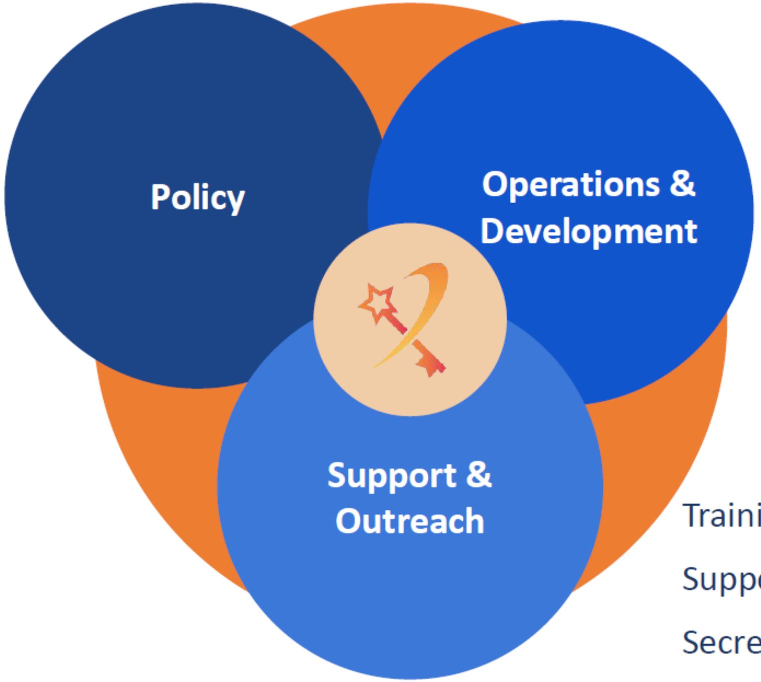
# We live in a federated world!

The UK Access Management Federation
FOR EDUCATION AND RESEARCH
A pilot project for WLCG
eduGAIN
IdP
CERN SSO
STS
MYPROXY
IOTA CA
VOMS
WLCG portal
X 509 VOMS
Normal X.509 service
Web
Identity Federation with the CERN SSO
PHASE 2
SWITCHaai
IGTF
grnet
-AAI
SIFULAN
IGTF
Interoperable Global Trust Federation
AP|EU|TAG
confia
AUSTRALIAN ACCESS FEDERATIO
SIR
Edugate
Gaining Access to Education
SURF
CONEXT
RCTSaai
REDE CIÊNCIA, TECNOLOGIA E SOCIEDADE
idem
aai
eduGAIN
RENATER
CONNECTEUR DE SAVOIRS
Colfire
Colombian Federation of Identity for Research and Education renata
GakuNin
FEIDE
TAAT
Comunidad Federada REUNA
CILogon
uakiri
NEW ZEALAND ACCESS FEDERATION
AAI @EduHr
arnes
eduID
SWAMID

slide inspiration: Licia Florio, NORDUNET

# eduGAIN

Implementation of eduGAIN
Future WG recommendations



Ongoing operations

New eduGAIN-OT

Evolution and duplication of core infrastructure

Training, Webinars , T&I town hall

Support and CSIRT

Secretariat, Business development

**78**

Identity Federations

**5100+**

Identity Providers

**3600+**

Service Providers

slide by: Maarten Kremers, SURF, for the GEANT 5-1 project

# But just identity federation with your home organisation is not enough

- Access services using **identities from their Home Organizations.**
- **Access** services **based on role(s)** users have in the **collaboration**. This info is not known to IdPs/eduGAIN.
- Secure integration of **guest identity solutions** and **support for stronger authentication** mechanisms.
- Requirement for **one persistent identity** across all the community's services when needed and **account linking**.
- **Web** and **non-web** resources
- **Hide complexity** of multiple IdPs/feds/At Auth/ technologies.

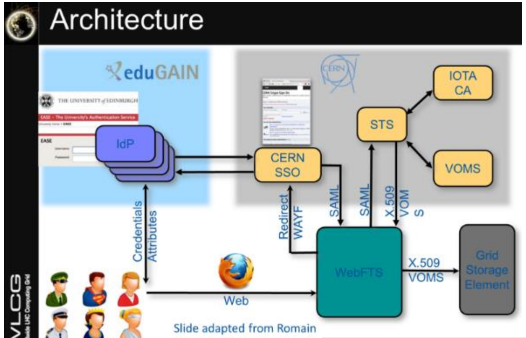
slide design: Licia Florio, NORDUNET

# Federation in research and e-Infrastructures: command-line and brokered access

For ‘CLI-based’ access and brokering (workflow management) for non-web services X.509 technology and `RFC3820 proxies’ are great … but end-user PKI is relatively complex:

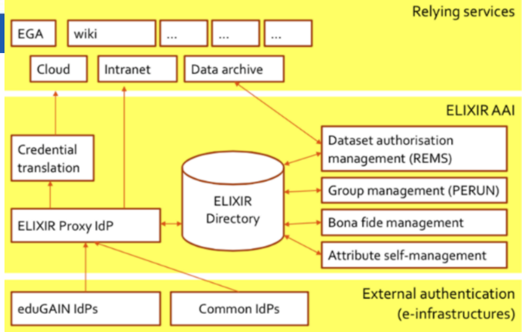
- Infrastructures move to hiding PKIX from the end-user and move to OIDC and Tokens
  - Fewer credentials to manage, appearing ‘simpler’ to the user
- Bridging and translation is a pragmatic approach for cases where PKIX worked better
  - Does not require major technical changes in existing R&E federations
  - Allows for community-centric identities-of-last-resort (or first resort, for that matter…)
  - Allows time for introduction of other technologies, such as OIDC and OAuth2 tokens
- Token translation in many infrastructures that use CLI or brokerage
  - Project MinE, EGI, …
  - translation in any way: SAML→OIDC, SAML→X509, X509→OIDC, X509→SAML, OIDC→X509

# Managing complexities of distributed identity sources

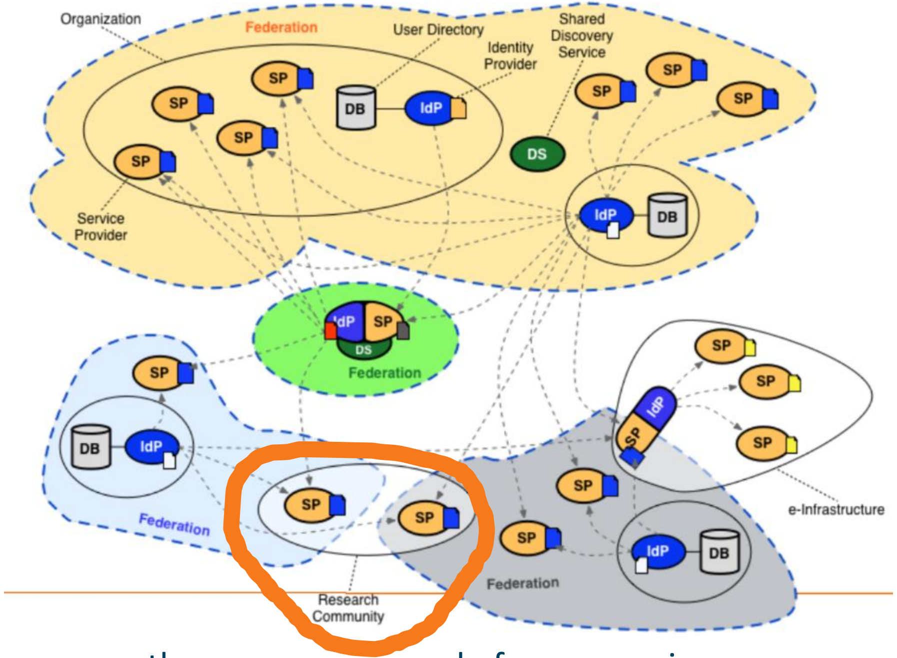


*WebFTS prototype*
*'FIM4R' in wLCG*
*Romain Wartel et al.*

*ELIXIR reference architecture*
*Mikael Linden et al.*



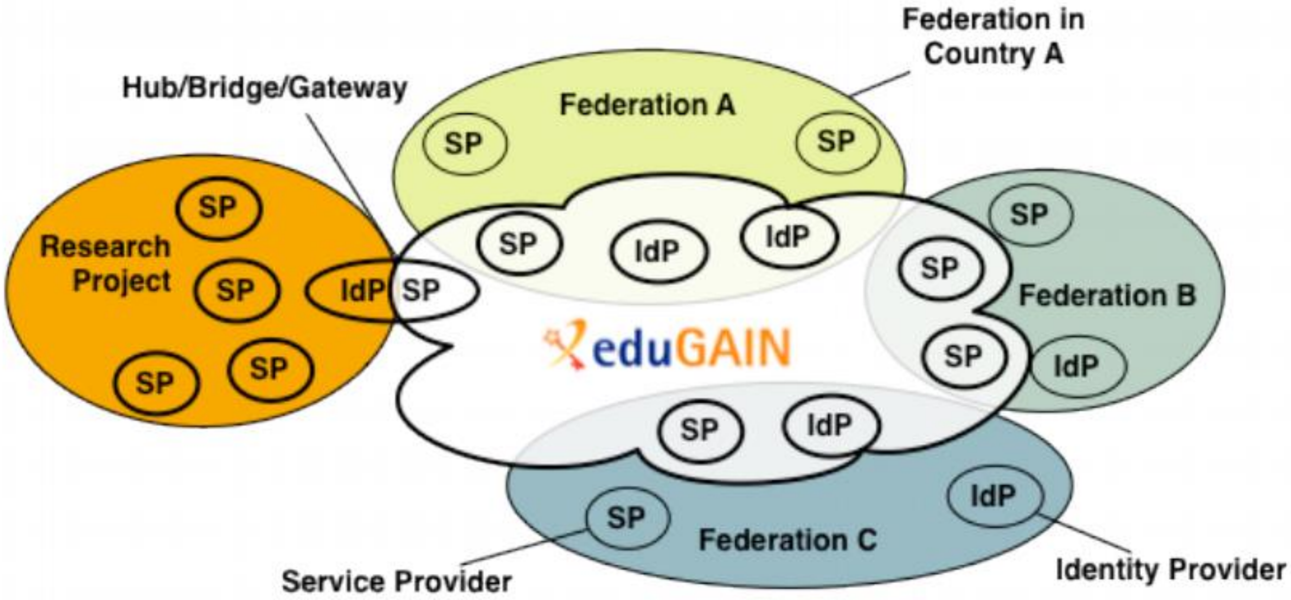
communities had either invented
their own 'proxy' model to abstract complexity



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

Community images: Romain Wartel, CERN; Mikael Linden, CSC; Lukas Hammerle, SWITCH

# Proxies



Federation with SP Proxy image by: SWITCH (CH)

# Federated access for research collaboration – AARC

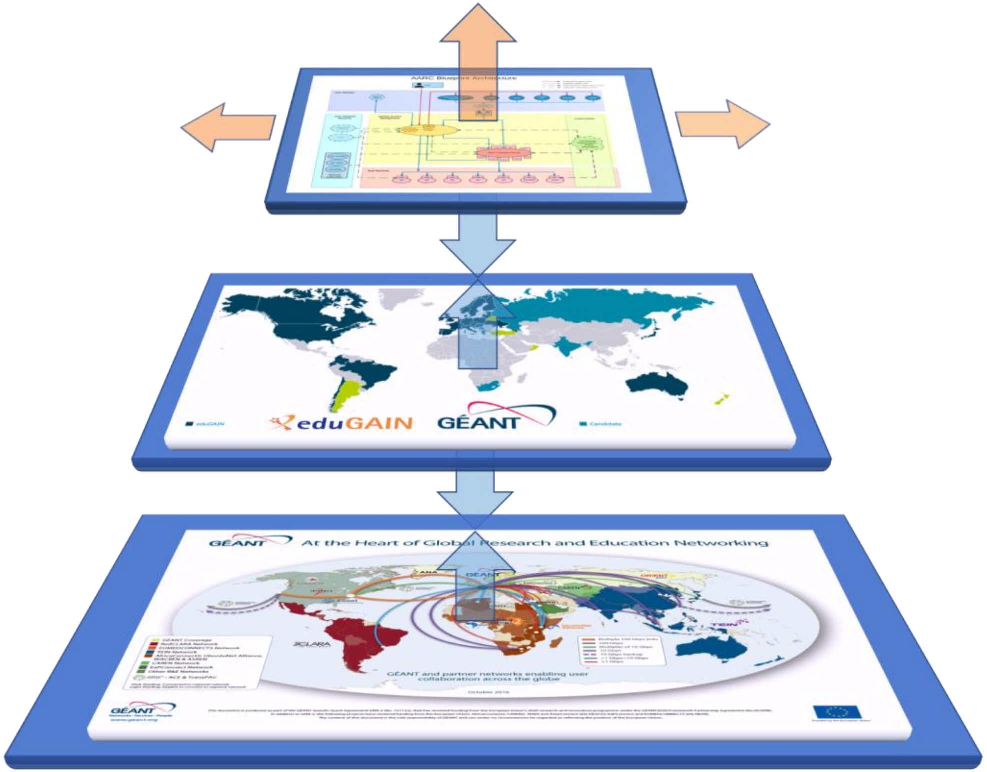
**Authentication and Authorization architecture for Research Collaboration**

*A set of building blocks on top of eduGAIN for international Research Collaboration*

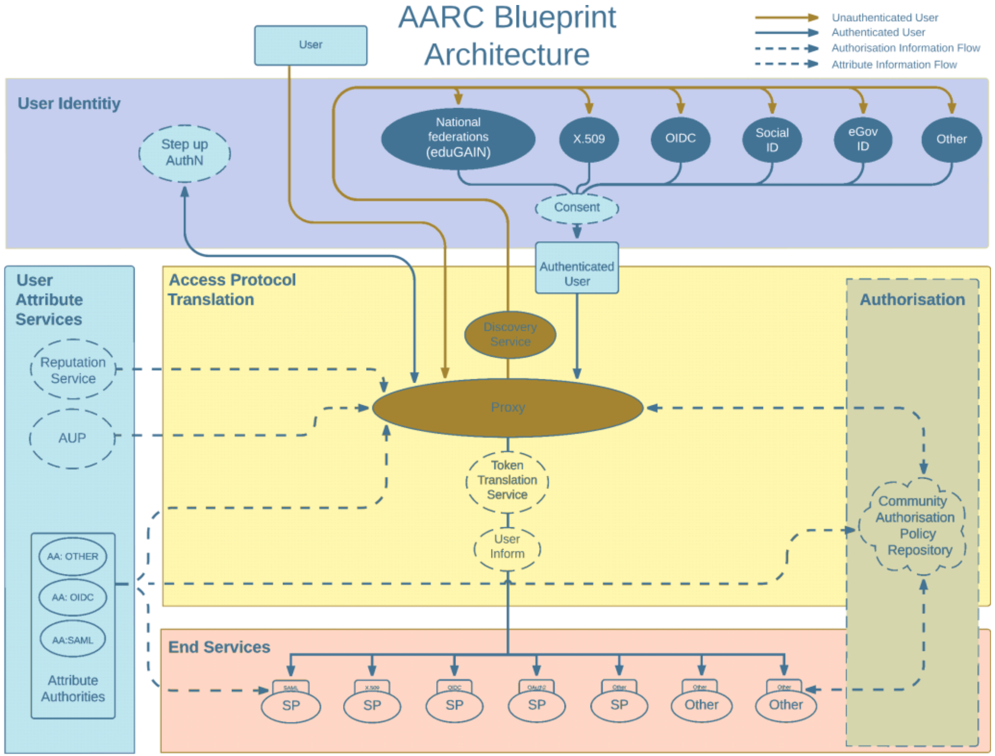
**eduGAIN and the Identity Federations**

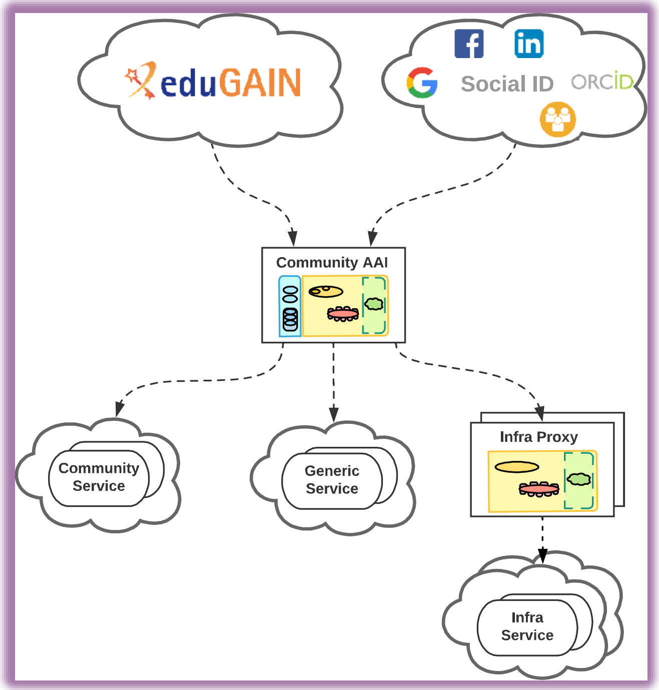
*Foundational federated access in R&E*

**Network connectivity**


AARC Blueprint Architecture
eduGAIN
GÉANT
GÉANT
At the Heart of Global Research and Education Networking
RedCLARA
GÉANT and partner networks enabling user collaboration across the globe
October 2016
GÉANT

# Most trust flows from the (research) community





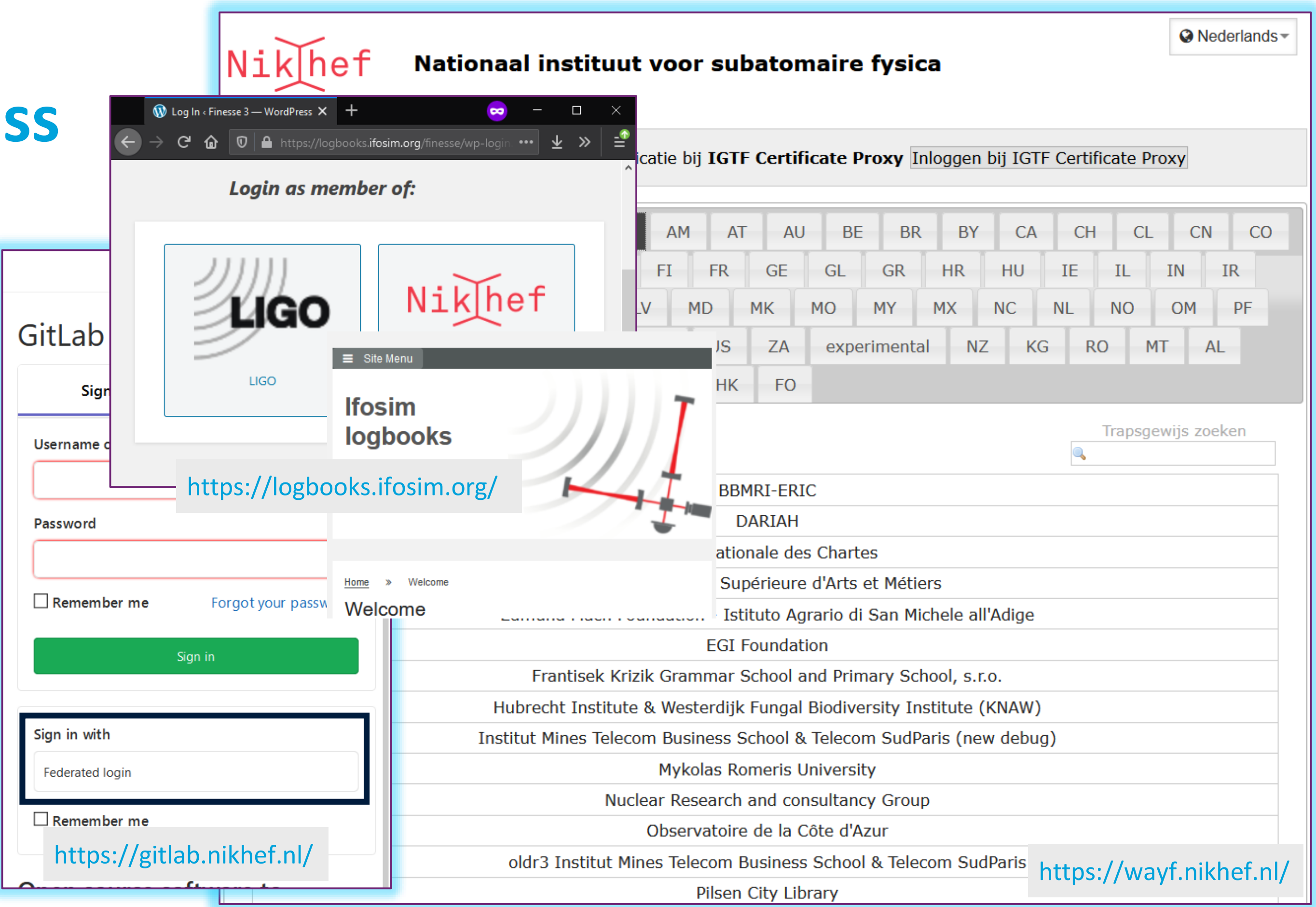
AARC Blueprint Architecture (2019) AARC-G045 https://aarc-community.org/guidelines/aarc-g045/; stacked proxies: EOSC AAI Architecture
EOSC Authentication and Authorization Infrastructure (AAI), ISBN 978-92-76-28113-9, http://doi.org/10.2777/8702

# Federated Access

Login via the
Nikhef service proxy
to gitlab, ifosim.org, …

*“Where are you from”*

discovery screen showing entities from the eduGAIN global interfederation



BBMRI-ERIC

DARIAH

…ationale des Chartes

…Supérieure d'Arts et Métiers

…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

Ifosim logbooks

Welcome

https://logbooks.ifosim.org/

https://gitlab.nikhef.nl/

https://wayf.nikhef.nl/

eduGAIN

ifosim federated AAI integration implementation by Mischa Sallé; per-country WAYF selection is a bespoke Nikhef WAYF feature

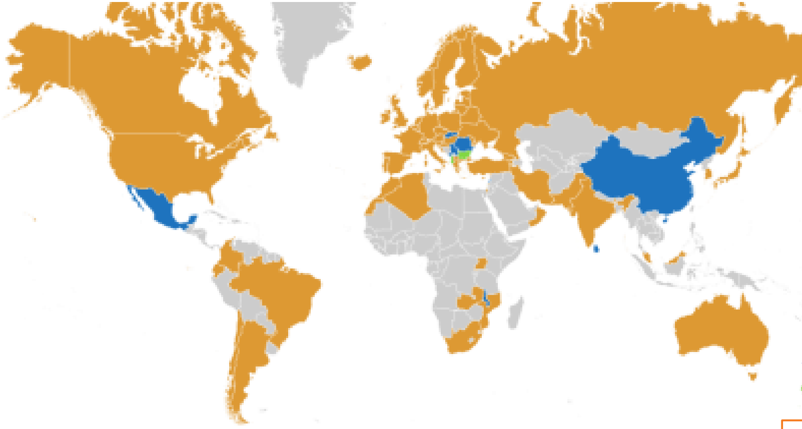
# But what did we now enable?

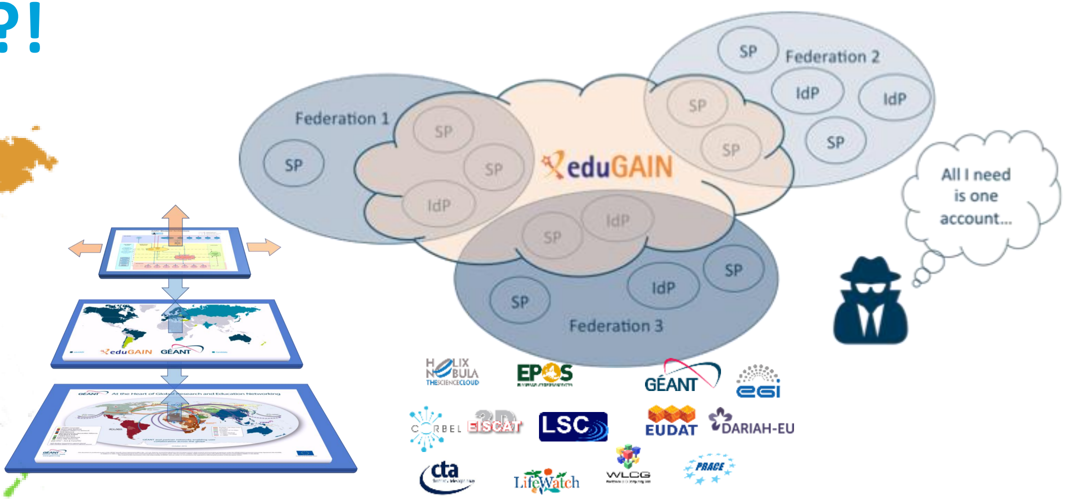
Collaborative security

Sirtfi

Testing resilience and Sirtfi v2

# 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 – 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.

SIRTFI
Security Incident Response Trust Framework for Federated Identity

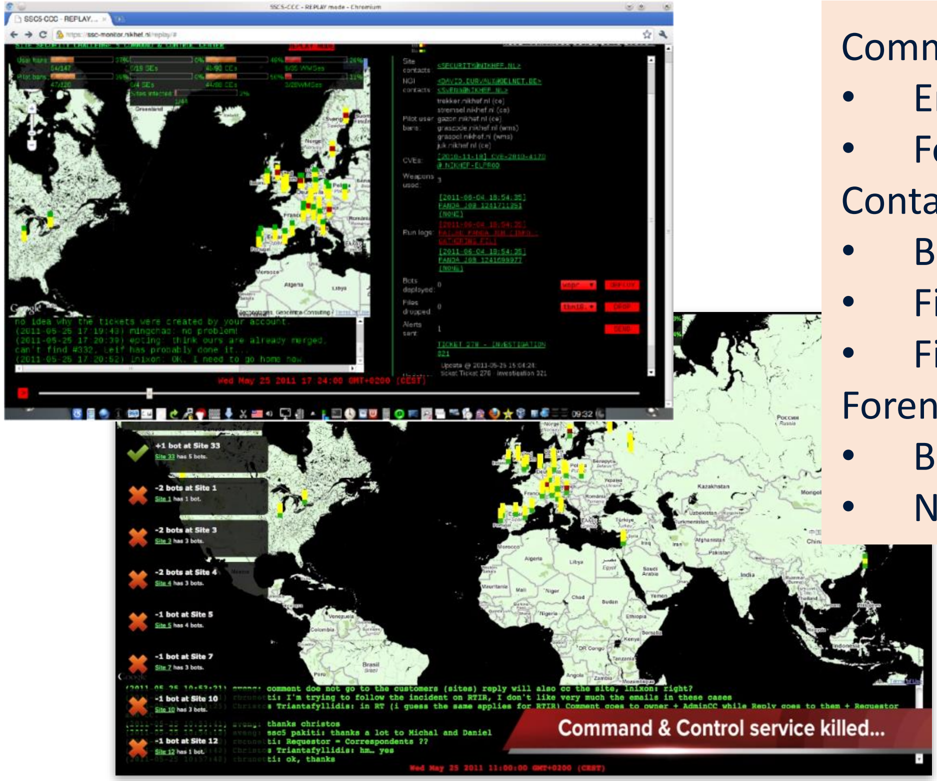
Defines a set of capabilities and roles associated with security incident response that an IdP or SP **organisation self-asserts**. The Sirtfi trust framework posits that organisations asserting conformance with these will coordinate their response to security incidents.

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

# A question of *when*, not *if*



```
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) lnixon: OK. I need to go home now.
Wed May 25 2011 17:24:00 GMT+0200 (CEST)
```

```
(2011-05-25 10:53:21) svenq: comment doe not go to the customers (sites) reply will also cc the site, lnixon: right?
chronetti: 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
svenq: thanks christos
svenq: ssc5 pakiti: thanks a lot to Michal and Daniel
chronetti: Requestor = Correspondents ??
Christos Triantafyllidis: hm… yes
chronetti: ok, thanks
Wed May 25 2011 11:00:00 GMT+0200 (CEST)
```

Command & Control service killed...

Communication:

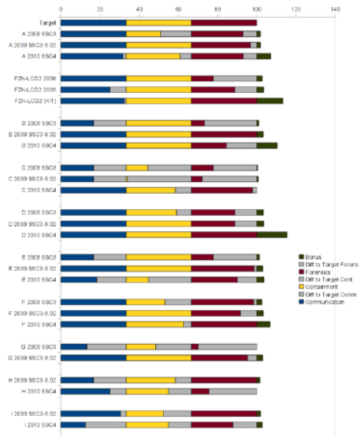
- Endpoints valid?
- Form/Content OK ?

Containment

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

Forensics

- Basic Forensics on binary
- Network traffic



Nikhef 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!

# A federated community security challenge

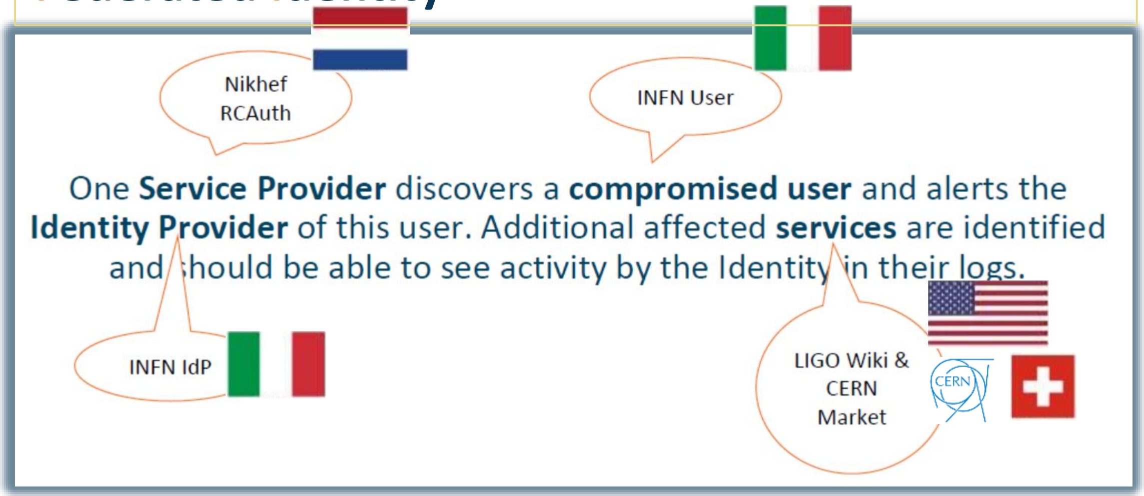
SIRTFI Security Incident Response Trust Framework for Federated Identity

AARC

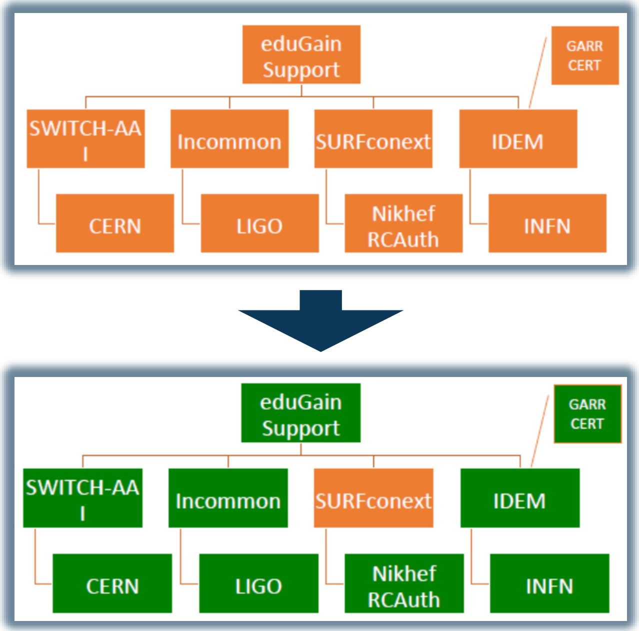
**Can we coordinate our collective R&E response?**

**‘challenges’ based on the *Sirtfi* contact model**

## Security Incident Response Trust Framework for Federated Identity



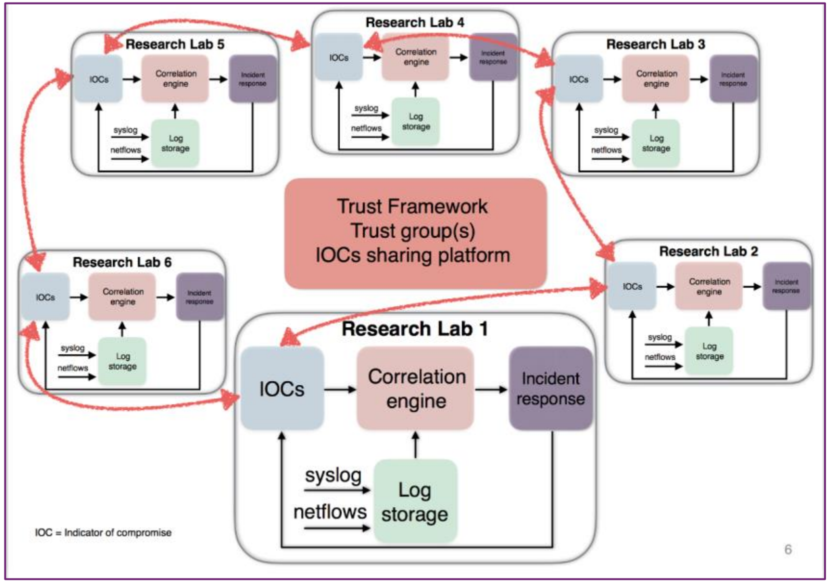
One **Service Provider** discovers a **compromised user** and alerts the **Identity Provider** of this user. Additional affected **services** are identified and should be able to see activity by the Identity in their logs.



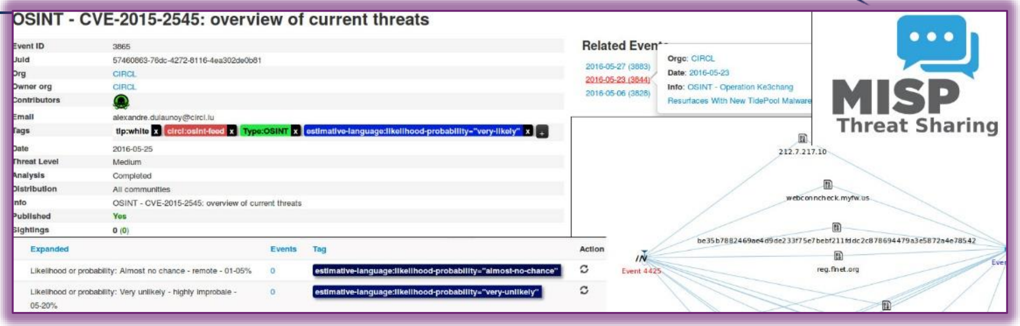
**parties involved in response challenge**

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

# Sharing threat intel – working with our community



IOC = Indicator of compromise



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

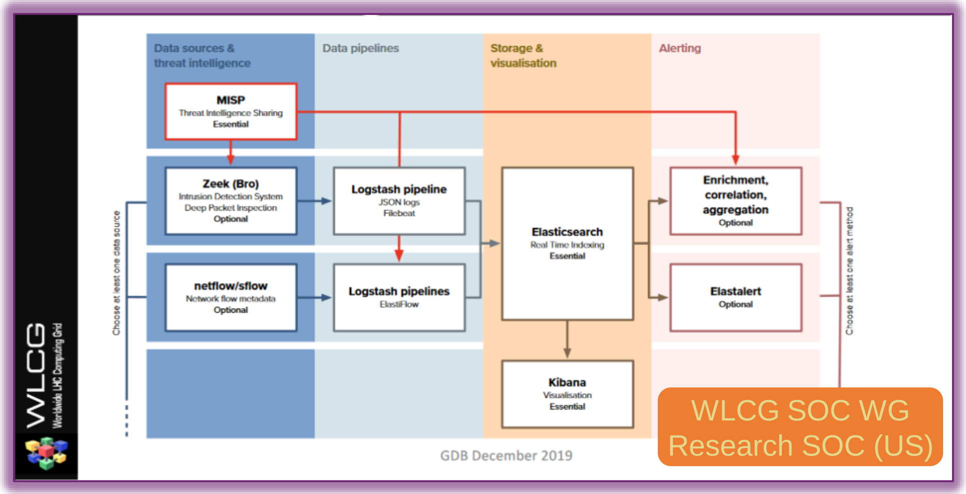
| | |
|---|---|
| Event ID | 3865 |
| Uuid | 57460863-76dc-4272-8116-4ea302de0b81 |
| Org | CIRCL |
| Owner org | CIRCL |
| Contributors | |
| Email | alexandre.dulaunoy@circl.lu |
| Tags | tlp:white x circl:osint-feed x Type:OSINT x estimative-language:likelihood-probability="very-likely" x + |
| Date | 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 Events

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" | |

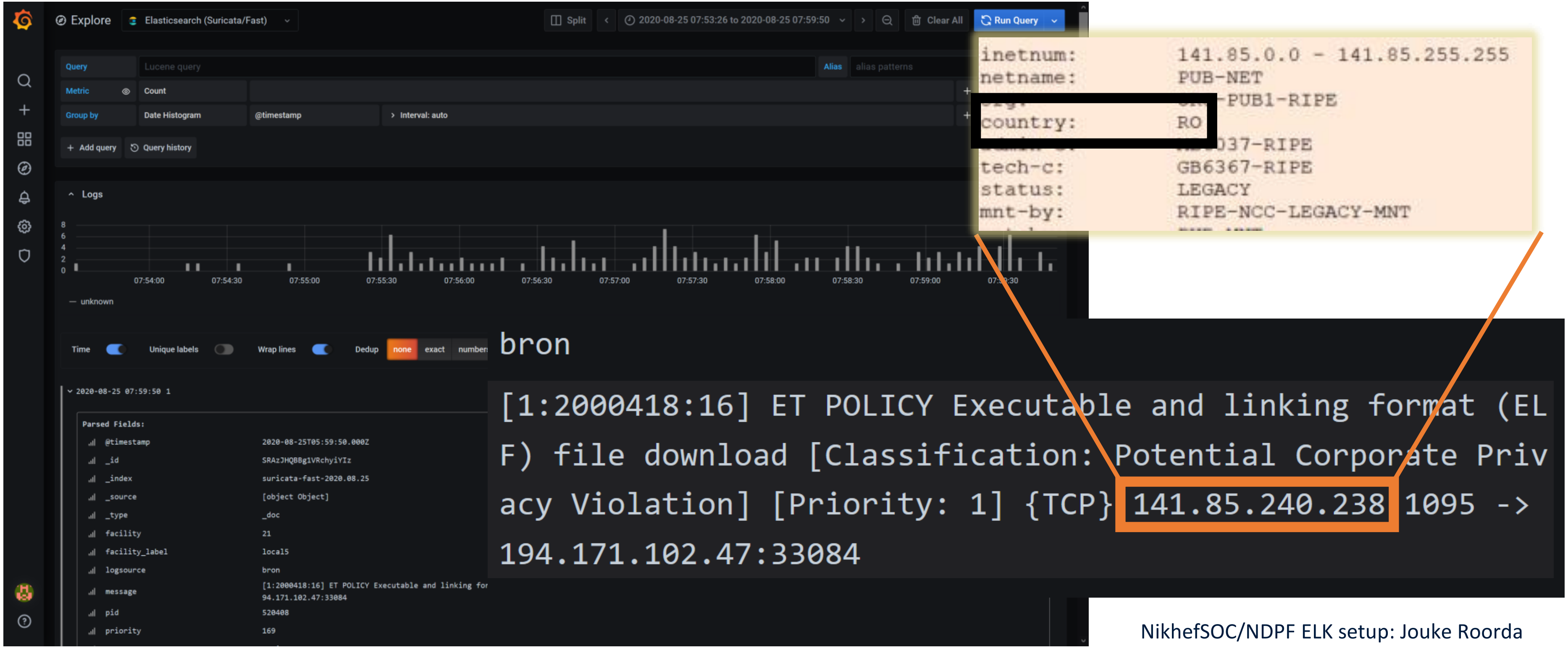


WLCG SOC WG
Research SOC (US)

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

# Nikhef SOC – NDPF traffic analysis

**many ‘false warnings’ when industry-standard (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/NDPF ELK setup: Jouke Roorda

# Scalable credential translation in the AARC BPA … building RCauth.eu

Leveraging federation for ubiquitous, compatible research and collaboration identity credentials

# Ingredients for credential minting & token translation

## eduGAIN (global R&E) *Entity Categories*

*Curated grouping of entities*

‘REFEDS R&S’

*this is a research service*

‘DP CoCo’

*abides by GDPR*

‘Sirtfi’

*cares for security response*

slower adoption process
adding identity assurance needs
action at all 60+ Feds & 4k+ IdPs

## e-Infrastructure IGTF Authentication Profiles

Common baseline and profiles *co-defined by relying parties*

user-centric ID harmonisation with unique global naming

‘BIRCH’

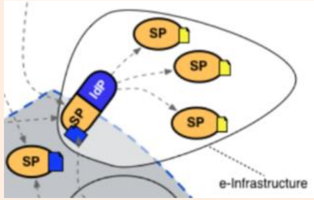
*real person with real name*

‘DOGWOOD’

*persistent linkable identifier*

research-specific user base

## Use of proxy bridging components



Identity and access ‘proxy’ harmonised eduGAIN IdPs

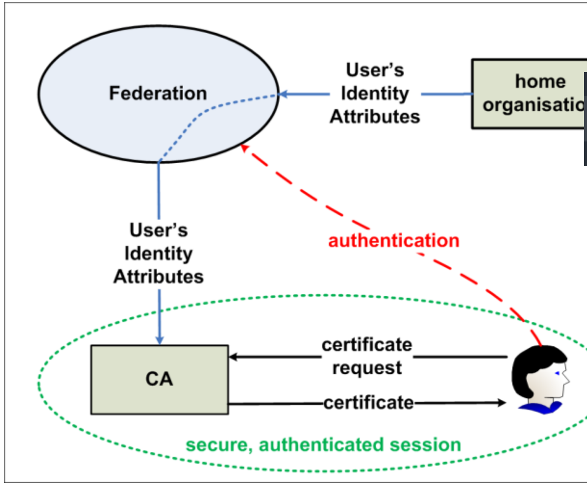
*based on entity categories*
*leverage Sirtfi and ‘R&S’*
*proxying is bi-directional*

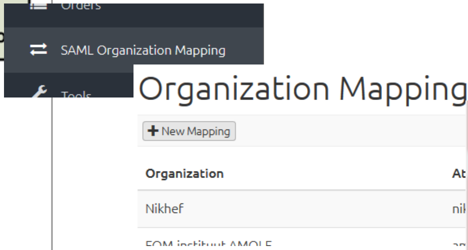
responsibility on the proxy operator

https://wiki.geant.org/display/AARC/Current+Status+of+SAML+Entity+Categories+Adoption
https://www.rcauth.eu/
https://aarc-community.org/

# Bridges and Token Translation Services

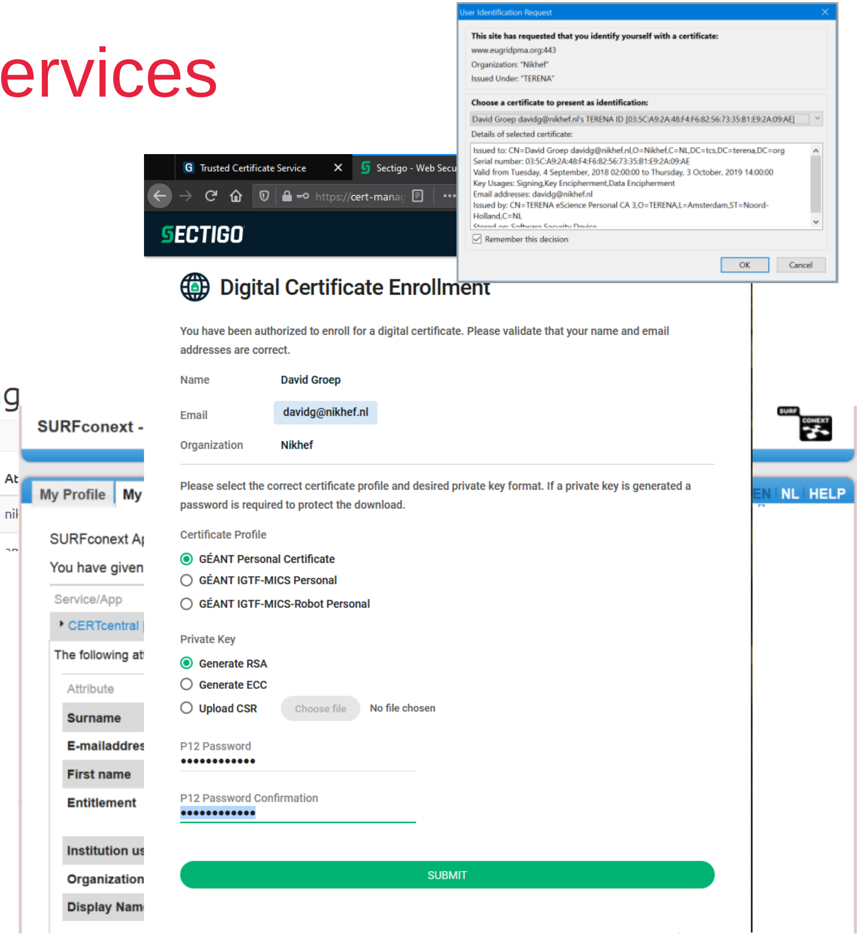
## GEANT Trusted Certificate Service


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



Organization Mapping

| Organization | At |
|---|---|
| Nikhef | nik |
| FOM instituut AMOLF | an |


SURFconext -

SURFconext Ap

You have given

Service/App

CERTcentral

The following at

| Attribute |
|---|
| Surname |
| E-mailaddres |
| First name |
| Entitlement |
| Institution us |
| Organization |
| Display Nam |

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
••••••••••••

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

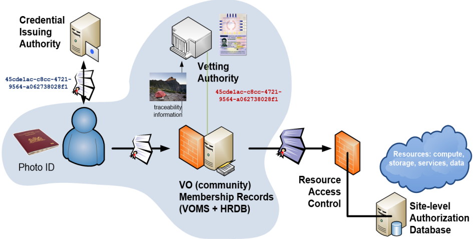
☑ Remember this decision

TCS (today: Sectigo) acts as SAML Service provider to eduGAIN: eligible authenticated users can obtain client certificate for access and delegation to services

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

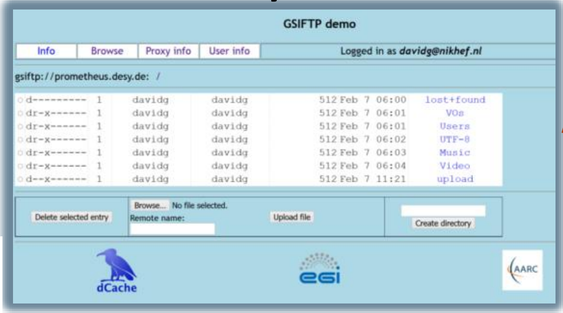
# A ‘CILogon-like’ Token Translations Service – RCauth.eu

- Ability to serve a large pan-European+ user base without national restrictions
  - without specific national participation requirement (serve sparsely distributed users)
- Use existing resources and e-Infrastructure services
  - no need for revise security model at resource centres or at infrastructure level
- Integrate science gateways and portals with minimal effort
  - only light-weight industry-standard protocols
- Support attribute-cert community membership services
  - support community membership via attribute certificates, also for science portal access to e-Infrastructure
- Concentrate service elements that require security expertise
  - not burden research communities with care for security-sensitive service components
  - keep a secure credential management model
  - coordinate compliance and accreditation



# In-line token translation services SAML-to-PKIX?

Community Science Portal



GSIFTP demo

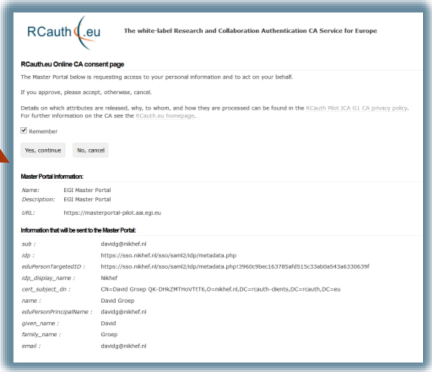
Logged in as *davidg@nikhef.nl*

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
```

No file selected.

Remote name:



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

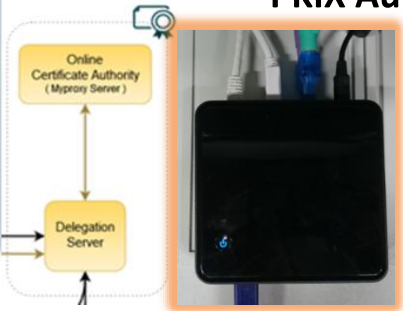
**Master Portal Information:**

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

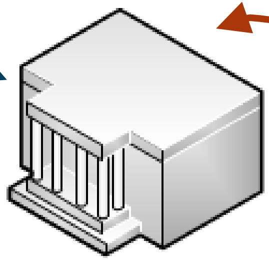
**Information that will be sent to the Master Portal:**

| | |
|---|---|
| *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-DHkZMTHoVTtT6,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



Accredited PKIX Authority

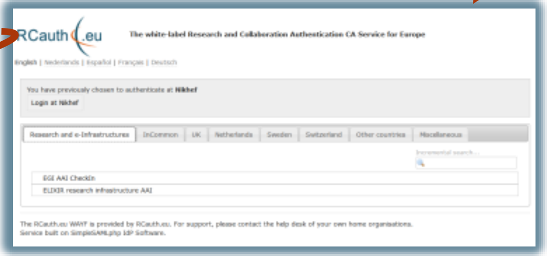


User Home Org

***or Infrastructure IdP***

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

REFEDS R&S
Sirtfi Trust



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

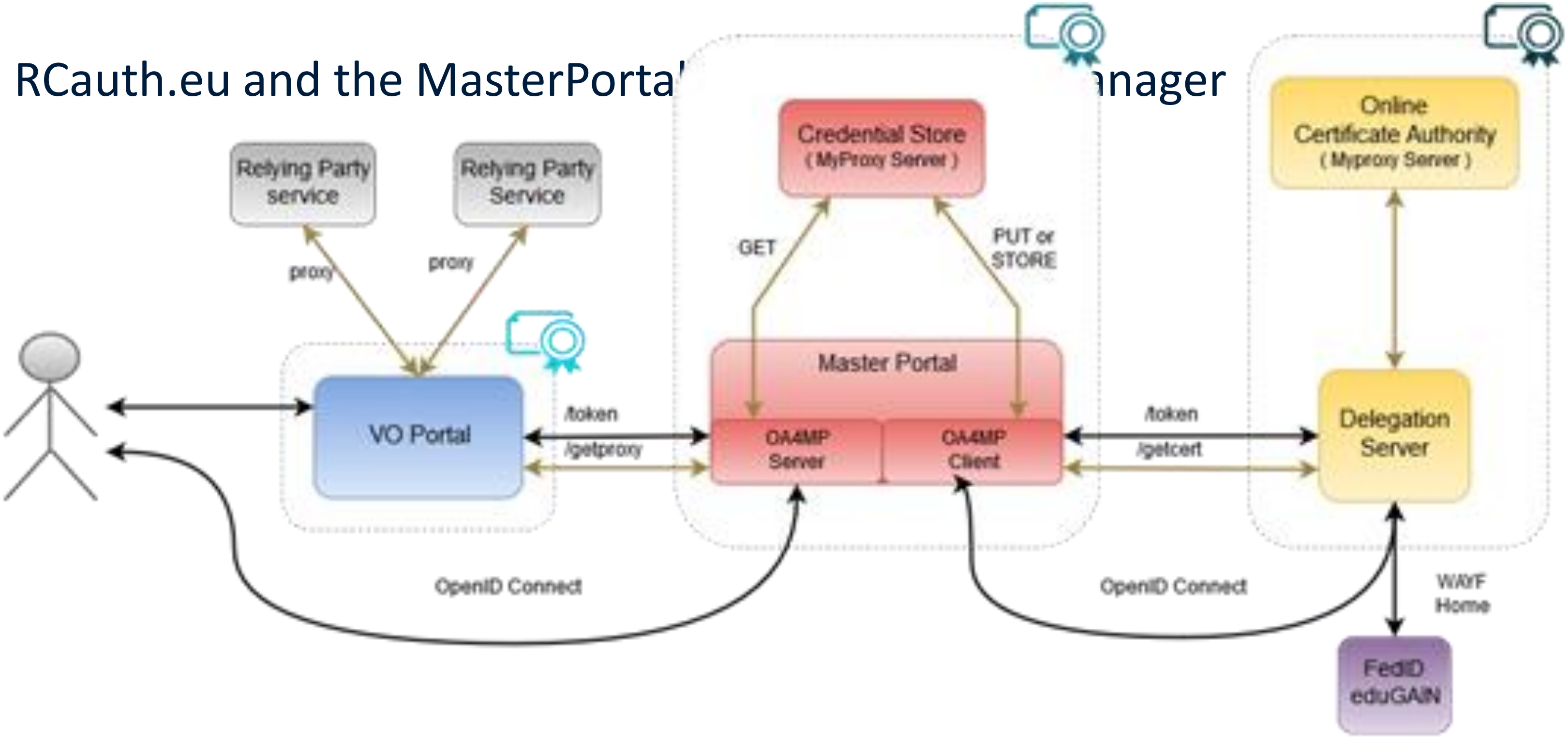
You have previously chosen to authenticate at **Nikhef**

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 / eduGAIN

**Built on CILogon and MyProxy**
**www.cilogon.org**

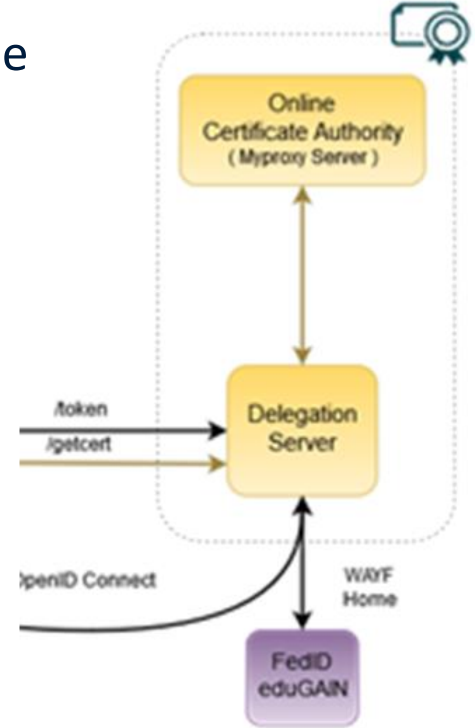
- RCauth.eu and the MasterPortal anager



# RCauth.eu – a white-label IOTA CA in Europe

RCauth.eu

- Cover as much as R&E Federated (Europe++) as possible
- Scoped to research and collaborative use cases
- In a scalable and sustainable deployment model



https://rcauth.eu/ https://rcdemo.nikhef.nl/

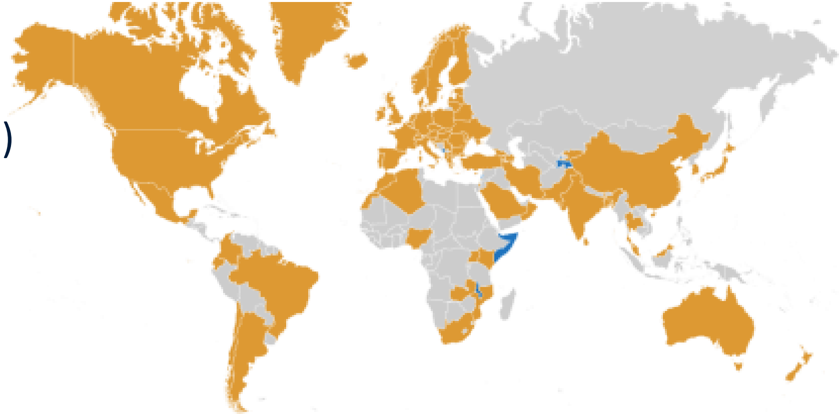
CILogon Service

Service inspired by and using components (such as the DS) from Jim Basney's CILogon, see https://www.cilogon.org/docs/20141030-basney-cilogon.pdf

# The joys of global interfederation

global IdPs in eduGAIN and
the quest for a reasonable, non-reassigned name

# Our Registration Authorities: the Federated IdPs



- distributed RAs: the *eligible IdPs*
  - connected through a Federated Identity Management System (FIMS)
  - primarily: ensemble of IdPs in eduGAIN that meet the policy requirements of this CA
- eligible applicants are all affiliated to an RA

**Three eligibility models**

1. Direct relationship CA-IdP, with agreement declaration
2. Rest of eduGAIN: – “Sirtfi” security incident response and OpSec capabilities plus
   – REFEDS “R&S section 6” non-reassigned identifiers and applicant name
   are required, and tested via statement in ‘meta-data’ and by releasing the proper attributes
3. within the Netherlands, SURFconext Annex IX* already ensures compliance for all IdPs
   *“IdPs within eduGAIN are deemed to have entered materially into an agreement with the CA”*

# 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”

# CommonName – the big challenge

Requirement

- Contain ‘a representation of the real name of the applicant’ as asserted by the IdP
- *the purely ‘opaque’ option is not very friendly to downstream services*

***Does the IdP give the attributed from which to construct the real name – globally?***

# commonName – should be readable element … … in printable 7-bit chars

‘REFEDS R&S’ gives a subset of attributes that should be released:

1. the *displayName* attribute from the IdP
2. the *givenName* attribute, followed by a space, followed by the *sn* attribute from the IdP
3. the *commonName* (cn) attribute from the IdP

but we need to make it printable in ASCII

We tried using *java.text.Normalizer.Form.NFD* and map the remainder to “X”, which gives:

| If IdP sends us this UTF-8 | Representation in CN RDN |
|---|---|
| Jőzsi Bácsi | Jozsi Bacsi |
| Guðrún Ósvífursdóttir | GuXrun Osvifursdottir |
| Χρηστος Κανελλοπουλος | XXXXXXX XXXXXXXXXXXX |
| 簡禎儀 | XXX |

**Oops!**

# Also Νικόλας Λιαμπότης might not like that … and I understand …

***java.text.Normalizer.Form.NFD*** **and ‘X-ing’ the rest particularly bad for**
Greeks, Bulgarians, Chinese, Georgians, Thai, Armenians, Serbians, …

***ICU - International Components for Unicode*** **(icu-project.org) appears to be better, but:**

- there are many options for transliteration
- some code points shared between different languages, that prefer different transliterations
- some code points are absent even in UTF-8 causing ambiguity

So we moved to using ICU, but even then the mapping is not trivial:

UTF-8 $\xrightarrow{\text{ICU}}$ Latin-1 $\xrightarrow{\text{ICU}}$ ASCII $\xrightarrow{\text{regex}}$ IA5String (we need PrintableString + “@” and minus [:/=])

thanks go to Mischa Sallé for the transliteration studies (and much more), ICU is available from https://icu.unicode.org/

# And straightforward translation is not always good

**Just Any-Latin fails for Slavonic unique “sh” sounds. E.g. for ‘Миша’**

- with *Any-Latin* becomes ‘Miša’ which then translates into ‘Misa’ after the Latin-Ascii

but quite some people called ‘Миша’ want to see ‘Mischa’, but not all, so you need

- first *Russian-Latin/BGN*, making it ‘Misha’, which is slightly better, then do *Any-Latin* (1-to-1)
- but “*Russian-Latin/BGN+Serbian-Latin/BGN*” is different from the reverse …

**First Any-Latin/BGN, then Any-Latin, to fix mapping to → š and the → s**

- Բարեւ աշխարհ → Barev a**sh**kharh (with the /BGN, to ensure the “sh”)
- ישראל → ysr'l (taken care of without the /BGN, otherwise the ש never makes it)

**And Unicode does not distinguish the *diaeresis* and the *umlaut***

- Günter Strauß ⟶ Gunter Strauss *should* have been ‘Guenter Strauss’
- Daniëlle ⟶ Danielle is good, you definitely don’t want ‘Danieelle’

*As the so for stability, we keep Any-Latin here and treat all as a diaeresis*

# For our multi-federated world, we ended up with

So the (for now) best combination seems to be the ordered transformation:

```
Transliterator.getInstance( "Russian-Latin/BGN;"+
        "Serbian-Latin/BGN;"+
        "Greek-Latin/UNGEGN;"+
        "[:Nonspacing Mark:] remove;"+
        "Any-Latin/BGN;“ +
        "Any-Latin;“ +
        "Latin-Ascii“
);

result.replaceAll("[^\\p{Lower}\\p{Upper}\\p{Digit} '()+,-.?@]",  "X");
```

*ordering to retain “ш” → “sh”*

*Fixes greek Λ adding a useless space*

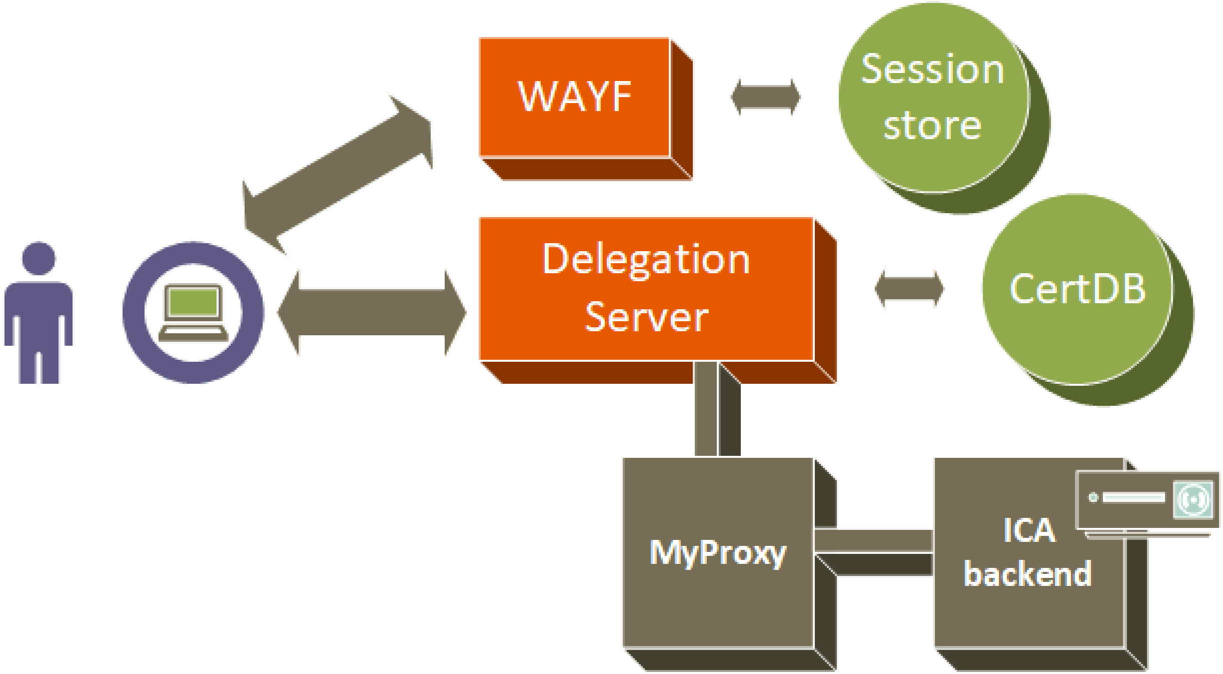
*Retain proper “sh” when coming from Armenian or Hebrew by /BGN first*

# What will we get?

```
$ java -cp icu4j-59_1.jar:. transliterate2 [...]
 "Jőzsi Bácsi" "Guðrún Ósvífursdóttir" \
 "Χρηστος Κανελλοπουλος" "簡禎儀“


Input:    Jőzsi Bácsi
Output:   Jozsi Bacsi
Input:    Guðrún Ósvífursdóttir
Output:   Gudrun Osvifursdottir
Input:    Χρηστος Κανελλοπουλος
Output:   Christos Kanellopoulos
Input:    簡禎儀
Output:   jian zhen yi
```

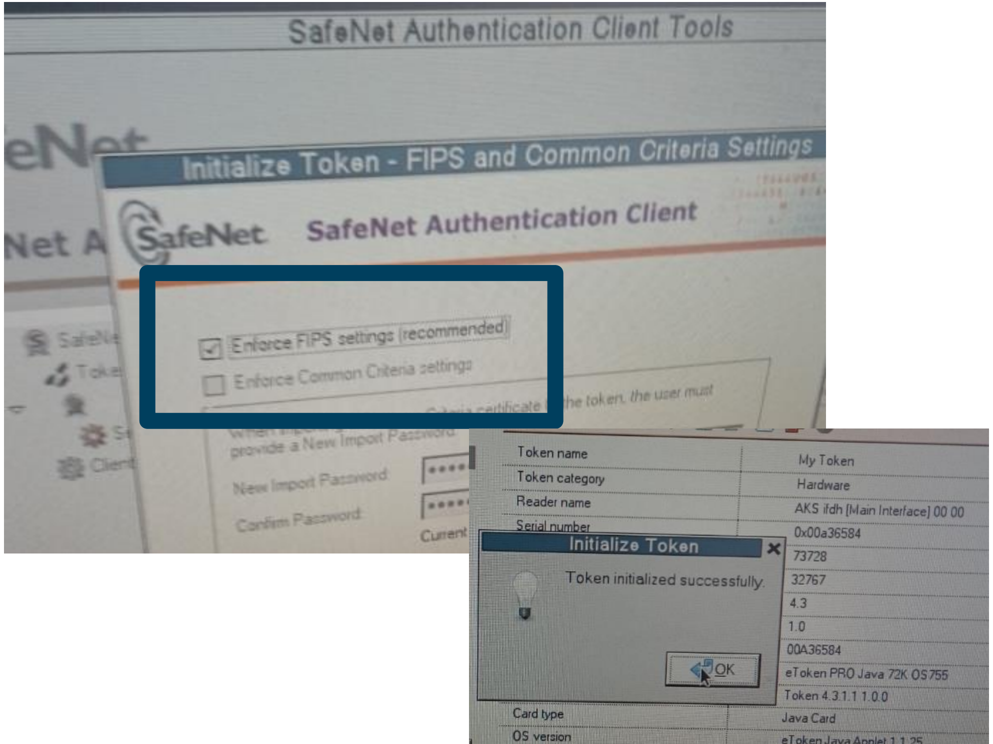
# Building the initial RCauth.eu

# From a single instance …



# A fully compliant ‘Heath Robinson’ CA

# It is on the HSM …



Initialize Token - FIPS and Common Criteria Settings

SafeNet Authentication Client

☑ Enforce FIPS settings (recommended)

□ Enforce Common Criteria settings

When importing [illegible] Criteria certificate [illegible] the token, the user must
provide a New Import Password

New Import Password ••••

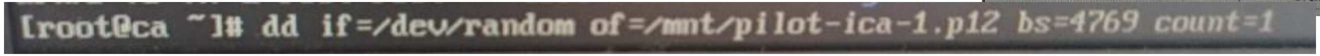
Confirm Password ••••

Current

| | |
|---|---|
| Token name | My Token |
| Token category | Hardware |
| Reader name | AKS ifdh [Main Interface] 00 00 |
| Serial number | 0x00a36584 |
| | 73728 |
| | 32767 |
| | 4.3 |
| | 1.0 |
| | 00A36584 |
| | eToken PRO Java 72K OS755 |
| | Token 4.3.1.1 1.0.0 |
| Card type | Java Card |
| OS version | eToken Java Applet 1.1.25 |

Initialize Token

Token initialized successfully.

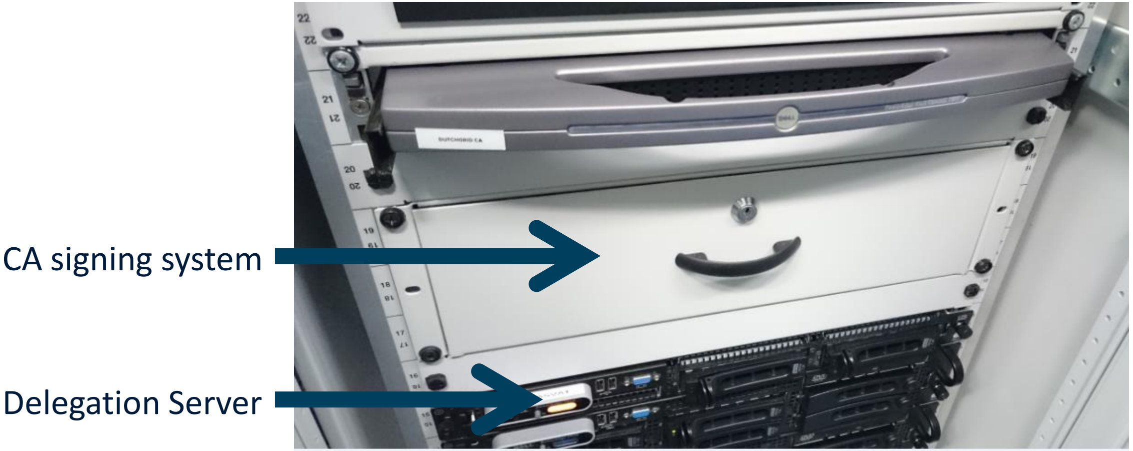


```
[root@ca ~]# dd if=/dev/random of=/mnt/pilot-ica-1.p12 bs=4769 count=1
```

# Physical controls

- Located at Nikhef, Amsterdam, NL
- Scientific Data Centre part of the NikhefHousing Facilities
- ID based access control, 24hr guard on-site
- CA and security systems in locked dedicated cabinet on 2nd floor

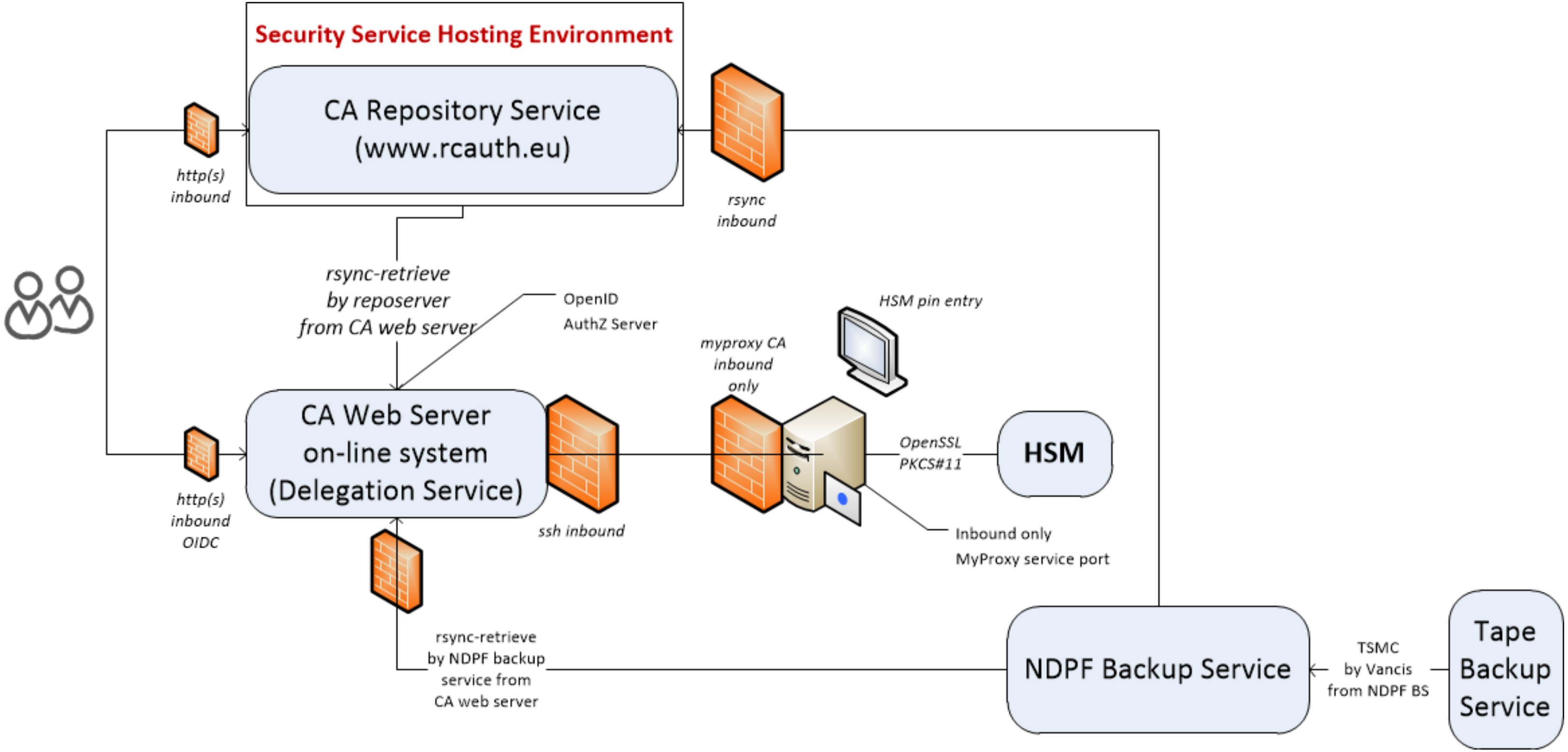
  On-line CA signing system in locked drawer



CA signing system

Delegation Server

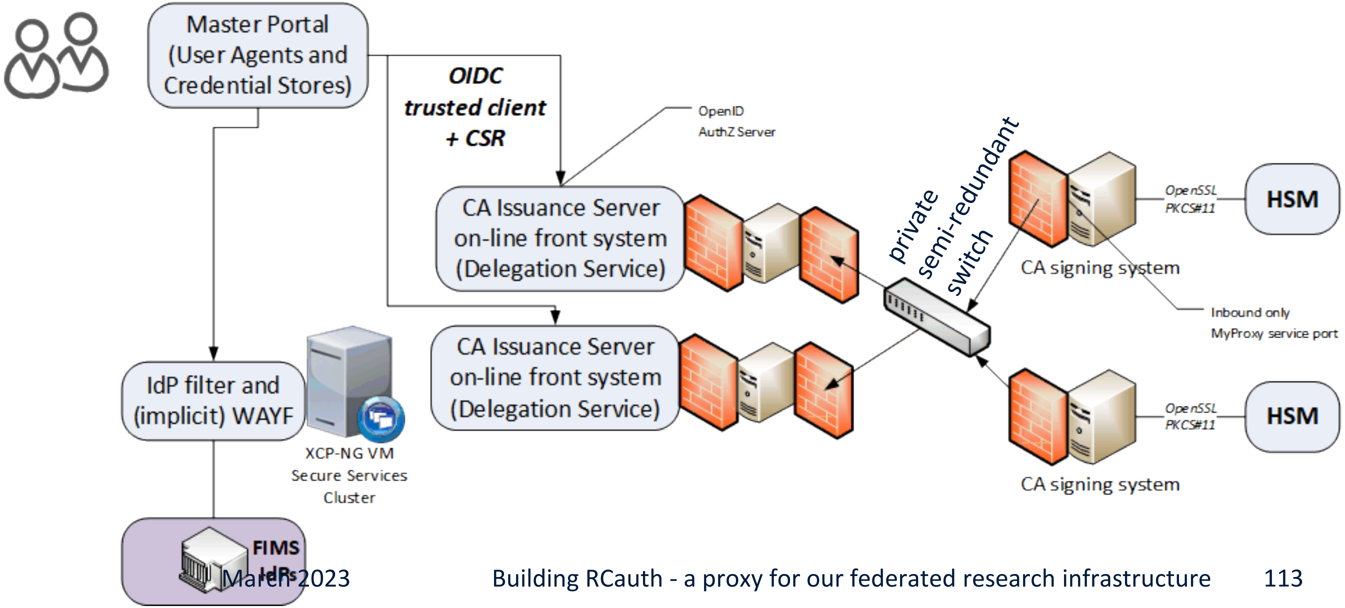
# Logical set-up


Security Service Hosting Environment
CA Repository Service (www.rcauth.eu)
http(s) inbound
rsync inbound
rsync-retrieve by reposerver from CA web server
OpenID AuthZ Server
HSM pin entry
myproxy CA inbound only
CA Web Server on-line system (Delegation Service)
http(s) inbound OIDC
ssh inbound
OpenSSL PKCS#11
HSM
Inbound only MyProxy service port
rsync-retrieve by NDPF backup service from CA web server
NDPF Backup Service
TSMC by Vancis from NDPF BS
Tape Backup Service

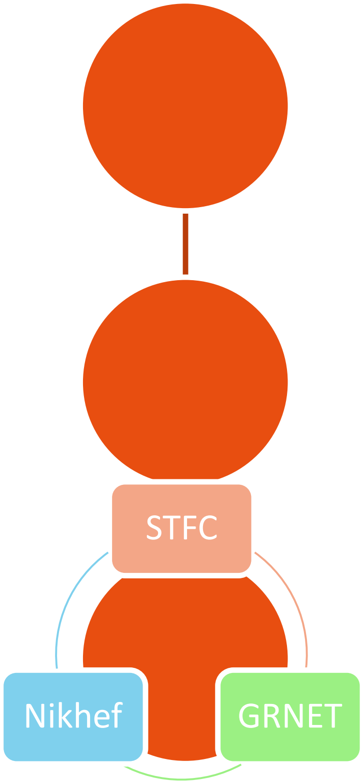
# A local highly-available setup at Nikhef Amsterdam

- Most ‘fault-prone’ components are
  - Intel NUC (single power supply)
  - HSM (can lock itself down, and the USB connection is prone to oxidation)
  - DS front-end servers (physical hardware, albeit with redundant disks and powersupplies)

**Eliminated first using ‘local HA’**


Master Portal (User Agents and Credential Stores)
OIDC trusted client + CSR
OpenID AuthZ Server
CA Issuance Server on-line front system (Delegation Service)
CA Issuance Server on-line front system (Delegation Service)
private semi-redundant switch
OpenSSL PKCS#11
HSM
CA signing system
Inbound only MyProxy service port
OpenSSL PKCS#11
HSM
CA signing system
IdP filter and (implicit) WAYF
XCP-NG VM Secure Services Cluster
FIMS IdPs

# Distributing RCauth.eu across three cooperating sites



Governance Board

**Representatives** (and one alternate) from each Materially Contributing Stakeholder
EGI.eu, EUDAT (ETFC), GÉANT, Nikhef (SURF)

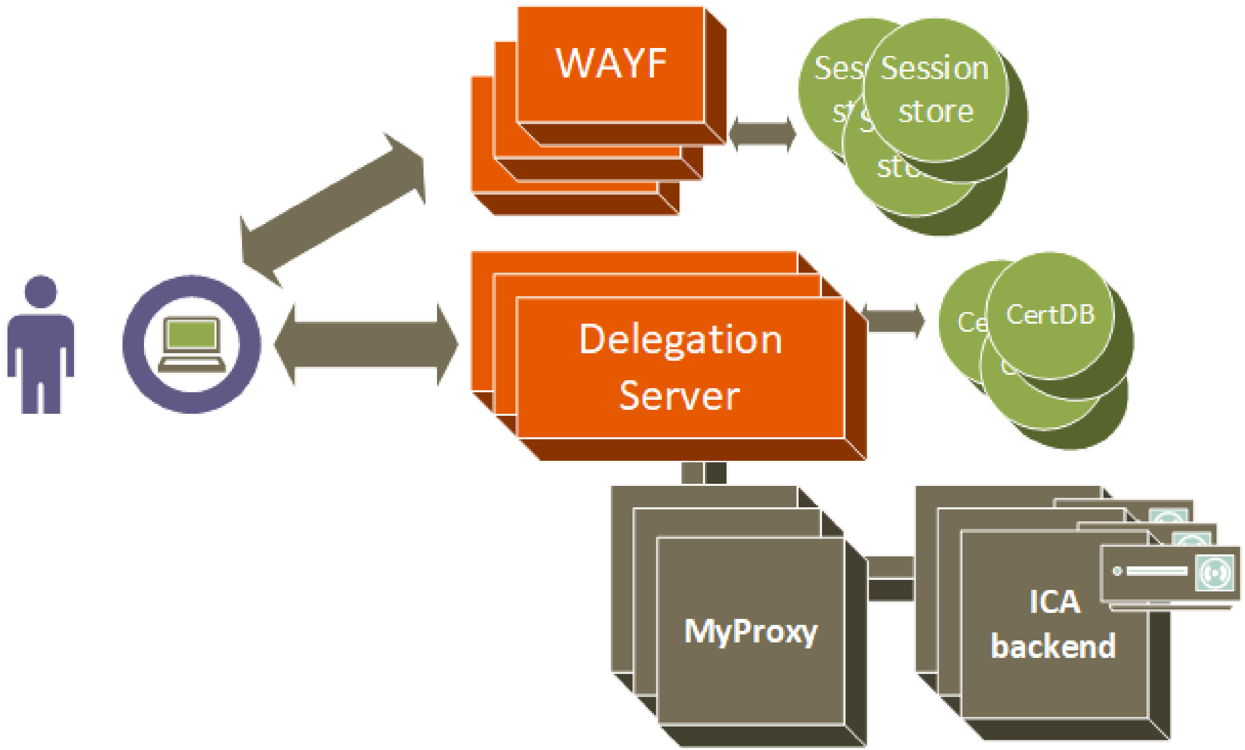
RCauth PMA

**Individuals** drawn from the wide community […] experts in the field of identity management for research and collaboration,
PKI technology and identity bridging

Ops Coordination Team

Operations people from each of the **hosting partners** with a (copy of) the RCauth.eu signing key,
and those partners otherwise involved in OPS

# … to a 3-fold continuously-consistent setup



# HA solutions

**Local high availability, three distinct providers?**

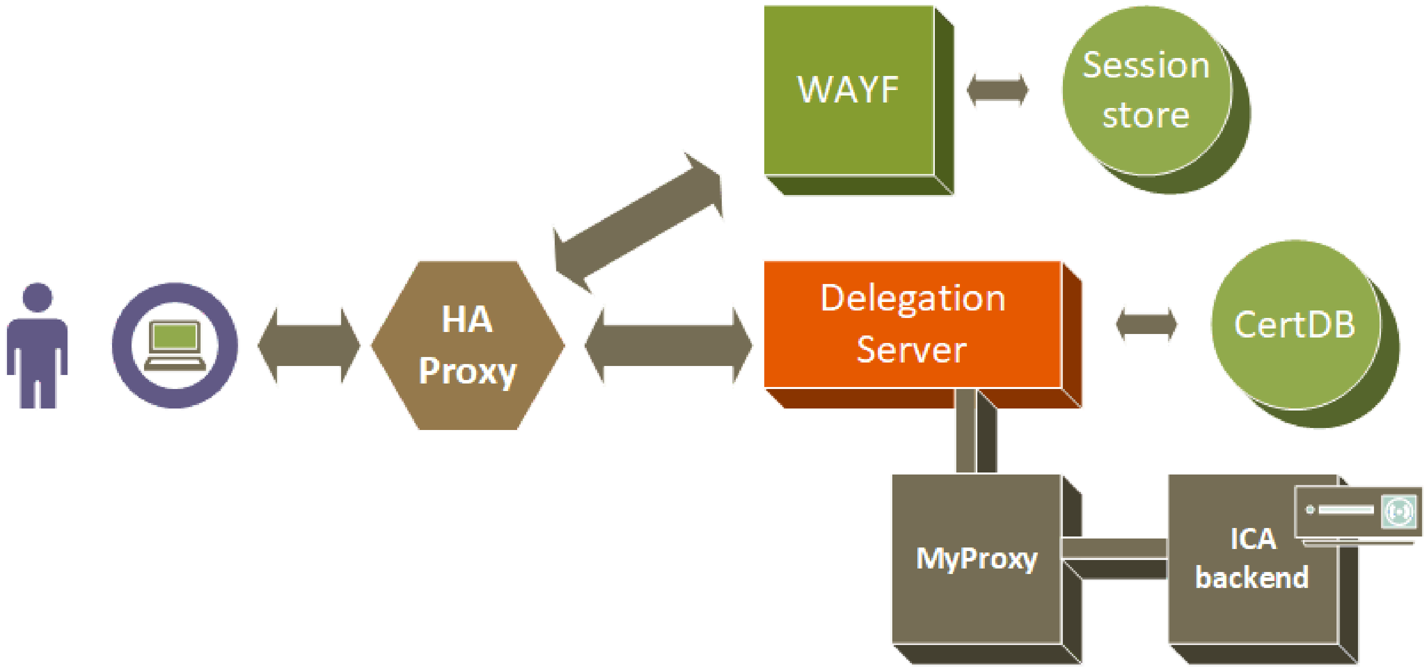
- pushes account linking burden to the relying parties/service providers
- users may have 3 credentials, which is confusing
- a single identifier would require ‘ensured’ database synchronization – no true independence

**DNS-based fail-over?**

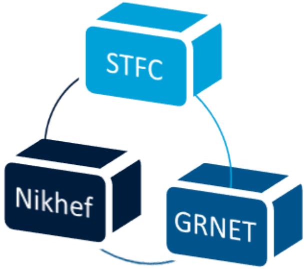
- the ‘trivial’ model relies on the client not to cache answers for long, ***and*** not to round-robin the DNS answers - since the WAYF and DS go together
- short TTL  limits reliance, since both service and domain name provider must be up
- ‘advanced’ DNS-based solutions (like for InAcademia) – with near-realtime updates of a distributed DNS may appear better, but still: need a overly-low TTL, and move the HA problem to the DNS provider (or ccTLD), rather than solve it

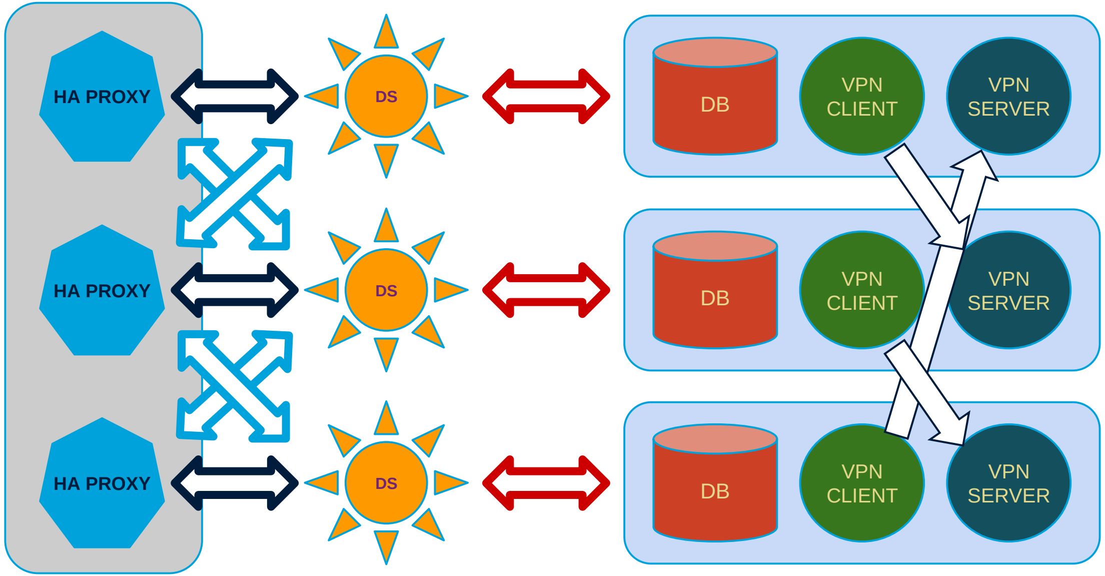
**So we looked at network-layer resilience, the ‘go-to’ solution for large CDN providers**

# Services at a site go up and down together  - adding HAproxy

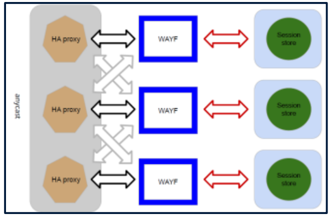


# Distributed RCauth service





*similar for the policy-filtering WAYF*

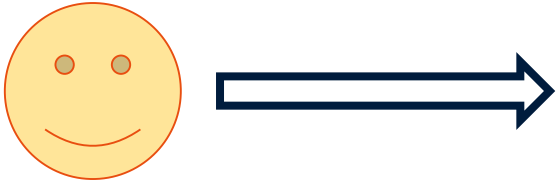


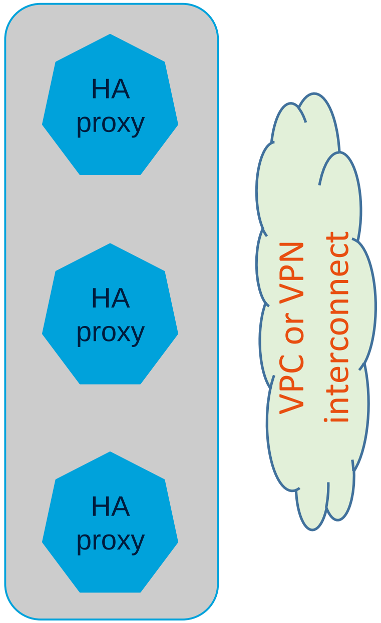
selected imagery: Mischa Sallé, Jens Jensen, Nicolas Liampotis

# A transparent multi-site setup

User

- connects to HA proxy at {wayf,pilot-ica-g1}.rcauth.eu
- HA proxy sends users to “closest” working service
- forward mainly to its own DS when available





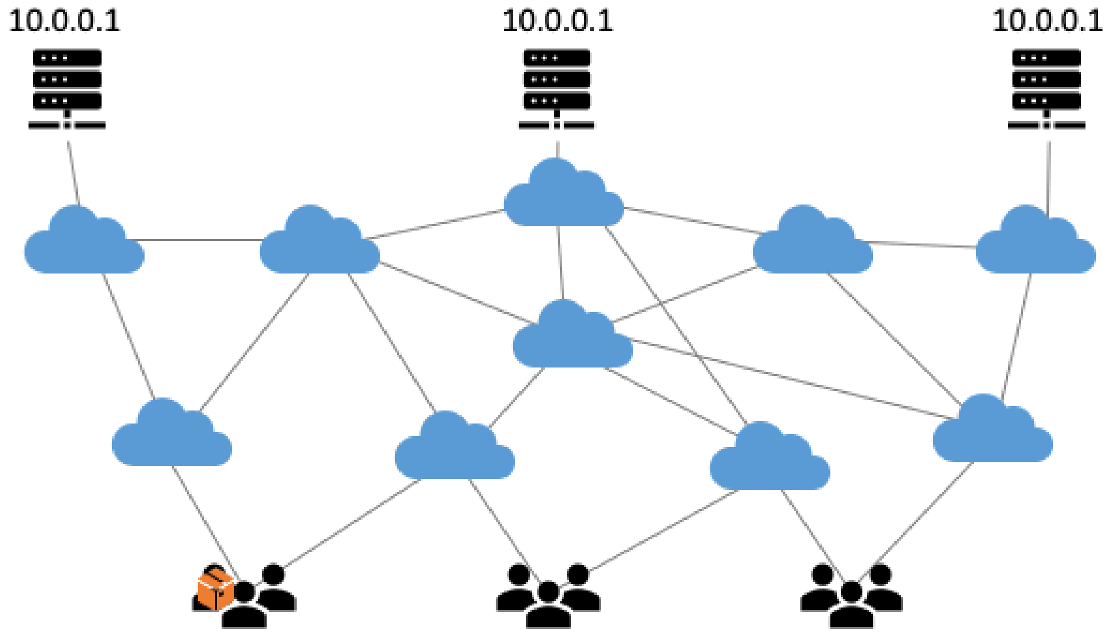
*If a HA loses its backend DS, can still route to another DS over VPC/VPN backend*

and wherever the user is, the service is at

- **2a07:8504:01a0::1**
- and **145.116.216.1** (for legacy IP users)

selected imagery: Mischa Sallé, Jens Jensen, Nicolas Liampotis

# Anycast: when the same place exists many times



## So we used

- 3 (for now: 2) sites
- one VM at each site exposing 2a07:8504:01a0::1
- smallest v6 subnet (/48)
- bird + a service probe
- each site’s own ASN
- some IRR DB editing
- IPv4 is similar, with a /24

*and some monitoring*

routing image: SIDNlabs - https://www.sidnlabs.nl/en/news-and-blogs/the-bgp-tuner-intuitive-management-applied-to-dns-anycast-infrastructure

# BIRD config and probes

- you need
- a health checker to drive the local BGP daemon
- a BGP talker, such as bird
- a very simple config

```
# Generated 2023-02-05 14:49:36.063331
# by anycast-healthchecker (pid=1299)
# 2001:db8::1/128 is a dummy IP Prefix.
# It should NOT be used and REMOVED
# from the constant.
define ACAST6_PS_ADVERTISE =
    [
        2001:db8::1/128,
        2a07:8504:1a0::1/128
    ];
```

```
include "/etc/bird.d/*.conf";

router id 194.171.98.77;

define ASN_OWN         = 65530;
define ASN_NEIGHBOUR   = 1104;
define ADDR_NEIGHBOUR4 = 194.171.98.94;
define ADDR_NEIGHBOUR6 = 2a07:8500:120:e011::1;

protocol device { scan time 10; }

protocol direct direct1 {
    interface "lo";
    ipv4 { import all; export none; };
    ipv6 { import all; export none; };
}

template bgp bgp_peers4 {
    local as ASN_OWN;
    ipv4 {
        import none;
        export filter match_route_filter;
    };
}
template bgp bgp_peers6 {
    local as ASN_OWN;
    ipv6 {
        import none;
        export filter match_route6_filter;
    };
}
protocol bgp BGP4 from bgp_peers4 { disabled no; neighbor ADDR_NEIGHBOUR4 as ASN_NEIGHBOUR; }
protocol bgp BGP6 from bgp_peers6 { disabled no; neighbor ADDR_NEIGHBOUR6 as ASN_NEIGHBOUR; }
```

# But what is ‘healthy’?

- Service status verification tool needed to ‘drive’ bird actions
- anycast_healthchecker by Pavlos Parissis

- with HAproxy
  on the front-end host
  on each site

```
Packager    : Mischa Sallé <msalle@nikhef.nl>
Vendor      : Pavlos Parissis <pavlos.parissis@gmail.com>
URL         : https://github.com/unixsurfer/anycast_healthchecker
Summary     : A healthchecker for Anycasted Services
Description :
Anycast-healthchecker monitors a service by doing periodic health
checks and based on the result instructs Bird daemon to either
advertise or withdraw the route to reach the monitored service. As
a result Bird will only advertise routes for healthy services.
```

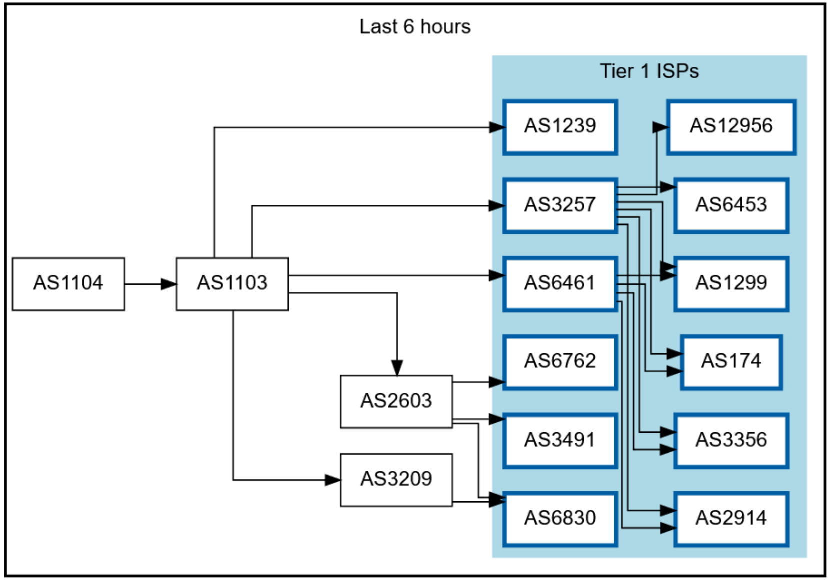
```
[haproxy]
check_cmd      = /usr/local/sbin/check_haproxy.sh
on_disabled    = withdraw
ip_prefix      = 145.116.216.1/32
[haproxy6]
check_cmd      = /usr/local/sbin/check_haproxy.sh
on_disabled    = withdraw
ip_prefix      = 2a07:8504:1a0::1/128
```

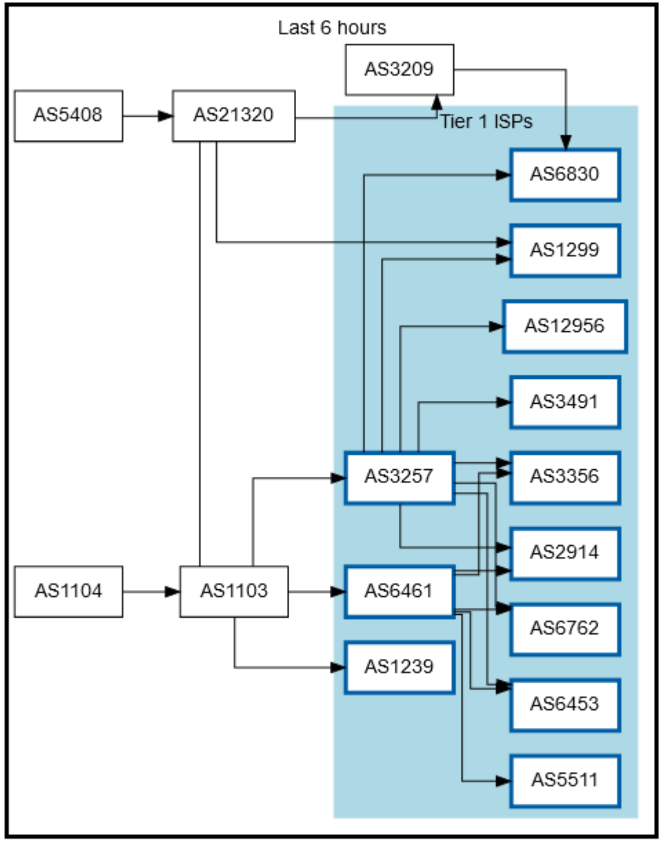
# Both Delegation Service and filtering WAYF should be up

- But since Nikhef also has local HA with two back-ends, either is OK!

```
# Checks WAYF backends, at least one should be up or starting
# i.e. in state 2 or 3 (see Section 9.3 Unix Socket commands in
# management.txt).
check_wayf()    {
    echo $state_cmd |\
        socat unix-connect:${haproxy_socket} stdio |\
        grep $wayf_pattern |\
        cut -d' ' -f${site_col},${state_col} |\
        while read wayf_site wayf_state
    do
        if [ "$wayf_state" -ge 1 -a "$wayf_state" -le 2 ];then
            # Found at least one up DS
            info "WAYF $wayf_site has state $wayf_state"
            return 1
        else
            warn "WAYF $wayf_site has state $wayf_state" >&2
        fi
    done
    return $((1-$?))
}
```

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

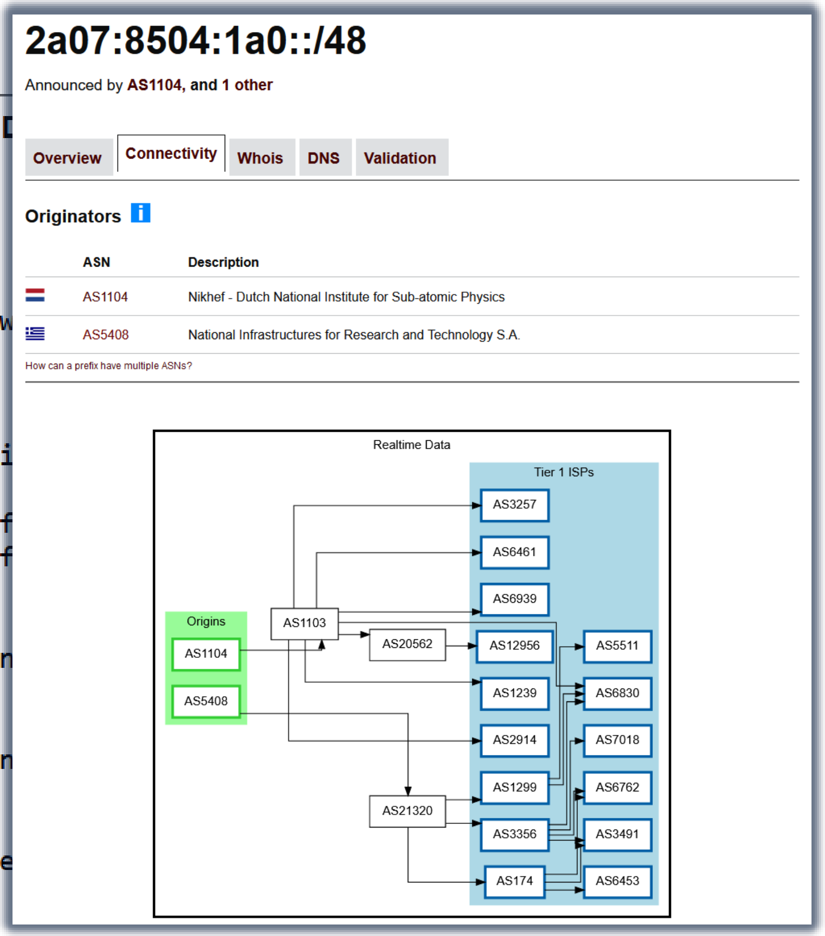




route maps: bgp.tools for 2a07:8504:1a0::/48 – IPv4 for 145.116.216.0/24 is similar – imagery from November 2022

# CERN Looking Glass Results - ee1

```
inet6.0: 155476 destinations, 303862 routes (155437 active, 0 holddow
+ = Active Route, - = Last Active, * = Both

2a07:8504:1a0::/48 *[BGP/170] 01:08:50, MED 20, localpref 10500
                      AS path: 20965 5408 I, validation-state: unveri
                    >  to 2001:798:99:1::39 via irb.200
                    [BGP/170] 4d 23:13:16, MED 20, localpref 10500, f
                      AS path: 1103 1104 I, validation-state: unverif
                    >  to fe80::1a2a:d300:140f:bdb0 via irb.20
                    [BGP/170] 6d 23:17:01, MED 20, localpref 10500
                      AS path: 2603 1103 1104 I, validation-state: un
                    >  to 2001:1458:0:9::2 via irb.2903
                    [BGP/170] 01:08:26, MED 25, localpref 10500
                      AS path: 559 20965 5408 I, validation-state: un
                    >  to 2001:1458:0:2c::2 via irb.2902
                    [BGP/170] 01:08:49, MED 10, localpref 10200
                      AS path: 174 174 21320 5408 I, validation-state
                    >  to 2001:978:2:2::2a:1 via irb.3811
```



## 2a07:8504:1a0::/48

Announced by **AS1104, and 1 other**

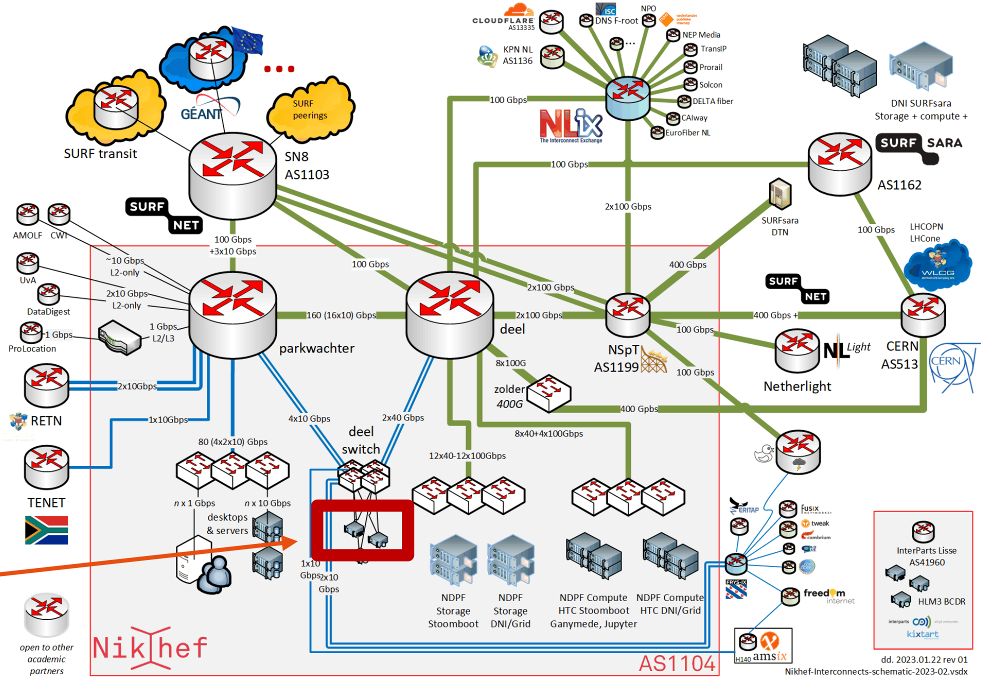
**Originators**

| | ASN | Description |
|---|---|---|
| 🇳🇱 | AS1104 | Nikhef - Dutch National Institute for Sub-atomic Physics |
| 🇬🇷 | AS5408 | National Infrastructures for Research and Technology S.A. |

How can a prefix have multiple ASNs?

# Shortest path, also when mixing with the default-free zone

```
[root@kwark ~]# traceroute -IA 145.116.216.1
traceroute to 145.116.216.1 (145.116.216.1), 30 hops max, 60 byte packets
 1  cmbr.connected.by.freedominter.net
      (185.93.175.234) [AS206238]
 2  connected.by.freedom.nl
      (185.93.175.240) [AS206238]
 3  et-0-0-0-1002.core1.fi001.nl.freedomnet.nl
      (185.93.175.208) [AS206238]
 4  as1104.frys-ix.net (185.1.203.66) [*]
 5  parkwachter.nikhef.nl
      (192.16.186.141) [AS1104]
 6  gw-anyc-01.rcauth.eu
      (145.116.216.1) [AS786/AS5408/AS1104]
```



*rcauth.eu HA proxy*

Route from home to RCauth.eu, from my home Freedom Internet ISP

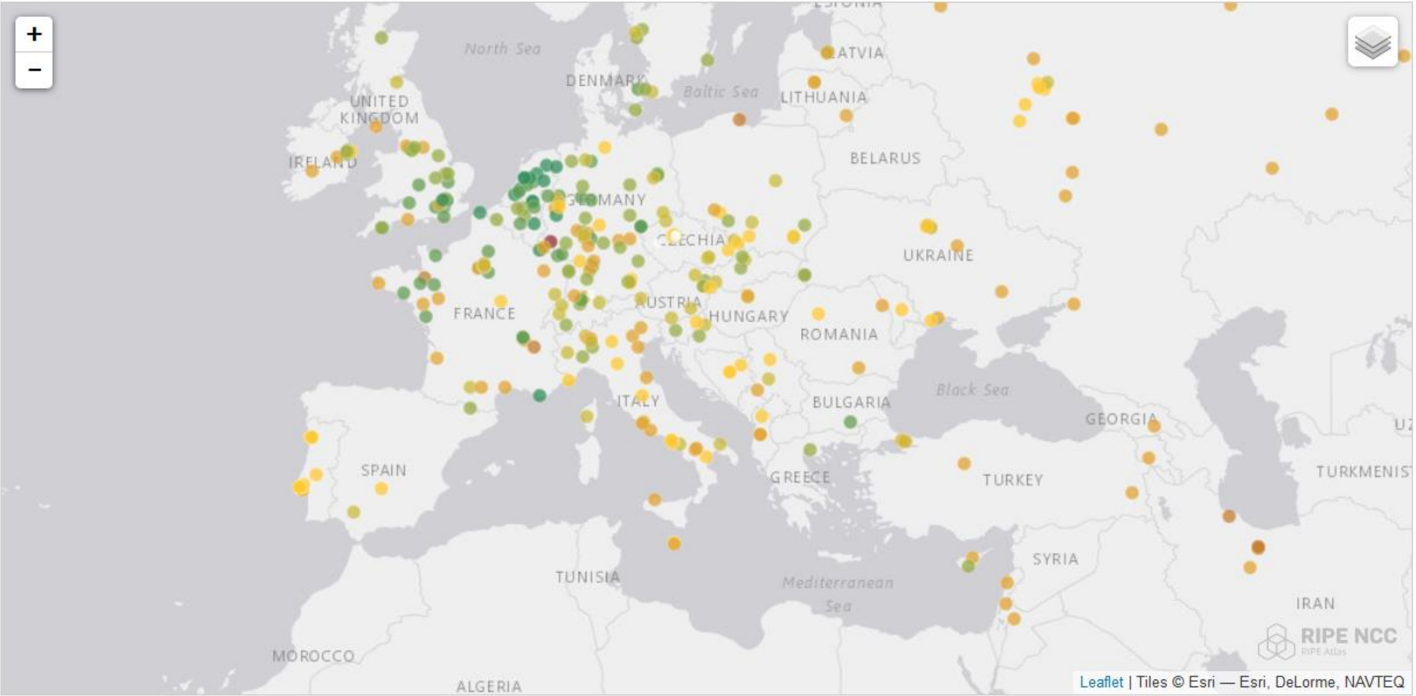
# Prerequisites are relatively simple

- **IPv6 /48 netblock** and legacy IPv4 /24
- your own, or a friendly, **ASN**
- a set of corresponding **IRR route objects**, and either none, or a correct RPKI
  (easily done in your local RIR registry: APNIC, RIPE, ARIN, AfriNIC, LACNIC)
- front-end service (**HAproxy**) for the Delegation Service and filtering WAYF
- **bird** (or quagga) with a service health checker

But you do not per-se need …

- a unique AS just for this anycast activity - it works equally well without it
- a balanced AS path length - unless you want load balancing as well as redundancy
- your own AS - if you have a friendly AS willing to re-announce your specific route

# And you get reasonable load balancing



< 10 ms: 29 | < 20 ms: 46 | < 30 ms: 59 | < 40 ms: 54 | < 50 ms: 64 | < 100 ms: 113 | < 200 ms: 91 | < 300 ms: 26 | > 300 ms: 5 | No Data: 0

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

# Other HA options

- Local HA with an HA proxy and pacemaker/CRM failover works
  on the local network – and can be meshed with two signing systems
  … this is used extensively (also active/passive) for other services at Nikhef

- DNS-based fast-failover – the method used for e.g. InAcademia
  automatic updating of DNS a distributed set of servers, auto-updating each other
  … does require that the DNS domain level operator remains available,
  since you need *very* short TTLs, and still your ccTLD/gTLD needs HA as well

- use dedicated HA links for the back-end database connection or ip-forwarding
  e.g. multiple redundant circuits over an MPLS cloud emerging at each site

# Current status of RCauth.eu

- All sites can sign production certificates
- DS databases cross-site replication using Galera over VPN
- HA CRL cross site synchronisation and issuance
- WAYF servers (GRNET and Nikhef)

# Reuse the RCauth experience

**All sources, Ansible playbooks, and materials are on GitHub**
**https://github.com/rcauth-eu**

- HA database and back-end VPN
  - 3-node peer-peer redundant VPN with automatic failover
  - extensible to >3, but then topology is less clear
- Web services
  - HAproxy stability and flexibility and coordinated ‘up-down’ status per site
- HAHAP | BGP Anycast
  - ‘bog-standard’ if service admins, cloud admins, and network people can collaborate and investigate incidents together

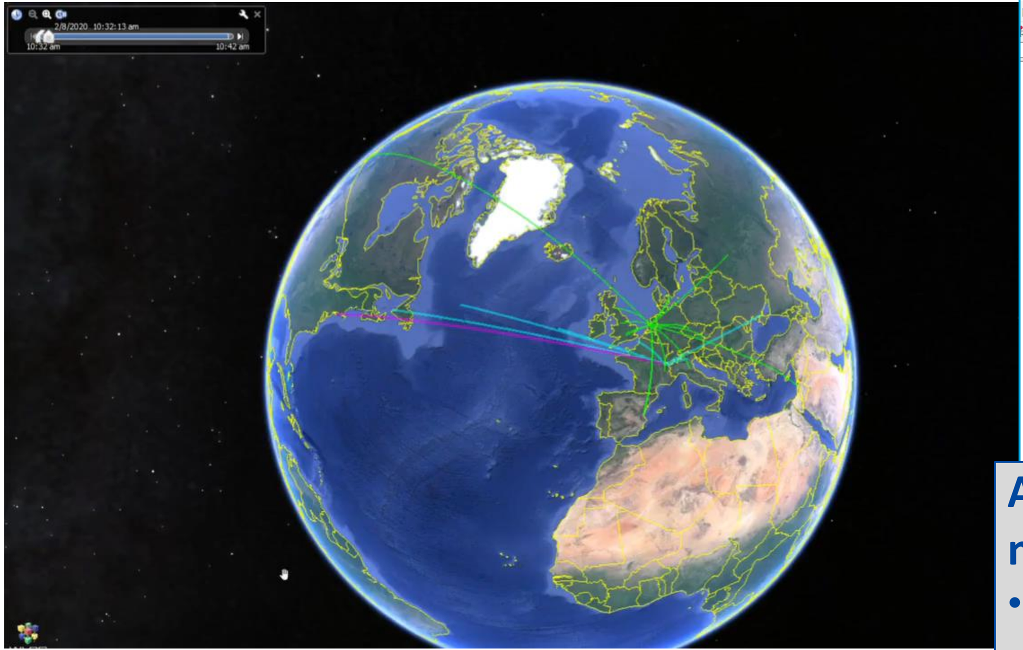
- secure credential sharing and moving shared secrets is still cumbersome in practice
  *‘the difference between theory and practice is that, in theory, there is no difference’*

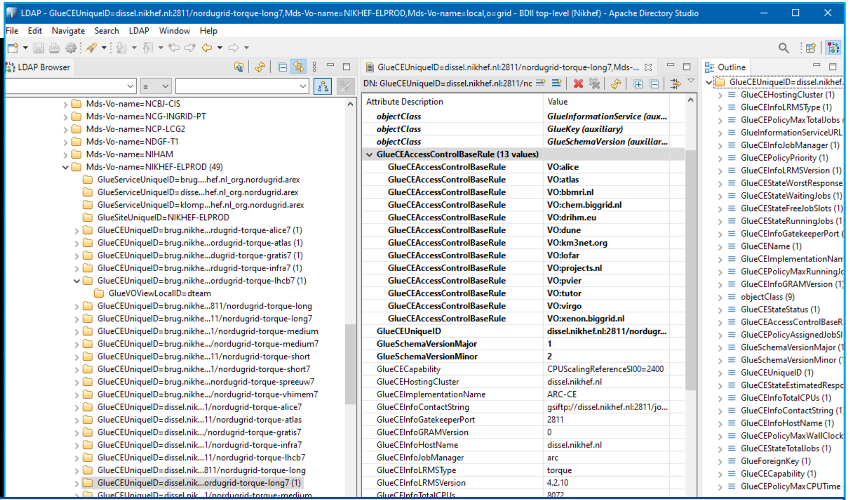
# Putting it back together again

Infrastructure for many communities
ESFRI Clusters, the EOSC, and the AARC TREE

Common patterns in scalability

# A global infrastructure of EGI, OSG and WLCG, …




LDAP - GlueCEUniqueID=dissel.nikhef.nl:2811/nordugrid-torque-long7,Mds-Vo-name=NIKHEF-ELPROD,Mds-Vo-name=local,o=grid - BDII top-level (Nikhef) - Apache Directory Studio

```
Mds-Vo-name=NCBJ-CIS
Mds-Vo-name=NCG-INGRID-PT
Mds-Vo-name=NCP-LCG2
Mds-Vo-name=NDGF-T1
Mds-Vo-name=NIHAM
Mds-Vo-name=NIKHEF-ELPROD (49)
  GlueServiceUniqueID=brug....hef.nl_org.nordugrid.arex
  GlueServiceUniqueID=disse...hef.nl_org.nordugrid.arex
  GlueServiceUniqueID=klomp...hef.nl_org.nordugrid.arex
  GlueSiteUniqueID=NIKHEF-ELPROD
  GlueCEUniqueID=brug.nikhe...rdugrid-torque-alice7 (1)
  GlueCEUniqueID=brug.nikhe...ordugrid-torque-atlas (1)
  GlueCEUniqueID=brug.nikhe...dugrid-torque-gratis7 (1)
  GlueCEUniqueID=brug.nikhe...rdugrid-torque-infra7 (1)
  GlueCEUniqueID=brug.nikhe...ordugrid-torque-lhcb7 (1)
    GlueVOViewLocalID=dteam
  GlueCEUniqueID=brug.nikhe...811/nordugrid-torque-long
  GlueCEUniqueID=brug.nikhe...11/nordugrid-torque-long7
  GlueCEUniqueID=brug.nikhe...1/nordugrid-torque-medium
  GlueCEUniqueID=brug.nikhe.../nordugrid-torque-medium7
  GlueCEUniqueID=brug.nikhe...11/nordugrid-torque-short
  GlueCEUniqueID=brug.nikhe...1/nordugrid-torque-short7
  GlueCEUniqueID=brug.nikhe...nordugrid-torque-spreeuw7
  GlueCEUniqueID=brug.nikhe.../nordugrid-torque-vhimem7
  GlueCEUniqueID=dissel.nik...1/nordugrid-torque-alice7
  GlueCEUniqueID=dissel.nik...11/nordugrid-torque-atlas
  GlueCEUniqueID=dissel.nik.../nordugrid-torque-gratis7
  GlueCEUniqueID=dissel.nik...1/nordugrid-torque-infra7
  GlueCEUniqueID=dissel.nik...11/nordugrid-torque-lhcb7
  GlueCEUniqueID=dissel.nik...811/nordugrid-torque-long
  GlueCEUniqueID=dissel.nik...ordugrid-torque-long7 (1)
  GlueCEUniqueID=dissel.nik...1/nordugrid-torque-medium
```

GlueCEUniqueID=dissel.nikhef.nl:2811/nordugrid-torque-long7,Mds-...

DN: GlueCEUniqueID=dissel.nikhef.nl:2811/nc

| Attribute Description | Value |
|---|---|
| ***objectClass*** | ***GlueInformationService (aux...*** |
| ***objectClass*** | ***GlueKey (auxiliary)*** |
| ***objectClass*** | ***GlueSchemaVersion (auxiliar...*** |
| **GlueCEAccessControlBaseRule (13 values)** | |
| **GlueCEAccessControlBaseRule** | **VO:alice** |
| **GlueCEAccessControlBaseRule** | **VO:atlas** |
| **GlueCEAccessControlBaseRule** | **VO:bbmri.nl** |
| **GlueCEAccessControlBaseRule** | **VO:chem.biggrid.nl** |
| **GlueCEAccessControlBaseRule** | **VO:drihm.eu** |
| **GlueCEAccessControlBaseRule** | **VO:dune** |
| **GlueCEAccessControlBaseRule** | **VO:km3net.org** |
| **GlueCEAccessControlBaseRule** | **VO:lofar** |
| **GlueCEAccessControlBaseRule** | **VO:projects.nl** |
| **GlueCEAccessControlBaseRule** | **VO:pvier** |
| **GlueCEAccessControlBaseRule** | **VO:tutor** |
| **GlueCEAccessControlBaseRule** | **VO:virgo** |
| **GlueCEAccessControlBaseRule** | **VO:xenon.biggrid.nl** |
| **GlueCEUniqueID** | **dissel.nikhef.nl:2811/nordugr...** |
| **GlueSchemaVersionMajor** | **1** |
| **GlueSchemaVersionMinor** | **2** |
| GlueCECapability | CPUScalingReferenceSI00=2400 |
| GlueCEHostingCluster | dissel.nikhef.nl |
| GlueCEImplementationName | ARC-CE |
| GlueCEInfoContactString | gsiftp://dissel.nikhef.nl:2811/jo... |
| GlueCEInfoGatekeeperPort | 2811 |
| GlueCEInfoGRAMVersion | 0 |
| GlueCEInfoHostName | dissel.nikhef.nl |
| GlueCEInfoJobManager | arc |
| GlueCEInfoLRMSType | torque |
| GlueCEInfoLRMSVersion | 4.2.10 |
| GlueCEInfoTotalCPUs | 8072 |

```
GlueCEUniqueID=dissel.nikhef.
  GlueCEHostingCluster (1)
  GlueCEInfoLRMSType (1)
  GlueCEPolicyMaxTotalJobs (
  GlueInformationServiceURL
  GlueCEInfoJobManager (1)
  GlueCEPolicyPriority (1)
  GlueCEInfoLRMSVersion (1)
  GlueCEStateWorstResponse
  GlueCEStateWaitingJobs (1)
  GlueCEStateFreeJobSlots (1)
  GlueCEStateRunningJobs (1
  GlueCEInfoGatekeeperPort (
  GlueCEName (1)
  GlueCEImplementationNam
  GlueCEPolicyMaxRunningJo
  GlueCEInfoGRAMVersion (1
  objectClass (9)
  GlueCEStateStatus (1)
  GlueCEAccessControlBaseR
  GlueCEPolicyAssignedJobSl
  GlueSchemaVersionMajor (1
  GlueSchemaVersionMinor (
  GlueCEUniqueID (1)
  GlueCEStateEstimatedRespo
  GlueCEInfoTotalCPUs (1)
  GlueCEInfoContactString (1
  GlueCEInfoHostName (1)
  GlueCEPolicyMaxWallClock
  GlueCEStateTotalJobs (1)
  GlueForeignKey (1)
  GlueCECapability (1)
  GlueCEPolicyMaxCPUTime
```

## An infrastructure with components matched to application needs

- systems architecture, compute (clusters), networking, storage, and application structure
- in a cost-efficient, and energy-efficient, way

BerkeleyDB Information System for EGI, from top-level BDII at ldap://bdii03.nikhef.nl:2170/o=grid; Earth visualization: https://dashb-earth.cern.ch/, Google Earth

# AARC AEGIS and the EOSC - Interconnecting communities

GÉANT

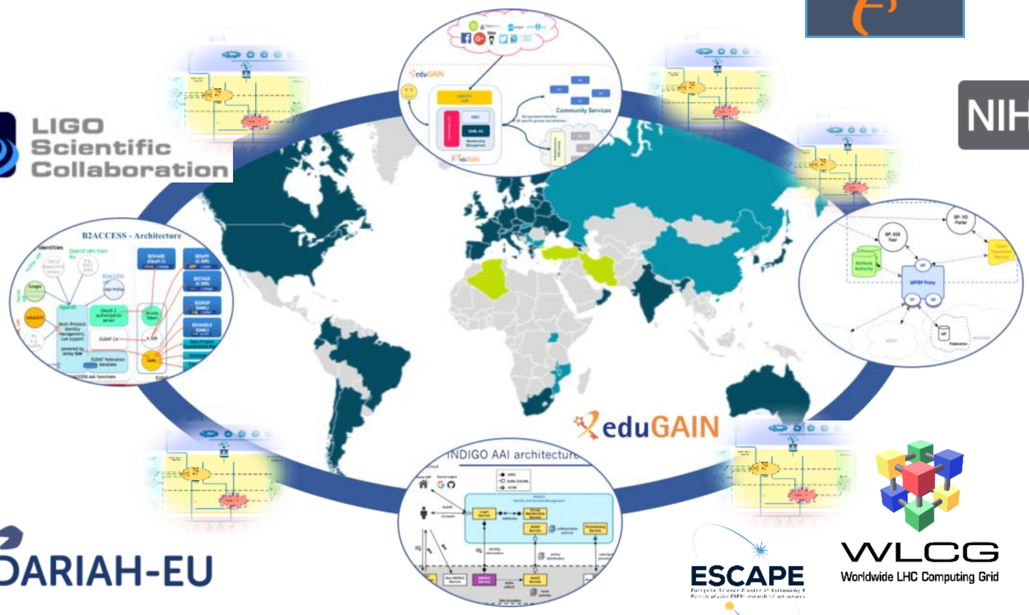
elixir

LSC LIGO Scientific Collaboration

NIH National Institutes of Health
*Turning Discovery Into Health*

EUDAT

egi



DARIAH-EU

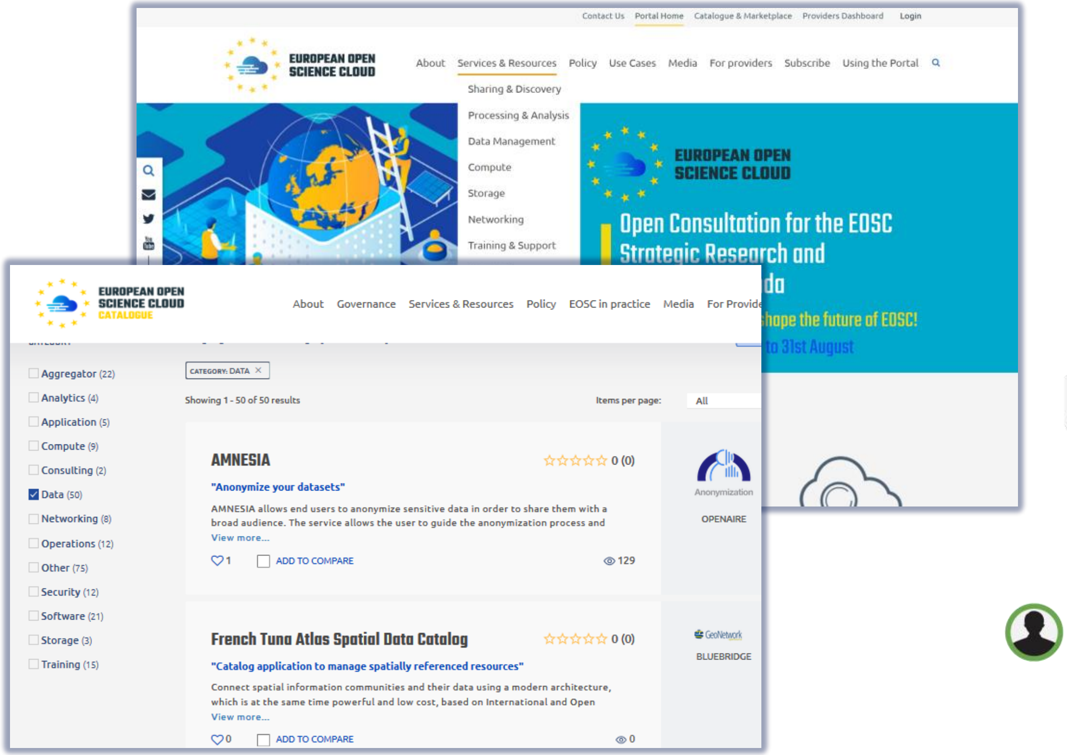
ESCAPE

WLCG
Worldwide LHC Computing Grid

INDIGO - DataCloud
Better Software for Better Science

https://aarc-community.org/about/aegis/

# EOSC: an ecosystem more than just services infrastructure



Contact Us | Portal Home | Catalogue & Marketplace | Providers Dashboard | Login

EUROPEAN OPEN SCIENCE CLOUD

About | Services & Resources | Policy | Use Cases | Media | For providers | Subscribe | Using the Portal

- Sharing & Discovery
- Processing & Analysis
- Data Management
- Compute
- Storage
- Networking
- Training & Support

EUROPEAN OPEN SCIENCE CLOUD

Open Consultation for the EOSC Strategic Research and ...da

...hape the future of EOSC!

...to 31st August

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"

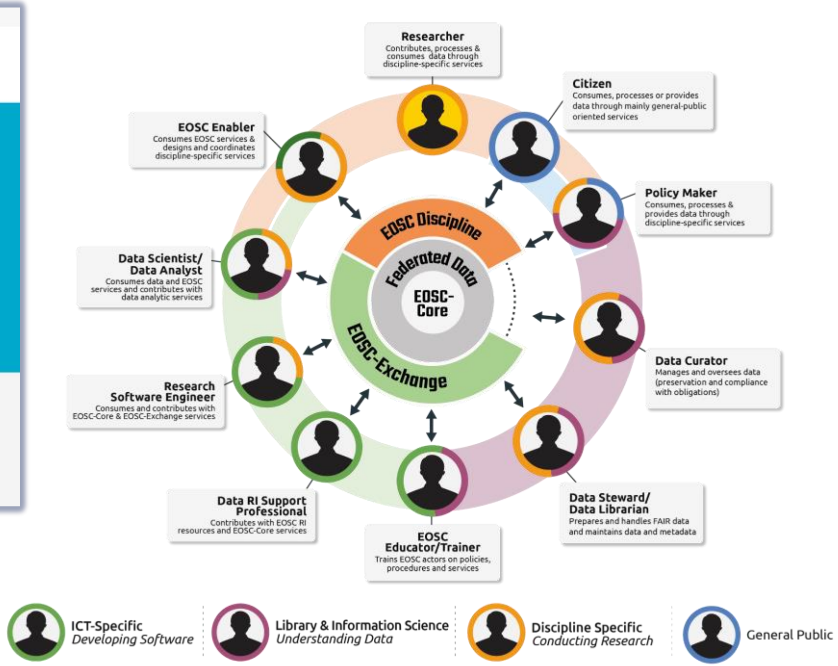
Connect spatial information communities and their data using a modern architecture, which is at the same time powerful and low cost, based on International and Open

View more...

♡0 □ ADD TO COMPARE 0

GeoNetwork

BLUEBRIDGE


Researcher
Contributes, processes & consumes data through discipline-specific services
Citizen
Consumes, processes or provides data through mainly general-public oriented services
EOSC Enabler
Consumes EOSC services & designs and coordinates discipline-specific services
Policy Maker
Consumes, processes & provides data through discipline-specific services
Data Scientist/ Data Analyst
Consumes data and EOSC services and contributes with data analytic services
EOSC Discipline
Federated Data
EOSC-Core
EOSC-Exchange
Data Curator
Manages and oversees data (preservation and compliance with obligations)
Research Software Engineer
Consumes and contributes with EOSC-Core & EOSC-Exchange services
Data RI Support Professional
Contributes with EOSC RI resources and EOSC-Core services
EOSC Educator/Trainer
Trains EOSC actors on policies, procedures and services
Data Steward/ Data Librarian
Prepares and handles FAIR data and maintains data and metadata
ICT-Specific
Developing Software
Library & Information Science
Understanding Data
Discipline Specific
Conducting Research
General Public

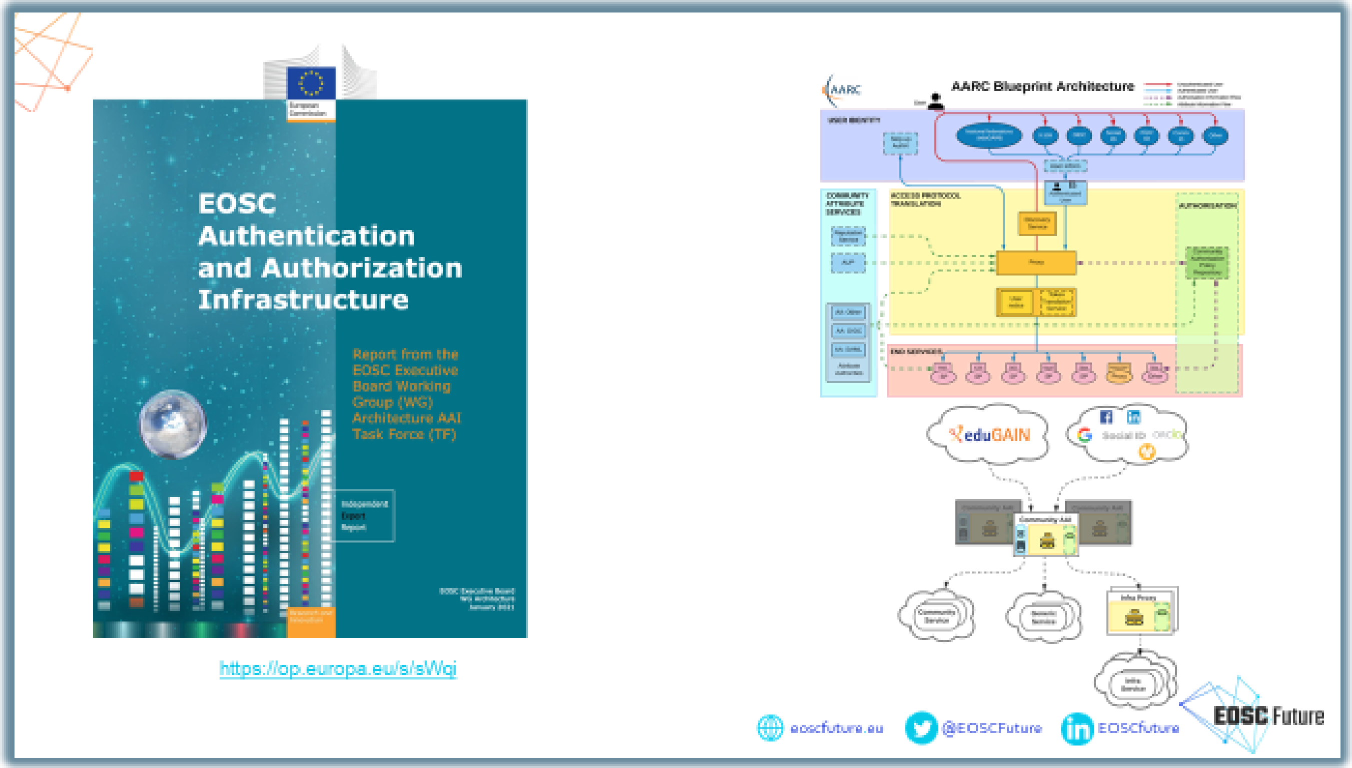
Circle diagram from Ignacio Blanquer’s ISGC 2022 keynote, Digital Skills for FAIR  and open science: doi.org/10.2777/59065; EOSC Portal (https://www.eosc-portal.eu/) by EOSChub

# Research Infra with AARC BPA proxies abound

SeaDataNet
BBMRI-ERIC®
WLCG
Worldwide LHC Computing Grid
DARIAH-EU
LIFE SCIENCE RI
EISCAT SCIENTIFIC ASSOCIATION
KM3NeT
SSHOC
social sciences & humanities open cloud
elixir
ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures
SKA
SQUARE KILOMETRE ARRAY
CLARIN
neic
PUHURI
cta
cherenkov telescope array
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) using the AAI proxy architecture. At the ~ bottom: (global) e-Infrastructures using the AARC BPA

# EOSC Authentication and Authorization Infrastructure



EOSC Authentication and Authorization Infrastructure

Report from the EOSC Executive Board Working Group (WG) Architecture AAI Task Force (TF)

Independent Expert Report

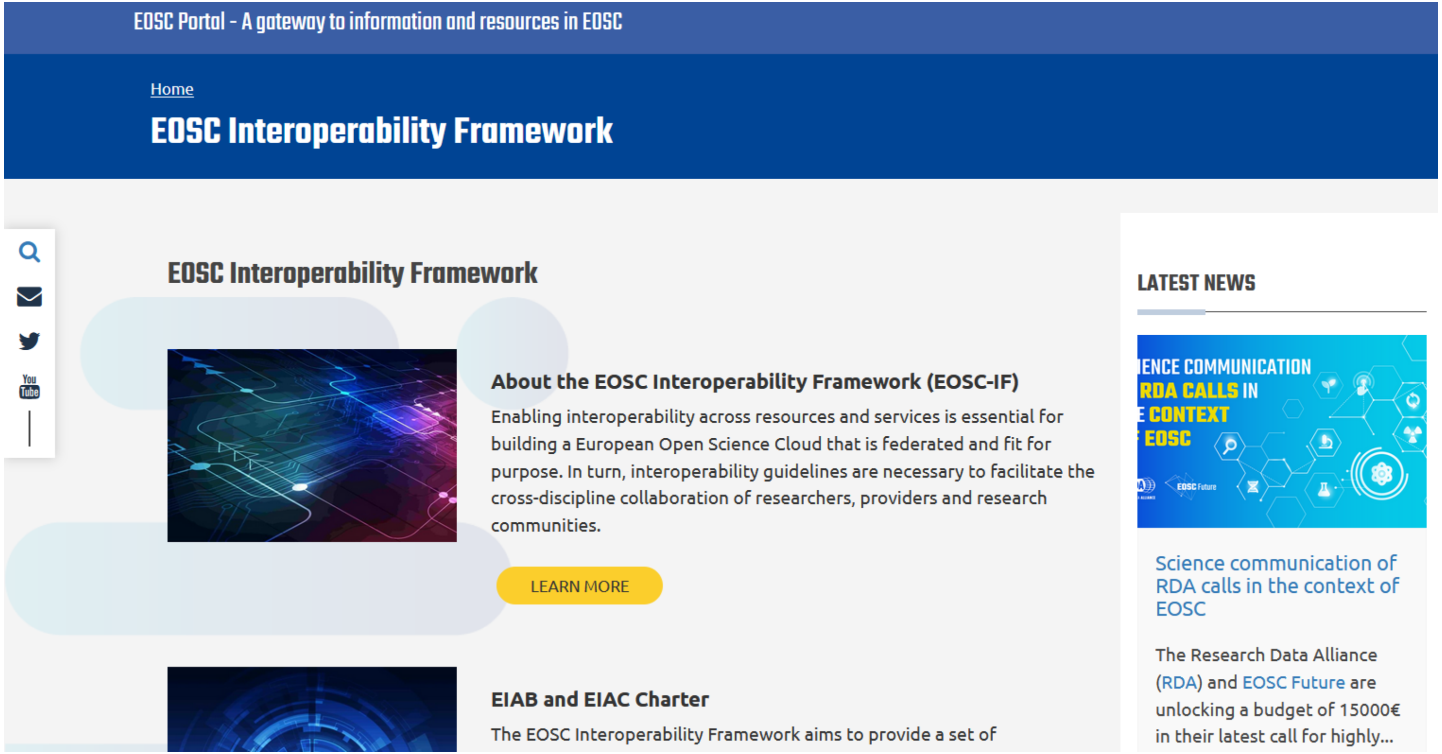
EOSC Executive Board WG Architecture January 2021

https://op.europa.eu/s/sWqi

eoscfuture.eu @EOSCFuture EOSCfuture

slide: Christos Kanellopoulos, GEANT, for EOSC Future WP7.3

# EOSC Interoperability Framework



EOSC Portal - A gateway to information and resources in EOSC

Home

EOSC Interoperability Framework

EOSC Interoperability Framework

**About the EOSC Interoperability Framework (EOSC-IF)**

Enabling interoperability across resources and services is essential for building a European Open Science Cloud that is federated and fit for purpose. In turn, interoperability guidelines are necessary to facilitate the cross-discipline collaboration of researchers, providers and research communities.

LEARN MORE

**EIAB and EIAC Charter**

The EOSC Interoperability Framework aims to provide a set of

LATEST NEWS

Science communication of RDA calls in the context of EOSC

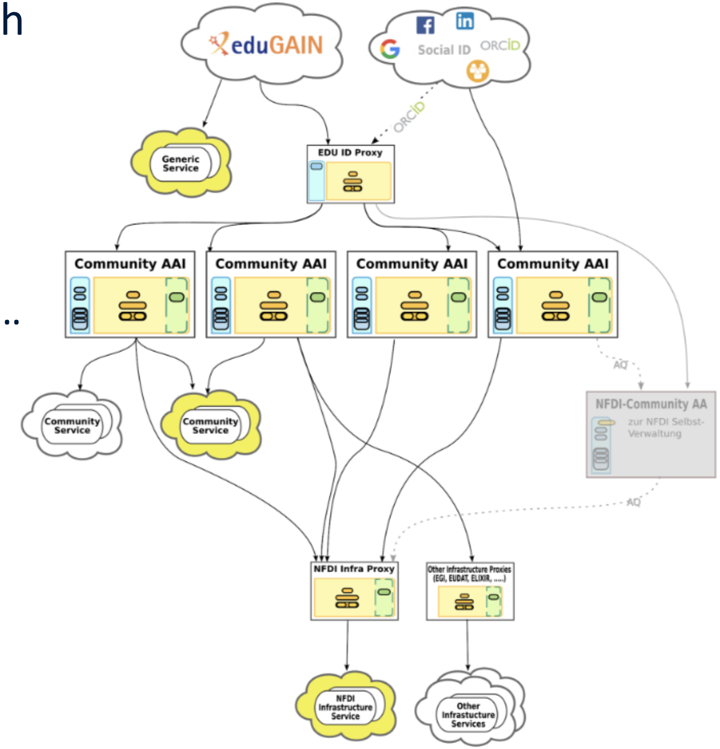
The Research Data Alliance (RDA) and EOSC Future are unlocking a budget of 15000€ in their latest call for highly...

https://eosc-portal.eu/eosc-interoperability-framework

# Composite AAIs – proxies beyond 'just' the EOSC

Proxy model supports harmonizing IdPs beyond research

- **eduID**-style identifiers
  - 'life-long learning' identifiers
  - independent student identifier for mobility & Erasmus-without-papers
  - eduGAIN-alignment is coming: eduid.nl, Swiss eduID, …
- **eIDAS** and government eID (e.g. DigID)
  - identity assurance step-up
- **ORCID** provides this service for research in general
  - since it persists, also very useful to allow researchers consistent access independent of home org ☺



Composite AAI image source: Christos Kanellopoulos (GEANT), Marcus Hardt (KIT)

# Collaboration and sharing is critical for research

"Authentication and Authorisation Infrastructures (AAIs) play a key role in enabling federated interoperable access to resources."

Proposed: **AARC Technical Revision to Enhance Effectiveness (AARC TREE)**

- define common strategies for the development and deployment of AAIs in the pan European Research Infrastructures
- improve access and sharing of scientific resources and
- improve interoperability among research infrastructure communities across the thematic areas

https://aarc-community.org/ - the AARC Technical Revision for Enhanced Effectiveness (AARC TREE) submitted as CSA to INFRA-DEV-01-05 on March 9, 2023

# Design Patterns in e-Infrastructures?

# So can we now discern a common pattern?

- Make central components passive and as stateless as possible
  - e.g. for fabric management, have central repository be a cacheable web service
  - although persistent storage obviously has to retain some state ☺

- Move complexity and volume requirements to the edge
  - the edge scales horizontally and scaling from 2+ is much easier than from 1→ 2

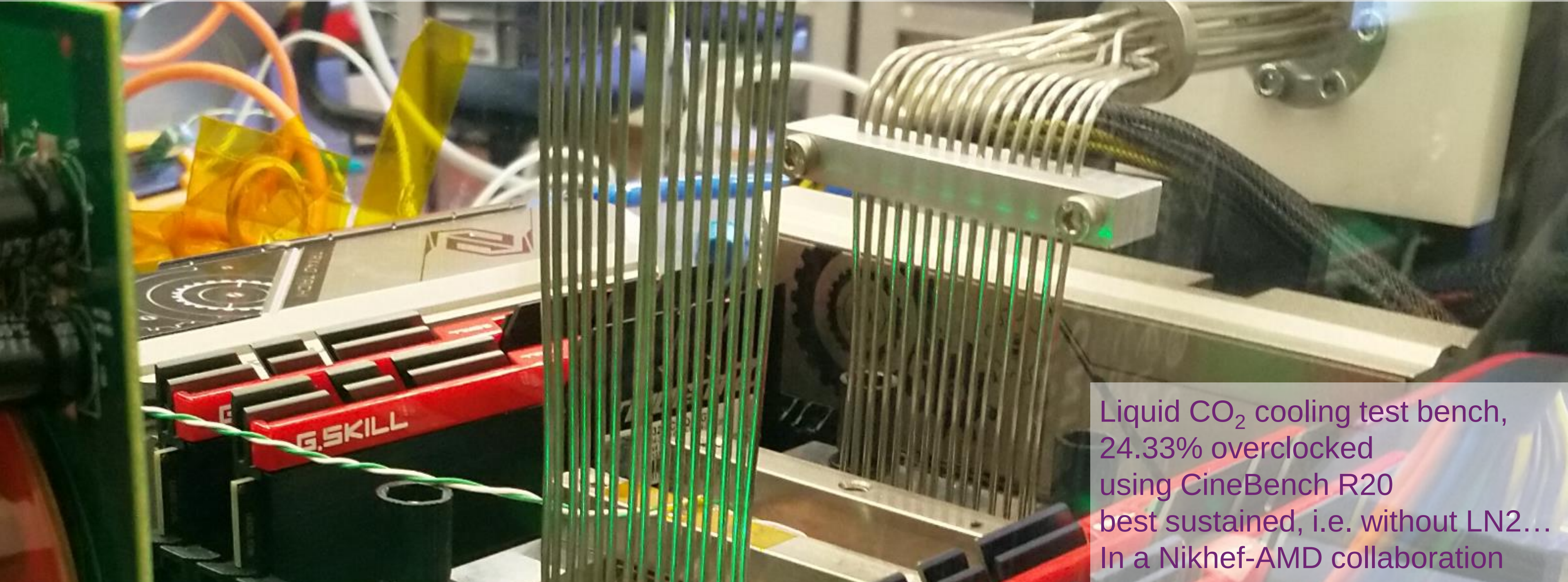
- You can move problems around, but it’s hard to actually *solve* them
  - e.g. lack of a single common interface implies one needs adaptors and plugins

- Scaling *collaboration and trust* federation is as complex as scaling systems
  - and beyond ‘Dunbar’s Number’, ~150, you will need some assessment and policy

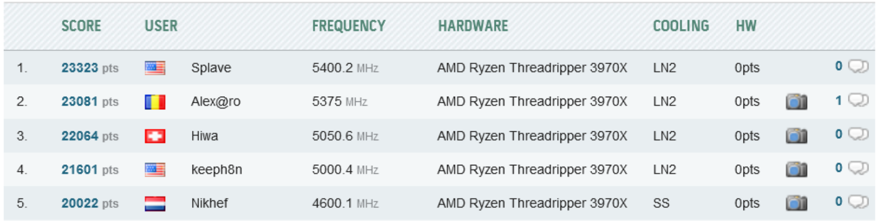
Nikhef

# … since some things are fun, but not quite *that* scalable …

AMD



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

This work has also been co-supported by projects that have received funding from the European Union's Horizon research and innovation programmes under Grant Agreement No. 856726 (GN4-3), 101017536 (EOSC Future), 777536 (EOSC-hub), 730941 (AARC2).

# Still here? Thanks!

*works presented here are part of many global collaborations. Thanks especially to Mischa Sallé, Tristan Suerink, Dennis van Dok, Mary Hester, Andrew Pickford, Jeff Templon, Roel Aaij, Arjen van Rijn, Krista de Roo (Nikhef), Nicolas Liampotis, Kyriakos Glinis (GRNET), Jens Jensen, Dave Kelsey, David Crooks (STFC RAL), Hannah Short (CERN), Uros Stevanovic and Marcus Hardt (KIT), Maarten Kremers (SURF), Licia Florio (GEANT and NORDUNET), Christos Kanellopoulos, Klaas Wieringa (GEANT), Tom Barton (Internet2), Tiziana Ferrari, Matt Viljoen, Baptise Grenir (EGI.eu), and the EGI, EUDAT, PRACE, GEANT, WLCG, REFEDS, WISE, IGTF, and FIM4R communities!*

David Groep
davidg@nikhef.nl
https://www.nikhef.nl/~davidg/presentations/
iD https://orcid.org/0000-0003-1026-6606

Maastricht University

Nikhef