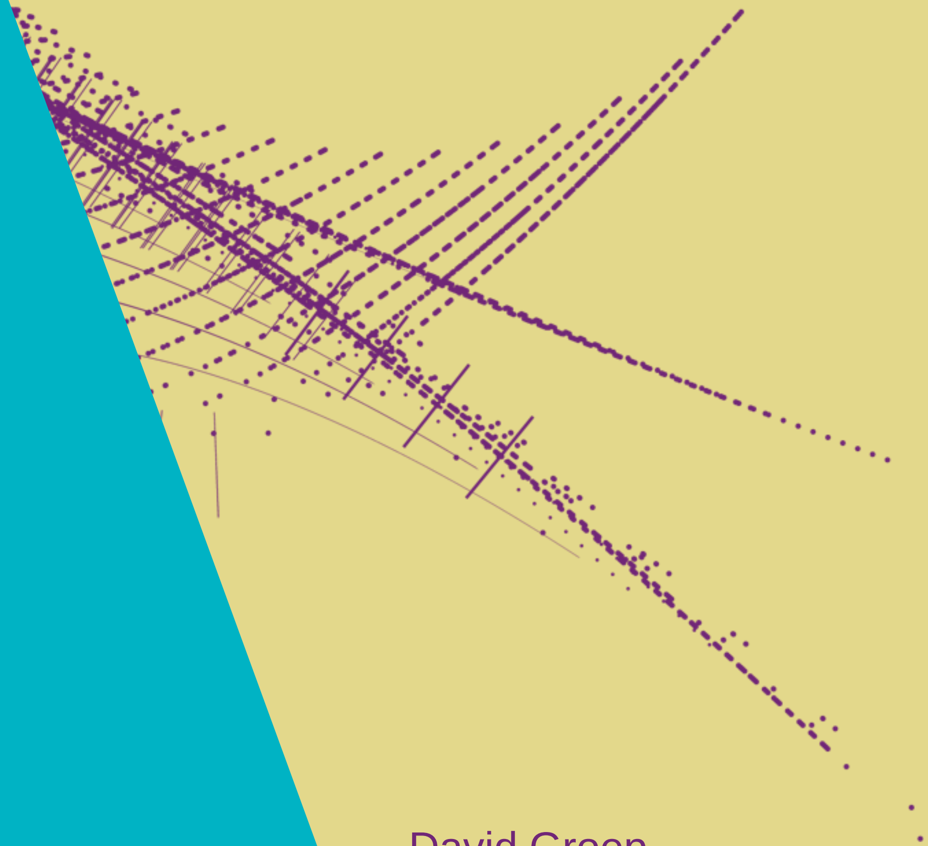
Nikhef

Maastricht University

Introduction to Nikhef,
Computing and Networking

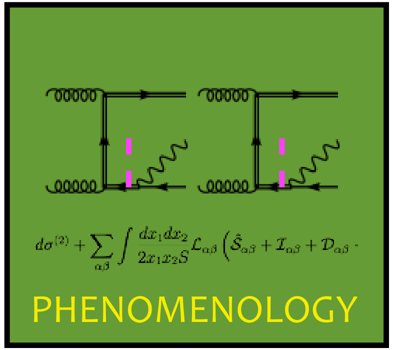
Welcome to Nikhef

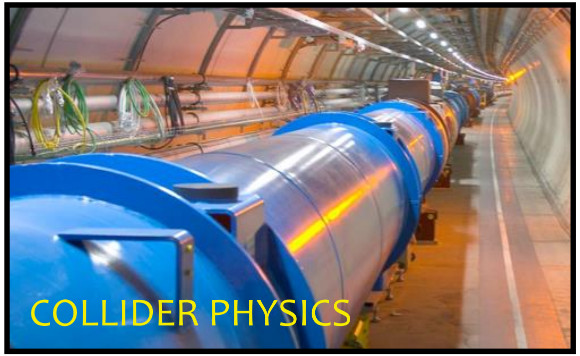


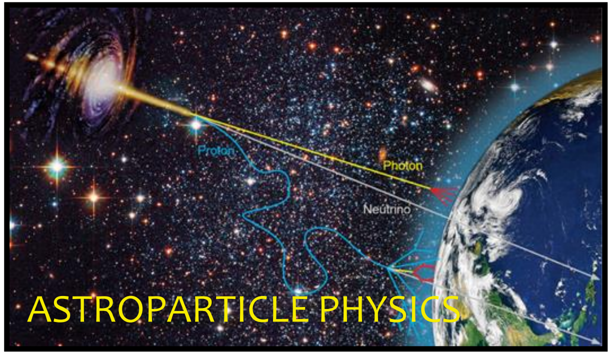
David Groep
March 2024

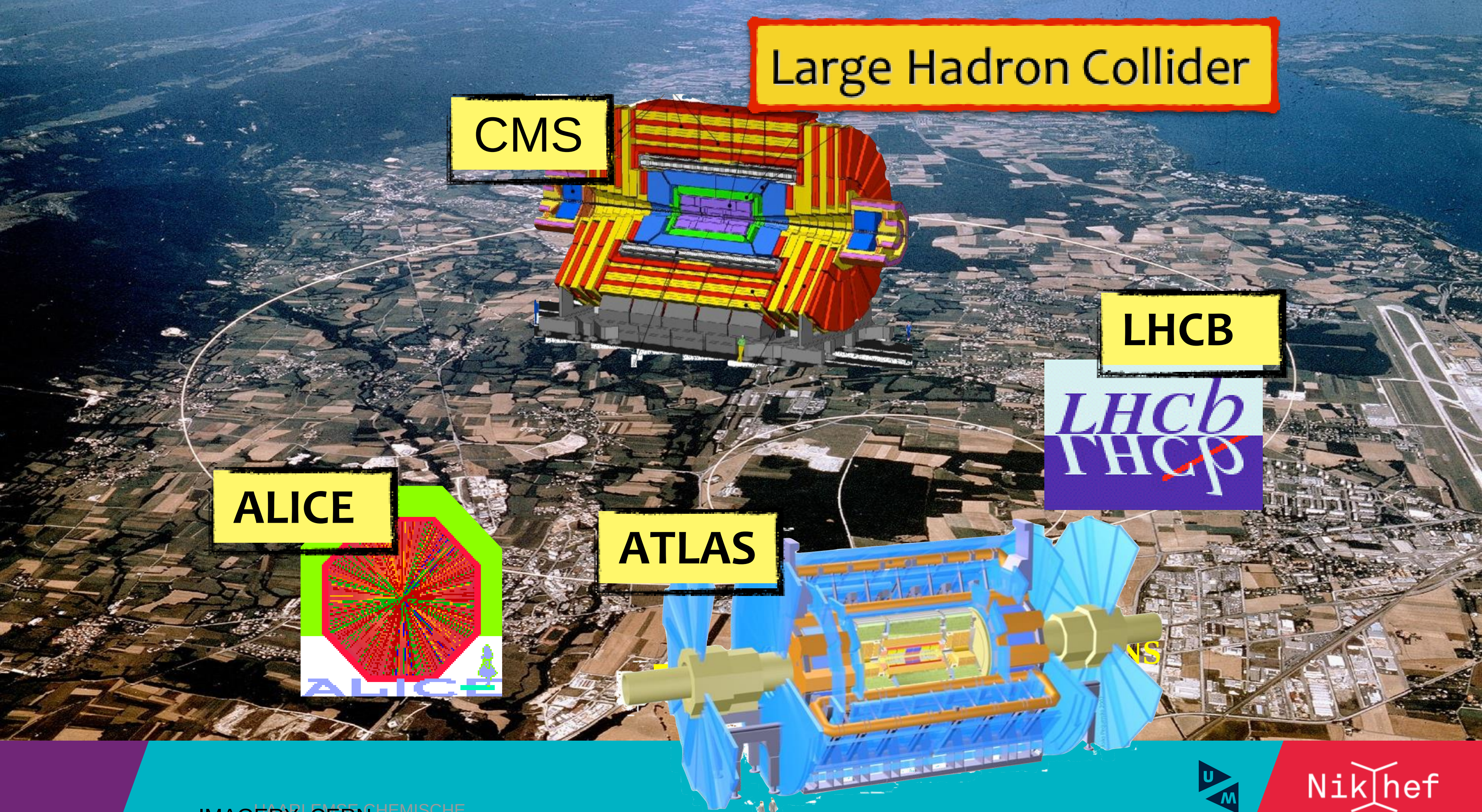
Verleggen van de grenzen van onze kennis

- **Accelerator-based particle physics**
  Experiments studying interactions in particle collision processes at particle accelerators, in particular at CERN;
- **Astroparticle physics**
  Experiments studying interactions of particles and radiation emanating from the Universe.

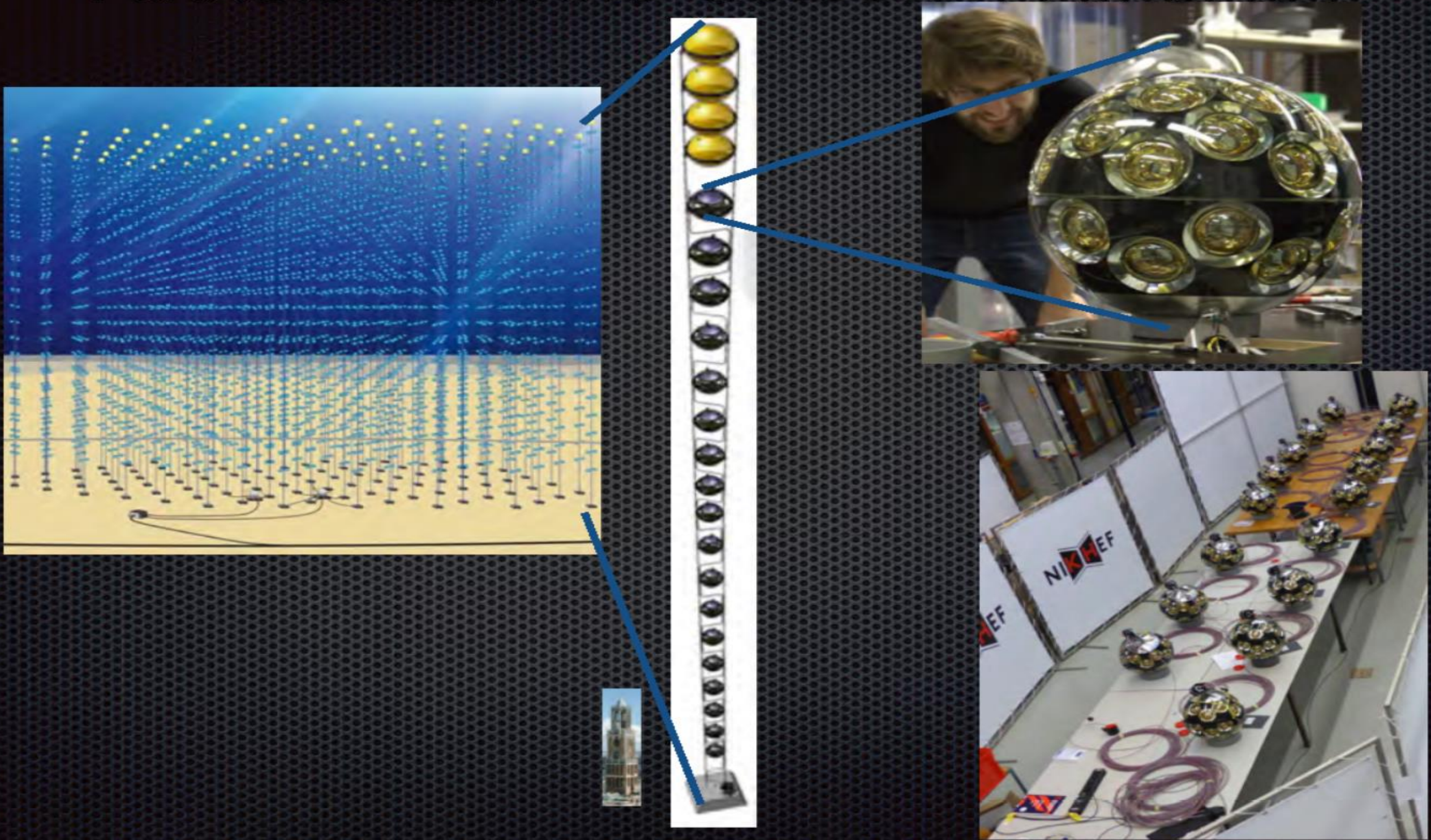




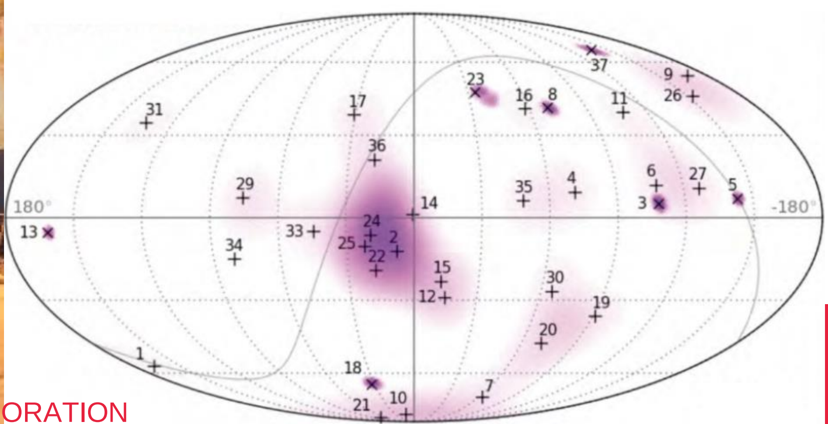



Large Hadron Collider
CMS
LHCB
LHCb
ALICE
ATLAS

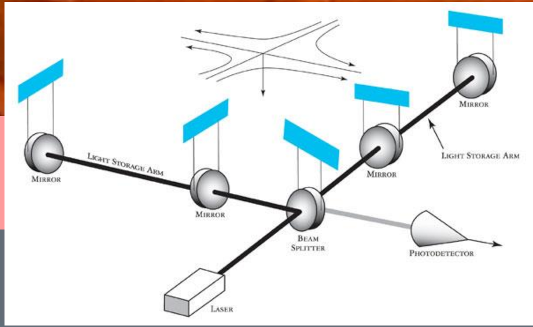
Nikhefs neutrino-detector: KM3NeT

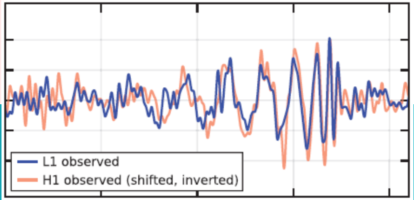


De Melkweg



IMAGERY: NIKHEF, NIOZ, KM3NET COLLABORATION



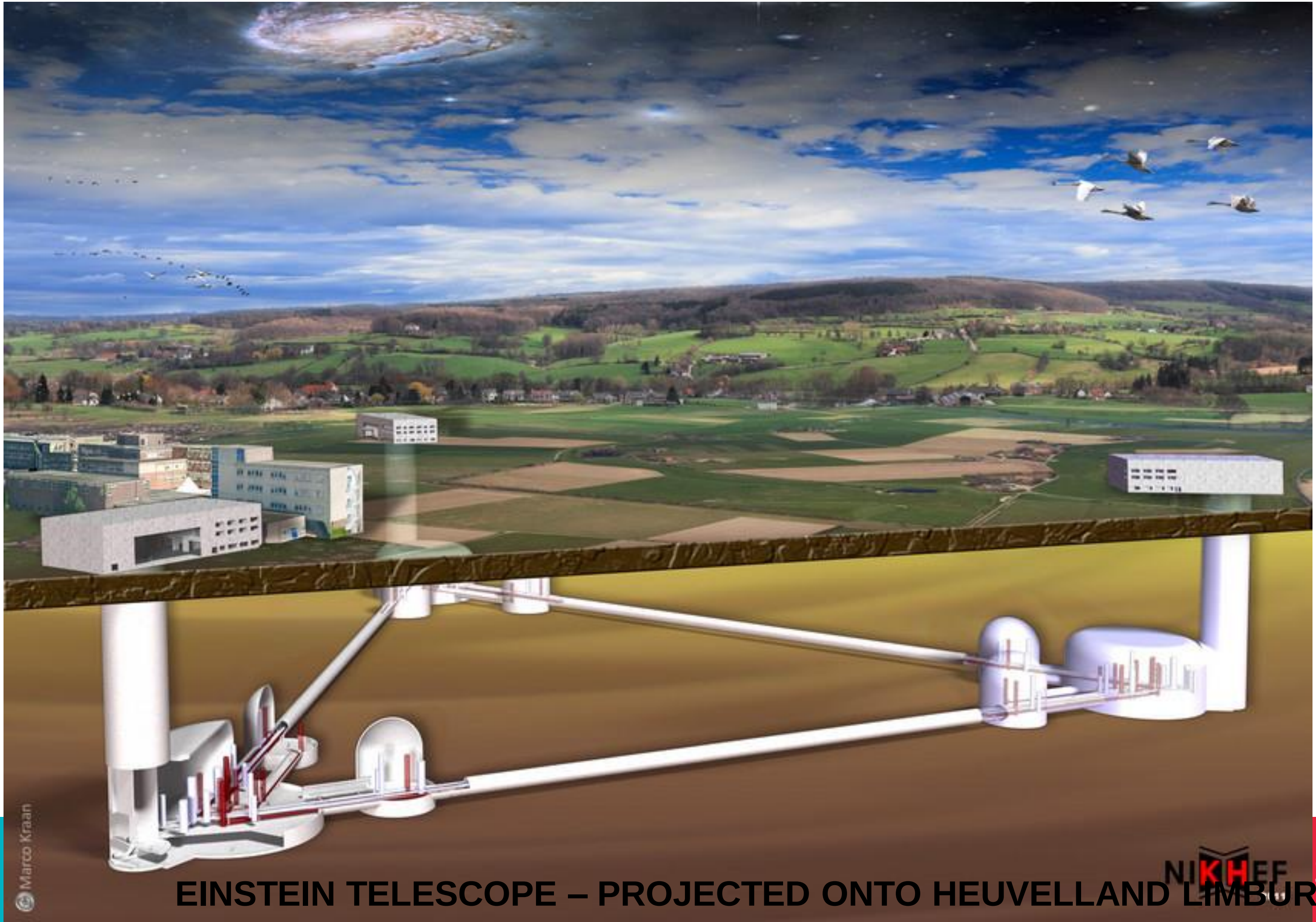


LSC LIGO Scientific Collaboration

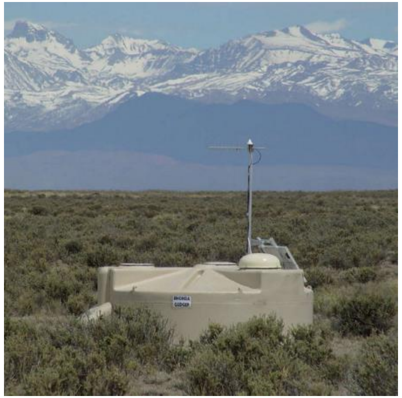
VIRGO

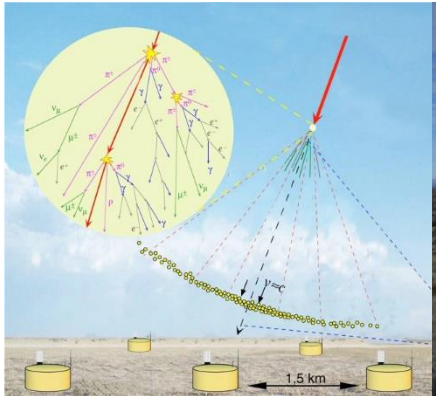
Nikhef

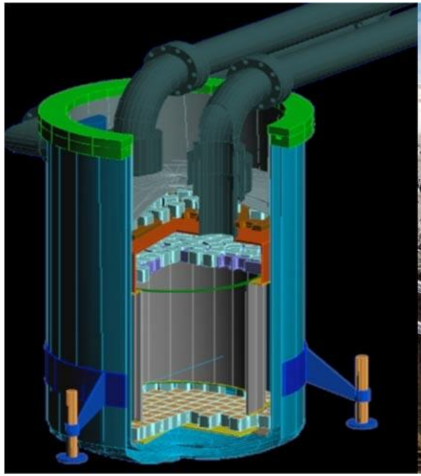
IMAGERY: GW-ASTRONOMY CO

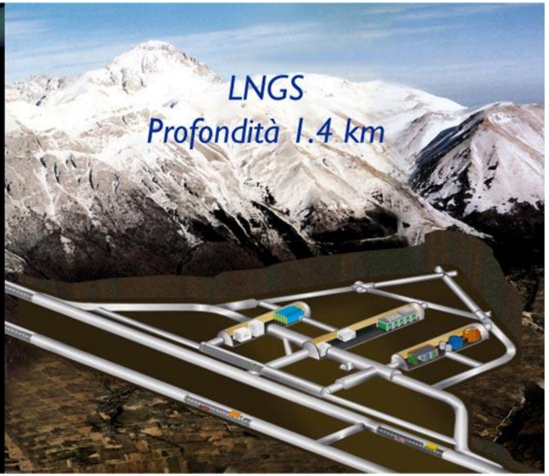


EINSTEIN TELESCOPE – PROJECTED ONTO HEUVELLAND LIMBURG, NL




1,5 km

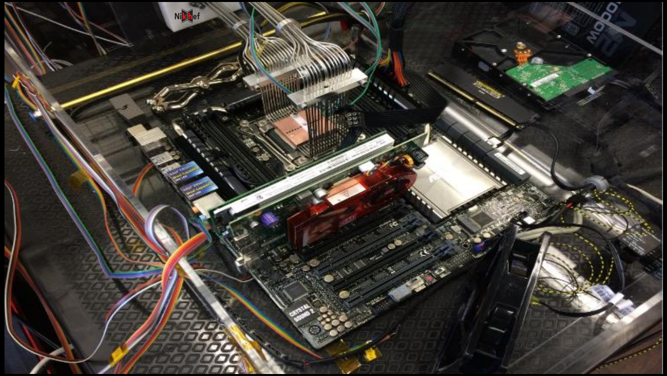



LNGS
Profondità 1.4 km

Nikhef collaboration – underpinning programmes

Detector R&D

Theoretical Physics



Physics Data Processing

PDP 'Physics Data Processing' activities and action lines

Scalable Computing & Algorithms

- algorithms for high-performance software
- data organisation
- accelerator throughput
- rethinking code: at small scale and on large scales

Research Infrastructure & Future Technologies

- NDPF processing facility
- National e-Infrastructure *coordinated by SURF*
- experimental next-gen systems engineering
- cross-tier global networks
- stressing & public/private collaborative design

Infrastructure for Secure Collaboration

- authentication and authorization protocols
- multi-domain federation
- global trust and identity
- access provisioning
- operational security

Data at the Large Hadron Collider at CERN

1964


VOLUME 13, NUMBER 16 PHYSICAL REVIEW LETTERS 19 OCTOBER 1964

BROKEN SYMMETRIES AND THE MASSES OF GAUGE BOSONS

Peter W. Higgs
Tait Institute of Mathematical Physics, University of Edinburgh, Edinburgh, Scotland
(Received 31 August 1964)

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

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

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

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_μ. Let us suppose that $V'(\varphi_0^2) = 0$, $V''(\varphi_0^2) > 0$; then spontaneous breakdown of U(1) symmetry occurs. Consider the equations [derived from (1) by treating $\Delta\varphi_1$, $\Delta\varphi_2$, and A_μ as small quantities] governing the propagation of small oscillations about the "vacuum" solution $\varphi_1(x) = 0$, $\varphi_2(x) = \varphi_0$:

$$\partial^\mu\{\partial_\mu(\Delta\varphi_1) - e\varphi_0 A_\mu\} = 0, \quad (2a)$$

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

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

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

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

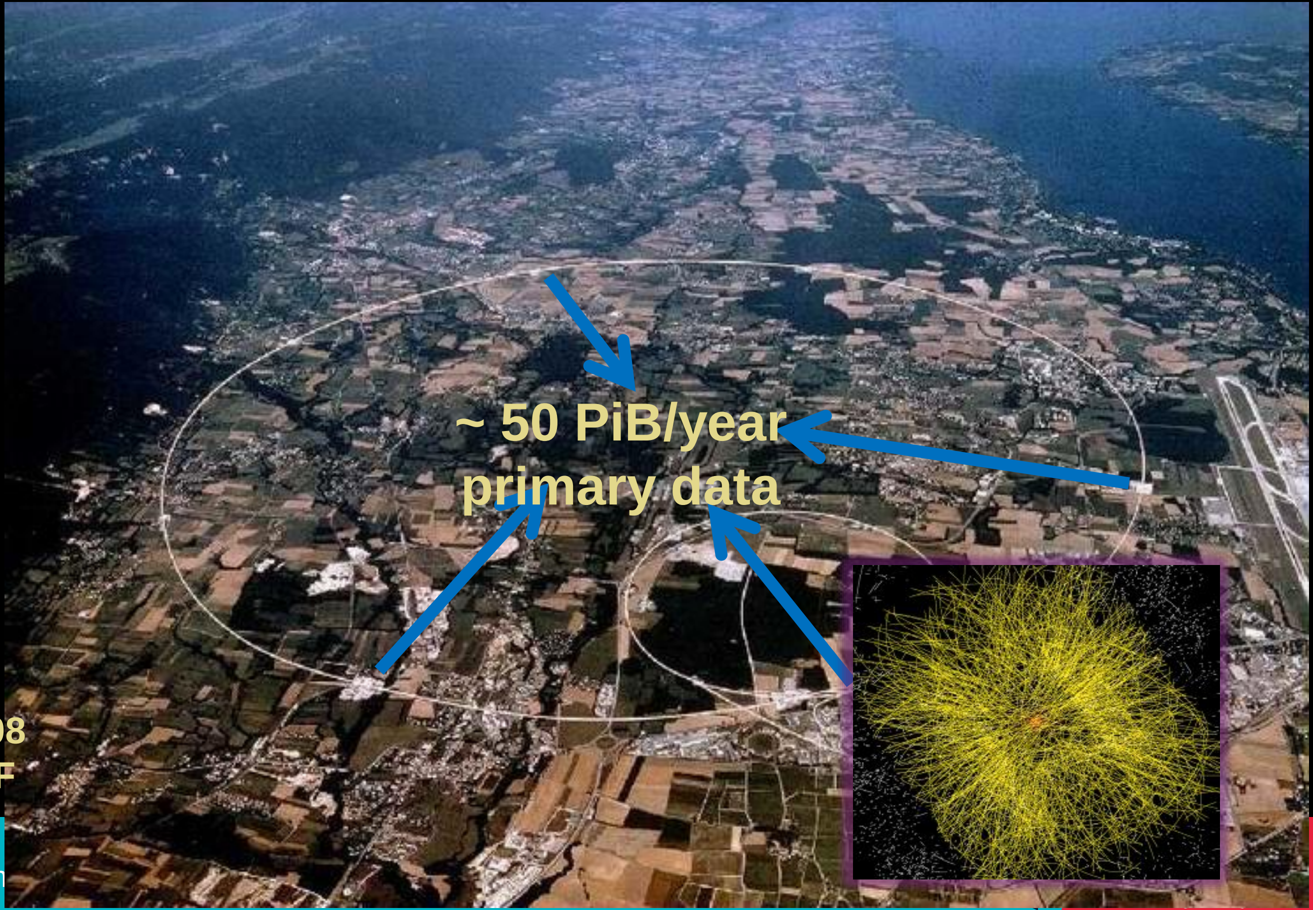
Equation (4) describes vector waves whose quanta have (bare) mass $e\varphi_0$. In the absence of the gauge field coupling ($e = 0$) the situation is quite different: Equations (2a) and (2c) describe zero-mass scalar and vector bosons, respectively. In passing, we note that the right-hand side of (2c) 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

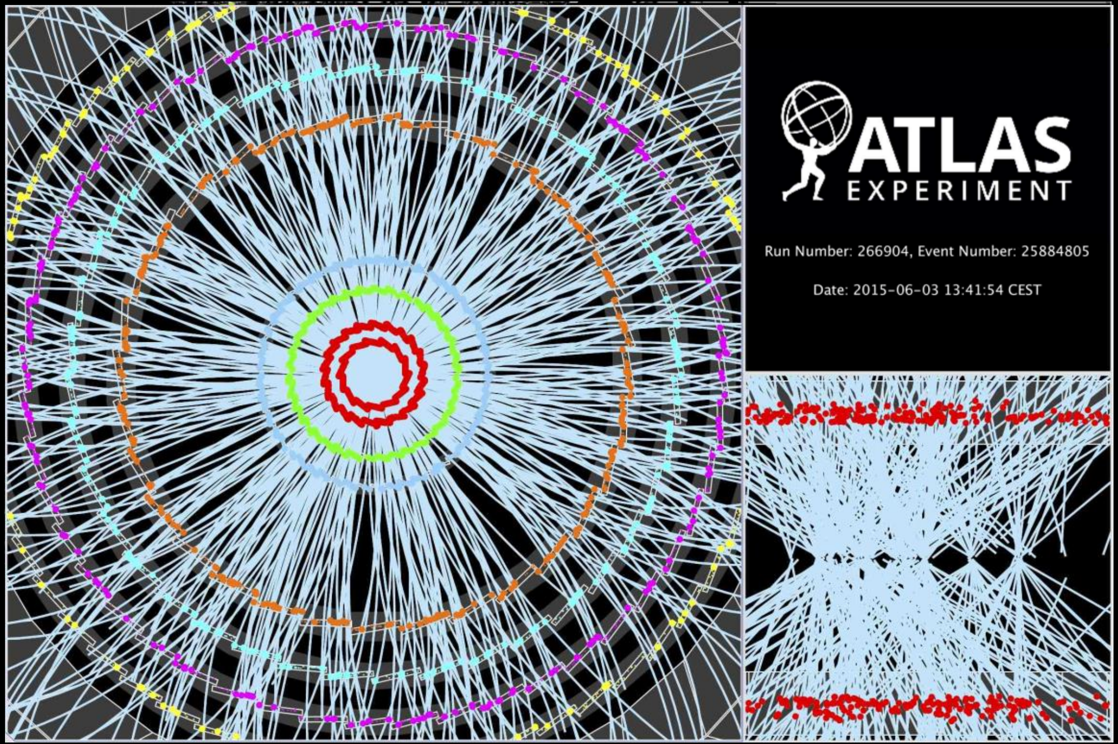
321

P. Higgs, Phys. Rev. Lett. 13, 508
16823 characters, 165kByte PDF



Computing on lots of data – 40 Mevents/sec

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

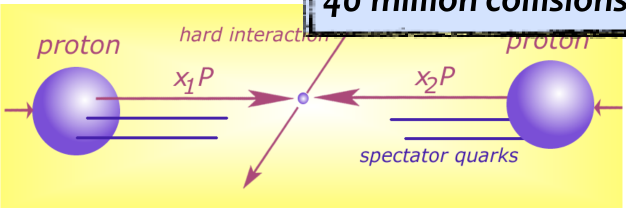


Display of a proton-proton collision event recorded by ATLAS on 3 June 2015, with the first LHC stable beams at a collision energy of 13 TeV; Event processing time: v19.0.1.1 as per Jovan Mitrevski and 2015 J. Phys.: Conf. Ser. 664 072034 (CHEP2015)

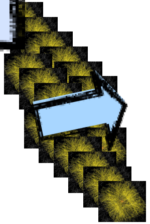
‘Big Science’ needs some computing …

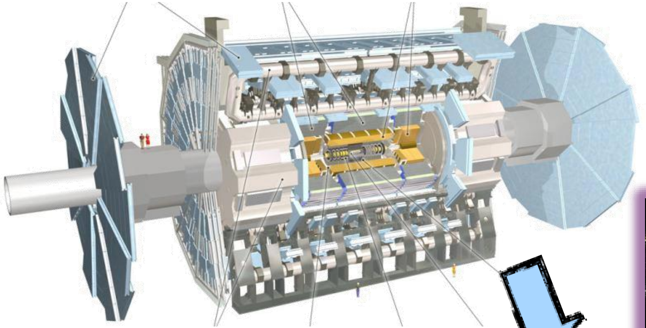
CERN CC B513, image: https://cds.cern.ch/record/2127440; tape library: CC-IN2P3 with LHC and LSST data; cabinets: Nikhef H234b

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

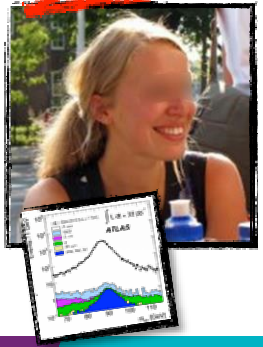


diagram adapted from Frank Linde; images: ATLAS collaboration, Nikhef. … and sorry for the GDPR-blur

Single CPU scaling stopped around 2004

limitation is power, not circuit size

and clock frequency is most 'power-hungry'
still some packages now @ TDP of 400W

multiple cores on the same die helped

AMD EPYC Genoa (Zen 4) has 96 cores on die
Intel Cascade Lake AP looked like a cludge
but now Sapphire Rapids appears better again

CPU design-performance gains left

predictive execution
out-of-order execution
on-die parallelism (multi-core)
pre-fetching and multi-tier caching
execution unit sharing ('SMT')
but at increased risk for security/integrity

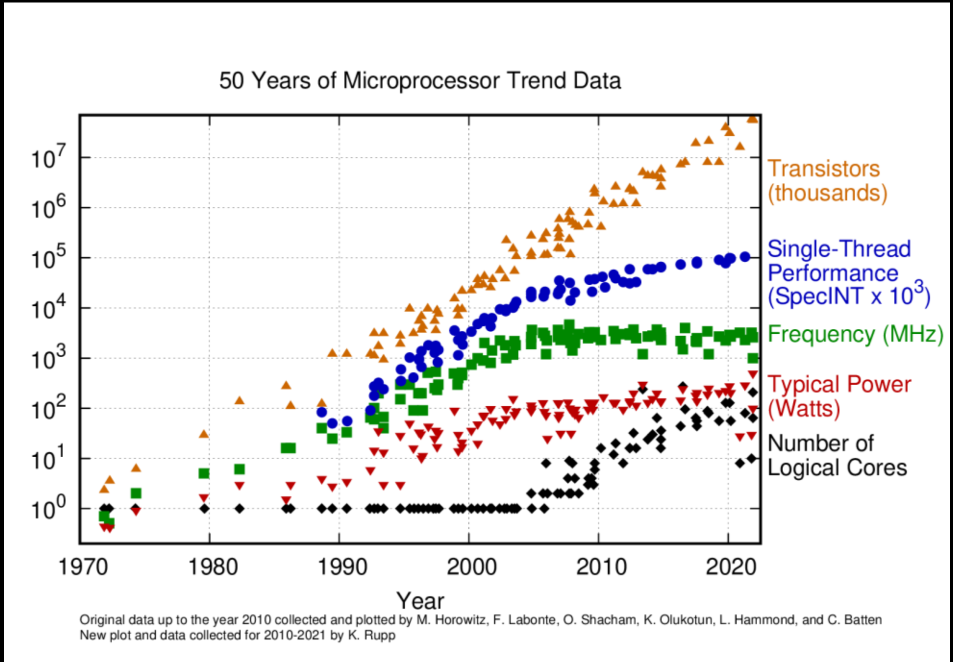
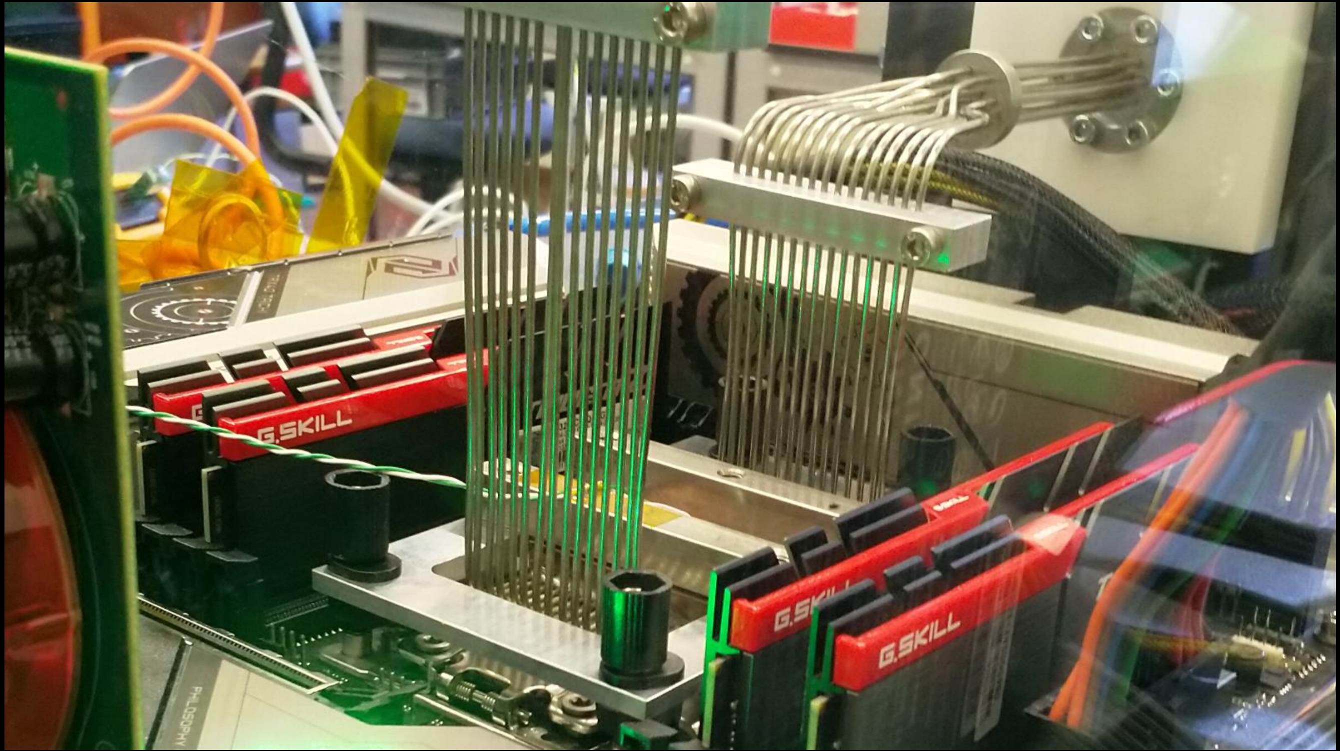


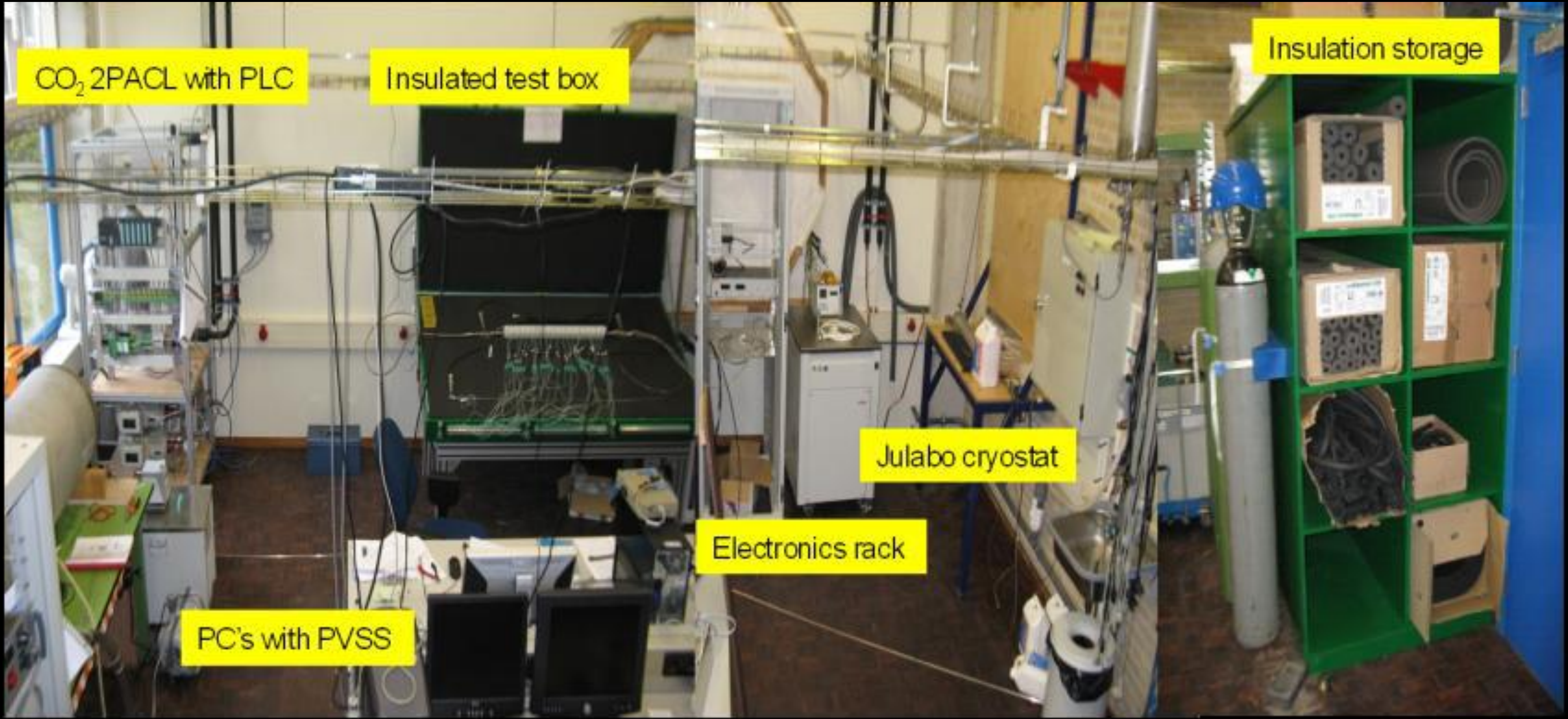
Image: K Rupp, https://github.com/karlrupp/microprocessor-trend-data

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



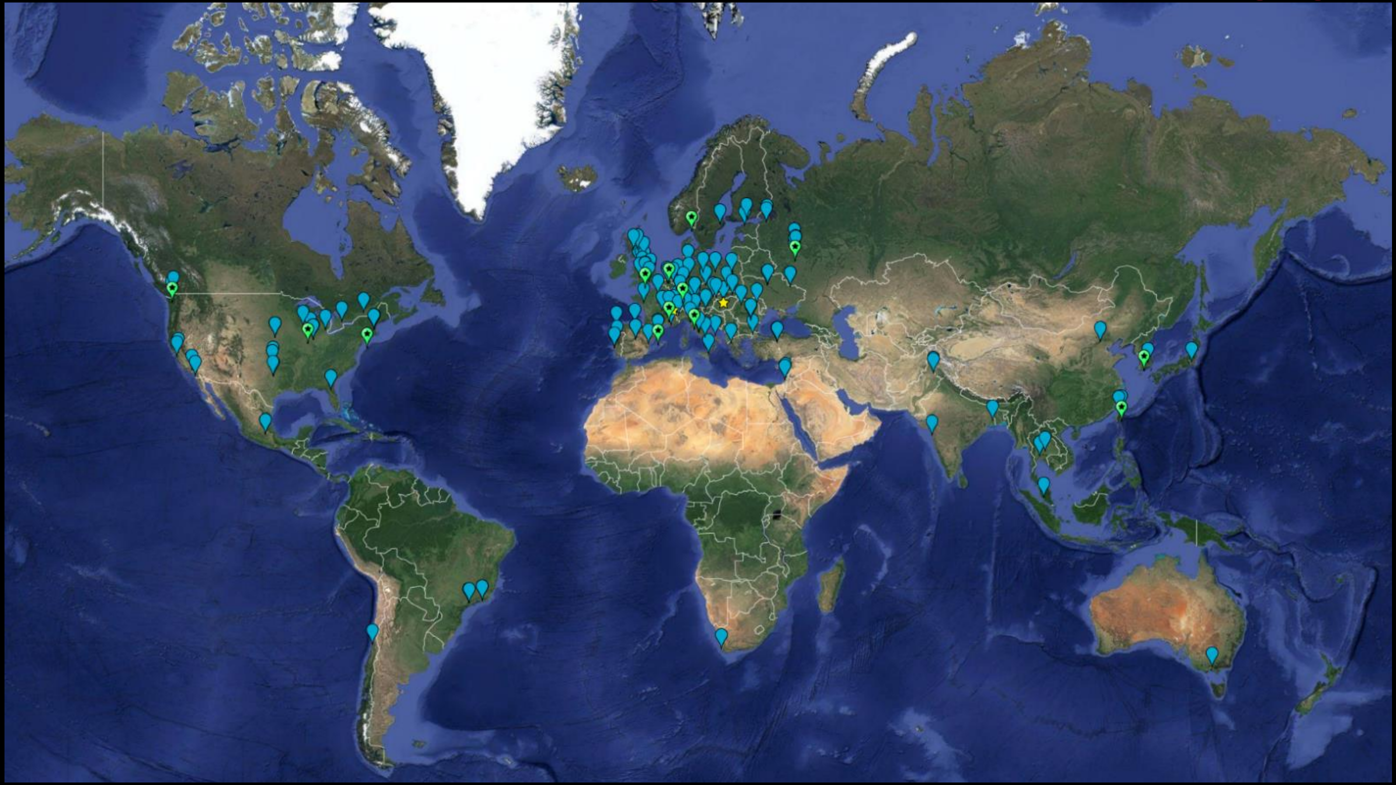
LCO2 cooling of an AMD Ryzen Threadripper 3970X [56.38 °C] at 4600.1MHz processor (~1.5x nominal speed) sustained, using the Nikhef LCO2 test bench system (https://hwbot.org/submission/4539341) - (Krista de Roo en Tristan Suerink)

… since you then need this around it …



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

With 20 000+ users, you need something global: WLCG!



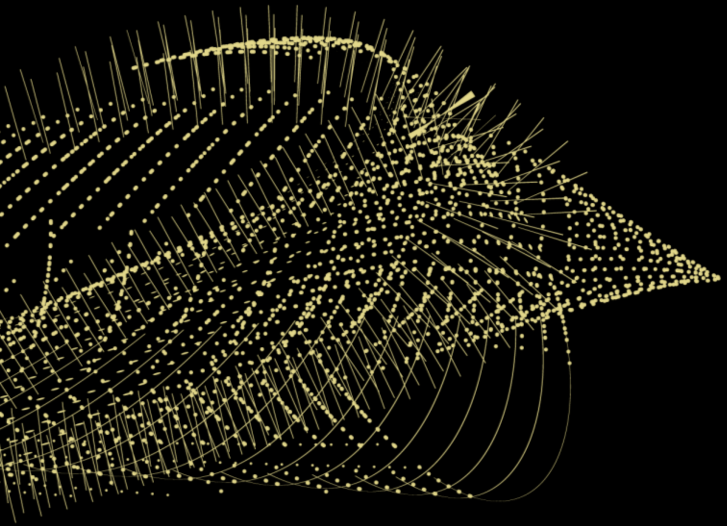
~ 1.4 million CPU cores
~ 1500 Petabyte
disk + archival

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

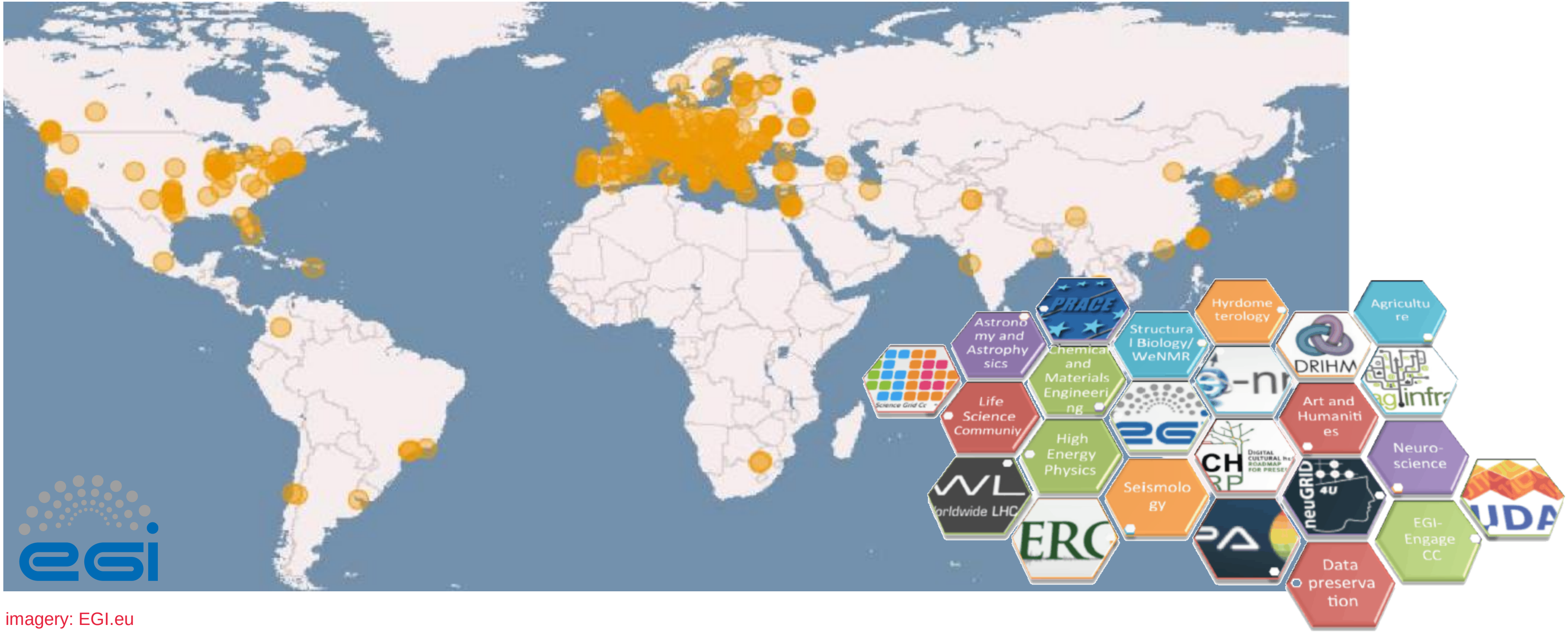


> 1 user group
> 1 organization

> 1 system
> 1 site

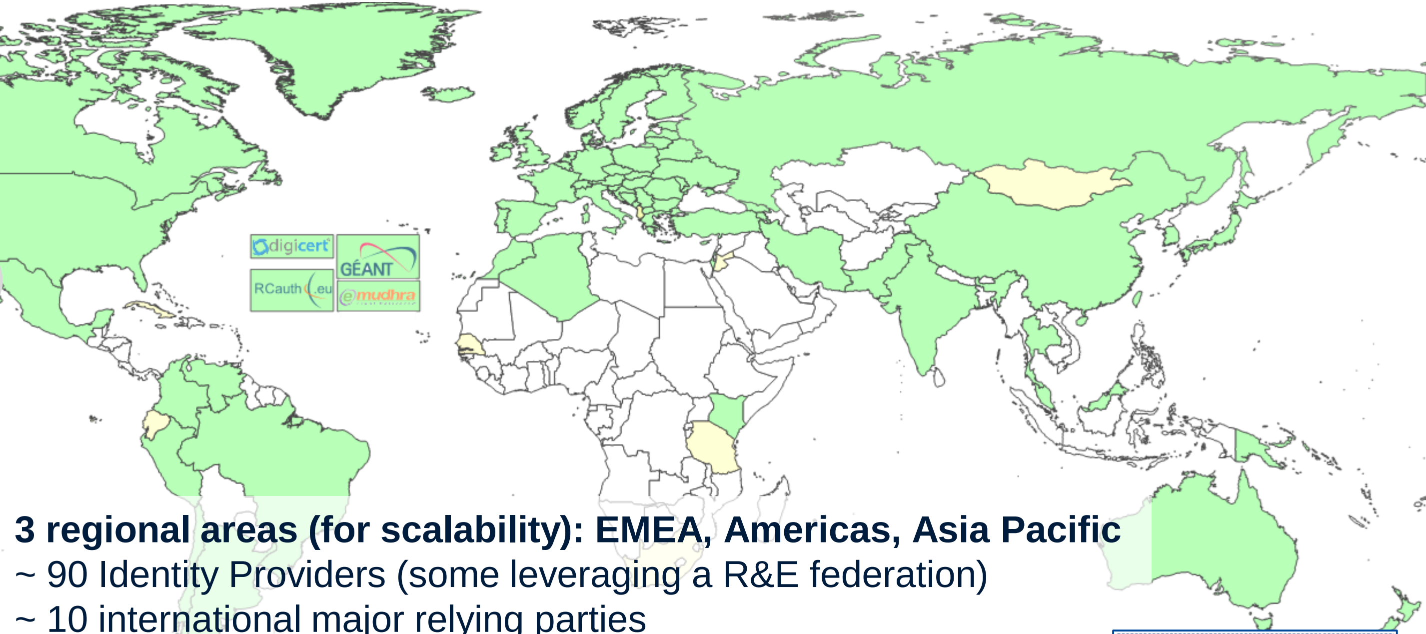
More of more than one …
@Nikhef

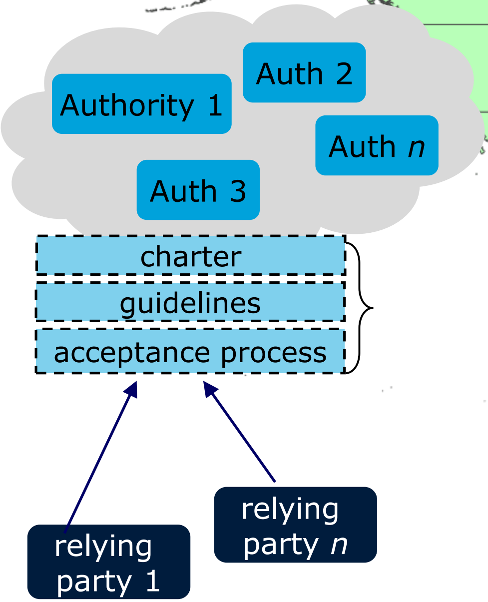
E-INFRASTRUCTURES: EGI, EUDAT, GEANT, PRACE, …



imagery: EGI.eu

Policy-bridged global federations for research computing





3 regional areas (for scalability): 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

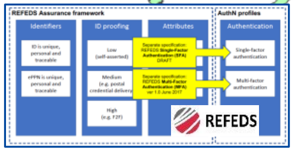
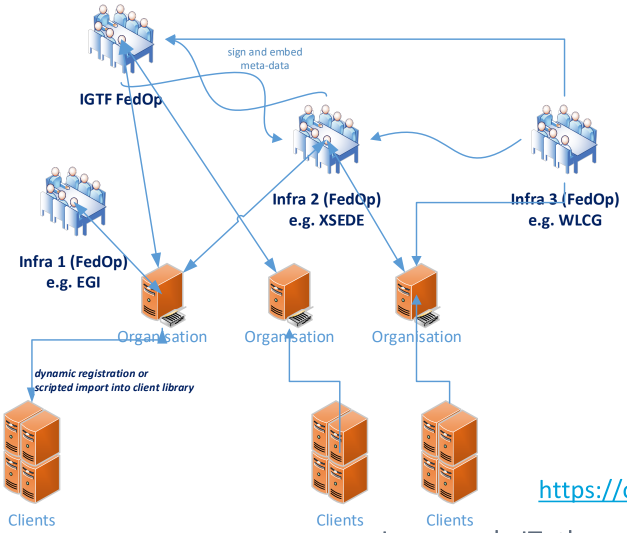


Image: Interoperable Global Trust Federation IGTF, https://igtf.net/; REFEDS Assurance Framework RAF: http://refeds.org/assurance, https://refeds.org/profile/mfa

OpenID Connect Federation

OIDC endpoints + trust policy data for registration can be federated in a meta-data feed

- makes OIDC ‘federatable’ (plain oidc is single OP)
- as for PKIX, can be technical or policy bridge
- delegated metadata makes ‘OIDC-fed’ scale in webscale scenarios



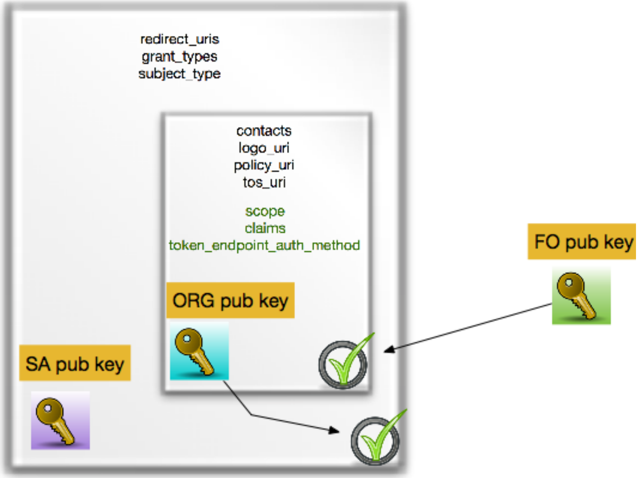
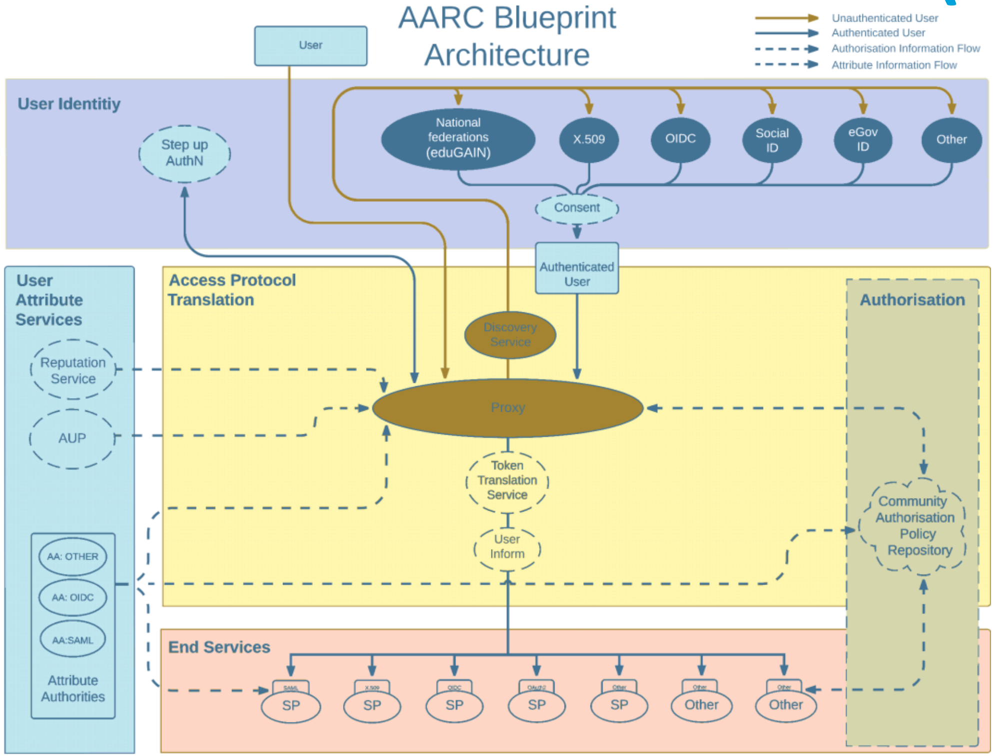
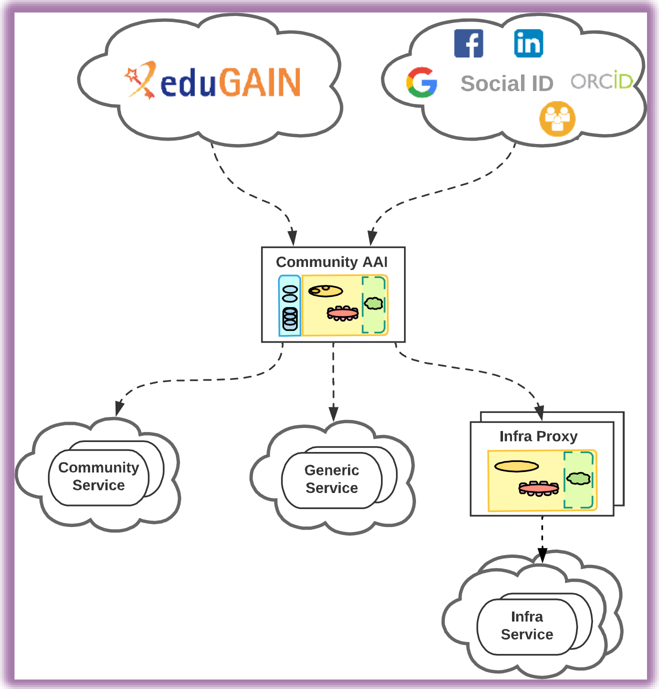


Image: Roland Hedberg, University of Umeå
OpenID Connect Fedrration: https://openid.net/specs/openid-connect-federation-1_0.html

Most trust flows from the (research) community

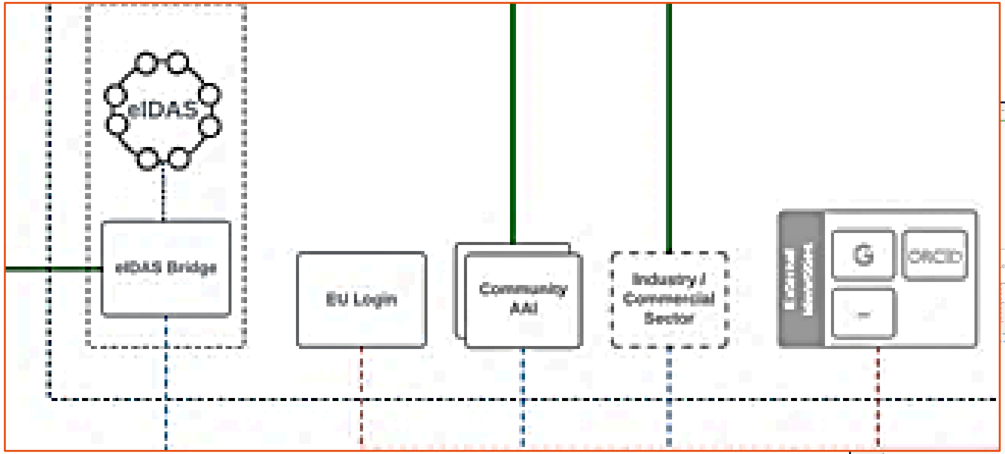


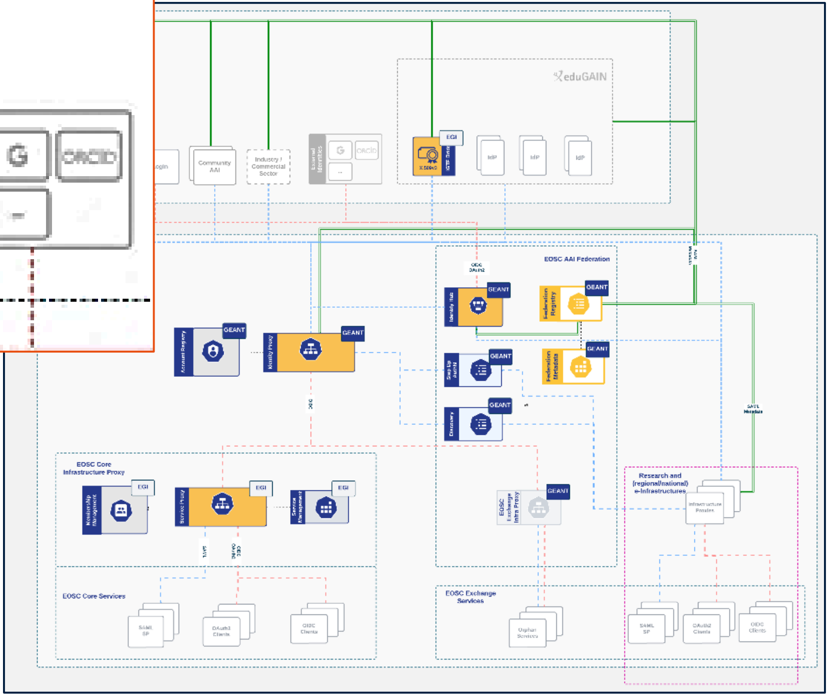


AARC Blueprint Architecture (2019) AARC-G045 https://aarc-community.org/guidelines/aarc-g045/; stacked proxies: EOSC AAI Architecture
EOSC Authentication and Authorization Infrastructure (AAI), ISBN 978-92-76-28113-9, http://doi.org/10.2777/8702

EOSC AAI Federation

***Identity assurance* brings the true value: authenticators are aplenty, and 'MFA' far less interesting than vetted identities. But HEI home IdPs seem reluctant to provide it …**

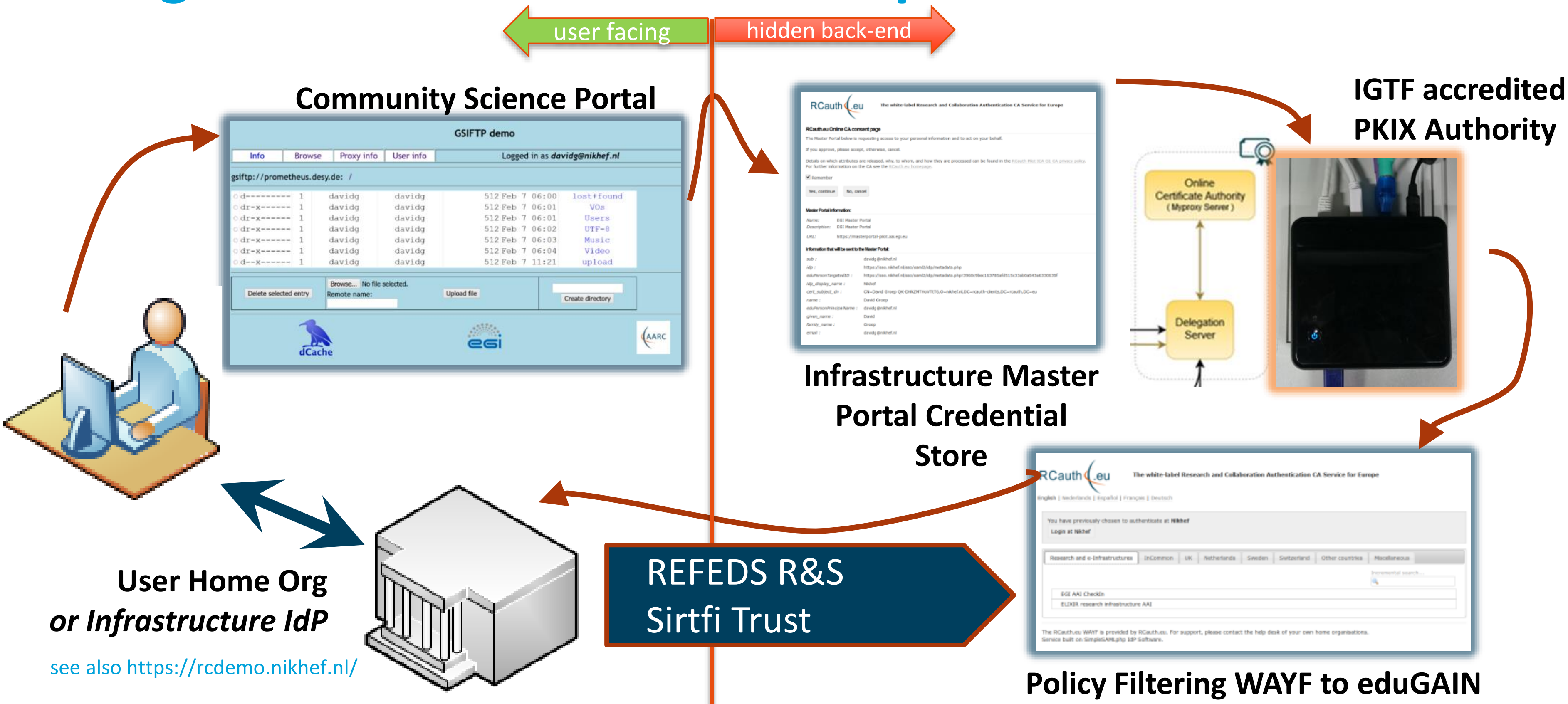




user identity comes 'with the user' from outside, mediated by the research community, ORCID, or from the home member state involved

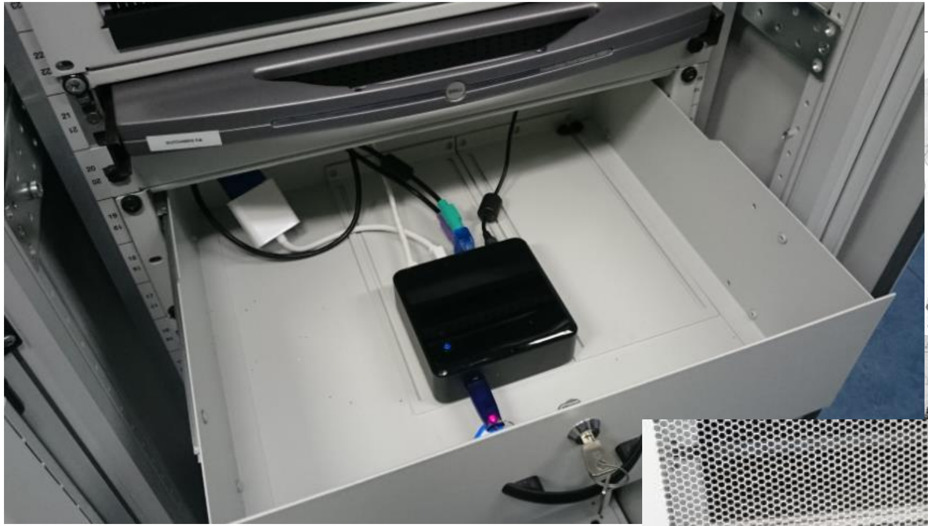
Image: EOSC AAI for the EOSC Core and Exchange Federation for the EOSC European Node by Christos Kanellopoulos, Nicolas Liampotis, David Groep (June 2023)

Bridges: token translation example ‘SAML’ to PKIX



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

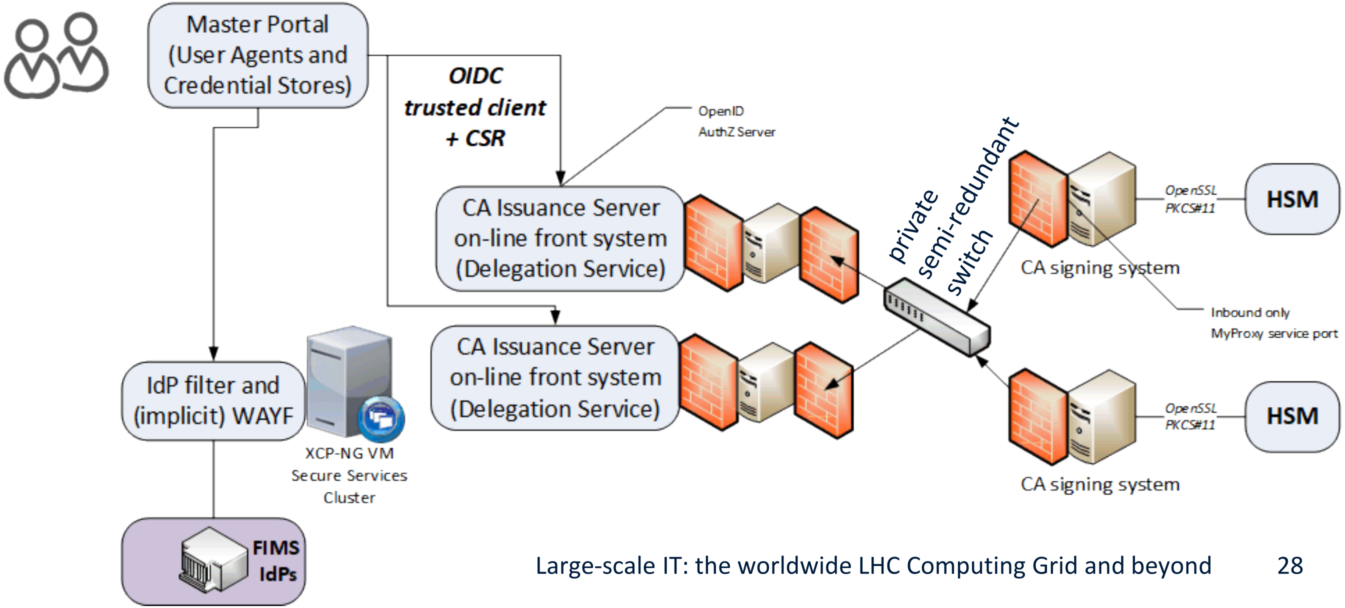
With a single, yet fully compliant, 'Heath Robinson' CA



A single-site locally-highly-available RCauth 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)

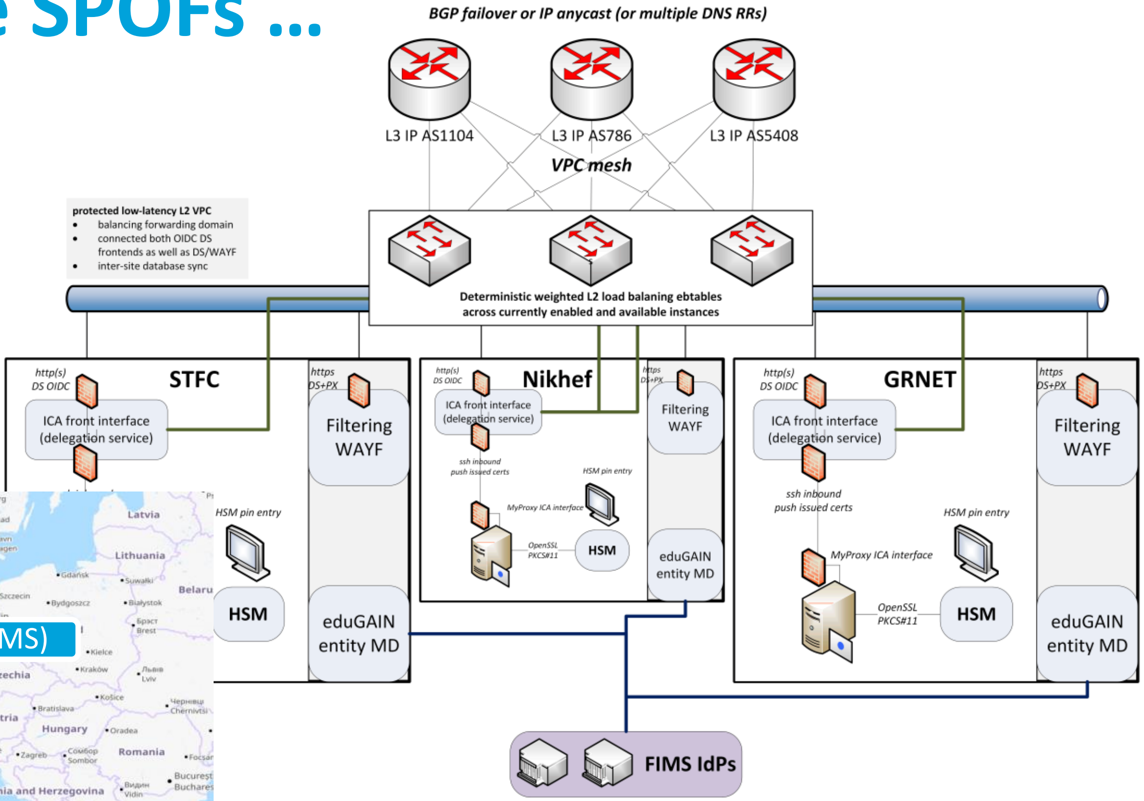
Eliminated SPOFs first using 'local HA'

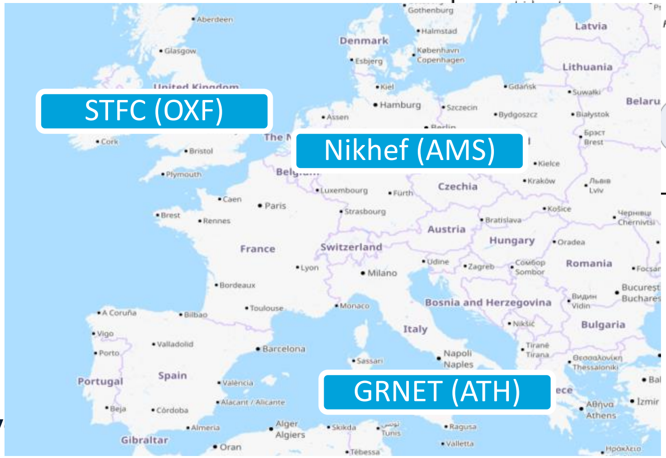


Since we do not like SPOFs …

Distributed High Availability setup

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



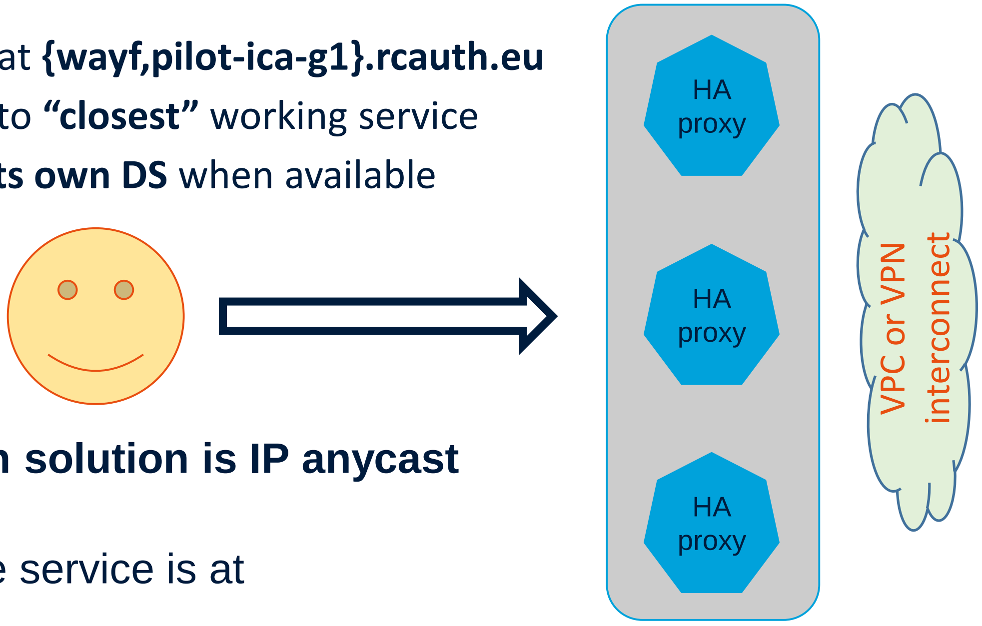


work supported by the EOSC Hub and EOSC Future Horizon Europe projects

A *transparent* multi-site setup is needed for the user

User

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



Straightforward proven solution is IP anycast

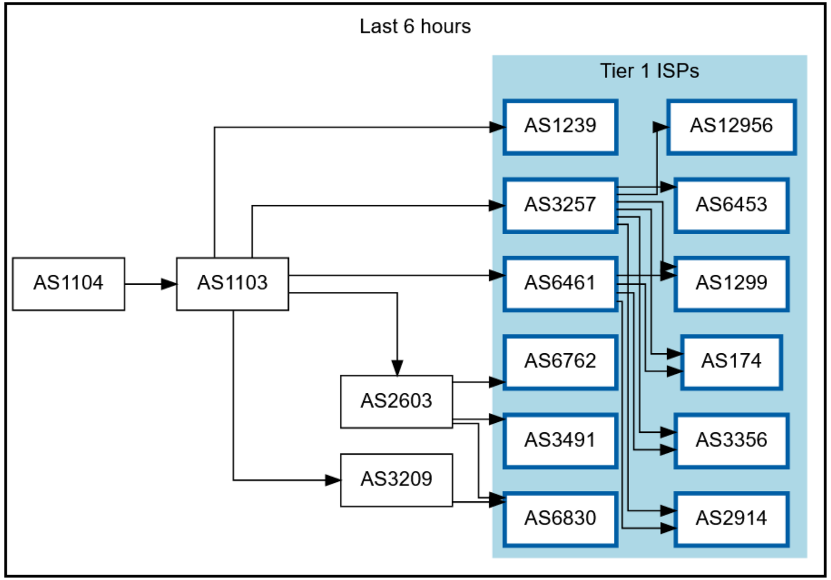
wherever the user is, the service is at

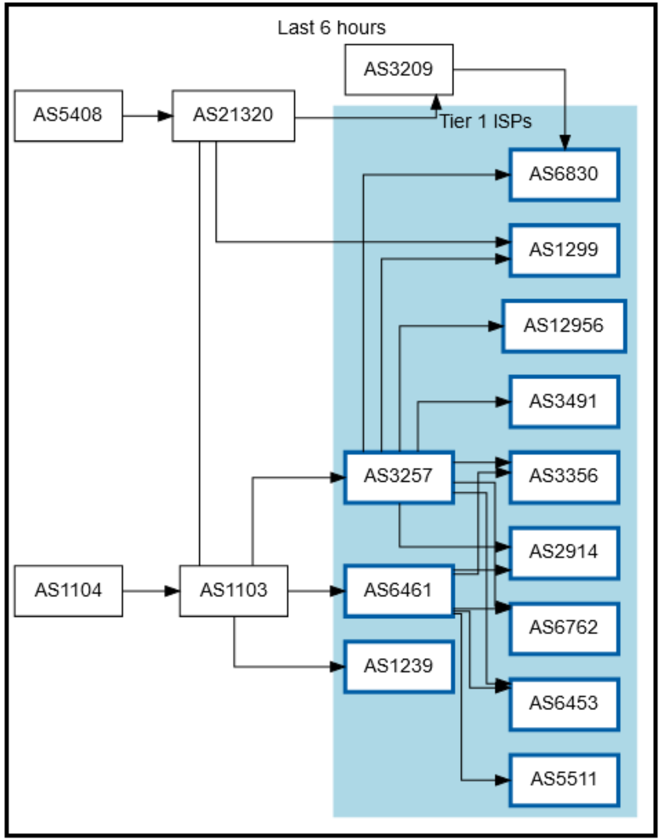
- **2a07:8504:01a0::1**
- or for legacy IP users at 145.116.216.1

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

selected imagery: Mischa Sallé, Jens Jensen, Nicolas Liampotis

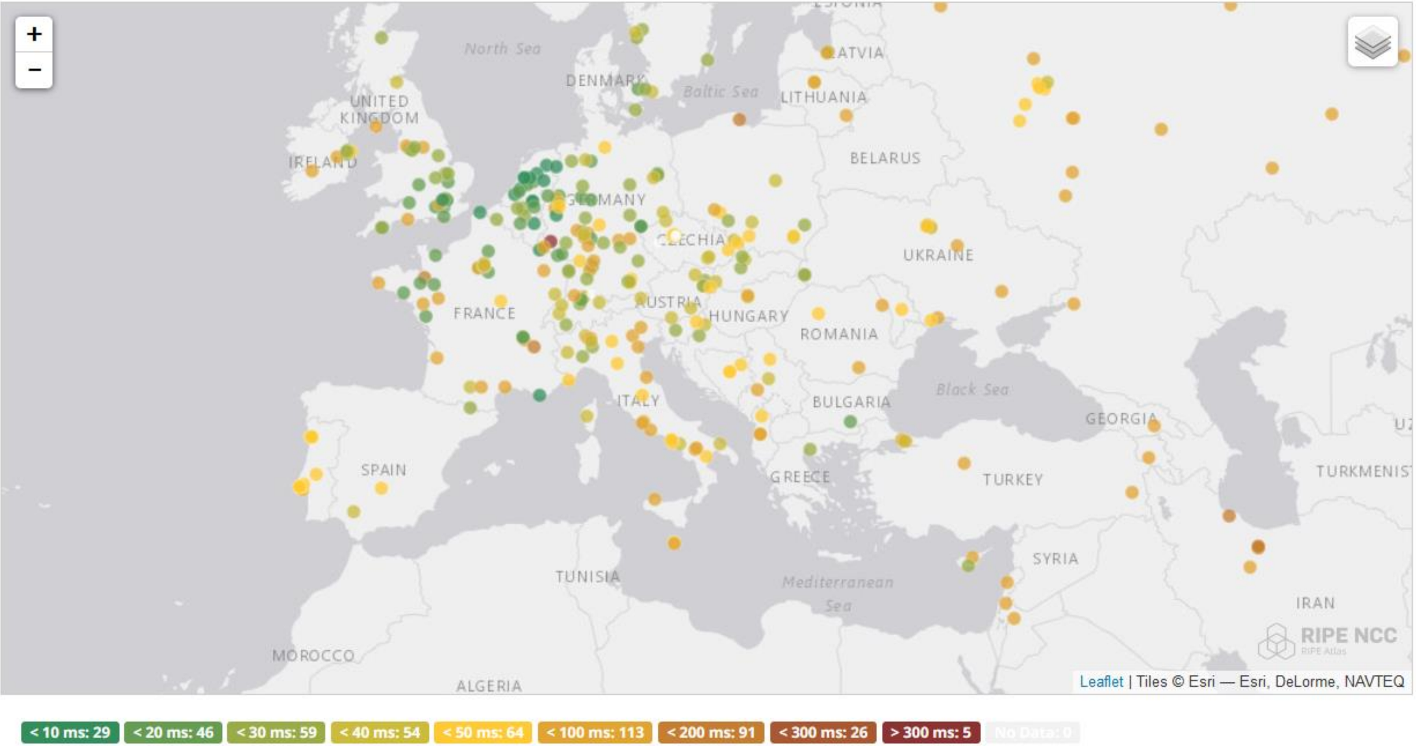
Getting 2a07:8504:1a0::/48 out there





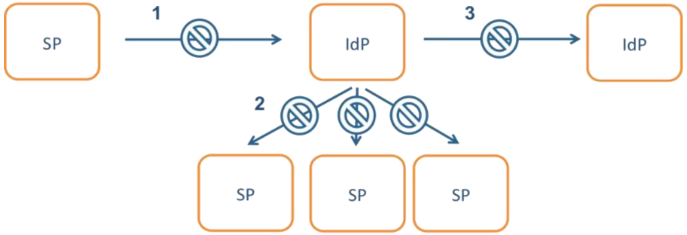
route maps: bgp.tools for 2a07:8504:1a0::/48 – IPv4 for 145.116.216.0/24 is similar – imagery from November 2022

And you get reasonable load balancing in Europe for free

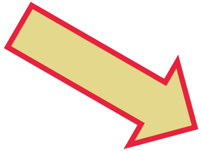


map: RIPE NCC RIPE Atlas - 500 probes, distributed across Europe (https://atlas.ripe.net/measurements/50949024/)

MANY PARTIES, SHARED SECURITY CHALLENGES

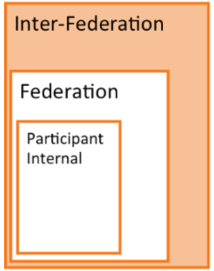


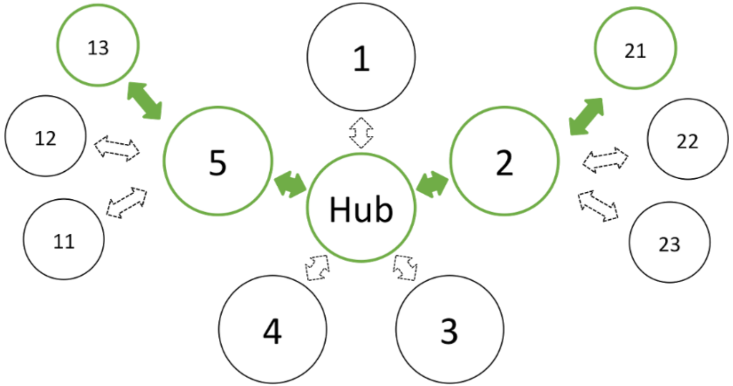
Incident Response Communication, communication blocks



Challenges

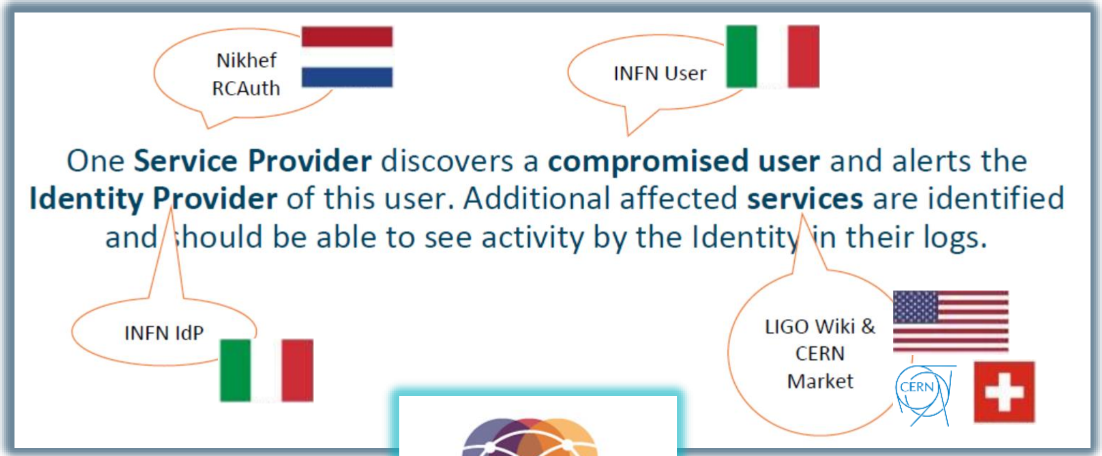
- IdP appears outside
  the service's security mandate
- Lack of contact or lack of trust in the IdP which to the SP is an unknown party
- IdP **fails to inform other affected** SPs, for fear of leaking data, of reputation, or just lack of interest and knowledge
- No established channels of communication, esp. not to federations themselves!





Inter-Federation Incident Response Communication

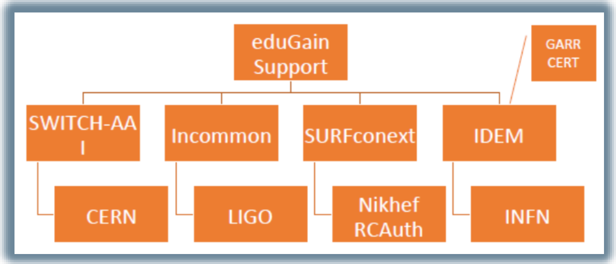
EXERCISES – COMMUNICATIONS AND ACTIONS

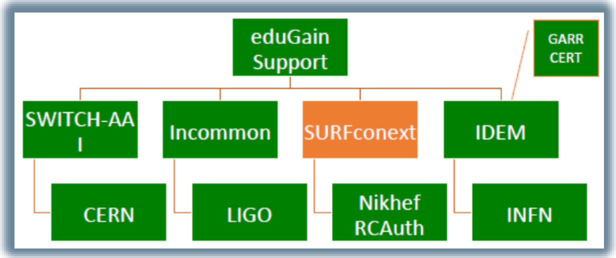


SIRTFI

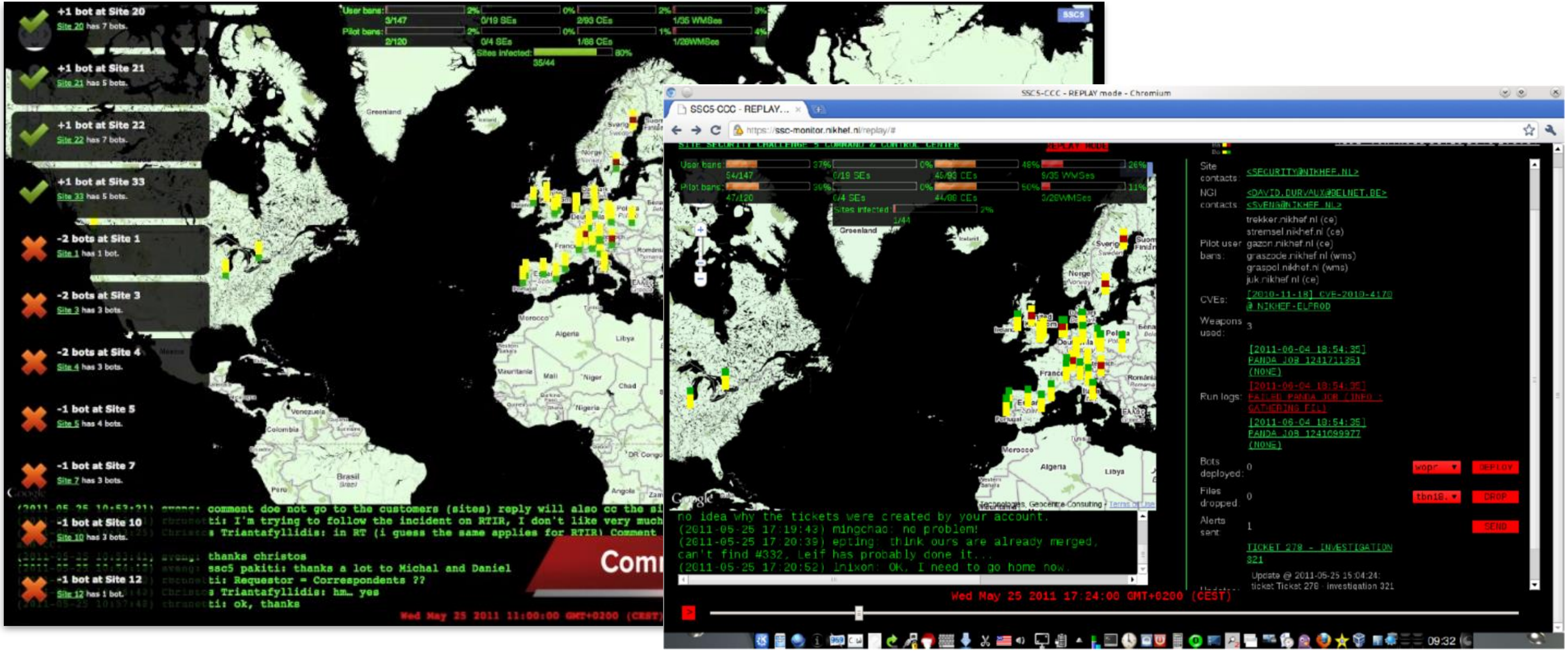
Security Incident Response Trust Framework for Federated Identity

parties involved in response challenge





SSC MONITORING



More than one: Nikhef Science Data Centre (‘234b’)

Physical farms: selecting the ‘worker nodes’

For HTC applications
– like WLCG, IGWN,
but also SKA, WeNMR – typically

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

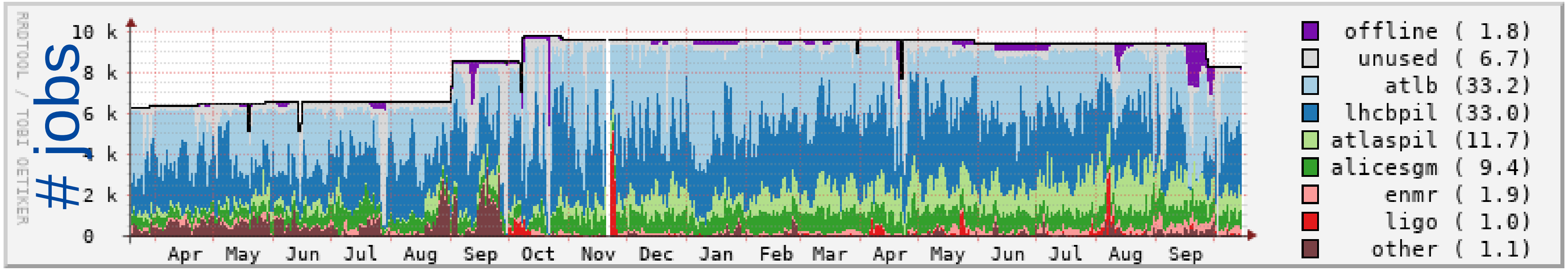
single-socket multicore systems are fine,
typical: 64-128 cores per system
network: 2x25Gbps
(+ ‘out of band’ management like IPMI)
memory: 8 GiB/core
local disk: 4TB NVME PCIe Gen4 x4
\+ space (physical + power) to add **GPU**

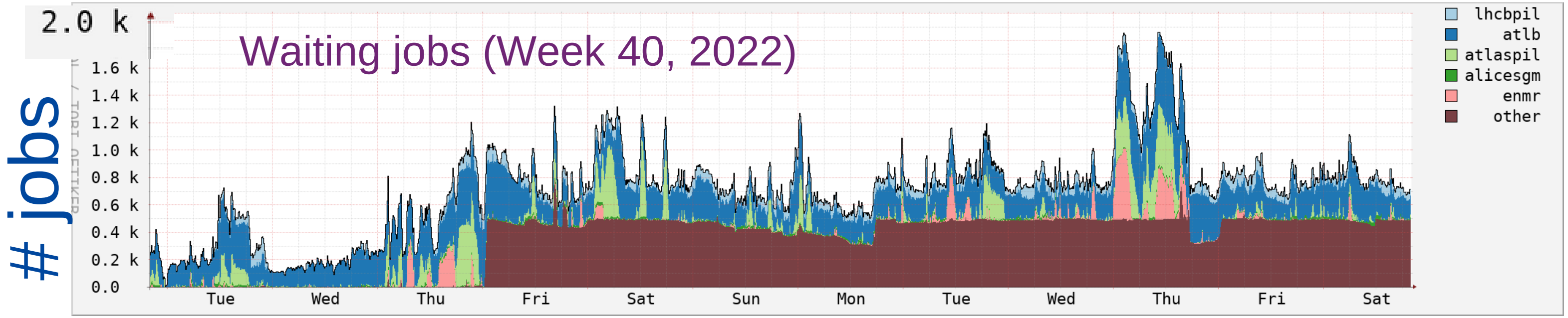
Image: Cluster ‘Lotenfeest’ at the Nikhef NDPF, acquired March 2020. Lenovo SR655 with AMD EPYC 7702P 64-Core single-socket

NDPF ‘WLCG and Dutch National Infra’ cluster

Running jobs:

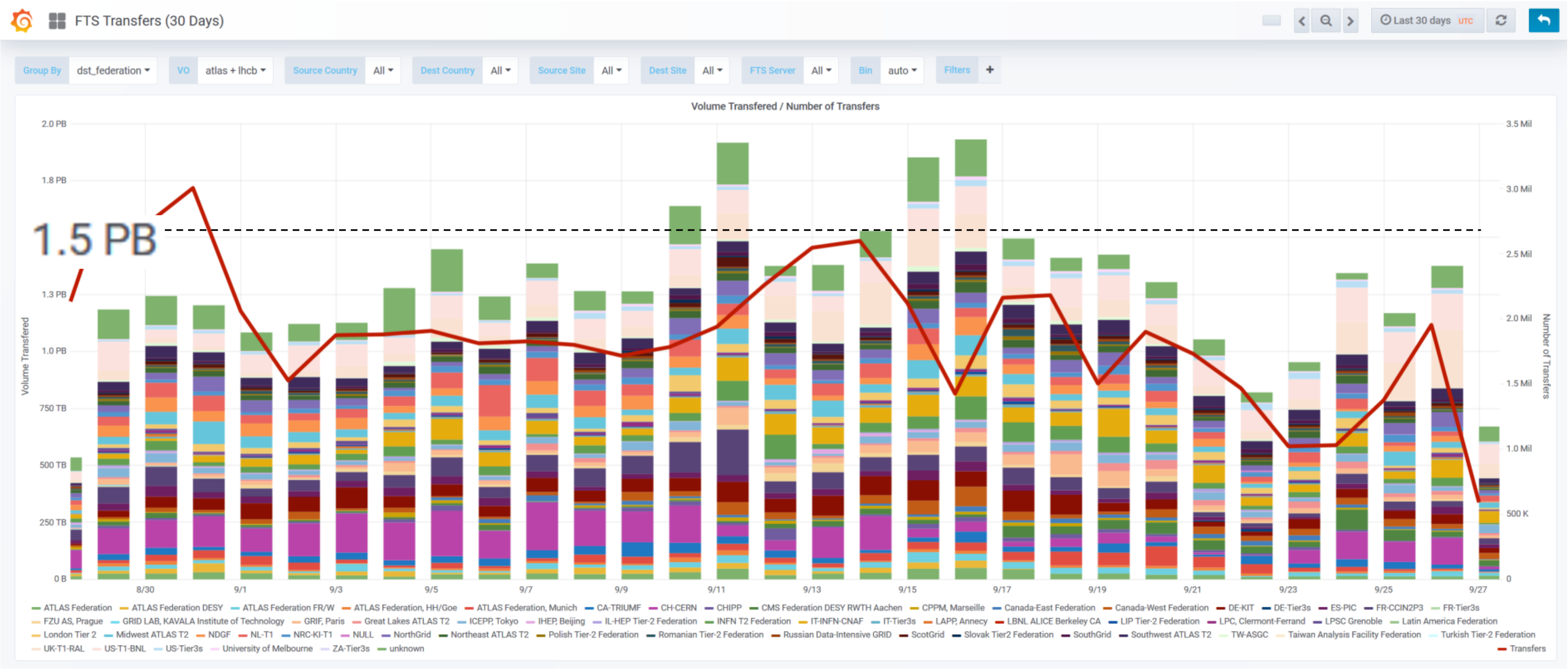
period: March 2021 .. October 2022





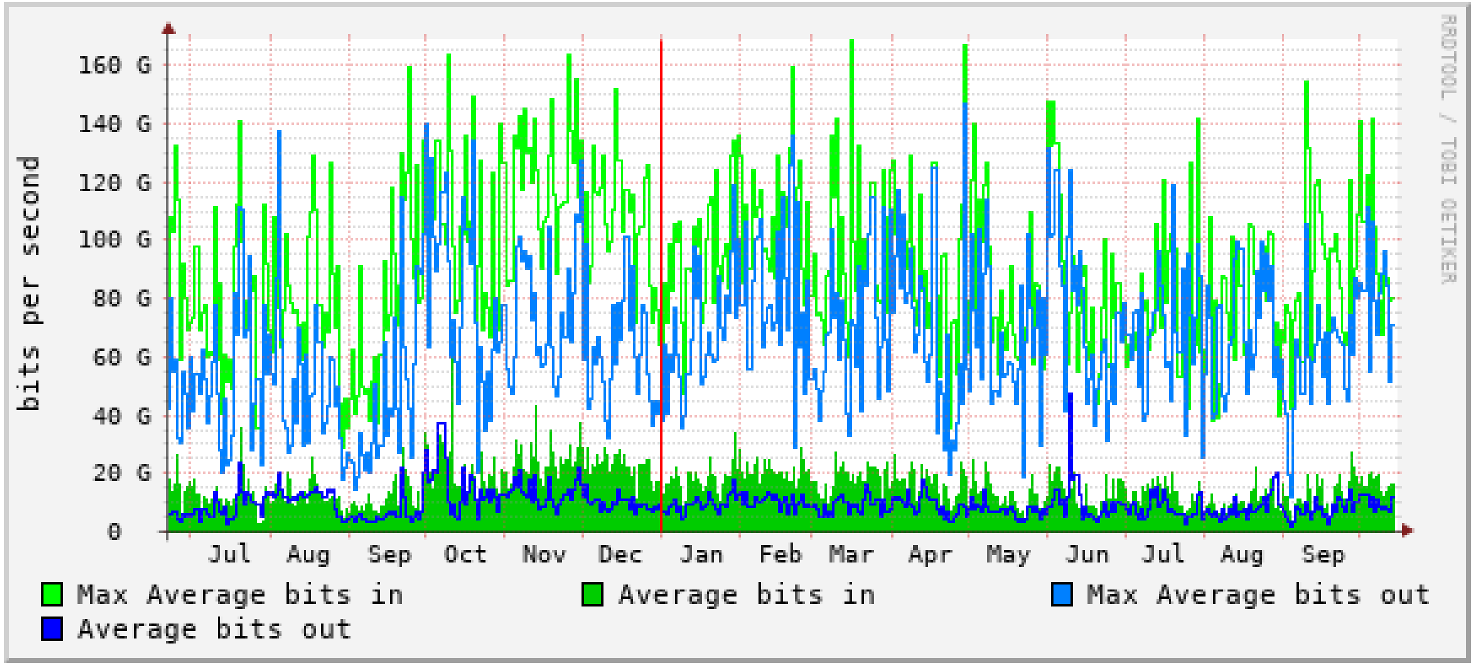
drainage event on Sept 27 are nodes being moved to the LIGO-VIRGO specific cluster; Source: NDPF Statistics overview, https://www.nikhef.nl/pdp/doc/stats/
‘other’ waiting jobs are almost all for the Auger experiment - GRISview images: Jeff Templon for NDPF and STBC

Global high throughput computing needs moving of data



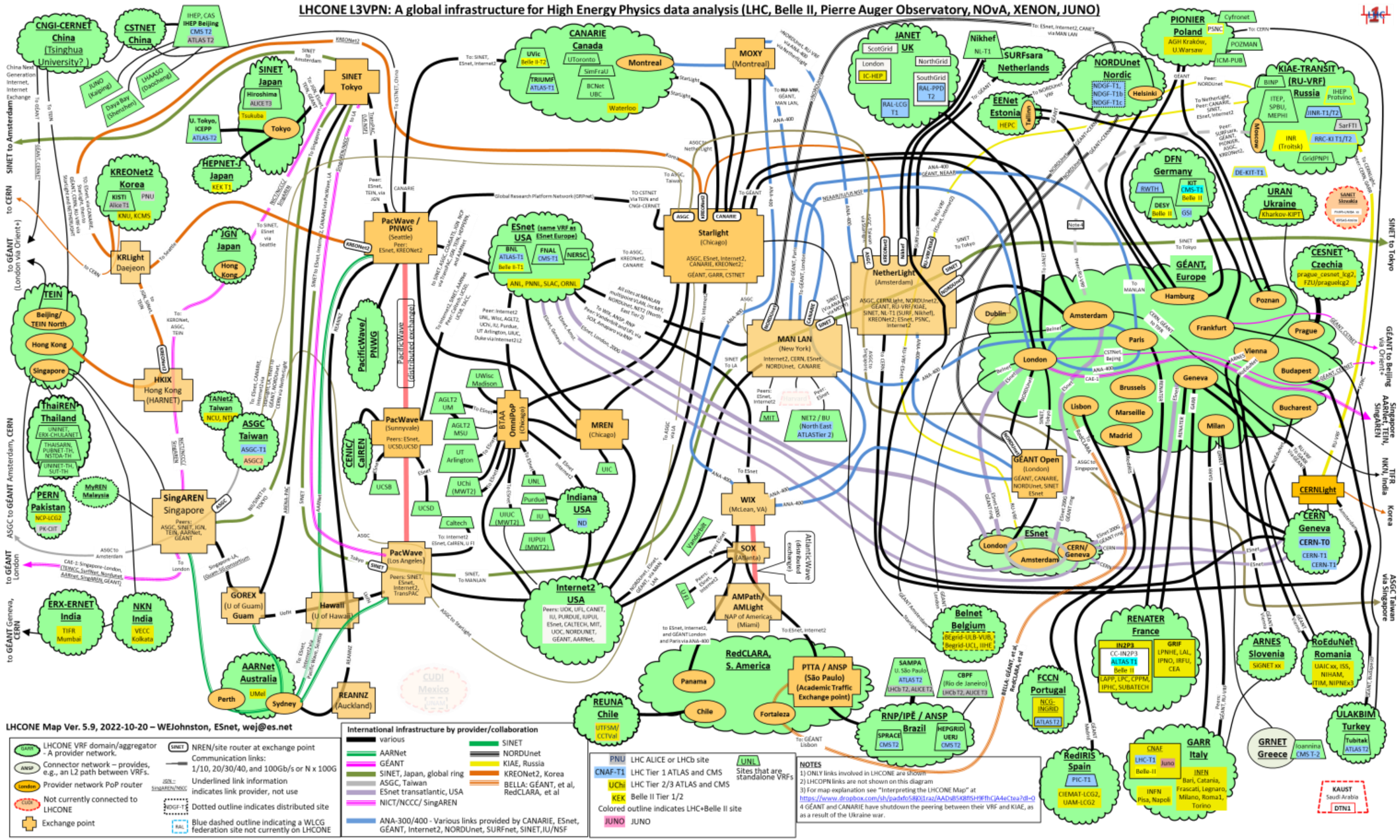
source: https://monit-grafana.cern.ch/d/000000420/fts-transfers-30-day ; data: November 2020 ; CERN FTS instance WLCG: daily transfer volume ATLAS+LHCb

Typical data traffic to and from the processing cluster



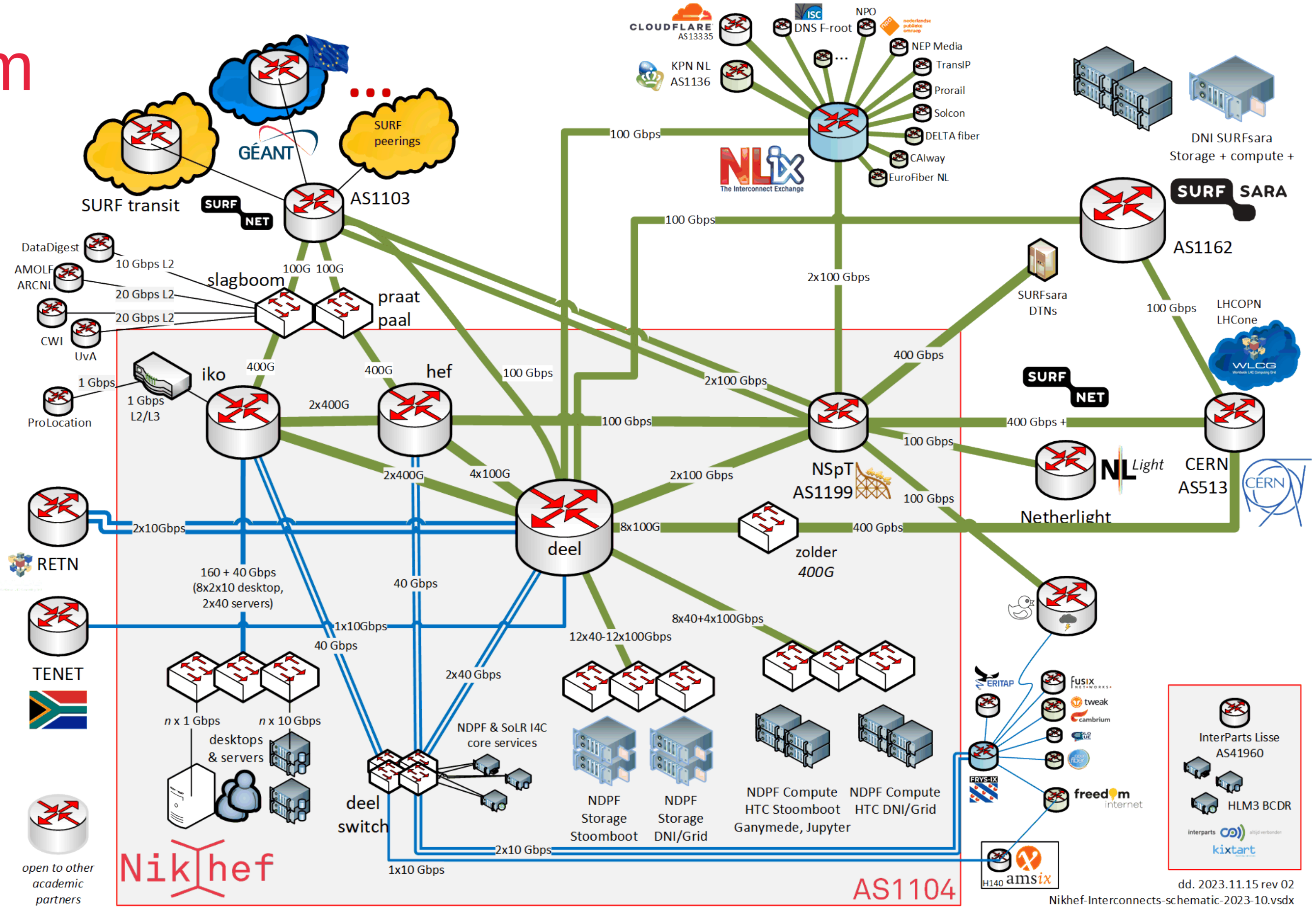
Source: Nikhef cricket graphs period June 2021 – October 2022 – aggregated (research) traffic to external peers from deelqfx – https://cricket.nikhef.nl/

LHCone



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

Just one random (smallish) autonomous system



AS1104

Exercising the network – sensor data and events

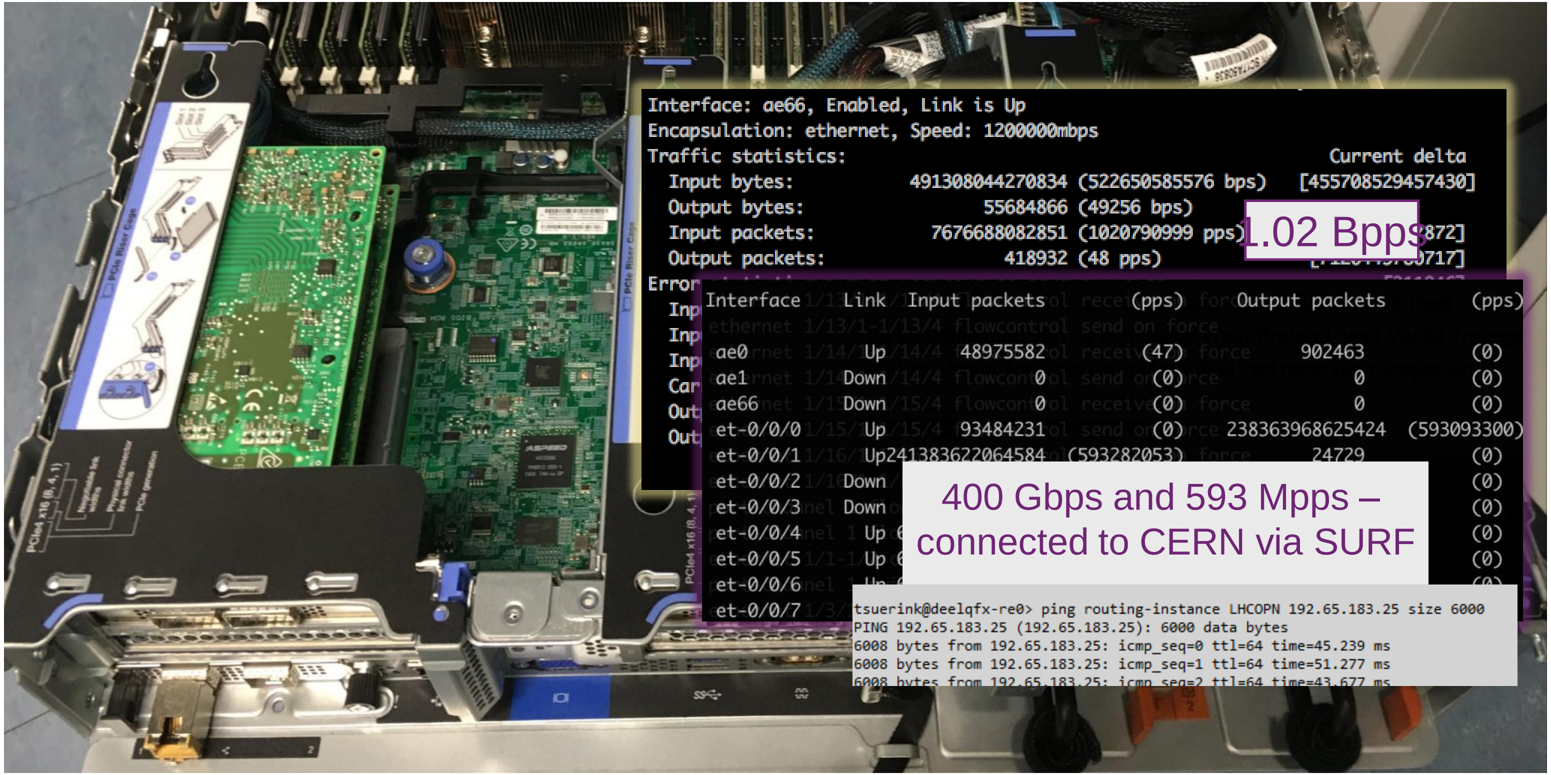
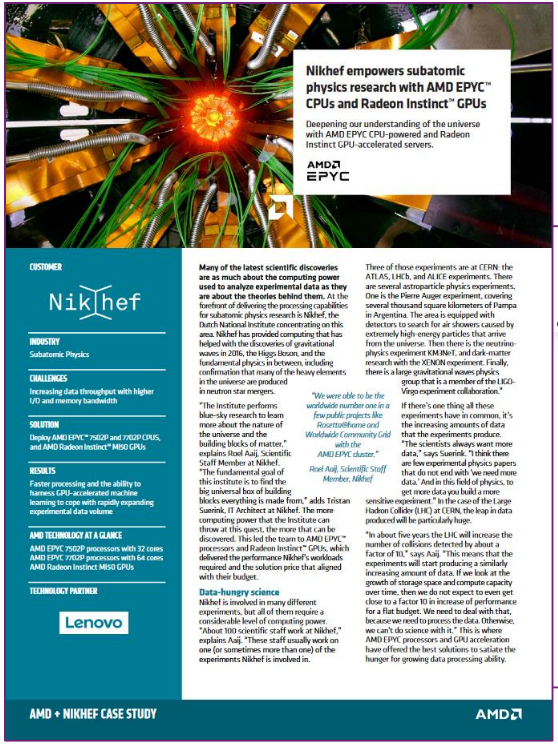


Image: ballenbak.nikhef.nl, Tristan Suerink

Innovation *on* infrastructure



Nikhef empowers subatomic physics research with AMD EPYC™ CPUs and Radeon Instinct™ GPUs

Deepening our understanding of the universe with AMD EPYC CPU-powered and Radeon Instinct GPU-accelerated servers.

AMD EPYC

CUSTOMER

Nikhef

INDUSTRY

Subatomic Physics

CHALLENGES

Increasing data throughput with higher I/O and memory bandwidth

SOLUTION

Deploy AMD EPYC™ 7502P and 7702P CPUS, and AMD Radeon Instinct™ MI50 GPUs

RESULTS

Faster processing and the ability to harness GPU-accelerated machine learning to cope with rapidly expanding experimental data volume

AMD TECHNOLOGY AT A GLANCE

AMD EPYC 7502P processors with 32 cores
AMD EPYC 7702P processors with 64 cores
AMD Radeon Instinct MI50 GPUs

TECHNOLOGY PARTNER

Lenovo

Many of the latest scientific discoveries are as much about the computing power used to analyze experimental data as they are about the theories behind them. At the forefront of delivering the processing capabilities for subatomic physics research is Nikhef, the Dutch National Institute concentrating on this area. Nikhef has provided computing that has helped with the discoveries of gravitational waves in 2016, the Higgs Boson, and the fundamental physics in between, including confirmation that many of the heavy elements in the universe are produced in neutron star mergers.

"The Institute performs blue-sky research to learn more about the nature of the universe and the building blocks of matter," explains Roel Aaij, Scientific Staff Member at Nikhef. "The fundamental goal of this institute is to find the big universal box of building blocks everything is made from," adds Tristan Suerink, IT Architect at Nikhef. The more computing power that the Institute can throw at this quest, the more that can be discovered. This led the team to AMD EPYC™ processors and Radeon Instinct™ GPUs, which delivered the performance Nikhef's workloads required and the solution price that aligned with their budget.

"We were able to be the worldwide number one in a few public projects like Rosetta@home and Worldwide Community Grid with the AMD EPYC cluster."

Roel Aaij, Scientific Staff Member, Nikhef

Data-hungry science

Nikhef is involved in many different experiments, but all of them require a considerable level of computing power. "About 100 scientific staff work at Nikhef," explains Aaij. "These staff usually work on one (or sometimes more than one) of the experiments Nikhef is involved in.

Three of those experiments are at CERN: the ATLAS, LHCb, and ALICE experiments. There are several astroparticle physics experiments. One is the Pierre Auger experiment, covering several thousand square kilometers of Pampa in Argentina. The area is equipped with detectors to search for air showers caused by extremely high-energy particles that arrive from the universe. Then there is the neutrino-physics experiment KM3NeT, and dark-matter research with the XENON experiment. Finally, there is a large gravitational waves physics group that is a member of the LIGO-Virgo experiment collaboration."

If there's one thing all these experiments have in common, it's the increasing amounts of data that the experiments produce. "The scientists always want more data," says Suerink. "I think there are few experimental physics papers that do not end with 'we need more data.' And in this field of physics, to get more data you build a more sensitive experiment." In the case of the Large Hadron Collider (LHC) at CERN, the leap in data produced will be particularly huge.

"In about five years the LHC will increase the number of collisions detected by about a factor of 10," says Aaij. "This means that the experiments will start producing a similarly increasing amount of data. If we look at the growth of storage space and compute capacity over time, then we do not expect to even get close to a factor 10 in increase of performance for a flat budget. We need to deal with that, because we need to process the data. Otherwise, we can't do science with it." This is where AMD EPYC processors and GPU acceleration have offered the best solutions to satiate the hunger for growing data processing ability.

AMD + NIKHEF CASE STUDY

AMD

FUNGIBLE

NIKHEF, SURF AND FUNGIBLE SET NEW BENCHMARK FOR THE WORLD'S FASTEST STORAGE PERFORMANCE

Companies Double Current Performance Record, Setting the New Bar at 6.55 Million Read IOPS

Image: Minister of Economic Affairs M. Adriaansens launched the Innovation Hub with Nikhef, SURF, Nokia and NL-ix, January 2023. Composite image from https://www.surf.nl/nieuws/minister-adriaansens-lanceert-testomgeving-voor-supersnelle-netwerktechnologie

Our science data flows are somebody else's DDoS attack

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,

Op de…
Een goe…
Examen…
Wat gebeurde er die nacht?
Vragen over dit artikel?

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

Door Rutger Otto

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

202 reacties

aanvallen en hoe de aanval uitgevoerd wordt.
Het 'red team' is verantwoordelijk voor de aanvallen, het 'blue team' voor de verdediging. Een van de partijen die aan de avond meedoet is Nikhef. Tristan, IT architect bij Nikhef, geeft aan "dat zij dit

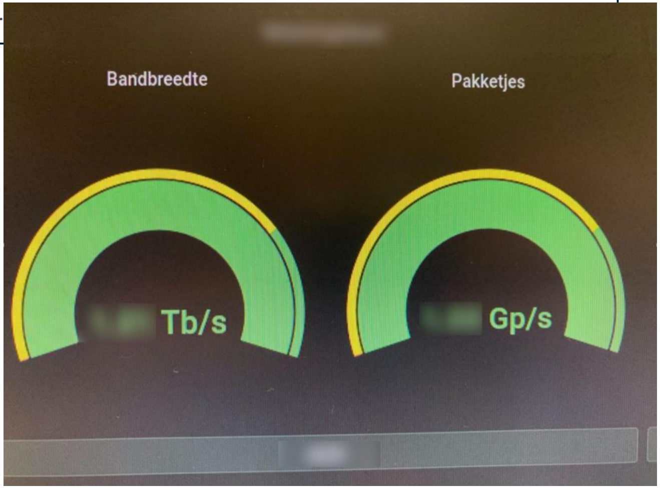
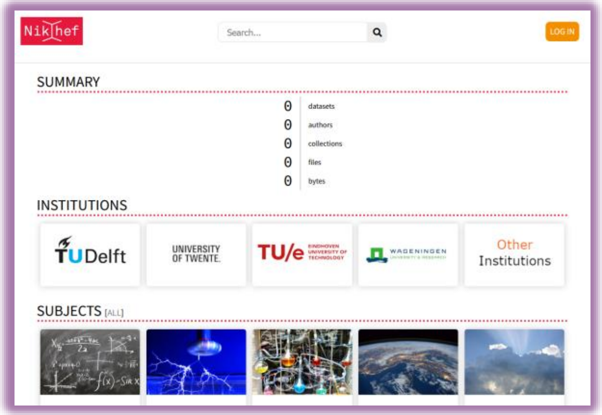


Image sources: belastingdienst.nl, rws.nl, nu.nl, werkentegennederland.nl

But what about some other digital competences?

We have the network (physical and in terms of expertise) for processing and analysis, but other elements are missing

- accessible analysis preservation for our on-prem experiments and R&D projects
- longer-term 'local' re-use in programmatic research
- complexity of managing ongoing large analyses
- data and software of non-collaboration-based outputs



Now preparing for **analysis preservation** and RDM through managed 'snapshots' as part of the analysis pipelines

- *linking Stoomboot dCache to a new institutional Research Data Management*

Power usage and efficient data centres

Nikhef scientific data centre (the ‘glass box’) designed for 400kW total use + cooling in 47 racks

De <u>snelste</u> CPU/GPU is voor ons niet altijd de beste (*sorry gamers & miners!*), want 5 jaar energie en beheer zijn even kostbaar als de server zelf

WKO: Warmte Koude Opslag

21% van het vermogen is nodig om te koelen, maar: we mogen 3500GJoule/jaar (~112 kWjaar, ~982 000 kWh) aan studenten tegenover leveren om ze warm te houden !

Let’s go on tour!

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

Maastricht University
Nikhef