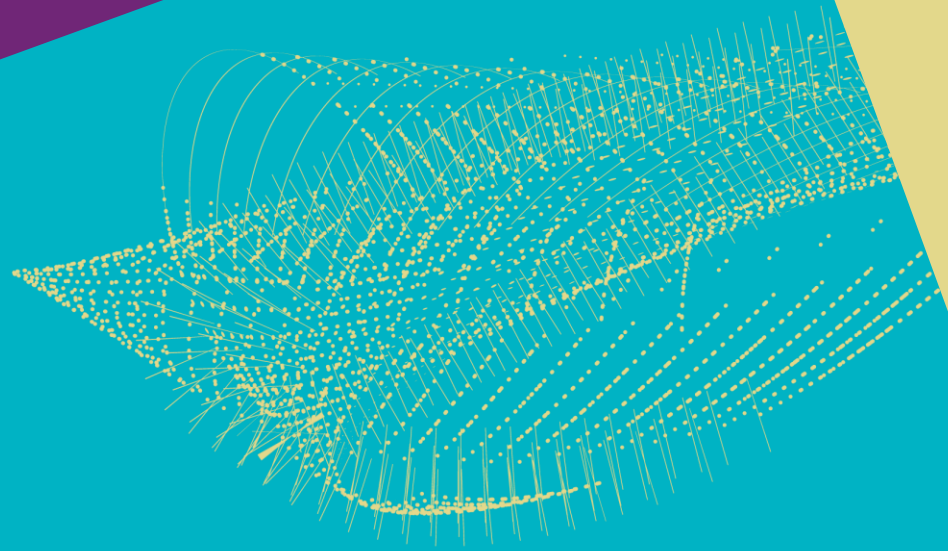
Nikhef



David Groep

Physics Data Processing

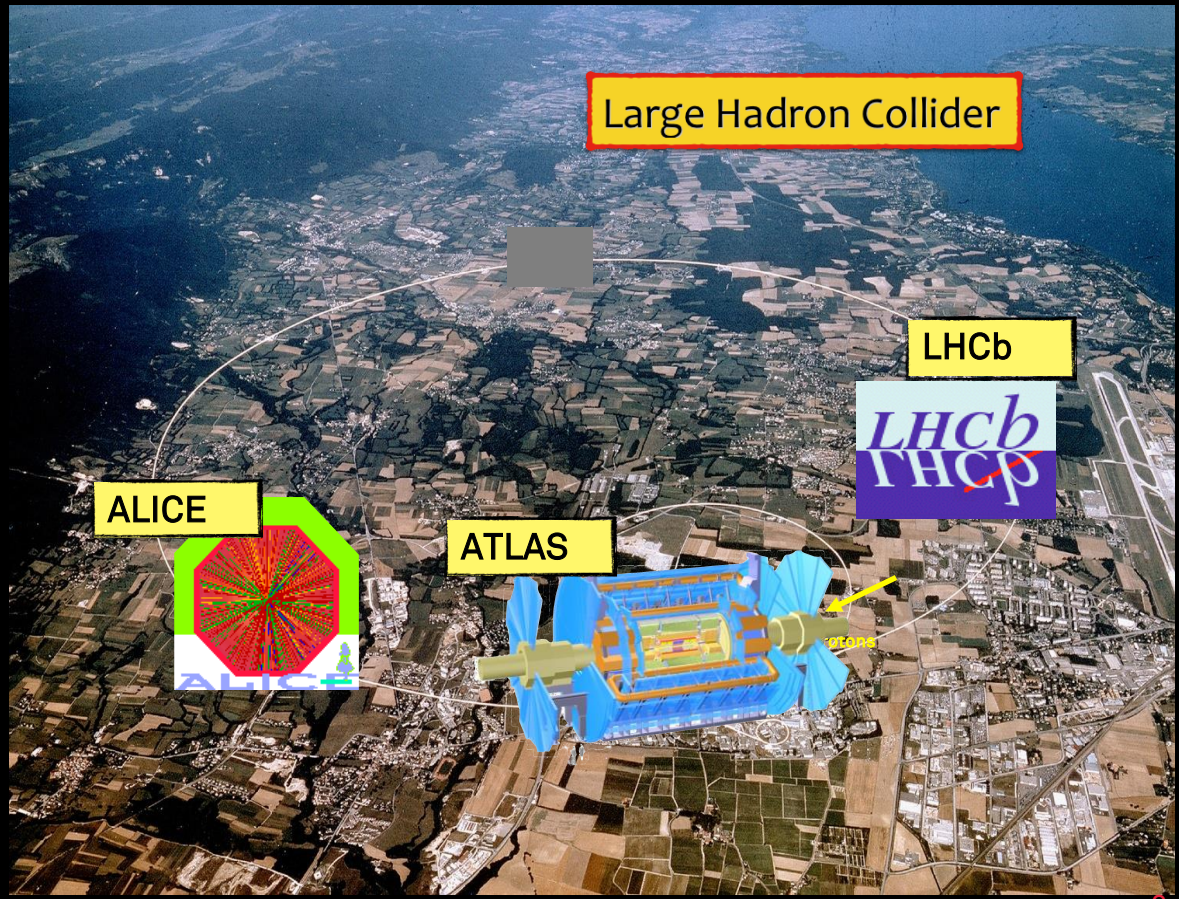
A Brief Overview

Nikhef – accelerator experiments

ATLAS
- *Higgs physics*
- Beyond the standard model

LHCb
- *rare decays*
- matter vs anti-matter

Alice
- *quark-gluon plasma*
- matter phase transitions



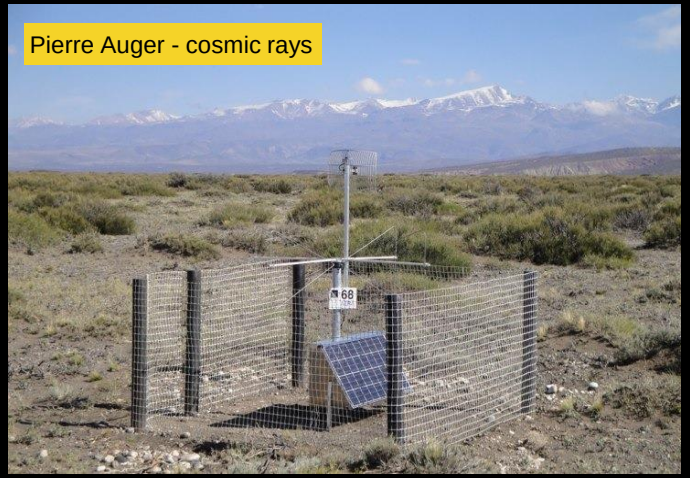
Nikhef – (astro)particle physics

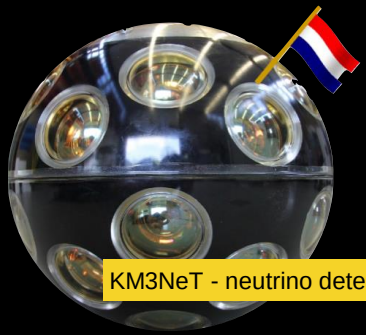
KM3NET
- neutrino telescope

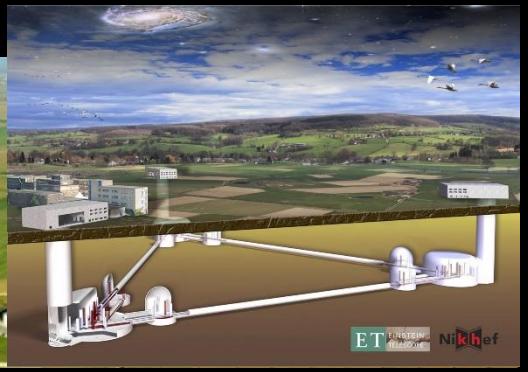
Virgo/LIGO/ET
- gravitational waves

Auger
- ultra-high energy cosmic rays

Xenon
- search for dark matter



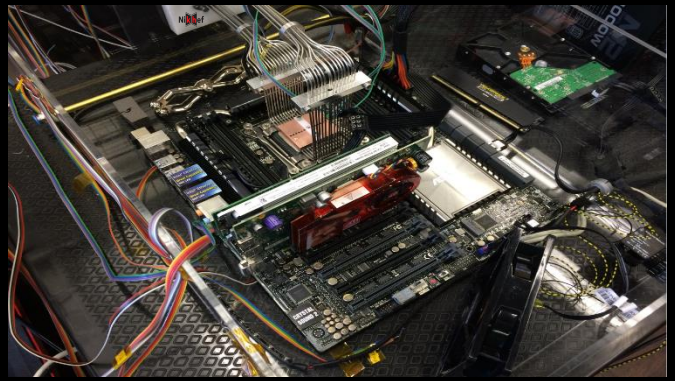




Nikhef collaboration – underpinning programmes

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

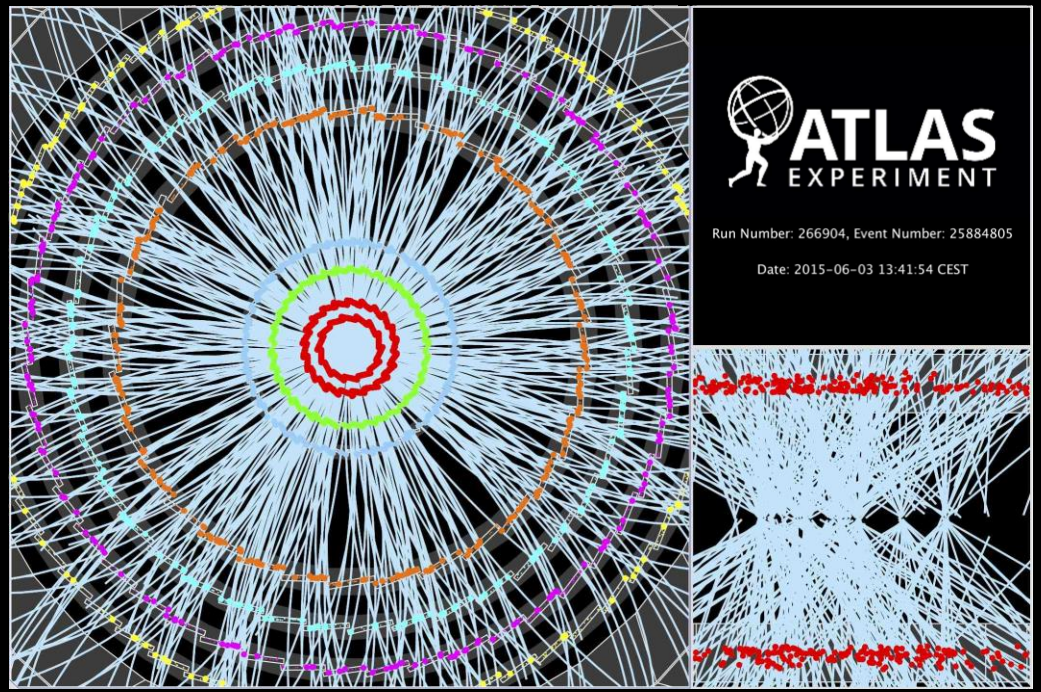
- NDPF processing facility
- National e-Infrastructure *coordinated by SURF*
- experimental next-gen systems engineering
- cross-tier global networks
- stressing & public/private collaborative design

Infrastructure for Secure Collaboration

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

Computing on data: 40 million per second

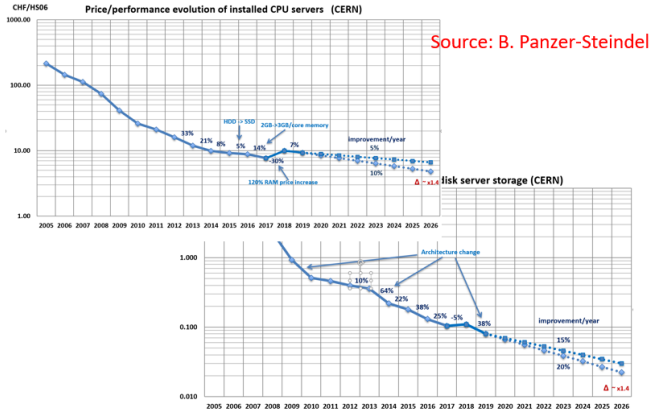
~ 10 seconds to a compute a single event at ATLAS (as an example) for 'jets' containing ~30 colllisions



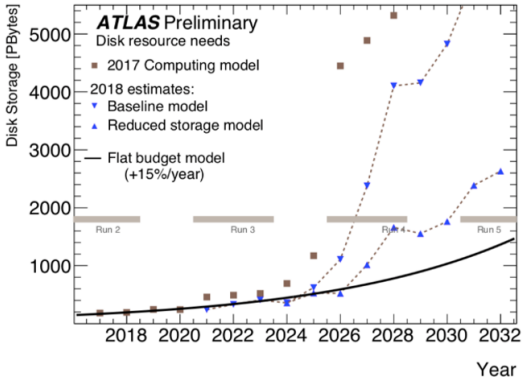
Display of a proton-proton collision event recorded by ATLAS on 3 June 2015, with the first LHC stable beams at a collision energy of 13 TeV
Event processing time: v19.0.1.1 as per Jovan Mitrevski and 2015 J. Phys.: Conf. Ser. 664 072034 (CHEP2015)

The High Luminosity Challenge

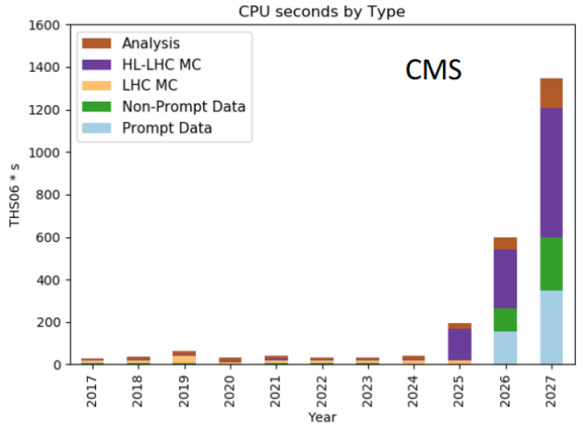
- The effort to reduce the gap between the estimations of needed and available resources is continuing
  - Changes in the computing models
  - Improvements in software performance
  - GPU computing more accessible
  - Market movements driving down CPU prices
- Still, we are not there yet!



Source: B. Panzer-Steindel



Source: D. Costanzo



Source: D. Lange

source: Andrea Sciaba et al. for the WLCG Resource Evolution WG, CHEP2019

Algorithms: beyond data organisation and throughput

High level C++ code → `if (abs(point[0] - origin[0]) > xhalfsz) return FALSE;`

Assembler instructions →

```
movsd 16(%rsi), %xmm0
subsd 48(%rdi), %xmm0   // load & subtract
andpd _2il0floatpacket.1(%rip), %xmm0 // and with a mask
comisd 24(%rdi), %xmm0 // load and compare
jbe ..B5.3     # Prob 43% // jump if FALSE
```

Same instructions laid out according to latencies on the Core 2 processor →

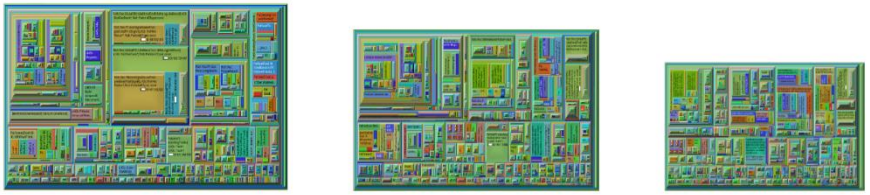
NB: Out-of-order scheduling not taken into account.

Cycle	Port 0	Port 1	Port 2	Port 3	Port 4	Port 5
1			load point[0]			
2			load origin[0]			
3						
4						
5						
6		subsd	load float-packet			
7						
8			load xhalfsz			
9						
10	andpd					
11						
12	comisd					
13						jbe

5 September 2007

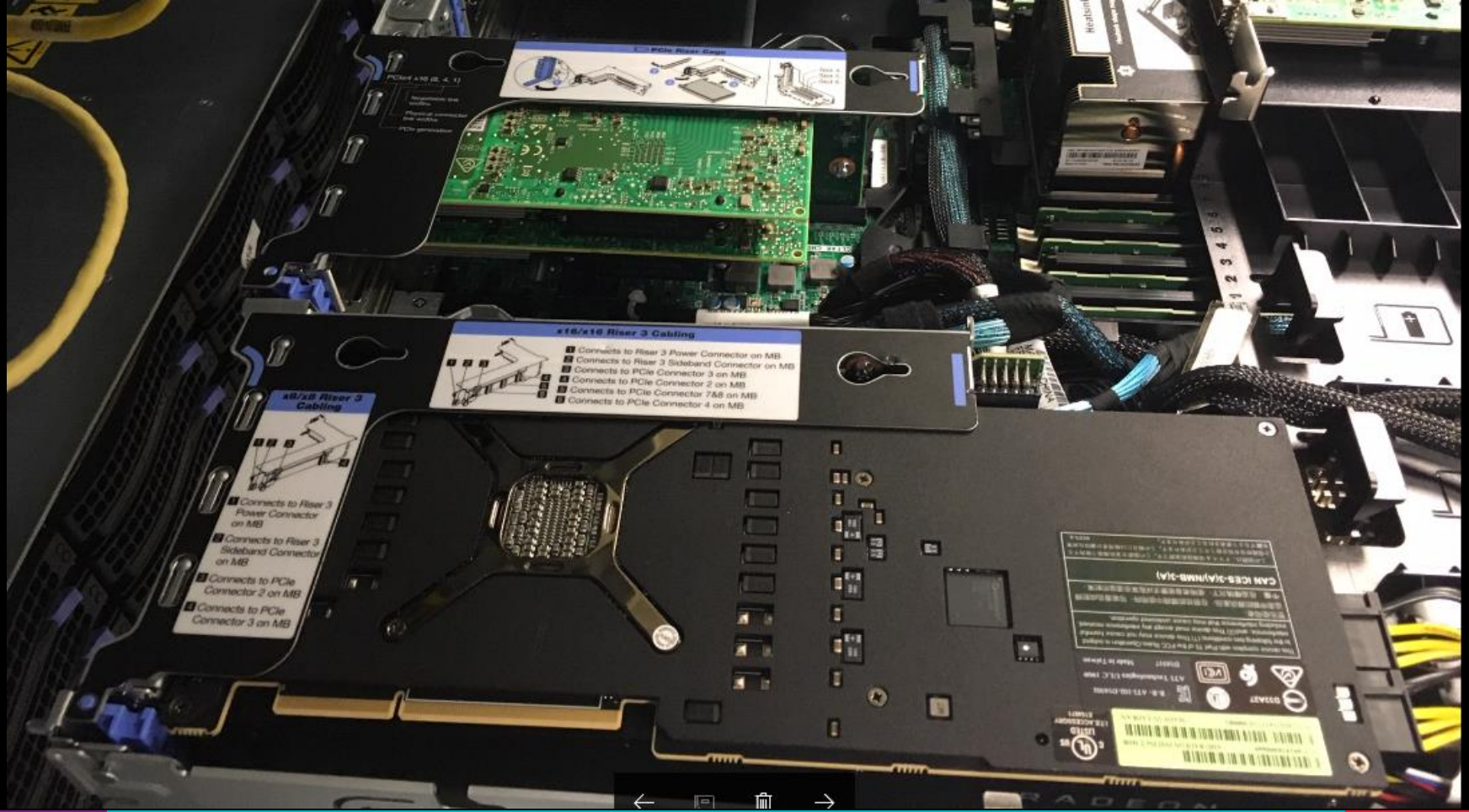
2007 Core 2 efficiency: Sverre Jarp, CHEP2007 (!)

plofkip.nikhef.nl



v45r1 v48r1 v48r1 (2015 reco)

Gerhard Raven et al, 2015 for LHCb



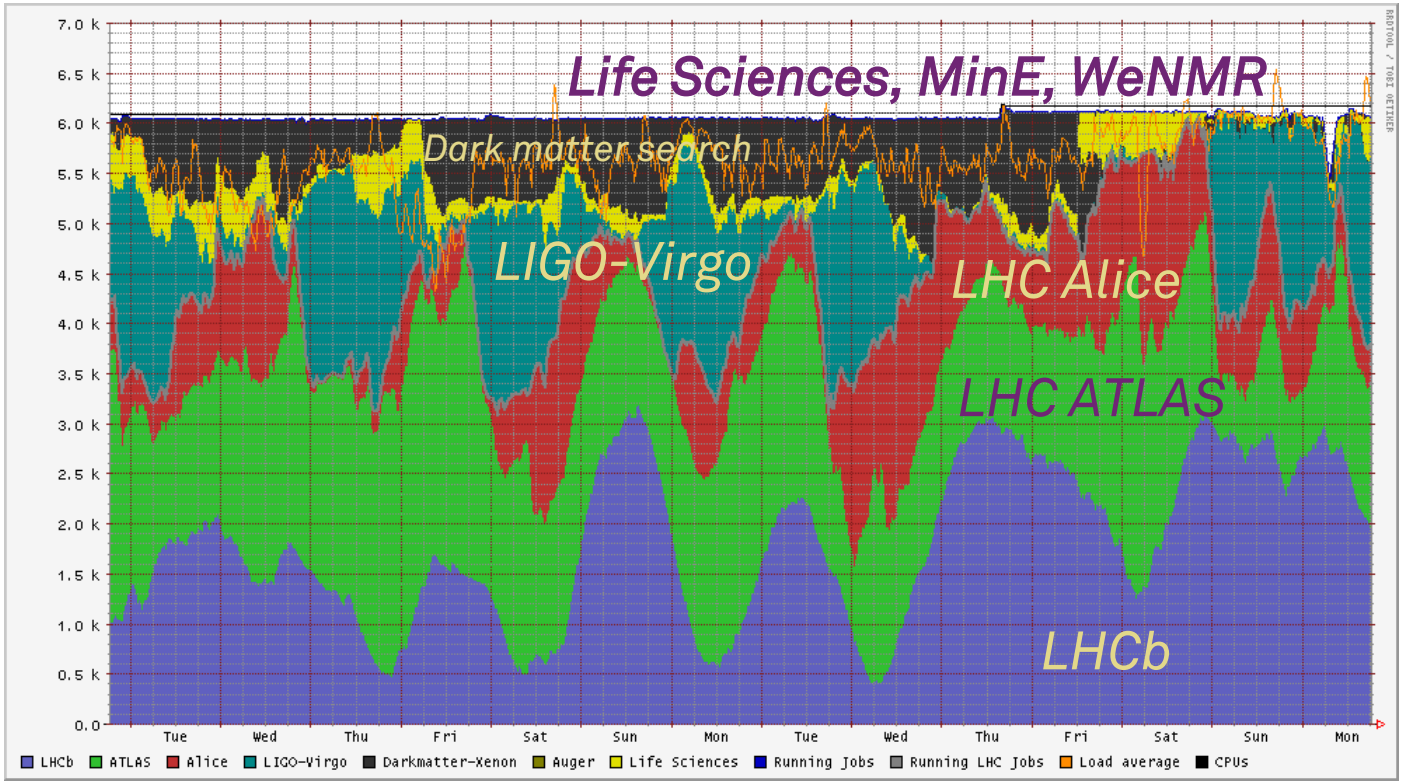
Prototype AMD MI50 GPU in the Nikhef Speeltuin last year

Image: Tristan Suerink

Nikhef Data Processing Facility – multi-community service

Services in the DNI

- compute in many operational sites (SURFsara, Nikhef, RUG-CIT)
- high-throughput storage at SURFsara and Nikhef,
- long-term storage at SURFsara,
- interconnected by SURFnet, and authentication by SURFcertificaten



NDPF voview short 1 October 2018

FEEDING DATA – TIERED STORAGE

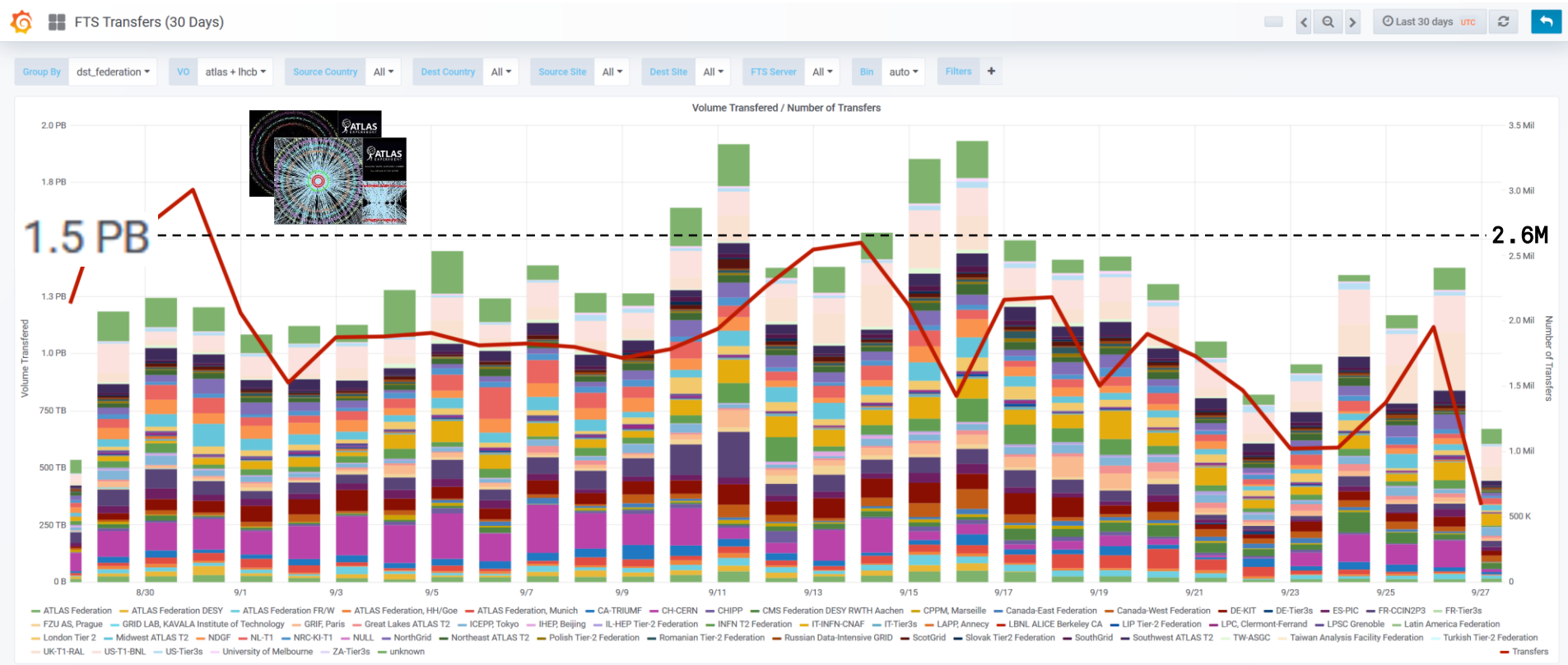
‘Hot’ data tier at ~ 5 PiB, ‘warm’ tier of ~ 1.6PiB, interconnect to nearline

12 MiB/s/TiB - about 4 PiB configured (both DNI/NL-T1 and Stoomboot)

‘hooikanon’: 240 (4x60) spindles, 12 TByte disks, 4x100Gbps network
4x IBM SL922 Power9 PPC servers, 8x NetApp E5700 controllers, 4 trays

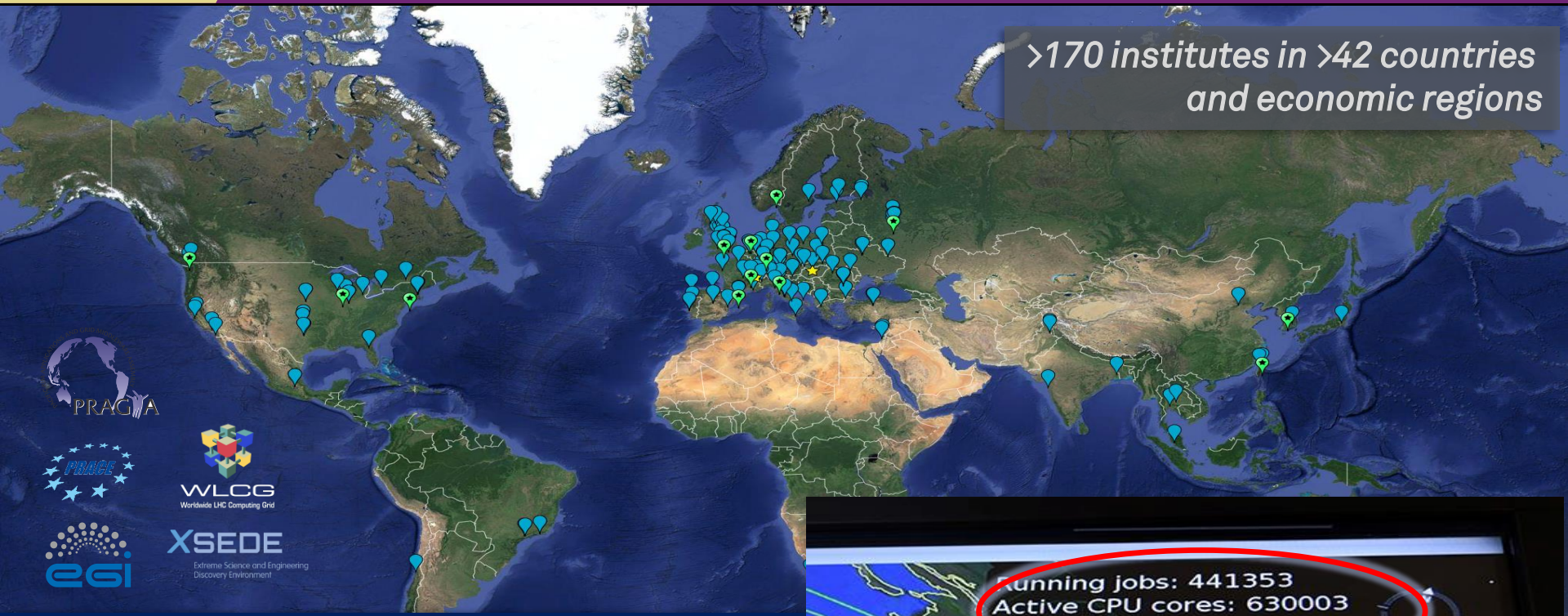
https://wiki.nikhef.nl/grid/100g_Storage_Box
hooikanon-01..04.nikhef.nl

Data Distribution: LHCOPN/LHCone

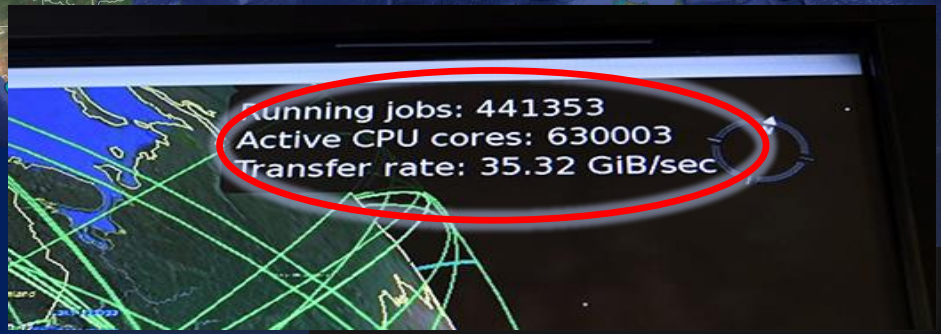


source: https://monit-grafana.cern.ch/d/000000420/fts-transfers-30-day on September 27, 2018

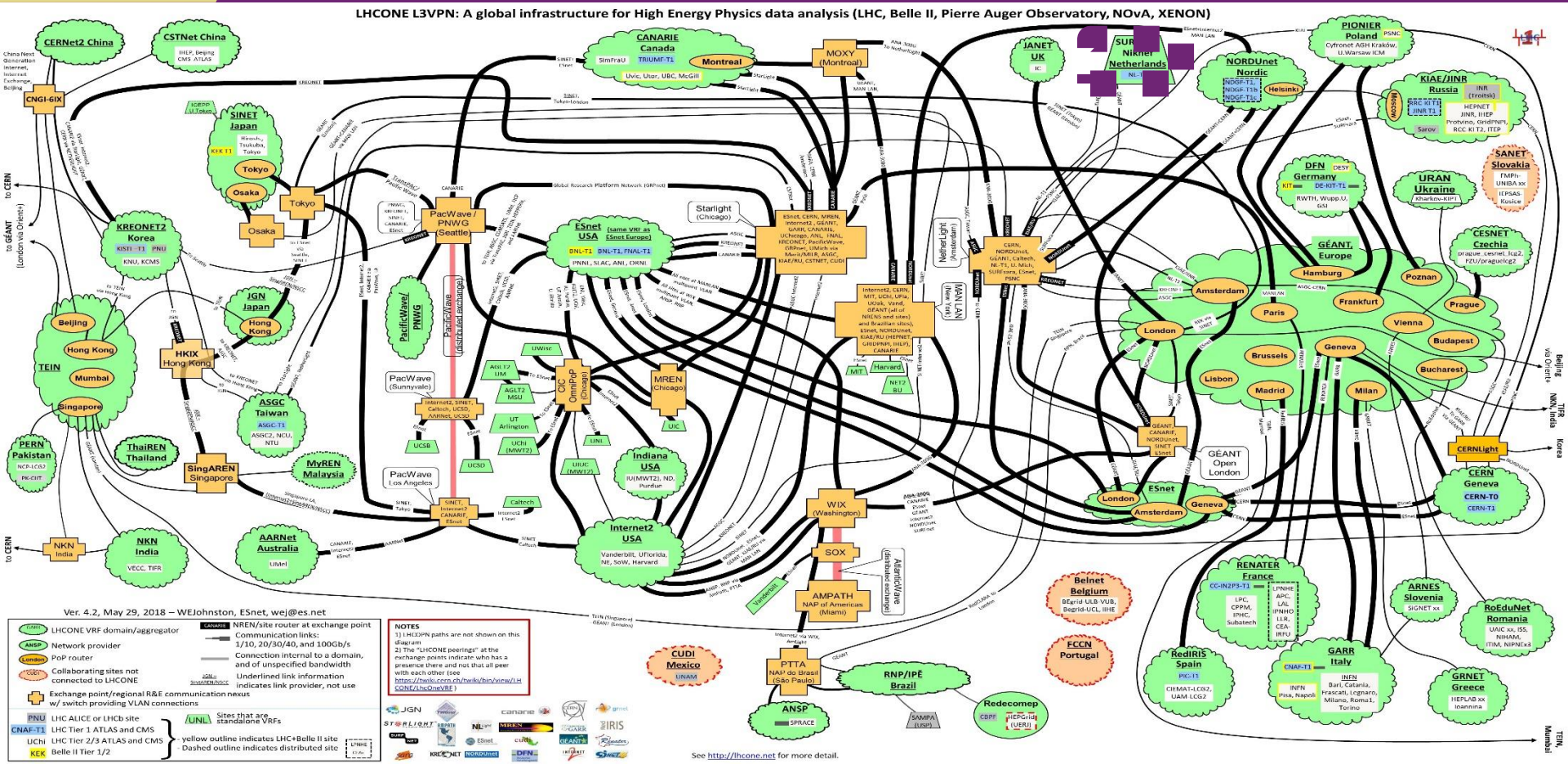
Globally distributed in EGI & WLCG



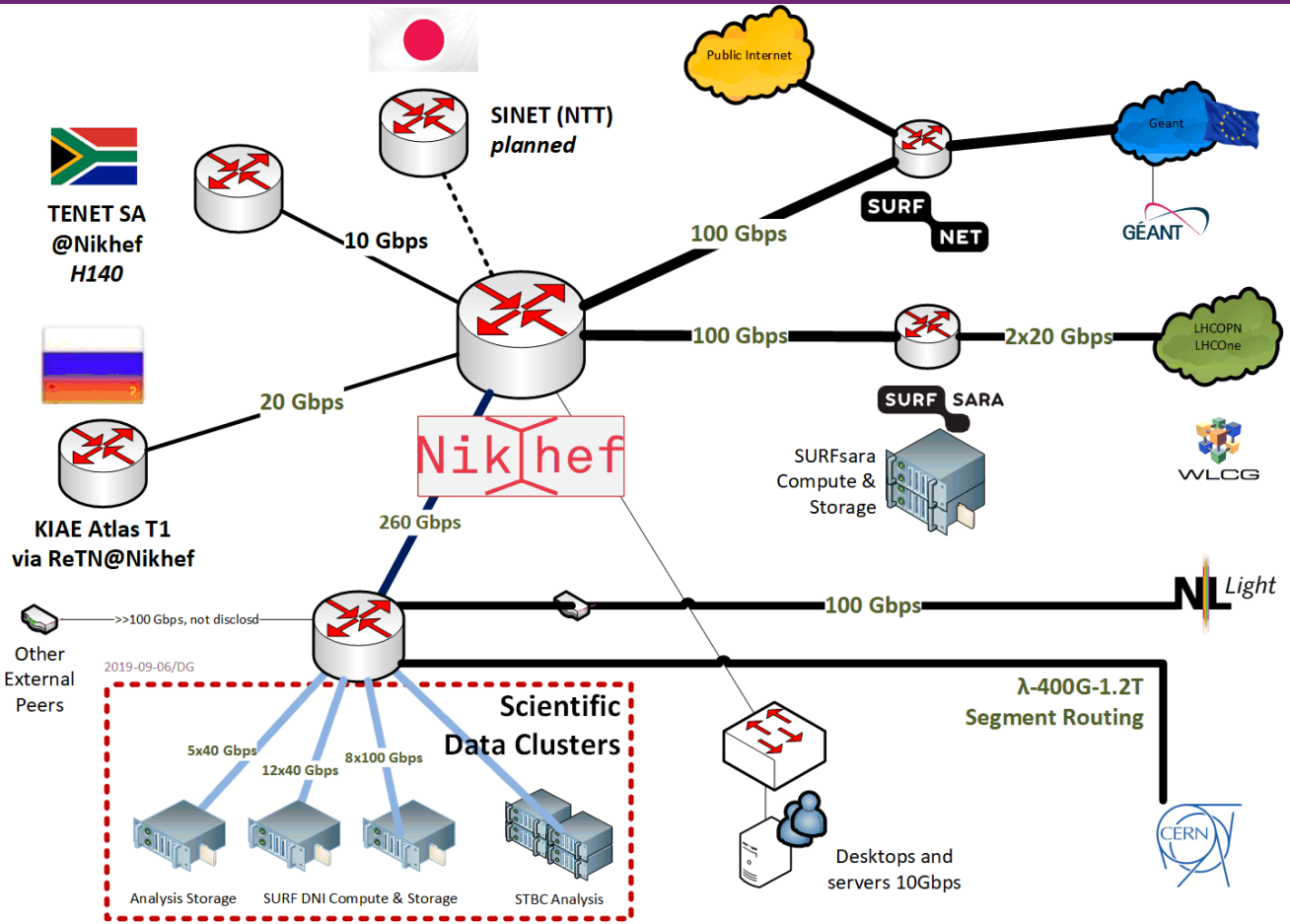
- *Computing* ~ 1,000,000 cores
- *On-line disks* > 310 PB
- *Archival* > 390 PB



‘LHC Open Network Environment’



Direct peering infra for ‘big data’



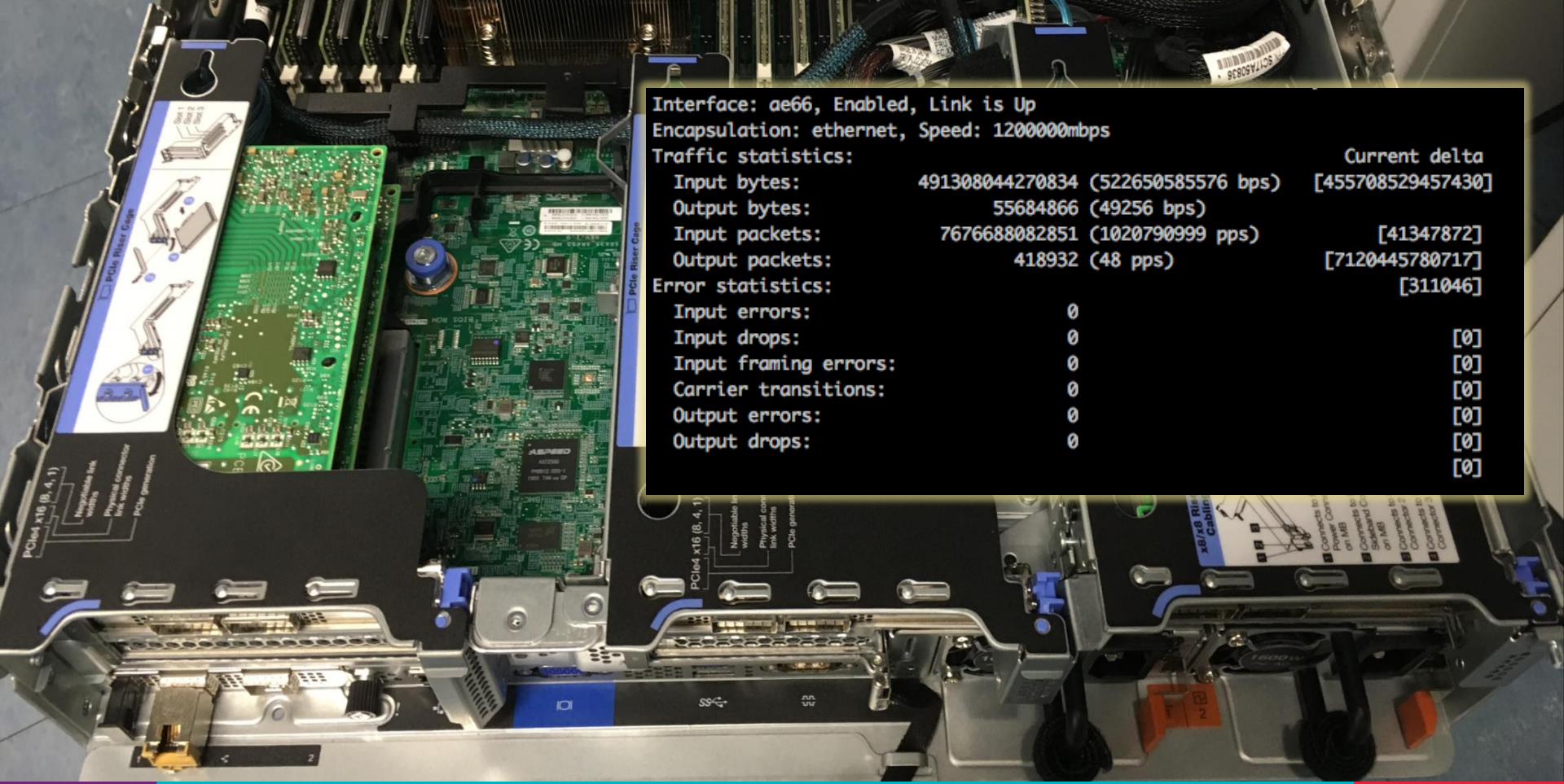
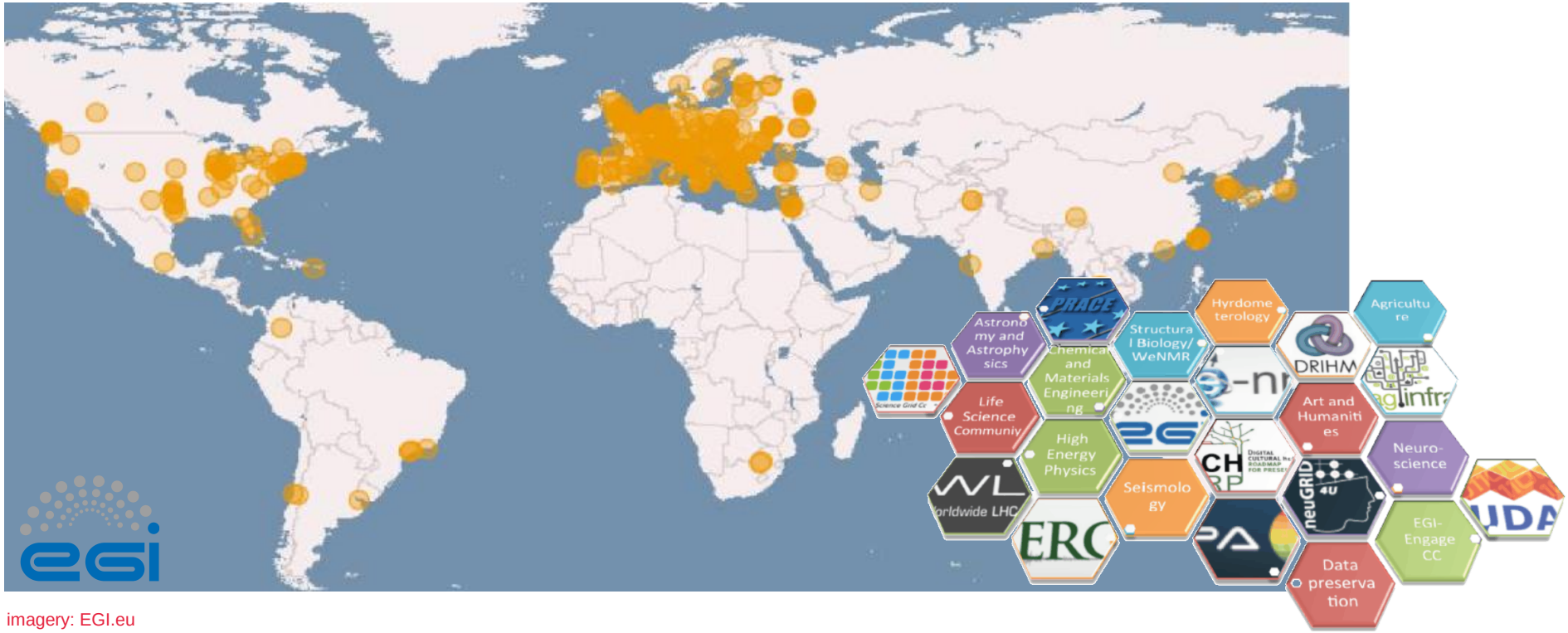


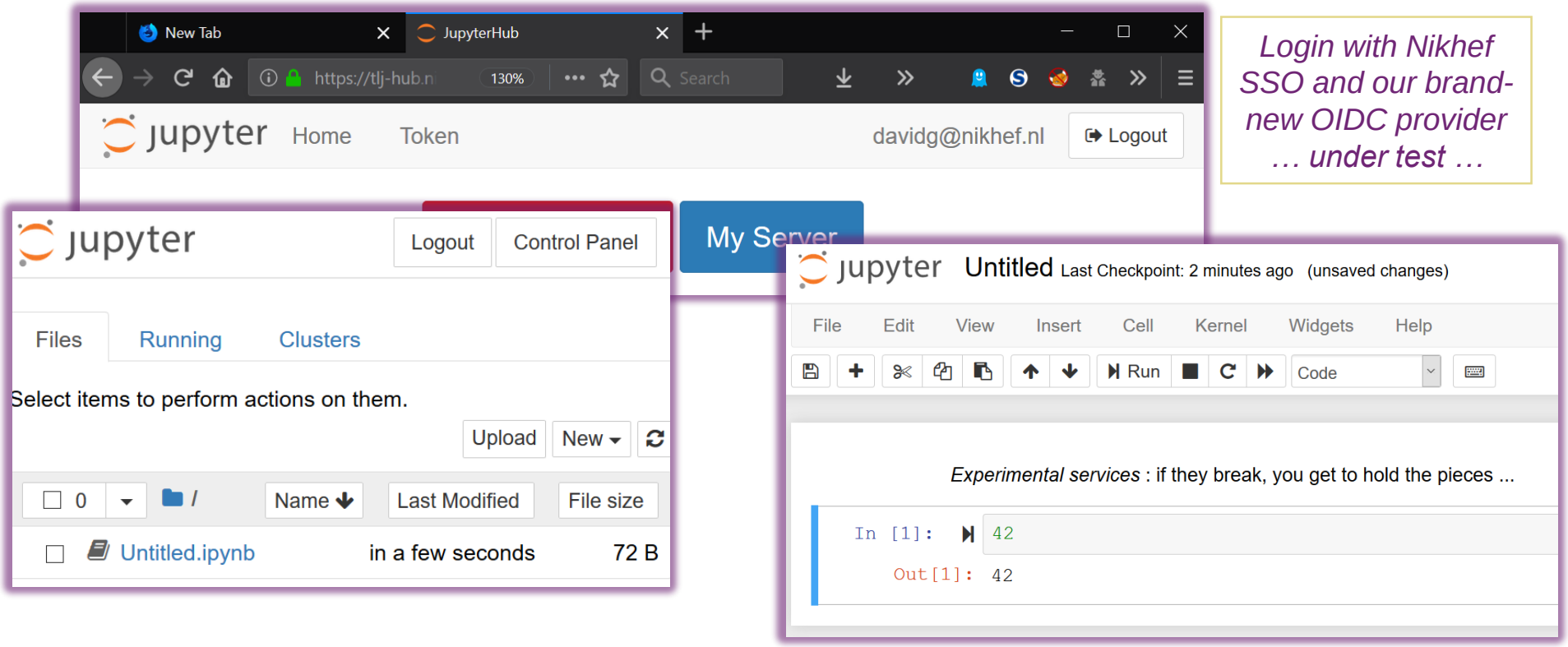
Image: Tristan Suerink

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



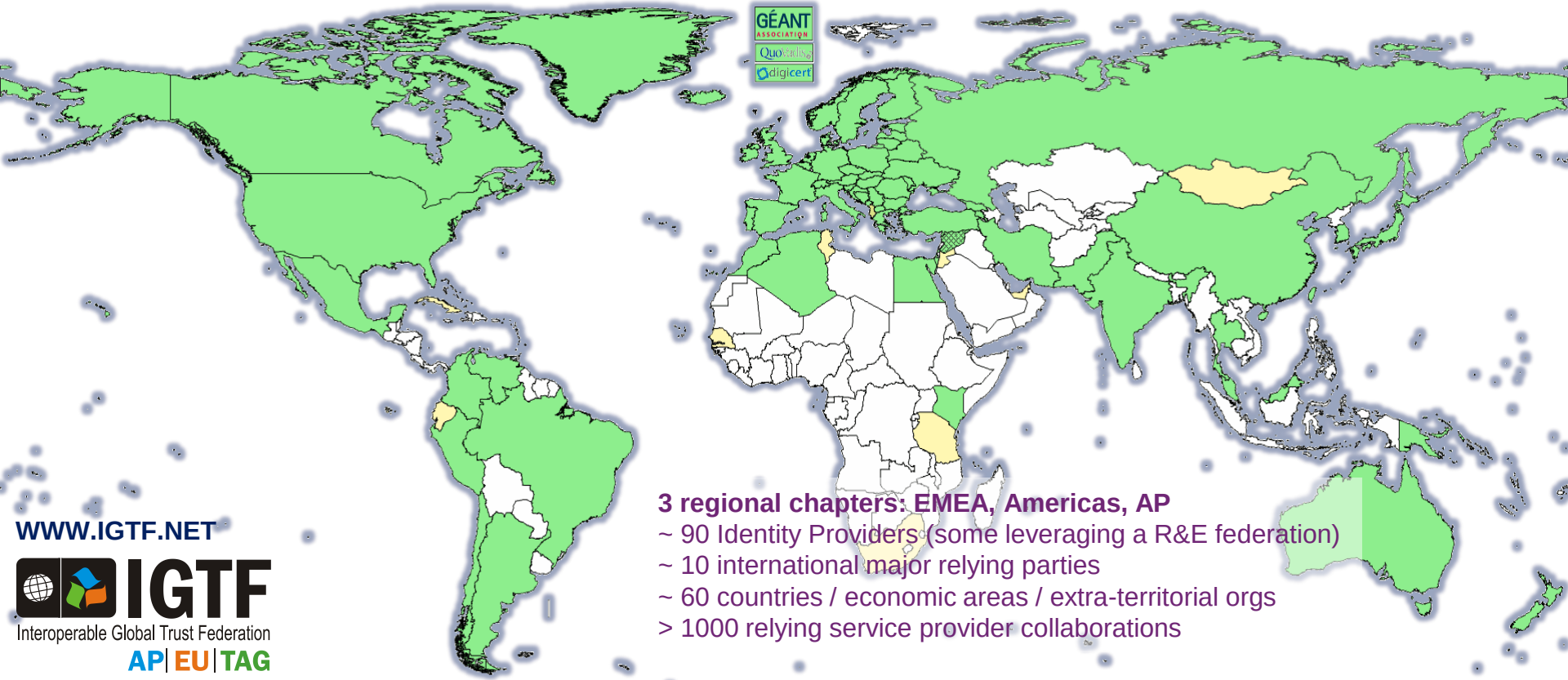
imagery: EGI.eu

LOCAL USERS, REMOTE USERS, …



Login with Nikhef SSO and our brand-new OIDC provider … under test …

INTEROPERABLE GLOBAL TRUST FEDERATION IGTF



WWW.IGTF.NET

IGTF
Interoperable Global Trust Federation
AP | EU | TAG

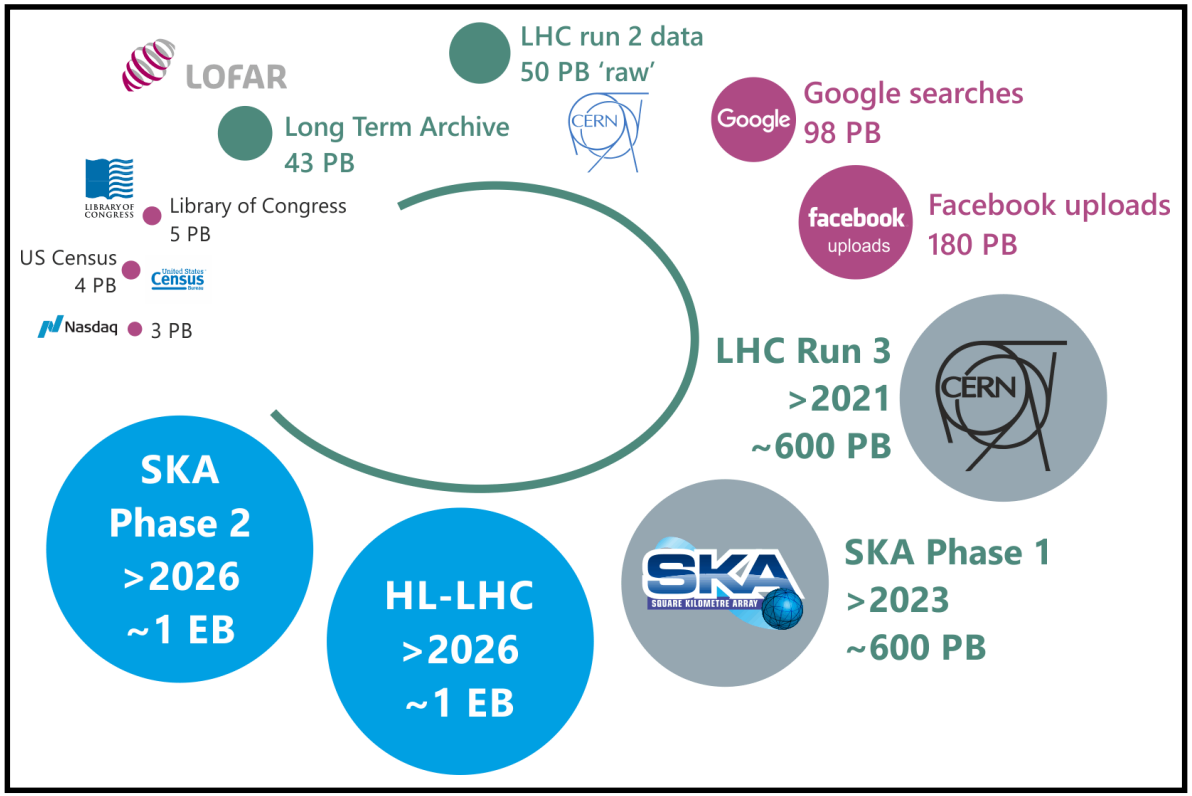
3 regional chapters: EMEA, Americas, AP
~ 90 Identity Providers (some leveraging a R&E federation)
~ 10 international major relying parties
~ 60 countries / economic areas / extra-territorial orgs
> 1000 relying service provider collaborations

DATA DELUGE - AND WITH DATA COMES PROCESSING

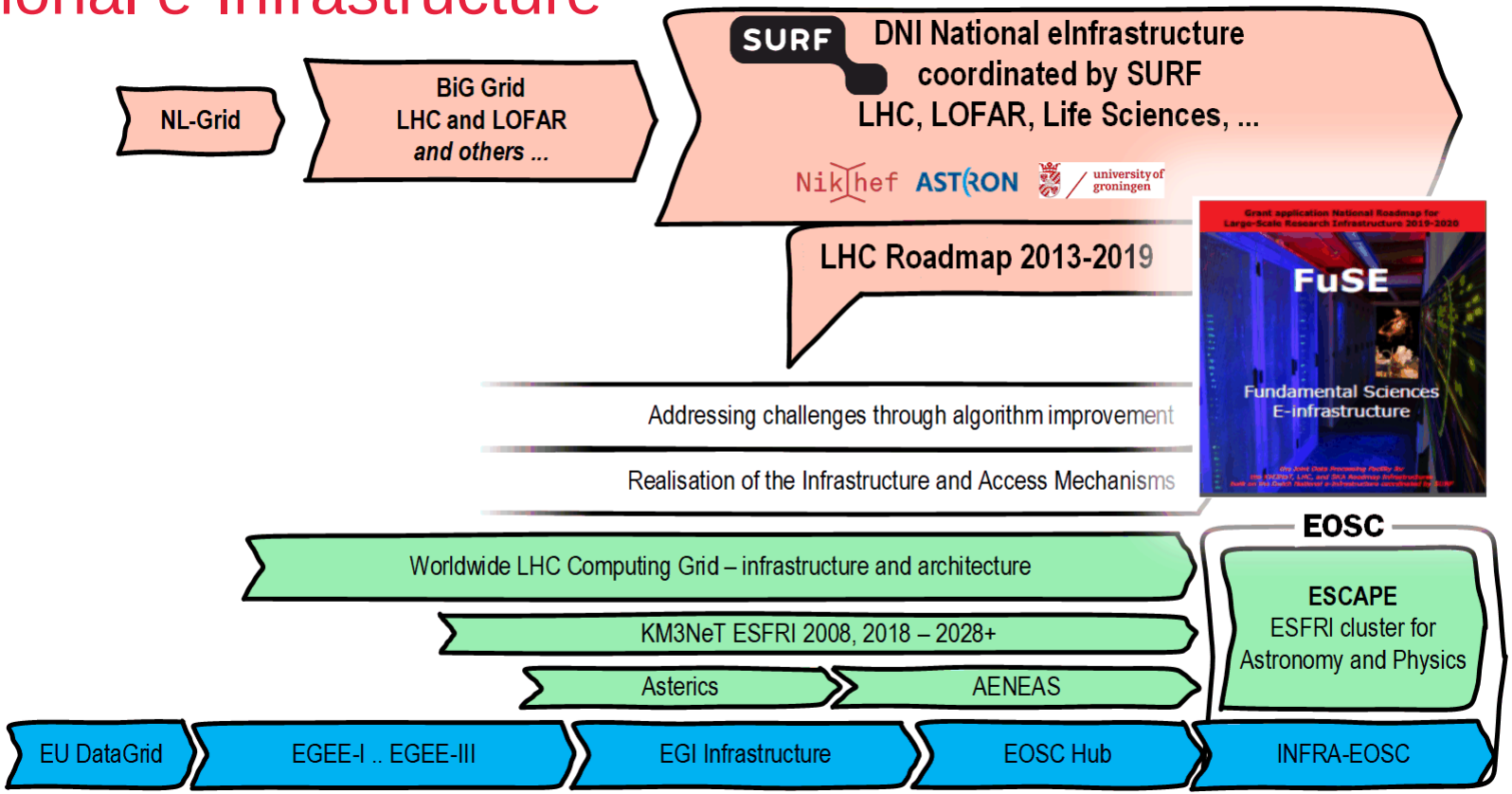
Challenges 2020-2025:

1. Bring ICT infrastructure to a next level
2. Optimize complex algorithms

Prepare now for even bigger challenges >2026



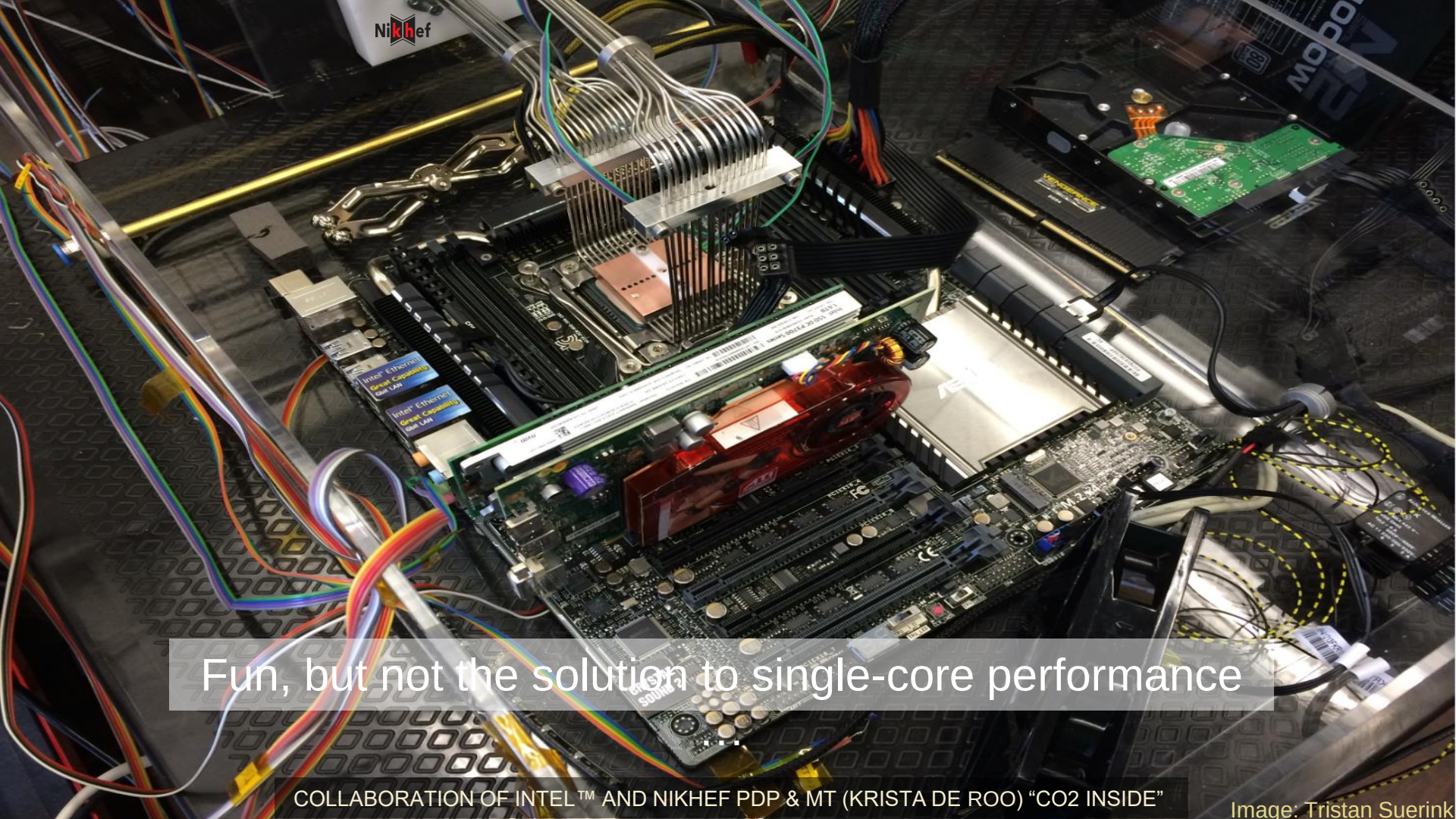
National e-Infrastructure

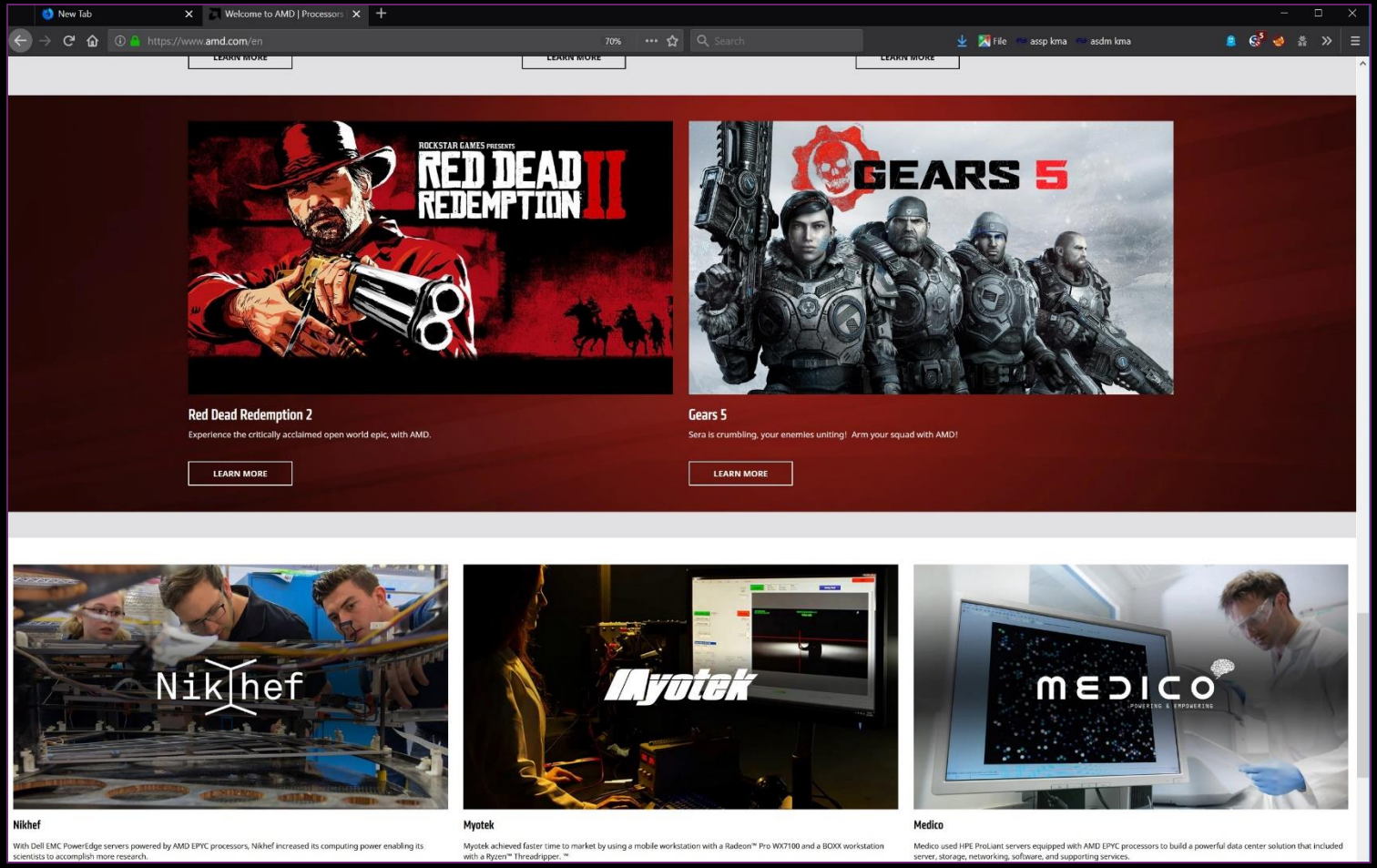


Nikhef

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

NIKHEF PDP - AN OVERVIEW

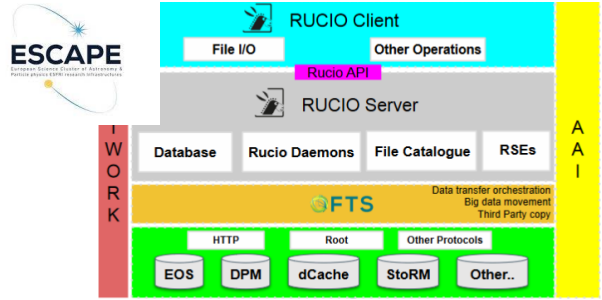

Nikhef
Fun, but not the solution to single-core performance
COLLABORATION OF INTEL™ AND NIKHEF PDP & MT (KRISTA DE ROO) “CO2 INSIDE”
Image: Tristan Suerink


New Tab
Welcome to AMD | Processors
https://www.amd.com/en
70%
Search
File
assp kma
asdm kma
LEARN MORE
LEARN MORE
LEARN MORE
ROCKSTAR GAMES PRESENTS
RED DEAD REDEMPTION II
GEARS 5
Red Dead Redemption 2
Experience the critically acclaimed open world epic, with AMD.
LEARN MORE
Gears 5
Sera is crumbling, your enemies uniting! Arm your squad with AMD!
LEARN MORE
Nikhef
Myotek
MEDICO
POWERING & EMPOWERING
Nikhef
With Dell EMC PowerEdge servers powered by AMD EPYC processors, Nikhef increased its computing power enabling its scientists to accomplish more research.
Myotek
Myotek achieved faster time to market by using a mobile workstation with a Radeon™ Pro WX7100 and a BOXX workstation with a Ryzen™ Threadripper. ™
Medico
Medico used HPE ProLiant servers equipped with AMD EPYC processors to build a powerful data center solution that included server, storage, networking, software, and supporting services.

SMALL PACKETS – BIG TROUBLE? 1.2 BPPS SYSTEM

But long-distance fat networks work fine for ‘big files’

- *bandwidth* x *delay* hampers remote random access
- which is why ESCAPE develops data lake caching



Some data is inherently ‘small’ and packetized

- ‘telemetry’ data
- remote distributed event processing

but sending many packets is far more challenging on network ASIC design – so let’s test it!

Graphic Data Lake, Simone Campana, ESCAPE WP2 DIOS
Image 1Bpps machine by Tristan ☺

OUR WORLD RECORD: 1 BILLION PACKETS PER SEC

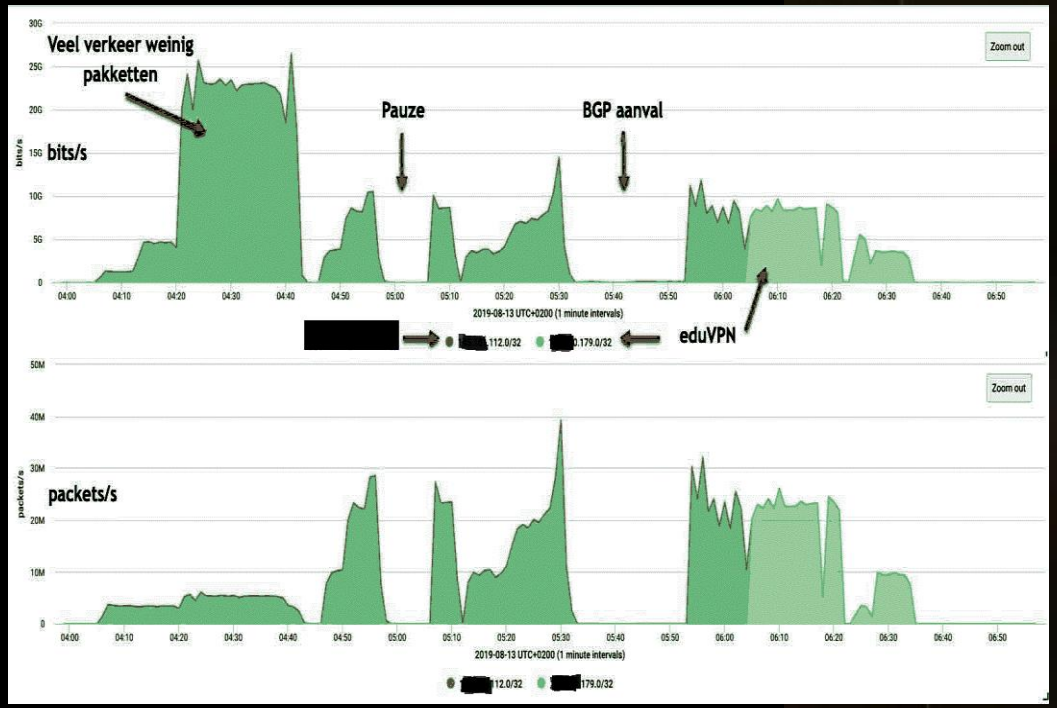
```
Interface: ae66, Enabled, Link is Up
Encapsulation: ethernet, Speed: 1200000mbps
Traffic statistics:                                          Current delta
  Input bytes:          491308044270834 (522650585576 bps)   [455708529457430]
  Output bytes:              55684866 (49256 bps)
  Input packets:          7676688082851 (1020790999 pps)            [41347872]
  Output packets:             418932 (48 pps)                  [7120445780717]
Error statistics:                                                    [311046]
  Input errors:                     0
  Input drops:                      0                                    [0]
  Input framing errors:             0                                    [0]
  Carrier transitions:              0                                    [0]
  Output errors:                    0                                    [0]
  Output drops:                     0                                    [0]
                                                                         [0]
```

14-06-2019: 1 billion pps *i.e.* 1 Gpps (and 522Gbps)

https://wiki.nikhef.nl/grid/1Bpps_Machine

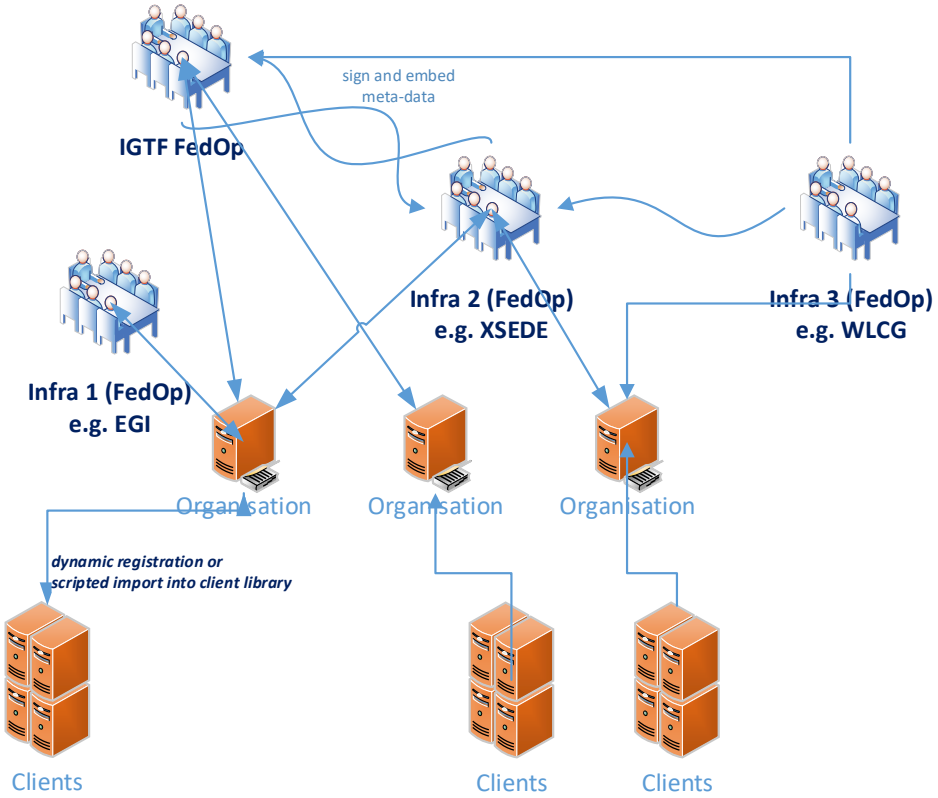
Image: Tristan Suerink

Our science data looks akin to a DDoS

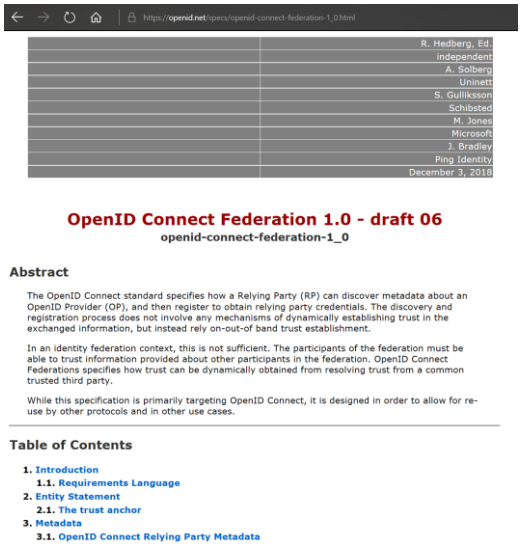


evaluating resilience to cyberattack – *in a cooperative way*

OIDC FED – TRUST IS TECHNOLOGY AGNOSTIC



OpenID Connect Federation: multilateral trust beyond GAFA



R. Hedberg, Ed.
independent
A. Solberg
Uninett
S. Gulliksson
Schibsted
M. Jones
Microsoft
J. Bradley
Ping Identity
December 3, 2018

OpenID Connect Federation 1.0 - draft 06

openid-connect-federation-1_0

Abstract

The OpenID Connect standard specifies how a Relying Party (RP) can discover metadata about an OpenID Provider (OP), and then register to obtain relying party credentials. The discovery and registration process does not involve any mechanisms of dynamically establishing trust in the exchanged information, but instead rely on-out-of band trust establishment.

In an identity federation context, this is not sufficient. The participants of the federation must be able to trust information provided about other participants in the federation. OpenID Connect Federations specifies how trust can be dynamically obtained from resolving trust from a common trusted third party.

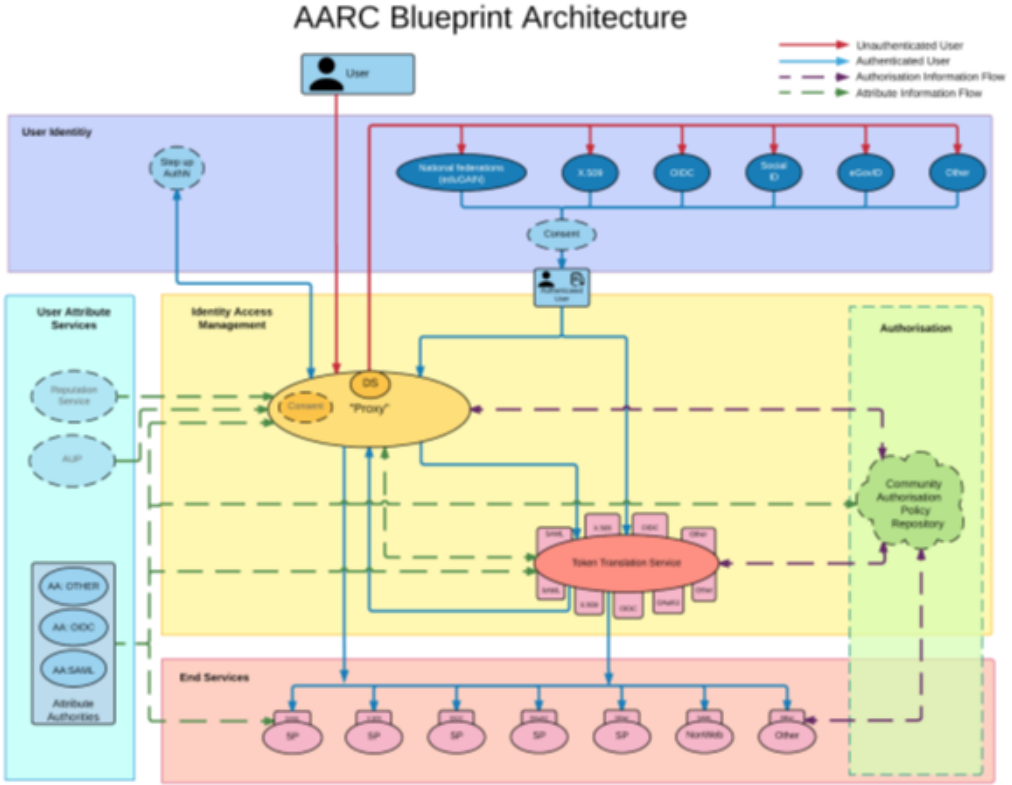
While this specification is primarily targeting OpenID Connect, it is designed in order to allow for re-use by other protocols and in other use cases.

Table of Contents

1. Introduction
   1.1. Requirements Language
2. Entity Statement
   2.1. The trust anchor
3. Metadata
   3.1. OpenID Connect Relying Party Metadata

see: openid.net → Specs & Dev info

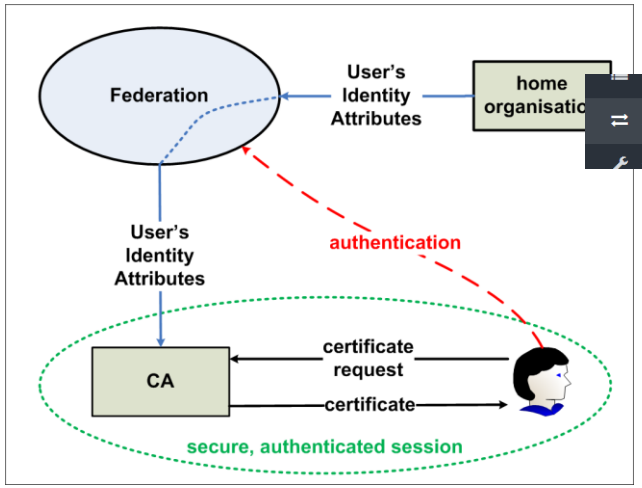
AARC BPA – COHERENCY BY PROXYING

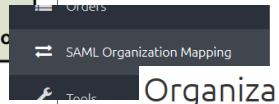


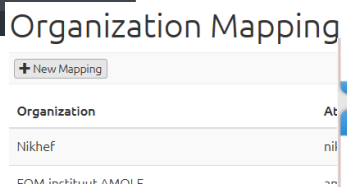
https://aarc-community.org/architecture/

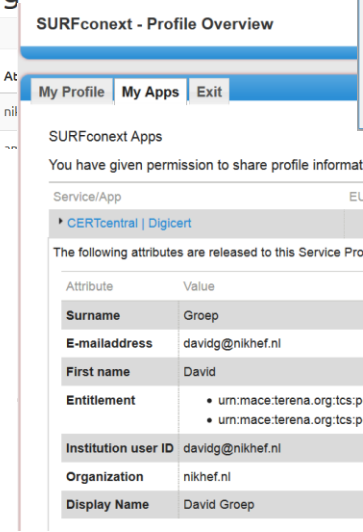
BRIDGES AND TOKEN TRANSLATION SERVICES

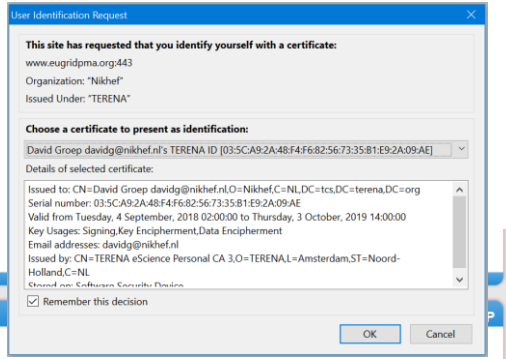
GEANT Trusted Certificate Service











DigiCert acts as a SAML Service provider to eduGAIN and eligible authenticated users can obtain their own certificate for access and delegation to services