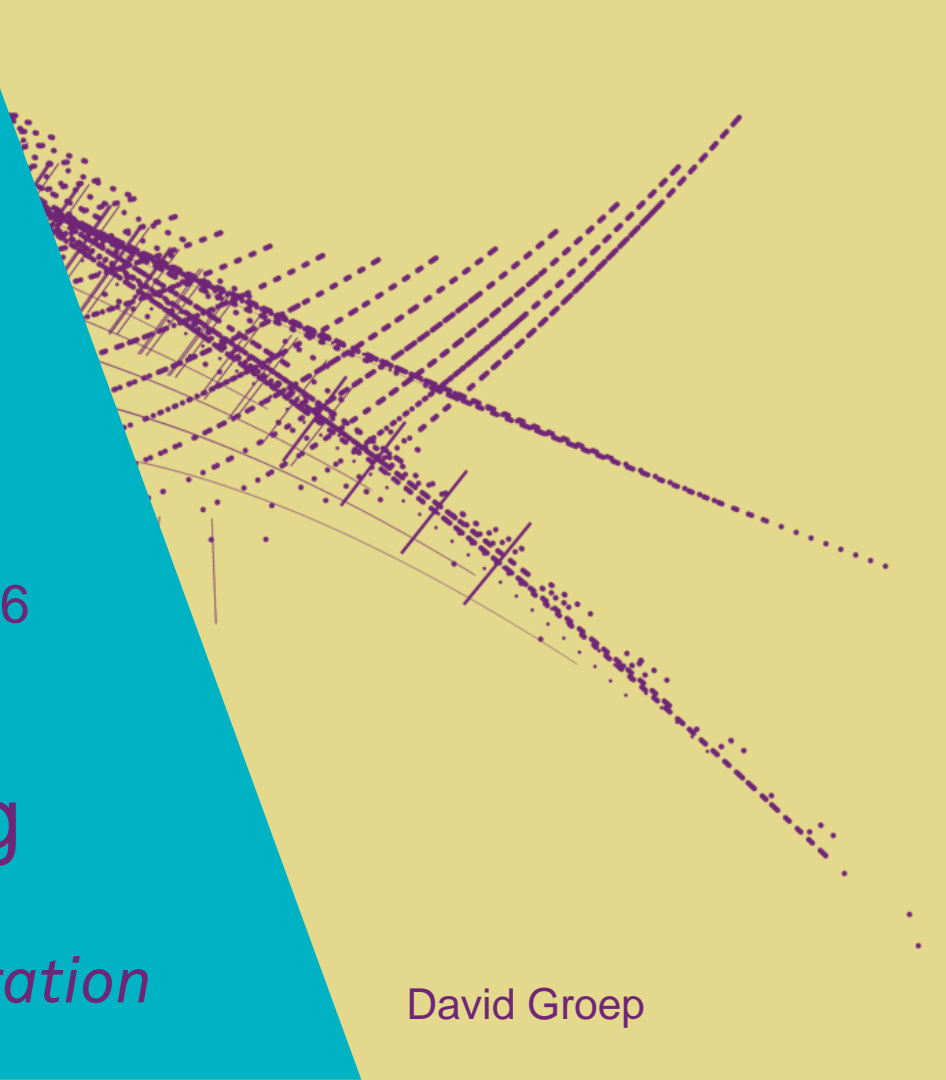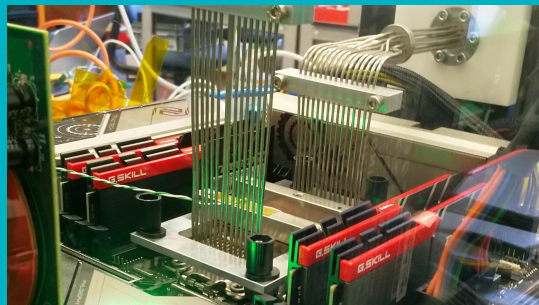scientific platform meeting January 2026

Physics Data Processing
accelerating 'time to science'
through computing and collaboration

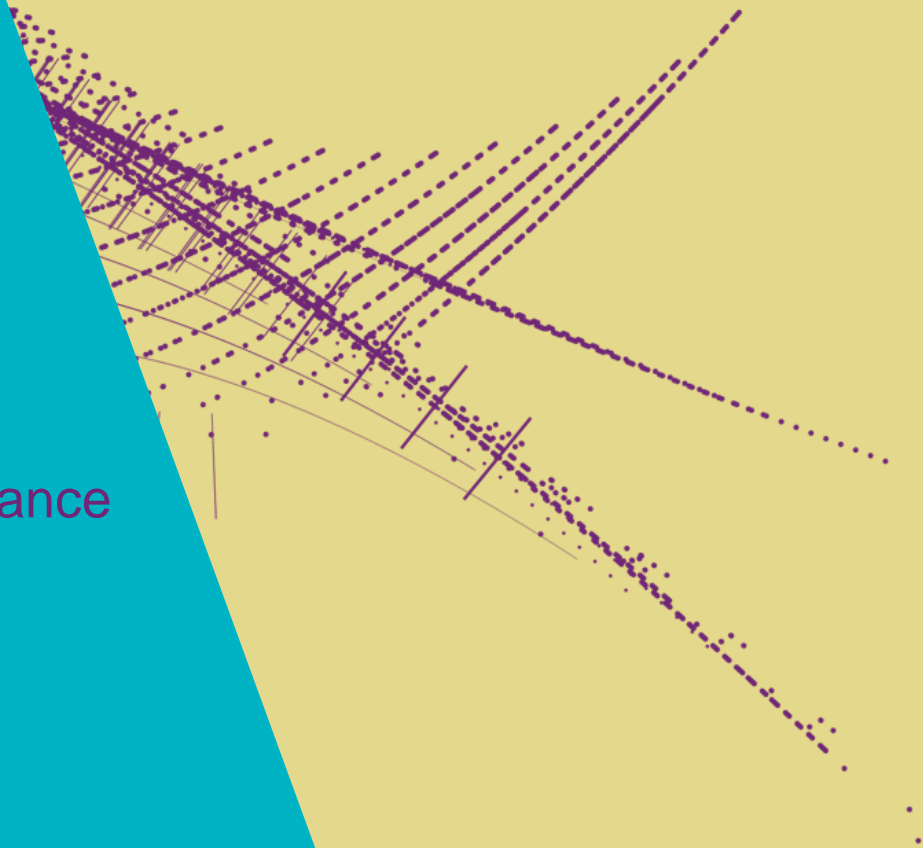David Groep

Today with a joint PDP-CT-PDP attendance

David Groep - PDP
Gerhard Raven - {LHCb, PDP}
Jeff Templon - PDP
Mary Hester - CT-PDP
Roel Aaij - PDP
Ronald Starink - CT-.*

25 years of European data driven computing at Nikhef
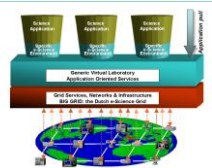# 1-1-2001 … 1-1-2026

+ 2 years of WTCW Virtual Laboratory!

accelerating 'time to science' with computing and collaboration

Nikhef

The pillars of Physics Data Processing

Algorithmic design patterns and software

- GPU & accelerator software, algorithms, heterogeneous processors optimisation
- algorithmic design R&D
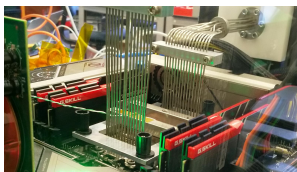
Infrastructure, network & systems co-design R&D

- building 'research IT facilities' co-design & development
- data science at volume scale
- research *on* IT infrastructure
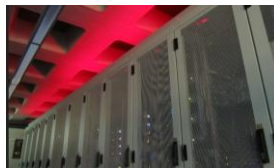
Infrastructure for trusted collaboration

- trust and identity for enabling communities
- managing complexity of collaboration mechanisms
- securing the infrastructure of our open science cloud

Infrastructure for scalable ML in HEP & APP

Inference optimisation for data throughput,
HTC/HEP and HPC/AI-factory convergence, hybrid workflows
"walking the 'staircase' for analysis and scheduled processing"
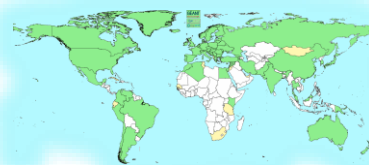
T&I, Federation, Research Infra Commons

accelerating 'time to science' through computing and collaboration

Nik|hef

Means and purpose of processing

- **Accelerate** time to results with **efficient use** of computing through algorithmic design, novel software methods, and optimisation of data driven workflows and models
- exploit **new systems design and integration:** heterogeneous architectures, innovative vendors, faster networks
- ensure physics **access to the broadest range of science services** with ease

Meanwhile

- ensure our computing capacity is sustainable … as *we fund it to a large extent from non-HEP/APP sources, and have a commitment to open science and collaboration*
- continue push for generic access of services … *important to keep individual RIs aligned*
- move more of our experiments to generic computing infrastructure … *so we can contribute compute more effectively, cheaper, better to our physics programmes*

accelerating 'time to science' through computing and collaboration

Nik|hef

Some computing, engineering and trust highlights

- seeing significantly increased **use of GPUs** in our analysis facility and for AI/ML

- Completely migrated to HTCondor for both local analysis and federated facility

- continued **94% occupancy** of our NL-federated Tier-1

- 'NL-T1' **compute for IWGN now >15%** of global capacity*

- Faster networks & more packets with 'interesting' effects

- Security and risk management **for EOSC EU Node**

- **AARC architecture (G080) and trust framework (PDK)** and with tokens for more than just IGWN and WLCG
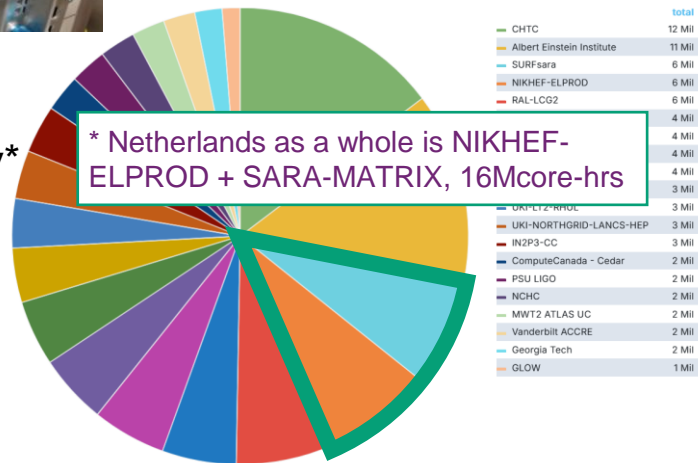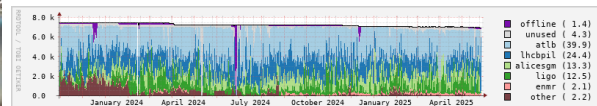
offline (1.4)
unused (4.3)
atlb (39.9)
lhcbpil (24.4)
alicesgm (13.3)
ligo (12.5)
enmr (2.1)
other (2.2)

* Netherlands as a whole is NIKHEF-ELPROD + SARA-MATRIX, 16Mcore-hrs

	total
CHTC	12 Mil
Albert Einstein Institute	11 Mil
SURFsara	6 Mil
NIKHEF-ELPROD	6 Mil
RAL-LCG2	6 Mil
	4 Mil
	4 Mil
	4 Mil
	3 Mil
UKI-LT2-RHUL	3 Mil
UKI-NORTHGRID-LANCS-HEP	3 Mil
IN2P3-CC	3 Mil
ComputeCanada - Cedar	2 Mil
PSU LIGO	2 Mil
NCHC	2 Mil
MWT2 ATLAS UC	2 Mil
Vanderbilt ACCRE	2 Mil
Georgia Tech	2 Mil
GLOW	1 Mil

accelerating 'time to science' through computing and collaboration

Nikhef

Last year's SAC observations

The PDP group has an ambitious remit, ranging from developing and supporting the use of modern algorithms on modern hardware, to the facility deployment at Nikhef (together with Nikhef Computing Technology), and essential collaborative infrastructure for scientific computing in international projects.
The team has established good relationships with the physics teams locally and at the university partners and has a focus on training and help-desk support for software issues.

As software is clearly critical (and frequently under-supported) across the institute's physics activities, we find the team size to be quite small for this role. Growing links to research software engineering teams at the university partners may be a means to broaden the support base for software and help address important issues of sustainability. Due consideration of software lifecycles should be given and new software technologies pursued. In particular, growing use of AI/ML is being observed and this means that the PDP group should also play an important role in both resource provisioning and skills training for new ML-based software techniques.

Resource delivery for international projects benefits hugely from stressing common interfaces and solutions, which we commend. Keeping in-house resources reduces risks and maintains vital expertise.

Report of the Nikhef Scientific Advisory Committee Meeting June 5-6, 2025

accelerating 'time to science' through computing and collaboration

'Keeping in-house resources…'

We need our operational capability to be relevant* in the ecosystem

- we *are* the 'staircase model' put into practice: scaling from STBC to DNI national and global resources – and influencing these by example Naturally aligned with the Research Infrastructure Commons

- we provide the 'rapid-response' capability of the innovation platforms and the (national) 'Experimental Technologies Platform' for SURF

outsourcing (or even off-site co-location) of facilities would jeopardize that role in innovation – and be detrimental to our talent case …

But we (cannot) do not do this alone

- MERGE proposal explicitly includes the Commons as mechanism to import and export expertise on AI/ML and infrastructure usage

accelerating 'time to science' through computing and collaboration

* thanks Wennink ☺
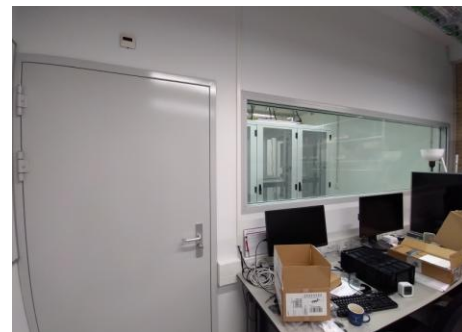
'… benefits hugely from stressing common interfaces & solutions'

Systems design for exploitation of scalable resources for scheduled processing and chaotic analysis: a combination of research & development + operations & support

In acquisition and real-time acceleration for 4D
- FASTTRACK accelerated R&D facility (WP5.1), and algorithmic work *in the experiments themselves* for WP5.2
- Jointly with SURF ETP innovation collaboration

In making HPC resources (and AI factories) a better fit for data intensive HEP/APP
- with SURF: workflows on Snellius preparing for 'Lorentz' tender – leveraging ATLAS & Alice
- EuroHPC Federation and the Tier-1→Tier-0 transition:
  systems design, AAI, Interfaces to Architecture Testbeds
- and alsi just the basics: helping promote the NWO 'Rekentijd' granting scheme

In developming a smooth transition 'staircase' path
- TDCC NES project on JupyterHub alignment Stoomboot and Snellius
- OSNL project on reproducible software at scale (with 4TU, TUDelft, NLeSC)

accelerating 'time to science' through computing and collaboration

Nikhef

Ecosystem approach to sustainability and resourcing

For our sustainability and impact

- aim to be present in each element of the Commons with a non-trivial contribution

- choose control-point elements to mitigate relatively small size

to ensure Nikhef can benefit from both national and global community resources.

'Large users' like us have
a real impact on this global ecosystem!

GORC IG: Typology and Definitions, https://doi.org/10.15497/RDA00087

Global Open Research Commons - Essential Elements

GOVERNANCE STRUCTURES

RULES OF PARTICIPATION AND ACCESS

ENGAGEMENT

HUMAN CAPACITY

STANDARDS

SUSTAINABILITY

COMPUTE, STORAGE, NETWORK, AAI

SERVICES & TOOLS

RESEARCH OBJECTS

V1.3, GORC:IG, 2024

Nikhef
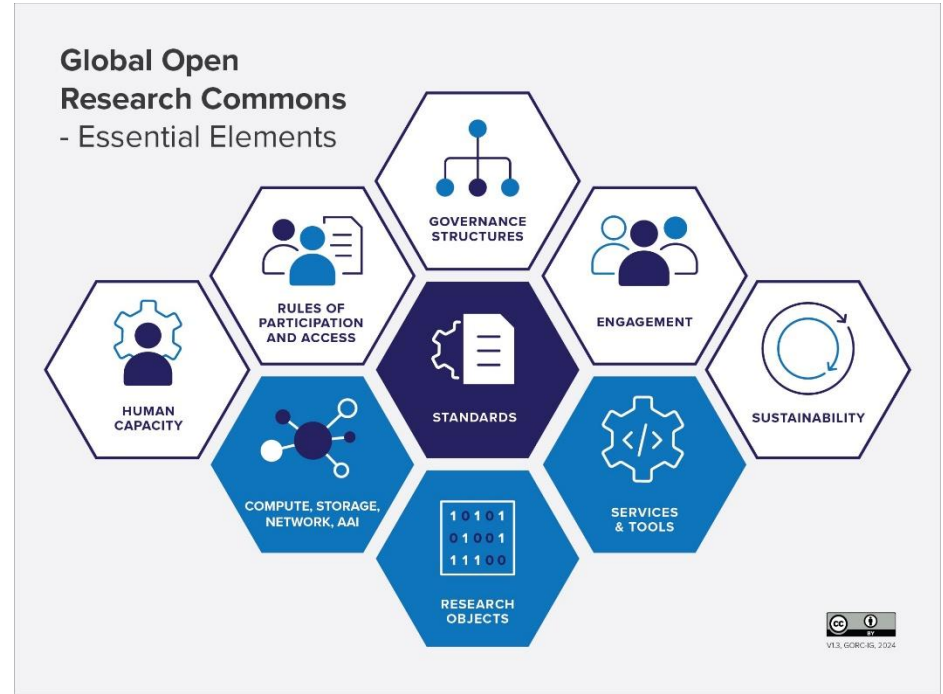
Innovation: '… growing use of AI/ML is being observed' …

New AI/ML R&D group brings activities together in *both* the experiments *and* for PDP

- links the 'data, physics, workflows, and facilities in the 2040s' role of PDP with stronger complement in the physics programmes on (AI/ML) software

- R&D on infrastructure extensions for AI/ML and accelerated compute generally

We will be evolving

- the physics staff and PoC hardware will be available starting mid-2026

- facility for inference (MERGE) still under review

While other aspects remain as relevant in the future

- systems R&D and our role as a 'Tier-1' facility for HEP/APP

- in a complex, multi-stakeholder world, mixing long and short term 'political' objectives

- accelerating time to results from 4D fast timing close to the detectors to the final plots ☺

accelerating 'time to science' through computing and collaboration

Balancing training & support with innovation R&D

The pathway to practical impact is via training and hands-on support
… with infinite demand and always limited capacity ☺

- computing courses need investment from two sides
- getting specific AI/ML workflows to run on both locally
  and on 'scale-up' systems from STBC → Snellius → LUMI
  → hopefully our federated HTC / Tier-1 in 2028+
- and also RDM and FAIR 'digital research process' courses

Here we lack the 'sweet spot' for trainers: it is beyond the helpdesk expertise,
and then our next layer is > level-3 engineering researcher support …
… but it's nice to be able to give training and see the effect of your R&D efforts!

accelerating 'time to science' through computing and collaboration

Nik|hef

Pathways via projects and collaborations

New initiatives and projects

- *strengthen* the strategic areas
- ensure *continuity* of research and infrastructure

future project pathways planned: MERGE, SHAKE, …

Public partner R&D engagement
AMD, Nokia/Cerebras, NL-ix, AmsIX…
Dutch national government

accelerating 'time to science' through computing and collaboration
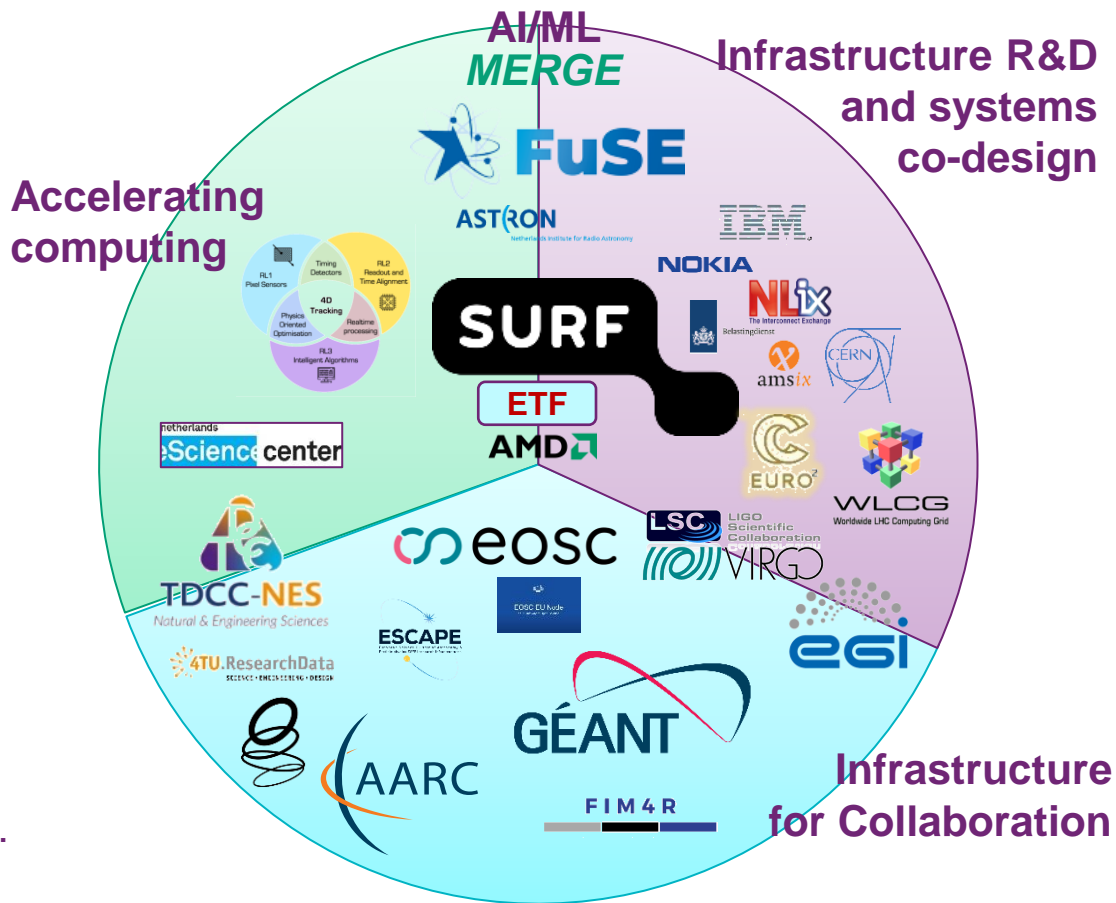
Weaknesses, threats … and mitigation scenarios

Weaknesses

- Continuity as 'cross-domain Tier-1 hub' based on project, not institutional, funding
  → fiche 'digital infrastructure', NWO-SURF report 'Reken er maar (niet meer) op'
  … MERGE is only a stopgap measure for this, as, again, it is an upgrade project

'the true effect of flat budget'
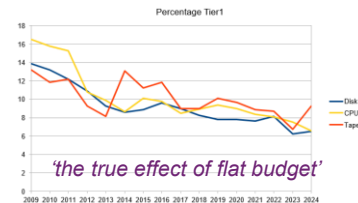
Threats

- Market developments in semicon ☹
  → phased and deferred acquisition, market research, and vendor engagement/co-innovation to get better conditions is the best we can do – but likely not enough for 2026-2027
- Divergence of HEP/HTC and EuroHPC/AI factories
  → co-design of new Tier-1 systems and ensure 'Lorentz' as national facility, collaborate to instigate change in EuroHPC (SURF, EuroCC, SHAKE)
- Balancing operations, support and research as algorithms (and AI) become more important
  → engage physics groups in software engineering (as per the SAC2025 report), more 'project' software engineering and computing/AI/ML training capacity in CT-PDP
- Risk of divergence WLCG from most 'others', even in our own domain, remains significant
  → invest in control points such as HPC CoEs and in AAI, policy, and 'OpSec'

accelerating 'time to science' through computing and collaboration

And some resource-centric ones … no realistic mitigation?

Weaknesses

- We provide a lot of research services, of all kinds, but disinvesting is not usually an option
  → we possible we can join up with others (SURF, OSG, other GWIs in the Commons, maybe NWO-I), but these in turn look back to us …

Threats

- Lack of people on the CT-(PDP) side in the longer term: with FASTTRACK, MERGE, and AI/ML the demand keeps growing.
  → the requested 'matching' engineering resources, esp. FASTTRACK, need to become available in CT

- Too many SPOFs, esp. in I4C, 'open science', and policy activities
  → 'dakpan Jeff' needs to take care of anchoring our 'horizontal' knowledge as well
    – makes for a very challenging job indeed!

accelerating 'time to science' through computing and collaboration

Opportunities and strength … (no mitigation scenarios ☺)

Strength
- Impact in our chosen focus areas: Accelerated computing, Infrastructure, Trust & Identity,
  → we are seen and proactively asked to partake in projects, and seen as trusted and active partner

Opportunities
- Acceleration advances and ML help us exploit new systems
  → alignment with EuroHPC (and AI factories) for HEP/APP use cases as the EuroHPC Federation forms
- Design of next gen national Tier-1 ('Lorentz' ?) facility and (to some extent) AI factory
  → improve suitability for HEP use cases through design choices
- Cross-programme AI/ML activity raises profile of computing in our programmes
  → may get us the audience awareness for the importance of computing and algorithms
- Heterogeneous and energy-efficient computing
  → more 'physics output per unit of .*'
- National Plan for Research Digitalisation ('fiche')
  → *if* granted, we can leverage far broader ecosystem at the institute *and* partner universities!

accelerating 'time to science' through computing and collaboration

Physics Data Processing
towards results unconstrained by computing

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

Re-strengthening the PDP in 2026

'working on the interface between systems research for AI/ML, data processing infrastructure for high-throughput computing, and accelerated computing - combining hands-on systems operation, direct engagement with user-physicists, combined with long-term strategic vision …'

'The new colleague will play an important part in designing and determining how Nikhef will continue to engage with and strengthen our national, European, and global computing ecosystems.

Key elements of the position include also the liaison with our national partners and our international collaborations at a strategic level, the acquisition of funding and management of projects, and shaping the 'computing systems aspects' driving the machine learning and AI/ML activities at Nikhef.

The candidate has a PhD in experimental sub-atomic physics or a very closely related field and extensive experience in the organization and use of data-intensive computing in large physics experiments. Systems operation experience in (at least one aspect of) high-throughput computing, data management systems, or networking, is expected.'

accelerating 'time to science' through computing and collaboration

Nikhef

Next Gen networks – connecting our data and people
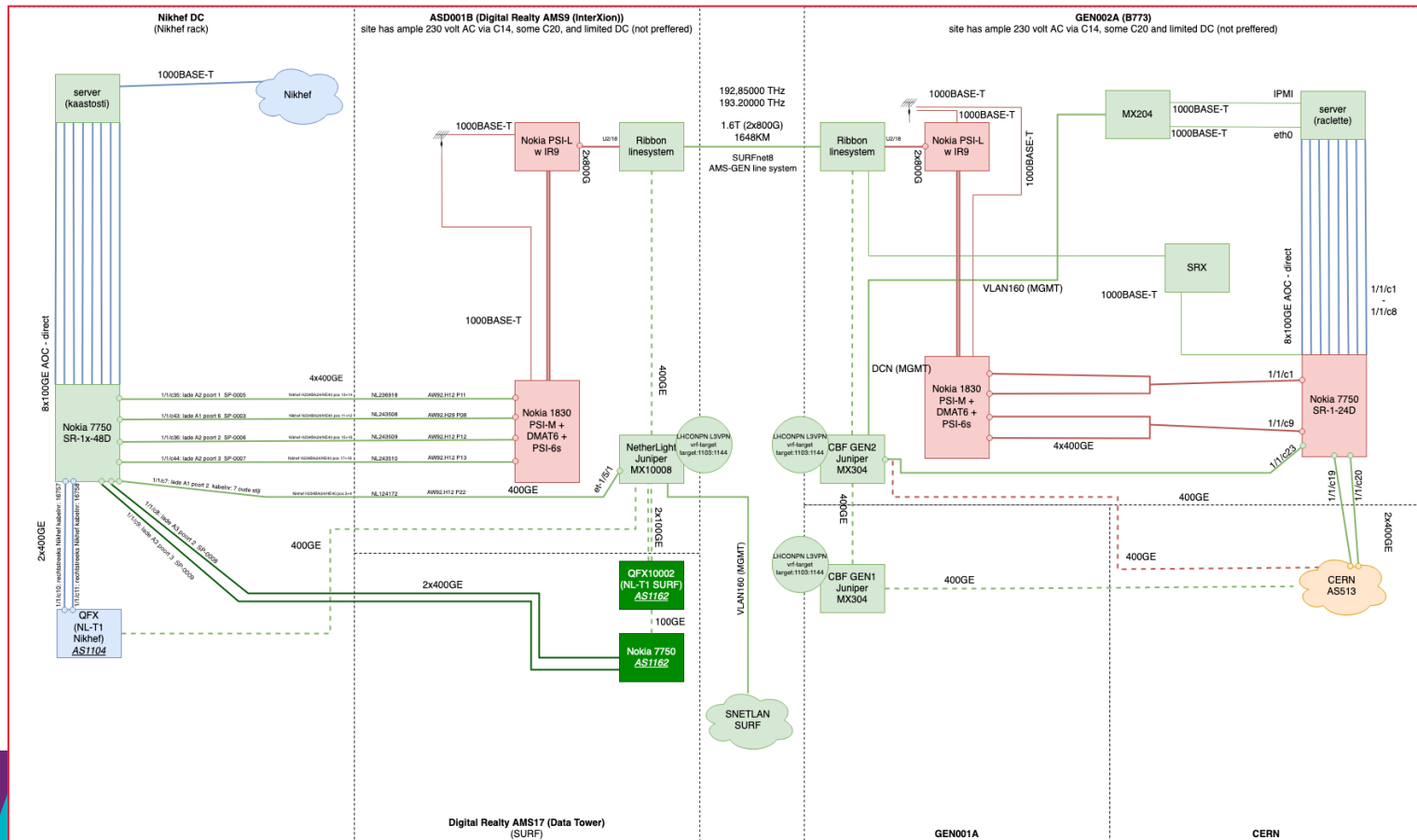
Foundational Principles for Digitalisation at Nikhef

Still needs explanation – despite digital sovereignty discussions. For both research and 'enterprise' ICT.
And 'if they are many, they are not principles', so we set only four:

1. **Institutional strategy and mission directs ICT decision making**
    ICT decisions are assessed based on the Nikhef strategic themes

2. **Collaboration as a core value**
    Nikhef stands for the whole of the Dutch community in (astro)particle physics and its European and global collaborations

3. **Shared public values and responsible technology**
    Nikhef employs, develops, and shapes technologies that preserve autonomy, justice and humanity, that builds on our academic sovereignty and integrity

4. **Digitalisation reflects the continuity in our research programmes**
    With research horizons measured in decades, ICT reflect this continuity in its choice of infrastructure, services, and information management, and in its human expertise

accelerating 'time to science' through computing and collaboration
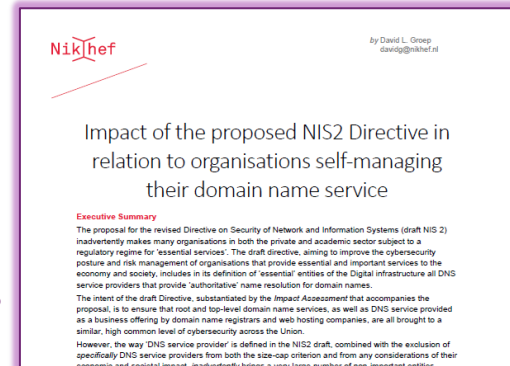
Policy development in ICT and our collaborative values

Our Research Infra (and Open Science) needs a *collaborative* 'values framework'
- frequently threated by increasingly 'corporate' approach to ICT services
- continuous remedial action needed at many levels: from European Commission and EP, down to even our 'own' centralized national institutes organization …

Continued vigilance on IT infra is part of PDP programme
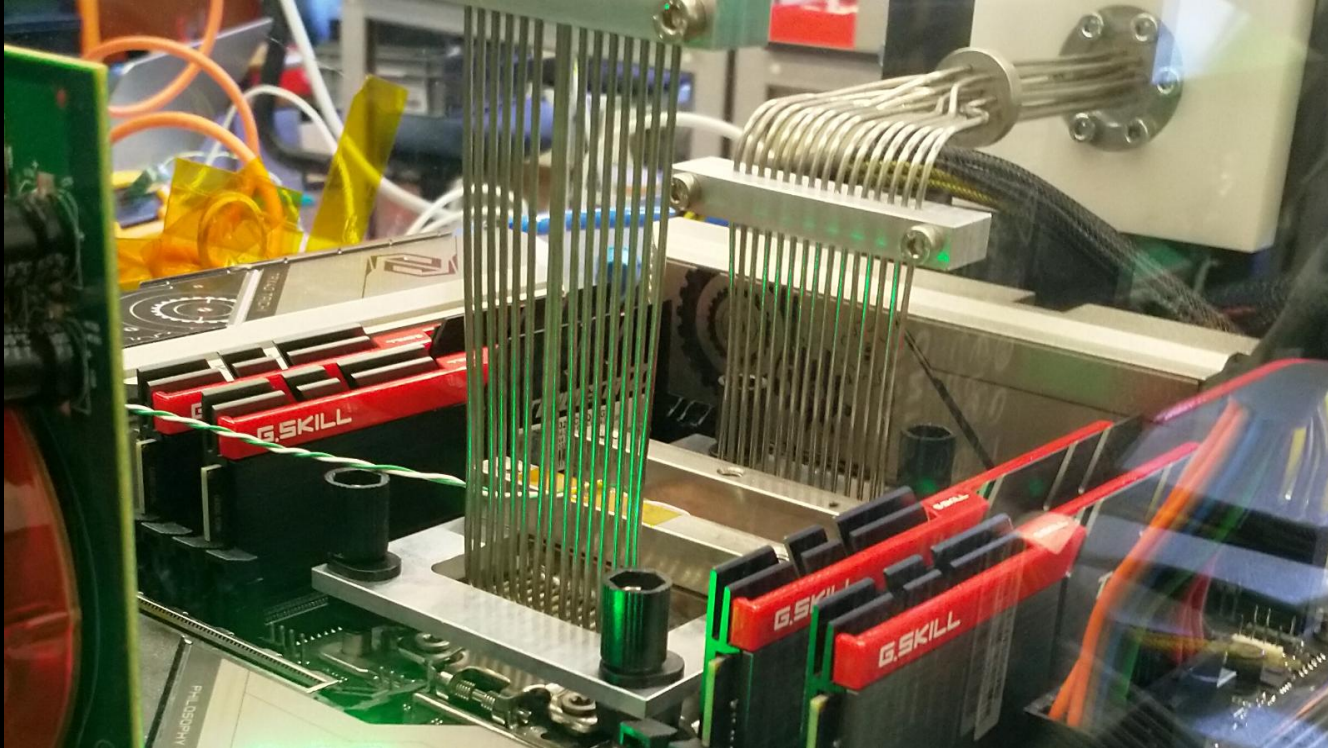to keep our research ICT infrastructure open, *e.g.* through:

- watch carefully and continuously NWO's push to 'corporate IT' and its impact on academic integrity & freedom;
- promote trustworthy federated access via AARC (globally), SRAM (our SURF national scheme), promote trust & identity with our university partners
- Digital sovereignty white papers advocating 'relevant' Netherlands & Europe
- build scalable security solutions rather than pay corporate 'ISO tick-box' providers

luckily many times in collaboration with GEANT, JISC, STFC, EGI, and our peer institutes in NL

Nik|hef

by David L. Groep
davidg@nikhef.nl

Impact of the proposed NIS2 Directive in relation to organisations self-managing their domain name service

Executive Summary
The proposal for the revised Directive on Security of Network and Information Systems (draft NIS 2) inadvertently makes many organisations in both the private and academic sector subject to a regulatory regime for 'essential services'. The draft directive, aiming to improve the cybersecurity posture and risk management of organisations that provide essential and important services to the economy and society, includes in its definition of 'essential' entities of the Digital infrastructure all DNS service providers that provide 'authoritative' name resolution for domain names.

The intent of the draft Directive, substantiated by the *Impact Assessment* that accompanies the proposal, is to ensure that root and top-level domain name services, as well as DNS service provided by specifically DNS service providers from both the size-cap criterion and from any considerations of their economic and societal impact, are all brought to a similar, high common level of cybersecurity across the Union.

However, the way 'DNS service provider' is defined in the NIS2 draft, combined with the exclusion of specifically DNS service providers from both the size-cap criterion and from any considerations of their economic and societal impact, inadvertently brings a very large number of non-important entities

https://doi.org/10.5281/zenodo.4629136

Nik|hef

Because we can … does not mean it's the scalable way ☺

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)

accelerating 'time to science' through computing and collaboration