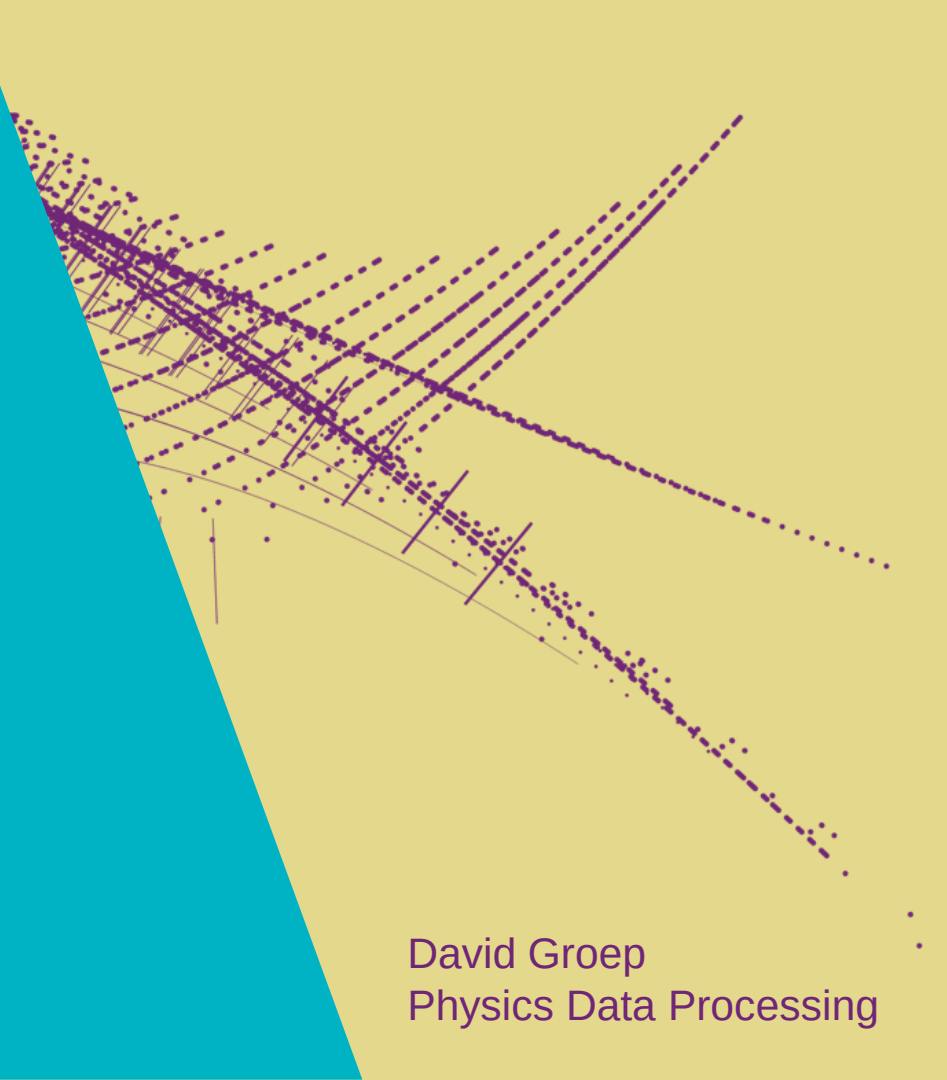
Nikhef

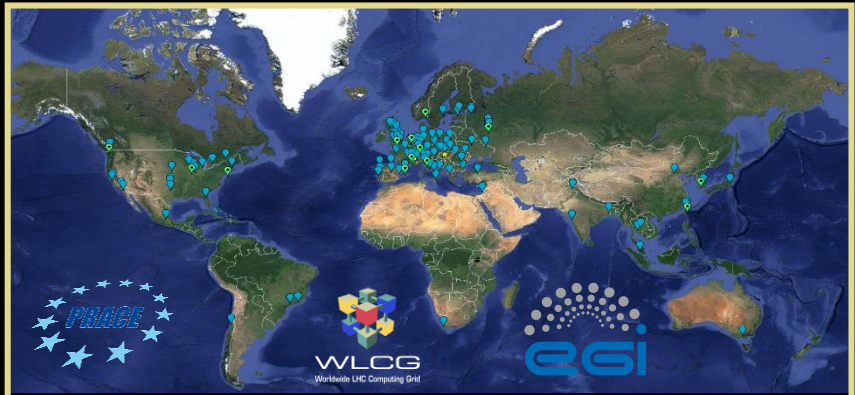


February 2022

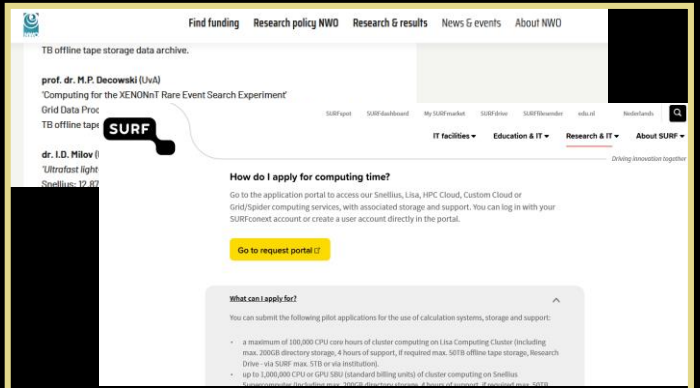
Computing resources and research in the Nikhef context

David Groep
Physics Data Processing

Nikhef and the federated e-Infrastructures resources

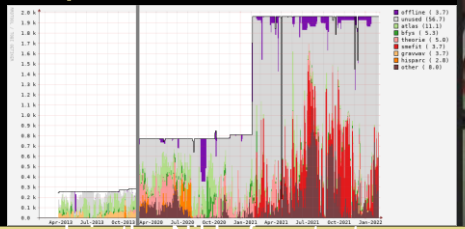


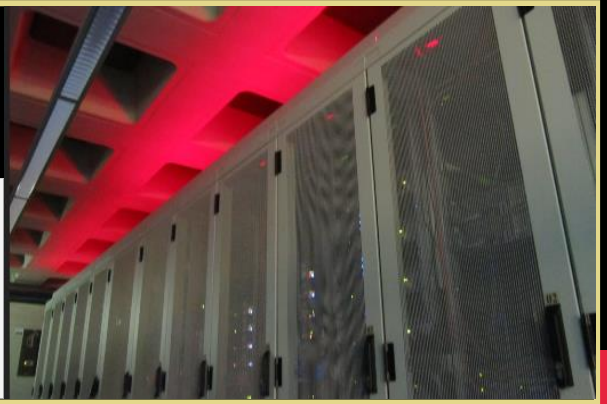
global T1 resources through Nikhef and SURF



SURF + Nikhef national resources

Nikhef-collaboration specific 'Stoomboot' compute & data services





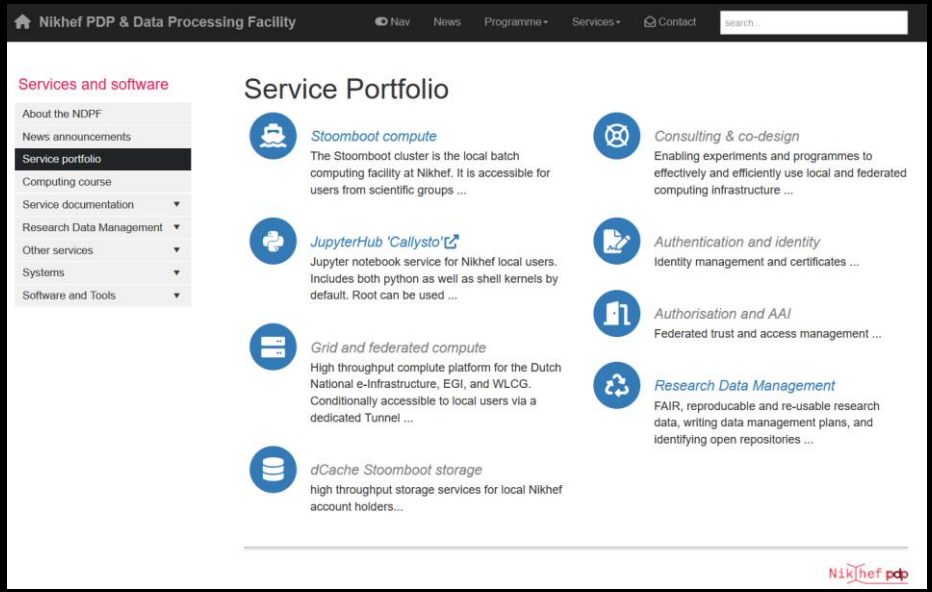
Computing resources and research in the Nikhef context

Images: CERN, SURF, Nikhef/D.Groep, Map sublayer: Google

Nikhef

Nikhef PDP & the Nikhef Data Processing Facility “NDPF”

Specialised resource services
- our ‘local’ analysis facility
- federated ‘Grid’ services for the Dutch National e-Infrastructure
- wLCG ‘Tier-1’ (Nikhef+SURF)



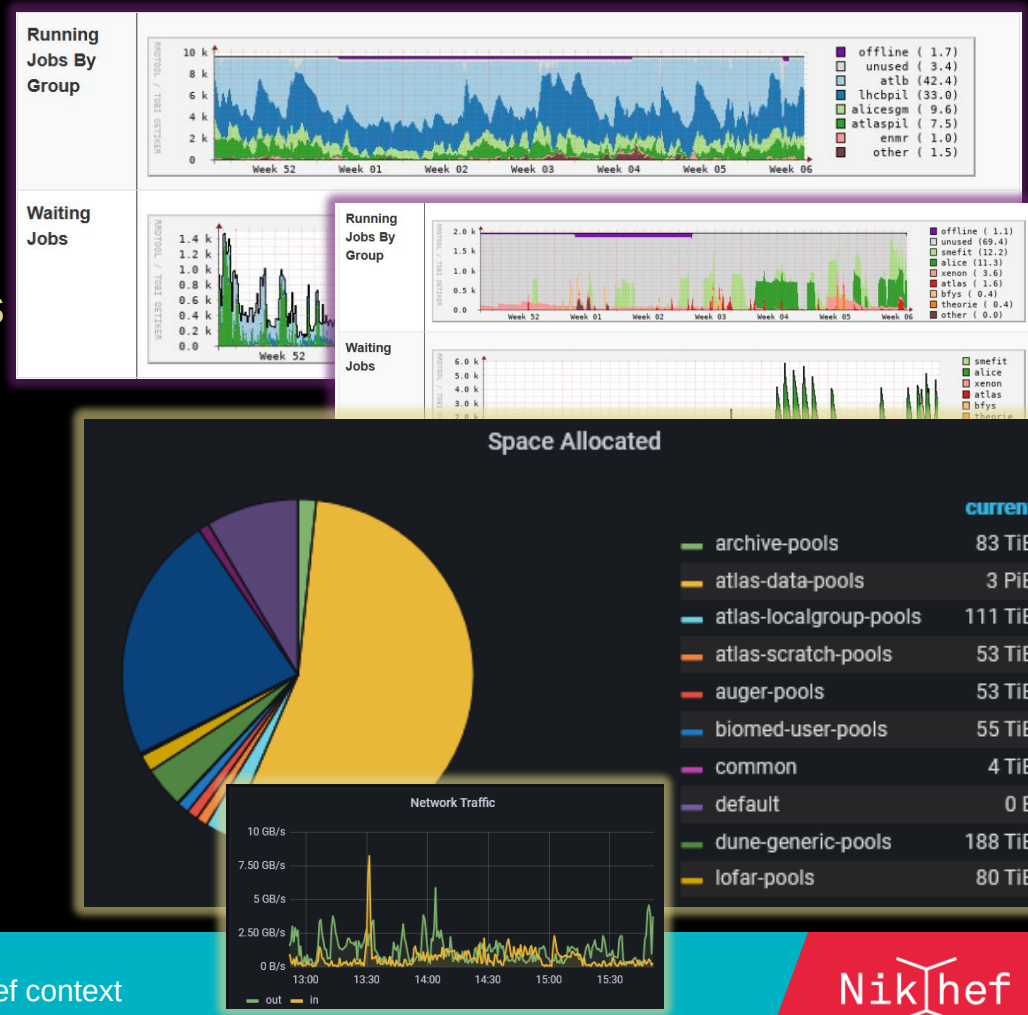
Specialised people are important!
- systems research engineering team (“DevOps+”, ~ 6FTE)
- linked to more generic ‘computing support’, but with advanced R&D role
- our software and standards engineering teams cross-fertilise this work
- link to research data management activities (@Nikhef: this is in PDP)

NDPF resources today

High Throughput Compute
- Grid/DNI & Stoomboot
- 9500+1500 cores and some GPUs

Scalable distributed storage
- ~7 PB federated + 3 PB local
- multiple storage qualities

Experimental facilities
- nationalespeeltuin.nl
- infrastructure innovation, with SURF SOIL and vendors



Images: Nikhef NDPF Prometheus and GRIDview data, February 2022, https://www.nikhef.nl/pdp/stats/

Stoomboot – our friendly local facility

A Nikhef collaboration ‘standard service’

- for anyone with a Nikhef account,
  i.e. registered users ‘with some plausible association’
- *resource priority* depends on fair sharing for compute, and
- shared for all Nikhef programmes (no experiment clusters)
- HPC and federation expertise **bundled** in
  joint DevOps team, the PDP research programme,
  and (access and research) software development team

Fully connected to Nikhef’s research network structure

- we postulate that an R&E institute has an R&E data network
- with maybe some office enclaves, not the other way round!

Image: Nikhef NDPF “Chocolade” compute nodes, AMS/H234b

Nikhef network topology as of February 2022

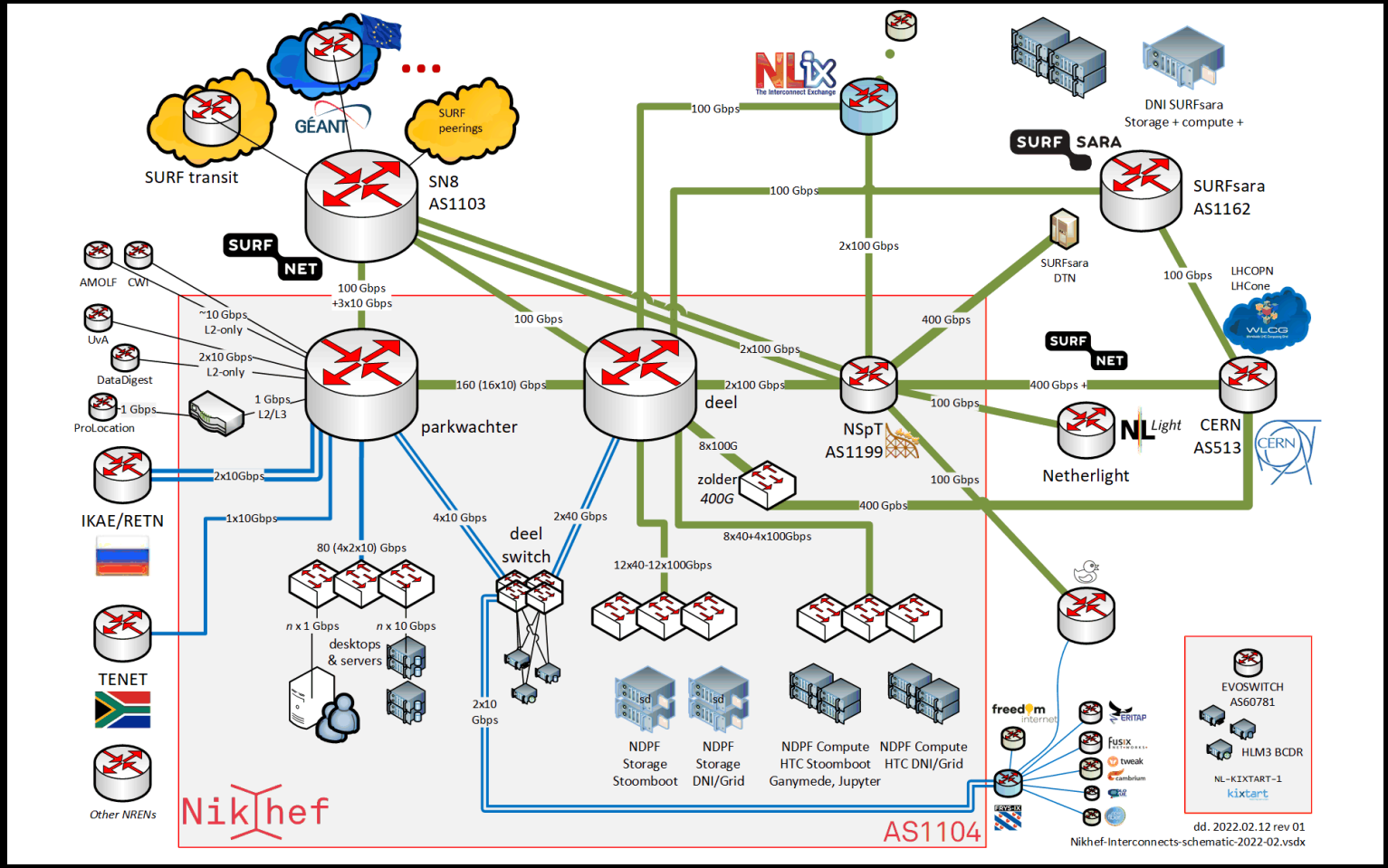
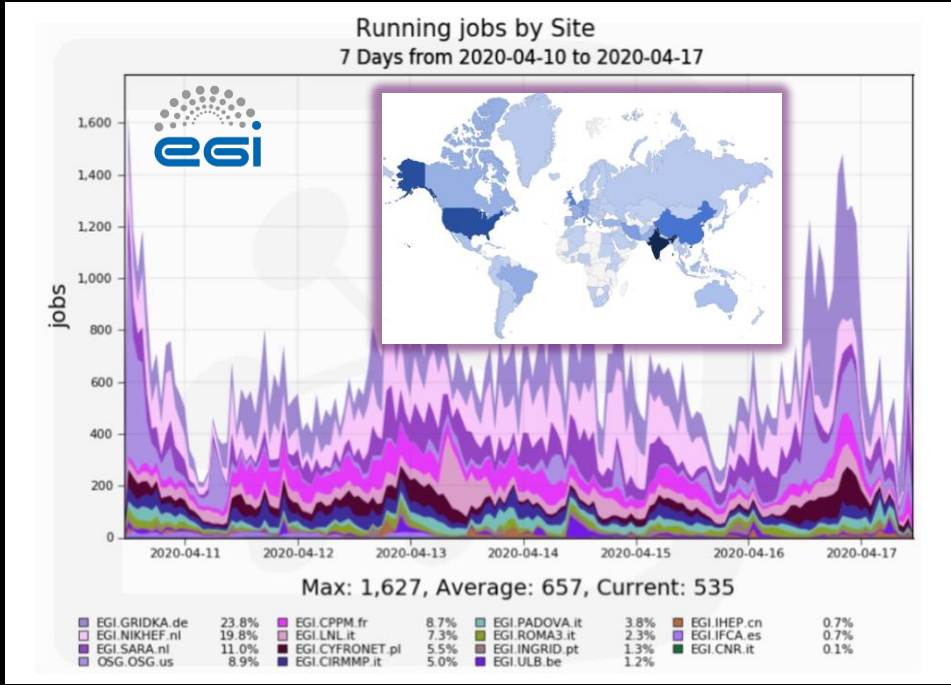


Image: the Nikhef Autonomous System and its interconnects, D.L. Groep/2022

Resource access modes

Modelled around ‘free at point of use’

- **local resources are always there** for those with ‘some reasonable link’ to Nikhef programme & collaboration
- access to the DNI, at Nikhef and at SURF, follow community centric access model - one Dutch qualified applicant is enough to bring in a global community
- application via SURF e-Infra portal (large requests via an NWO route)



positive effects of collaborative research more than warrant this model

image from: https://www.egi.eu/use-cases/research-infrastructures/wenmr-a-worldwide-e-infrastructure-for-nmr-60348-2/
Haddock User Map WeNMR portal - Alexandre Bonvin, Bijvoet Centrum voor Biomoleculair Onderzoek, Universiteit Utrecht

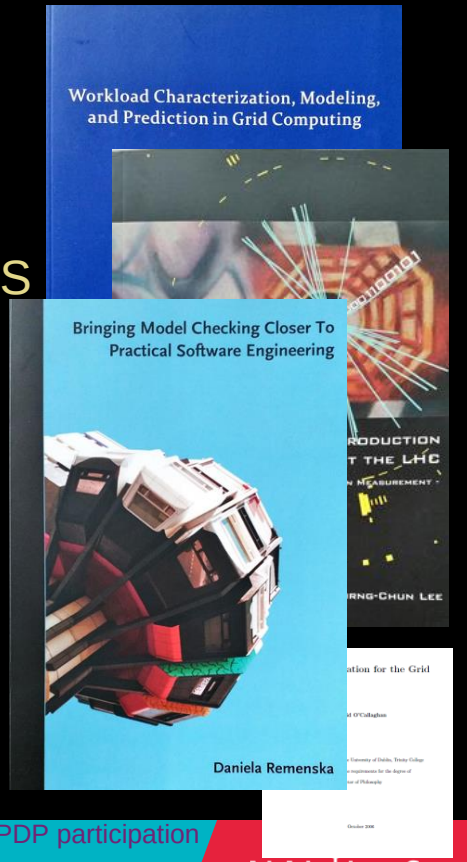
Beyond machines - collaborative computing research

the (Nikhef) e-infra, with its diverse user base and large global applications, as a computer science and modelling proving ground

Some past computer science research PhD projects

- Data intensive workflow scheduling research validated in ATLAS *(Hurng-Chun Lee, UvA – with ATLAS)*
- Estimated run time prediction in near-full occupancy scenarios on Nikhef WLCG and DNI clusters *(Hui Li, LIACS – with PDP)*
- Model checking for distributed scheduling systems, using LHCb DIRAC as the object *(Daniela Remenska, VU CS – with PDP and LHCb)*

Many international examples as well: Assurance models for federated infrastructures (Daniela Pöhn and Jule Ziegler at LRZ); Trust modelling (David O'Callaghan, TCD)



Images: front covers PhD thesis from Nikhef and/or with key Nikhef PDP participation

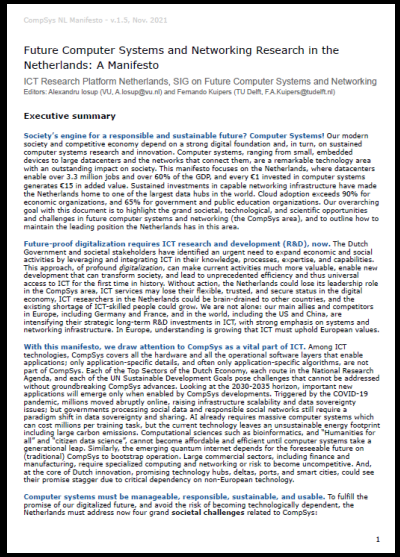
More is interesting, and much is needed

Opportunities abound:

- LHCb has an open data policy for joint associate collaborators (for ML &c)
- projects (*e.g.* NWA NextGrasPP) ML & GPU co-design, QC (with SURF and linked to e.g. CERN QTI)
- we co-signed the CompSys NL manifesto for a good reason!
- …

… where Nikhef facilities and our federated, global, e-Infra system, and our systems-innovation lab ‘Speeltuin’, provides opportunity, a study object, and capacity

… *and for educational purposes, we have often donated decommissioned hardware which of course works just fine, but the energy-performance-rack space ratios are no longer efficient for the DNI and WLCG use cases*


CompSys NL Manifesto - v.1.5, Nov. 2021

Future Computer Systems and Networking Research in the Netherlands: A Manifesto

ICT Research Platform Netherlands, SIG on Future Computer Systems and Networking
Editors: Alexandru Iosup (VU, A.Iosup@vu.nl) and Fernando Kuipers (TU Delft, F.A.Kuipers@tudelft.nl)

Executive summary

Society's engine for a responsible and sustainable future? Computer Systems! Our modern society and competitive economy depend on a strong digital foundation and, in turn, on sustained computer systems research and innovation. Computer systems, ranging from small, embedded devices to large datacenters and the networks that connect them, are a remarkable technology area with an outstanding impact on society. This manifesto focuses on the Netherlands, where datacenters enable over 3.3 million jobs and over 60% of the GDP, and every €1 invested in computer systems generates €15 in added value. Sustained investments in capable networking infrastructure have made the Netherlands home to one of the largest data hubs in the world. Cloud adoption exceeds 90% for economic organizations, and 65% for government and public education organizations. Our overarching goal with this document is to highlight the grand societal, technological, and scientific opportunities and challenges in future computer systems and networking (the CompSys area), and to outline how to maintain the leading position the Netherlands has in this area.

Future-proof digitalization requires ICT research and development (R&D), now. The Dutch Government and societal stakeholders have identified an urgent need to expand economic and social activities by leveraging and integrating ICT in their knowledge, processes, expertise, and capabilities. This approach, of profound *digitalization*, can make current activities much more valuable, enable new development that can transform society, and lead to unprecedented efficiency and thus universal access to ICT for the first time in history. Without action, the Netherlands could lose its leadership role in the CompSys area. ICT services may lose their flexible, trusted, and secure status in the digital economy, ICT researchers in the Netherlands could be brain-drained to other countries, and the existing shortage of ICT-skilled people could grow. We are not alone: our main allies and competitors in Europe, including Germany and France, and in the world, including the US and China, are intensifying their strategic long-term R&D investments in ICT, with strong emphasis on systems and networking infrastructure. In Europe, understanding is growing that ICT must uphold European values.

With this manifesto, we draw attention to CompSys as a vital part of ICT. Among ICT technologies, CompSys covers all the hardware and all the operational software layers that enable applications; only application-specific details, and often only application-specific algorithms, are not part of CompSys. Each of the Top Sectors of the Dutch Economy, each route in the National Research Agenda, and each of the UN Sustainable Development Goals pose challenges that cannot be addressed without groundbreaking CompSys advances. Looking at the 2030-2035 horizon, important new applications will emerge only when enabled by CompSys developments. Triggered by the COVID-19 pandemic, millions moved abruptly online, raising infrastructure scalability and data sovereignty issues; but governments processing social data and responsible social networks still require a paradigm shift in data sovereignty and sharing. AI already requires massive computer systems which can cost millions per training task, but the current technology leaves an unsustainable energy footprint including large carbon emissions. Computational sciences such as bioinformatics, and "Humanities for all" and "citizen data science", cannot become affordable and efficient until computer systems take a generational leap. Similarly, the emerging quantum internet depends for the foreseeable future on (traditional) CompSys to bootstrap operation. Large commercial sectors, including finance and manufacturing, require specialized computing and networking or risk to become uncompetitive. And, at the core of Dutch innovation, promising technology hubs, deltas, ports, and smart cities, could see their promise stagger due to critical dependency on non-European technology.

Computer systems must be manageable, responsible, sustainable, and usable. To fulfill the promise of our digitalized future, and avoid the risk of becoming technologically dependent, the Netherlands must address now four grand **societal challenges** related to CompSys:

1

Manifasto: https://bit.ly/ManifestoCompSysNL