Dreaming of a CS lab

The DACS FSE Computing Sciences Lab as a Fieldlab Demonstarator

Concepts, capabilities, connections

David Groep, January 2025

Why the CS lab?

Initially triggered by 'basic' short-term needs

- new Bachelor Computer Science,
  with strong engineering component (esp. in Y2 & Y3)
- Large-scale IT and robotics, also in DSAI Year-3

But with wide use for {master, PhD} thesis work and cross-discipline and cross-organisation *research* as well:

- HPC simulation and modelling , ML models, LLM applications, and collaborating with research output 'as a service'
- science and engineering beyond DACS:
  ET pathfinder & LHCb for GWFP, the MSP programme, …
- data science in other faculties: need for 'student sandboxes'
- large multi-modal 4D imaging (e.g. for NWO ENW XL FASTER with Ron Heeren), and EU Horizon and Euregio collaborations

Period 1	Period 2	Period	Period 4	Period 5	Period
Computer Networks;	Software Engineering & Architectures;	P R O J E C T	Embedded Programming;	IT Management & Privacy	P R O J E C T
Introduction to Artificial Intelligence;	Principles of Programming Languages		Computer Security;	Numerical Methods	
Intelligent User Interfaces	1 out of the two elective modules:		Parallel Programming	1 out of the two elective modules	
	- Intelligent Interaction: course Image & Video Processing + Project 2-1: Human-Computer Interaction - Artificial ...			- High-Performance Computing: course High-Performance Computing + Project 2-2: High-Performance Computing Cyber-security and IoT: choose the course between Information security or ubiquitous computing IoT + Project 2-2: Cybersecurity IoT	

Year 3

Period 1*	Period 2*	Period 3	Period 4	Period 5	Period 6
Digital Society;	Computer Security;	P R O J E C T	Operations Research Case Studies **		BACHELOR'S THESIS
Game Theory;	Software and Systems Verification;		Intelligent Systems;		
Semantic Web;	Logic for Artificial Intelligence;		Data Analysis		
Recommender Systems;	Parallel Programming;				
Robotics and Embedded Systems;	Large Scale IT and Cloud Computing;				
Introduction to Quantum Computing	Introduction to Bio-Informatics;				
PROJECT			BACHELOR'S THESIS		

Maastricht University | DACS

DACS FSE Computing Sciences Lab as a Fieldlab Demonstrator

Suddenly ... CSlab

And then there was Sectorplan 1
... whose funds needed to be **spent in 2024**

'Wereldwijd verbonden vanuit een stevige basis in onze Euregio'

Research and education are inherently collaborative, and 'services exist' and are being used .. whether we know about them or not:

- European Open Science Cloud, 'ESFRI' research infrastructures, Erasmus Without Papers, EU Identity Wallets, SURF-NWO Rekentijd & Snellius, EOSC EU Node 'free VMs for faculty'

Images: UM Strategic Programme 2022-2026, SURF ETP
slide captures: www.nikhef.nl/~davidg/p/ for UM-CI office & LsIT course

Translating 'wereldwijd verbonden' education & research as a technical systems and network design

1. 'we can *use* existing services from outside': AAI interoperability, network connectivity and bandwidth, devolution of responsibility and subsidiarity of concern

2. 'we can *contribute* in collaborations in education and research': we can offer services, support pan-European student mobility in our programme, and provide innovative lab environments that attract students and make our programmes unique

3. 'we teach and train our students to understand, design, build, study, work with, and experiment with interconnected services and systems that are globally connected'

Of course we should do that in a safe, scalable, and transparent way,
and at the same time build and use our own science and engineering R&E concepts in an agile way

With a view towards the future … a decadal vision of common reference platform model, usable for DACS & FSE, ET, … for FASTER 4D-imaging, … for SBE, FASoS, and UM-wide?

Technical inspiration

- DACS CS-Lab concept *for today* focusses on systems and networks
    - those act a foundation for more specific lab settings: robotics, XR, security
    - focus on education labs, but equally applicable to research networks

- design around micro-segmented networks and inter-subnet controls
    - so you could build multiple labs on a partly-shared infrastructure
    - leveraging the ESnet (US DoE) 'ScienceDMZ' concept, and
    - on the UvA Security and Network Engineering master lab
      ('every student a machine on the public internet
      and they are *entirely responsible* for this box')
    - with inspiration from SURF Experimental Technologies
      Platform and Nikhef open innovation lab

- for CS and engineering education (extensible to research), and
  not targeting enterprise services, **nor** for now sensitive data services
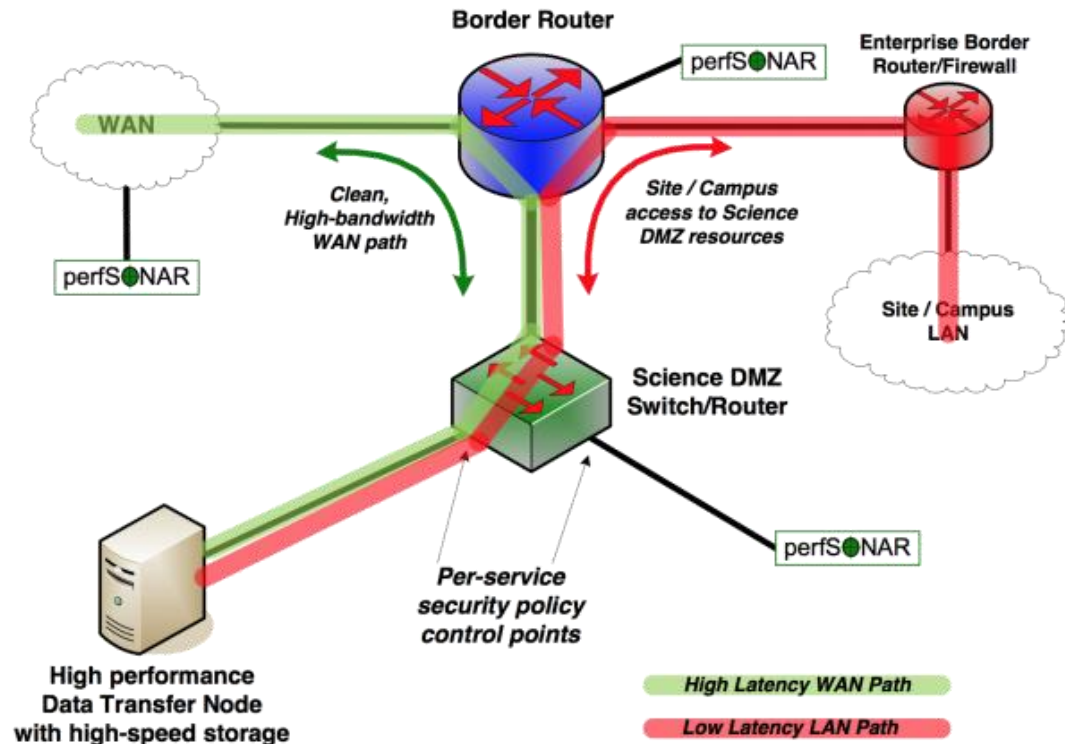
'ScienceDMZ'

Predicable performance
and data access for research

*'where research services,
data, and researchers meet'*

- lossless transfers and latency hiding
- elements horizontally scalable
- security zoning/segmentation
  protects services and data
- outside enterprise perimeter

for a different type of traffic: science data transfers rather than eyeball networks

Image and 'ScienceDMZ' concept promulgated by ESnet (see fasterdata.es.net)
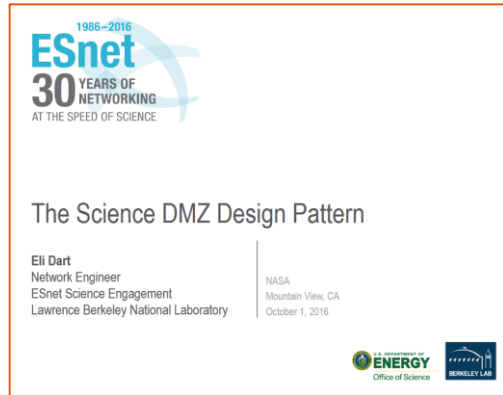

Maastricht University | DACS

The Value Of Routine Performance

- It's important to get to where high performance is normal

- No magic, no arcana, things just normally work – for petabytes of data

- DOE HPC facilities now easily shuffle around hundreds of terabytes
  - Some people have smaller data sets too
  - But the point is that it's normal and routine

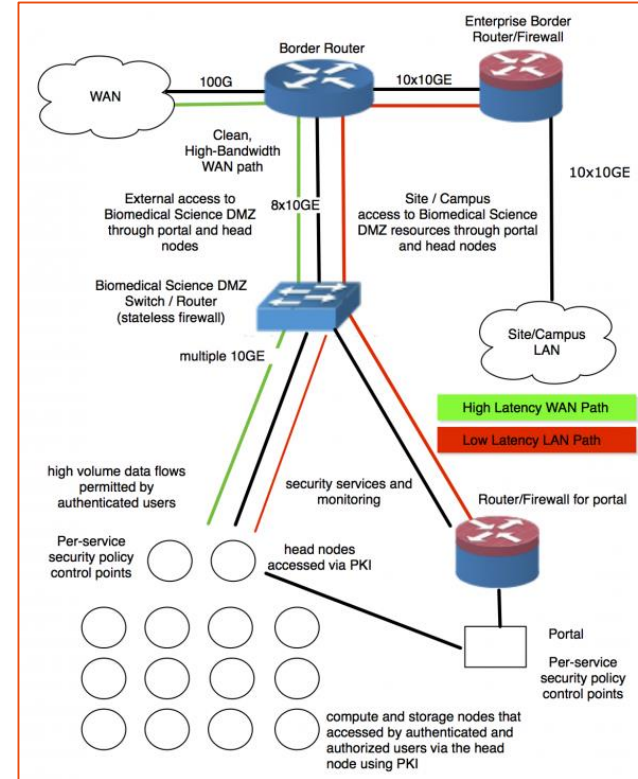- What follows is one specific example, chosen because of some specific features

ESnet

Maastricht University | DACS

DACS FSE Computing Sciences Lab as a Fieldlab Demonstrator

Labs as a design pattern …
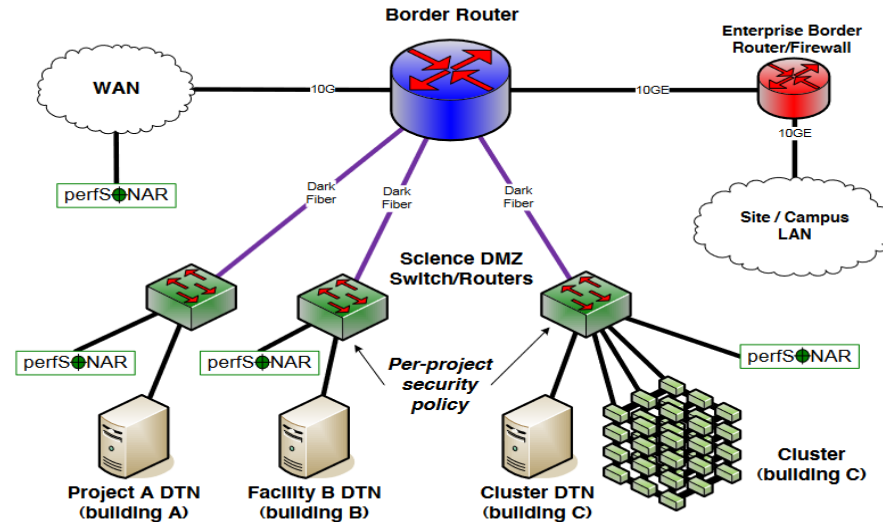

High Latency WAN Path

Low Latency LAN Path

Left: components in a medical sciences DMZ
from https://crd.lbl.gov/news-and-publications/news/2017/esnets-science-dmz-design-could-help-transfer-protect-medical-research-data/

And once you have multiple FieldLabs

Multiple Science DMZs – Dark Fiber

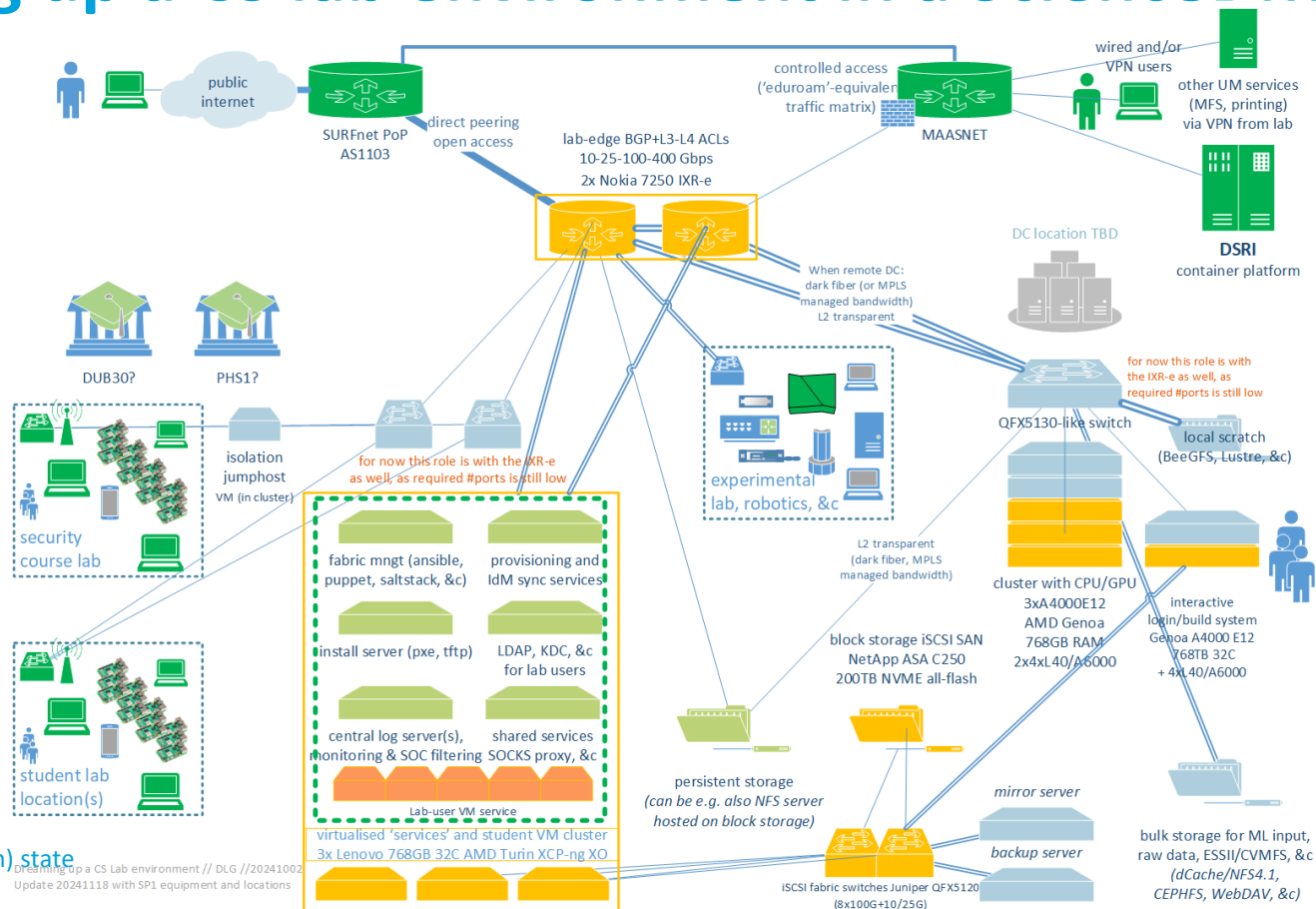15 – ESnet Science Engagement (engage@es.net) - 7/13/17

Dreaming up a CS lab environment in a ScienceDMZ

SOLL and IST (yellow, green) state
of the CSlab foundation

Dreaming up a CS Lab environment // DLG //20241002
Update 20241118 with SP1 equipment and locations

CS Lab – the 2024 foundation (<20 December 2024)

Virtualised base infrastructure: fabric management, logging, IAM bridging (to SURF SRAM or locally with a SAML SP), proxy and gateway services for sub-labs, ~300+ student sandbox VMs, XCP-ng & XenOrchestra managed with SAML or OIDC auth

- 3x SR635v3 (each: 786GB RAM, 2x25G data network, 2x25G iSCSI, 64 logical cores 9355P)

Storage back-end service: block storage for VMs (supporting NFSv4 via VM with linux)

- NetApp C250 with 200 TByte net NVME all-flash capacity, 4x100G iSCSI

HPC starter cluster: AI/ML courses, reasoning, HPC modelling, interactive & batch

- 4x ASUS ESC4000A (each: 3.25GHz 64 logical core EPYC 9355P, 768 GB RAM, 2x16TB NVME, 2x25G data network), 3 of these systems have 4 Nvidia L40 GPUs each (so 12 GPUs in total)
- one node as the 'interactive' validation and compilation node (scheduling will be in a VM)

Network: access gateway for standard edge connectivity (BGP, MPLS, 802.1q, EVPN, L3/L4 ingress and egress ACLs and policers) and iSCSI SAN fabric switches (25/100G)

- pair of Nokia 7250 IXR-e, each: 2xQSFP28 100G, 8xSFP28 25G, 24xSFP+ 10G (or 1G)
- pair of Juniper QFX 5120's, each: 8xQSFP28 100G, 28xSFP28 25/10/1G

A lab for fieldlabs

It is also a technology lab & computing sciences engineering innovation lab

- networking from commonly used protocols like evpn-vpls/mpls
to high-speed udp science data transport or alternative tcp stacks,
or segment routing and connecting to LHCone at 400Gbps to 1.6Tbps?
- storage, compute, acceleration, cluster computing, virtualisation, orchestration, automation, AuthN & AuthZ infrastructures and IAM, …
- power usage, application efficiency, divergence of 2019/424 to actual usage?

'Lab' aspect need not to stop at layer-7

- ITSRM, SMS & FitSM/20k, … to federated services and Tier-2/1/0 scaleout
- red-teaming in resilience exercises with SURF/NMDDoS – or for ourselves?

A collaborative demonstrator

The Fieldlab & ScienceDMZ concepts are readily generalizable
- both a technology lab (networking, computing, storage, IoT, security)
- as well as a infrastructure and systems lab for SMS and ITSRM concepts

So let's try and use this for an integrated approach already now for the CSLab
- linked to current and evolving ICT at UM as a whole
- ITSRM for the CSLab in line with UM IT security risk management

look at this also as a playground for new ITSM ideas and experiments
– this is a 'harmless' lab with forgiving users and local flexibility ☺

The most important ingredients …

- **Creativity**
- **Flexibility**
- **Desire to experiment**

CSLab equipment was selected for the usual HA design, but

- can, may, and should be used also for experiments
- a common playground for students, researchers, and ICT engineers
- diversity in equipment, components, client devices, IoT stuff, and tools
  is intentional for experimentation

Other labs may have a different scope and purpose, but all share the top-3 elements!

Getting more hands-on

... as we have to start making choices soon
for mounting, numbering, labelling, port speeds, filters, and much more

Collaboration … because it's absolutely needed

In our interactions with 'the world', UM must appear as a coherent whole

- Network circuit provisioning to SURF
- Internet number resources (AS, IPv6 and IPv4 space) from SN-LIR (1xAS,/47 v6, /18 v4)
- L1 and L2 network transport across campus building locations
- RFC1918 number resources coordination
  (so users can have access to both UM and lab even when both VPN and lab are up)
- Housing and secured access, racking and stacking, power feeds and PDUs
- jumphost to the lab OoBM network (may be virtual, UM internal VPN access only) during commissioning (to prevent having to sit in DEB1 all the time ☺)
- Collaboration services like SURF SRAM also needs UM central coordination
- Incident response coordination & Sirtfi

even though DACS can initiate and prepare SURF staff for the requests,
SURF really wants these to come from our institutional contact persons

Collaboration … because it's just better together

- IAM for linking lab services to UM identity
  an experiment: *CSLab end-users shall not have any lab-specific re-usable credentials,*
  but use only SAML, OIDC, or Wallet VCs to authenticate, or ssh pubkeys for login/access

- release of attributes and provisioning interfaces for some lab services
  like: which students are in which course,
  and hence need to be granted access to which parts of the lab?

- prevent duplication of SMS supporting assets
  that the CSLab would need to set up but can be usefully shared:
  protected GitLab for the CMDB, Jira/Confluence-like service, …

Collaboration ... because we're all human

- coordinated cable numbering in shared DC facilities like DEB1

- alignment on e.g. 802.1q identifiers because it prevents confusion in human communication (CSlab needs ~50 out of 4093 for its segmentation)

- joint access to documentation and SMS information

- work out a way for things sent to the UM central helpdesk to bounce misdirected questions back to the lab

- periodic technical meetings & chat, since working together is nice,
  ... and because we all forget to ask for the right stuff sometimes!

More discussion time …
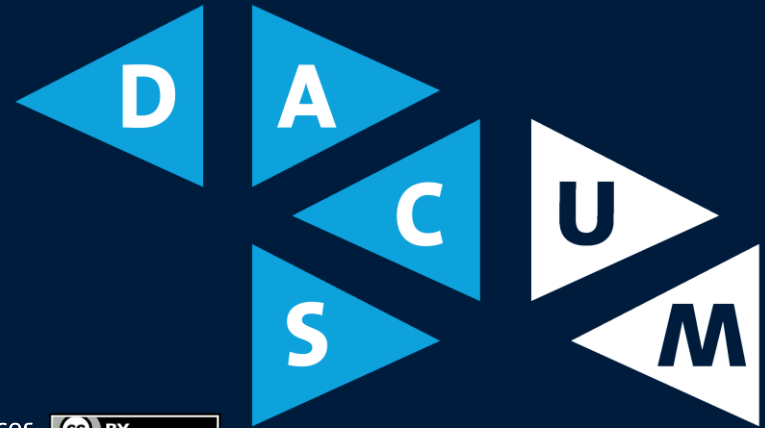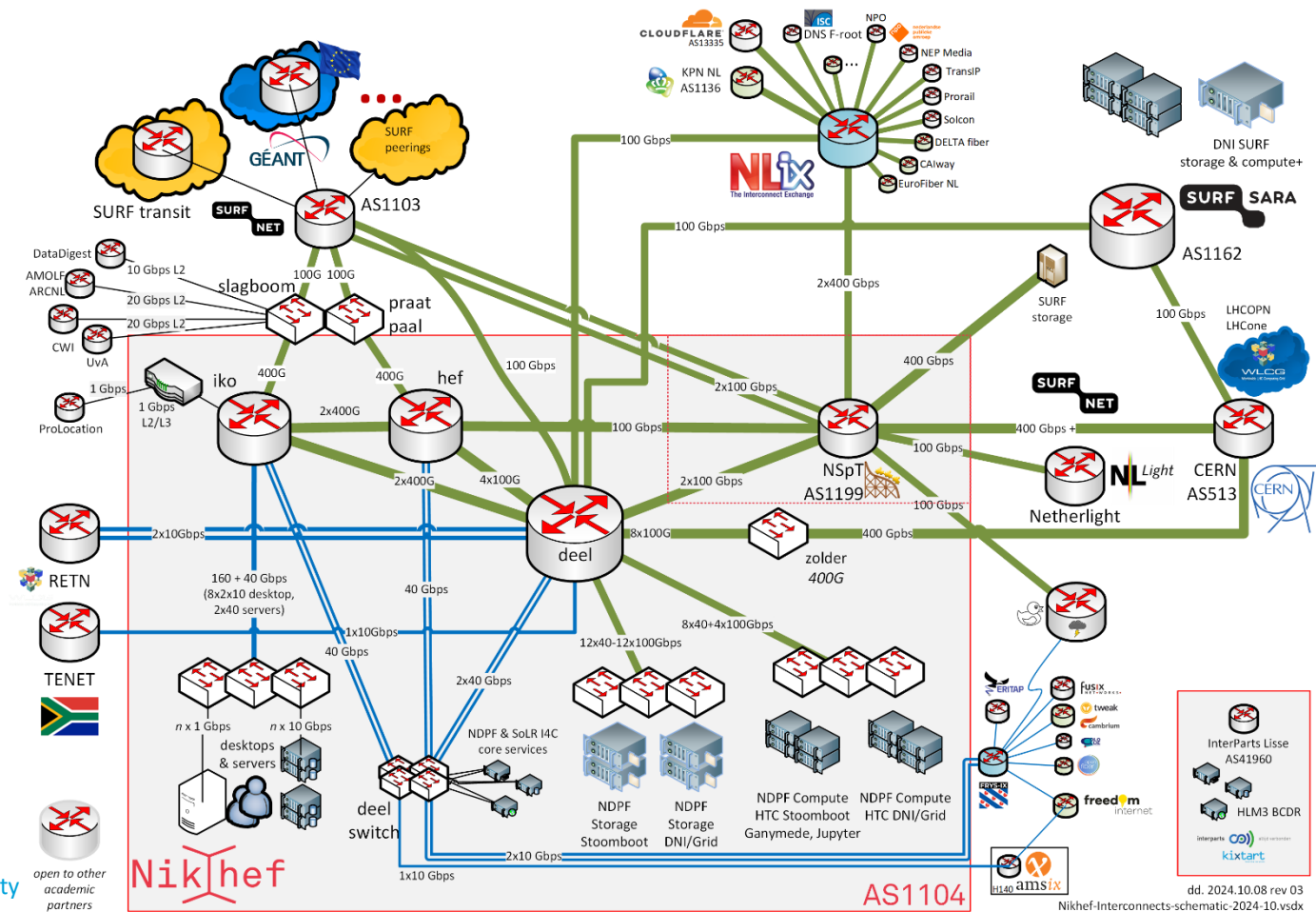
David Groep
david.groep@maastrichtuniversity.nl
https://www.nikhef.nl/~davidg/presentations/
https://orcid.org/0000-0003-1026-6606

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

Nikhef

Maastricht University | Department of Advanced Computing Sciences

AS1104

Maastricht University | DACS

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

Labs and more labs!

Nikhef SP105: H234b (L) and H234f (R)