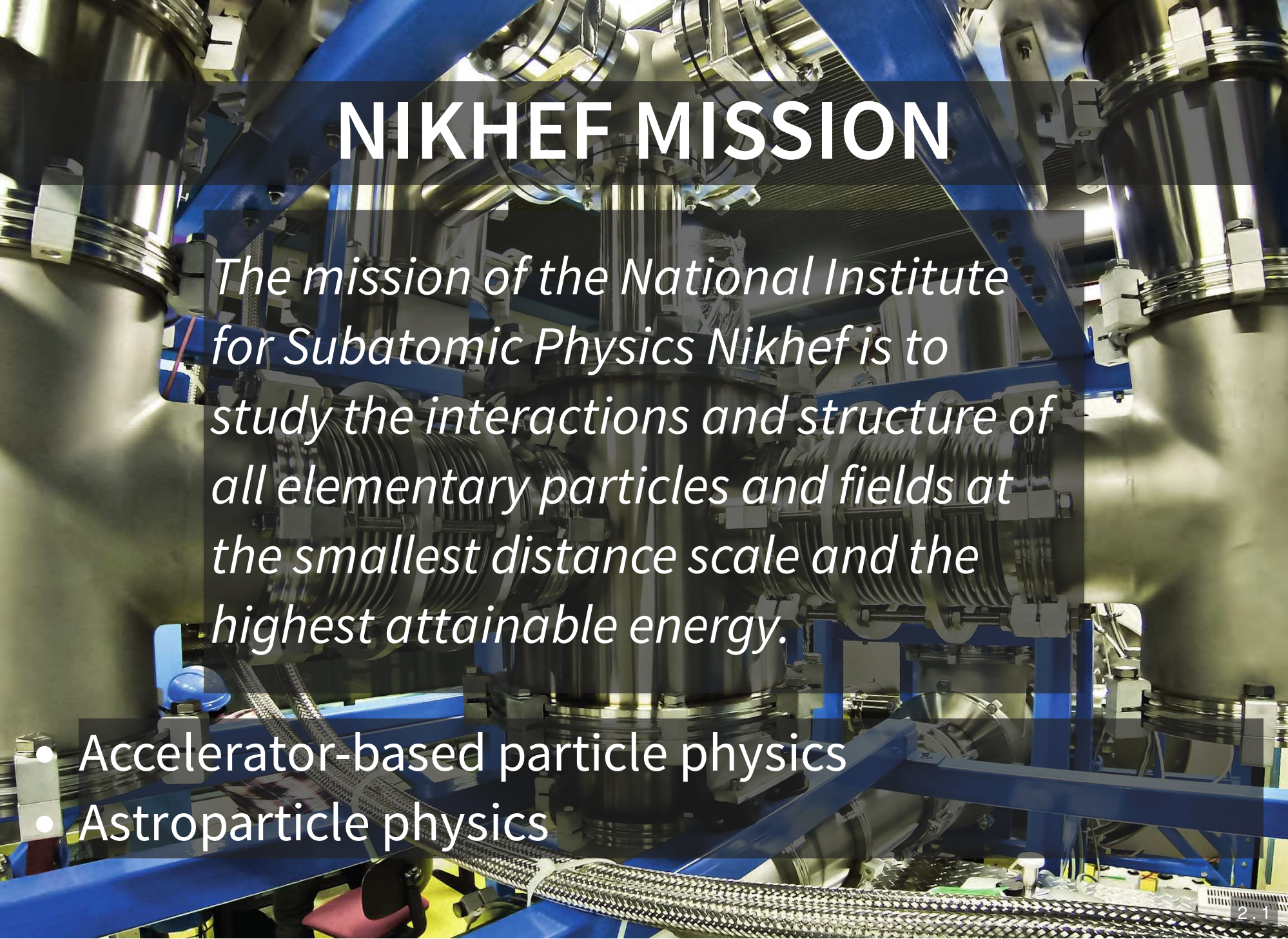
NIKHEF SITE REPORT

DENNIS VAN DOK

2019-03-26



NIKHEF MISSION

The mission of the National Institute for Subatomic Physics Nikhef is to study the interactions and structure of all elementary particles and fields at the smallest distance scale and the highest attainable energy.

- Accelerator-based particle physics
- Astroparticle physics

ORGANISATION

Radboud University

Nikhef

rijksuniversiteit groningen

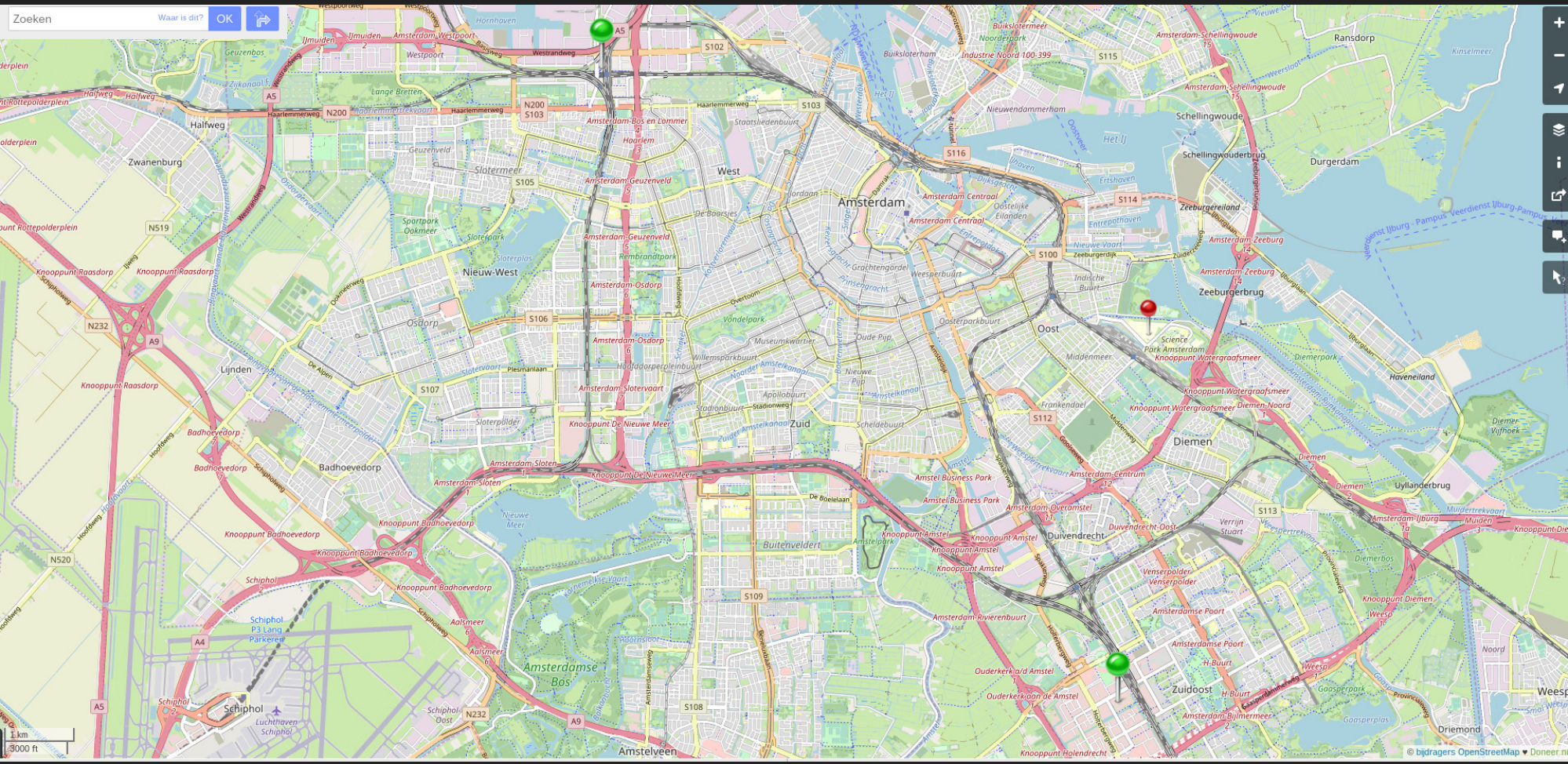
University of Amsterdam

Maastricht University

NWO | Instituten

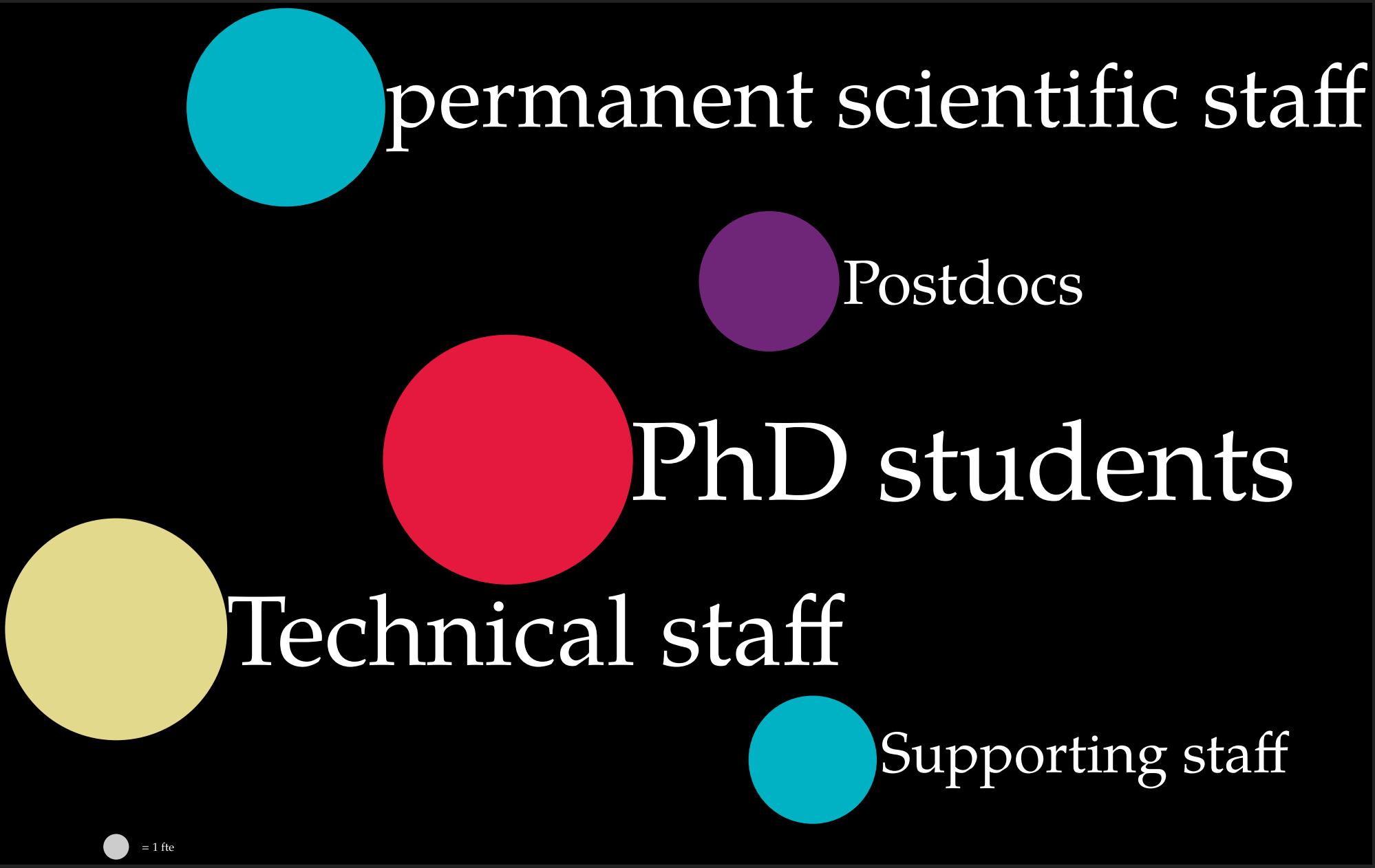
Utrecht University

VU VRIJE UNIVERSITEIT AMSTERDAM


Zoeken
Waar is dit?
OK
Amsterdam
West
Nieuw-West
Oost
Zuid
Zuidoost
Noord
Diemen
Duivendrecht
Ransdorp
Schellingwoude
Durgerdam
Zeeburgereiland
Zeeburgerbrug
Halfweg
Zwanenburg
Slotermeer
Osdorp
Lijnden
Badhoevedorp
Schiphol
Schiphol P3 Lang Parkeren
Amsterdamse Bos
Buitenveldert
Amstelveen
Driemond
Uylanderbrug
Haveneiland
Science Park Amsterdam
Amsterdam Centraal
Knooppunt De Nieuwe Meer
Knooppunt Badhoevedorp
Knooppunt Amstel
Knooppunt Watergraafsmeer
Knooppunt Diemen
Knooppunt Holendrecht
Amsterdam-Sloten
Amsterdam-Osdorp
Amsterdam-Geuzenveld
Amsterdam-Bos en Lommer
Amsterdam-Zeeburg
Amsterdam-Centrum
Amsterdam-Rijnkanaal
Amsterdam-Amstelkanaal
Amsterdam-Slotervaart
Amsterdam-IJburg
Amstel Business Park
Amsterdamse Poort
Venserpolder
Duivendrecht-Oost
Ouderkerk a/d Amstel
Ouderkerk aan de Amstel
Amsterdam-Rijnenburg
Amsterdam-Overamstel
Amsterdam-Buikermeer
Het IJ
Oostvaarderplassen
Nieuwendammerham
Buiksloterham
Industrie Noord 100-399
Noorderpark
Kinselmeer
Diemerbos
Diemerpark
Diemer Vijfhoek
Diemen-Noord
Frankendael
Middenmeer
Oosterparkbuurt
Weesperbuurt
Grachtengordel
Rembrandtpark
Vondelpark
Museumkwartier
Oude Pijp
Nieuwe Pijp
Apollobuurt
Stadionbuurt
Scheldebuurt
Willemsparkbuurt
Hoofddorpleinbuurt
Staatsliedenbuurt
Haarlem
Westpoort
IJmuiden
Geuzenbos
Lange Bretten
Sportpark Ookmeer
Sloterpark
Sloterplas
Sloterpolder
Nieuwe Meer
Hornhaven
Entrepothaven
Oostelijke Eilanden
Indische Buurt
Veerdienst IJburg - Pampus
Veerdienst IJburg - Muiden
A5
A4
A9
A10
N200
N232
N519
N520
S102
S103
S105
S106
S107
S108
S109
S112
S113
S114
S115
S116
S100
1 km
3000 ft
© bijdragers OpenStreetMap ♥ Doneer nu

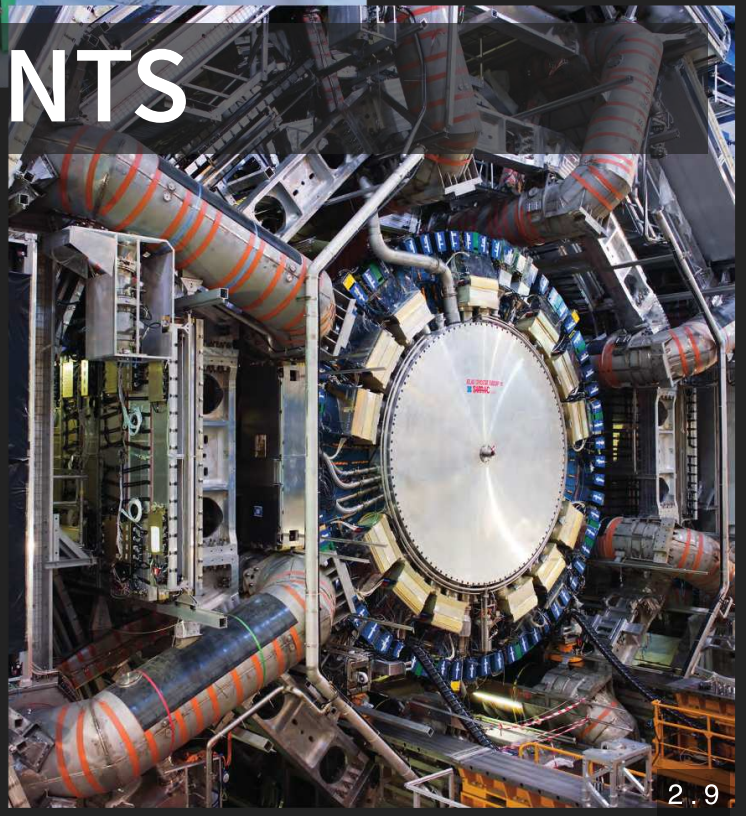
Nikhef

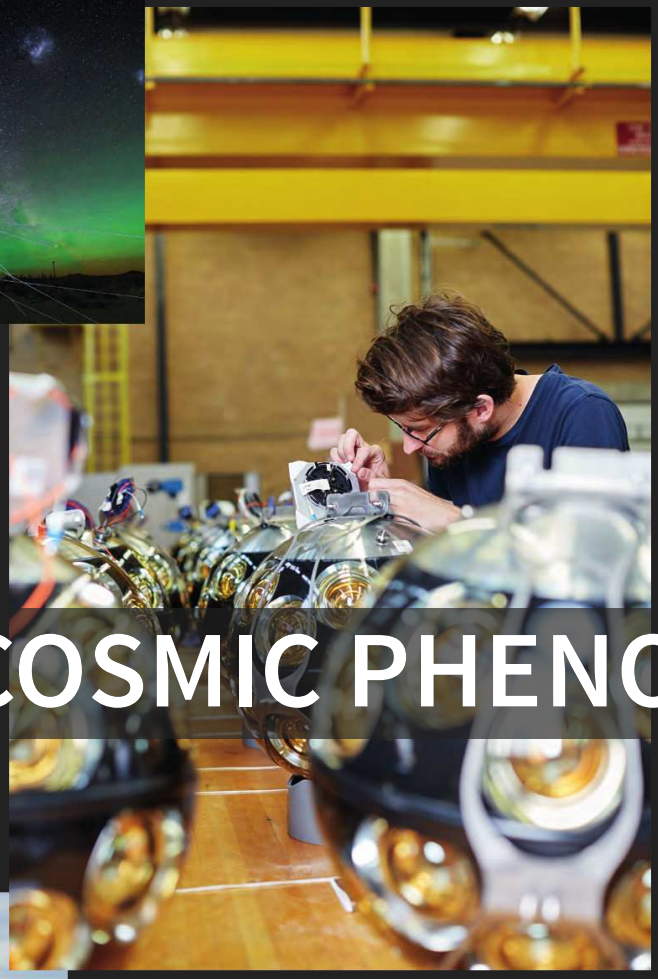
PERSONNEL

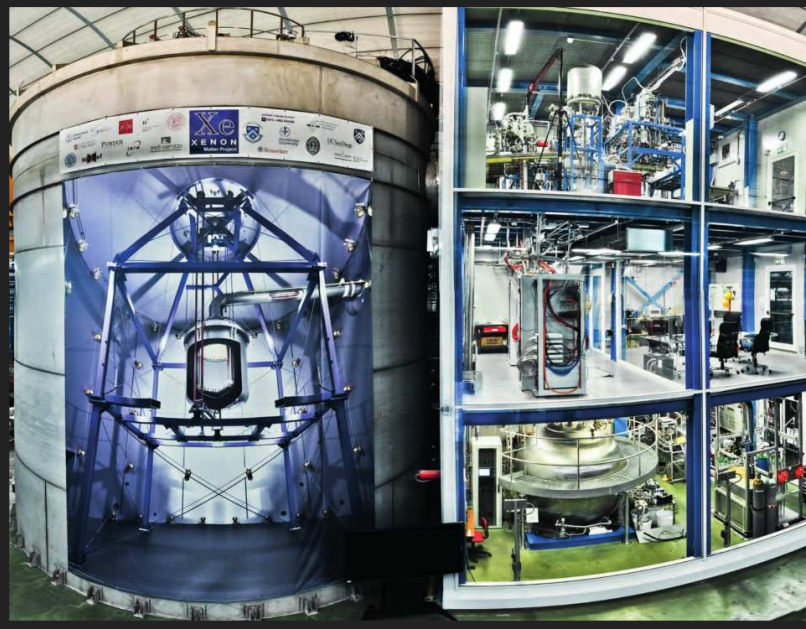

permanent scientific staff
Postdocs
PhD students
Technical staff
Supporting staff
= 1 fte

PROJECTS

LHC EXPERIMENTS







COSMIC PHENOMENA

OTHER PROGRAMS

- Theory programs
- Enabling programs
  - Detector Research and Development
  - Physics Data Processing

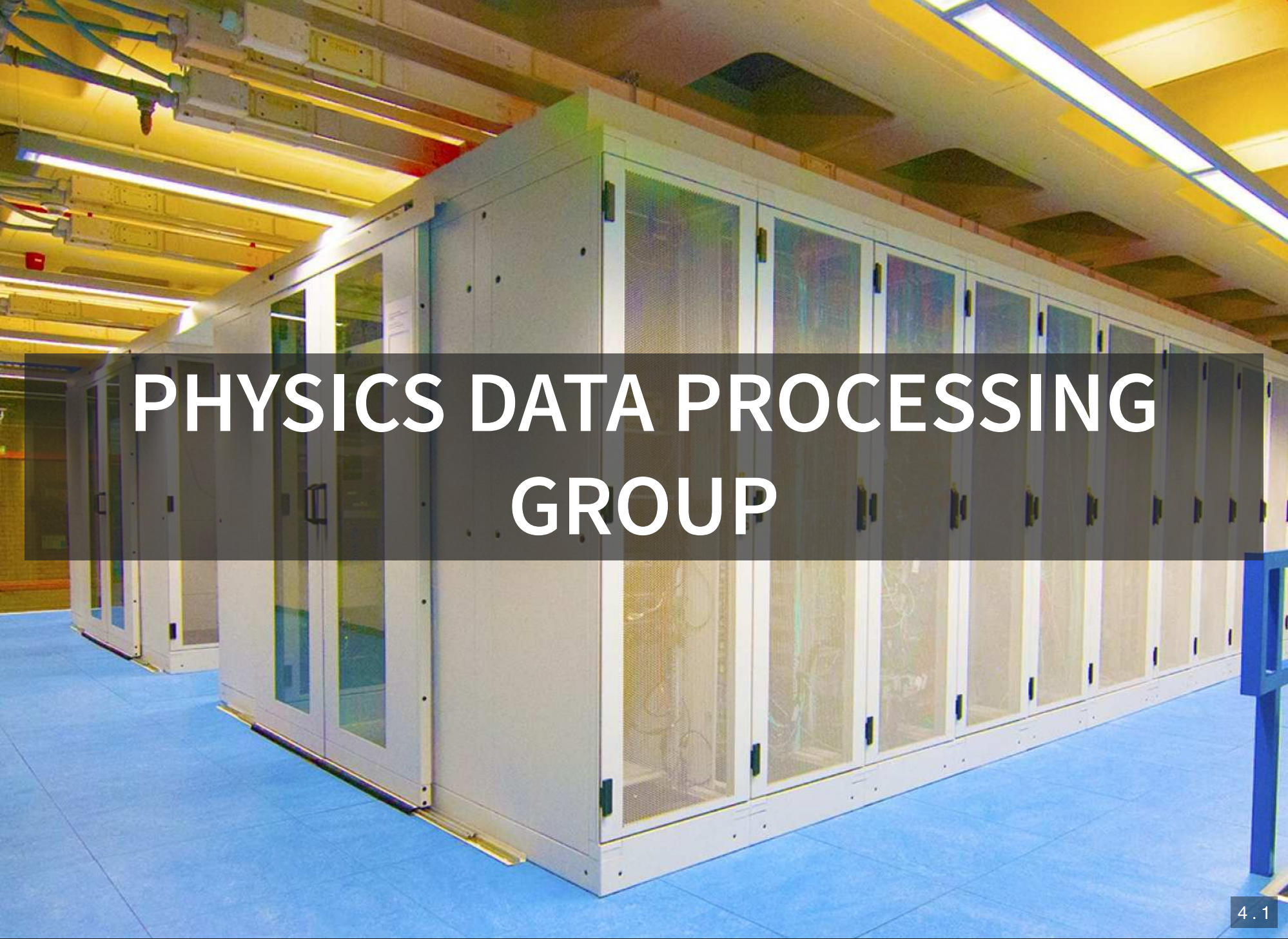
COMPUTER TECHNOLOGY

- Office automation
- Project support (detector software engineering)
- Scientific computing (a.k.a. Physics Data Processing or PDP)

PEOPLE

New sysadmin in the PDP group: **Mary Hester** joined in Februari.

The Computer Technology group hired a new webmaster: **Roel Roomeijer**. joined the system administration team in March.


PHYSICS DATA PROCESSING
GROUP

WLCG TIER-1

Nikhef-PDP is half of the Netherlands Tier-1.

The other half is SURFsara.

Together we also participate in the DNI, the Dutch National Infrastructure for data/compute intensive sciences.

- LOFAR (radio astronomy)
- Project MinE (genetic causes of motor neuron disease ALS)
- LIGO/VIRGO

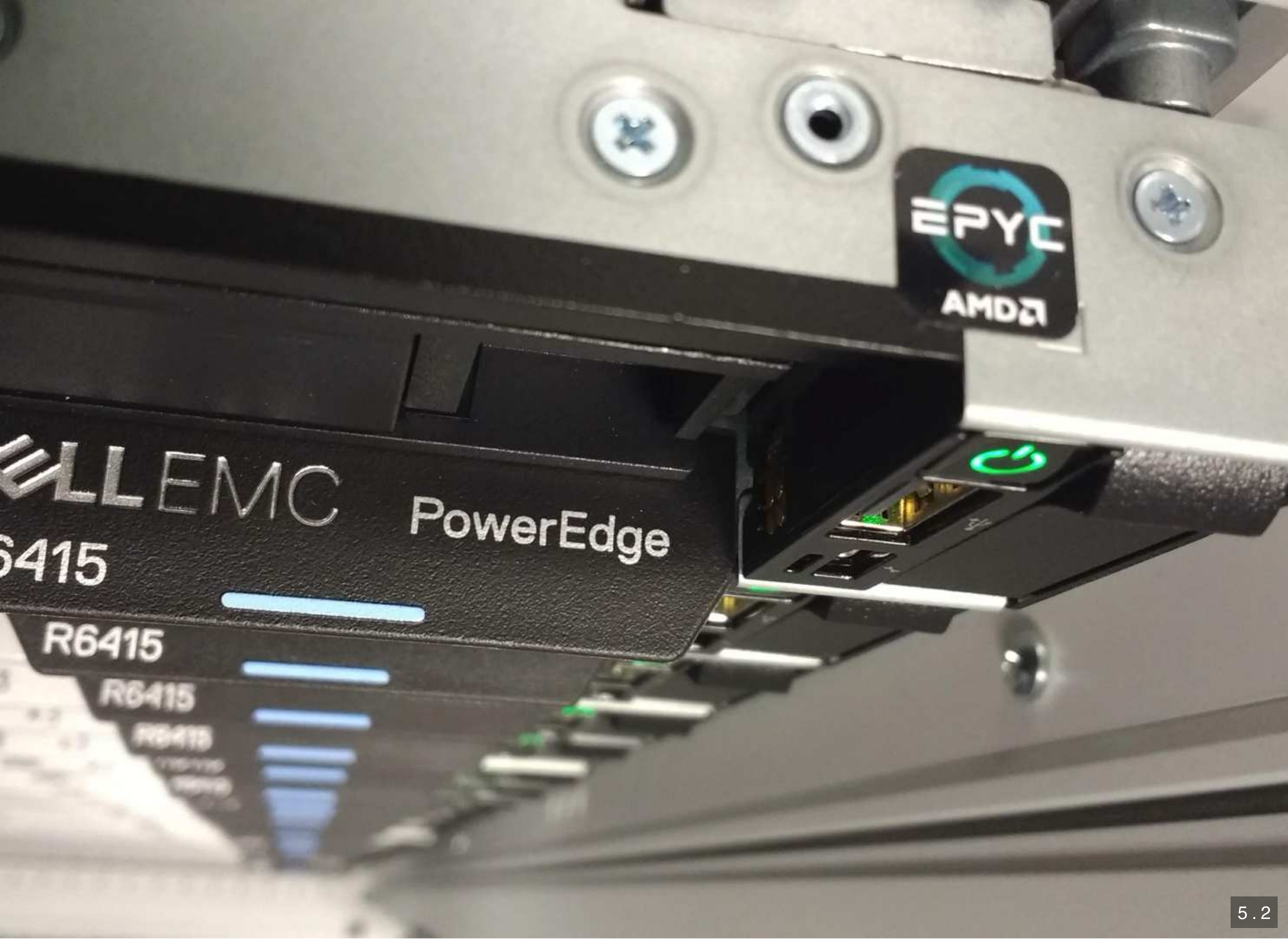
COMPUTING LAST YEAR

SURFSARA

(thanks to Onno Zweers)

- first non-HEP pilots for high throughput openstack platform;
- LOFAR download servers to be replaced with a solution based on dCache macaroons;
- SKA: started running network tests with Australia & South Africa;
- New storage dashboards for VOs;
- CEPH as backend underneath dCache (v4.2): so far, performance not yet production-ready.

NEW HARDWARE


EPYC
AMD
LLEMC
PowerEdge
6415
R6415
R6415
R6415

AMD EPYC

81 DELL R6415 nodes (3 racks) with one AMD EPYC 7551P 32-Core Processor.

- Two racks for NIKHEF-ELPROD compute cluster.
- One rack replaces the local stoomboot cluster (more about that later).

Price/performance is good.

CPU	HEPSPEC06/core	€/core
Intel(R) Xeon(R) Gold 6148	19.57	315
AMD EPYC 7551P	14.94	247

Tried both with and without hyperthreading. Seemed to provide little benefit so turned off now.

Single socket system doesn't suffer cache coherence penalty.

Fast 3.2 TB NVMe SSDs for local storage

GRID STORAGE

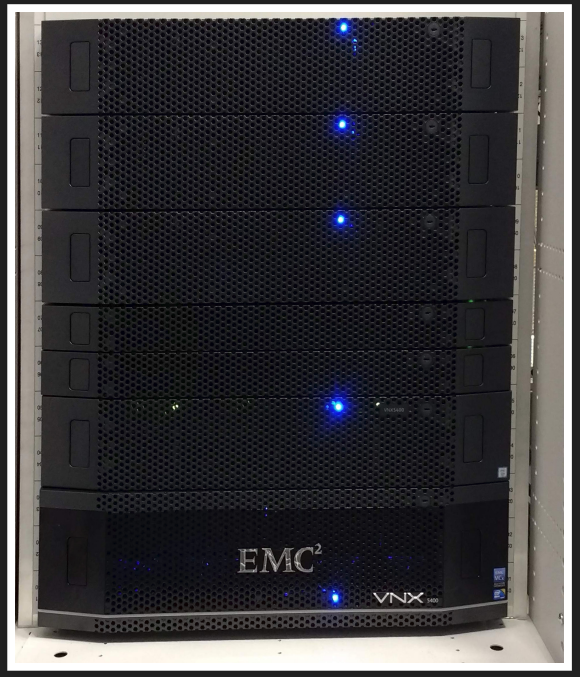
/DATA

Old Hitachi storage system replaced by Netapp.

- Initially tried a Fujitsu Eternus DX200, but this did not work because it could not handle LDAP+TLS.
- Currently a dual NetApp FAS8200 with 2κ 300TB Disk and 2κ 9TB SSD
- Redundancy but no back-up.

/PROJECT: EMC VNX5400

This was bought for the /project storage systems; important data that has an external backup. Data is available on Unix and Windows systems, so mixed mode of Unix permissions and Windows ACLs.



Not too many vendors offer this combination. In fact, the EMC is the last of its kind to offer it.

OFF-SITE BACKUP

Partnership with University of Groningen led to a geographically separated off-site backup solution.

Hardware and software managed and maintained by Groningen; component replacements done by local team.

Much more economical than the previous commercial offering.

Nikhef

DEVELOPMENTS

PRIVATE CLOUD

Setting up a high-throughput private compute cloud with Openstack.

Project was not getting much traction. Several avenues were explored; lack of long-term stability in the software development played a role.

Currently given a higher priority due to internal demand for alternatives to classic cluster computing. Nikhef as a lab is trying *agile* as a way of managing projects and the cloud project is likewise approached as agile.

SALT CONFIGURATION MANAGEMENT

Progress has been made to bring more systems under salt's control. LDAP, DNS, dCache.

Integrated generation of Icinga checks.

Not replacing the legacy grid services which may not be with us much longer.

DCACHE WITH DISTRIBUTED NFS ON LOCAL CLUSTER

High-throughput clustered local file access.

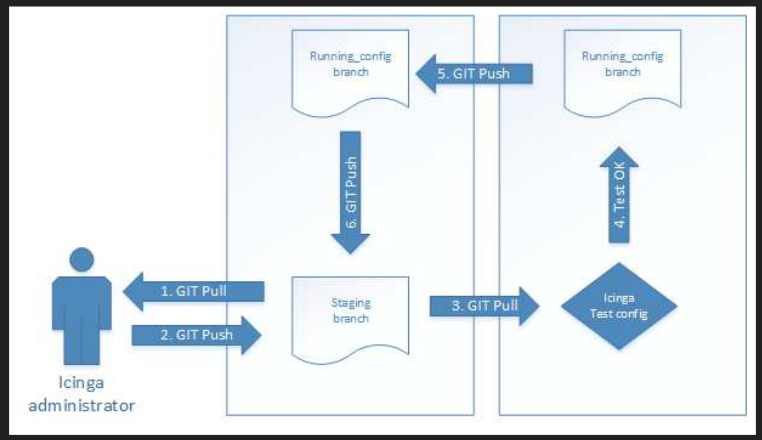
Much appreciated by users

But…see below.

ICINGA FOR IT SERVICES

- Migration from Nagios to Icinga2
  - Old Nagios managed by single person difficult to get new hosts and probes deployed.
  - In the new system everyone should be able to add/edit probes for their hosts.

Standard Icinga2 install;
configuration on Gitlab server;
server will install and test new config automatically.



Nikhef

INTERESTING TIMES

EMC ACL ISSUES

Transferring ACLs from our older systems overloaded the write buffer which locked op the entire volume.

A workaround was eventually found but it took months for DELL EMC to have the correct engineer on the case and to come up with a good diagnosis and a workable solution.

The problem was reported in July, a resolution came in October.

MEMORY ISSUES

A second issue involved the memory consumption of tcp NFS connections. These turned out to be significantly higher than for udp, leading to a system crash and downtime.

This issue took even longer to resolve; the initial problem was found in Spring 2018; the conclusion was that the machine was sold with the wrong specs. A solution was proposed in December which means that in April we will have a planned downtime for a system upgrade.

DCACHE NFS

- original cluster had 8 cores/node, and 1 Gbps
- new nodes have 25 Gbps and 32 cores
- Possible bugs in NFS client implementation in Linux kernel
- Possible bugs in distributed NFS in dCache

The result was too much traffic through the NFS door and hanging clients.

Resolution:

- temporarily reinstall part of cluster with SLC6
- disable dCache NFS access on other nodes
- stress test installation of dCache 5
- Planned upgrade in April

Nikhef

FUTURE PLANS

ROLL OUT A HIGH-THROUGHPUT CLOUD

The vision of this is still blurry. It is hard to get a good sense from the users what they would want. Finding a representative pilot community is important.

If possible, we will virtualize the current legacy grid capacity and explore its elasticity.

CREAM ON THE WAY OUT

Considering alternatives:

- ARC
- HTCondor (adopting HTCondor for more than just batch jobs seems to have benefits)

400 GB NETWORK TESTS

Talk to Tristan.

DPM/DCACHE

Trying to minimize the legacy Knowledge maintainance (technical debt); since dCache is now used at Nikhef, having both dCache and DPM in one lab makes no sense. We're going to phase out DPM and set up dCache for grid storage.

SEE YOU IN AMSTERDAM
HEPiX Amsterdam, 14–18 October 2019