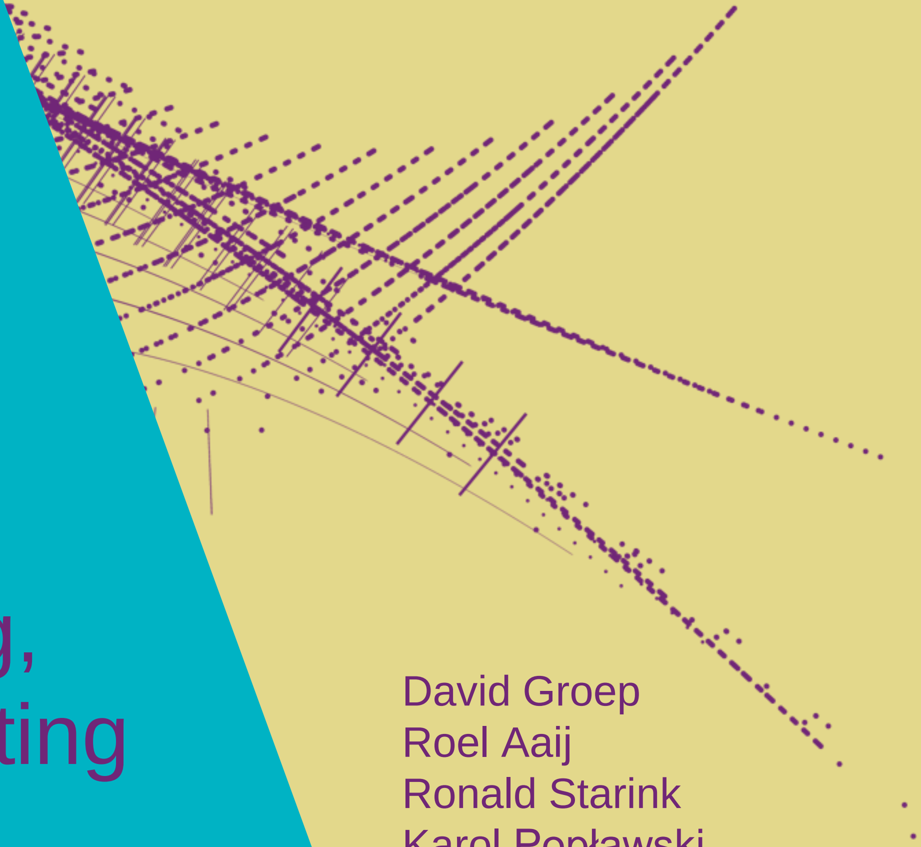
Nikhef

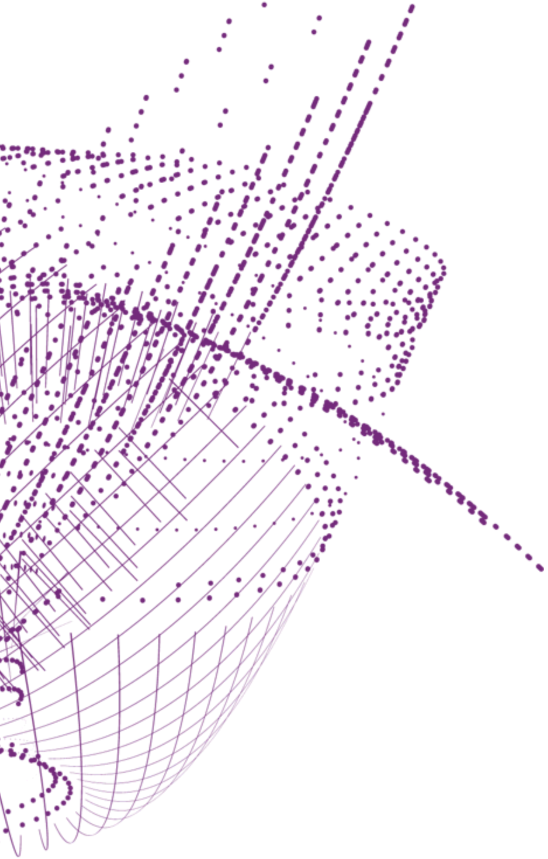
Nikhef SEP panel visit November 2023

Physics Data Processing, Engineering and Computing

accelerating 'time to results' through computing and collaboration



David Groep
Roel Aaij
Ronald Starink
Karol Popławski
Osama Karkout



David Groep | PL Physics Data Processing

Physics Data Processing & ‘CT-PDP’ engineering

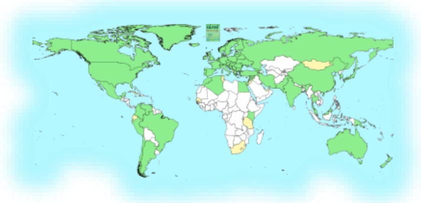
PDP: Computing as research and instrumentation

validated through our real-life applications

Physics Data Processing programme lines

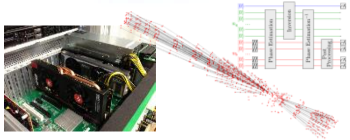
1. infrastructure, network & systems research

- **building** ‘research IT facilities’ through co-design & development
- big data science innovation: research **next gen IT infrastructure**



2. infrastructure for trusted collaboration

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



3. algorithmic design patterns - *more in Roel’s introduction*

- **GPU accelerated** computing, **Quantum Computing**, AI and **Machine Learning**

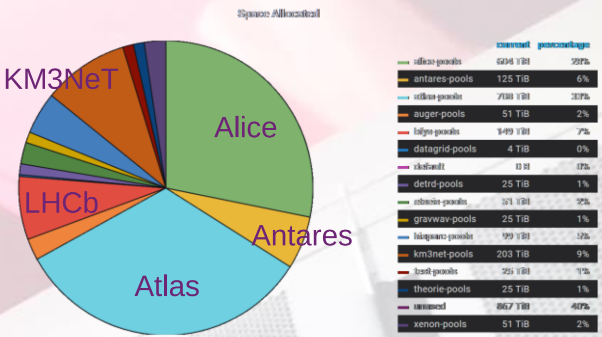
Images: H234b data centre; IGTF trust map; plofkip.nikhef.nl; Davide Nicotra, Maastricht and Nikhef, arXiv:2308.00619v2 quant-ph for JInst.

Physics Data Processing and CT-PDP engineering effort

2.5	staff	(**David Groep**, **Roel Aaij**, Jeff Templon)
1	postdoc	(Maarten van Veghel)
~ 11	engineers: DevOps, research software, RDM, Collaboration / Trust & Identity *with organic embedding in the Computer Technology engineering group*	

With wide range of activity leads, including for example **LHCb RTA Reco convener**, **SURF innovation expert group chair**, **EOSC** Security Coordinator, **AEGIS** Trust and Identity policy lead, **Interoperable Global Trust Federation** chair, Dutch **National Infrastructure Executive**, lead for the **Thematic DCC Natural and Engineering Sciences**, and members in board & committees for the (global) e-Infrastructure landscape: GEANT GCC, PC-GWI, CieDO…

Infrastructure for Research

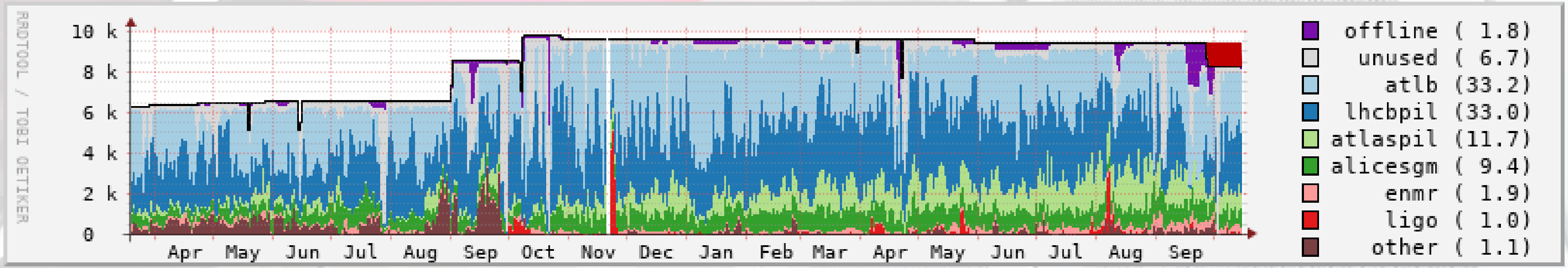


	current	percentage
alice-pools	504 TiB	23%
antares-pools	125 TiB	6%
atlas-pools	708 TiB	33%
auger-pools	51 TiB	2%
lhcb-pools	149 TiB	7%
datagrid-pools	4 TiB	0%
default	0 B	0%
detrd-pools	25 TiB	1%
[illegible]	51 TiB	2%
gravwav-pools	25 TiB	1%
[illegible]	99 TiB	5%
km3net-pools	203 TiB	9%
test-pools	25 TiB	1%
theorie-pools	25 TiB	1%
unused	867 TiB	40%
xenon-pools	51 TiB	2%

High-Through Compute (HTC) + HT Storage

- **National e-Infrastructure** coordinated by SURF
  LHC NL-T1, IGWN, KM3NeT, Xenon, DUNE, WeNMR, MinE ...
- **'Stoomboot' local analysis facility** + IGWN cluster & 'submit node'

~ 12 000 cores (total), 13 PByte storage – very competitive *w.r.t.* commercial cloud

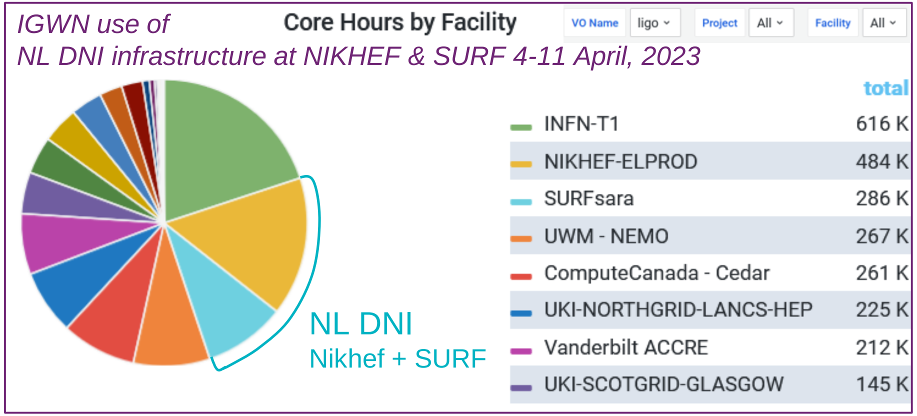


Occupancy: NDPF DNI processing facility in the period March 2021 .. October 2022. Top-right: storage capacity allocated in DNI Nikhef segment

Common solutions are essential for our ‘national’ facility

Alignment of common e-Infrastructure
and shared use by experiments
(LHC, GW, KM3NeT, DUNE, Xenon, ...)

- *common solutions,* since bespoke systems for each experiment do not scale for Nikhef (or NL)
- efficient sharing of both hardware and DevOps effort
- synergy with other domains helps sustainable funding



IGWN use of
NL DNI infrastructure at NIKHEF & SURF 4-11 April, 2023

Core Hours by Facility

VO Name: ligo · Project: All · Facility: All

	total
🟩 INFN-T1	616 K
🟨 NIKHEF-ELPROD	484 K
🟦 SURFsara	286 K
🟧 UWM - NEMO	267 K
🟥 ComputeCanada - Cedar	261 K
🟦 UKI-NORTHGRID-LANCS-HEP	225 K
🟪 Vanderbilt ACCRE	212 K
🟪 UKI-SCOTGRID-GLASGOW	145 K

NL DNI
Nikhef + SURF

Continuation of our long-term strategy - from EU DataGrid in 2000 onwards:

- drives collaborative efforts we work with for identity management and **common protocols globally**:
  ACCESS-CI and CILogon (US) – key players for LIGO, DUNE, and US-ATLAS
- common **processing framework** development (e.g. together with KM3NeT in its INFRADEV project)

Data on Tue April 11th 2023 from the OSG accounting for the LIGO VO for past week (with also SURFsara fully in production)
https://gracc.opensciencegrid.org/d/9u1-Q3vVz/cpu-payload-jobs?orgId=1&var-ReportableVOName=ligo&var-Project=All&var-Facility=All&var-Probe=All&var-interval=1d&from=1680566400000&to=1681257600000

Innovation *on* infrastructure

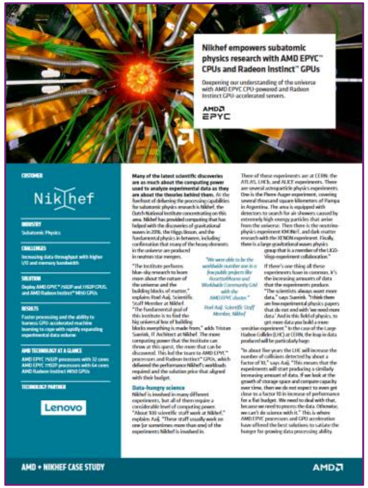
- Network-to-systems integration
- Storage throughput & parallelism
- Systems integration design and tuning

FUNGIBLE

Contact Us

NIKHEF, SURF AND FUNGIBLE SET NEW BENCHMARK FOR THE WORLD'S FASTEST STORAGE PERFORMANCE

Companies Double Current Performance Record, Setting the New Bar at 6.55 Million Read IOPS



Nikhef empowers subatomic physics research with AMD EPYC™ CPUs and Radeon Instinct™ GPUs

Deepening our understanding of the universe with AMD EPYC CPU-powered and Radeon Instinct GPU-accelerated servers.

AMD EPYC

CUSTOMER

Nikhef

INDUSTRY

Subatomic Physics

CHALLENGES

Increasing data throughput with higher I/O and memory bandwidth

SOLUTION

Deploy AMD EPYC™ 7502P and 7702P CPUS, and AMD Radeon Instinct™ MI50 GPUs

RESULTS

Faster processing and the ability to harness GPU-accelerated machine learning to cope with rapidly expanding experimental data volume

AMD TECHNOLOGY AT A GLANCE

AMD EPYC 7502P processors with 32 cores
AMD EPYC 7702P processors with 64 cores
AMD Radeon Instinct MI50 GPUs

TECHNOLOGY PARTNER

Lenovo

Many of the latest scientific discoveries are as much about the computing power used to analyze experimental data as they are about the theories behind them. At the forefront of delivering the processing capabilities for subatomic physics research is Nikhef, the Dutch National Institute concentrating on this area. Nikhef has provided computing that has helped with the discoveries of gravitational waves in 2016, the Higgs Boson, and the fundamental physics in between, including confirmation that many of the heavy elements in the universe are produced in neutron star mergers.

"The Institute performs blue-sky research to learn more about the nature of the universe and the building blocks of matter," explains Roel Aaij, Scientific Staff Member at Nikhef. "The fundamental goal of this institute is to find the big universal box of building blocks everything is made from," adds Tristan Suerink, IT Architect at Nikhef. The more computing power that the Institute can throw at this quest, the more that can be discovered. This led the team to AMD EPYC™ processors and Radeon Instinct™ GPUs, which delivered the performance Nikhef's workloads required and the solution price that aligned with their budget.

> "We were able to be the worldwide number one in a few public projects like Rosetta@home and Worldwide Community Grid with the AMD EPYC cluster."
>
> Roel Aaij, Scientific Staff Member, Nikhef

Data-hungry science

Nikhef is involved in many different experiments, but all of them require a considerable level of computing power. "About 100 scientific staff work at Nikhef," explains Aaij. "These staff usually work on one (or sometimes more than one) of the experiments Nikhef is involved in. Three of these experiments are at CERN: the ATLAS, LHCb, and ALICE experiments. There are several astroparticle physics experiments. One is the Pierre Auger experiment, covering several thousand square kilometers of Pampa in Argentina. The area is equipped with detectors to search for air showers caused by extremely high energy particles that arrive from the universe. Then there is the neutrino-physics experiment KM3NeT, and dark matter research with the XENON experiment. Finally, there is a large gravitational waves physics group that is a member of the LIGO-Virgo experiment collaboration."

If there's one thing all these experiments have in common, it's the increasing amounts of data that the experiments produce. "The scientists always want more data," says Suerink. "I think there are few experimental physics papers that do not end with 'we need more data.' And in this field of physics, to get more data you build a more sensitive experiment." In the case of the Large Hadron Collider (LHC) at CERN, the leap in data produced will be particularly huge.

"In about five years the LHC will increase the number of collisions detected by about a factor of 10," says Aaij. "This means that the experiments will start producing a similarly increasing amount of data. If we look at the growth of storage space and compute capacity over time, then we do not expect to even get close to a factor 10 in increase of performance for a flat budget. We need to deal with that, because we need to process the data. Otherwise, we can't do science with it." This is where AMD EPYC processors and GPU acceleration have offered the best solutions to satiate the hunger for growing data processing ability.

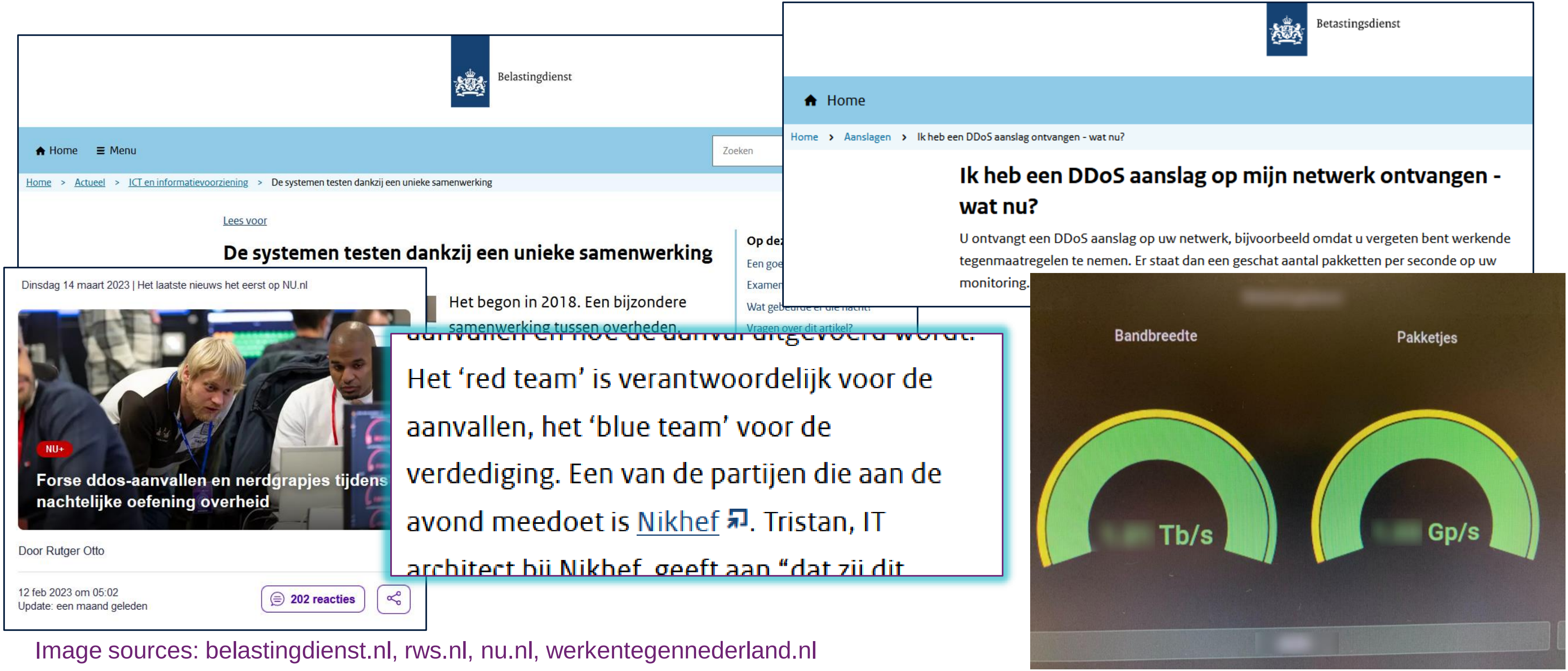
AMD + NIKHEF CASE STUDY

AMD

- early **engineering engagement with vendors** to build us suitable systems
- **co-design** of our national HPC systems ('Snellius')
- **data-intensive compute** with DPUs, or on-NIC FPGAs?
- networks > 800Gbps, >1 Bpps (today: 400G to CERN)

Image: Minister of Economic Affairs M. Adriaansens launched the Innovation Hub with Nikhef, SURF, Nokia and NL-ix, January 2023. Composite image from https://www.surf.nl/nieuws/minister-adriaansens-lanceert-testomgeving-voor-supersnelle-netwerktechnologie; Bluefield Hackathon by Nvidia/Mellanox; *abbreviations*: **DPU**: Data Processing Unit; **on-NIC FPGAs:** on-network interface card field programmable gate arrays; **pps**: packets per second

Our science data flows are somebody else’s DDoS attack



Belastingdienst

Home ≡ Menu Zoeken

Home > Actueel > ICT en informatievoorziening > De systemen testen dankzij een unieke samenwerking

Lees voor

De systemen testen dankzij een unieke samenwerking

Het begon in 2018. Een bijzondere samenwerking tussen overheden.

Op de
Een goe
Examen
Wat gebeurde er die nacht?
Vragen over dit artikel?

Belastingdienst

Home

Home › Aanslagen › Ik heb een DDoS aanslag ontvangen - wat nu?

Ik heb een DDoS aanslag op mijn netwerk ontvangen - wat nu?

U ontvangt een DDoS aanslag op uw netwerk, bijvoorbeeld omdat u vergeten bent werkende tegenmaatregelen te nemen. Er staat dan een geschat aantal pakketten per seconde op uw monitoring.

Dinsdag 14 maart 2023 | Het laatste nieuws het eerst op NU.nl

NU+

Forse ddos-aanvallen en nerdgrapjes tijdens nachtelijke oefening overheid

Door Rutger Otto

12 feb 2023 om 05:02
Update: een maand geleden

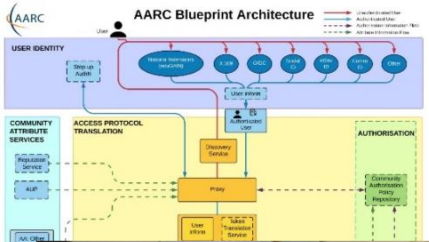
202 reacties

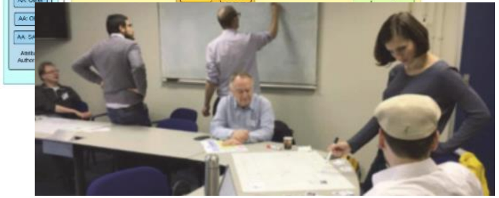
aanvallen en hoe de aanval uitgevoerd wordt.

Het ‘red team’ is verantwoordelijk voor de aanvallen, het ‘blue team’ voor de verdediging. Een van de partijen die aan de avond meedoet is Nikhef. Tristan, IT architect bij Nikhef geeft aan “dat zij dit

Image sources: belastingdienst.nl, rws.nl, nu.nl, werkentegennederland.nl

Infrastructure for Collaboration

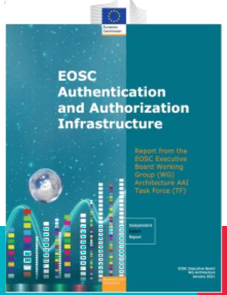




Target impactful areas in architectures for 'AAI' & 'OpSec'

- **authentication & authorization for research collaboration**
  - **AARC** project & community: GEANT Framework projects, R&E federation, identity and credentialing services, EOSC Future, …
  - recently awarded: **EOSC Core** security and **AARC-TREE**
  - **policy frameworks for interoperability** for data protection and global seamless service access
  - continuous **technical evolution** driving IGWN, WLCG in line with AARC and global AAI architecture

- embedding data **processing needs** of our experiments in the (EOSC) landscape
  - **EOSC** Interoperability Framework: Security Baseline, AAI Architecture
  - **eduGAIN** Operational Security for the global R&E inter-federation service
  - **EGI** Advanced computing for research federation**, GEANT community**


EOSC
Authentication and Authorization Infrastructure
Report from the EOSC Executive Board Working Group (WG) Architecture AAI Task Force (TF)

AAI: Authentication and Authorization Infrastructure; **OpSec**: Operational Security (incident response); **AARC**: Authentication and Authorisation for Research Collaboration community and projects**; EOSC**: European Open Science Cloud; **EGI** and **GEANT**: pan-European e-Infrastructure and network collaborations; **IGWN**: International Gravitational Waves Observatory Network;

Collaboration: Research Data Management beyond ‘FA’

FAIR for **live data**, in large volumes, from ‘FA’ towards the ‘I’ and ‘R’

- not *that* many disciplines with really **voluminous data**
  - so nationally join forces with those who do: ASTRON (SRCnet), KNMI (earth observation, seismology), *&c*
- work with those who care about ***software*** to bring data to life on ***infrastructure***
  - NLeSC, 4TU.RD/TUDelft, CWI, and with those who ensure the *infrastructure*: SURF
- for our own analyses and the local (R&D) experiments we work towards *continuous deposition* of **re-usable** data and software:
  - **Thematic Digital Competence Centre** for the Natural and Engineering Sciences
  - **co-develop Djehuty RDM** repository software link ‘Stoomboot’ analysis cluster storage

March 2, 2023

Workshop on FAIR for Physical Objects – Summary and Conclusions

Author: Esther Plomp On the first of March 2023, the...

Read article →

February 22, 2023

Submit your idea for TDCC-NES project program funding

We would like to invite researchers in the scientific domain...

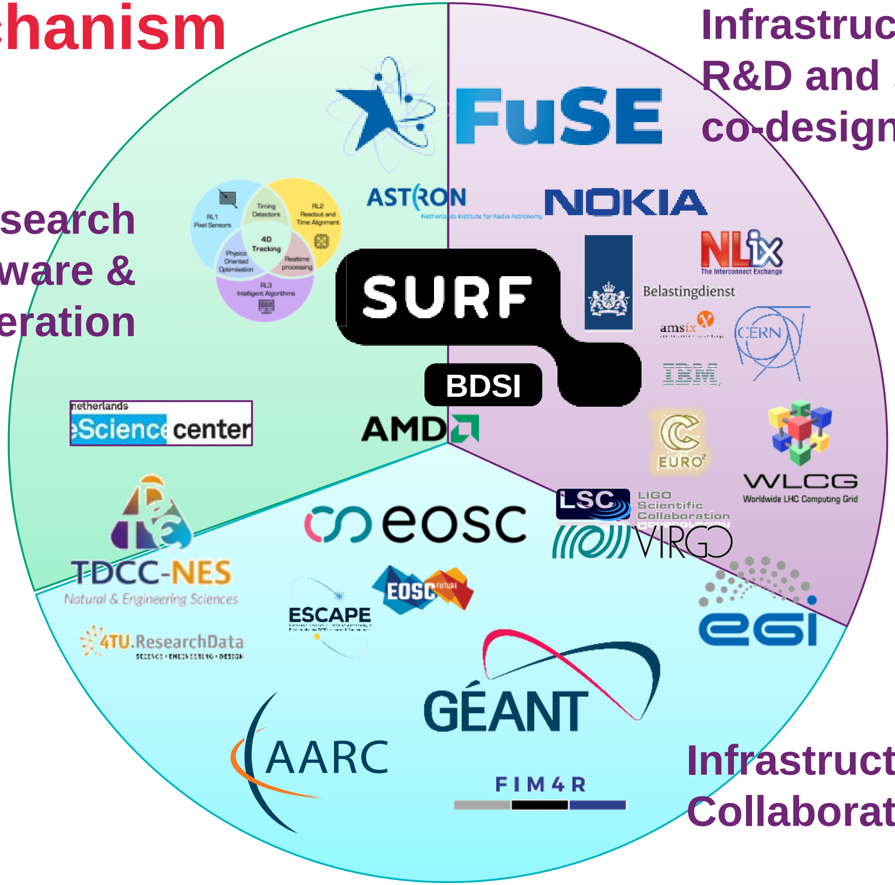
Read article →

PDP strategic projects mechanism

Join initiatives and projects that

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

project pathways include
SURF innovation, GN5-*, AARC-TREE, EOSC Core, LHC4D (planned), …



Research Software & Acceleration

Infrastructure R&D and system co-design

Infrastructure for Collaboration

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

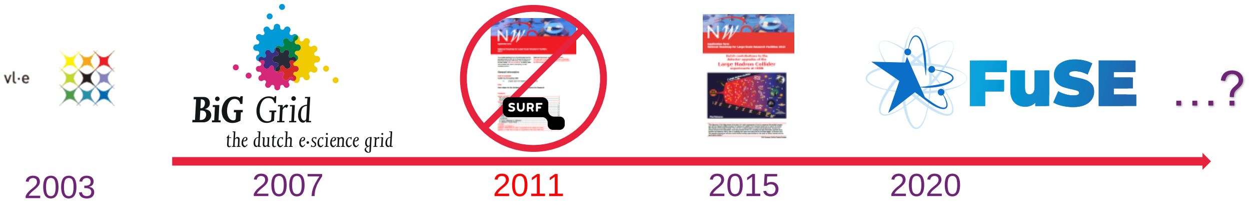
Sustained infrastructure for advanced computing?

Data processing: a **persistent need** for all our experiments. And ‘we’ are not alone!

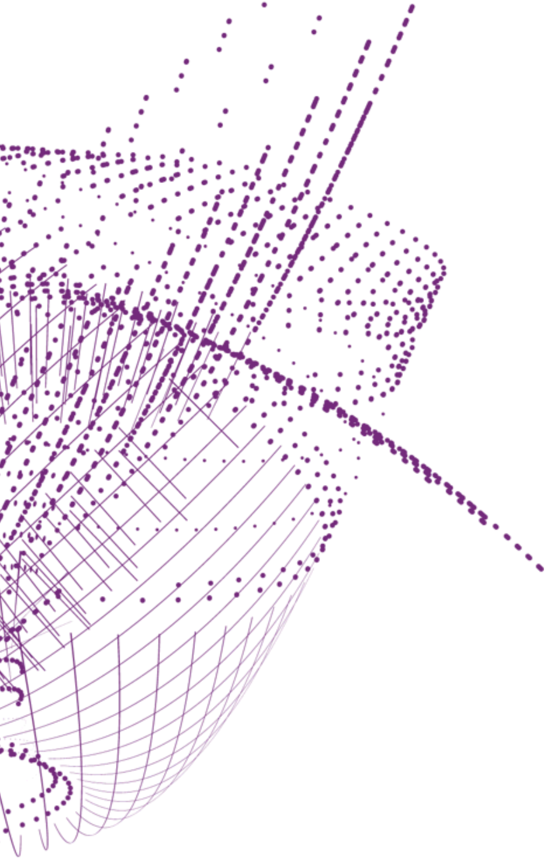
- many disciplines: ESCAPE (our experiments plus astronomy), bio-informatics, health, SSH, …

  need infrastructure to exploit the collected data with **long-term ICT capabilities**

- project-based funding for ‘upgrades’ of infrastructure not the appropriate way of funding **persistent requirements** … but only thing we have at the moment:



timeline: ICES-KIS, BIG and GWI projects for computing capacity, by submission date



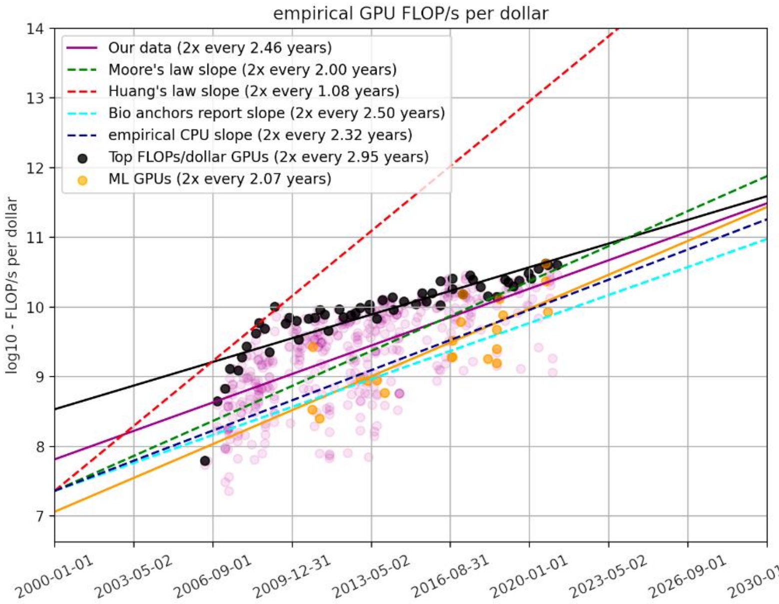
Roel Aaij | Senior Scientific Researcher PDP

Algorithms and Acceleration

Efficient and Scalable Computing For HEP

- CPU-only is not going to be affordable;
  In other words: compute accelerators offer more physics for less money
- Most mature compute accelerators: GPUs

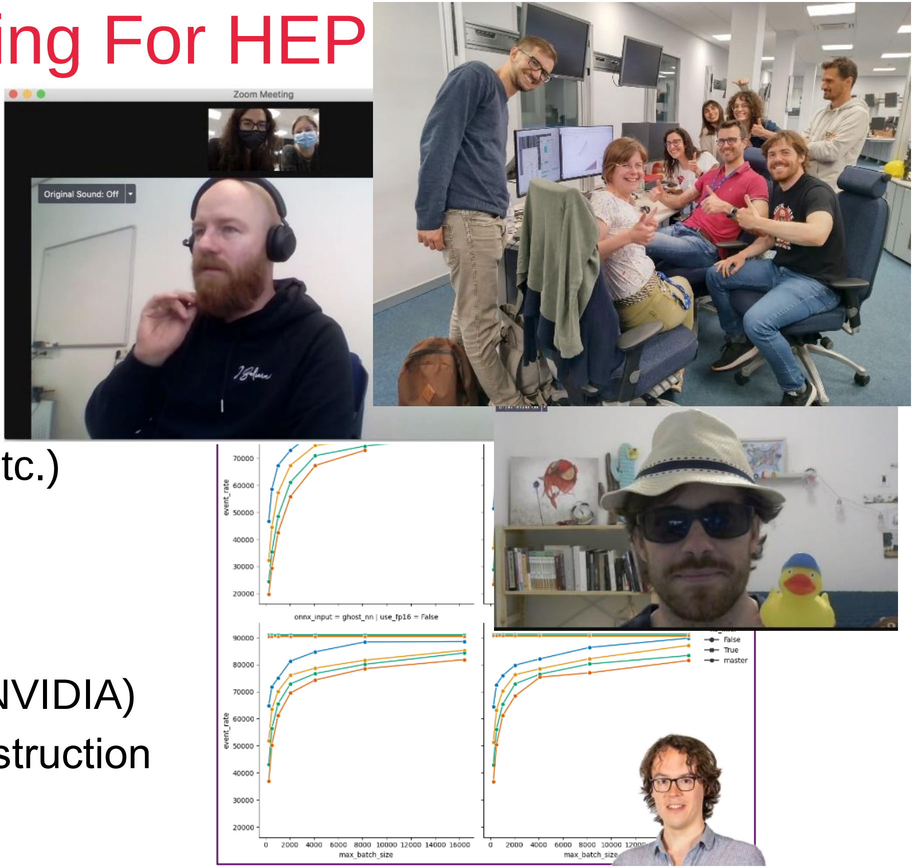
- How to program?
- How to optimize at algorithm and system level?
- How to integrate in frameworks and infrastructure?
- How to do this efficiently in terms of people's time?
- How to maintain?
- What will future (compute) accelerators look like?
- How to provide career perspective to people who focus on (accelerated) software?



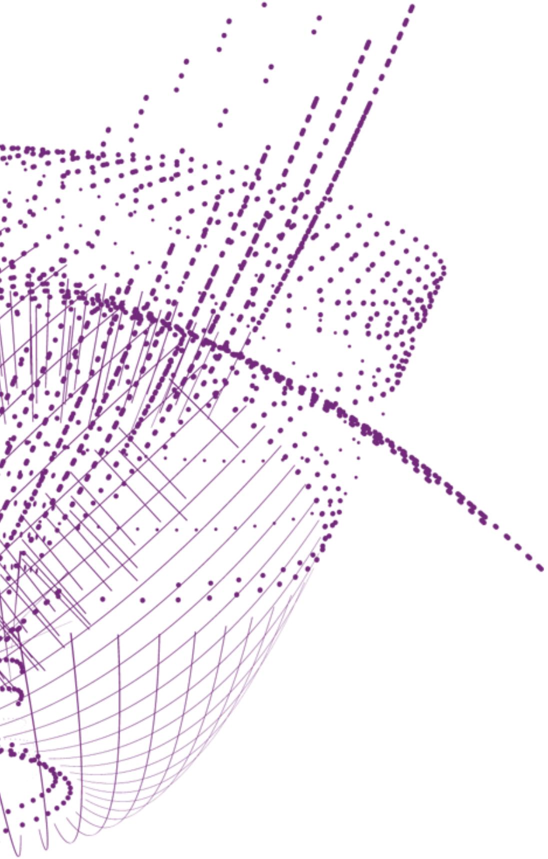
https://epochai.org/blog/trends-in-gpu-price-performance

Efficient and Scalable Computing For HEP

- LHCb’s first-stage GPU trigger (Allen)
  - Bespoke application with all-custom kernels
  - 4 TB/s of detector data on 400 GPUs
  - Nominal luminosity next year
  - Focused on integration (DAQ, LHCb stack, etc.)
- - From idea to R&D to production in 5 years
- With NLeSC: fast ML inference
  - Using standard format (ONNX) and libraries
- User (software) support for GPUs (AMD and NVIDIA)
- FASTER: computing for HL-LHC & ‘4D’ reconstruction



https://github.com/LHC-NLeSC/run-allen-run



Ronald Starink | TGL Computing Technology

Computing group at Nikhef

CT organisation

CT-B: system administration & service desk

- General support for ICT, end user support
- Staff: 7

CT-PDP: dedicated support for the PDP programme

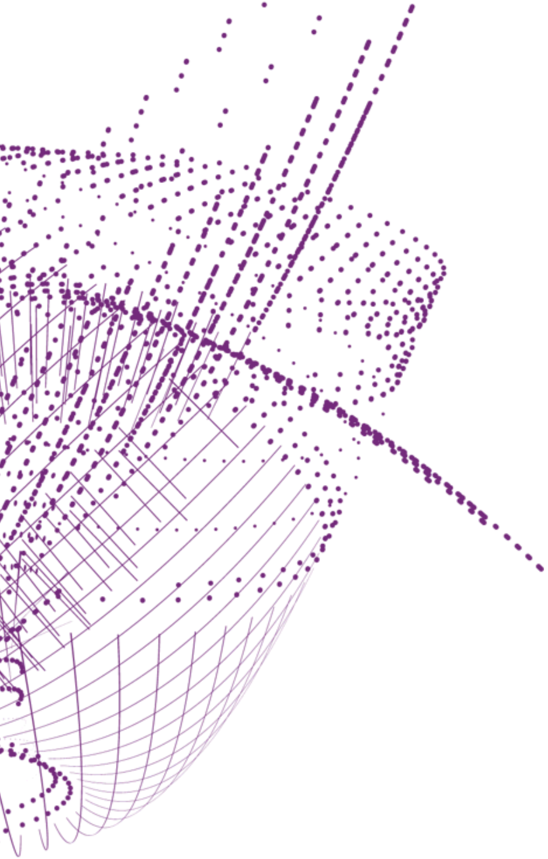
- Software & infrastructure innovation
- Staff: 11
- (→ presentations DG and RA)

CT-PO: support for projects by experiments

- Software engineering: slow controls, data acquisition, analysis framework support
- Staff: 6
- (→ presentation KP)

Challenges

- General: recruitment in a competitive labour market
- CT-B: balance flexibility ↔ standardization
- CT-PO: matching resource with experiments' needs (time, expertise)



Karol Popławski | Software Engineer CT-PO

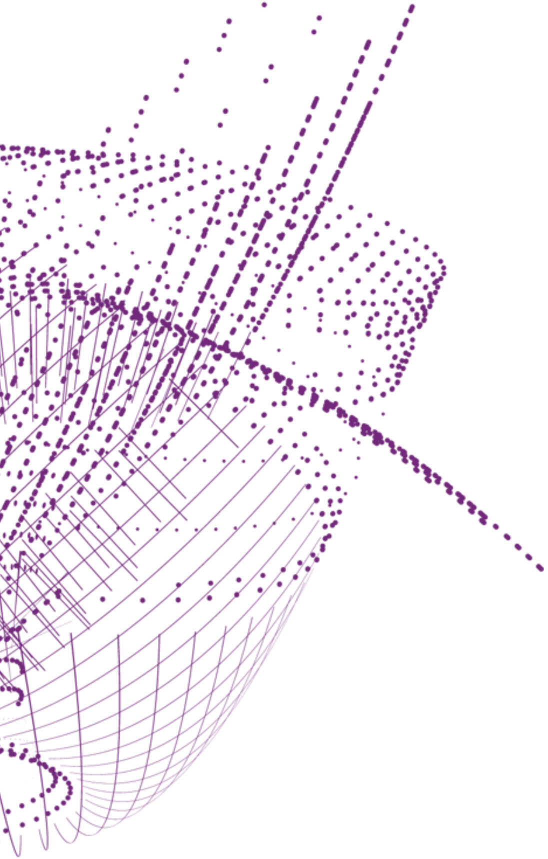
Project engineering & support

KP – Software, Controls and more

Barrel Alignment	DAQ of ~5800 channels
MDT DCS Module	1200 chambers: T, B-field, electronics monitoring & configuration
SciFi tracker	FEE configuration with DB, granularity up to 1.5mln thresholds
Scrum Master	MT, ET, R&D

CT-PO

Henk	FELIX
	MDM firmware
Ton	SciFi FEE calibration
	PTOLEMY
Kostis	Future 4D tracking with White Rabbit

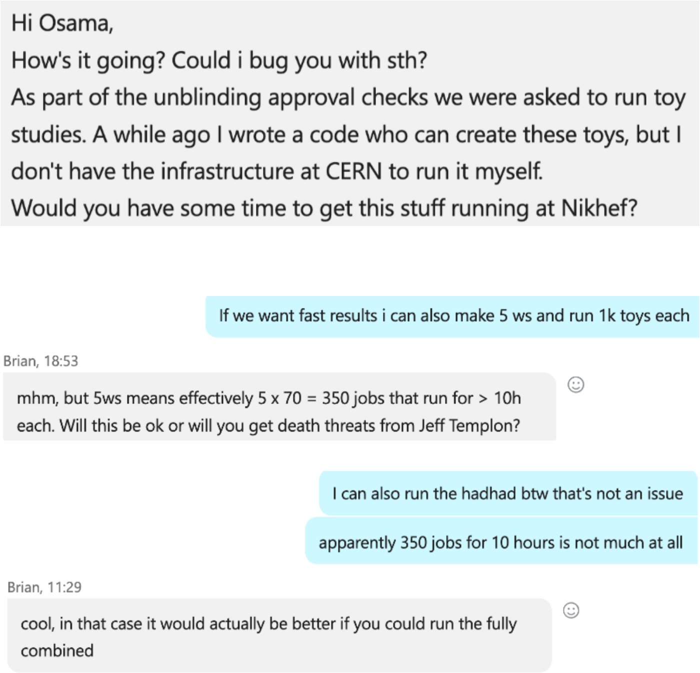


Osama Karkout | happy ATLAS & stoomboot user

How I use stoomboot

Data Analysis

designed for data analysis:
economic & efficient



Hi Osama,

How's it going? Could i bug you with sth?

As part of the unblinding approval checks we were asked to run toy studies. A while ago I wrote a code who can create these toys, but I don't have the infrastructure at CERN to run it myself.

Would you have some time to get this stuff running at Nikhef?

If we want fast results i can also make 5 ws and run 1k toys each

Brian, 18:53

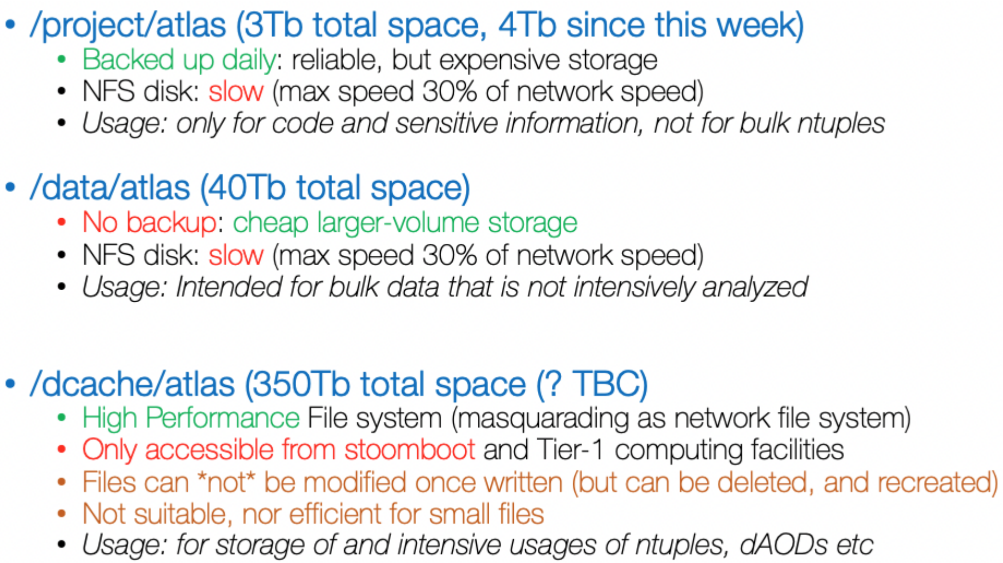
mhm, but 5ws means effectively 5 x 70 = 350 jobs that run for > 10h each. Will this be ok or will you get death threats from Jeff Templon?

I can also run the hadhad btw that's not an issue

apparently 350 jobs for 10 hours is not much at all

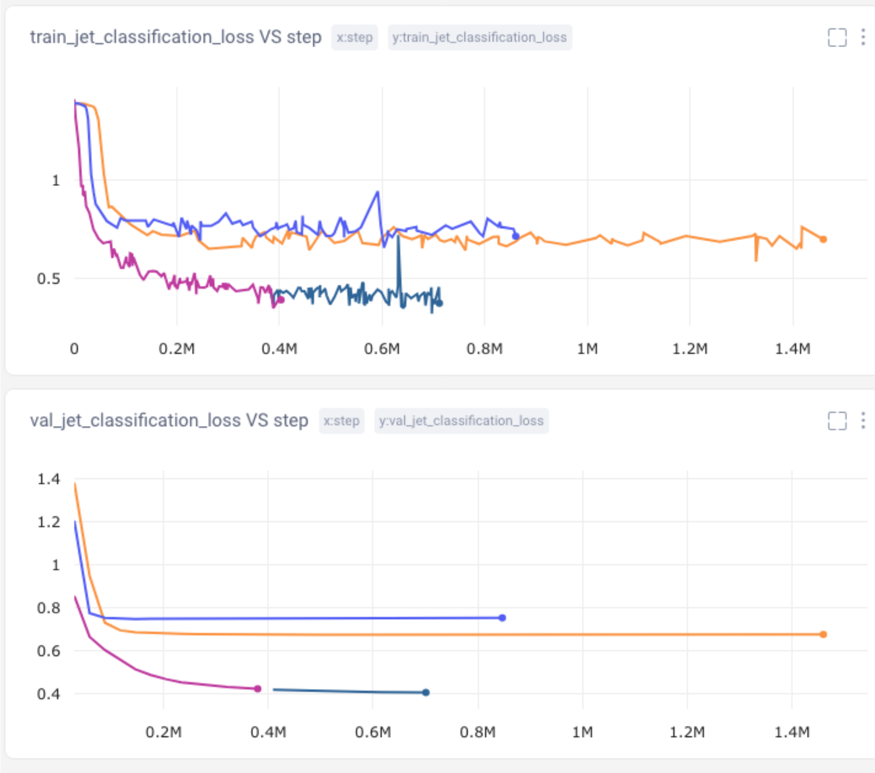
Brian, 11:29

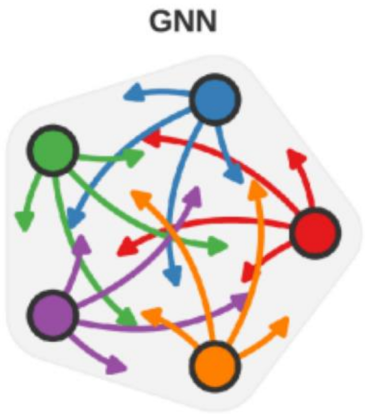
cool, in that case it would actually be better if you could run the fully combined

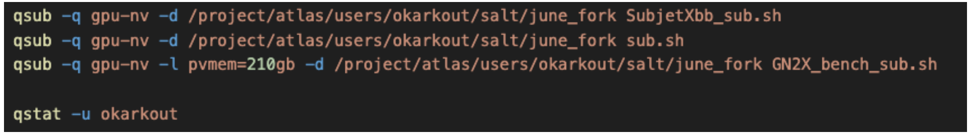


- /project/atlas (3Tb total space, 4Tb since this week)
  - Backed up daily: reliable, but expensive storage
  - NFS disk: slow (max speed 30% of network speed)
  - *Usage: only for code and sensitive information, not for bulk ntuples*
- /data/atlas (40Tb total space)
  - No backup: cheap larger-volume storage
  - NFS disk: slow (max speed 30% of network speed)
  - *Usage: Intended for bulk data that is not intensively analyzed*
- /dcache/atlas (350Tb total space (? TBC)
  - High Performance File system (masquarading as network file system)
  - Only accessible from stoomboot and Tier-1 computing facilities
  - Files can *not* be modified once written (but can be deleted, and recreated)
  - Not suitable, nor efficient for small files
  - *Usage: for storage of and intensive usages of ntuples, dAODs etc*

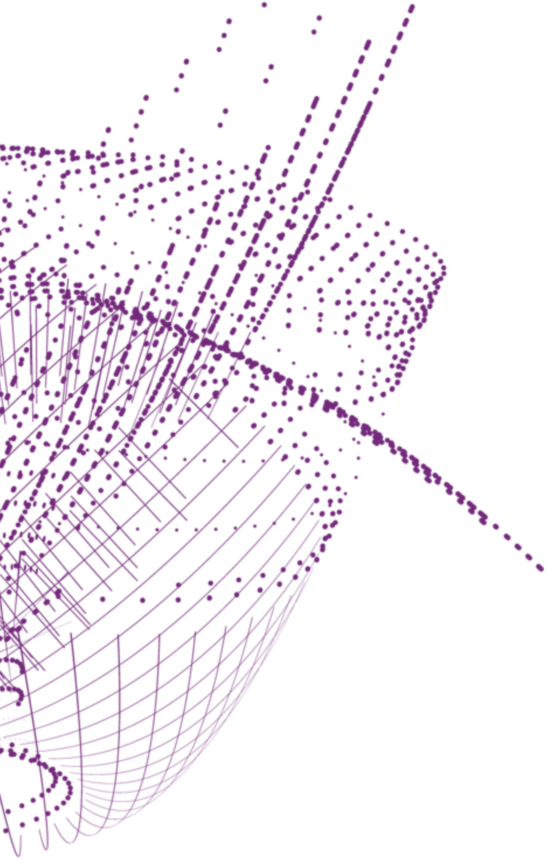
Neural Network training: ATLAS GNNs and Transformers







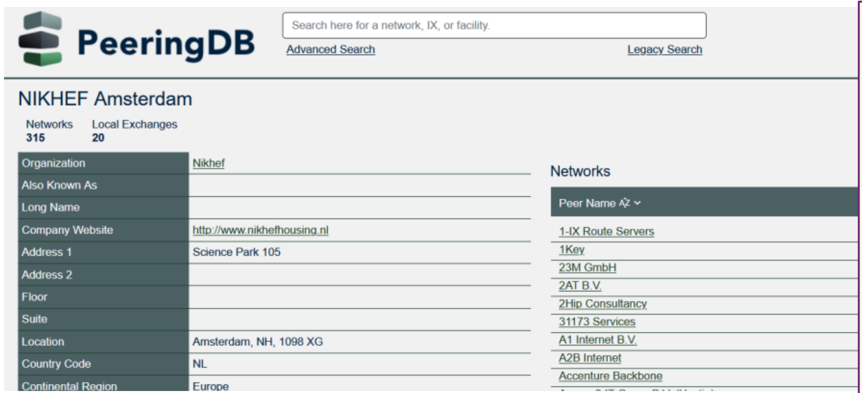
```
qsub -q gpu-nv -d /project/atlas/users/okarkout/salt/june_fork SubjetXbb_sub.sh
qsub -q gpu-nv -d /project/atlas/users/okarkout/salt/june_fork sub.sh
qsub -q gpu-nv -l pvmem=210gb -d /project/atlas/users/okarkout/salt/june_fork GN2X_bench_sub.sh

qstat -u okarkout
```



Arjen van Rijn
David Groep

Data centre visit and networks

Our datacenter is all about connectivity



PeeringDB

Search here for a network, IX, or facility.

Advanced Search · Legacy Search

NIKHEF Amsterdam

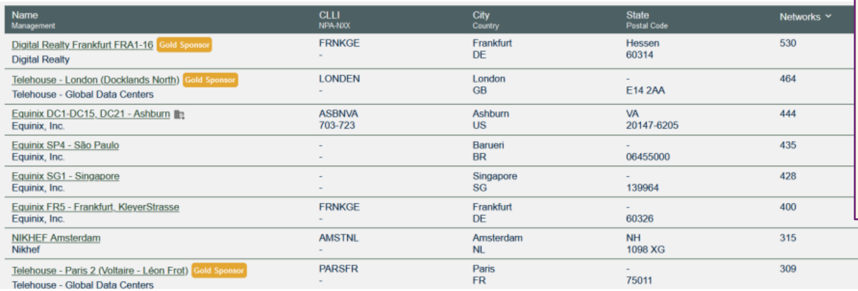
Networks	Local Exchanges
315	20

Organization	Nikhef
Also Known As	
Long Name	
Company Website	http://www.nikhefhousing.nl
Address 1	Science Park 105
Address 2	
Floor	
Suite	
Location	Amsterdam, NH, 1098 XG
Country Code	NL
Continental Region	Europe

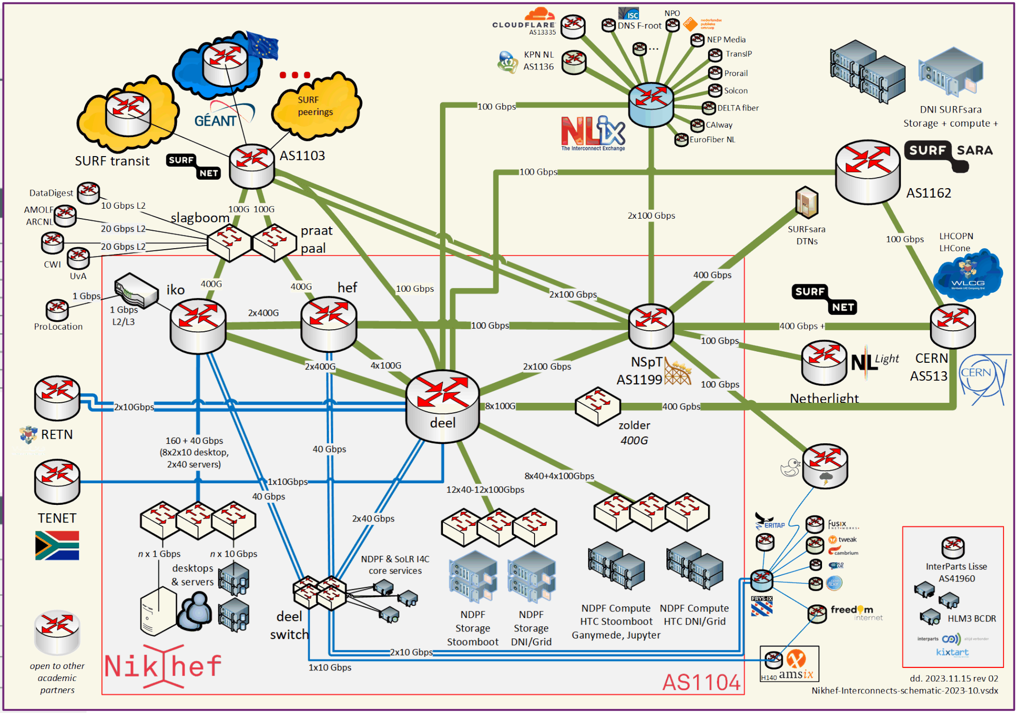
Networks

Peer Name A-Z
1-IX Route Servers
1Key
23M GmbH
2AT B.V.
2Hip Consultancy
31173 Services
A1 Internet B.V.
A2B Internet
Accenture Backbone

#7 on the worldwide peering list



Name Management	CLLI NPA-NXX	City Country	State Postal Code	Networks
Digital Realty Frankfurt FRA1-16 Gold Sponsor Digital Realty	FRNKGE -	Frankfurt DE	Hessen 60314	530
Telehouse - London (Docklands North) Gold Sponsor Telehouse - Global Data Centers	LONDEN	London GB	- E14 2AA	464
Equinix DC1-DC15, DC21 - Ashburn Equinix, Inc.	ASBNVA 703-723	Ashburn US	VA 20147-6205	444
Equinix SP4 - São Paulo Equinix, Inc.	- -	Barueri BR	- 06455000	435
Equinix SG1 - Singapore Equinix, Inc.	- -	Singapore SG	- 139964	428
Equinix FR5 - Frankfurt, KleyerStrasse Equinix, Inc.	FRNKGE -	Frankfurt DE	- 60326	400
NIKHEF Amsterdam Nikhef	AMSTNL -	Amsterdam NL	NH 1098 XG	315
Telehouse - Paris 2 (Voltaire - Léon Frot) Gold Sponsor Telehouse - Global Data Centers	PARSFR -	Paris FR	- 75011	309



Summary data of our datacentre

Maximum power: 2,7 MW

Power Usage Efficiency (PUE): 1,3

Residual heat used for:

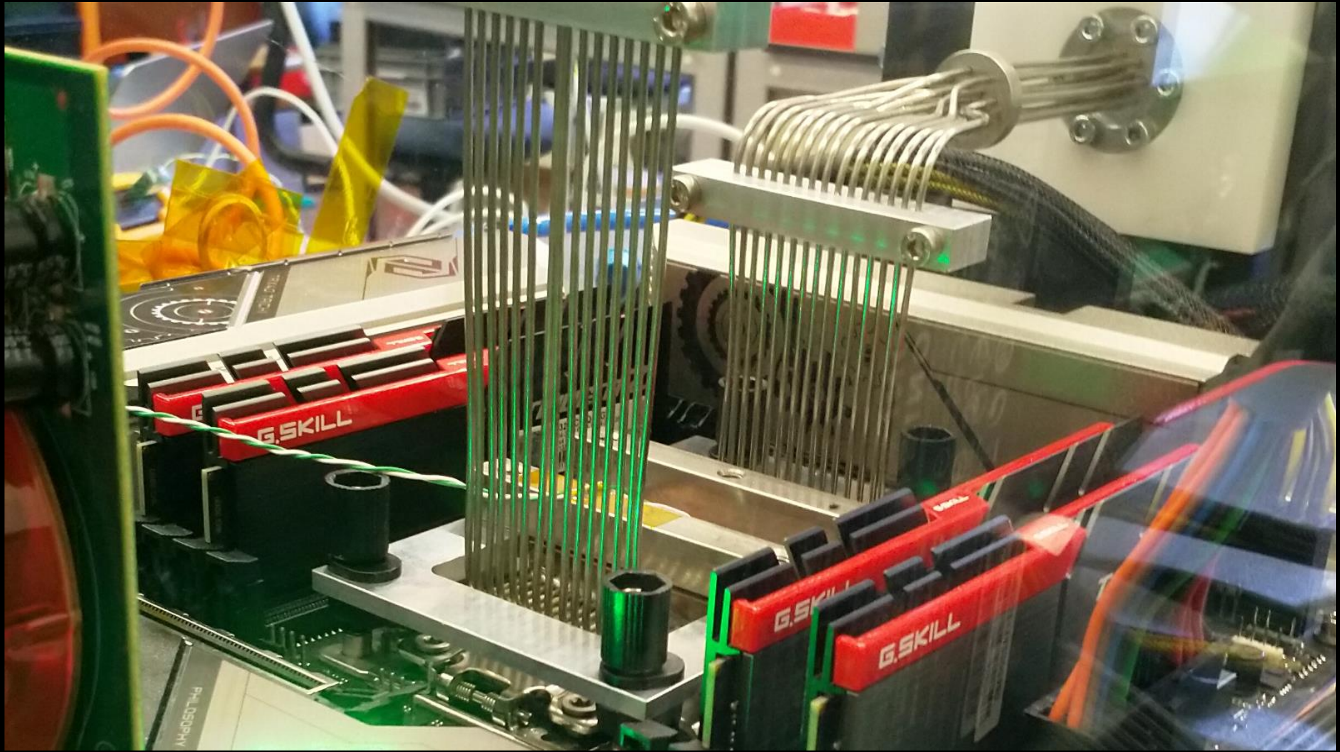
- Nikhef building
- Amsterdam University College
- Student housing

Room	Purpose	# racks max	IT power average per rack (kW)	IT power total (kW)	IT power currently (kW)
H234b	Scientific computing	47	6,25	300	180
H140	Nikhef Housing	282	1,9	660	528
H142	Nikhef Housing (extension)	112	4,0	540	10

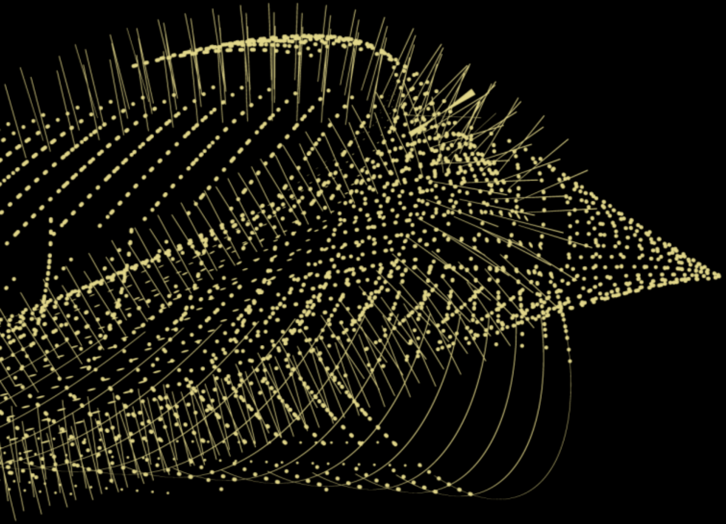
Financial figures (annually):

- Turnover currently ~5 M€; will grow to about 6,5 M€;
- Running costs will grow to 3 – 3,5 M€ (energy costs!);
- Depreciation of extension: 1 M€ annually until end 2028;
- Net result: 2 – 3 M€

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

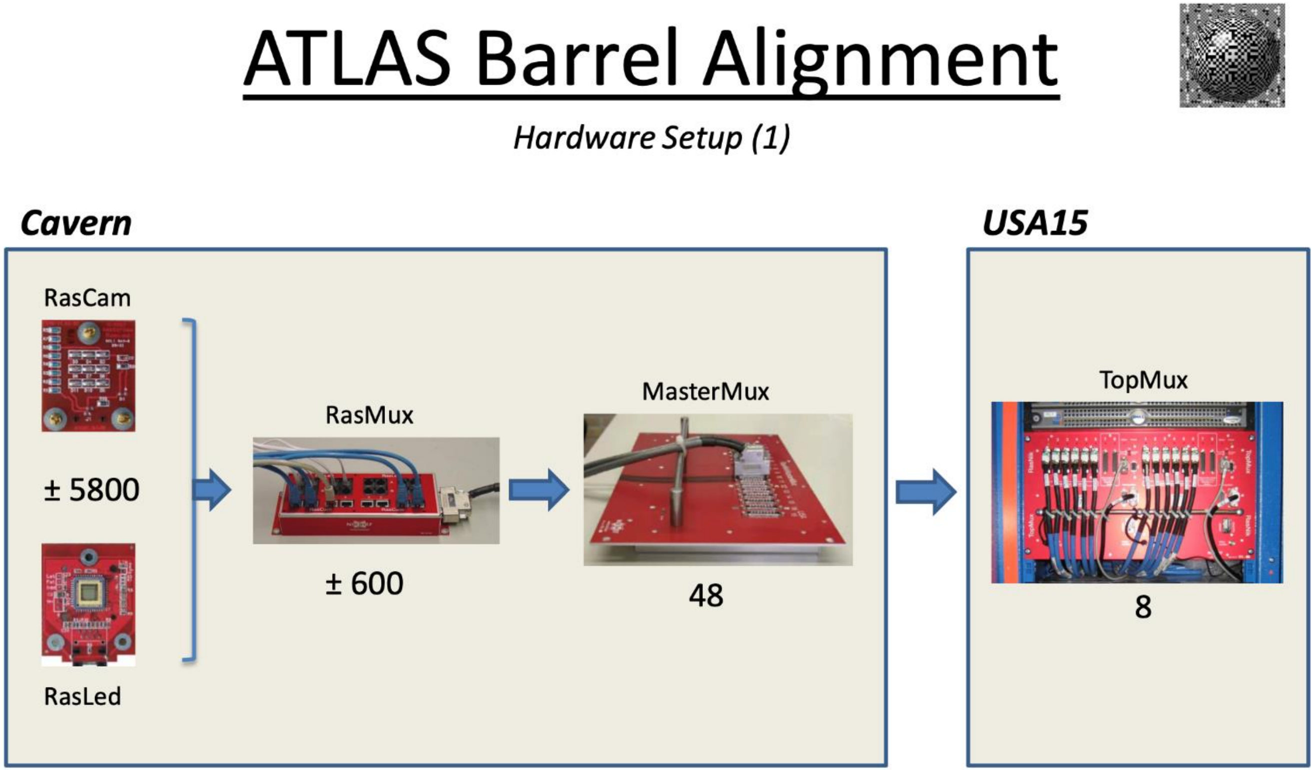


LCO2 cooling of an AMD Ryzen Threadripper 3970X [56.38 °C] at 4600.1MHz processor (~1.5x nominal speed) sustained, using the Nikhef LCO2 test bench system (https://hwbot.org/submission/4539341) - (Krista de Roo en Tristan Suerink)



Supplementary materials

Backup – Balign



ATLAS Barrel Alignment

Hardware Setup (1)

Cavern

RasCam

± 5800

RasLed

RasMux

± 600

MasterMux

48

USA15

TopMux

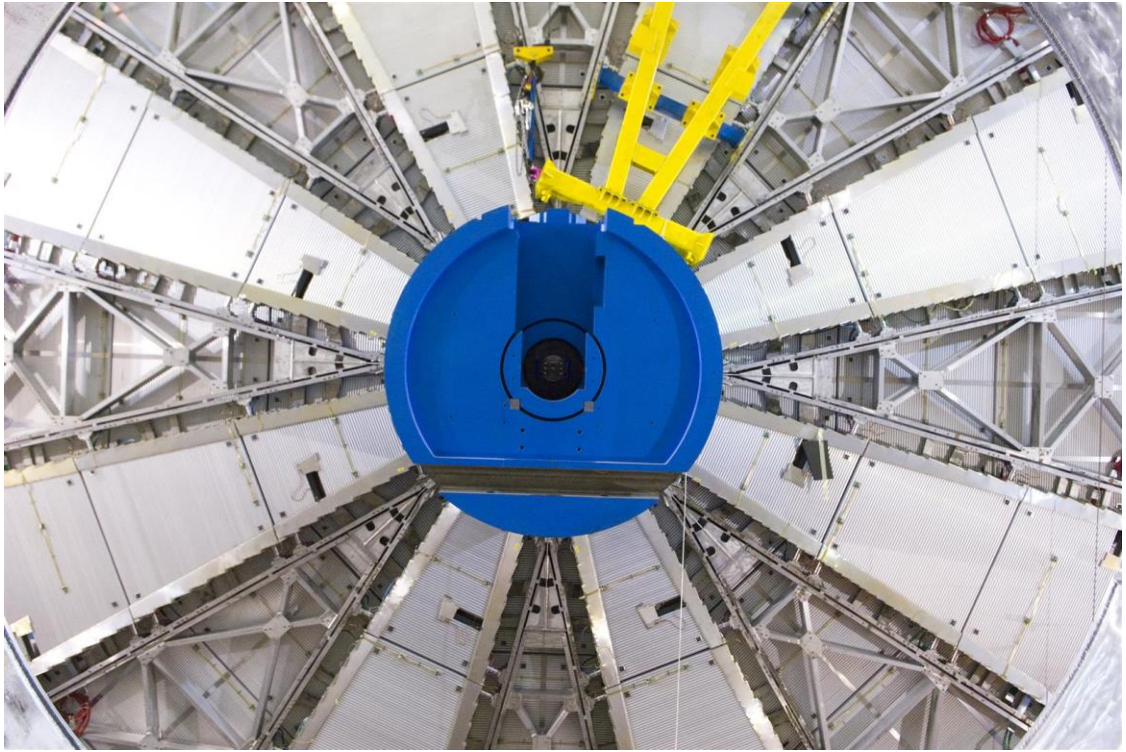
8

Source: Future Barrel Alignment by Robert Hart

Backup – MDM

1,171 chambers with total 354,240 tubes (3 cm diameter, 0.85-6.5 m long)

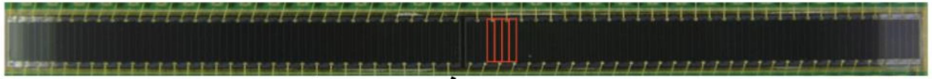
Tube resolution 80 μm



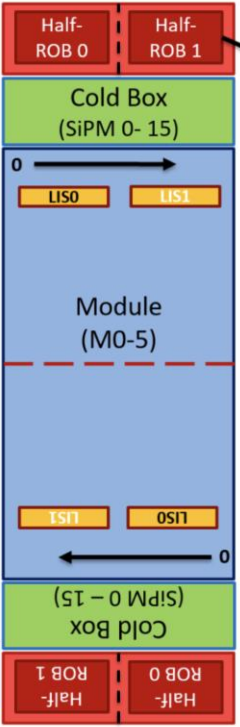
Source: https://atlas.cern/Discover/Detector/Muon-Spectrometer

Backup – SciFi

SciFi - Channels



128-channel SiPM (Hamamatsu)



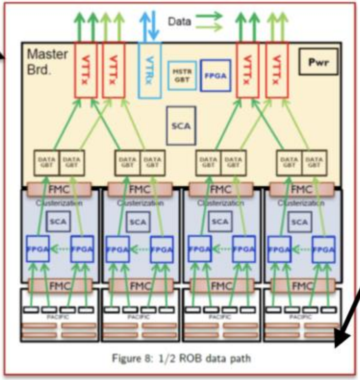
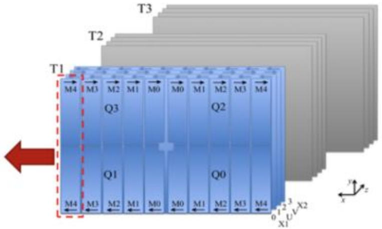


Figure 8: 1/2 ROB data path



512 HalfROBs:

- 512 MBs
- 2048 CBs
- 2048 PBs
- 8192 ASIC chips on PBs
- e.g. 524288 SiPM channels with 3 thresholds

Source: SciFi Collaboration