Provisioning of Grid computing resources in the Norwegian Research and Education Cloud

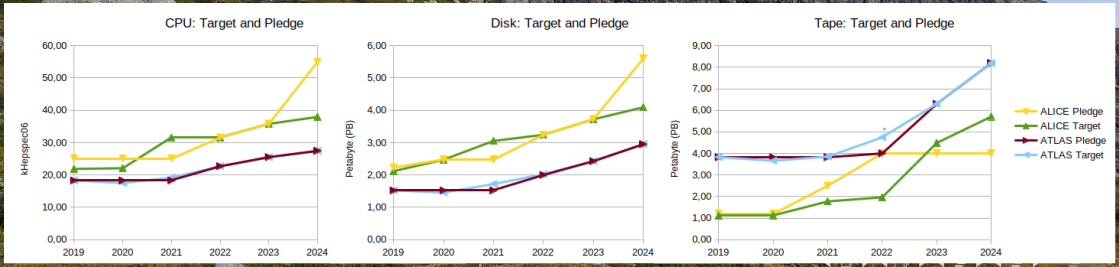
Matthias Richter[1], Emmanuel E. Moutoussamy[2], Terje Furenes[2] Raymond A Kristiansen[2], Maiken Pedersen[3]

[1] Department of Physics and Technology, University of Bergen
[2] IT Department, University of Bergen
[3] IT Department, University of Oslo

Grid computing demands in Norway

Two research activities in Norway connected to LHC experiments: ALICE and ATLAS

	Experiment	CPU	Disk	Tape
Pledged in 2024	ALICE	55.00 kHS06	6.6 PB	4 PB
	ATLAS	27.46 kHS06	2.05 PB	8.19 PB



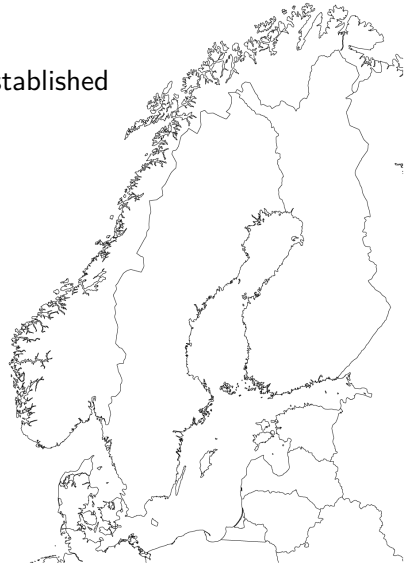
NT1 - Common Nordic Grid site

For the Nordic countries a common Nordic Tier 1 has been established in 2006 and operated as Nordic Data Grid Facility (NDGF)

In 2012, NDGF has been put under the umbrella of Neic, the Nordic e-Infrastructure Collaboration, as Neic Tier1 (NT1) to eventually develop a broader scope beyond WLCG.

Key features:

- Common data storage disk and tape
- Computing resources provided as individual sites.
- 6 Nordic sites and 2 non-Nordic federation additions
- Sites own and run the hardware, usually national funding
- Provides managed services (batch systems, tape archives, etc)



National e-Infrastructure in Norway

Large Scale HPC

Sigma2 AS has strategic responsibility for and manages the national e-infrastructure for large-scale data- and computational science in Norway.

It is now a subsidiary of **Sikt**, the Norwegian Agency for Shared Services in Education and Research.

Collaboration with universities to form **NRIS**, the Norwegian research infrastructure services.

sigma2 NRIS

Medium Size Cloud Service

Norwegian Research and Education Cloud (**NREC**) provides fast, standardized servers and storage for the Norwegian higher education sector.

NREC

- Both options have been considered for operating the Norwegian NT1 resources.

Norwegian Research and Education Cloud - NREC

- National service providing Infrastructure as a service
- Jointly operated by Universities of Bergen and Oslo
- Built on Opensource software, simple setup, economic operation
- Distributed administration team, many people involved part-time, total $\sim$ 5 FTEs
- Openstack as backbone to build the service
- Infrastructure is 100% code-defined
- Easy sharing of infrastructure
- Self-Service infrastructure

Compute

Block Storage

Multiple Regions

Success factors for hosting Grid infrastructure in NREC

Studies have been and are regularly carried out to estimate operational cost, so far most of the arguments are in favor of NREC

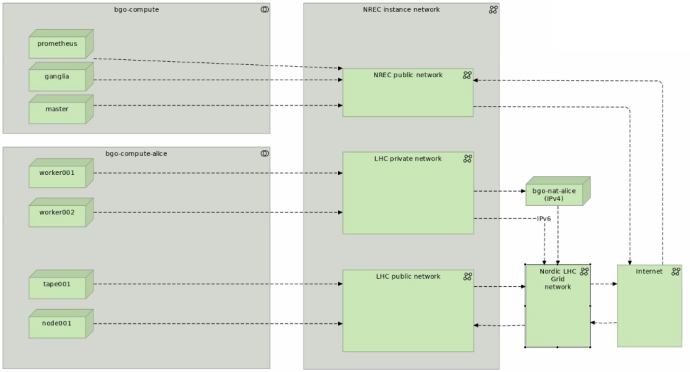
- Basic and stable minimal solution.
- Operational cost are lower
- No expensive (and unused) HPC components.
- Existing tape storage at IT departments computing centers, easy integration
- Well established communication, small team, agile spirit

Installed infrastructure for Norwegian NT1 contribution,

▷ funded by infrastructure project of Research Council of Norway

	Region	since		
25	Bergen	2019	Dell PowerEdge R7425	2x AMD 7551 Epyc
20	Bergen	2022	Dell PowerEdge R6525	2x AMD 7543 Epyc
15	Oslo	2020	Supermicro 2124 BT	2x AMD 7552 Epyc
4	Oslo	2020	Supermicro 2124 BT	2x AMD 7002 Epyc
24	Bergen	2019	Dell PowerEdge R740XD storage server	16x 10 TB SATA
25	Bergen	2022	Dell PowerEdge R740XD storage server	16x 12 TB SATA

NREC infrastructure BGO



- Compute workers and dCache pool nodes are running on dedicated hardware
- All other service instances, like VO Box, monitoring instances running on common hardware
- Connection to Nordic LHC network, all instances have IPv6

Storage servers and tape robot are in the same data center close to the compute servers and connected via dedicated 25 Gb/s network.

Administration of virtual instances

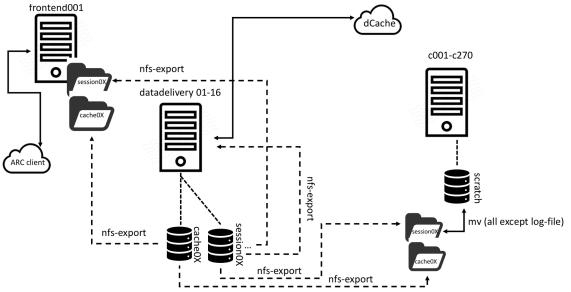
Very simple setup:

- The whole plan of the infrastructure is stored as code in Gitlab
- Virtual instances are provisioned using Terraform through the Openstack CLI
- Virtual instances are configured and updated via Ansible playbooks
  - sharing of infrastructure setup via code
  - make use of commenly available tools and playbooks, e.g. setup of SLURM, CVMFS, squid, prometheus exporters

Behind the scene: NREC administrators are creating dedicated flavors and make sure that instances end up on the correct hardware

Computing resources

- Computing resources are operated in both BGO and OSL regions
  - ▷ OSL 68 virtual instances providing 2170 cores
  - ▷ BGO 325 virtual instances providing 5200 cores
- Both sites provide SLURM clusters to provide compute nodes to Grid middleware
- ALICE setup in Bergen is running JAliEn which is managing whole nodes
- ATLAS setup in Oslo is running ARC which also handles data staging more flexible



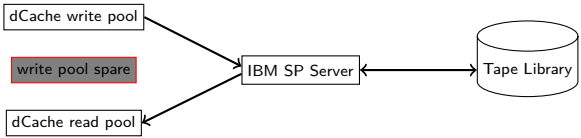
dCache Disk

The Nordic Tier 1 is operating a distributed dCache cluster

- Only dCache disk storage in Bergen is integrated into NREC
- ATLAS storage is outside NREC, more network traffic in and out NREC
- installed storage servers are managed as CEPH cluster
- NREC provides CEPH storage as block storage volumes
- Dedicated disk pool nodes for integration to dCache mount the CEPH volumes

dCache Tape storage

- Integrated into NREC in BGO region for the ALICE tape storage
- Dedicated 10 Gb/s link, write at 800 GB/s with two parallel drives
- 4 drives allow simultaneous mount
- 4 PB, upgraded by additional 6 PB by end of 2024
- Specific hypervisor servers to host dCache pool instances with 10 TB SSD cache
- 2 write pools, one active, one failover
- hot spare slot for new instance
- Management is interplay of site admin, dCache admin and IBM SpectrumProtect operator



Monitoring CPU utilization

- Slurm worker nodes are mostly allocated to JAliEn, but sometimes no matching jobs can be found
- Certain error conditions lead to JobAgents not properly starting, e.g. CVMFS outage, or something stuck due to a network outage
- we see fluctuating utilization of cores resulting in less CPU hours being effectively used for grid productions

▷ started a project to gather all relevant data from site monitoring, job traces and grid monitoring to understand job behaviour and optimize execution

Conclusions

- NREC is providing self-serviced e-Infrastructure in Norway
- NREC is hosting dedicated hardware for provisioning grid resources
- Cost-efficient operation of hardware funded by infrastructure projects of Research Council of Norway
- Using cloud technology perfectly separates different administration tasks
- Stable, secure and scalable infrastructure