



# HIP Site Report

Roope Jukkara<sup>2</sup>, Sami Lehti<sup>1</sup>, Tomas Lindén<sup>1</sup>, Mikael Myllymäki<sup>1</sup>,  
Ville Salmela<sup>2</sup>, Tommi Tervo<sup>2</sup>

Helsinki Institute of Physics<sup>1</sup>, CSC - IT Center for Science<sup>2</sup>



17.10.2023

HEPiX Fall 2023 Workshop, 16-20 October 2023, Victoria, Canada



Helsinki Institute of Physics (HIP) participates in the LHC experiments ALICE, CMS and TOTEM. HIP collaborates with CSC - IT Center for Science on providing WLCG resources. The ALICE resources are part of the Nordic distributed Tier-1 resource NDGF and the CMS resources form a CMS Tier-2 called T2\_FI\_HIP. The HIP dCache storage was recently upgraded and the raw capacity of the new storage is 6 760 TB, which is more than three times as large as the previous system. The new storage is located about 500 km north of the location of the previous system. This site report will mainly consist of the dCache storage upgrade.



# Contents



- 1 HIP and CSC
- 2 ALICE and CMS computing
- 3 dCache upgrade
- 4 LUMI integration for HEP usage
- 5 Funding application
- 6 Summary



The **Helsinki Institute of Physics** (HIP) is a physics research institute operated jointly by the

- University of Helsinki
- Aalto University
- University of Jyväskylä
- Lappeenranta-Lahti University of Technology
- Tampere University
- The Finnish Radiation and Nuclear Safety Authority (STUK) is a fixed-term interim member since 2018.

HIP participates in the LHC experiments ALICE, CMS and TOTEM.

**CSC** - IT Center for Science in Finland provides computational resources for research. The CSC headquarters are in Espoo in Southern Finland, but the main machine room is in Kajaani some 500 km north of Helsinki. CSC hosts and maintains the HIP ALICE and CMS dCache storage.



## ALICE computing in Finland is part of the **NDGF dTier-1**

- **CPU** 10.7 kHS06 2023 pledge at CSC, Kajaani, Open Stack cluster **alice-tron** running ALIEN on cPouta, CSC grant
- **Disk** 1.7 PB 2023 pledge at CSC, Kajaani

## The **HIP CMS Tier-2 project**

- Supports the LHC Run 3 Finnish computing needs
- Prepares for the HL-LHC computing developments and requirements
- Provides a WLCG Tier-2 site **T2\_FI\_HIP** for CMS and for the HIP CMS-group
  - **CPU** 25 kHS06 2023 pledge, **kale-cms**, OpenHPC, ARC-CE, in Viikki, Helsinki
  - **Disk** 2.7 PB 2023 pledge at CSC, Kajaani
  - **CMS services** in Kumpula, Helsinki

## Personnel:

- HIP: CMS T. Lindén, S. Lehti, M. Myllymäki
- CSC: dCache Roope Jukkara, Ville Salmela and Tommi Tervo
- CSC/NDGF: E. Edelman, ALICE coordinator



# dCache upgrade



- Old dCache: 8 HPE Apollo 4510 G9 4U servers, total raw size 2176 TB, in Espoo
- New dCache: Funded by Research Council of Finland FIRI grant
- The tendering process was run by the University of Helsinki
- Minimized total investment and operational costs by requiring a compact system
- The acquisition was based on the NDGF dCache pool recommendations
- The RCoF spending time limits and the worldwide delivery problem delayed the acquisition, the switch was delivered at the very end of 2022
- New dCache located in CSC Kajaani data center
- Remotely operated from Espoo by CSC personnel
- Multi Protocol Label Switching (MPLS) is used to reserve bandwidth for LHC-traffic from the shared network, to bypass firewalls and to connect to the Nordic LHC-network

## Disk Space Usage

CellName	DomainName	Total Space/MiB	Free Space/MiB	Precious Space/MiB	Layout (precious sticky free)
alice_hpc2n_uumu_se_r03	n-a38_hpc2n_uumu_se_Domain	13628750	0	0	
alice_hpc2n_uumu_se_w03	n-a40_hpc2n_uumu_se_Domain	13651488	13627913	0	
alice_hpc_ku_d8_w04	dcache09_hpc_ku_d8_2Domain	30408704	1228	0	
alice_hpc_ku_d8_w04	dcache08_hpc_ku_d8_1Domain	30408704	30408703	0	
alice_hpc_ku_d8_w05	dcache10_hpc_ku_d8_1Domain	9728000	9726290	0	
alice_uib_no_r01	tape001_dcache_ibc_uib_no_Domain	9437184	438	0	
alice_uib_no_w01	tape002_dcache_ibc_uib_no_Domain	9437184	9437184	0	
alice_uib_no_w02	tape003_dcache_ibc_uib_no_Domain	9437184	1447	0	
atlas_hpc2n_uumu_se_r03	n-a34_hpc2n_uumu_se_Domain	13579841	784	0	
atlas_hpc2n_uumu_se_w03	n-a36_hpc2n_uumu_se_Domain	13312800	4479	0	
atlas_hpc_ku_d8_w04	dcache08_hpc_ku_d8_2Domain	30408704	6002	0	
atlas_hpc_ku_d8_w04	dcache09_hpc_ku_d8_1Domain	30408704	30408703	0	
atlas_hpc_ku_d8_w05	dcache11_hpc_ku_d8_1Domain	9728000	10508	0	
atlas_nuc_lia_se_r03	rdglap006_serenity_nuc_lia_se_Domain	11264000	1705	0	
atlas_nuc_lia_se_w03	rdglap005_serenity_nuc_lia_se_Domain	13779000	2192	3616081	
atlas_uio_no_r01	dcache-tape02_grid_uio_no_Domain	10485760	4444	0	
atlas_uio_no_w01	dcache-tape01_grid_uio_no_Domain	11524446	0	0	
csc_f_041	arctic11_csc_f_1Domain	416015625	280708726	0	
csc_f_042	arctic12_csc_f_1Domain	416015625	2802380869	0	
csc_f_043	arctic13_csc_f_1Domain	416015625	2811063502	0	
csc_f_044	arctic14_csc_f_1Domain	416015625	281232984	0	
csc_f_045	arctic15_csc_f_1Domain	416015625	279769310	0	

Figure: The new ALICE pools as part of the NDGF dCache system.

Cell Info		Space Info		Request Info
Cell	Domain	Total Space (MiB)	Free Space (MiB)	Precious Space (...)
<input type="text" value="Cell"/>	<input type="text" value="Domain"/>			
hip_pool_011	hip-pool11_csc_f_Domain	396.7 TiB	364.3 TiB	32.4 TiB
hip_pool_012	hip-pool12_csc_f_Domain	396.7 TiB	364.5 TiB	32.2 TiB
hip_pool_013	hip-pool13_csc_f_Domain	396.7 TiB	364 TiB	32.7 TiB
hip_pool_015	hip-pool15_csc_f_Domain	396.7 TiB	363.9 TiB	32.8 TiB
hip_pool_014	hip-pool14_csc_f_Domain	396.7 TiB	364.3 TiB	32.5 TiB
hip_pool_016	hip-pool16_csc_f_Domain	396.7 TiB	364.2 TiB	32.5 TiB
hip_pool_017	hip-pool17_csc_f_Domain	396.7 TiB	363.4 TiB	33.4 TiB
hip_pool_018	hip-pool18_csc_f_Domain	396.7 TiB	364.2 TiB	32.5 TiB

Figure: The new CMS dCache pools.



# dCache upgrade



- Disk servers: 13 DELL PowerEdge R740xd2 2U, RAID 6, XFS
- dCache head nodes 2 DELL PowerEdge R6515 servers hosting VMs
- Raw total size 6760 TB, an increase by more than a factor of 3
- LAN 40 Gb/s  $\rightarrow$  2\*25 Gb/s
- WAN 40 Gb/s  $\rightarrow$  100 Gb/s
- OS is RHEL 9 with RedHat free (time limited) licenses for CERN use
- ALICE installation: Update NDGF Ansible scripts, install new pools, drain old pools, migrate to new pools
- CMS installation was done from scratch for a clean new configuration
- CMS installation is currently on dCache v8.2.33
- Protocols: WebDAV and XRootD
- Tokens are supported on the new system
- The old system uses about 6 kW of power and the new one 5-6 kW
- The old dCache will be moved to Kumpula for interactive usage



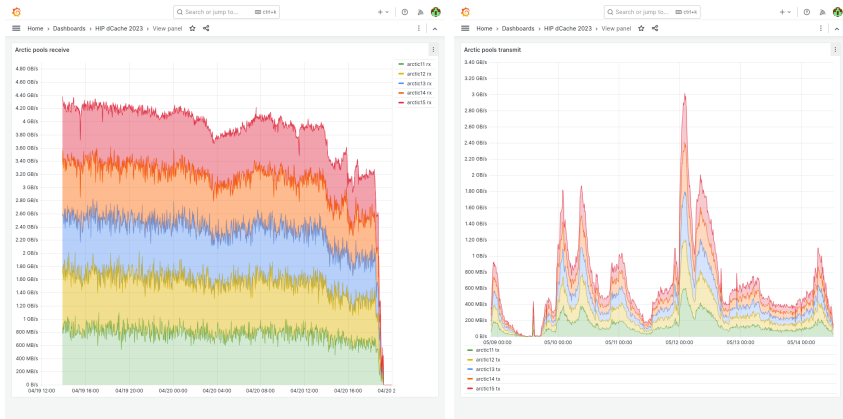


Figure: The five ALICE pools have reached **4,2 GB/s** writing and **3 GB/s** reading.

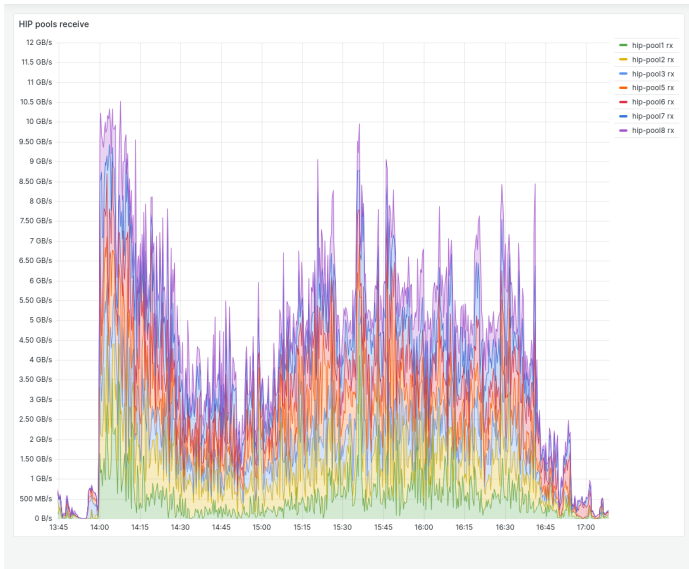


Figure: The eight CMS pools have reached more than **10 GB/s** writing.



Figure: The CMS HammerCloud jobs CPU efficiency for a  $\approx 10$  km CE-SE distance.

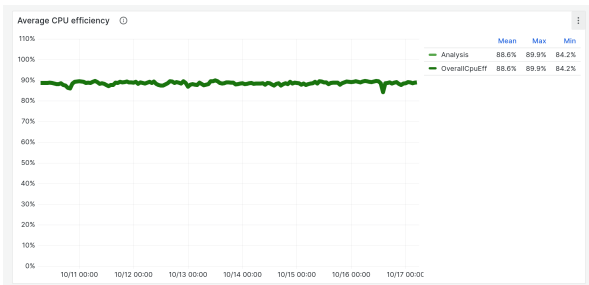


Figure: The CMS HammerCloud jobs CPU efficiency for a  $\approx 500$  km CE-SE distance.



# LUMI integration for HEP usage



**LUMI** (#3 on Top 500 list) HEP challenges: lacks local disk and CVMFS

We have three LUMI related projects:

- NLCG task force Lumi pilot working on demonstrating ALICE, ATLAS, CMS and LHCb running on the LUMI supercomputer, Install CVMFS, ARC, and a local disk replacement on LUMI, Abdulrahman Azab, Tomas Lindén, Gianfranco Sciacca, Mattias Wadenstein
  - A demonstration setup is being built on cPouta
  - CSCS experiences with the similar HPE ALPS HPC system are applied
- Exploring the Use of AMD GPUs for High Performance Computing in the CMS Reconstruction, porting CMSSW CPU-code to the AMD MI250X GPU:s, Andrea Bocci, Eric Cano, Tomas Lindén, Abdulla Mohamed, Felice Pantaleo
  - Code porting is in progress, a fat CMSSW container has been built
- LUMI-as-a-Service, testing tools developed for HPC usage on LUMI, Tommaso Boccali, Tomas, Lindén, Dainele Spiga
  - An Aptainer container with CVMFS as non root has been created
  - First CMS Glidein jobs have been run manually on LUMI



## Funding application

- HIP participated in the 2023 FCCI RCoF FIRI application for CPU capacity
- Decisions will be given in December 2023 or January 2024



# Summary



- The new ALICE and CMS dCache storage has increased capacity, bandwidth and performance
- The new dCache storage will be enough for Run 3
- For HL-LHC a storage upgrade will be needed the next time
- Work is in progress to make use of the LUMI supercomputer and potentially other HPC systems
- An application to acquire new CPU capacity has been made to the RCoF