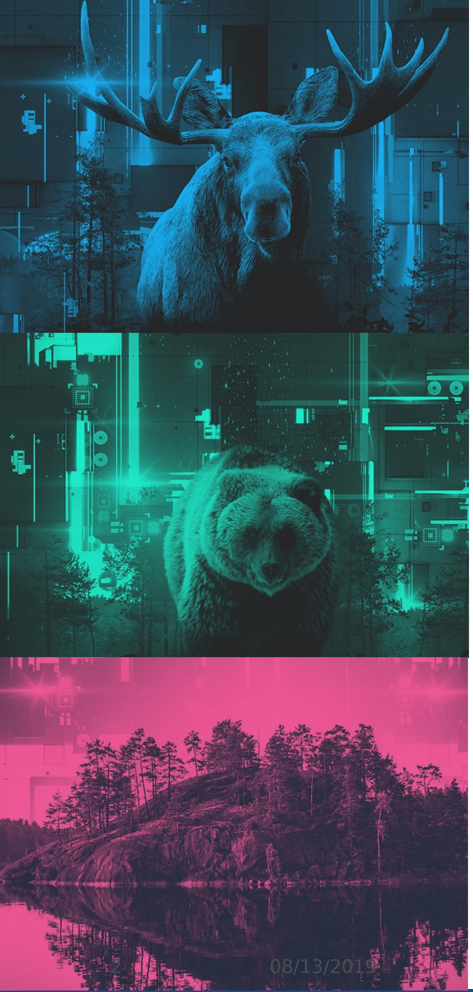
CSC

ICT Solutions for Brilliant Minds

New computing and data environment



What is CSC upcoming computing and data environment

- Something familiar… but much more powerful
  - **Puhti** - Supercomputer with Intel CPUs
- Something a bit different
  - **Puhti-ai** – Supercomputer with GPUs
  - **Mahti** - Supercomputer with AMD CPUs
- Something new
  - **Allas** – Large storage system with access and usage possibilities beyond traditional filesystem.

Scientific drivers for new infrastructure

New challenges

Large scale simulations

- For example climate change, space weather, fusion reactors, astronomical phenomena, particle physics

Mid-scale simulations

- For example materials science, energy technology, GIS

Data-intensive computing

- For example computational econometrics, bioinformatics, language research

Data-intensive computing for sensitive data

- For example medical research, register research

Artificial intelligence

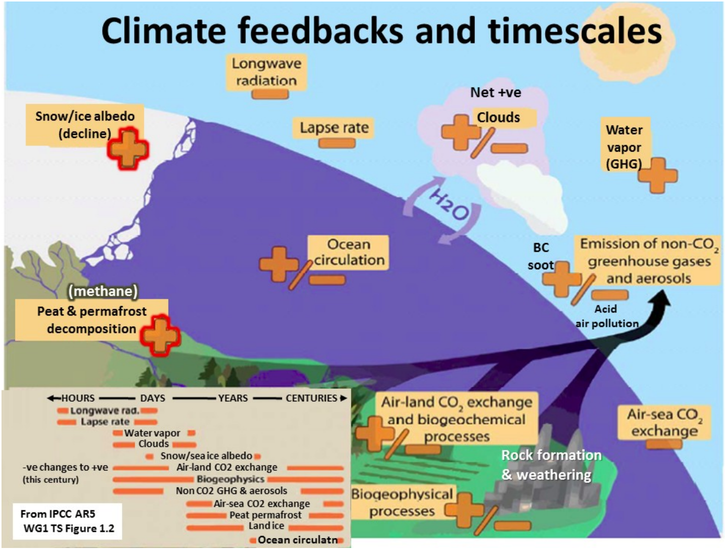
- For example natural language research, business applications, computer vision

Internet of Things (IoT) and data streams

- For example satellites, weather stations, sensor networks

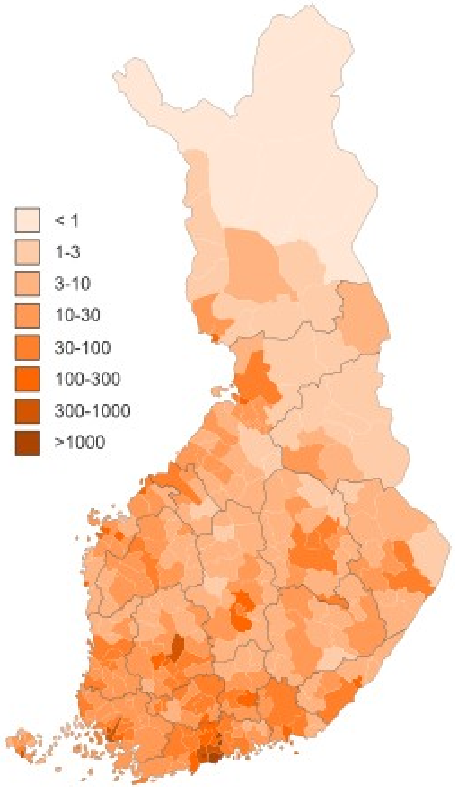
Real-world examples - medium-scale simulations & data-intensive computing & data streams

- Atmospheric feedback mechanisms
  - In study of climate change, understanding feedback mechanisms is crucial for because they may either amplify or diminish the effect. Therefore these are key for determining the climate sensitivity and future climate.
  - This research involves various environmental measurements, satellite data and multiscale modelling. The multiscale models research start from nanoscale (quantum chemistry), reaching out towards global atmospheric models.
  - PI: Prof. Markku Kulmala, University of Helsinki



Real-world examples - data-intensive computing with sensitive data

- The Sequencing Initiative Suomi
  - One of the largest human sequencing initiatives in Europe.
  - Whole-genome and whole-exome sequence data from Finnish samples can be combined with decades of data gathered in the Finnish health and welfare sector, enabling breakthroughs in personalized health care.
  - The researchers have already unraveled genetic components affecting growth and health of newborns, leading to new health recommendations, as well as repurposing of drug molecules to offer new hope for cancer patients.
  - PI: Prof. Aarno Palotie, University of Helsinki

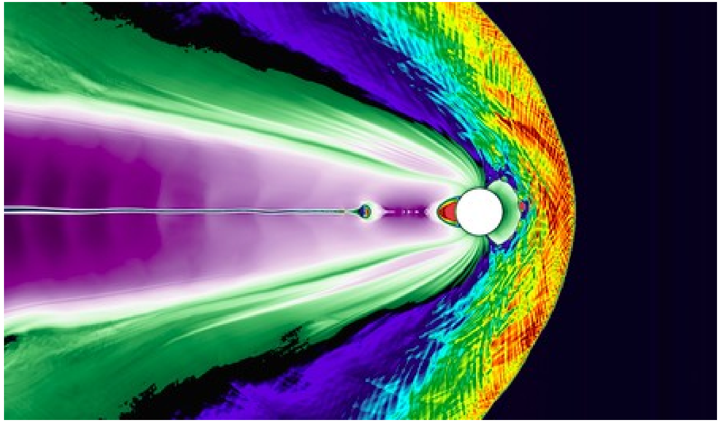


Real-world examples - data streams & data-intensive computing with sensitive data

- Whole-genome analysis of cancers and patients
  - Unravel the genetic components of human cancer susceptibility
  - Develop computational methods to fully benefit from the massive influx of data from multiple high-throughput, whole-genome scale experiments.
  - The goal is to translate these findings into clinical benefits, such as novel approaches in cancer risk prediction, prevention, diagnosis, and treatment.
  - PI: Prof. Lauri Aaltonen, University of Helsinki

Real-world examples - large-scale simulations

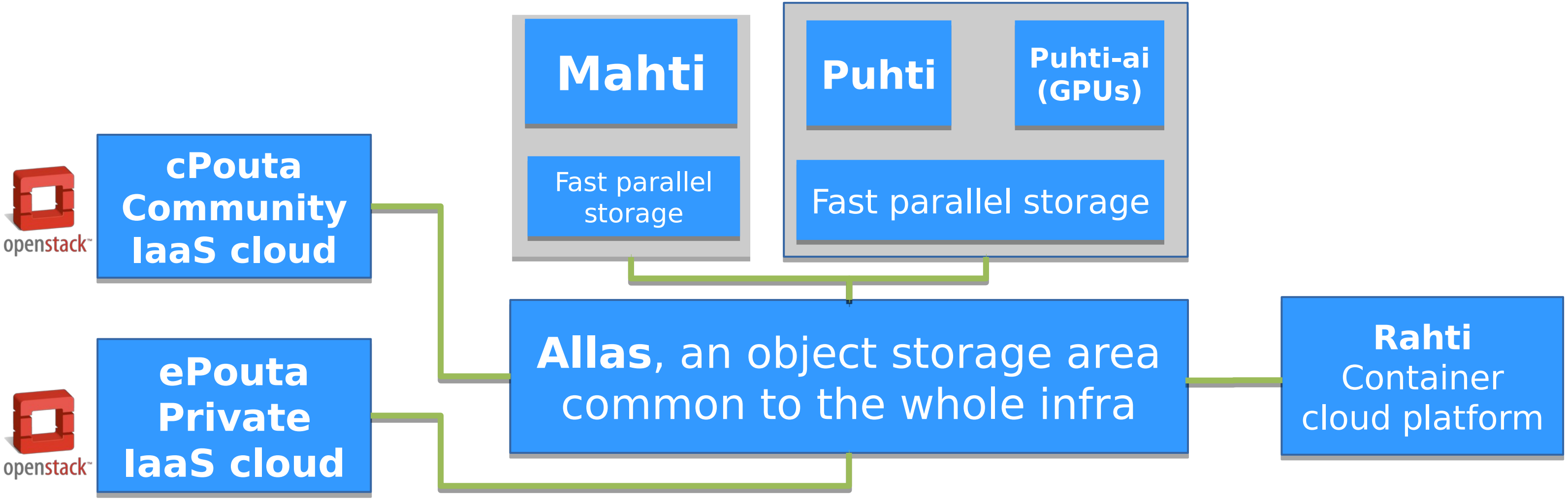
- World’s most accurate model on space weather, Vlasiator
  - Simulate the entire near-Earth space at a global scale using the 6D kinetic hybrid-Vlasov approach, to study fundamental plasma processes ((reconnection, particle acceleration, shocks) and gain a deeper understanding of space weather.
  - PI: Prof. Minna Palmroth, University of Helsinki



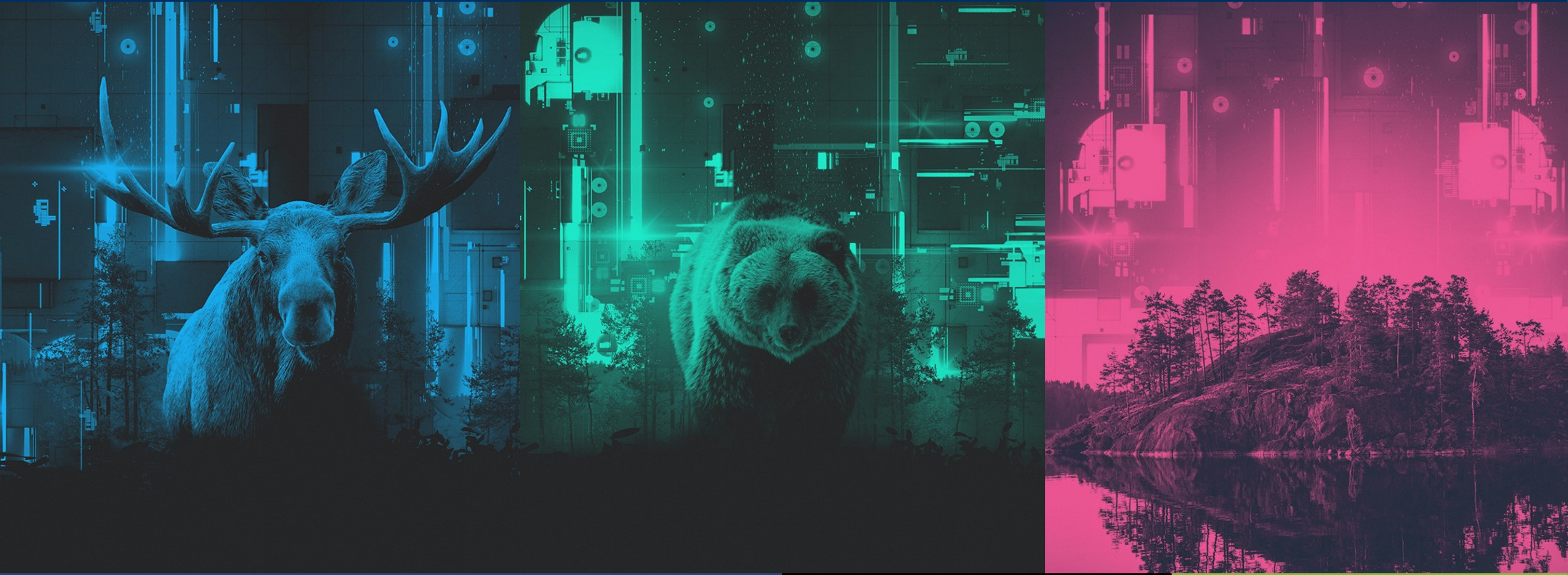
The new Finnish research infrastructure for data management and computing

Balanced HPC ecosystem for supporting the six drivers

Heterogeneous, workload-optimized node architecture, support for complex workflows, datasets-as-a-service and containerization



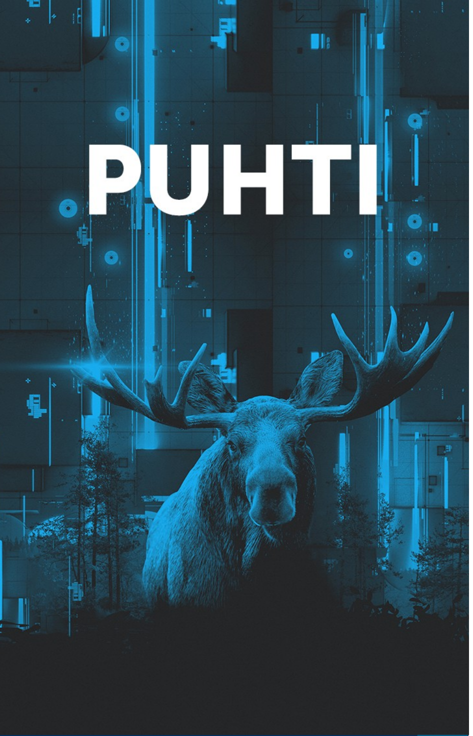
Technical details



In customer use end of Q1 / 2020

Mahti - Supercomputer

- 1404 compute nodes with next generation AMD Rome CPUs
  - Two 64 core CPUs per node
  - 256 GB of memory per node
  - About 180 000 cores in total
- Infiniband HDR interconnect between nodes
  - 200 GB/s bandwidth
- Especially for large scale simulations, but also for other use cases
- Over 8 petabytes of work disk for data under active use

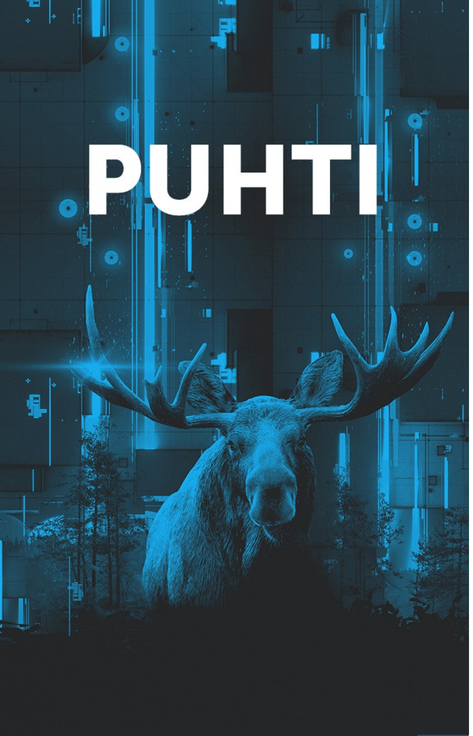


In customer use end of Q3 / 2019

Puhti - computing cluster

- New Intel Cascade lake CPU architecture
- 682 nodes: each having 2 Intel Xeon Gold 6230 2,1 GHz processors with 20 computing cores
- 40 cores in each node, 27280 cores in total.
- Floating point performance is 2.0 Petaflops
- Infiniband HDR interconnect between nodes
  - First machine in the world !
- 4 Petabytes work disk for data under active use
  - Shared with GPU partition
- Operational on August 2019

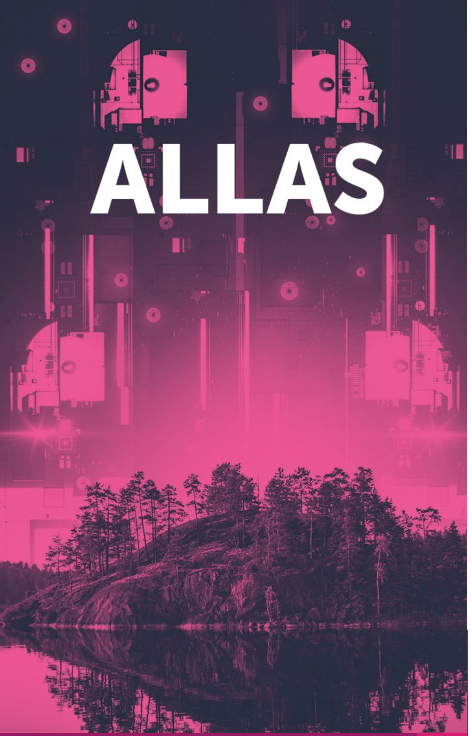
Node configuration in Puhti



In customer use end of Q3 / 2019

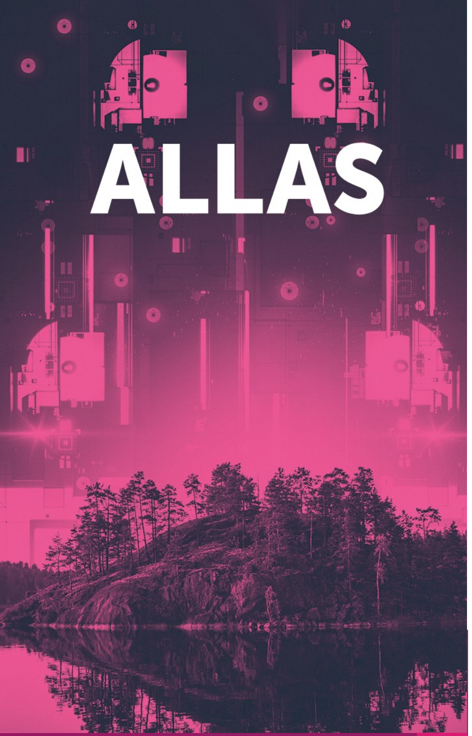
Node type	Memory	Number of nodes	Interconnect
Cascade Lake	192 GB	532	HDR100
Cascade Lake + 3.2 TB NVMe disk	384 GB	40	HDR100
Cascade Lake	384 GB	92	HDR100
Cascade Lake	768 GB	12	HDR100
Cascade Lake	1,5 TB	6	HDR100
AI partition nodes			
Cascade Lake + 4 Nvidia Volta V100 each with 32 GB of memory + 3.2 TB NVMe disk	384 GB	80	Dual rail HDR100

Allas – object storage: what it is for?



In customer use end of Q3 / 2019

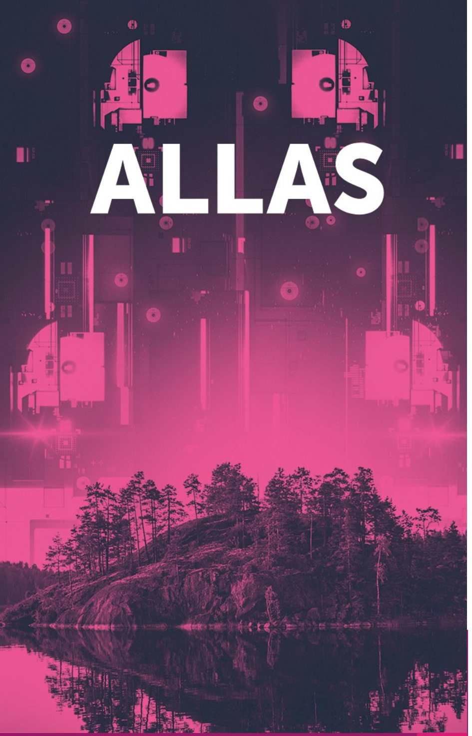
- Allas is new storage service for all computing and cloud services
  - Meant for data during project lifetime
  - Default quota 1 TB / Project.
  - Possible to upload data from personal laptops or organisational storage systems into Allas
  - Available in Taito, Puhti and Mahti
  - Data can also be shared via Internet



In customer use end of Q3 / 2019

Allas - storage

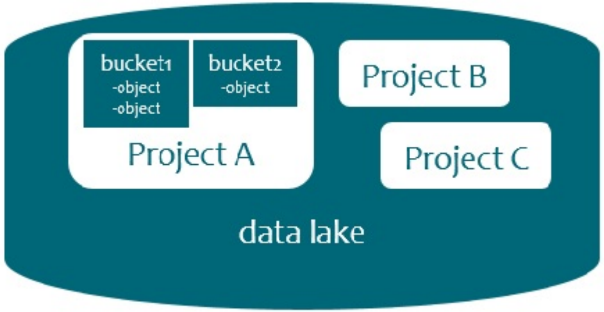
- **12 Petabytes** of storage space for data stored over the life-time of a project
- Default quota is 1 TB / project.
- Object storage based on open source CEPH
- An object is stored in multiple servers so a disk or server break does not cause data loss.
  - There is no backup i.e. if a file is deleted, it cannot be recovered
  - Data cannot be modified in the object storage – data is immutable.
- Rich set of data management features to be built on top of it, initially S3 and Swift APIs supported

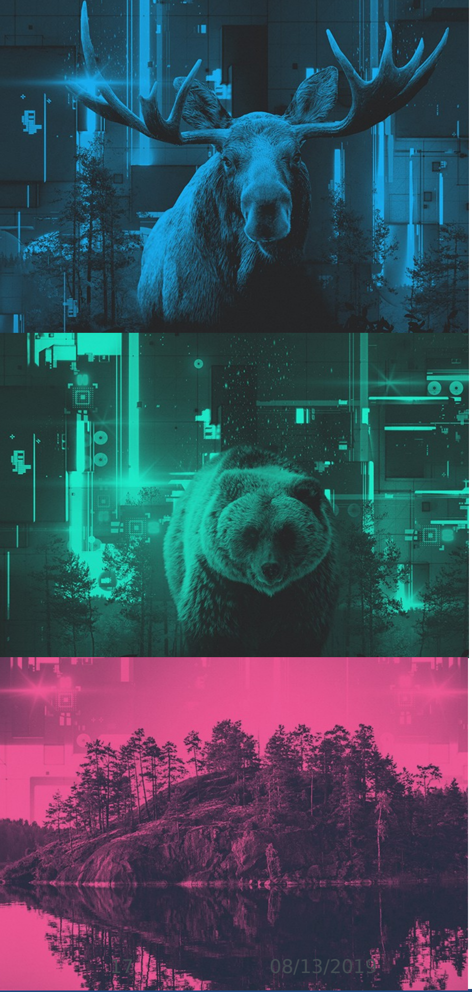


In customer use end of Q3 / 2019

Allas – object storage: terminology

- Storage space in Allas is provided **per CSC project**
- Project space can have multiple **buckets**
  - Only one level of hierarchy of buckets (no **buckets** within **buckets**)
- Data is stored as **objects** within a **bucket**
  - Blobs of data, can be anything (generally, **object** = file)
- Name of the **bucket** must be unique within Allas
- **Objects** can have metadata
  - Metadata are a key-value pairs, e.g. ”content=shoe”
  - Pseudofolders inside a **bucket** via metadata

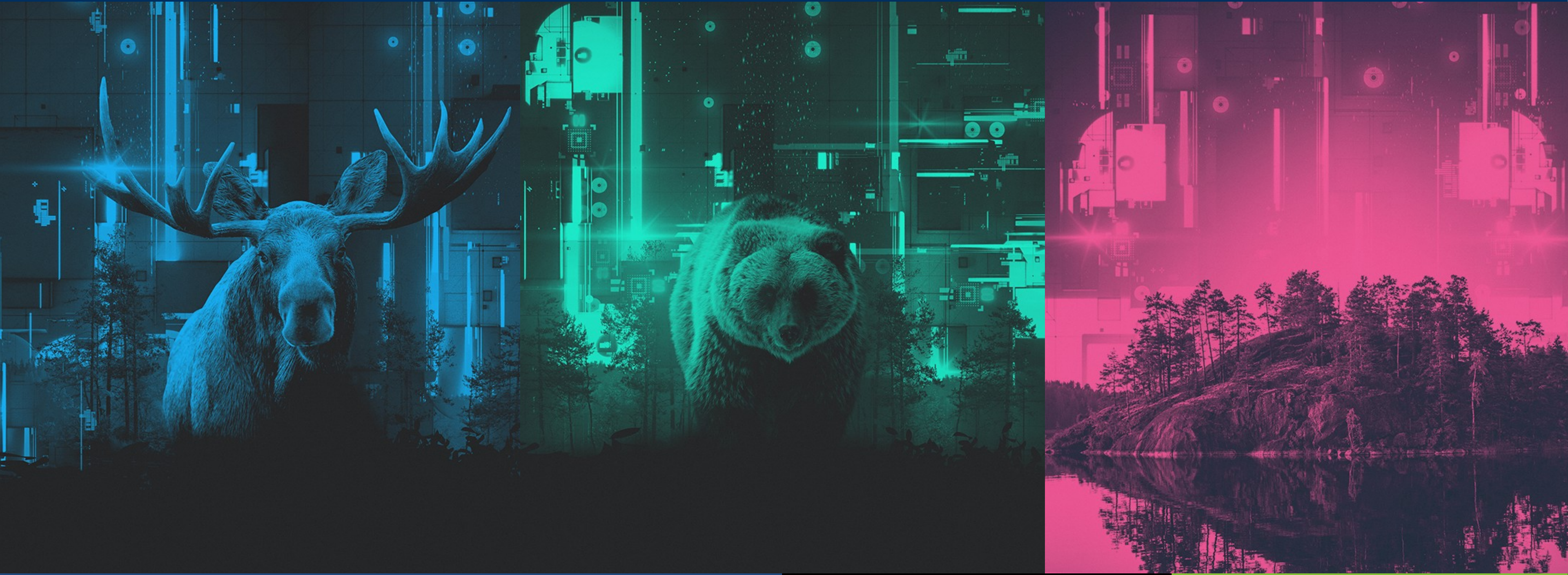


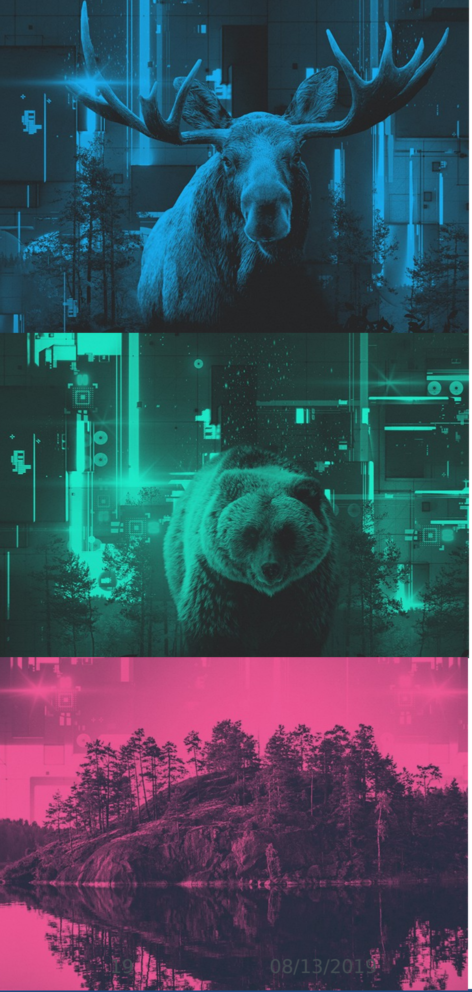


Installation and decommission schedule

- April – July 2019: Puhti and Allas installation and acceptance testing
- August 2019: Puhti and Allas available for customers
  - Both CPU and GPU partitions of Puhti
- August 2019: Sisu decommissioned
- Late 2019 – early 2020: Mahti installation and acceptance testing
- End of 2019: Taito decommissioned
- Early 2020: Mahti available for customers

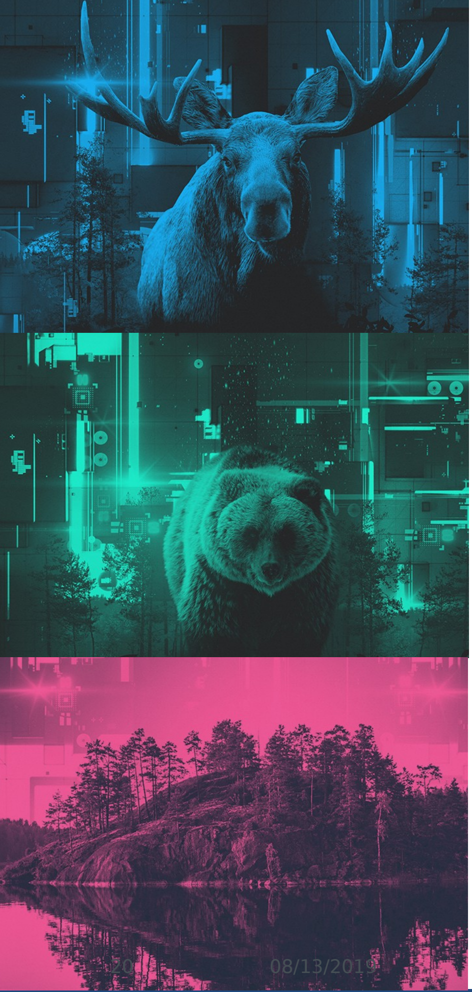
What changes in the new infrastructure?





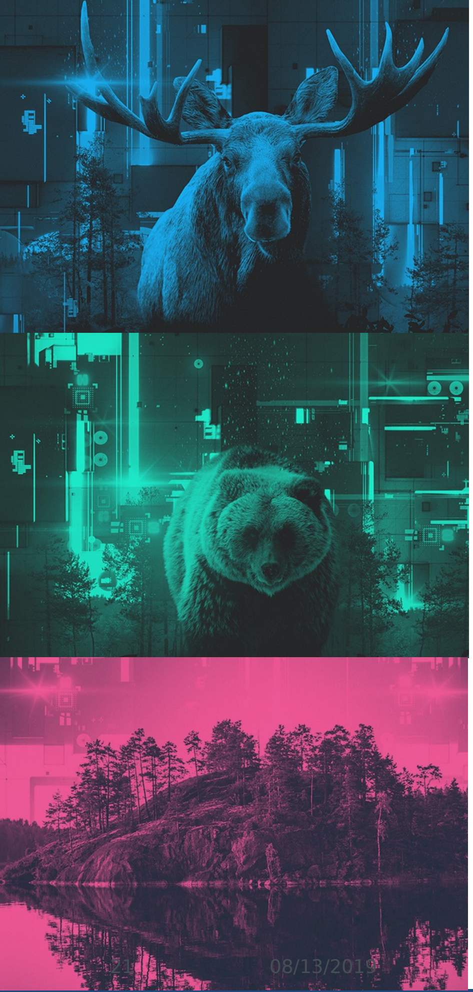
Access to new systems and data migration

- All users need to apply for new services via CSC customer portal
  - **Puhti** and **Allas** access can be already applied at my.csc.fi
- Users need to transfer their data from current $HOME, $WRKDIR and project directories to new infrastructure
  - Data should be first migrated to **Allas** and then to supercomputers when needed
  - Detailed instructions and guidance from CSC will be available
- Current data in Taito and Sisu remains available until end of 2019



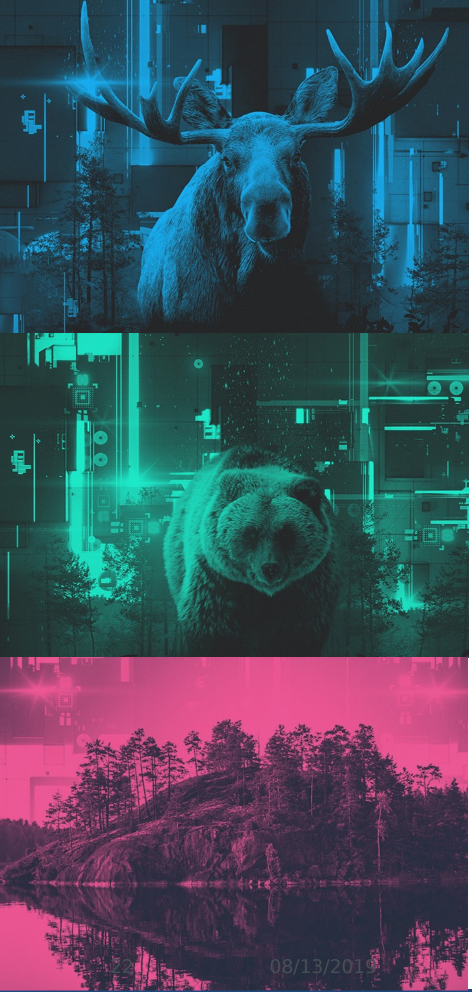
Storage in new infrastructure

- Allas is the common storage platform for all systems
  - project duration storage and sharing of data (3-5- years)
    - More long term storage in other services (CSC FAIR data, university services, …)
  - Data cannot be accessed from supercomputers with standard Unix commands (ls, cp, mv, etc.) but one needs to use object storage specific tools (Initially S3 and Swift interfaces).
  - For computations data needs to be typically copied to/from supercomputer with these tools
  - Data can be shared over the internet and accessed e.g. with web browser from local workstation
  - Storage is provided **per project,** default quota 1 TB
  - Used storage space **consumes billing units**
  - No backups



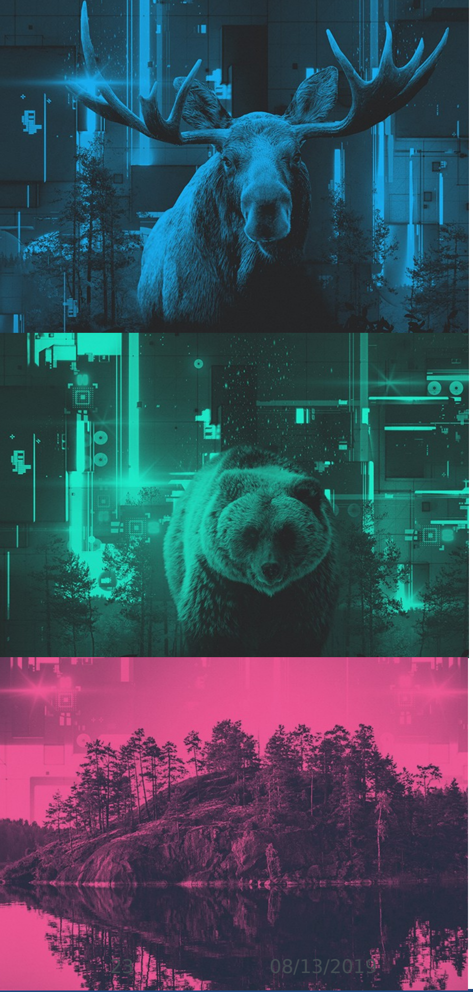
Storage in new infrastructure

- Puhti and Mahti have separate file systems with common structure
  - HOME: **User specific** directory for small data. Default quota 10 GB.
  - APPL: **Project specific** directory, for example for sharing projects own application codes. Default quota 50 GB.
  - SCRATCH: **Project specific** area for temporary data, i.e. intermitten simulation results. Default quota 1 TB
    - Similar to $WRKDIR in current systems (however, $WRKDIR was user specific)
    - Used storage space **consumes billing units**
    - Automatic cleaning: Files will be deleted 90 days from last access, relevant data should be moved to **Allas**
  - No backups in any of the storage areas



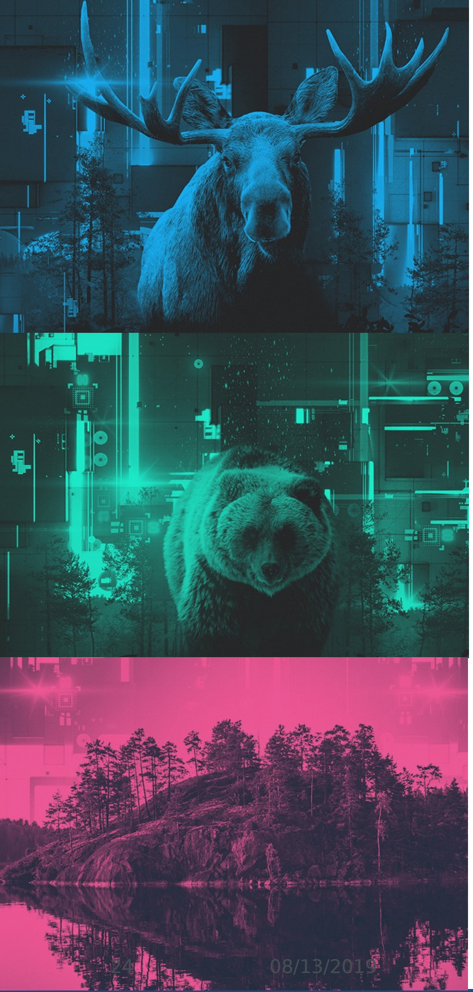
Using pre-installed applications

- Scientific software offering remains mostly the same as in Taito
- Similar **module** system as previously
  - i.e. **module load gromacs** or **module load r-env**
- Optimum runtime parameters (number of CPU cores etc.) most likely different from current systems



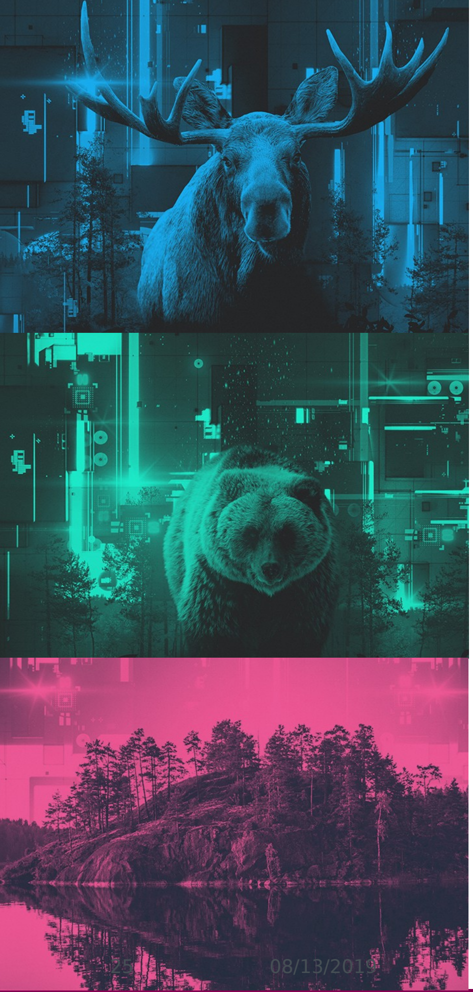
Batch jobs

- Similar SLURM batch job queuing system as previously
- New queues and policies (number and type of nodes, running times)
- Recommended to write new batch job scripts from scratch



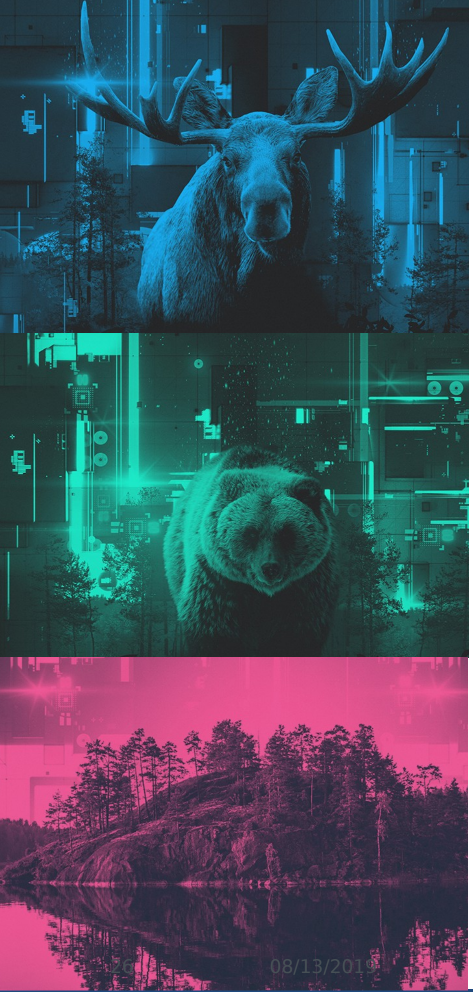
Software environment for building own applications

- New software stack
  - GNU and Intel compilers
  - Various high-performance libraries
  - HPC-X (OpenMPI based) and MPICH MPI libraries
- Applications should be rebuild
  - Configure scripts, Makefiles etc. may need modifications
- Applications should be installed in APPL disk area
  - Easier sharing for the whole project



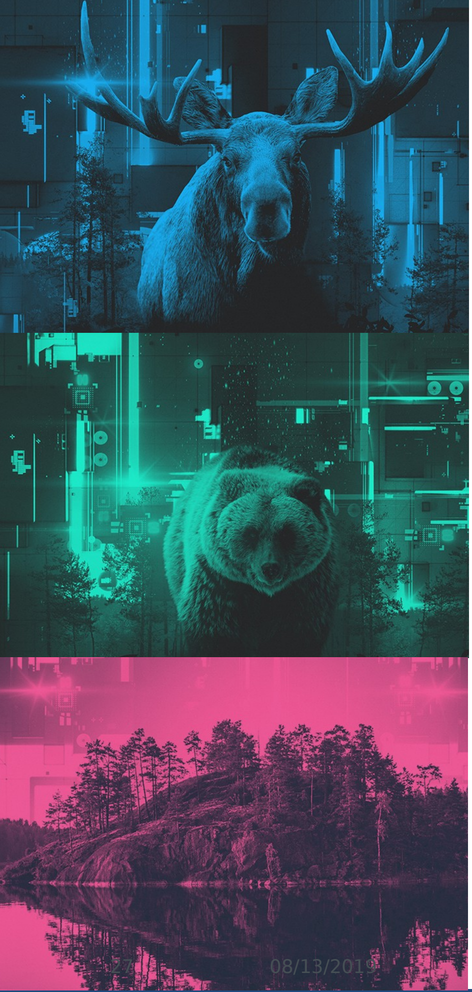
Developing applications by yourself

- New performance analysis and debugging tools
- Efficient use of Mahti might require further optimisation work e.g. hybrid OpenMP/MPI parallelization
- Own applications should be installed in APPL disk area
  - Easier sharing for the whole project



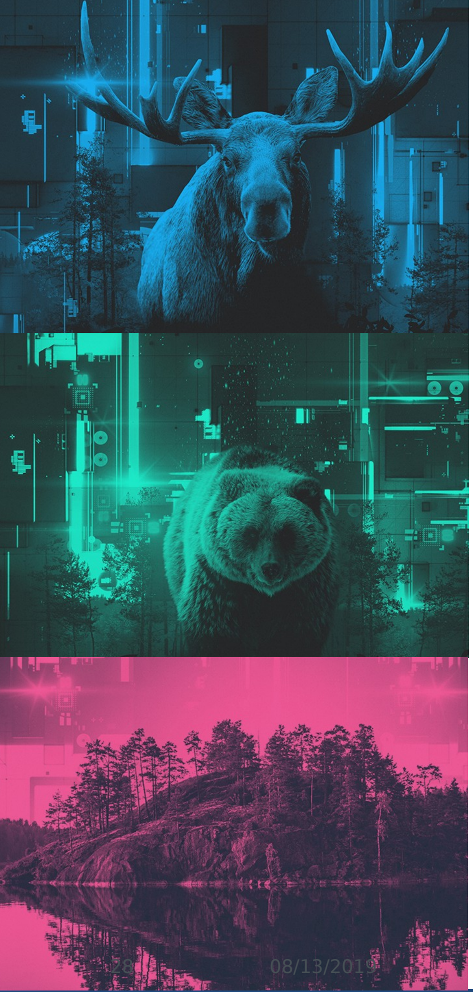
Training

- Several workshops and webinars about new infrastructure in 2019-2021
  - Both basic and advanced level
- First advanced porting and optimisation workshop for Puhti in June 2019.
- https://www.csc.fi/web/training/-/puhti_workshop_june_2019



Important dates for users

- August 2019: Access to **Puhti** and **Allas** open. Request access to **Puhti** and **Allas**
- August 2019: Sisu decommissioned
  - Data can still be accessed via Taito
- January 1st 2020: Taito decommissioned
- Early 2020: Access to **Mahti** open
- End of 2019: Current data in Taito and Sisu no longer available



Questions?

Up-to-date information about timetables, relevant changes for users etc. at

research.csc.fi/dl2021-utilization