Life Science Compute Cluster

Annual Report 2023

Life Science Compute Cluster
University of Vienna
Biology Building
Djerassiplatz 1
1030 Vienna
Austria

https://lisc.univie.ac.at
contact.lisc@univie.ac.at

Contact:
Thomas Rattei
Centre for Microbiology and Environmental Systems Science
+43 1 4277 91280

Overview

The Life Science Compute Cluster (LiSC) is a high-performance computing infrastructure for bioinformatics and computational life science.

According to the European model for High-Performance Computing, the LiSC is a Tier-2 cluster: it provides computing capacity that is available at a single research institution. The main difference to larger, Tier-1 computing facilities such as the Vienna Scientific Cluster (VSC), is the equipment with a rich, flexible and up-to-date bioinformatics software repository and the availability of major biological databases on-site. These installations allows typical users to analyse their data without any software installation, just by using the pre-installed tools and databases.

The Life Science Compute Cluster is friendly to non-specialists, and can be used for data analysis as well as method development and evaluation.

Three organisational units of the University of Vienna finance the Life Science Compute Cluster:

- Centre for Microbiology and Environmental Systems Science
- Faculty of Chemistry
- Faculty of Life Science

User accounts are free of charge. Accounts are available for students and employees of the University of Vienna and their scientific collaborators.

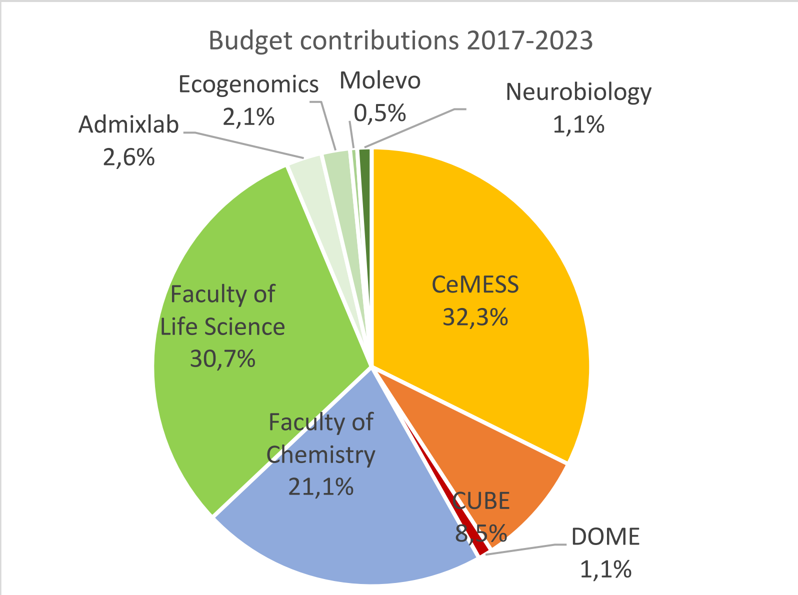
Main decisions for development and operation of the LiSC, such as the configuration of new hardware, management of compoutational workload and operation of file systems, are discussed and decided in bi-annual user meetings. The Division of Computational Systems Biology (CUBE) at the Centre for Microbiology and Environmental Systems Science (CeMESS) operates and administers the LiSC.

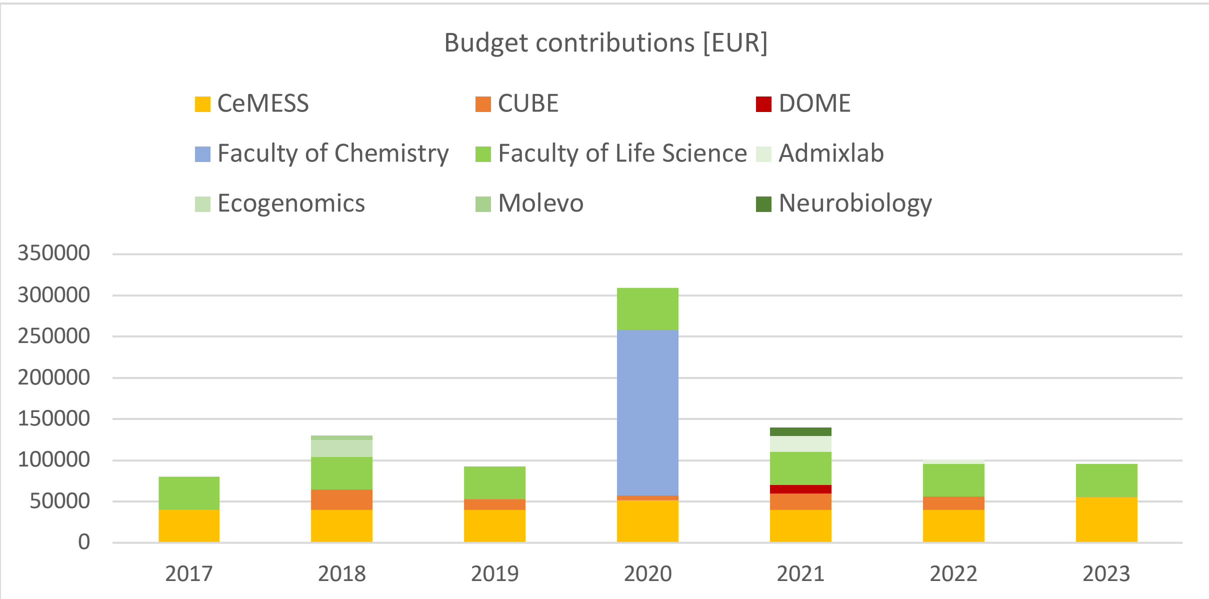
LiSC in 2023

Hardware budget

In 2023, the LiSC has received a hardware budget 95.000 EUR, consisting of two contributions of 55.000 EUR by the Centre for Microbiology and Environmental Systems Science and 40.000 EUR by the Faculty of Life Sciences, respectively.

From 2017-2023, all annual hardware budgets sum up to 306.600 EUR. They were mainly provided by the three involved faculties, and by additional contributions of workgroups with specific computational needs, mainly from the startup budgets of new groups.





Summary of LiSC hardware budget contributions 2017-2023

Personnel

The LiSC is scientifically and technically coordinated by Thomas Rattei, head of the Division of Computational Systems Biology (CUBE) at the Centre for Microbiology and Environmental Systems Science (CeMESS). The LiSC is operated by the CUBE system administrator (1 Full time equivalent technician position). In 2023, the 1 FTE of the CUBE system administrator was filled by two 20h/week positions:

- Johann Dorn (software and user support)
- Michael Neumayer (hardware and network)

Housing and power consumption

The LiSC is installed in 6 racks of the server room in the University Biology Building (UBB). Depending on the actual computational load, it consumes 20..30 kW electrical power directly, for network, storage and computing. The amount of additional energy for cooling by the central facilities of the UBB was not separately monitored in 2023. Costs for housing and energy consumption are covered by the University of Vienna and are not part of the LiSC hardware budget.

We aim at a minimal environmental footprint of the LiSC. We therefore consider energy efficiency as an important criterion for the selection of new hardware components.

Typical breakdown of the LiSC power consumption at medium computational load (from https://dashboard.lisc.univie.ac.at/)

Compute nodes and interactive nodes

At the end of 2023, the LiSC provided 79 compute nodes and two interactive nodes for file transfer, interactive computing, testing and management of compute jobs.

Compute nodes

Nodes	CPU cores	GPUs	RAM (GB)	Local diskspace (TB)	Number of nodes
A 01-30	16		128	5	29
B 01-16	20		314	5	16
C 01-02	40		1024	7	2
D 01-16	64		512	7	16
G 01-02	24	6	512	3	2
G 03-05	64	2	512	44	3
H 01-04	128		2048	11	4
H 05-10	128		2048	14	6
Total	2144	18	21952	494	79

Interactive nodes

Node type	CPU cores	GPUs	RAM (GB)	Local diskspace (TB)	Number of nodes
Login	64	1	512	44	2

Summary of LiSC compute nodes and interactive nodes

Virtual environment cluster

The highly available virtual environment cluster of LiSC consists of three nodes running Proxmox Virutal Environment (PVE) and two Fibre-Channel based Storage Area Network (SAN) fabrics that connect the three nodes with one shared RAID storage unit. In 2023, the LiSC hosted up to 35 virtual machines that provide scientific applications, databases, web servers and technical services. The virtual environment cluster is available to all LiSC users. It hosts several workgroup servers, which are used e.g. for development and testing of new bioinformatic resources. Further requests are welcome.

Host type	CPU cores	RAM (GB)	Diskspace (TB)	Number of systems
Proxmox VE	32	768	0,48 SSD, 1 HDD	3
Shared storage			20	1

Summary of nodes and storage of the LiSC virtual environment cluster

Network filesystems

LiSC provides five main network filesystems for:

- centrally installed applications (/lisc/app),
- user home directories (/lisc/user),
- recent project data (/lisc/project),
- archived project data (/lisc/archive),
- short-term intermediate (scratch) data (/lisc/scratch).

Volume name	Type	Diskspace (TB)	Max files (million)	Backup
/lisc/app	NAS (zfs)	30	1000	Daily
/lisc/user	NAS (zfs)	30	1000	Daily
/lisc/project	Parallel (beegfs)	400	100	Daily
/lisc/archive	Parallel (beegfs)	200	10	Daily
/lisc/scratch	Parallel (beegfs)	1100	1000	None

The two Network-attached storage (NAS) filesystems /lisc/app and /lisc/user are hosted by two identical nodes. In case of hardware failures, each node can run both filesystems after manual failover. Both filesystems are optimized for high availability and good performance for high numbers of small files.

Host type	CPU cores	RAM	HDD space	Number of nodes
nas 01-02	32	128G	27T	2

Summary of NAS storage nodes

Three parallel beegfs file systems host scientific project data and intermediate computing results. These filesystems are optimized for high (/lisc/project and /lisc/archive) and very high (/lisc/scratch) throughput and provide low latency for a medium number of files. Filesystem metadata are always stored on mirrored SSD or NVME drives. Filesystem storage targets are either located on external Fibrechannel RAID arrays, external Fibrechannel JBOD arrays, or internal RAID arrays.

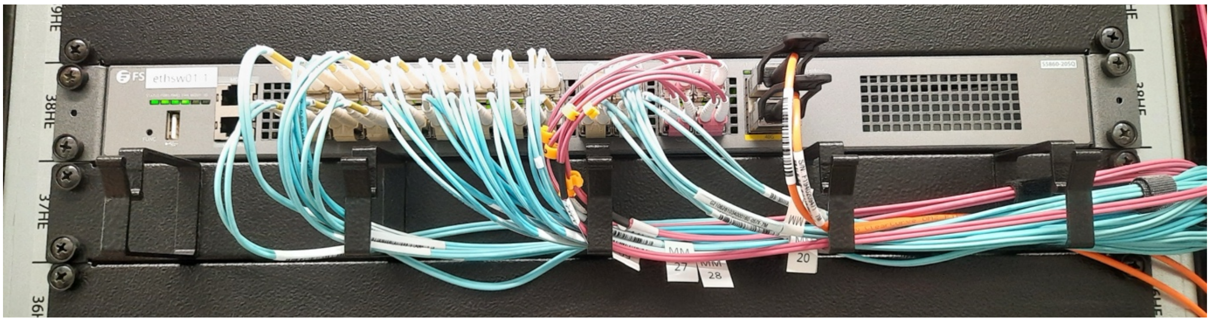
Host type	CPU cores	RAM	SSD/NVME space	HDD space	Number of nodes
fcs 01-04	20	64G	500G		4
beegfs 01-16	48	64G	360..760G	66T	16
FC RAID				200T	2
FC JBOD/ZFS				100T	2

Summary of beegfs storage nodes

Network connectivity

The connection of the LiSC to external networks is provided and managed by the Central IT services (ZID) of the University of Vienna. It comprises ethernet links with 1Gbit/s, 10Gbit/s and 25Gbit/s. The two interactive nodes and the virtual environment cluster nodes represent the main data transfer hubs of the LiSC. Each of these servers are connected to the University network through bonded 2x25Gbit/s ethernet ports. These provide up to 50Gbit/s data transfer rate per server on multiple connections.

The internal network infrastructure of the LiSC consists of an ethernet network and an infiniband network. The internal ethernet network consists of a firewall cluster with two nodes, a core-switch, with 2 stacked and therefore fully redundant switches and 8 access-switches. Each access-switch is connected to each of the 2 members of the core-switch stack with 2 redundant 10GbE-Links.



One of the chassis of the internal ethernet core switch

The internal infiniband network provides low latency and high throughput for the network attached storage and for the parallel filesystems. All systems are connected with 56Gb/s FDR dual lane cards, of which the compute nodes use one. The infiniband network uses a partial fat tree with two stem switches and 6 edge switches.

Infiniband networks need short cables and connect at the back of the switches

Scheduling of computational workload

All 79 compute nodes of the LiSC are controlled by the central workload scheduler slurm. The compute nodes are grouped into three slurm partitions:

- basic (compute nodes suitable for most jobs): node-[a-d]
- gpu (compute nodes that provide Graphical Processing Units - GPUs): node-g
- himem (compute nodes that provide 2TB of main memory): node-h

The configuration of slurm ensures fair use of the LiSC by user and group fair-share weights as well as user-based cpu*runtime limitations.

Centrally installed software

In 2023, the LiSC was running Oracle Linux 8 as operating system. Login and compute nodes were installed by using automatic deployment techniques and provide a software package selection suitable for workstation and software development usage scenarios.

Scientific applications are centrally installed in multiple versions using two typical frameworks for configuration and dependency management:

- environment modules
- conda environments

The LiSC team adds newly available versions of installed software automatically within one week. The LiSC encourages its users to request central installations of additional software and versions through the helpdesk, as central installation avoids redundancy and thus ensures reproducibility of computational analyses and workflows.

As an increasing number of workflows and applications are available as software containers (e.g. the widely used structure prediction software AlphaFold), the LiSC support the use of rootless docker and singularity containers.

Monitoring and response to failures

The operation of all LiSC hardware, network and software components is monitored by a cluster of two monitoring servers. These collect data into >41.000 items and indicate problems through >43.000 triggers. In case of problems, these triggers send alerts to the LiSC team. We solve most of the problems during our office hours (Mon-Fri 9-17). If severe problems are indicated, we try to solve them as soon as possible also outside the office hours.

Compute nodes as these node-d systems send operational data to the monitoring servers, which trigger alerts in case of problems

Users and groups in 2023

In the end of 2023, 419 users from 39 groups were using the LiSC. User groups include students and research groups from three faculties: Centre for Microbiology and Environmental Systems Science, Faculty of Life Science and Max Perutz Labs. Students are assigned to their research groups, as soon as they are supervised in a project, e.g. a lab rotation, MSc thesis or PhD thesis. The group "students" therefore contains students who participate in a course that uses the LiSC, but are not yet affiliated with a research group.

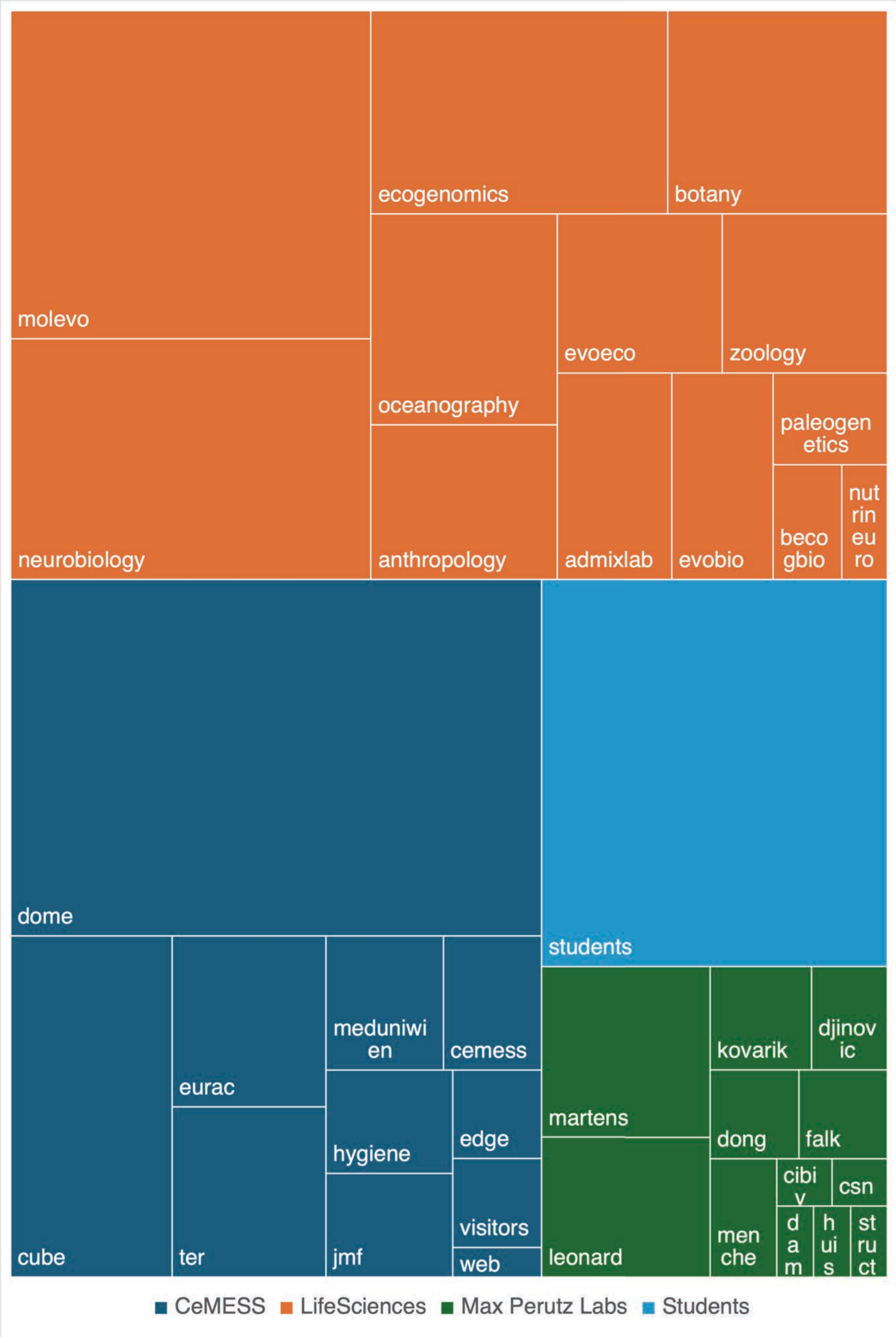
In 2023, 12 MSc courses have used the LiSC, including:

- 300362 UE Applied Data Analysis (2023S)
- 301224 SE Bioinformatics for Microbial Metagenomics (2023W)
- 301169 SE Applied Machine Learning for biological problems (2023W)
- 300108 UE Genome analysis for Ecologists (2023W)
- 300106 UE Analyses of single cell transcriptome data (2023W)

LiSC user accounts expire automatically, if not actively extended. The numbers mentioned above therefore represent users who actively use the LiSC or have used it in the same year.

Typical breakdown of user activity at LiSC on a Monday morning. The numbers include users who have been deleted during the last year.

(from https://dashboard.lisc.univie.ac.at/)



Treemap of LiSC users and groups by 2023. Group areas represent numbers of users.

Compute node utilization in 2023

Resource utilization and job waiting times

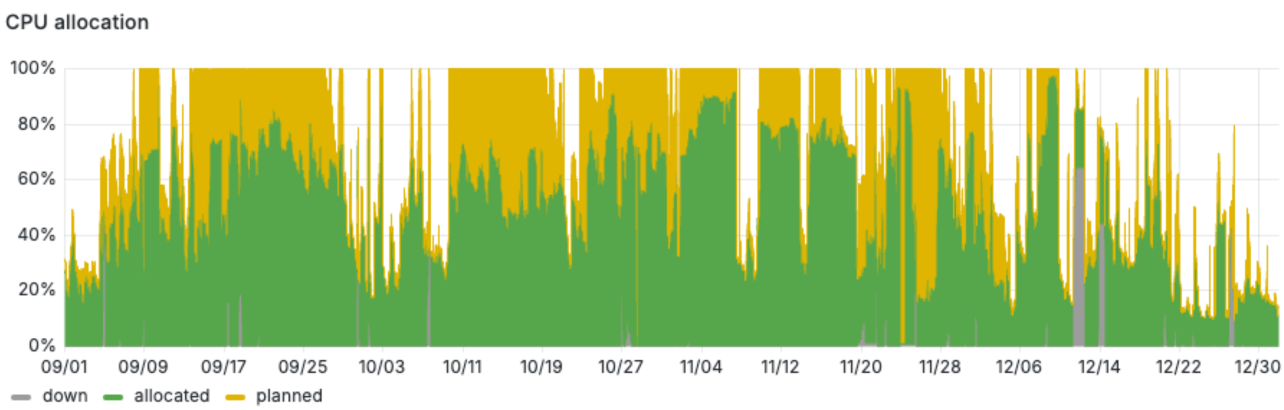
The LiSC monitors the usage of CPU cores, main memory (RAM) and GPUs. During the entire year 2023, the compute nodes were operational for 98.4% of the time. The average utilization of CPU, RAM and GPU was different in 2023:

- 48% of CPU
- 33% of RAM
- 70% of GPU

We have observed high variability in the compute node utilization, but only few periods of significantly increased job waiting times.

In general, the job waiting times for CPU jobs depend on the requirements of the job. In 2023, we have observed these average waiting times from job submission to job start:

- 50% of the jobs start within 1.4 hours
- 90% of the jobs start within 3.1 hours
- 99% of the jobs start within 9.2 hours
- Longest waiting time of a job was 21.7 days



CPU allocation during Q4/2023. This example demonstrates how much the utilization of compute nodes varies over time (from https://dashboard.lisc.univie.ac.at/)

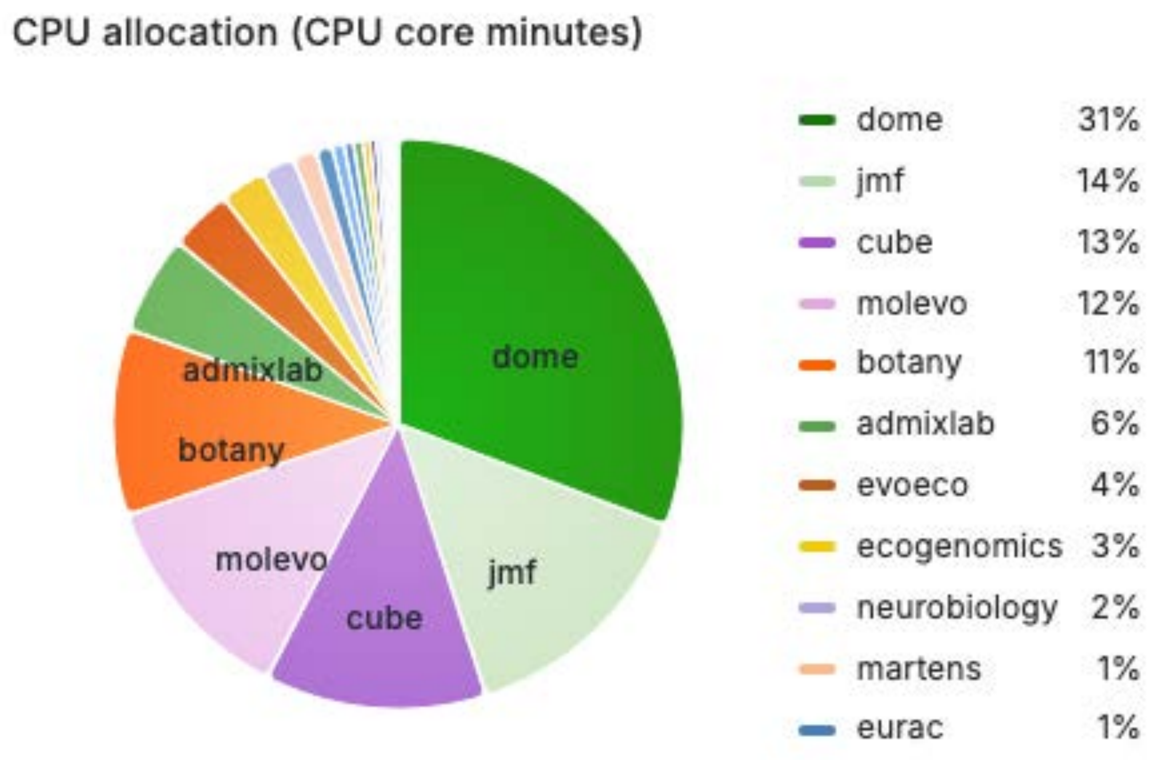
Waiting times for GPU jobs have sometimes been much longer than for CPU jobs, due to larger job series and the low number of GPU nodes.

Runtime of jobs and scheduling throughput

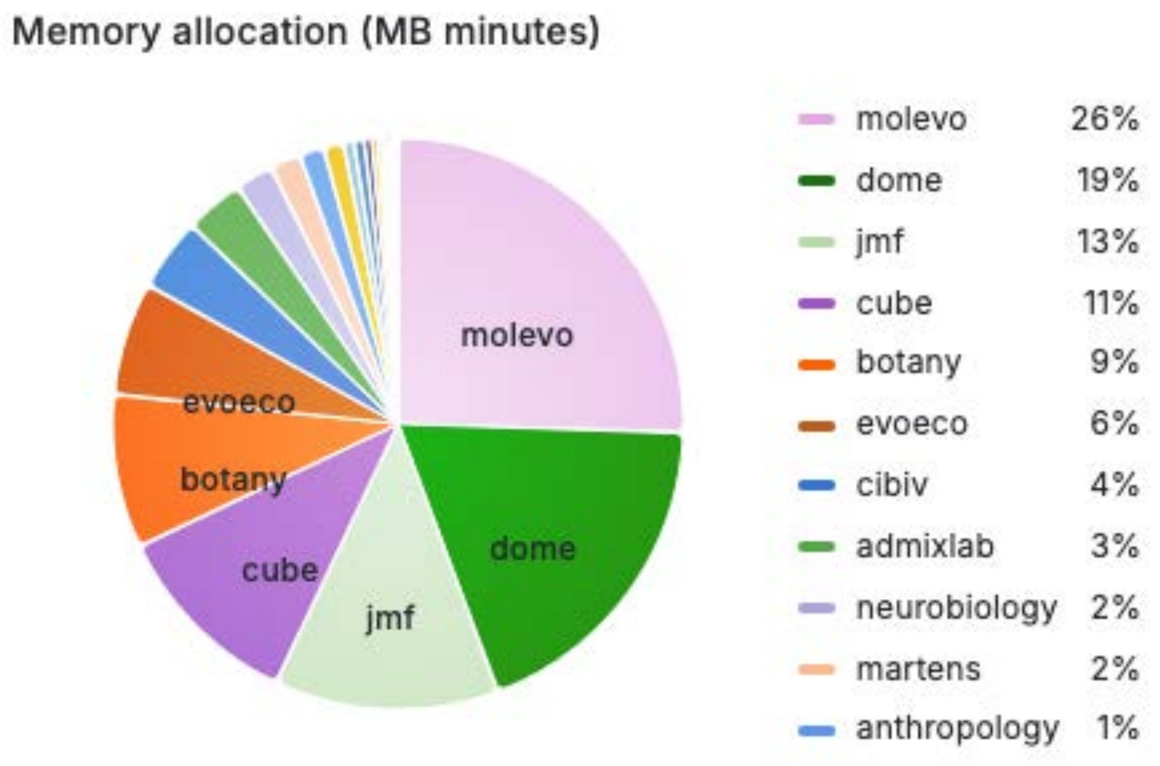
Runtimes of jobs range from minutes to weeks, with an optimum of 1 hour per job. Typically, the LiSC schedules 50 jobs per hour. However, we have observed a maximum of 61858 scheduled jobs per hour in 2023.

Usage of resources by group

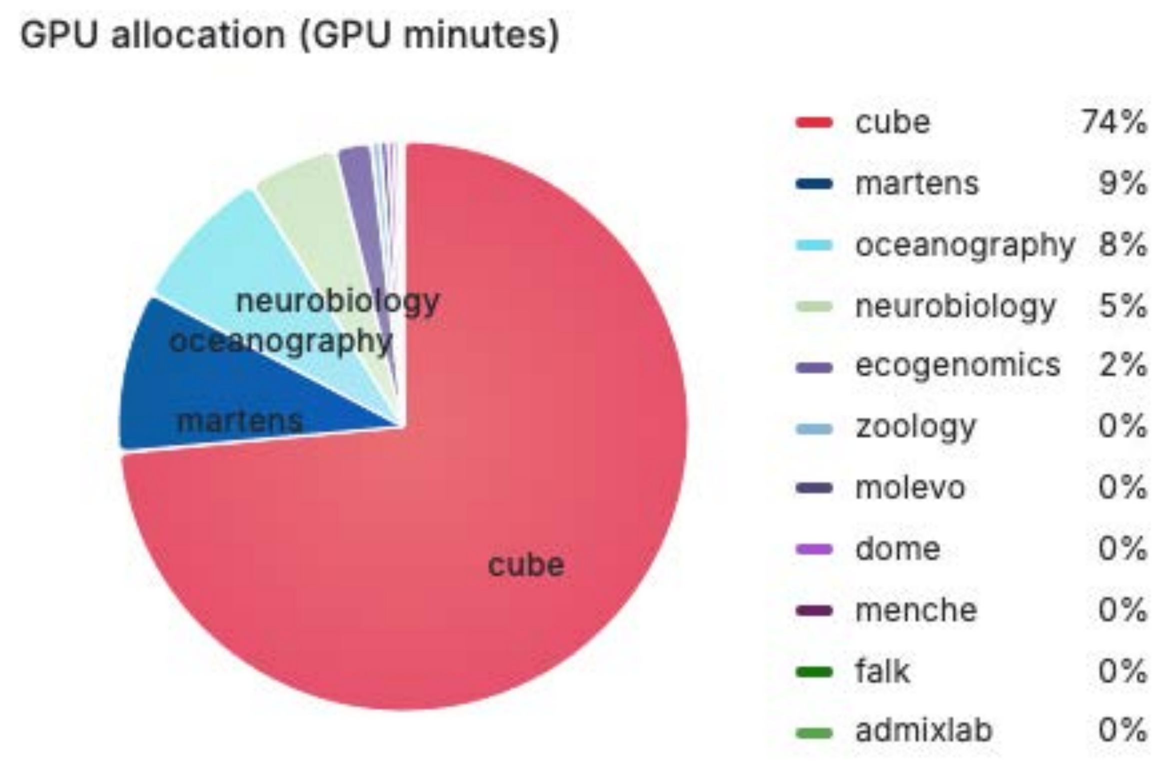
The usage of CPU, RAM and CPU varies by group (see figures below). This is due to different group size and because of different computational needs of the groups. Overall, the usage of LiSC corresponds well with the budgetary contributions of its stakeholders in 2023 (see p. 4).



CPU allocation by group in 2023 (from https://dashboard.lisc.univie.ac.at/)



RAM allocation by group in 2023 (from https://dashboard.lisc.univie.ac.at/)

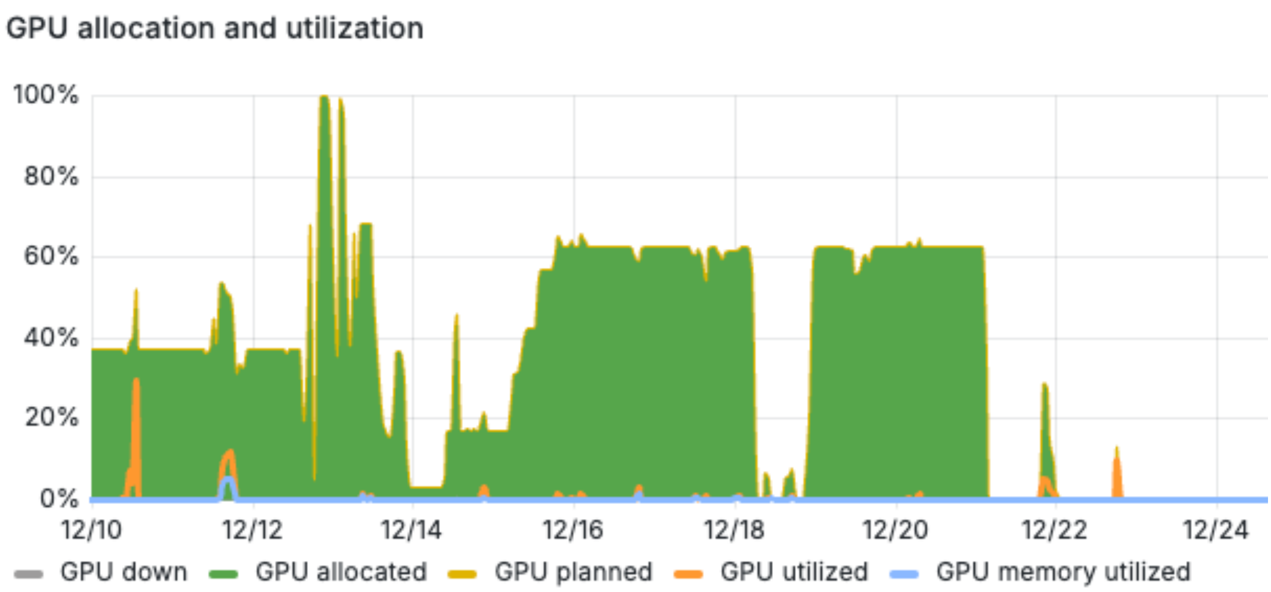


GPU allocation by group in 2023 (from https://dashboard.lisc.univie.ac.at/)

Scheduling and accounting of GPUs

Graphical processing units (GPU) are used by an increasing number of jobs. Typical examples are AlphaFold and other AI based tools. Most HPC infrastructures therefore need to increase the number of GPU nodes. The high costs for GPUs make detailed monitoring and efficient scheduling of GPUs essential for economic operation and growth of HPC systems. In 2023 we have improved the scheduling of GPUs as generic resources in two ways: exclusive use and shared use.

As the slurm workload scheduler does not provide actual GPU accounting data, we have developed a custom software for the monitoring and statistical accounting of GPU use in compute jobs. The GPU accounting resulted in valuable usage data. Most GPU-enabled software at LiSC use GPUs only for short periods during their runtime. A typical example is again Alphafold, which typically uses GPUs during ~1% of the job runtime only. The results overall indicated that most GPU-enabled software in Life Sciences requires a high CPU:GPU ratio and allows shared GPU use by default. Slurm at LiSC has been configured and users have been instructed accordingly.

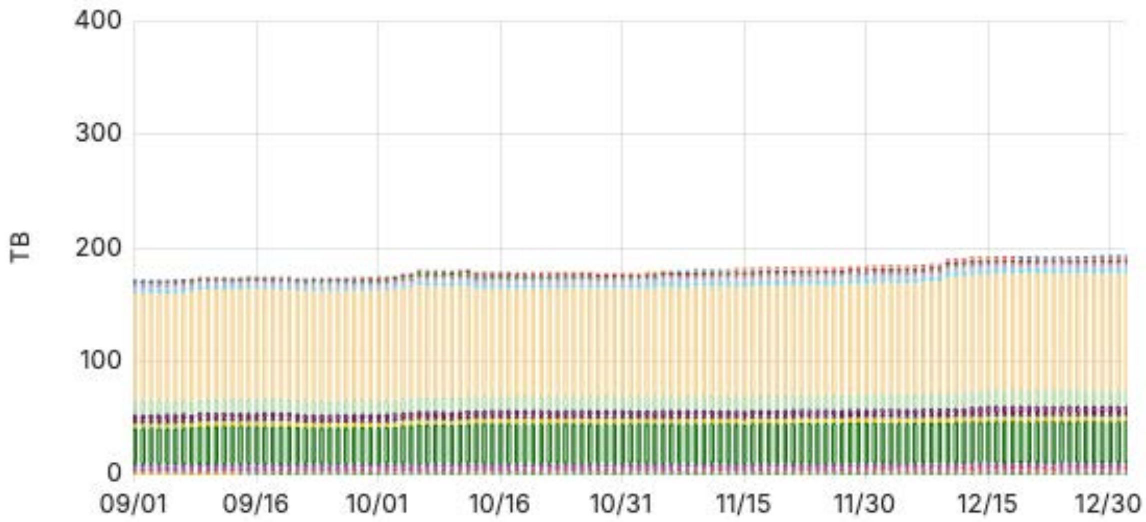


GPU allocation and actual utilization in December 2023. Most GPU-enabled software at LiSC use GPUs only for short periods during their runtime. (from https://dashboard.lisc.univie.ac.at/)

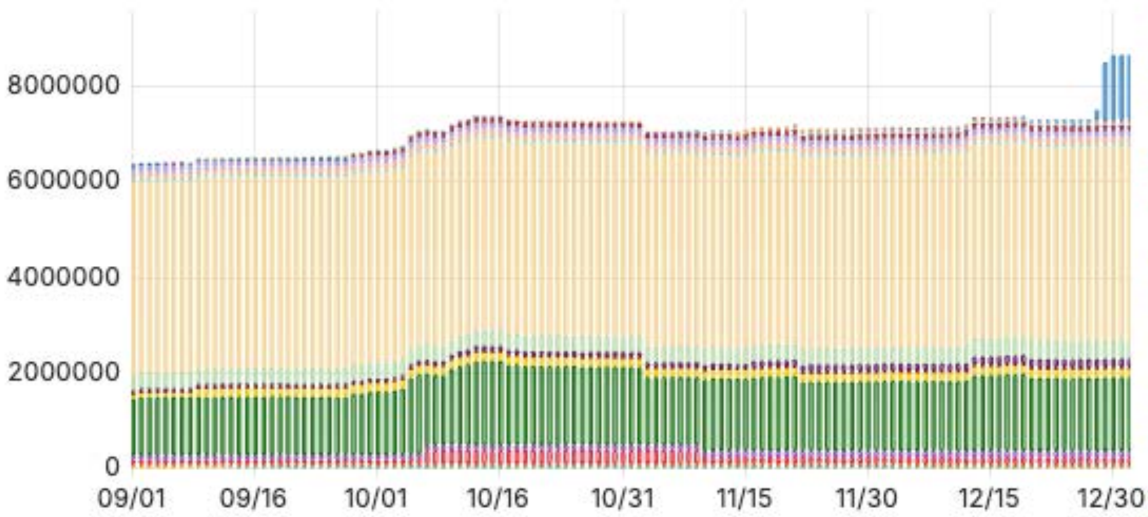
Data storage in large filesystems in 2023

Usage of /lisc/project

The parallel filesystem /lisc/project stores project data of recent projects and was about 50% filled in 2023.



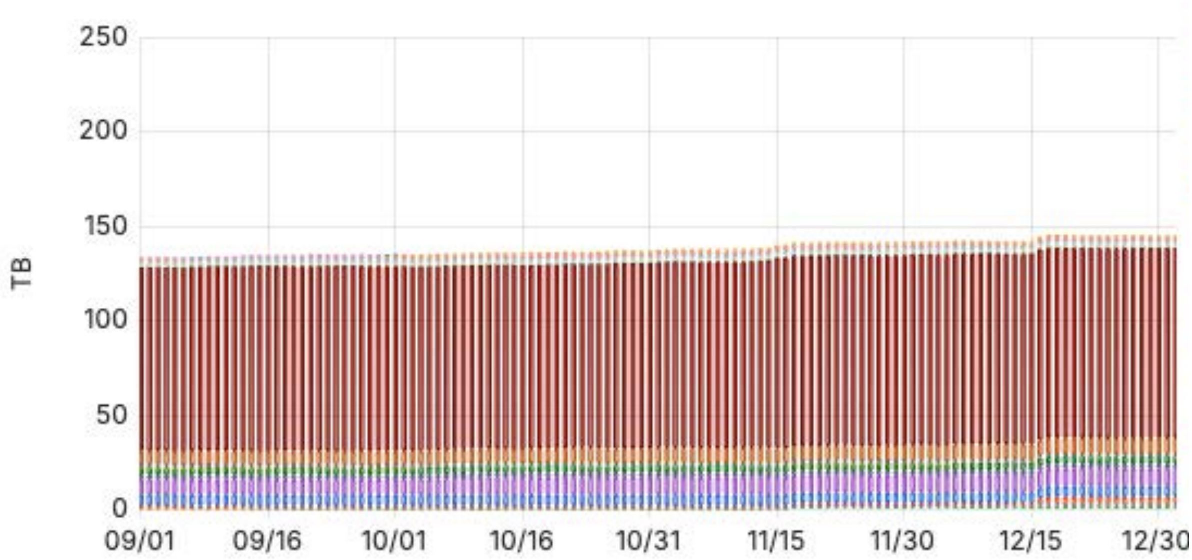
Diskspace usage in /lisc/project colored by group in Q4/2023 (from https://dashboard.lisc.univie.ac.at/)



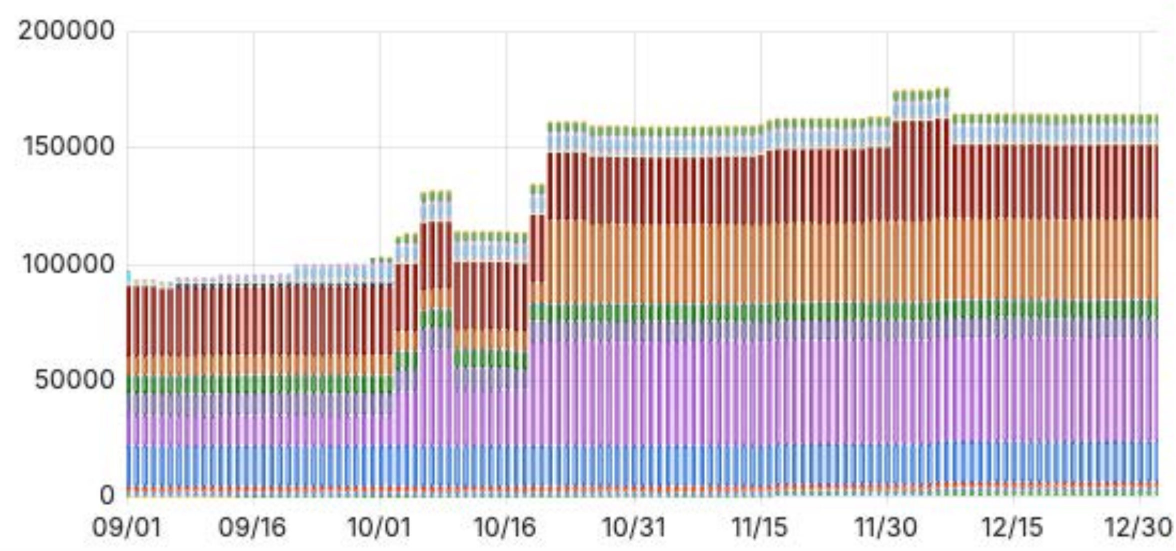
Number of files in /lisc/project colored by group in Q4/2023 (from https://dashboard.lisc.univie.ac.at/)

Usage of /lisc/archive

The parallel filesystem /lisc/archive stores project data of older projects and was about 60% filled in 2023.



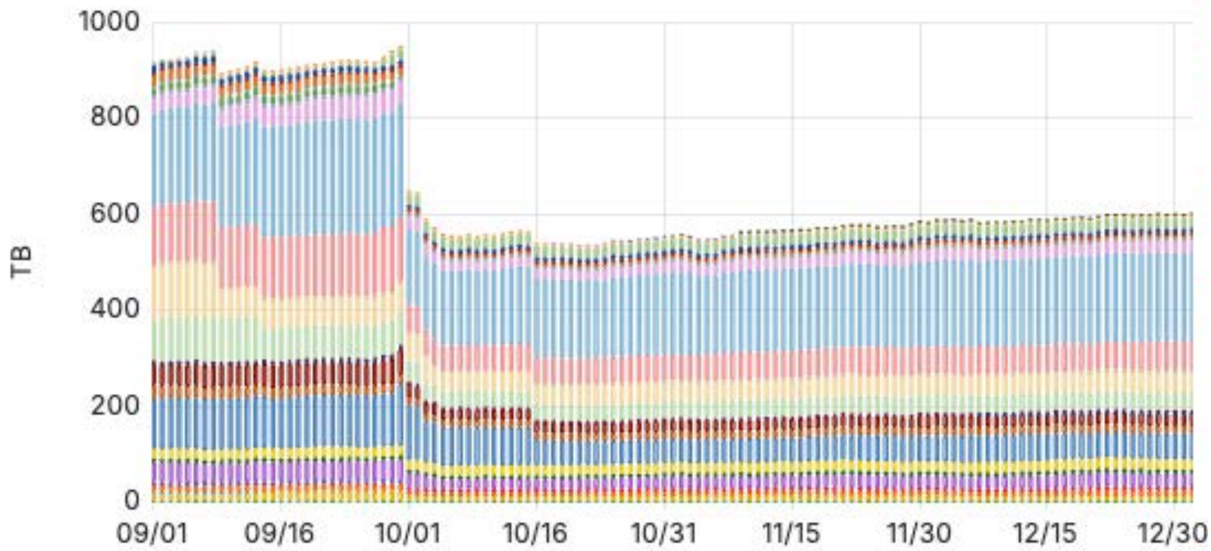
Diskspace usage in /lisc/archive colored by group in Q4/2023 (from https://dashboard.lisc.univie.ac.at/)



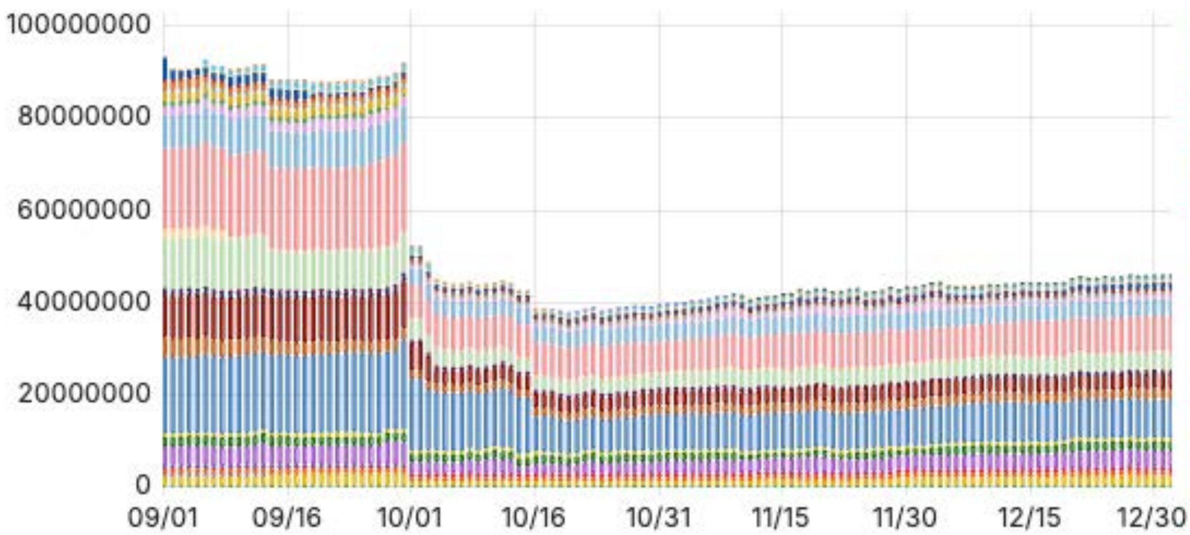
Number of files in /lisc/archive colored by group in Q4/2023 (from https://dashboard.lisc.univie.ac.at/)

Usage of /lisc/scratch

The parallel filesystem /lisc/scratch stores short-term intermediate data, e.g. the output of compute jobs. It was rebuilt in July-September 2023 and was about 60% filled in the end of 2023.



Diskspace usage in /lisc/scratch colored by group in Q4/2023 (from https://dashboard.lisc.univie.ac.at/)



Number of files in /lisc/scratch colored by group in Q4/2023 (from https://dashboard.lisc.univie.ac.at/)

LiSC user meetings, training and support in 2023

User meetings

The first user meeting in 2023 took place on May 5. About 25 users participated in person and online. We have presented and discussed the recent LiSC usage, onboarding of new user groups, changes on the user and job administration, the new helpdesk software, recommendations for long-running compute jobs, upcoming new infrastructure for data storage (DataLife project) and the LiSC hardware extensions planned for 2023.

The second user meeting in 2023 took place on October 11. About 20 users participated in person and online. We have presented and discussed new LiSC hardware 2023, node name changes and the hardware outlook for 2024, LiSC services for teaching and learning, LiSC dashboards, changes in slurm scheduling, best practice examples for compute jobs, and gave a status update on DataLife and LiSC storage.

LiSC training

The LiSC has offered several training activities to its users in 2023:

- 4 introduction and refresher webinars for new and returning users:
    - March 1, 2023
    - June 7, 2023
    - October 12, 2023
    - December 6, 2023
- Individual introductions for users of new groups ("onboarding" support for groups)
- Snakemake user and developer meeting as part of the CUBE lab meeting on June 16, 2023

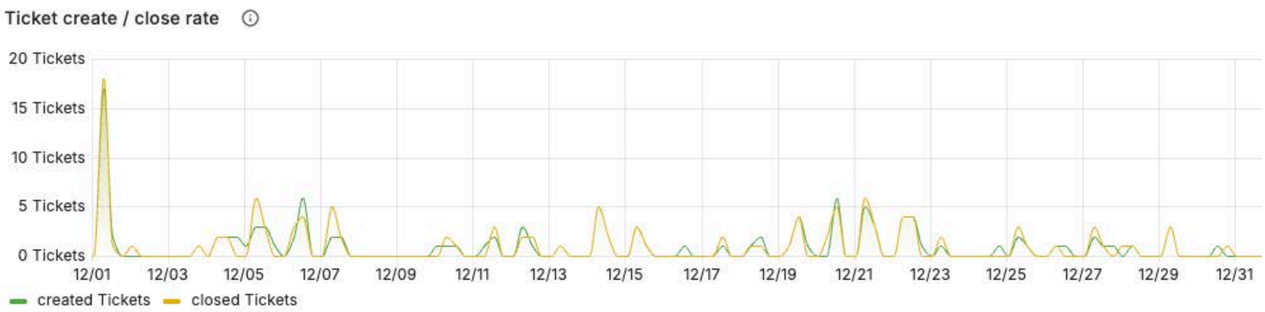
LiSC helpdesk

The LiSC helpdesk can be reached by email (contact.lisc@univie.ac.at) or by using the web form on http://lisc.univie.ac.at/helpdesk. In total, we have handled 928 requests in 2023. Top categories have been:

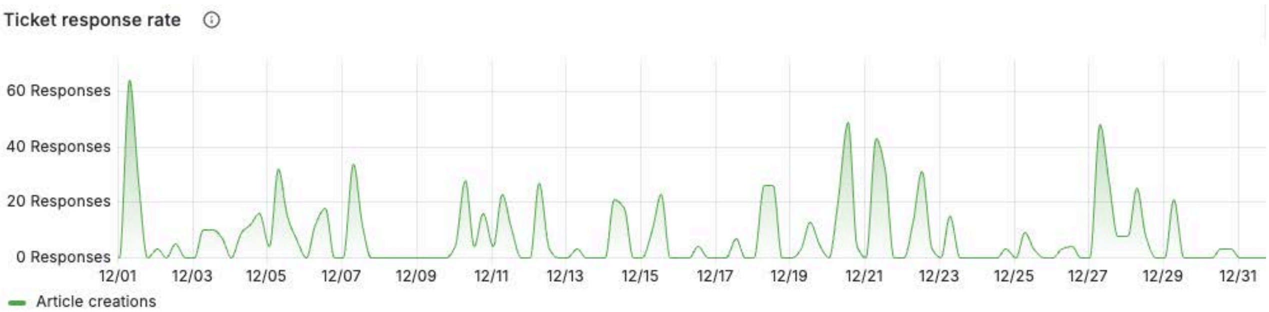
1. LiSC access requests
2. Slurm job time extensions
3. Software installation requests
4. Disk quota requests

The helpdesk receives about 5 requests per day, which require about 20 responses until solution. The maximum of 18 requests and 64 responses was on Dec 1, 2023.

Very most of the requests in 2023 could be solved at the same day, often within minutes or few hours. This is reflected in the close alignment of ticket create and close operations in the figure below.



Number of created and closed helpdesk tickets per day in December 2023 (from https://dashboard.lisc.univie.ac.at/)



Number of helpdesk responsed per day in December 2023 (from https://dashboard.lisc.univie.ac.at/)

Scientific contributions and impact in 2023

The LiSC is an essential infrastructure for research projects of the involved groups. A Google scholar search for "Life Science Compute Cluster" during the period 2023-2023 resulted in 25 publications and preprints.

Five selected highlights are:

- Huang Z et al. Three amphioxus reference genomes reveal gene and chromosome evolution of chordates. Proc Natl Acad Sci U S A. 2023 Mar 7;120(10):e2201504120. doi: 10.1073/pnas.2201504120.
- Pawar H et al. Ghost admixture in eastern gorillas. Nat Ecol Evol 2023 Sep;7(9):1503-1514. doi: 10.1038/s41559-023-02145-2.
- Dharamshi J E et al. Gene gain facilitated endosymbiotic evolution of Chlamydiae. Nat Microbiol 2023 Jan;8(1):40-54. doi: 10.1038/s41564-022-01284-9.
- Neukirchen S et al. Stepwise pathway for early evolutionary assembly of dissimilatory sulfite and sulfate reduction. ISME J 2023 Oct;17(10):1680-1692. doi: 10.1038/s41396-023-01477-y.
- Maire J et al. Colocalization and potential interactions of Endozoicomonas and chlamydiae in microbial aggregates of the coral Pocillopora acuta. Sci Adv 2023 May 19;9(20):eadg0773. doi: 10.1126/sciadv.adg0773.

Unfortunately, not all users acknowledge the LiSC in their publications. The number of 25 publications in 2023 is therefore only part of its scientific impact. Along with the DataLife project (see below), we will improve the monitoring of the scientific impact of our infrastructure.

In addition to peer-reviewed publications, we therefore measure the scientific contribution of LiSC by the number of students and PostDocs, who have been using the LiSC in 2023:

- 75 MSc students
- 164 PhD students
- 87 PostDocs

Outlook and plans for 2024

Hardware and software plans

In 2024, we will improve the GPU computing capacity of the LiSC. We envision two strategies to this end:

1. We will enable current CPU nodes for GPU dependent workflows, such as AlphaFold, by adding one GPU per node (probably Tesla T4). If the first user meeting in 2024 agrees, this will finally result in 18 GPU enabled basic compute nodes (node-c[01-02] and node-d[01-16]).
2. If the first user meeting in 2024 agrees, we will consolidate our existing 5 GPU nodes (node-g[01-05]) with 8 NVIDIA A30 into 4 dual-A30 nodes with NVLINK.
3. If the first user meeting in 2024 agrees, we will buy additional nodes with multiple GPUs from the latest generation.

We will upgrade the LiSC to Oracle Linux 9 in Q1/2024. This will provide newer versions of Linux kernel, system libraries and user environment.

We will test solutions for a more energy efficient operation of the compute nodes (e.g. automatic node shutdown and startup, controlled by the slurm workload scheduler).

Storage and research data management

Together with the Max Perutz Laboratories, we currently develop a new infrastructure for research data storage and processing. This project, DataLife, will establish about 10 PB of data storage. We will connect the storage with high bandwith to the analytical facilities and to computing clusters, including LiSC and Vienna Scientific Cluster (VSC). Most importantly, we will implement a new service and software framework for easy and efficient management of the data on DataLife storage. Our aim is that all data on DataLife will be FAIR (findable, accessible, interoperable, and reusable).

Links

LiSC homepage:

https://lisc.univie.ac.at

LiSC helpdesk:

https://lisc.univie.ac.at/helpdesk

LiSC dashboards:

https://dashboard.lisc.univie.ac.at
(Only accessible from the network of the University of Vienna)