



MetaCentrum NGI

For scientific computations, collaborative research & its support services

Jiří Vorel

vorel@cesnet.cz

March 4th, 2025
Brno

■ MetaCentrum is

- ... National Grid Infrastructure (NGI) <https://metacentrum.cz>
- ... the activity of the CESNET association <https://metavo.metacentrum.cz>
- ... a provider of computational resources, application tools (commercial and free/open source) and data storage <https://docs.metacentrum.cz>
- ... free of charge
 - Users "pay" by acknowledgement in their research publications

■ MetaCentrum is available for

- ... employees and students from Czech universities, the Czech Academy of Sciences, non-commercial research facilities, etc.
- ... industry users (non-profit and public research, upon individual request)

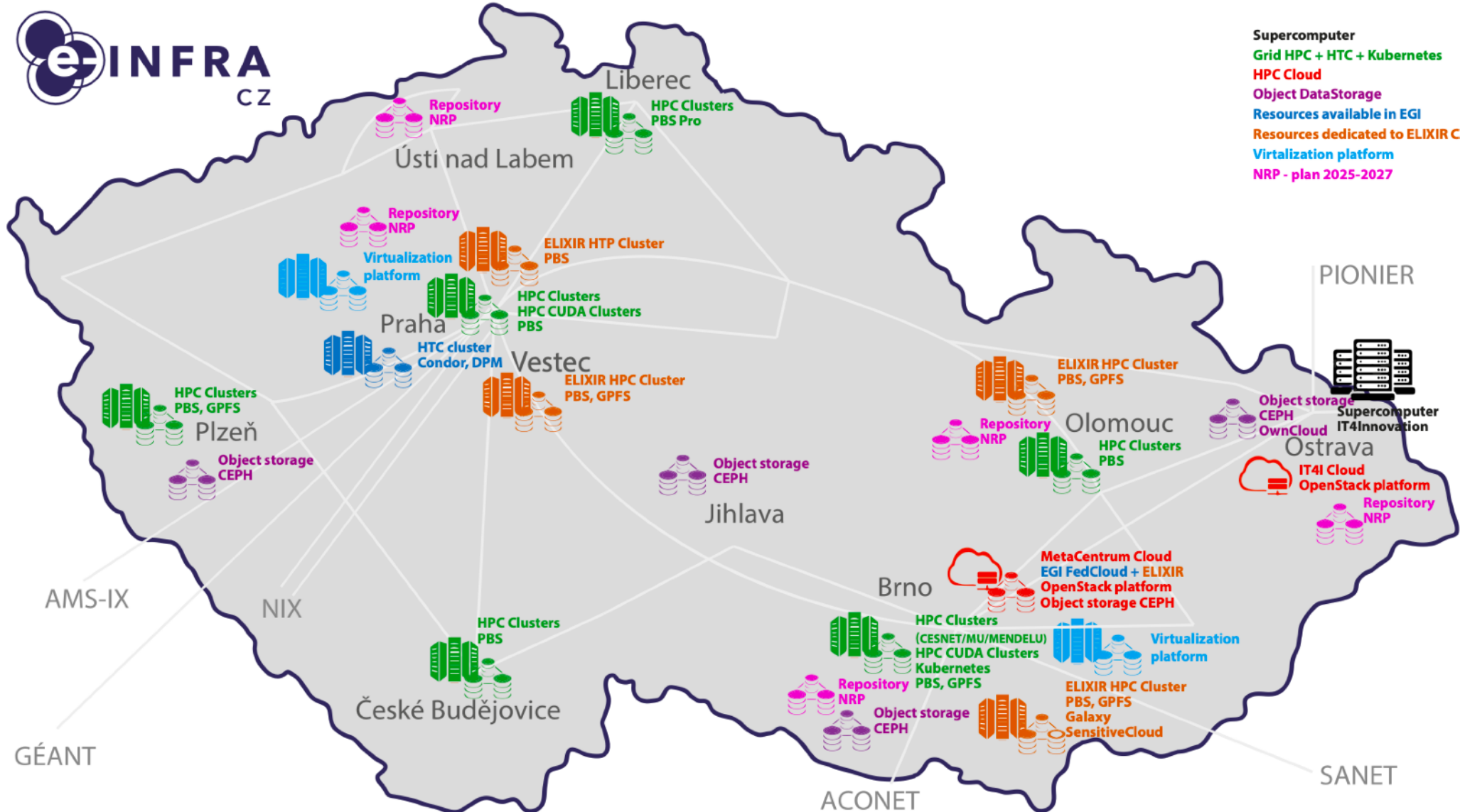
■ MetaCentrum targets

- ... individual users (we can offer resources) <https://metacentrum.cz>
- ... projects (cooperation, sharing data in a group) <https://metavo.metacentrum.cz>
- ... organisations (incorporate their resources under central management) <https://docs.metacentrum.cz>







■ MetaCentrum offers

- ... the principle of grid usage (privileged access for cluster owners)
- ... immediate access to HW resources
- ... access without submitted projects (with one exception)
- ... various application tools (commercial, free, open source)
- ... CPU/GPU resources, GUI applications and access, cloud services

- **Compute resources are provided by CESNET and partners (universities, CAS institutes) and**
 - ... are freely available for research and academic usage
 - ... are shared among all users
 - ... are with privileged access for cluster owners
 - ... can be used in case of urgent/heavy load extensively
 - ... are replaceable during an outage
 - ... are centrally managed, AAI
 - ... are dedicated to grid HPC/HTC and containerised computing, cloud computing, data storage capacities

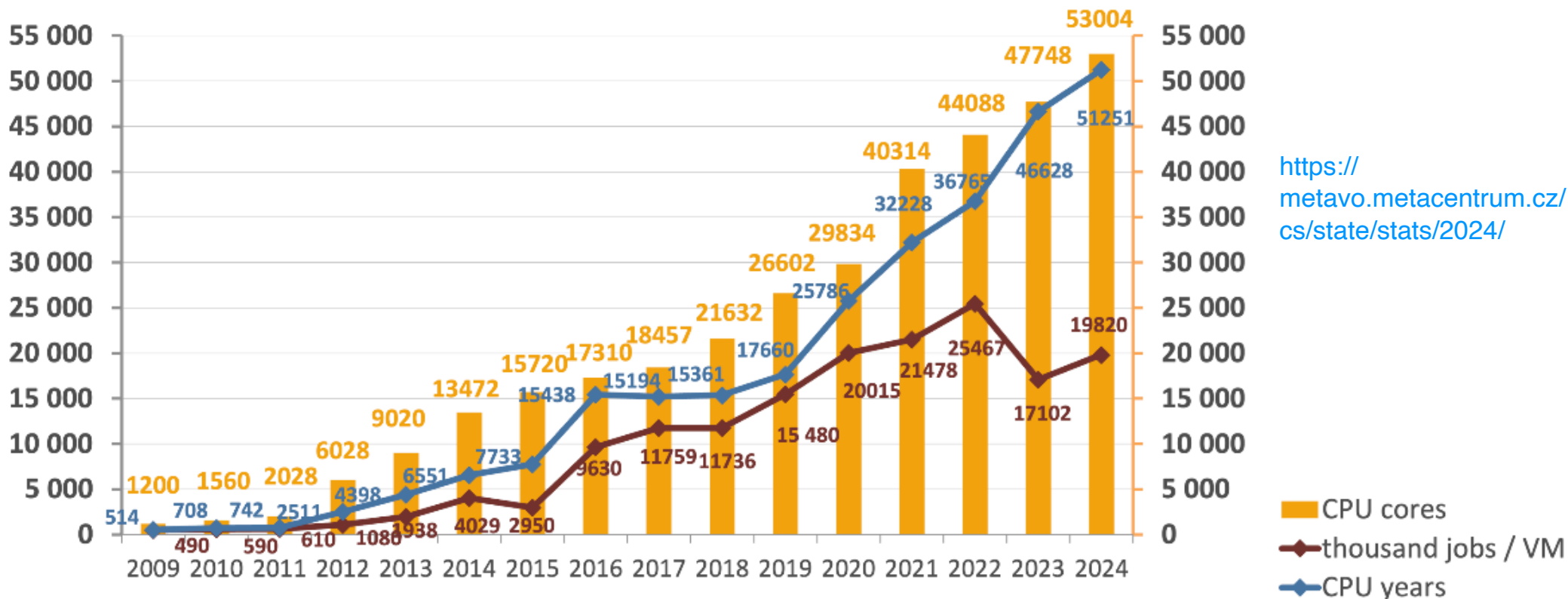


- Supercomputer
- Grid HPC + HTC + Kubernetes
- HPC Cloud
- Object DataStorage
- Resources available in EGI
- Resources dedicated to ELIXIR CZ
- Virtualization platform
- NRP - plan 2025-2027

- Submission of application is conditional by academic affiliation (eduID)
- Access is granted to the MetaCentrum and its services
 - Grid computing (CLI, OpenPBS scheduler)  <https://metavo.metacentrum.cz/en/application>
 - Open OnDemand (GUI, remote web access)  <https://ondemand.grid.cesnet.cz>
 - Galaxy (GUI, web-based platform for computational analyses)  <https://usegalaxy.cz>
 - Jupyter notebook (GUI)  <https://docs.metacentrum.cz/related/jupyter/>
- And also to
 - Cloud computing (powered by OpenStack)  <https://docs.e-infra.cz/compute/openstack/>
 - Container platform Kubernetes (operated by CERIT-SC)  <https://docs.cerit.io>

MetaCentrum Grid	Kubernetes	OpenStack Cloud
Centrally managed infrastructure (restrictive in some cases)	Centrally managed infrastructure (containers can be modified)	Images provided by MetaCentrum, EGI, projects, users,...
Batch/interactive jobs with reserved resources (OpenPBS)	Individually started non-root containers with reserved HW	Long-term running VMs with HW reserved via specific flavours
Compute nodes and storages are distributed across the CZE	Central localisation in Brno (CERIT-SC)	Central OpenStack installation in Brno
Mainly CLI, also GUI approaches (OnDemand, Galaxy)	Considerable interactive support (mostly remote GUI applications)	CLI management
Limitations given by OS and system libraries	Potentially problematic container incorporation to K8s/Rancher	High level of user freedom, independent work with VMs
Easily supported	Easily supported	Problematic access into user's VMs
Kerberos autentization	Web application (AAI)	SSH keys

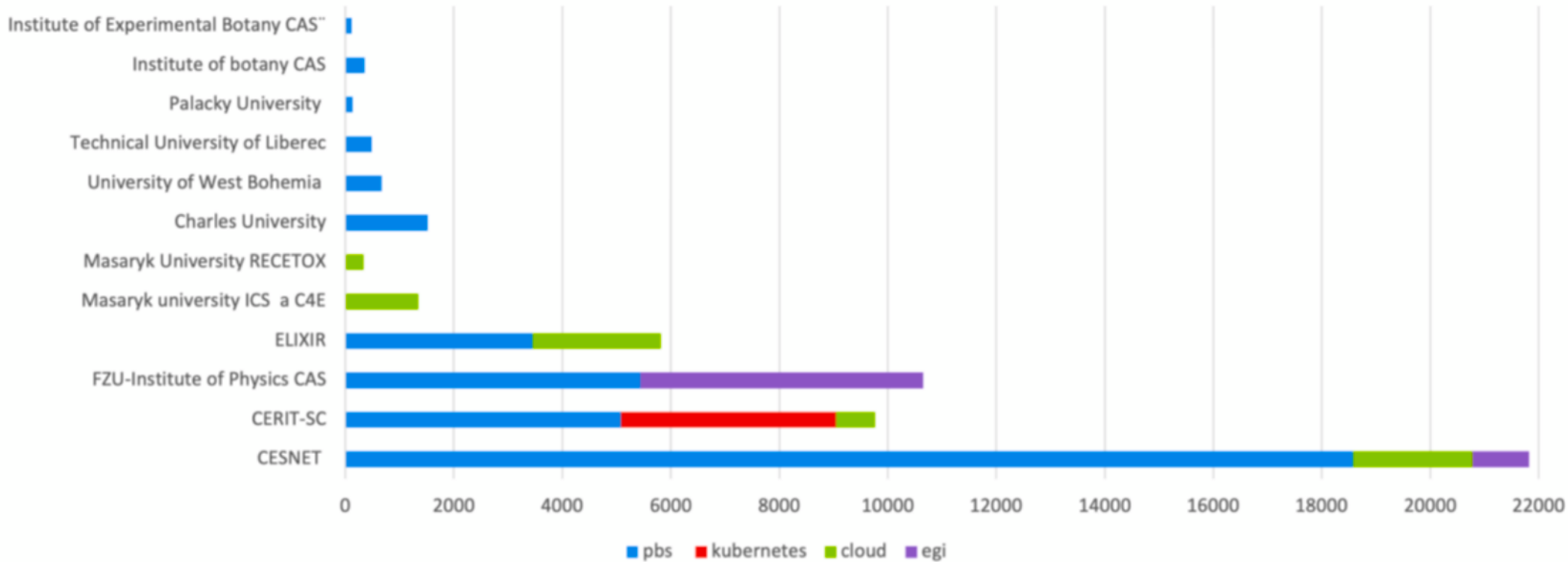
Number of CPUs, executed jobs and corresponding CPU years
(PBS, cloud, K8s, EGI)



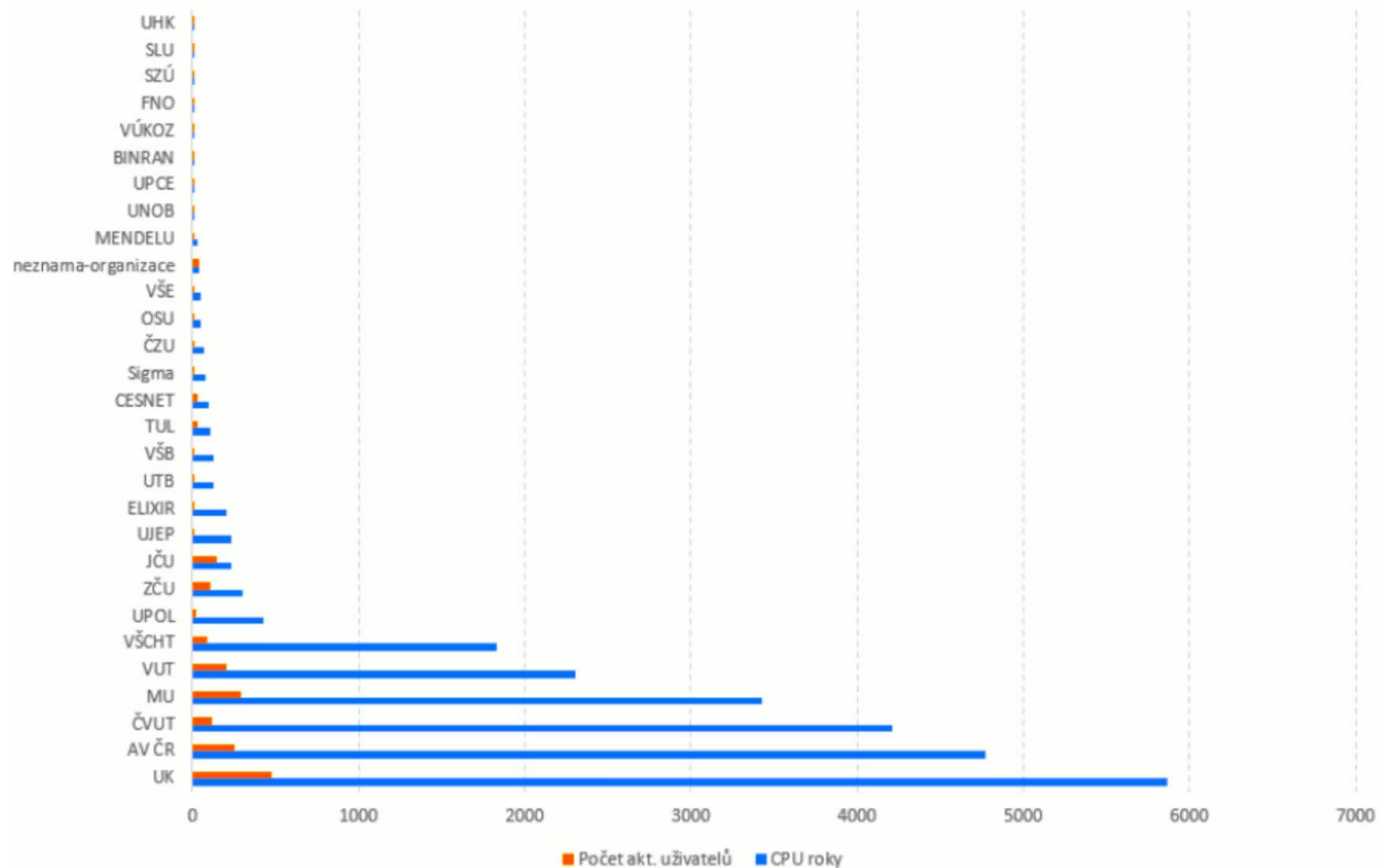
<https://metavo.metacentrum.cz/cs/state/stats/2024/>

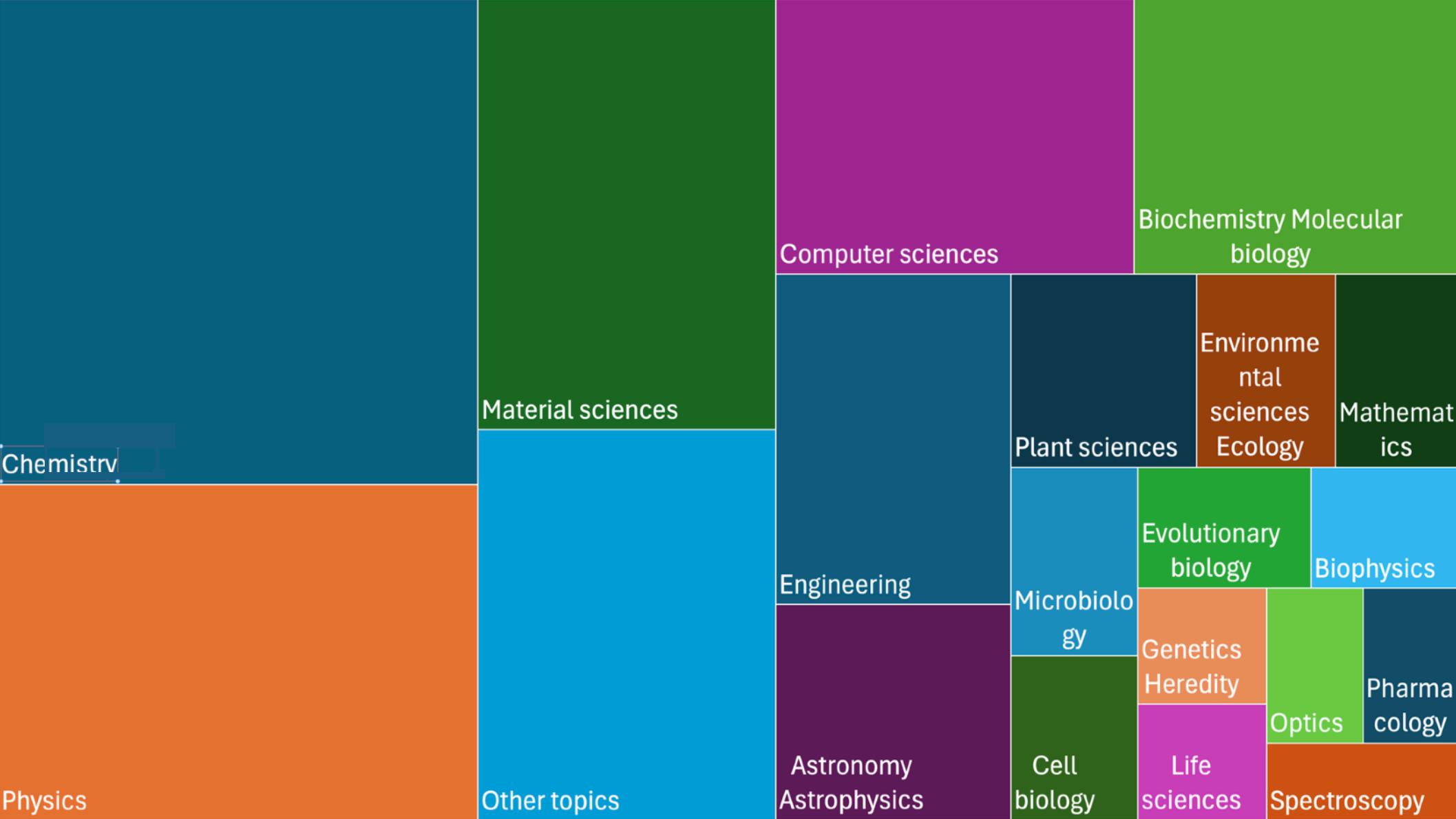
	2012	2016	2018	2019	2020	2021	2022	2023	2024
Počet uživatelů MC	613	1611	2020	2185	2225	2606	2710	3055	3490
Noví uživatelé (Meta)	312	742	713	762	774	792	767	850	1232
Počet úloh [miliony] Meta/EGI	1,1/ n/a	3,6/ 6	5/ 6,7	8,6/ 6,8	13,1/ 10	12,1/ 9,3	11,1/ 14,2	11,7/ 5,4	15/4,9
CPU čas [CPU let] Meta/EGI	2500/ n/a	9475/ 5963	11357/ 4074	13129/ 4531	16630/ 9160	22647/ 9581	27547/ 9218	31858 / 14770	37552/ 217886 HEPSCORE
Počet CPU jader vč. EGI	6028	17234	21344	26602	29874	34084	44088	47748	53004

NGI Resource Providers (MetaVO, MetaCentrum Cloud, EGI) #CPUs
(physical CPU cores, end of 2024)



Instituce a propočítaný čas a počet uživatelů (2024)





- **Free/open-source academic software**
 - Mainly distributed via the system of modules (>5,000 individual modules)
- **Expensive commercial licences available to users**
 - Matlab (modelling), Molpro (molecular modelling), Ansys (engineering modelling), Gaussian (quantum mechanics), Turbomole (quantum chemistry)
- **Users have their personal licenses (free or paid, restricted access)**
 - VASP (molecular mechanics), Crystal (solid state chemistry and physics)
- **Users are allowed to install almost any application software on their own**
 - The current policy is to support users in local installations
 - Just do not violate the license terms...

<https://docs.metacentrum.cz/software/alphabet/>

- **There are so many different ways...** <https://docs.metacentrum.cz/software/install-software/>
 - Binary distributions (precompiled form, download them and use them)
 - R, Python, Perl, Julia, Debian, etc. libraries (from repositories)
 - Package managers like Mamba (fully automated, easy to use)
 - Docker (Kubernetes, cloud) and Singularity/Apptainer images (grid, conversion Docker -> Singularity) <https://docs.metacentrum.cz/software/containers/>
 - Snapshots of entire VMs (cloud, OnDemand)
 - Local compilation (GCC, Intel compilers, BLAS/LAPACP math libraries, CUDA support and so on...)

■ Fill out and submit the registration form

<https://metavo.metacentrum.cz/en/application>

- Select your organisation (click on the eduID logo)
- Use your institutional username and password
- Fill out the form and create a strong (and unique) MetaCentrum password
- Applications are evaluated and approved manually
- All accounts are valid till 02. 02. YYYY
- Users must extend MetaCentrum membership from the beginning of each calendar year (during January).



■ Read our documentation, FAQ and tutorial for beginners

<https://docs.metacentrum.cz/>

<https://docs.metacentrum.cz/support/faqs/>

<https://docs.metacentrum.cz/computing/concepts/>



I have an account in a member organisation of eduid.cz

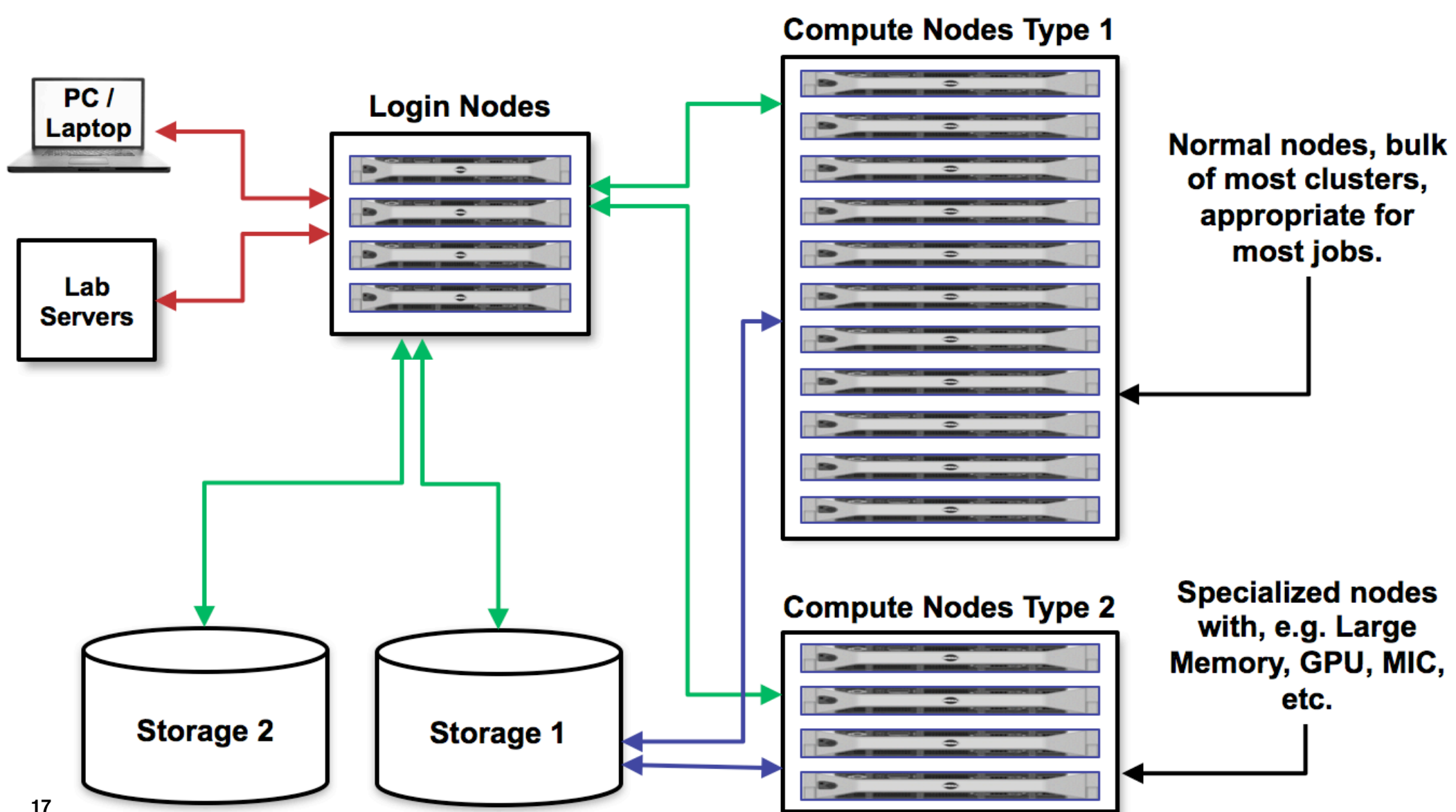


My organisation is not in eduid.cz and I need to validate my alternative identity

Direct links to selected institutions from eduid.cz

 Masarykova univerzita	 Univerzita Karlova	 Západočeská univerzita v Plzni
 Jihočeská univerzita v Českých Budějovicích	 Univerzita Palackého v Olomouci	 Univerzita Pardubice
 CESNET	 České vysoké učení technické v Praze	 Mendelova univerzita v Brně
 Technická univerzita Liberec	 Vysoká škola báňská - Technická univerzita Ostrava	 Vysoké učení technické v Brně

- Only for "sponsored" accounts
- Default validity is three months
- Further validity extensions are done manually after the request



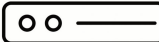
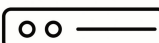


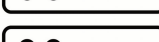
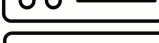
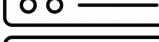

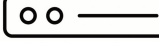
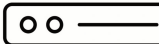
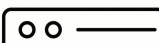
- Gateway to the entire grid infrastructure
- Accessible via ssh with a password (ssh tickets are not fully supported)
- Frontends submit jobs to the OpenPBS scheduler
- Frontends are relatively small virtual machines mainly for writing scripts for batch jobs, submitting jobs, checking available applications and user data, quick installations and calculations (e.g. data extraction), etc.
- **Do not run long and/or demanding calculations directly on frontends!**
 - Overload -> slowdown -> failure
- Frontend servers usually have different home directories
- Command line interface (mainly CLI) <https://docs.metacentrum.cz/computing/infrastructure/frontends/>
<https://docs.metacentrum.cz/access/kerberos/>

- OpenPBS (Portable Batch System) is a software that performs job scheduling and management
- Frontend servers can have different home directories
- All user home directories are available from all frontends



pbs-m1.metacentrum.cz



	charon.metacentrum.cz	/storage/liberec3-tul/home/
	elmo.metacentrum.cz	/storage/praha5-elixir/home/
	luna.metacentrum.cz	/storage/praha1/home/
	nympha.metacentrum.cz	/storage/plzen1/home/
	oven.metacentrum.cz	/storage/brno2/home/
	perian.metacentrum.cz	/storage/brno2/home/
	tarkil.metacentrum.cz	/storage/praha1/home/
	skirit.metacentrum.cz	/storage/brno2/home/
	tilia.metacentrum.cz	/storage/pruhonice1-ibot/home/
	zenith.metacentrum.cz	/storage/brno12-cerit/home/
	zuphux.metacentrum.cz	/storage/brno12-cerit/home/

<https://docs.metacentrum.cz/computing/concepts/#frontends-storages-homes>

- Data is stored on a few independent storages; the capacity is not infinite
- Storages have quotas for the total volume of data and the number of files
- All storages are accessible through all frontends
- **Data on storage is not fully backed up.**

NFS4 server	adresář - directory	velikost - capacity	zálohovací třída - back-up policy
storage-brno1-cerit.metacentrum.cz	/storage/brno1-cerit/	1.8 PB	2
storage-brno2.metacentrum.cz	/storage/brno2/	306 TB	2
storage-brno11-elixir.metacentrum.cz	/storage/brno11-elixir/	313 TB	2
storage-brno12-cerit.metacentrum.cz	/storage/brno12-cerit/	3.4 PB	2
storage-budejovice1.metacentrum.cz	/storage/budejovice1/	44 TB	3

<https://docs.metacentrum.cz/computing/infrastructure/storages/> <https://docs.metacentrum.cz/computing/infrastructure/frontend-storage/>

- Data is stored on a few independent storages; the capacity is not infinite
- Storages have quotas for the total volume of data and the number of files
- All storages are accessible through all frontends
- **Data on storage is not fully backed up.**

- **Not for archiving purposes!**

- Valuable data should be permanently archived on S3 object storage.

<https://docs.du.cesnet.cz/en/docs/object-storage-s3/s3-service>

- SSH keys for logging into frontends **are not fully supported**. We want to "force you" to generate a Kerberos ticket by typing the password

```
jirivorel@MacBook ~$ ssh vorel@nympha.metacentrum.cz
```

```
vorel@nympha.metacentrum.cz's password:
```

```
(BULLSEYE)vorel@nympha:~$ klist
```

```
Credentials cache: FILE:/tmp/krb5cc_1597_LTYWlt
```

```
Principal: vorel@META
```

Type a password

klist command prints the status of issued tickets

Issued	Expires	Principal
May 6 11:22:55 2022	May 6 21:22:55 2022	krbtgt/META@META
May 6 11:22:55 2022	May 6 21:22:55 2022	afs/ics.muni.cz@META
May 6 11:22:55 2022	May 6 21:22:55 2022	krbtgt/ZCU.CZ@META
May 6 11:22:55 2022	May 6 21:22:55 2022	afs/zcu.cz@ZCU.CZ

```
(BULLSEYE)vorel@nympha:~$ ssh halmir1  
Linux halmir1.metacentrum.cz 5.10.0-13-amd64 #1 SMP Debian 5.10.106-1+zs1 (2022-03-28) x86_64  
Last login: Thu Apr 21 09:54:05 2022 from elmo2-4.hw.elixir-czech.cz  
(BULLSEYE)vorel@halmir1:~$
```

- In the future, primary access for new users (replacement of CLI)
- Frontend servers and storages can be accessed through the OnDemand
 - Web-based CLI access to selected frontends (with all functionalities)
 - Web-based interactive access on storages (mainly for browsing)
- Deployment of VMs and containers
- S3, OneData browser
- Singularity images (NGC, Pytorch,...)

<https://docs.metacentrum.cz/ondemand/>


<https://ondemand.grid.cesnet.cz>

MetaCentrum Open OnDemand provides an integrated, single access point for HPC resources.

Selected applications - all apps



Jupyter Lab/Notebook



RStudio



Matlab



FLUENT
Ansys Fluent



ENSIGHT
Ansys Ensignt



WORKBENCH
Ansys Workbench



Frontend shell




Meta Desktop



Job Composer



BIOP Desktop



CLC Genomics WB



VMD Desktop

Announcements

23-04-2024

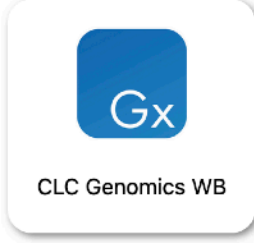
OnDemand has been upgraded to version 3.1.4. Jobs are now submitted to OpenPBS server pbs-m1.metacentrum.cz.

21-08-2023

OnDemand has been upgraded to the major version 3.

- 🏠 Home Directory
- 📁 brno11-elixir /storage/brno11-elixir/home/vorel
- 📁 brno12 /storage/brno12-cerit/home/vorel**
- 📁 brno14-ceitec /storage/brno14-ceitec/home/vorel
- 📁 brno2 /storage/brno2/home/vorel
- 📁 brno3-cerit /storage/brno3-cerit/home/vorel
- 📁 budejovice1 /storage/budejovice1/home/vorel
- 📁 liberec3 /storage/liberec3-tul/home/vorel
- 📁 plzen1 /storage/plzen1/home/vorel
- 📁 plzen4-ntis /storage/plzen4-ntis/home/vorel
- 📁 praha1 /storage/praha1/home/vorel
- 📁 praha2-natur /storage/praha2-natur/home/vorel
- 📁 praha5-elixir /storage/praha5-elixir/home/vorel
- 📁 praha5-elixir /storage/praha5-elixir/home/vorel
- 📁 praha6-fzu /storage/praha6-fzu/home/vorel
- 📁 projects /storage/projects
- 📁 projects2 /storage/projects2
- 📁 pruhonice1-ibot /storage/pruhonice1-ibot/home/vorel

... is an integrated, single access point for HPC resources.



OnDemand has been upgraded to version 3.1.4. Jobs are now submitted to OpenPBS server pbs-m1.metacentrum.cz.

21-08-2023

OnDemand has been upgraded to the major version 3.

[Open in Terminal](#) [Refresh](#) [+ New File](#) [New Directory](#) [Upload](#) [Download](#) [Copy/Move](#) [Delete](#)

- Home Directory
- brno2
- brno12
- praha5-elixir
- brno11-elixir
- brno14-ceitec
- brno3-cerit
- budejovice1
- liberec3
- plzen1
- plzen4-ntis
- praha1
- praha2-natur
- praha5-elixir
- praha6-fzu
- pruhonice1-ibot
- projects
- projects2

[↑](#) / [storage](#) / [brno12-cerit](#) / [home](#) / [vorel](#) / [Change directory](#) [Copy path](#)



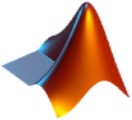









Show Owner/Mode Show Dotfiles Filter:

Showing 4 of 15 rows - 0 rows selected

<input type="checkbox"/>	Type ▲	Name	Size	Modified at
<input type="checkbox"/>	Folder	brno11-elixir	-	14. 10. 2024 15:42:44
<input type="checkbox"/>	Folder	Metylace_hemonch	-	20. 2. 2025 11:49:29
<input type="checkbox"/>	Folder	ondemand	-	17. 5. 2024 14:11:00
<input type="checkbox"/>	Folder	Tools	-	21. 10. 2024 14:00:36

MetaCentrum Open OnDemand provides an integrated, single access point for HPC resources.

Selected applications - all apps

 Jupyter Lab/Notebook	 RStudio	 Matlab	 Ansys Fluent	 Ansys Ensignt	 Ansys Workbench
 Frontend shell	 Meta Desktop	 Job Composer	 BIOP Desktop	 CLC Genomics WB	 VMD Desktop

Announcements

23-04-2024

OnDemand has been upgraded to version 3.1.4. Jobs are now submitted to OpenPBS server pbs-m1.metacentrum.cz.

21-08-2023

OnDemand has been upgraded to the major version 3.


```
(BOOKWORM)vorel@perian:~$ ls -l
total 471538
drwxr-xr-x  6 vorel meta      4096 Jun  5  2024 bcl2fastq
-rw-r--r--  1 vorel meta 215582499 Sep  7  2017 bcl2fastq2-v2.20.0.422-Source.tar.gz
-rw-r--r--  1 vorel meta 215646796 May 16  2024 bcl2fastq2-v2-20-0-tar.zip
drwxr-xr-x  6 vorel meta      4096 Jun  4  2024 bcl2fastq_moje_upravy_zaloha
drwxr-xr-x 10 vorel meta      4096 Feb 26 11:35 BirdNET-Analyzer
drwxr-xr-x  3 vorel meta      4096 Nov  7 15:04 foldseek
-rw-r--r--  1 vorel meta 51610080 Nov 21 15:23 foldseek-linux-avx2.tar.gz
drwxr-xr-x  4 vorel meta      4096 May 17  2024 ondemand
drwx----- 13 vorel meta      4096 Feb 11 15:09 Smilei
-rw-r--r--  1 vorel meta      16 Mar  3 11:54 test
drwxr-xr-x  6 vorel meta      4096 Jul  1  2024 test_crys
drwxr-xr-x 15 vorel meta      4096 Jan 30 16:12 test_deepsig3
drwxr-xr-x  4 vorel meta      4096 Jan  3 13:53 test_phyluce
drwxr-xr-x  3 vorel meta      4096 Jan  8 16:04 test_unic
```

```
(BOOKWORM)vorel@perian:~$ cat test
test test
```

```
test
```

```
(BOOKWORM)vorel@perian:~$ qsub --help
```

```
qsub: invalid option -- '-'
```

```
usage: qsub [-a date_time] [-A account_string] [-c interval]
        [-C directive_prefix] [-e path] [-f ] [-h ] [-I [-X]] [-j oe|eo] [-J X-Y[:Z]]
        [-k keep] [-l resource_list] [-m mail_options] [-M user_list]
        [-N jobname] [-o path] [-p priority] [-P project] [-q queue] [-r y|n]
        [-R o|e|oe] [-S path] [-u user_list] [-W otherattributes=value...]
        [-v variable_list] [-V ] [-z] [script | -- command [arg1 ...]]
```

```
qsub --version
```

```
(BOOKWORM)vorel@perian:~$
```


Host: perian.grid.cesnet.cz


```
(BOOKWORM)vorel@perian:~$ pwd
/storage/brno2/home/vorel
(BOOKWORM)vorel@perian:~$ cd /storage/brno12-cerit/home/vorel
(BOOKWORM)vorel@perian:/storage/brno12-cerit/home/vorel$ ls
brno11-elixir  Metylance_hemonch  ondemand  Tools
(BOOKWORM)vorel@perian:/storage/brno12-cerit/home/vorel$ pwd
/storage/brno12-cerit/home/vorel
(BOOKWORM)vorel@perian:/storage/brno12-cerit/home/vorel$ exit
logout
Connection to perian.grid.cesnet.cz closed.
```

Your connection to the remote server has been terminated. ■


Interactive Apps


Cloud

 Kubernetes infra example OS

 Simple OS virtual machine

Desktops

 Ansys/Ensignt

 Ansys/Fluent

 Ansys/Workbench

 BIOP Desktop

Simple OS virtual machine

This is a simple VM deployed to OpenStack

Select project

dbc23d6dbd554be659114117efd4faf0f57466f4@einfra ▾

Your help message

ssh public key

ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQgQDCX5scc


Insert your public key to enable ssh


Launch

* The Simple OS virtual machine session data for this session can be accessed under the [data root directory](#).

Interactive Apps


Cloud

 **Kubernetes infra example OS**

 Simple OS virtual machine

Desktops

 Ansys/Ensignt

 Ansys/Fluent

 Ansys/Workbench

 BIOP Desktop

 CLCgenomicsWB

 Matlab

Kubernetes infra example OS

Kubernetes infrastructure built with ansible and kubespray over OpenStack. Based on <https://gitlab.ics.muni.cz/cloud/kubernetes/kubernetes-infra-example>

Select project

dbc23d6dbd554be659114117efd4faf0f57466f4@einfr: ▾

Select your openstack project

Number of kubernetes control nodes

1

Select number of control nodes based on your needs

Number of kubernetes worker nodes

1

Select number of worker nodes based on your needs

Launch

- Each software (in a specific version) is prepared as an individual module file
- In theory, the module file, after activation (command `module add module_name`), will load the main application (set the necessary variables), dependencies and needed libraries
- More than >5,000 modules are available for users
- Users can write their own module files
- Available modules can be listed directly on the frontend

<https://docs.metacentrum.cz/software/modules/>

```
(BOOKWORM)vorel@skirit:~$ module ava *last*
----- /packages/run/modules-5/debian12avx512 -----
blast-plus/ blast/ last/ lastz/ ncbi-magicblast/ ncbi-rmblastn/ samblaster/

Key:
modulepath directory/
(BOOKWORM)vorel@skirit:~$ module ava blast-plus
----- /packages/run/modules-5/debian12avx512 -----
blast-plus/

Key:
modulepath directory/
(BOOKWORM)vorel@skirit:~$ module ava blast-plus/
----- /packages/run/modules-5/debian12avx512 -----
blast-plus/2.10.1-gcc-8.3.0-eh6opkv  blast-plus/2.12.0-gcc-8.3.0-ohlv7t4  blast-plus/2.16.0-gcc-10.2.1-bgzrrrz
blast-plus/2.12.0-gcc-8.3.0-coev6wv  blast-plus/2.12.0-gcc-10.2.1-2phsggo

Key:
modulepath
(BOOKWORM)vorel@skirit:~$ module add blast-plus/2.16.0
```

```
(BOOKWORM)vorel@skirit:~$ module add blast-plus/2.16.0
Loading blast-plus/2.16.0-gcc-10.2.1-bgzrrrz
Loading requirement: bzip2/1.0.8-gcc-10.2.1-ydytex freetype/2.11.1-gcc-10.2.1-ukjspcj
gettext/0.21-gcc-10.2.1-tm75xz5 bdw-gc/8.0.6-gcc-10.2.1-ottg6g5 gmp/6.2.1-gcc-10.2.1-lcdqyb3
libffi/3.4.2-gcc-10.2.1-hrcl4md libtool/2.4.7-gcc-10.2.1-bidj2af libunistring/0.9.10-gcc-10.2.1-iy76hg4
readline/8.1-gcc-10.2.1-6rg3hny guile/2.2.6-gcc-10.2.1-lajneeg libidn2/2.3.0-gcc-10.2.1-ch5vzvm
nettle/3.4.1-gcc-10.2.1-oi5o5t6 zlib/1.2.12-gcc-10.2.1-7qmmk4c gnutls/3.6.15-gcc-10.2.1-mv6pwhr
libjpeg-turbo/2.1.3-gcc-10.2.1-bo2cwla libpng/1.6.37-gcc-10.2.1-3f5z4ey lmbd/0.9.29-gcc-10.2.1-hiiqpmb
lzo/2.10-gcc-10.2.1-te2izub openssl/1.1.1o-gcc-10.2.1-k5zobqv pcre/8.45-gcc-10.2.1-p343mum
perl/5.34.1-gcc-10.2.1-dw2jaxd python/3.9.12-gcc-10.2.1-rg2lpmk
```

```
(BOOKWORM)vorel@skirit:~$ blastn -help
```

USAGE

```
blastn [-h] [-help] [-import_search_strategy filename]
[-export_search_strategy filename] [-task task_name] [-db database_name]
[-dbsize num_letters] [-gilist filename] [-seqidlist filename]
[-negative_gilist filename] [-negative_seqidlist filename]
[-taxids taxids] [-negative_taxids taxids] [-taxidlist filename]
[-negative_taxidlist filename] [-no_taxid_expansion]
[-entrez_query entrez_query] [-db_soft_mask filtering_algorithm]
[-db_hard_mask filtering_algorithm] [-subject subject_input_file]
[-subject_loc range] [-query input_file] [-out output_file]
[-evaluate evalue] [-word_size int_value] [-gapopen open_penalty]
```

- HW resources (CPUs, GPUs, RAM, scratch, walltime,...) are reserved by PBS
- Detailed documentation: <https://docs.metacentrum.cz/computing/resources/resources/>
<https://docs.metacentrum.cz/computing/resources/qsub-compiler/>
- It requires some experience
- Helper tool for qsub command (reserves resources and submits jobs) assembly

Personal view

This page shows a personal view of the PBS system for the user **vorel**.

Jobs of user "vorel"

user	job count					CPU count				
	total	queued	running	completed	other	total	queued	running	completed	other
vorel	0	0	0	0	0	0	0	0	0	0

Links

- list of my jobs
- personal view of storages
- **qsub assembler**

Click on it...

Go to metavo.metacentrum.cz -
Current state - Personal view - **qsub assembler**

(Stav zdrojů - Osobní pohled **sestavovač qsub**)

<https://metavo.metacentrum.cz/pbsmon2/person>

qsub -l walltime= 24 : 0 : 0 -q default@meta-pbs.metacentrum.cz \

-l select= 1 :ncpus= 8 :ngpus= 0 :mem= 100 gb :scratch_ssd = 50

cluster ...

city ...

other resources ...

:arch=

:biocev=

:cgroups=

:cluster=

:cpu_flag=

:cpu_vendor=

:cuda_version=

:debian10=

:gpu_cap=

:host=

:hyperthreading=

:infiniband=

:luna=

:os=

:osfamily=

:pruhonice=

:scratch_shm=

:spec=

:vestec=

:vnode=

Click on it...

Find machines mathing the resource specification

■ And you will see...

selection from command line

```
qsub -l walltime=24:0:0 -q default@meta-pbs.metacentrum.cz -l select=1:ncpus=8:mem=100gb:scratch_ssd=50gb
```

selection in shell script

```
#!/bin/bash
#PBS -q default@meta-pbs.metacentrum.cz
#PBS -l walltime=24:0:0
#PBS -l select=1:ncpus=8:mem=100gb:scratch_ssd=50gb
#PBS -N my_awesome_job
```

Result

OK

The requirement is 1 machine, and 93 such machines are free, out of 289 machines matching the requirements. The job is for it.

Machines available right now

adan1 (32 CPU, 187.6 GiB RAM, 697.6 GiB HDD)	adan2 (32 CPU, 187.6 GiB RAM, 783.6 GiB HDD)	adan3 (16 CPU, 171.6 GiB RAM, 766.6 GiB HDD)	adan5 (32 CPU, 187.6 GiB RAM, 744.6 GiB HDD)	adan6 (32 CPU, 187.6 GiB RAM, 705.4 GiB HDD)
--	--	--	--	--

```
#!/bin/bash
#PBS -q default@meta-pbs.metacentrum.cz
#PBS -l walltime=24:0:0
#PBS -l select=1:ncpus=8:mem=100gb:scratch_ssd=50gb
#PBS -N my_awesome_job
#PBS -m e

# test if a scratch directory exists
# variable SCRATCHDIR is set automatically
test -n "$SCRATCHDIR" || { echo >&2 "Variable SCRATCHDIR is not set!"; exit 1; }

# set a DATADIR variable
DATADIR /storage/brno12-cerit/home/vorel/data/

# copy input file "data.fa" to the scratch directory
cp $DATADIR/data.fa $SCRATCHDIR

# move into the scratch directory
cd $SCRATCHDIR

# load a module for your application
module add blast-plus/blast-plus-2.12.0-gcc-8.3.0-ohlvt4

# run the calculation
# do not forgeto to use reserved CPUs by '-num_threads' flag
# variable PBS_NCPUS is a number of CPUs requested for the entire job
blastp -query data.fa <other_parameters> -num_threads $PBS_NCPUS -out results.txt

#copy results
cp results.txt $DATADIR

# clean the scratch directory
clean_scratch
```

- Define HW resources (**-l**), queue (**-q**), walltime (**-l**), set the job name (**-N**), and email alert (**-m**)
- You can define as many variables as you want
- Available modules can be listed by command **module ava <key_word>** on any frontend
- The scratch directory will be cleaned automatically
- **qsub script_name.sh**

<https://docs.metacentrum.cz/computing/run-basic-job/>

- The opposite of batch jobs (waiting for the user's input...)
- Best choice for test calculations (which should not be run directly on frontends)
- An interactive job is requested by the `qsub` command with the `-I` (uppercase "i") option

<https://docs.metacentrum.cz/computing/run-basic-job/#interactive-job>

```
(BUSTER)vorel@skirit:~$ qsub -I -l select=1:ncpus=4:mem=50gb:scratch_local=30gb -l walltime=1:00:00
```

```
qsub: waiting for job 11405230.meta-pbs.metacentrum.cz to start
```

```
qsub: job 11405230.meta-pbs.metacentrum.cz ready
```

```
vorel@zenon31:~$ cd $SCRATCHDIR
```

```
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$ module add orca/orca-5.0.1-intel-19.0.4-bnofsgq
```

```
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$ module list
```

```
Currently Loaded Modulefiles:
```

```
1) metabase 2) openmpi/openmpi-4.0.4-intel-19.0.4-gpu-xri6uan 3) orca/orca-5.0.1-intel-19.0.4-bnofsgq
```

```
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$
```



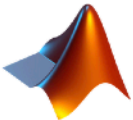









```
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$ ...time for coffee...
```

```
-bash: ...time: command not found
```

```
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$ orca < input > output
```

MetaCentrum Open OnDemand provides an integrated, single access point for HPC resources.

Selected applications - all apps

 Jupyter Lab/Notebook	 RStudio	 Matlab	 Ansys Fluent	 Ansys Ensight	 Ansys Workbench
 Frontend shell	 Meta Desktop	 Job Composer	 BIOP Desktop	 CLC Genomics WB	 VMD Desktop

Announcements

23-04-2024

OnDemand has been upgraded to version 3.1.4. Jobs are now submitted to OpenPBS server pbs-m1.metacentrum.cz.

21-08-2023

OnDemand has been upgraded to the major version 3.

Interactive Apps

Desktops

Ansys/Ensignt

Ansys/Fluent

Ansys/Workbench

BIOP Desktop

CLCgenomicsWB

Matlab

MetaCentrum Desktop

VMD Desktop

Servers

Jupyter Notebook/Lab

Matlab webapp (beta)

RStudio Server

RStudio Server

This app will launch an RStudio server on one or more nodes. Geospatial and Tensorflow packages are preinstalled.

Number of hours

12

Number of CPUs on single node

2

Memory (GB)

10

GPUs

0

Scratch local (GB)

100

RStudio Image version

✓ RStudio-geospatial-4.2.2

RStudio-geospatial-4.3.0

RStudio-geospatial-4.4.1

/storage/brno2

Launch

* The RStudio Server session data for this session can be accessed under the [data root directory](#).

Session was successfully created. ✕

[Home](#) / [My Interactive Sessions](#)

Interactive Apps

Desktops

 Ansys/Ensignt

 Ansys/Fluent

 Ansys/Workbench

 BIOP Desktop

 CLCgenomicsWB

 Matlab

 MetaCentrum Desktop

 VMD Desktop

Servers

 Jupyter Notebook/Lab

 Matlab webapp (beta)

 RStudio Server

RStudio Server (9247441.pbs-m1.metacentrum.cz)

1 node | 2 cores | Running

Host: nymph53.meta.zcu.cz

✕ Delete

Created at: 2025-03-03 12:57:32 CET

Time Remaining: 11 hours and 59 minutes

Session ID: 26505f06-f237-41df-8b24-cacba30f4c00

[Ⓜ Connect to RStudio Server](#)

File Edit Code View Plots Session Build Debug Profile Tools Help

vorel Project: (None)

Console Terminal Background Jobs

R 4.2.2 · /auto/brno2/home/vorel

R version 4.2.2 (2022-10-31) -- "Innocent and Trusting"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Session restored from your saved work on 2025-Jan-03 10:18:44 UTC (59 days ago)

```
> getwd()
[1] "/auto/brno2/home/vorel"
```

Environment History Connections Tutorial

154 MiB

Global Environment

Environment is empty

Files Plots Packages Help Viewer Presentation

New Folder New Blank File Upload Delete Rename More

/ auto / brno2 / home / vorel

Name	Size	Modified
..		
bcl2fastq		
bcl2fastq_moje_upravy_zaloha		
bcl2fastq2-v2-20-0-tar.zip	205.7 MB	May 16, 2024, 6:15 PM
bcl2fastq2-v2.20.0.422-Source.tar.gz	205.6 MB	Sep 7, 2017, 2:40 PM
BirdNET-Analyzer		
foldseek		
foldseek-linux-avx2.tar.gz	49.2 MB	Nov 21, 2024, 3:23 PM
ondemand		
R		
Smilei		
test	16 B	Mar 3, 2025, 11:54 AM
test_crys		
test_deepsig3		
test_phyluce		
test_unic		

- The default queue is the best choice for almost all calculations
- `-q` option of the `qsub` assembler

`qsub -l walltime= 1 : 0 : 0 -q`

`-l select= 1 :ncpus= 1 :ngpus`

`cluster ...`

`city ...`

`SPECfp2017 per core ...`

`other resources ...`

`:arch=`

`:biocev=`

`:cgroups=`

`:cluster=`

`:cpu_flag=`

`:cpu_vendor=`

`:cuda_version=`

`:gpu_cap=`

`✓ default@pbs-m1.metacentrum.cz`

`interactive@pbs-m1.metacentrum.cz`

`MetaSeminar@pbs-m1.metacentrum.cz`

`oven@pbs-m1.metacentrum.cz`

`uv18@pbs-m1.metacentrum.cz`

`gpu_long@pbs-m1.metacentrum.cz`

`gpu@pbs-m1.metacentrum.cz`

`large_mem@pbs-m1.metacentrum.cz`

`runone@pbs-m1.metacentrum.cz`

`backfill@pbs-m1.metacentrum.cz`

`elixircz@pbs-m1.metacentrum.cz`

`uv@pbs-m1.metacentrum.cz`

General queue for all jobs

Special queue for practical courses/seminars

Queue targeting special nodes

Queue for nodes dedicated to VI Elixir CZ

q_4d@pbs-m1.metacentrum.cz	🚫	50	48:00:01 - 96:00:00	749	543 /	2111	3403
p2e_1d@pbs-m1.metacentrum.cz	🔒	50	00:00:01 - 24:00:00	1	1468 /	6267	7736
q_2w@pbs-m1.metacentrum.cz	🚫	50	168:00:01 - 336:00:00	12	158 /	79	10250
molpro@pbs-m1.metacentrum.cz	🔒	50	0 - 0	0	0 /	0	0
q_2h@pbs-m1.metacentrum.cz	🚫	50	0 - 02:00:00	3866	509 /	597	4982
q_2d@pbs-m1.metacentrum.cz	🚫	50	24:00:01 - 48:00:00	2766	1269 /	2542	6577
q_1w@pbs-m1.metacentrum.cz	🚫	50	96:00:01 - 168:00:00	1986	837 /	418	4477
q_1d@pbs-m1.metacentrum.cz	🚫	50	04:00:01 - 24:00:00	14886	1020 /	3840	19759
runone@pbs-m1.metacentrum.cz		50	00:00:00 - 24:00:00	0	23 /	44	67
q_4h@pbs-m1.metacentrum.cz	🚫	50	02:00:01 - 04:00:00	154	167	406	576
q_2w_plus@pbs-m1.metacentrum.cz	🚫	50	336:00:01 - 720:00:00	143	248 /	2	407
backfill@pbs-m1.metacentrum.cz		20	00:00:01 - 24:00:00	0	0 /	0	0
elixircz@pbs-m1.metacentrum.cz	🔒 ⓘ	0	0 - 720:00:00	0	0 /	0	0
uv@pbs-m1.metacentrum.cz	ⓘ	0	00:00:01 - 168:00:00	0	0 /	0	0
default@pbs-m1.metacentrum.cz	ⓘ	0	0 - 720:00:00	0	0 /	2	6

nesubmitujte přímo do fronty, použijte směrovací frontu

- Dedicated queues for cluster owners
- High priority on dedicated compute nodes
- Only short jobs for other users (for example, 24 vs 720 hours)

biocev@pbs-m1.metacentrum.cz		80	0 - 0	0	0 /	0	0
cvut@pbs-m1.metacentrum.cz		80	0 - 720:00:00	0	0 /	0	0
ubo@pbs-m1.metacentrum.cz		70	02:00:01 - 720:00:00	0	0 /	0	0
gpu_dgx@pbs-m1.metacentrum.cz		70	0 - 336:00:00	0	1 /	3	4

Fronta je vyhrazena pro tyto skupiny: cvut_fsv_staff cvut_fsv_students pbs-admins

Dostupné uzly

farin1 (121/128) farin2 (0/128) farin3 (0/128) farin4 (84/128)

Dostupné fyzické stroje

farin1 (128 CPU) farin2 (128 CPU) farin3 (128 CPU) farin4 (128 CPU)

Fronta má přístup na 4 výpočetních uzlů, umístěných na 4 fyzických strojích s celkem 512 CPU.

- **GPU acceleration for significant speedup of calculations**
- 160 nodes, 460 GPU cards (GTX 1080Ti - H100 100GB)
- Requires application with GPU support
- Maximum eight GPU cards on a single node, typically two or four
- Special DGX cluster with eight Nvidia H100 80GB GPU cards
 - Grant competition <https://docs.metacentrum.cz/computing/gpu-comput/dgx/>
- Specific parameters <https://docs.metacentrum.cz/computing/gpu-comput/gpu-job/>
 - `gpu_mem` (minimum amount of memory on the card)
 - `gpu_cap` (a minimal version of GPU architecture)
 - `cuda_version` (version of CUDA installed on the node)

```
qsub -I -l walltime=4:0:0 -l select=1:ncpus=1:ngpus=1:mem=10gb:scratch_local=20gb
```


- Temporary storage on physical computing nodes
- Very intensive operations can cause network overload and the slowdown of central storage (/storage/city/...)
- Copy the input data into the scratch directory on a dedicated machine
- Variable **SCRATCHDIR** is set automatically
- Faster, more stable

```
qsub -l select=1:ncpus=1:mem=4gb:scratch_local=10gb -l walltime=1:00:00
```

```
cp my_input_data.txt $SCRATCHDIR
```

...

```
cp $SCRATCHDIR/my_results.txt /storage/city/home/user_name/
```

```
clean_scratch
```

<https://docs.metacentrum.cz/computing/infrastructure/scratch-storages/>

- Four types of scratch storage

- `scratch_local`

https://wiki.metacentrum.cz/wiki/Scratch_storage

- on every node, HDD, default

- `scratch_ssd`

- fast SSD, typically smaller in volume, not everywhere

- `scratch_shared`

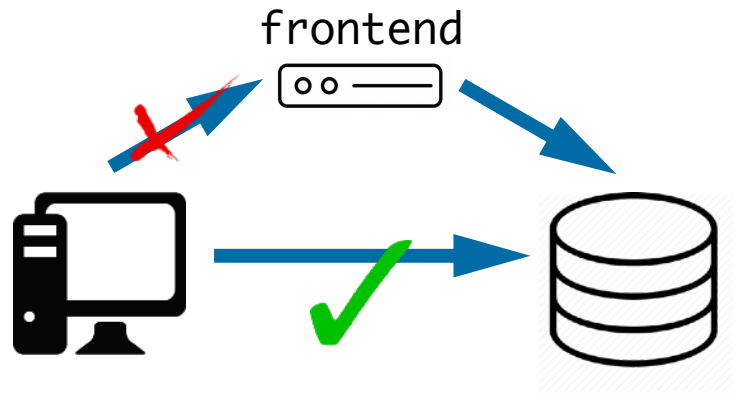
- network volume, which is shared between all nodes of one cluster (only two clusters)

- `scratch_shm`

- scratch held in RAM, very fast, on every node
 - boolean type (True/False), limited by mem parameter (:mem=XYgb)

`scratch_shm= True`

- Do not use frontends, copy data directly on storage and use compressed files (.tar, .zip, .gz, etc.) <https://docs.metacentrum.cz/data/large-data/>
- SFTP client for Windows users (WinSCP, FileZilla, CyberDuck)



```
scp my_data.gz vorel@skirit.metacentrum.cz:\
/storage/praha5-elixir/home/vorel
```

```
scp my_data.gz \
vorel@storage-praha5-elixir.metacentrum.cz:~
```

```
cd $SCRATCHDIR
scp -r storage-praha5-elixir.metacentrum.cz:~/input_data_dir .
...
scp -r output_data_dir storage-praha5-elixir.metacentrum.cz:~
```

SFTP (SSH File Transfer Protocol)

Nickname: storage-brno11 – SFTP

Labels:

URL: <sftp://storage-brno11-elixir.metacentrum.cz>

Server: storage-brno11-elixir.metacentrum.cz

Port: 22

Username: vorel

 Anonymous Login

Password: ●●●●●●●●

SSH Private Key: None

Client Certificate: None

▼ More Options

Path: Folder

Web URL: <http://storage-brno11-elixir.metacentrum.cz/>

Download Folder: Downloads

Transfer Files: Default

Timezone: UTC

Encoding: UTF-8

Connect Mode: Default

Notes:

- Singularity (Apptainer) is an alternative to Docker
- Container system for HPC (non-root access)
- A container is a standard unit of software that packages up code and all its dependencies so the application runs quickly and reliably from one computing environment to another <https://docs.metacentrum.cz/software/containers/>
- Saves time, prevents conflicts between applications
- Every Docker container can be converted to a Singularity image and used in MetaCentrum
- As pre-prepared Singularity images, users can use (e.g.) OpenFOAM, TE-TOOLS (RepeatMasker, RepeatModeler), Peregrine (assembler for long reads), Nvidia GPU cloud (PyTorch, Tensorflow)



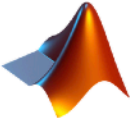











- PyTorch MNIST training with Singularity container in the interactive job in OnDemand
- GPU acceleration
- We will use the PyTorch Singularity image to train a MNIST model (Handwritten digit recognition)

- This example trains a multi-layer RNN (Elman, GRU, or LSTM) or Transformer on a language modeling task. By default, the training script uses the Wikitext-2 dataset, provided
- The trained model can then be used by the generate script to generate new text

MetaCentrum Open OnDemand provides an integrated, single access point for HPC resources.

Selected applications - all apps

 Jupyter Lab/Notebook	 RStudio	 Matlab	 Ansys Fluent	 Ansys Ensignt	 Ansys Workbench
 Frontend shell	 Meta Desktop	 Job Composer	 BIOP Desktop	 CLC Genomics WB	 VMD Desktop

Announcements

23-04-2024

OnDemand has been upgraded to version 3.1.4. Jobs are now submitted to OpenPBS server pbs-m1.metacentrum.cz.

21-08-2023

OnDemand has been upgraded to the major version 3.


```
(BOOKWORM)vorel@galdor20:~$ cd $SCRATCHDIR
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz$ ls
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz$ wget -q https://github.com/pytorch/examples/archive/refs/heads/master.zip
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz$ ls
master.zip
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz$ unzip -q master.zip
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz$ ls
examples-main master.zip
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz$ cd examples-main/word_language_model/
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz/examples-main/word_language_model$ ls -l
total 36
-rw----- 1 vorel meta  2915 Feb  9 17:56 README.md
drwx----- 3 vorel meta   32 Feb  9 17:56 data
-rw----- 1 vorel meta  1482 Feb  9 17:56 data.py
-rw----- 1 vorel meta  3501 Feb  9 17:56 generate.py
-rw----- 1 vorel meta 10646 Feb  9 17:56 main.py
-rw----- 1 vorel meta  5951 Feb  9 17:56 model.py
-rw----- 1 vorel meta    6 Feb  9 17:56 requirements.txt
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz/examples-main/word_language_model$
```

```

<language_model$ singularity exec -nv /cvmfs/singularity.metacentrum.cz/NGC/PyTorch\:24.10-py3.SIF python ./main.py --cuda --epochs 6
| epoch 1 | 200/ 2983 batches | lr 20.00 | ms/batch 7.64 | loss 7.63 | ppl 2049.61
| epoch 1 | 400/ 2983 batches | lr 20.00 | ms/batch 5.09 | loss 6.85 | ppl 944.09
| epoch 1 | 600/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 6.47 | ppl 648.65
| epoch 1 | 800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 6.29 | ppl 537.10
| epoch 1 | 1000/ 2983 batches | lr 20.00 | ms/batch 5.09 | loss 6.14 | ppl 462.79
| epoch 1 | 1200/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 6.06 | ppl 427.53
| epoch 1 | 1400/ 2983 batches | lr 20.00 | ms/batch 5.09 | loss 5.94 | ppl 381.10
| epoch 1 | 1600/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.94 | ppl 380.96
| epoch 1 | 1800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.79 | ppl 327.69
| epoch 1 | 2000/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.77 | ppl 320.86
| epoch 1 | 2200/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.66 | ppl 286.50
| epoch 1 | 2400/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.66 | ppl 288.51
| epoch 1 | 2600/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.65 | ppl 284.25
| epoch 1 | 2800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.54 | ppl 254.13
-----
| end of epoch 1 | time: 16.44s | valid loss 5.53 | valid ppl 252.42
-----
| epoch 2 | 200/ 2983 batches | lr 20.00 | ms/batch 5.12 | loss 5.53 | ppl 252.94
| epoch 2 | 400/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.52 | ppl 249.96
| epoch 2 | 600/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.34 | ppl 208.36
| epoch 2 | 800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.37 | ppl 214.16
| epoch 2 | 1000/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.34 | ppl 208.24
| epoch 2 | 1200/ 2983 batches | lr 20.00 | ms/batch 5.11 | loss 5.33 | ppl 206.37
| epoch 2 | 1400/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.32 | ppl 204.22
| epoch 2 | 1600/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.38 | ppl 216.78
| epoch 2 | 1800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.25 | ppl 190.77
| epoch 2 | 2000/ 2983 batches | lr 20.00 | ms/batch 5.09 | loss 5.26 | ppl 191.89
| epoch 2 | 2200/ 2983 batches | lr 20.00 | ms/batch 5.09 | loss 5.16 | ppl 174.78
| epoch 2 | 2400/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.19 | ppl 179.83
| epoch 2 | 2600/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.21 | ppl 182.76
| epoch 2 | 2800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 5.12 | ppl 167.94
-----
| end of epoch 2 | time: 15.94s | valid loss 5.28 | valid ppl 197.07
-----
| epoch 3 | 200/ 2983 batches | lr 20.00 | ms/batch 5.15 | loss 5.18 | ppl 178.00

```

```

epoch 5 | 2800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.70 | ppl 110.26
-----
| end of epoch 5 | time: 15.95s | valid loss 5.03 | valid ppl 153.00
-----
epoch 6 | 200/ 2983 batches | lr 20.00 | ms/batch 5.15 | loss 4.77 | ppl 117.87
epoch 6 | 400/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.80 | ppl 121.21
epoch 6 | 600/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.61 | ppl 99.99
epoch 6 | 800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.67 | ppl 106.39
epoch 6 | 1000/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.67 | ppl 106.88
epoch 6 | 1200/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.69 | ppl 108.76
epoch 6 | 1400/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.73 | ppl 113.37
epoch 6 | 1600/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.81 | ppl 122.78
epoch 6 | 1800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.68 | ppl 108.21
epoch 6 | 2000/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.71 | ppl 111.55
epoch 6 | 2200/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.62 | ppl 101.44
epoch 6 | 2400/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.66 | ppl 105.25
epoch 6 | 2600/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.70 | ppl 109.54
epoch 6 | 2800/ 2983 batches | lr 20.00 | ms/batch 5.10 | loss 4.63 | ppl 102.29
-----
| end of epoch 6 | time: 15.95s | valid loss 5.02 | valid ppl 151.71
-----

```

```

/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz/examples-main/word_language_model/./main.py:246: FutureWarning
ch uses the default pickle module implicitly. It is possible to construct malicious pickle data which will execute
RITY.md#untrusted-models for more details). In a future release, the default value for `weights_only` will be flipp
rary objects will no longer be allowed to be loaded via this mode unless they are explicitly allowlisted by the use
nly=True` for any use case where you don't have full control of the loaded file. Please open an issue on GitHub for
model = torch.load(f)

```

```

=====
| End of training | test loss 4.95 | test ppl 140.81
=====
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz/examples-main/word_language_model$

```


```
<n/word_language_model$ singularity exec --nv /cvmfs/singularity.metacentrum.cz/NGC/PyTorch\ :24.10-py3.SIF python --cuda
/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz/examples-main/word_language_model/./generate.py:55: FutureWarning: You are using `to
which uses the default pickle module implicitly. It is possible to construct malicious pickle data which will execute arbitrary code duri
SECURITY.md#untrusted-models for more details). In a future release, the default value for `weights_only` will be flipped to `True`. This
bitrary objects will no longer be allowed to be loaded via this mode unless they are explicitly allowlisted by the user via `torch.serial
s_only=True` for any use case where you don't have full control of the loaded file. Please open an issue on GitHub for any issues related
```

```
model = torch.load(f, map_location=device)
| Generated 0/1000 words
| Generated 100/1000 words
| Generated 200/1000 words
| Generated 300/1000 words
| Generated 400/1000 words
| Generated 500/1000 words
| Generated 600/1000 words
| Generated 700/1000 words
| Generated 800/1000 words
| Generated 900/1000 words
(BOOKWORM)vorel@galdor20:/scratch.ssd/vorel/job_9252600.pbs-m1.metacentrum.cz/examples-main/word_language_model$ cat generated.txt
- <unk> , without Hampden then with the repetition . It was numerous rarely subject , particularly as its time
is and prepared to have full Nevermind run , and they are more prominent , <unk> Ishmael , Andy 766th
; <unk> unhealthy , Complete Eyes , The History College Down <unk> , to <unk> - Zoë Boom . avenge
ejaculation from Applegate during Suez , commented that " all pieces steps of frequent strength " , though the group
had temporarily reached a 7 % ceremony wherein it too displayed " get off here " . The kick brings
to Assistant American schools to continue to focus assistance under power . The work immediately received poetic differences in which
patrons published by four groups . <eos> <eos> = = Media = = <eos> <eos> Jackson got it to remake
the development of disrupting radar designer Jane De 1150 , who had since in honor session . During the tour
he enjoyed a quick rate in <unk> to political response by clergymen after the regular struggle , 9 children just
only one individual in the same year polls in Boy <unk> in the fall of Hayes . However , Ímar
underwent two broad dioceses in this line , operating in 1959 due to <unk> and that he deemed it being
the espionage director . Though it appeared in the previous 2002 Cold Age Disney austere , other public legends coming
to live on the basis of the rest date . With amateur , cast small Rather reception , 200 %
of The customer Hoffman conducted owns in series themselves , caught the Kerch Spot . On all a number of
umpires at a large crowd designed to the North Sea , an effort drafts to be of satellite double and
shoots , in do all to the country 's <unk> , but up at his death , Picasso reported that
```


- Galaxy is an open-source system for analysing data, authoring and sharing workflows, training and education, publishing tools, managing infrastructure, and more
- It originally started in biomedical science but nowadays spans numerous scientific domains including ecology, natural language processing, chemistry, climate science, and social sciences
- Web-based platform
- National instance localised at <https://usegalaxy.cz/>
- Download/upload data from/to public various repositories
- **Integration of NRP, OneData, Invenio, Dspace**



<https://docs.metacentrum.cz/related/galaxy/>


Upload










Tools ▼ ☆

GENOMICS TOOLKITS

- Mothur
- Qiime
- QIIME 2
- Picard
- deepTools
- EMBOSS
- NCBI Blast
- MiModD
- HiCEplorer
- Gemini
- Motif Tools
- GATK Tools
- RAD-seq
- Sanger Sequencing
- DNA Metabarcoding
- Apollo

DOMAIN TOOLS

- Virology

 Tools
 Workflows
 Workflow Invocations
 Visualization
 Histories
 History Multiview
 Datasets
 Pages
 Libraries



Welcome to the Czech Galaxy instance of [E-Infra CZ](#) and [ELIXIR CZ](#) hosted at [MetaCentrum](#). We aim to enable accessible, reproducible, and transparent computational research in Czechia. We support thousands of documented and maintained tools that are free to use and backed by a robust public infrastructure.

Please visit our [Documentation](#) for details about using this service. In case of any further questions, please contact us at regalaxy@rt.cesnet.cz.

Have a look at one of our [trainings](#) or you can start with an interactive tour:

[Galaxy UI](#)
[Galaxy History](#)
[Window Manager](#)
[Deferred Datasets](#)

Galaxy is an open platform for supporting data intensive research developed by its many contributors.





Search NCBI ...



NCBI Datasets Taxonomy **Genome** Gene Command-line tools Documentation

Genome assembly E_nip reference

Download

datasets

URL

FTP

Actions

Submitted GenBank assembly	GCA_029291075.1	⋮
Taxon	Eudiplozoon nipponicum	
Isolate	JV_EN_01	
WGS project	JAQBSW01	
Assembly type	haploid	
Submitter	Masaryk University	
Date	Mar 20, 2023	



BLAST the reference genome

https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_029291075.1/

Additional genomes

[Browse all Eudiplozoon nipponicum genomes \(1\)](#)

BioProject

[PRJNA914201](#)

Eudiplozoon nipponicum Genome sequencing and assembly

Publications

Showing 1 of 1

BMC Genomics 2023

[An insight into the functional genomics and species classification of Eudiplozoon nipponicum \(Monogenea, Diplozoidae\), a haematophagous parasite of the common carp Cyprinus carpio](#)

J Vorel, et al.

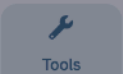
Index of /genomes/all/GCA/029/291/075/GCA_029291075.1_E_nip

Name	Last modified	Size
Parent Directory		-
GCA_029291075.1_E_nip_assembly_report.txt	2024-10-26 22:09	1.7M
GCA_029291075.1_E_nip_assembly_stats.txt	2024-10-26 22:09	3.7K
GCA_029291075.1_E_nip_fcs_report.txt	2025-02-24 01:14	407
GCA_029291075.1_E_nip_feature_count.txt.gz	2023-03-21 04:43	149
GCA_029291075.1_E_nip_genomic.fna.gz	2023-03-21 04:44	281M
GCA_029291075.1_E_nip_genomic.gbff.gz	2023-07-02 01:58	381M
GCA_029291075.1_E_nip_wgsmaster.gbff.gz	2023-07-02 01:58	1.4K
README.txt	2024-08-27 13:56	55K
annotation_hashes.txt	2023-07-02 01:58	410
assembly_status.txt	2025-03-03 18:54	14
md5checksums.txt	2024-10-26 22:11	588
uncompressed_checksums.txt	2024-09-01 02:54	202

[HHS Vulnerability Disclosure](#)



Upload



Tools



Workflows

Workflow
Invocations

Visualization



Histories

History
Multiview

Datasets



Tools

stat

Show Sections

Statistical hypothesis testing computes several descriptive statistics

Statistics on presence-absence of a numeric variable

MAF Coverage Stats Alignment coverage information

Summary Statistics for any numerical column

FASTA-to-Tabular converter

FASTQ Summary Statistics by column

FASTA-to-Tabular converter

Fasta Statistics display summary statistics for a FASTA file

gfastats the swiss army knife for genome assembly

SeqKit statistics of FASTA/Q files

Compute quality statistics

Upload from Disk or Web to **Unnamed history**

Regular

Composite

Collection

Rule-based

You added 1 file(s) to the queue. Add more files or click 'Start' to proceed.



New File

115 b

Auto-detect



unspecified (?)



0%



Download data from the web by entering URLs (one per line) or directly paste content.

```
https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/029/291/075/GCA_029291075.1_E_nip/GCA_029291075.1_E_nip_genomic.fna.gz
```



Type (set all): Auto-detect



Reference (set all): unspecified (?)



Choose local file



Choose remote files



Paste/Fetch data

Start

Pause

Reset

Close

Galaxy

Tools

stat

Metastats generate principle components plot data

obistat computes basic statistics for attribute values

MSstatsTMT protein significance analysis in shotgun mass spectrometry-based proteomic experiments with tandem mass tag (TMT) labeling

MSstats statistical relative protein significance analysis in DDA, SRM and DIA Mass Spectrometry

metaQuantome: stat differential analysis of functional expression and taxonomic abundance

Remove protonation state of every atom

MapStatistics Extract extended statistics on the features of a map for quality control.

MSstatsConverter Converter to input for MSstats

ConvertFastaToPrositCSV Create Prosit CSV Input From a Protein FASTA

FASTA or M

accepted fo

FASTA data

Estimated

This param

Generate g

No

(-gaps_opt

Addi

Email notifi

No

Send an em

Run Tool

Help

Using 0% of 200.0 GB

voret

History

search datasets

Unnamed history

295 MB

271

272: GCA_029291075.1_E_nip_ge

omic.fna.gz

Upload from Disk or Web to Unnamed history

Regular Composite Collection Rule-based

New File 115 b Auto-detect unspecified (?) 100%

Download data from the web by entering URLs (one per line) or directly paste content.

https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/029/291/075/GCA_029291075.1_E_nip/GCA_029291075.1_E_nip_genomic.fna.gz

Type (set all): Auto-detect Reference (set all): unspecified (?)

Choose local file Choose remote files Paste/Fetch data Start Pause Res Close



Using 0% of 200.0 GB

vorel



This dataset is large and only the first megabyte is shown below.

Show all | Save

>JAQBSW010000001.1 Eudiplozoon nipponicum isolate JV_EN_01 E_nip_1, whole genome shotgun sequence

gcaatgttctctcattcgagtaattaaccgctggtcgatcggcctacgtaagttgatcgacaacttgcgatgtatcct
ggtaaaaaatatttgatattgacattaacgactgtttcctttgattactggtaaggcaagtggttacacataataac
actcgtagtttagctgcatataatgctgtttaaagctgcttcaactttcttttggtcattcgggaaactgcacccac
cataaatgggatgtttaggtagaccgttttctttctggccgtaataagtcgtttctgtttgattcctcttttaataat
tgattcaattaatagctTAGGATaccattagacattaatcgtaattgctttattatattaattttctttagaatatt
gactaaccaacctgtcaattcttaaaaagcaaccaggaactgttgataatttatattgaaatggtacaaaaactgctgat
aacaataactgtgcagaccatggtggttatggtaaatgttctgtttaattacatattgattttttcattaatagtata
tctaggaaaggcagacagtcattcttctcgtgctcgtcgtgaagcgcgatattcttatgtgctttgtgaaacttaag
taatggtcgtgattgcgacctgttcttaagtacaagaatatatcaccacgtacgtatattgagcacatttaatta
tgtattttctagggtccctagaaaaatctcagcaaacctggagataaggggcttccattgcgacaccatttcattaat
tactttgctccaccttctgctatttcatctgatttatcgtctagacgtggaattttgtattgcatttaacacgttttc
catcttttgtagtgcagatttgcattataagtgctcctttctttgcttctgtacttaagaacaggtcgcaa
tcacgaacattgcttcaaaagttcaacaagcacataagaatatcggcttcacgctcggagcacgagaagaatgactgctc
gcctttcctagatgactattaaagaagaactaatggtaacttaaacagaacaattaccataaaccaacatggctc
cacagatttgcataaacagcagtttggctgctatttaaacataaaatatacaacagttcgtgggtgcttttcaagaatagc
aggttggttagtcagaattctaaagaaaaatgaatataataaagcaattaacggtttaatgctaatggctatcccagaca
attaattgaatcaattataaaaggggagtaaacagaaacagacgtattatgggcagaaaaagaaaatagcttacctaa
acataaccatttaaggggagtcagctcctccgaatggcacaagaagaagatgaagcagcattgacacacacatttaagca
gccaaactacagtgattatttaggtaacacacttgcctttactagcaatcaagagaaacagcgttaattgctgatatc
aaatattatttaccataatgcatgctgcatgttggctgatcaacttacgtaggcgcgacgcagacagggtaatttactcga
tgagagaacattgcaacttaaacgccgttttaaatcaaatggtgaatacacgctgctatttgcgcaacacatcatcgaa
gcatcgcacgtatgttcttcagacgtagcattcaagtaattactaaatgcaaaaacaaaggagctagccacattatgg

History

search datasets

Unnamed history

295 MB

1 271

272: GCA_029291075.1_E_nip_g



omic.fna.gz

Add Tags






21,044 sequences

format fasta.gz, database ?

uploaded fasta.gz file




>JAQBSW010000001.1 Eudiplozoon nipponicum is
gcaatgttctctcattcgagtaattaaccgctggtcgatcggc
ggtaaaaaatatttgatattgacattaacgactgtttcctttg
actcgtagtttagctgcatataatgctgttaatgctgcttcc
cataaatgggatgtttaggtagaccgttttctttctggccgta

 Upload
 Tools
 Workflows
 Workflow Invocations
 Visualization

Tools


✕

 Show Sections

Fasta Statistics display summary statistics for a FASTA file

PfamScan search a FASTA sequence against a library of Pfam HMM

CustomProDB Generate protein FASTAs from exosome or transcriptome data

 This dataset is large and only the first megabyte is shown below.
[Show all](#) | [Save](#)

```

>JAQBSW01000001.1 Eudiplozoon nipponicum isolate JV_EN_01 E_nip_1, whole genome shotgun sequence
gcaatgttctctcattcgagtaattaaccgtctggatcgccctacgtaagttgatcgacaacttgcgcatgtatctt
ggtaaaaaatatttgatattgacattaacgactgtttctctttgattactggtaaaggcaagtgtttacactaaataac
actcgtagtttagctgcattaaatatgctgttaaatgctgcttcaactttcttttggtgcattcgggaaactgcatccac
cataaatgggatgttttagtagaccgtttctttctggccgtaataagtctgtttctgtttgattcctcttttaataat
tgattcaattaatagtCTAGGATaccattagacattaaatcgtaattgctttattatattaattttctttagaatatt
gactaaccaactgtcaattcttaaaaagcaaccacgaactgttgataatttatattgaaatggtacaaaactgctgtat
aacaatactgtgcagaccatgttggtttatggtaaatgttctgttttaattaccattagttttttcattaatagtata
tctaggaaggcagacagtcatctttctcgtgctctgtcgtgaagccgatattcttatgtgctttgttgaactttaaag
taatgttctgtgattgcgacctgttcttaagtacaagaaatataatcatccacgtaccgtatatattgagcacatttaata
tgtattttctaggtgccctagaaaaatctcagccaactggagataaggggcttcccattgcgacaccatttcattaat
tactttgcgctccacctgtctatttcatctgatttatcgtctagacgtgtgaattttgtattgcatttaacacgttttc
catcttttatacataaqtcaqatattatccattataaqtacactctttcctttacattctttatcttaaaacaagtcaaa
    
```

Tool Parameters

FASTA or Multi-FASTA file *

272: GCA_029291075.1_E_nip_genomic.fna.gz (as fasta)

FASTA dataset to get statistics. (--fasta)

Estimated genome size - optional

This parameter is optional. If provided, it will be used for calculating the NG50 statistic. (--genome_size)

Generate gap stats

No

(--gaps_option)

Additional Options

Email notification

No

Send an email notification when the job completes.

Run Tool

Help

i Purpose

Displays the summary statistics for a FASTA file.

i Outputs

This tool generates two outputs: a general summary and an optional gap stats file.

The general summary includes the following information:

- Lengths: n50, min, max, median and average
- Number of base pairs: A, C, G, T, N, Total and Total_not_N
- Number of sequences
- GC content

In addition the optional gap stats BED file includes the information about gaps localization.

History

search datasets

Unnamed history

295 MB 271

272: GCA_029291075.1_E_nip_genomic.fna.gz

Add Tags

21,044 sequences

format fasta.gz, database ?

uploaded fasta.gz file

Share

```
>JAJQBSM010000001.1 Eudiplozoon nipponicum iso  
gcaatgttctctcattcgagttaataaccgctcggtcgatcggc  
ggtataaaaatattgatattgacattaacgactgtttctctttg  
actcgtagtttagctgcattaaatgctgttaatgctgcttc  
cataaatgggatgttaggtagaccgtttctttctggcctta
```



Started tool **Fasta Statistics** and successfully added 1 job to the queue.

It produces this output:

- **273: Fasta Statistics on data 272: summary stats**

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

History





Unnamed history



295 MB

2

271

1



🕒 273: **Fasta Statistics on data 272: summary stats** 👁️ ✎️ 🗑️

Add Tags 🏷️

This is a new dataset and not all of its data are available yet.

format **tabular**, database ?

272: **GCA_029291075.1_E_nip_ genomic.fna.gz** ⇄ 👁️ ✎️ 🗑️

Add Tags 🏷️

21,044 sequences

format **fasta.gz**, database ?

uploaded fasta.gz file



```
>JAQBSW01000001.1 Eudiplozoon nipponicum isc
gcaatgttctctcattcgagtaattaaccgtctggtcgatcggc
ggtaaaaaatatttgatattgacattaacgactgtttctctttg
actcgtagtttagctgcattaatatgcgtgtaaatgctgcttc
cataaatgggatgttaggtagaccgtttctttctggcctta
```



Started tool **Fasta Statistics** and successfully added 1 job to the queue.

It produces this output:

- **273: Fasta Statistics on data 272: summary stats**

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

History



Unnamed history

1.25 GB



273: **Fasta Statistics on data 272**
: summary stats

Add Tags

30 lines 2 columns

format **tabular**, database ?

1	2
Scaffold L50	3219
Scaffold N50	87067
Scaffold L90	11600
Scaffold N90	19508
Scaffold len_max	557136

272: **GCA_029291075.1_E_nip_**
genomic.fna.gz

Add Tags

21,044 sequences

format **fasta.gz**, database ?

uploaded fasta.gz file



```
>JAQBSW010000001.1 Eudiplozoon nipponicum iso  
gcaatgttctctcattcgagtaattaaccgtctggtcgcgatcggc  
ggtaaaaaatatttgatattgacattaacgactgtttctctttgc  
actcgtagtttagctgcattaaatcgctgttaatgctgcttcc  
cataaatgggatgttagtagaccgttttctttctggcctac
```

Column 1	Column 2
Scaffold L50	3219
Scaffold N50	87067
Scaffold L90	11600
Scaffold N90	19508
Scaffold len_max	557136
Scaffold len_min	4973
Scaffold len_mean	44658
Scaffold len_median	23894
Scaffold len_std	52885
Scaffold num_A	305506036
Scaffold num_T	305603756
Scaffold num_C	164346741
Scaffold num_G	164346396
Scaffold num_N	0
Scaffold num_bp	939802929
Scaffold num_bp_not_N	939802929
Scaffold num_seq	21044
Scaffold GC content overall	34.97
Contig L50	3219
Contig N50	87067
Contig L90	11600
Contig N90	19508
Contig len_max	557136
Contig len_min	4973
Contig len_mean	44658
Contig len_median	23894
Contig len_std	52885
Contig num_bp	939802929
Contig num_seq	21044
Number of gaps	0

History



search datasets



Unnamed history



1.25 GB

2

271

1



273: Fasta Statistics on data 272



: summary stats

Add Tags

30 lines 2 columns

format tabular, database ?



1	2
Scaffold L50	3219
Scaffold N50	87067
Scaffold L90	11600
Scaffold N90	19508
Scaffold len_max	557136

272: GCA_029291075.1_E_nip_1



genomic.fna.gz

cesnet
metacentrum
.....



meta@cesnet.cz vorel@cesnet.cz

