Leibniz-Rechenzentrum
der Bayerischen Akademie der Wissenschaften

LRZ Introduction to Linux Cluster and Compute Cloud

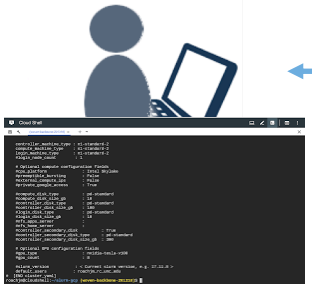
02.07.2019 | Dr. Shaila Rössle-Blank

Agenda

- 10:00 - 11:00 Introduction to LRZ hardware
- 11:00 - 11:15 Break
- 11:15 - 13:00 Linux Cluster Hands-On
  - login
  - module system
  - slurm
  - Conda environment
  - Jupyter notebook
- 13:00 - 13:30 Lunch
- 13:30 - 15:00: Compute Cloud Hands-On
  - How to create a VM in the LRZ Computing Cloud
  - Conda environment
  - Examples

Important

- https://doku.lrz.de
- https://doku.lrz.de/display/PUBLIC/High+Performance+Computing



LRZ account

ssh lxlogin#.lrz.de
(5,6,7,8,10)

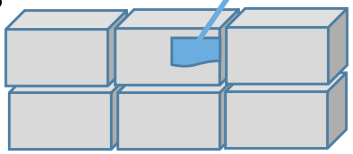
Command line

```
#!/bin/bash
## a job which uses 1 core of a node
#SBATCH -o /home/hpc/.../.../myjob.%j.%N.out
#SBATCH -D /home/hpc/.../.../mydir
#SBATCH -J Jobname
#SBATCH --get-user-env
#SBATCH --clusters=serial
#SBATCH --partition=serial_mpp2
#SBATCH --mail-type=end
#SBATCH --mem=800mb
#SBATCH --cpus_per_task=1
#SBATCH --mail-user=xyz@xyz.de
#SBATCH --export=NONE
#SBATCH --time=24:00:00

source /etc/profile.d/modules.sh
cd mydir
./myprog.exe
```

Batch script

Commands for code execution, input directory, specify number/type of nodes, length of run, output directory, module, partition,...

Login Nodes
Each login node is shared by many users

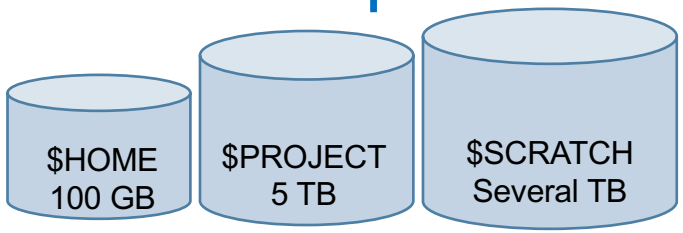


file editing, data backup, job submission

Run jobs by submitting your batch script to the compute nodes using SLURM (sbatch)

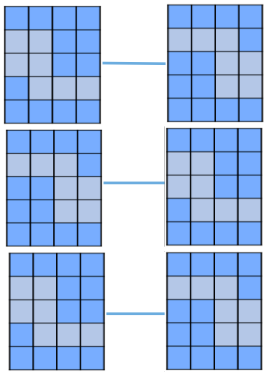
- Teramem – single node with 6TB main memory
- CoolMUC-2 – 28 nodes (serial and parallel processing)
- CoolMUC-3 – 64 many-core processors (parallel and vector processing)
- IvyMUC - 8-way Ivy Bridge-based nodes (parallel processing)

It is the responsibility of the user to backup important data.
$HOME- snapshot and tape backup



Shared File System

Fast connection



Compute Nodes

The job is submitted to a queue and will wait in line until nodes are availabe. Queues are managed by SLURM (job scheduler) that enables jobs to run efficiently.

Access to the Linux Cluster

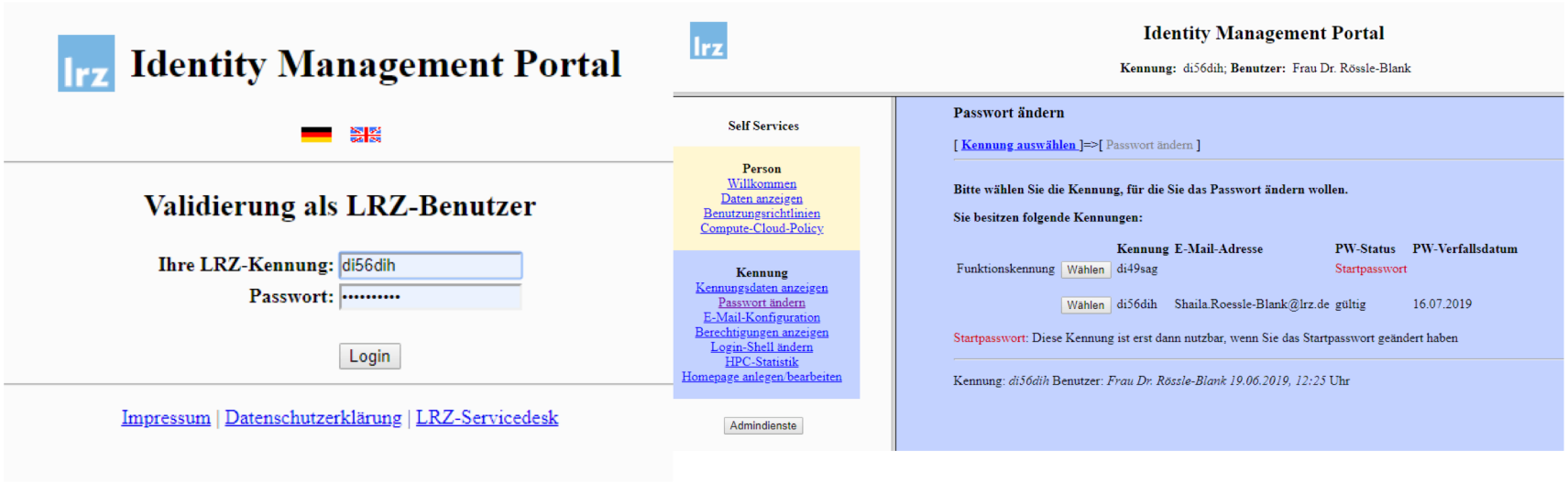
https://doku.lrz.de/display/PUBLIC/Linux+Cluster

Get an account (Scientists and students from Munich Universities as well as Bavarian Universities)

1. Contact the responsible **master user** located at your institution. If you don't know who your master user is, please contact the head of your department (group, institute), who will be able to point you to your master user.
2. Your master user can create a new LRZ account for you through the LRZ Identity Management Portal.
3. You will receive an email from the LRZ user support that your account has been activated.
4. Access to the LRZ Linux-Cluster will be possible within one day.

Change password

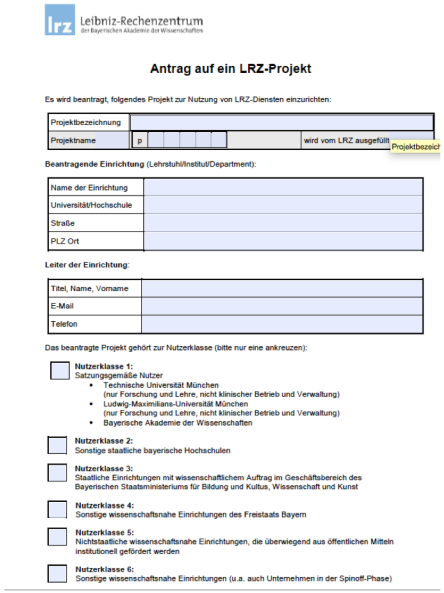
https://idmportal.lrz.de/r/entry.pl



Access to the Linux Cluster

https://doku.lrz.de/display/PUBLIC/Linux+Cluster

If your institute or group has **no master user**, your group can apply for a new LRZ project. You have to fill out this two page PDF (only available in German): Antrag auf ein LRZ Projekt.



Leibniz-Rechenzentrum
der Bayerischen Akademie der Wissenschaften

Antrag auf ein LRZ-Projekt

Es wird beantragt, folgendes Projekt zur Nutzung von LRZ-Diensten einzurichten:

Projektbezeichnung	
Projektname	p — wird vom LRZ ausgefüllt

Projektbezeich

Beantragende Einrichtung (Lehrstuhl/Institut/Department):

Name der Einrichtung	
Universität/Hochschule	
Straße	
PLZ Ort	

Leiter der Einrichtung:

Titel, Name, Vorname	
E-Mail	
Telefon	

Das beantragte Projekt gehört zur Nutzerklasse (bitte nur eine ankreuzen):

☐ **Nutzerklasse 1:**
Satzungsgemäße Nutzer
- Technische Universität München
(nur Forschung und Lehre, nicht klinischer Betrieb und Verwaltung)
- Ludwig-Maximilians-Universität München
(nur Forschung und Lehre, nicht klinischer Betrieb und Verwaltung)
- Bayerische Akademie der Wissenschaften

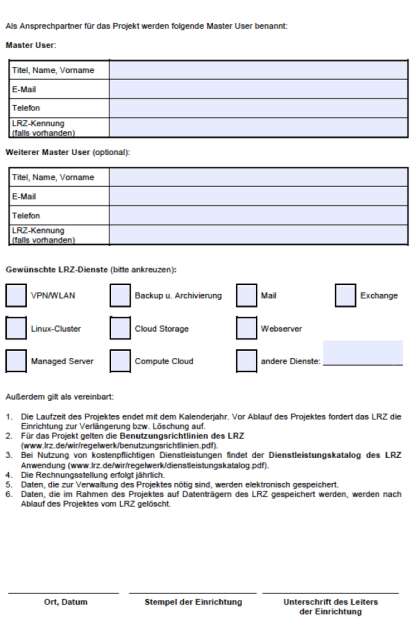
☐ **Nutzerklasse 2:**
Sonstige staatliche bayerische Hochschulen

☐ **Nutzerklasse 3:**
Staatliche Einrichtungen mit wissenschaftlichem Auftrag im Geschäftsbereich des Bayerischen Staatsministeriums für Bildung und Kultus, Wissenschaft und Kunst

☐ **Nutzerklasse 4:**
Sonstige wissenschaftsnahe Einrichtungen des Freistaats Bayern

☐ **Nutzerklasse 5:**
Nichtstaatliche wissenschaftsnahe Einrichtungen, die überwiegend aus öffentlichen Mitteln institutionell gefördert werden

☐ **Nutzerklasse 6:**
Sonstige wissenschaftsnahe Einrichtungen (u.a. auch Unternehmen in der Spinoff-Phase)



Als Ansprechpartner für das Projekt werden folgende Master User benannt:

Master User:

Titel, Name, Vorname	
E-Mail	
Telefon	
LRZ-Kennung (falls vorhanden)	

Weiterer Master User (optional):

Titel, Name, Vorname	
E-Mail	
Telefon	
LRZ-Kennung (falls vorhanden)	

Gewünschte LRZ-Dienste (bitte ankreuzen):

☐ VPN/WLAN ☐ Backup u. Archivierung ☐ Mail ☐ Exchange
☐ Linux-Cluster ☐ Cloud Storage ☐ Webserver
☐ Managed Server ☐ Compute Cloud ☐ andere Dienste: _____

Außerdem gilt als vereinbart:

1. Die Laufzeit des Projektes endet mit dem Kalenderjahr. Vor Ablauf des Projektes fordert das LRZ die Einrichtung zur Verlängerung bzw. Löschung auf.
2. Für das Projekt gelten die **Benutzungsrichtlinien des LRZ** (www.lrz.de/wir/regelwerk/benutzungsrichtlinien.pdf).
3. Bei Nutzung von kostenpflichtigen Dienstleistungen findet der **Dienstleistungskatalog des LRZ** Anwendung (www.lrz.de/wir/regelwerk/dienstleistungskatalog.pdf).
4. Die Rechnungsstellung erfolgt jährlich.
5. Daten, die zur Verwaltung des Projektes nötig sind, werden elektronisch gespeichert.
6. Daten, die im Rahmen des Projektes auf Datenträgern des LRZ gespeichert werden, werden nach Ablauf des Projektes vom LRZ gelöscht.

Ort, Datum | **Stempel der Einrichtung** | **Unterschrift des Leiters der Einrichtung**

Access to the Linux Cluster

https://doku.lrz.de/display/PUBLIC/Linux+Cluster

By getting your account, a HOME directory and a HOME contingent are automatically created (with identifier contingent = project contingent).

Current default quotas for a project are:

- $HOME: 100 GB per project
- $WORK: 1000 GB per project
- $SCRATCH: several Tbyte (temp)

All $HOME, $PROJECT and $SCRATCH file system mounts have the general structure

/<mountpoint>/<group>/<user>

It is the responsibility of the user to store important data.

$HOME- snapshot and tape backup

Login

https://doku.lrz.de/display/PUBLIC/Linux+Cluster

ssh -Y lxlogin5.lrz.de -l *userID*	Haswell (CoolMUC-2) login node
ssh -Y lxlogin6.lrz.de -l *userID*	Haswell (CoolMUC-2) login node
ssh -Y lxlogin7.lrz.de -l *userID*	Haswell (CoolMUC-2) login node
ssh -Y lxlogin8.lrz.de -l *userID*	KNL Segment (CooMUC-3) login node
ssh -Y lxlogin10.lrz.de -l *userID*	Ivy Bridge (IvyMUC) login node

No password

https://doku.lrz.de/display/PUBLIC/ssh+-+Secure+Shell+on+LRZ+HPC+Systems

Hands-on: Using Packages

Open a Terminal → Debian GNU/Linux

$ sudo su

$ apt update

$ apt install ssh

$ apt install python3

Using a package in your Computer

$ python3.6

Using a package in the Linux Cluster

$ export DISPLAY=localhost:0

$ ssh –Y lxlogin6.lrz.de

$ xeyes

$ python3.6 → ERROR

$ module available

$ module available python

$ module load python/3.6_intel

$ python3.6

Modules

https://confluence.lrz.de/display/PUBLICDRAFT/Environment+Modules

Environment Modules, or short **Modules** are the means by which most of the installed scientific software is made available on HLR systems.

The LRZ Linux Cluster provide users with the possibility to load and unload complete environments for compilers, libraries and software packages by a single command.

The distinct advantage of the modules approach is that the user is no longer required to explicitly specify paths for different software versions nor need to try to keep the related environment variables coordinated. With the modules approach, users simply "load" and "unload" modules to control their environment.

Module

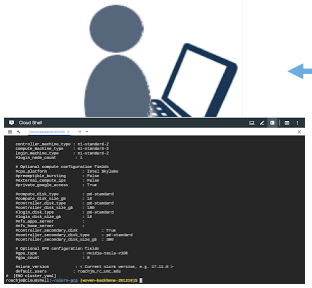
Use the **module** command to manage modules.

- **module list:** This lists all the modules which are currently loaded into the users' environment, in the order in which they were loaded.
- **module available:** This option lists all the modules which are available to be loaded. Notice that most have version numbers associated with them.
- **module show:** This option requires a module name as an argument. It will display the changes to the environment that would be done if the module were loaded.
- **module load:** This option will make the version of the package available for your use. Without version number, it lownloads the default package
- **module unload: T**he loaded version will be removed even if not explicitly specified the version number

Hands-on: Module

```
$ module list

$ module available cellranger

$ module show cellranger/2.2.0

$ module load cellranger/2.2.0

$ module list

$ module unload cellranger
```

Will also load the bcl2fastq/2.20



LRZ account

ssh lxlogin#.lrz.de
(5,6,7,8,10)

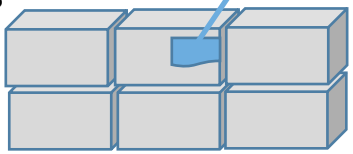
Command line

```
#!/bin/bash
## a job which uses 1 core of a node
#SBATCH -o /home/hpc/.../.../myjob.%j.%N.out
#SBATCH -D /home/hpc/.../.../mydir
#SBATCH -J Jobname
#SBATCH --get-user-env
#SBATCH --clusters=serial
#SBATCH --partition=serial_mpp2
#SBATCH --mail-type=end
#SBATCH --mem=800mb
#SBATCH --cpus_per_task=1
#SBATCH --mail-user=xyz@xyz.de
#SBATCH --export=NONE
#SBATCH --time=24:00:00

source /etc/profile.d/modules.sh
cd mydir
./myprog.exe
```

Batch script

Commands for code execution, input directory, specify number/type of nodes, length of run, output directory, module, partition,...

Login Nodes
Each login node is shared by many users

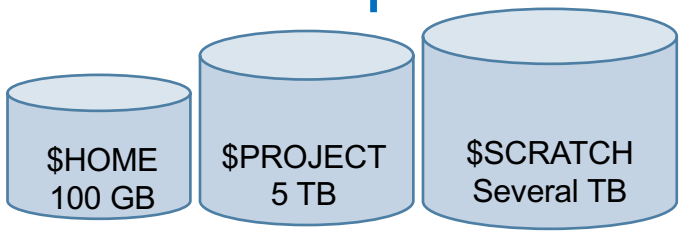


file editing, data backup, job submission

Run jobs by submitting your batch script to the compute nodes using SLURM (sbatch)

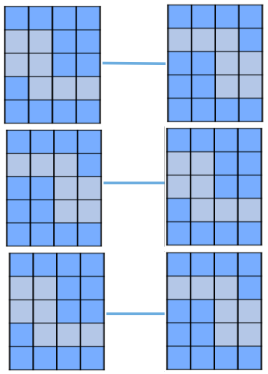
- Teramem – single node with 6TB main memory
- CoolMUC-2 – 28 nodes (serial and parallel processing)
- CoolMUC-3 – 64 many-core processors (parallel and vector processing)
- IvyMUC - 8-way Ivy Bridge-based nodes (parallel processing)

It is the responsibility of the user to backup important data.
$HOME- snapshot and tape backup



Shared File System

Fast connection



Compute Nodes

The job is submitted to a queue and will wait in line until nodes are availabe. Queues are managed by SLURM (job scheduler) that enables jobs to run efficiently.

SLURM: Batch Scheduling System

- The LRZ Linux Clusters use a batch scheduling system called *SLURM* to handle the queuing, scheduling, and execution of jobs.

- Users generally submit jobs by writing a job script file and submitting the job to Slurm with the sbatch command.

- The sbatch command takes a number of options (some of which can be omitted or defaulted). These options define various requirements of the job, which are used by the scheduler to figure out what is needed to run your job, and to schedule it to run as soon as possible, subject to the constraints on the system, usage policies, and considering the other users of the cluster.

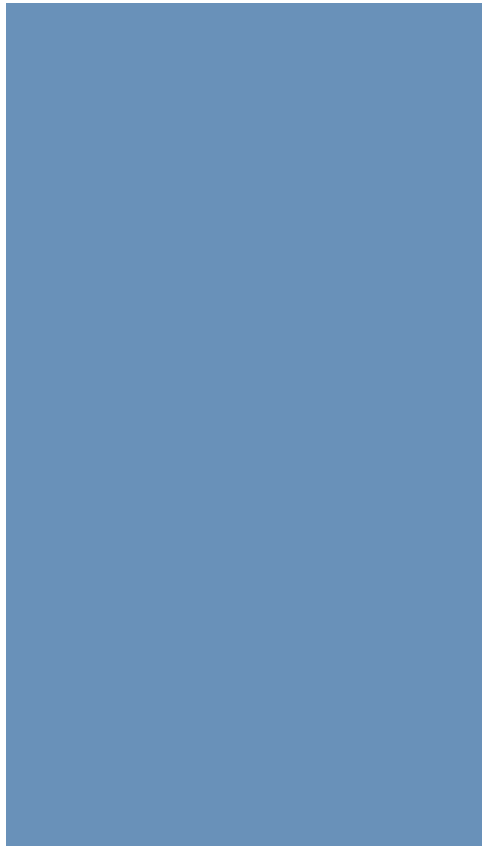
- It is also possible to submit an interactive job, but that is usually most useful for debugging purposes.

Creating a Job Script

https://doku.lrz.de/display/PUBLIC/Running+serial+jobs+on+the+Linux-Cluster

- One option for running a job on the HPC cluster is to set up a job script. This script will request cluster resources and list, in sequence, the commands that you want to execute. A job script is a plain text file that can be edited with a UNIX editor such as vi, nano, or emacs.

- To properly configure a job script, you will need to know the general script format, the commands you wish to use, how to request the resources required for the job to run, and, possibly, some of the Slurm environmental variables.



SLURM Commands

https://doku.lrz.de/display/PUBLIC/SLURM+Workload+Manager

command	description
sacct	report job accounting information about active or completed jobs
salloc	allocate resources for a job in real time (typically used to allocate resources and spawn a shell, in which the **srun** command is used to launch parallel tasks)
sbatch	submit a job script for later execution (the script typically contains one or more **srun** commands to launch parallel tasks)
scancel	cancel a pending or running job
sinfo	reports the state of partitions and nodes managed by Slurm (it has a variety of filtering, sorting, and formatting options)
squeue	reports the state of jobs (it has a variety of filtering, sorting, and formatting options), by default, reports the running jobs in priority order followed by the pending jobs in priority order
srun	used to submit a job for execution in real time

SLURM Job Script Structure

https://doku.lrz.de/display/PUBLIC/SLURM+Workload+Manager

```
#!/bin/bash                                          # interpreter

#SBATCH –J JobName
#SBATCH -D /.../projectID/userID/dirName
#SBATCH -o /home/.../hpc/.../myjob.%j.out            # name of the output
#SBATCH --get-user-env                               # set user environment properly
#SBATCH --mail-type=end
#SBATCH --mail-user=userEmail                        # email to receive notifications
#SBATCH --export=NONE

#SBATCH --clusters=ivymuc / serial /inter
#SBATCH --partition=ivymuc_batch / serial_batch / mpp2_inter
#SBATCH --nodes=1
#SBATCH --cpus-per-task=1
#SBATCH --mem=up to 57000mb      # maximum memory the job can be used when cluster serial
#SBATCH --time=24:00:00          # estimated run time

source /etc/profile.d/modules.sh                     #initialize the module system
module use -a /lrz/sys/share/modules/extfiles        #see in module available package
module load package

./myprog.exe
```

- Interpreter
- Job/user information
- Specifying Resources
- Job commands
  - Module system
  - Executable

How to make/run a script

- Cell Ranger
- https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
- https://doku.lrz.de/display/PUBLIC/Cell+Ranger

- NAMD
- https://www.ks.uiuc.edu/Research/namd/2.13/ug/
- https://doku.lrz.de/display/PUBLIC/NAMD

- LAMMPS
- https://lammps.sandia.gov/doc/Manual.html
- https://doku.lrz.de/display/PUBLIC/lammps
- https://rcc.uchicago.edu/docs/software/applications/lammps/index.html

Hands-on: sbatch

cd /lrz/sys/courses/LRZ_Intro>

```
#!/bin/bash
#SBATCH -J script.sh
#SBATCH -D /home/hpc/pr28fa/di56dih/testrun/blast_test
#SBATCH -o /home/hpc/pr28fa/di56dih/testrun/blast_test/blast.%j.out
#SBATCH --get-user-env
#SBATCH --mail-type=end
#SBATCH --mail-user=di56dih@lrz.de
#SBATCH --export=NONE
#SBATCH --clusters=ivymuc
#SBATCH --nodes=1
#SBATCH --time=08:00:00
source /etc/profile.d/modules.sh
module use -a /lrz/sys/share/modules/extfiles
module load blast

blastp -query /home/hpc/pr28fa/di56dih/testrun/blast_test/blast.input
 -db /home/hpc/pr28fa/di56dih/testrun/blast_test/zebrafish
-out /home/hpc/pr28fa/di56dih/testrun/blast_test/results.txt
```

Hands-on: Job submission

- cd course
- ls
- vi script.sh
- sbatch script.sh # The job script is submitted to the queue.
- squeue -M *clusterName* -u userID # Prints the list of current jobs.
- sacct –M *clusterName* –j *JOBID* –o jobid,partition,user,start,end,elapsed,maxrss

Conda

Conda is an environment management for any language—*Python*, R, Ruby, Lua, Scala, Java, JavaScript, C/ C++, FORTRAN.

Advantages

- Install a package yourself without waiting for an administrator to install all the dependencies and create a module for your program.

- Conda is very simple to use but gives you administrator-like power, allowing you to install packages that would normally require elevated privileges.

- To use Conda you must first create

a new Conda environment.

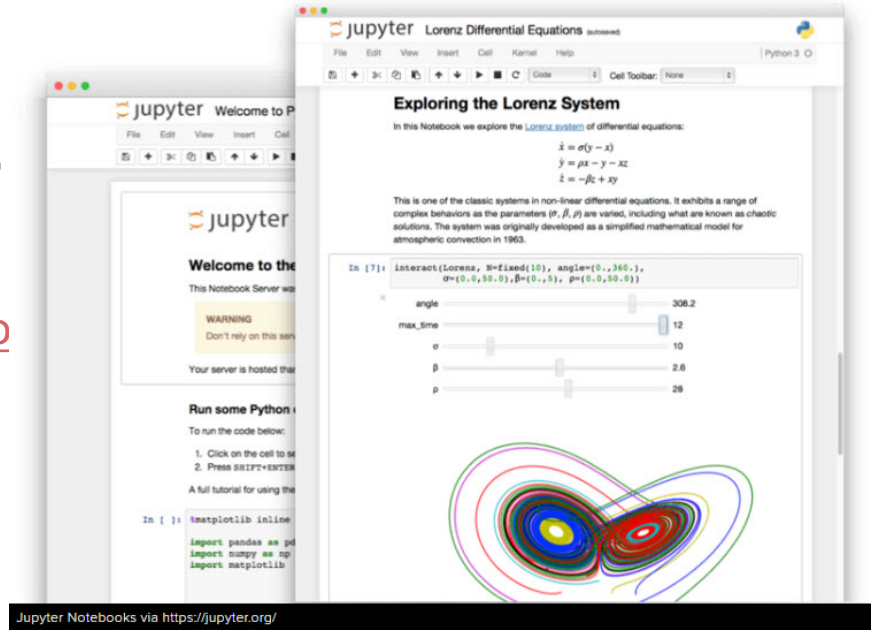
  - conda create -n <environment_name>

Jupyter Notebook

http://doku.lrz.de/display/PUBLIC/Jupyter+Notebook+on+the+Linux+Cluster

- Jupyter notebook allows a user to interactively code, execute and visualize programs inside of a web browser. While the notebook supports multiple languages, most often it is used with Python.
- It is a web-service where you can run any code through a browser interface
- To use Jupyter Notebook in the LRZ Linux Cluster, instructions on:

  https://doku.lrz.de/display/PUBLIC/Jupyter+Notebook+o



Hands-on: Jupyter notebook

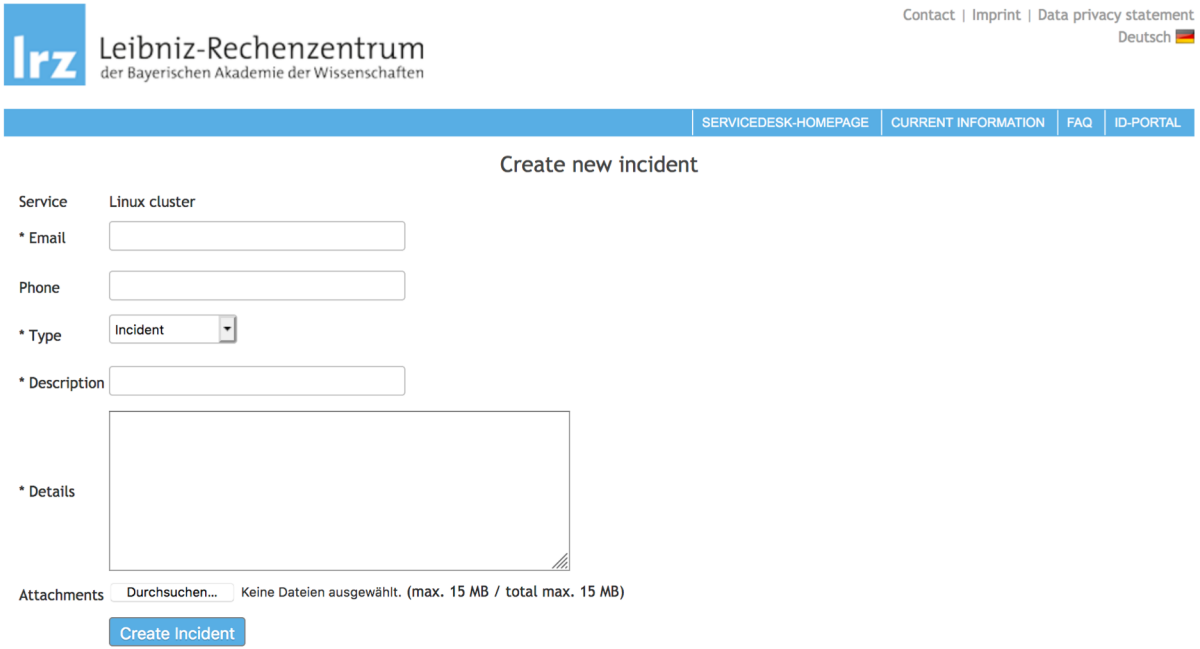
- Go to
  - *https://doku.lrz.de/display/PUBLIC/Jupyter+Notebook+on+the+Linux+Cluster*
- Login
  - *qsh –Y lxlogin8.lrz.de*
  - *cd LinuxCourse*
- Create a Conda environment
- Open the example

Ticket

https://doku.lrz.de/display/PUBLIC/Support+for+HPC+and+Big+Data

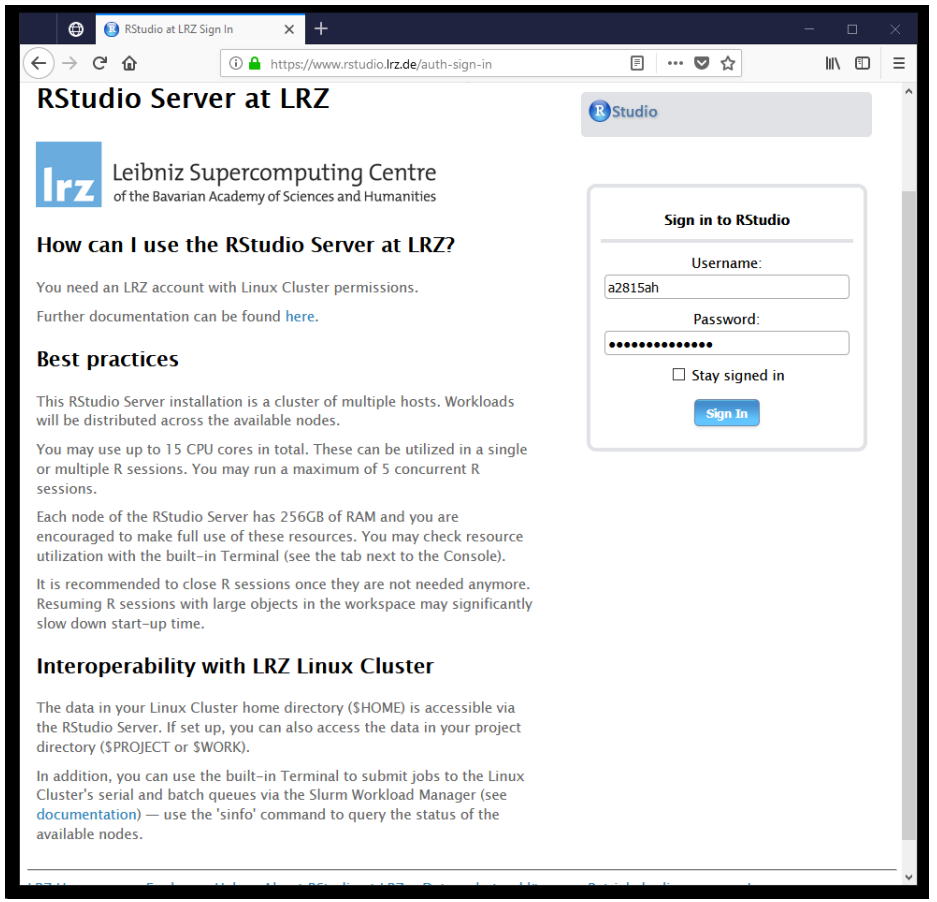
If you have a specific service request, complain or an incident on a specific LRZ system please click to submit an incident to the service desk:

https://servicedesk.lrz.de/en/simplesubmit/23



RStudio

https://www.rstudio.lrz.de/auth-sign-in

- up to 15 CPU cores in total
- single or multiple R sessions
- maximum of 5 concurrent R sessions
- max 256 GB RAM
- $HOME, $PROJECT or $WORK

It is recommended to close R sessions once they are not needed anymore.