



PP&B – Computer Support Group

FRITZ-HABER-INSTITUT  
MAX-PLANCK-GESELLSCHAFT



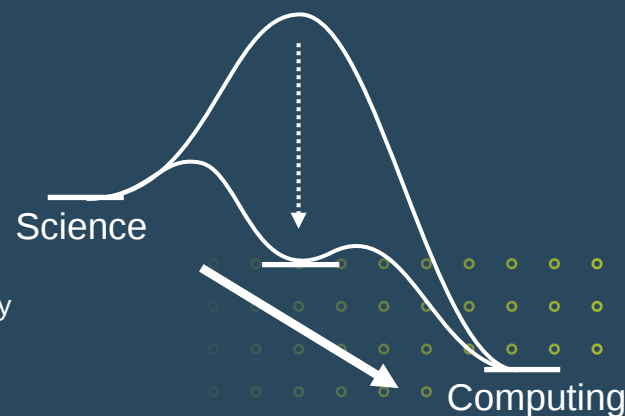
# HIGH PERFORMANCE COMPUTING AT FHI

## THE VIPER SUPERCOMPUTER & THE FHI COMPUTE CLOUD – AN OVERVIEW

Simeon D. Beinlich<sup>1,\*</sup> and Heinz Junkes<sup>1</sup>

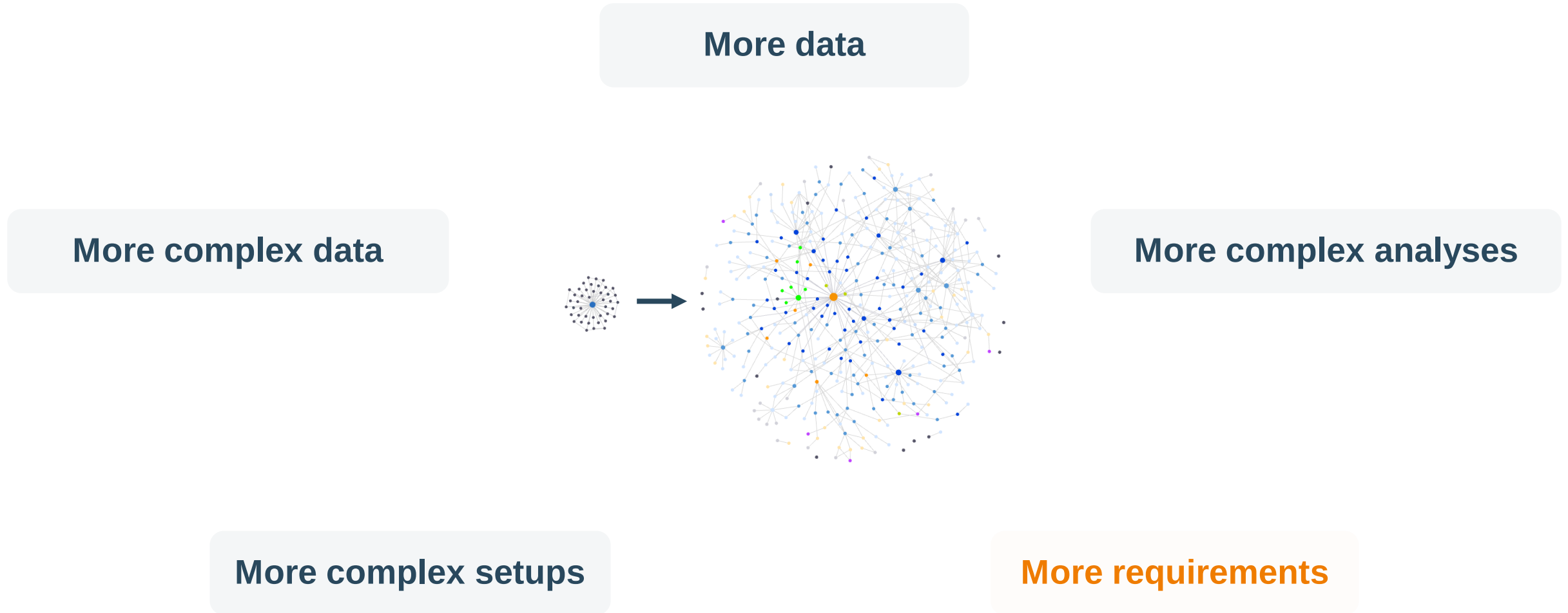
<sup>1</sup> PP&B – Computer Support Group  
Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin, Germany

\*beinlich@fhi.mpg.de



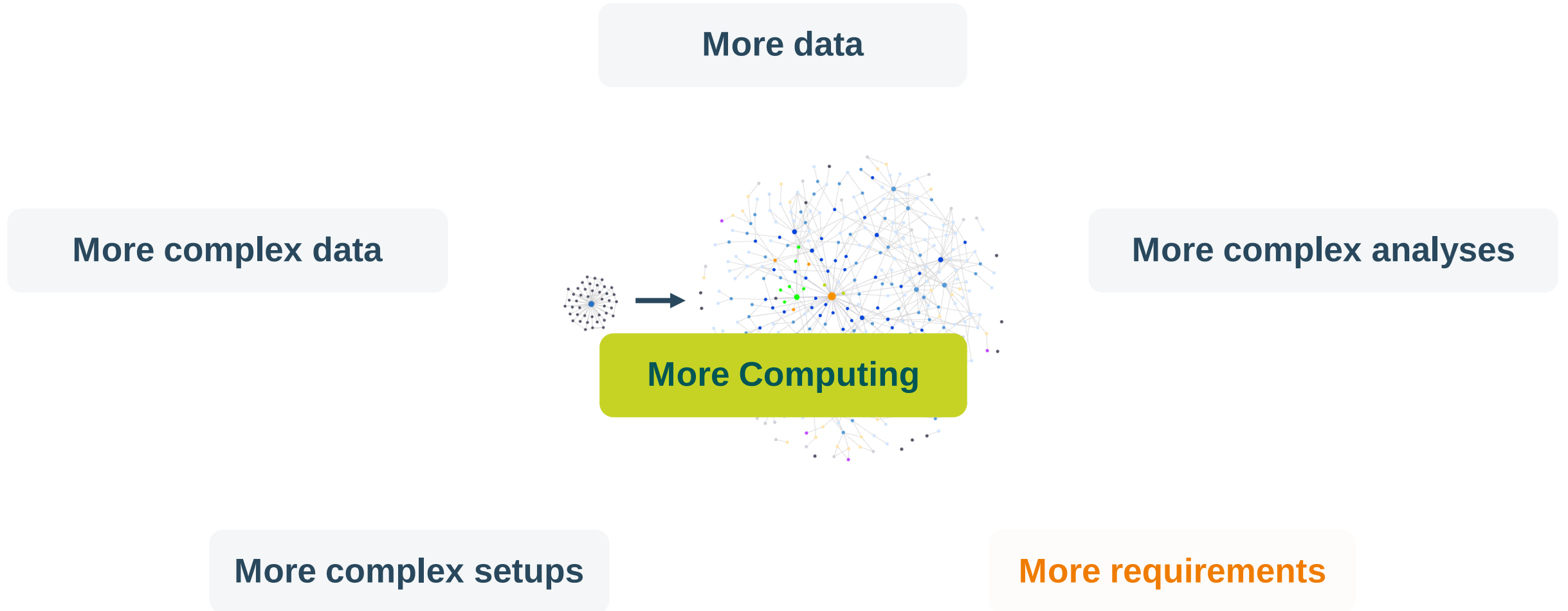


# COMPUTING REQUIREMENTS IN SCIENCE





# COMPUTING REQUIREMENTS IN SCIENCE





# COMPUTING REQUIREMENTS IN SCIENCE



Laptop / Desktop

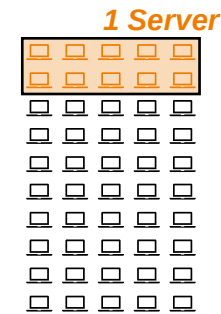


Computing Infrastructure



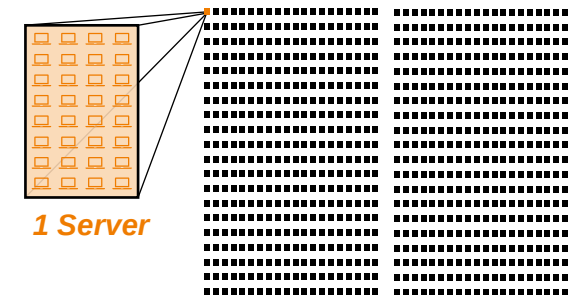
Laptop

Run 24/7 for 1 month



FHI COMPUTE  
x 50

Run 3 days (1 server)



VIPER  
x  
25000

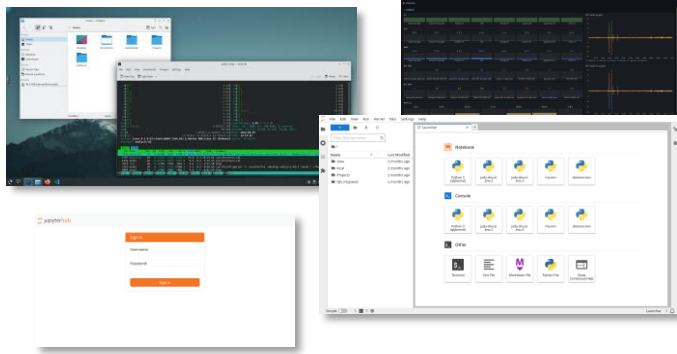
Run 22 hours @ 1 server  
or 2 hours @ 10 servers



# HPC INFRASTRUCTURE AT FHI (& MPG)



# HPC INFRASTRUCTURE AT FHI (& MPG)



FHI  
(PPB)



MPG  
(MPCDF)

## FHI COMPUTE Servers

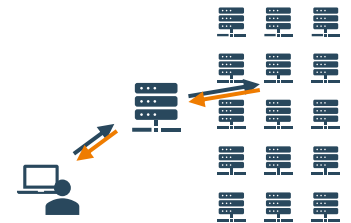
- Large Compute Servers (6x):  
Large CPU + Memory + GPU
- **Virtual Desktops** (Graphical Applications)
- **Jupyterhub** (Python & Matlab Server)
- Terminal Access
- **Direct Execution** (no Scheduling)
- User friendly



*Anything that requires more than a desktop PC*

## VIPER Supercomputer

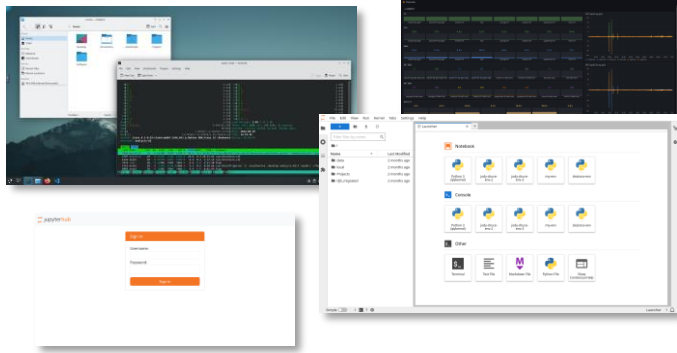
- Huge Compute Servers (768x):  
Huge CPU + Memory + GPU
- Terminal Access
- **Job Submission** (Scheduling)
- Advanced



*Large Scale Analyses / Processing  
& Ab-Initio Simulations*



# HPC INFRASTRUCTURE AT FHI (& MPG)



FHI  
(PPB)



MPG  
(MPCDF)

## FHI COMPUTE Servers

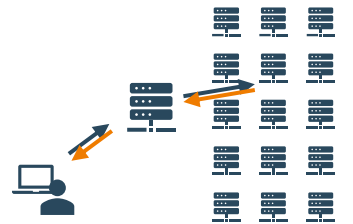
- Large Compute Servers (6x):  
Large CPU + Memory + GPU
- **Virtual Desktops** (Graphical Applications)
- **Jupyterhub** (Python & Matlab Server)
- Terminal Access
- **Direct Execution** (no Scheduling)
- User friendly



*Anything that requires more than a desktop PC*

## VIPER Supercomputer

- Huge Compute Servers (768x):  
Huge CPU + Memory + GPU
- Terminal Access
- **Job Submission** (Scheduling)
- Advanced

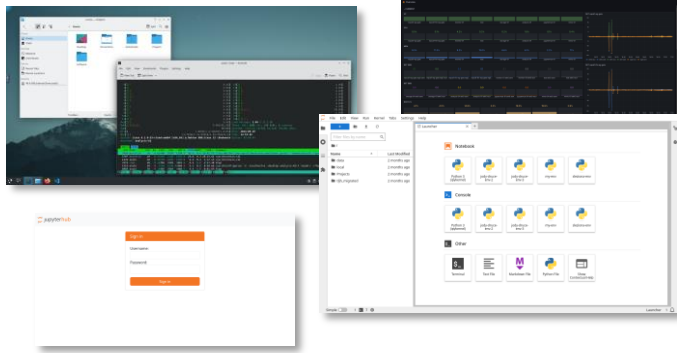


*+ RAVEN, SuperMUC, JUWELS, ...*  
*(department-level / grant-based)*

*Large Scale Analyses / Processing  
& Ab-Initio Simulations*



# HPC INFRASTRUCTURE AT FHI (& MPG)



FHI  
(PPB)

## FHI COMPUTE Servers

- Large Compute Servers (6x):  
Large CPU + Memory + GPU
- **Virtual Desktops** (Graphical Applications)
- **Jupyterhub** (Python & Matlab Server)
- Terminal Access
- **Direct Execution** (no Scheduling)
- User friendly



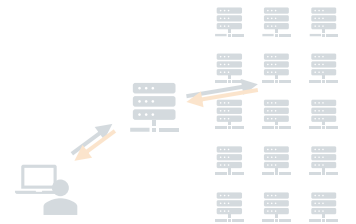
*Anything that requires more than a desktop PC*



MPG  
(MPCDF)

## VIPER Supercomputer

- Huge Compute Servers (768x):  
Huge CPU + Memory + GPU
- Terminal Access
- **Job Submission** (Scheduling)
- Advanced



*+ RAVEN, SuperMUC, JUWELS, ...*  
*(department-level / grant-based)*

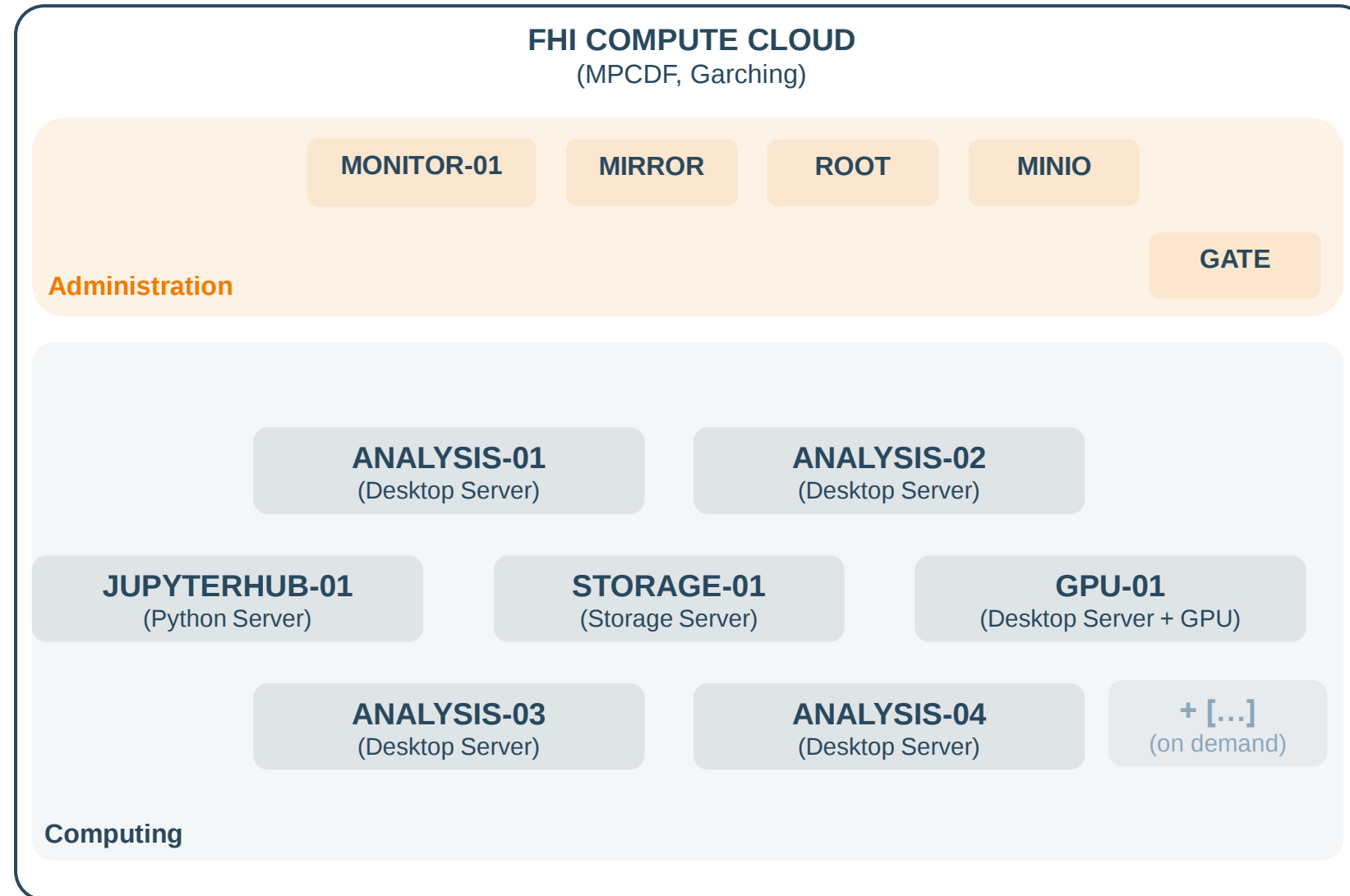
*Large Scale Analyses / Processing  
& Ab-Initio Simulations*



# FHI COMPUTE SERVERS

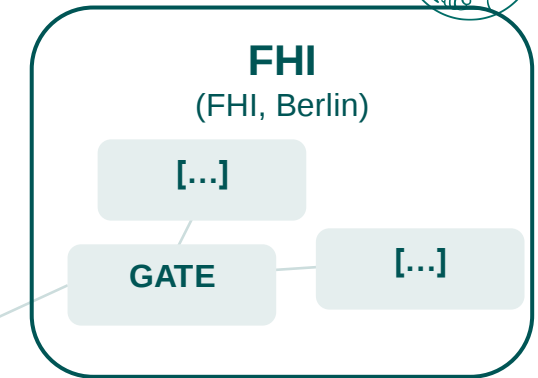
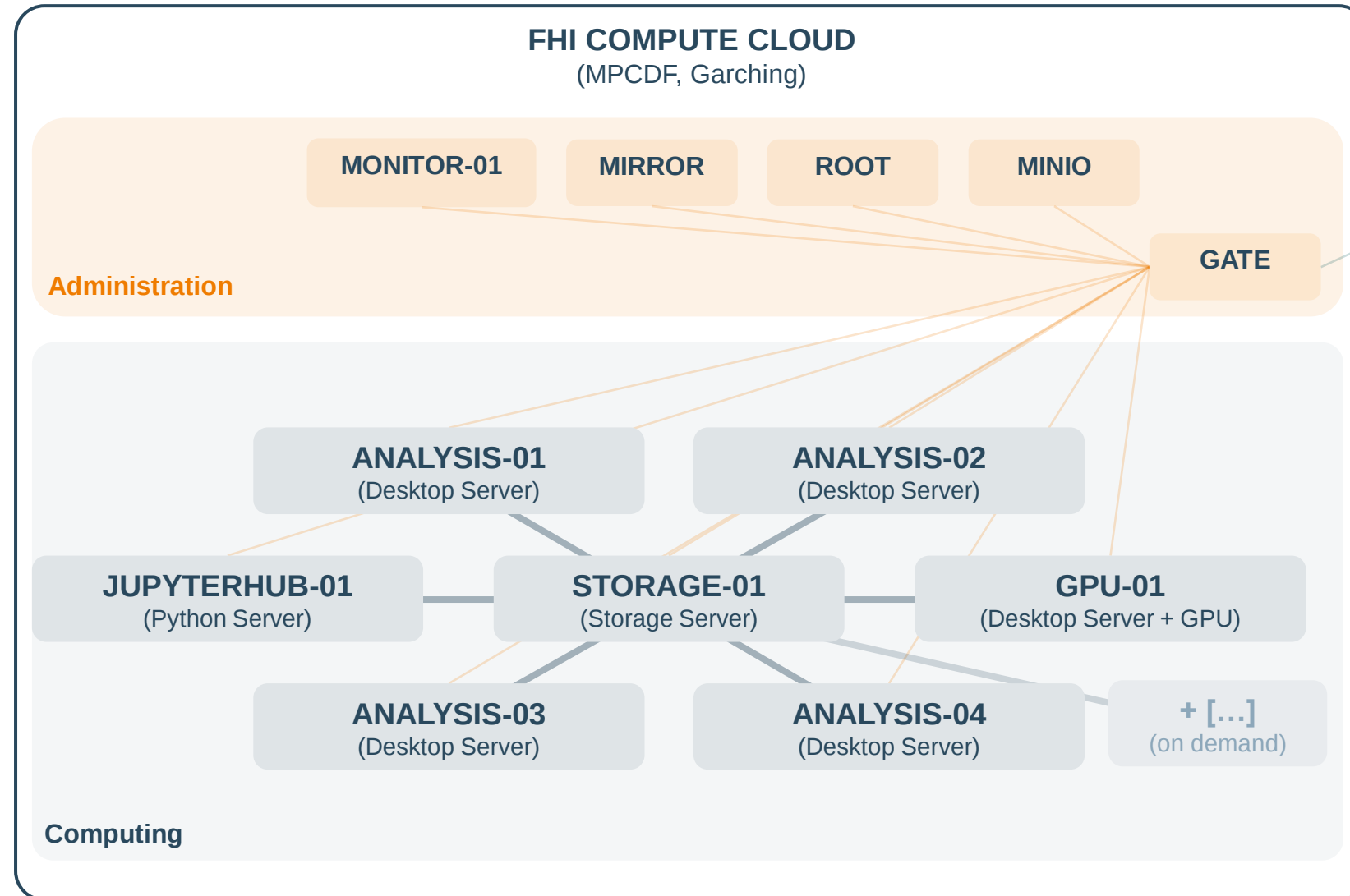


# FHI COMPUTE SERVERS – OVERVIEW



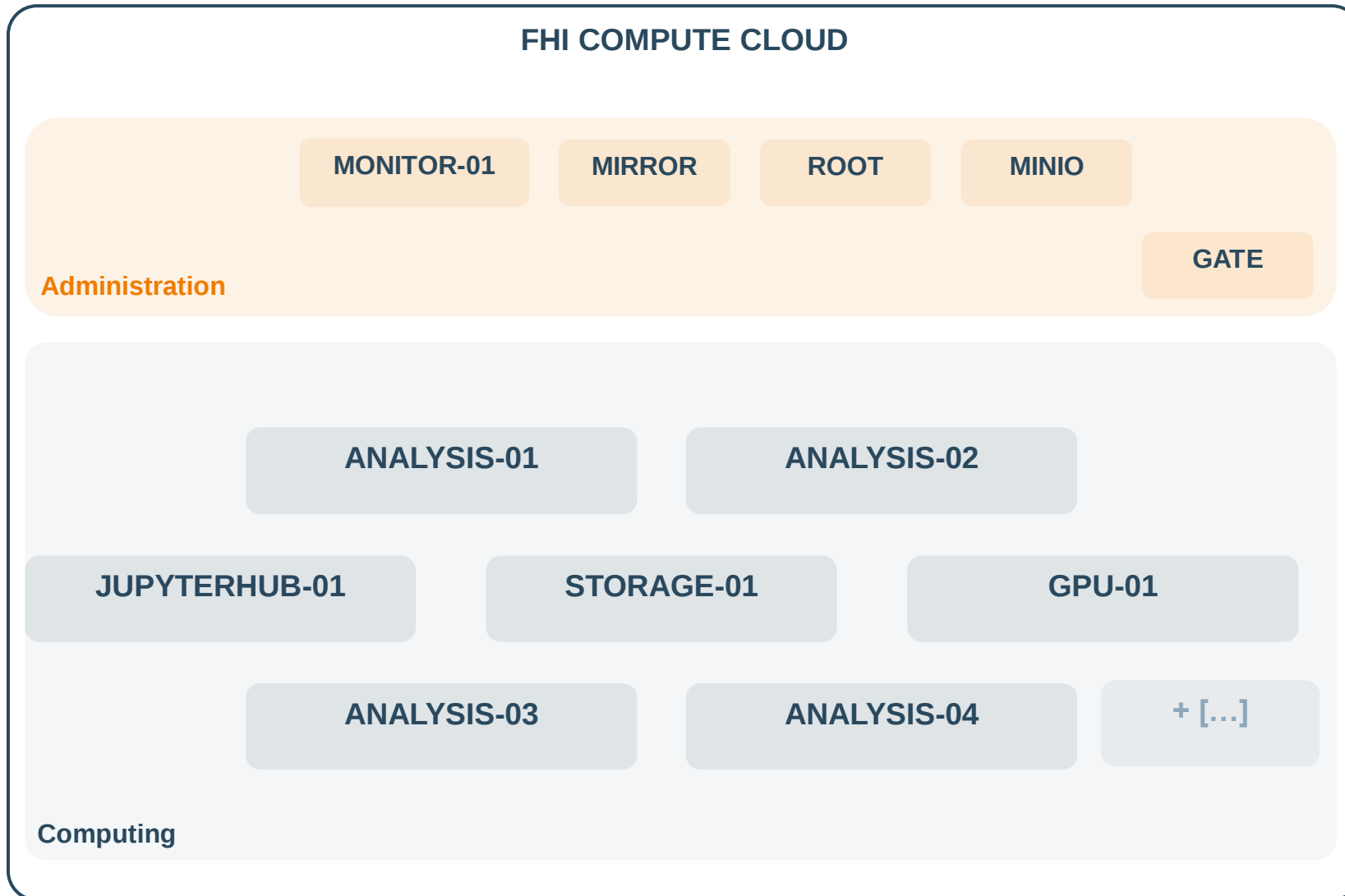


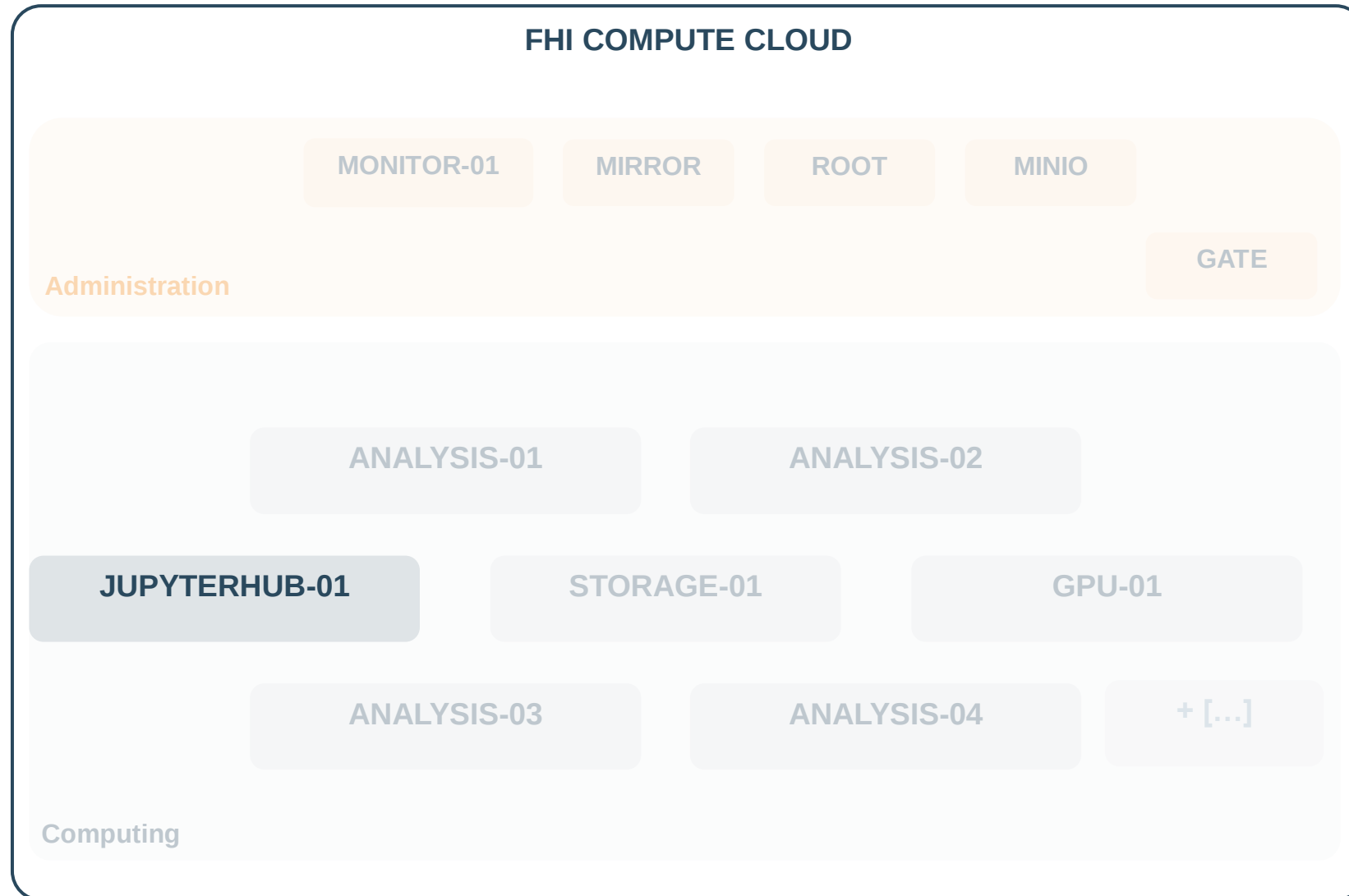
# FHI COMPUTE SERVERS – OVERVIEW



- 6 Compute Servers  
up to 32c128g + NV Tesla A100
- Full access from FHI network
  - Storage (Windows shares)
  - Desktop / terminal access
- Central storage
- Modular, scalable setup
- PPB administrated (full control)
- Custom software on request
- Shared project accounts possible
- No MPCDF account required
- Direct execution (no scheduler)

*User friendly*







# JUPYTERHUB-01 – PYTHON & MATLAB SERVER (AND MORE)

## Use case

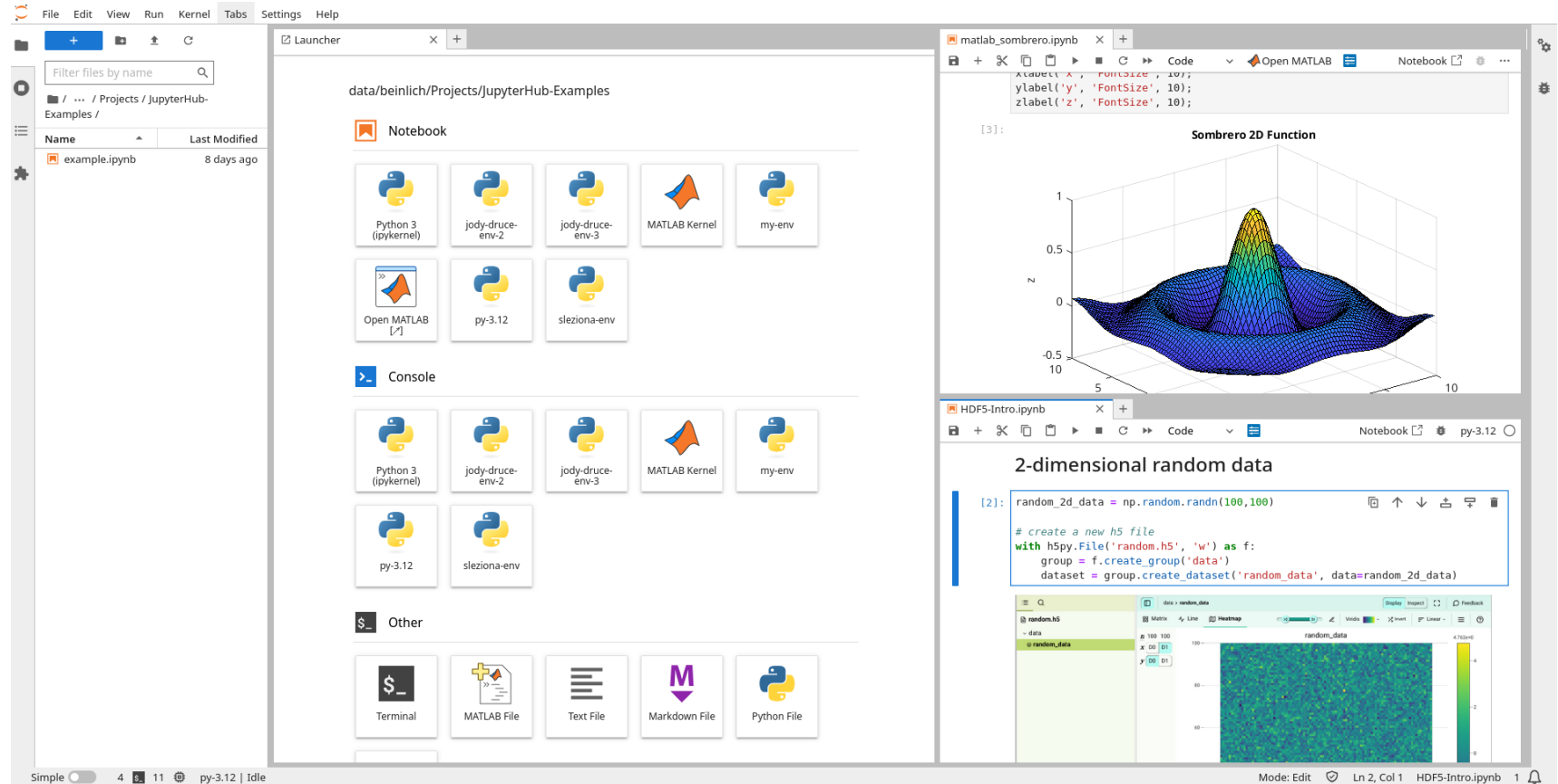
- Python / Matlab Shells
- Python / Matlab Scripts
- Python / Matlab Notebooks  
*“interactive script with notes”*
- Terminal

## Resources

- 32 cores, 128GB RAM
- Central, FHI accessible storage
  - 2TB Home
  - 10TB Data
- Fast local Storage (1TB)

## Advantages

- Large resources
- No local installation required (browser/terminal access)
- Custom setups for single users
- Multiple, independent Python setups
- Uniform setups for multiple users (e.g. for groups)
- **Directly use ready-made notebooks!**



## Requires:

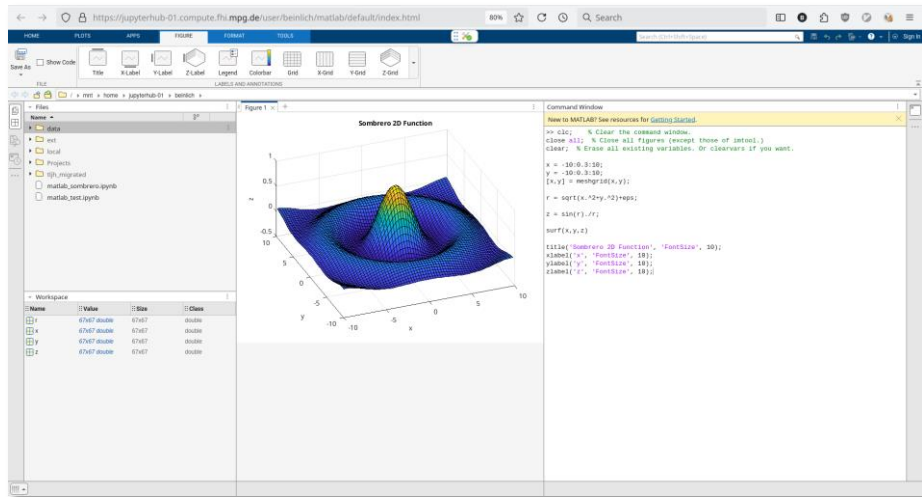
- Compute Account
- Browser (*Chrome recommended*)

## Documentation:

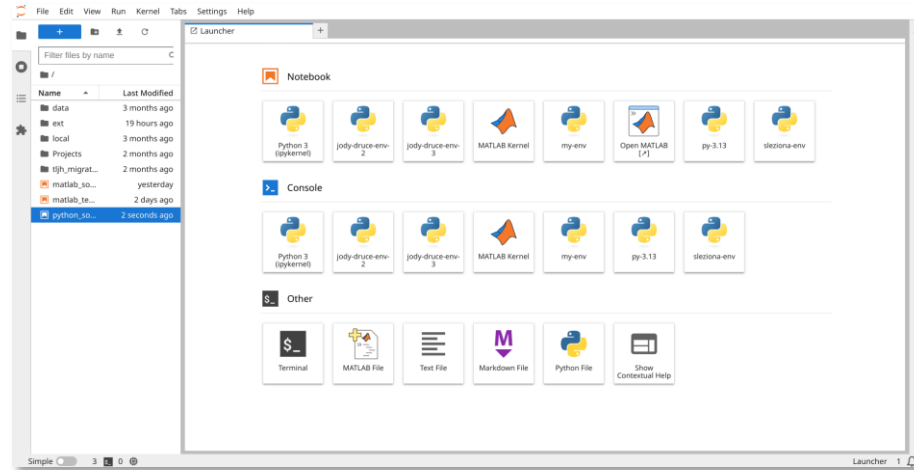
- PPBwiki (small): <https://ppbwiki.fhi.mpg.de/Main/Compute-Jupyterhub>
- GIT Wiki (large): <https://gitlab.fhi.mpg.de/fhi-hpc/jupyterhub-compute/-/wikis/HOME>



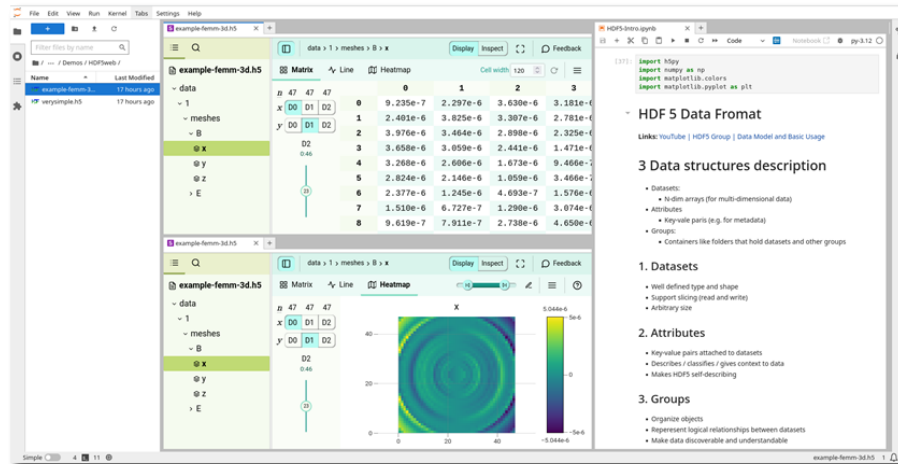
# JUPYTERHUB-01 – PYTHON & MATLAB SERVER (AND MORE)



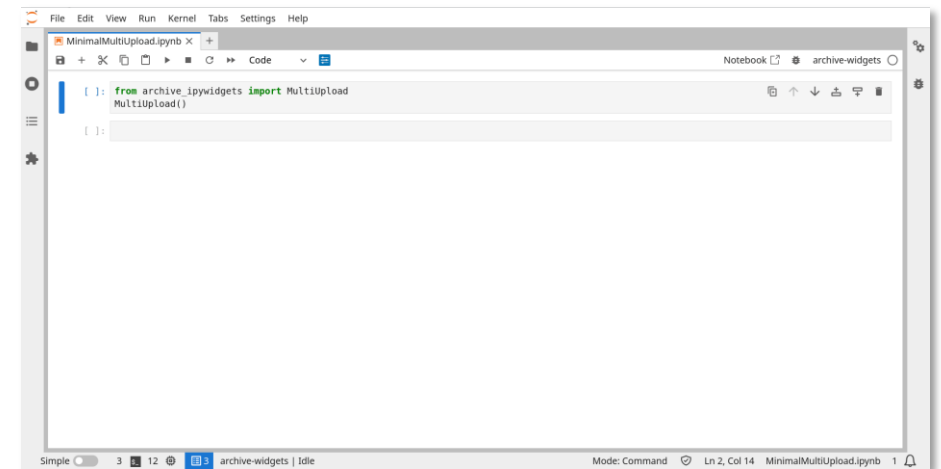
Matlab in a Browser



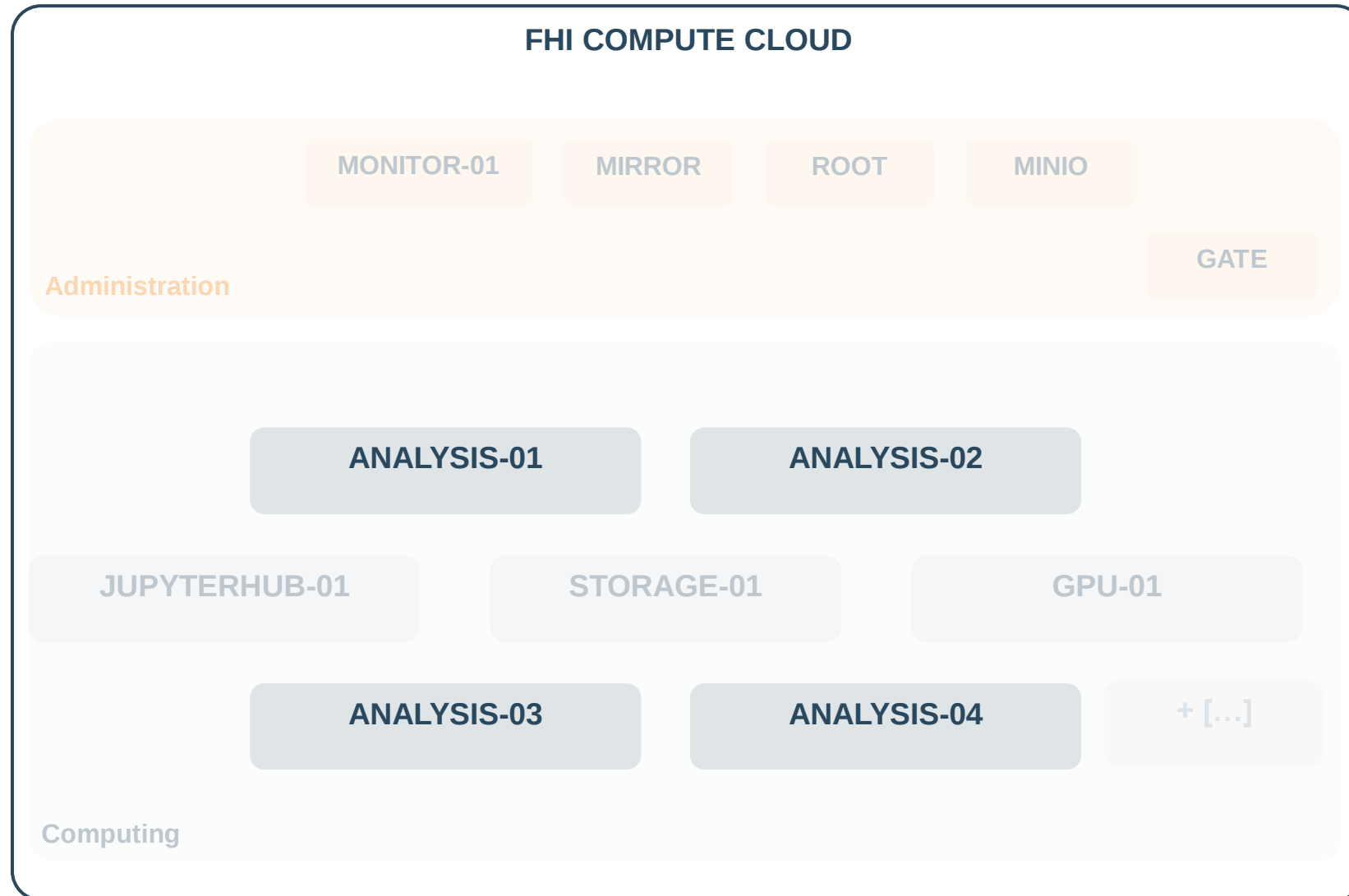
Python in a Browser



HDF5web in a Browser



Custom Widgets (e.g. Archive)





# ANALYSIS-0X – VIRTUAL DESKTOP SERVERS

## Use case

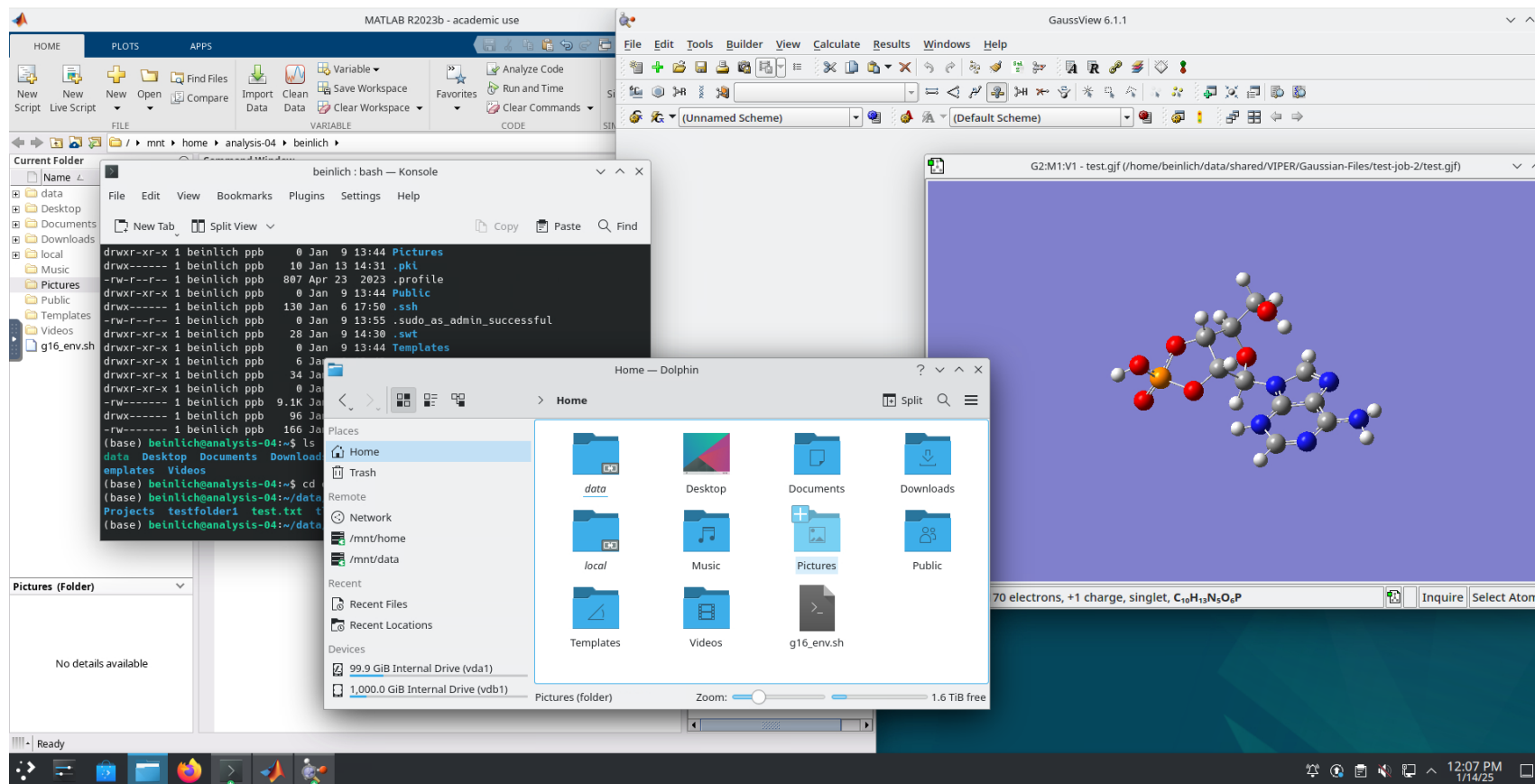
- Graphical Applications (MATLAB, Gaussview, VESTA, *maybe no video editing...*)
- Terminal Applications (DFT, Molecular Dynamics)

## Resources

- 32 cores, 128GB RAM (max.)
- Central, fhi accessible storage
  - 2TB Home
  - 10TB Data
- Fast local Storage (1TB)

## Advantages

- Large resources
- Graphical Applications
- No local installation required (browser sufficient)
- Any Linux-compatible software (debian)
- No terminal / programming experience required.

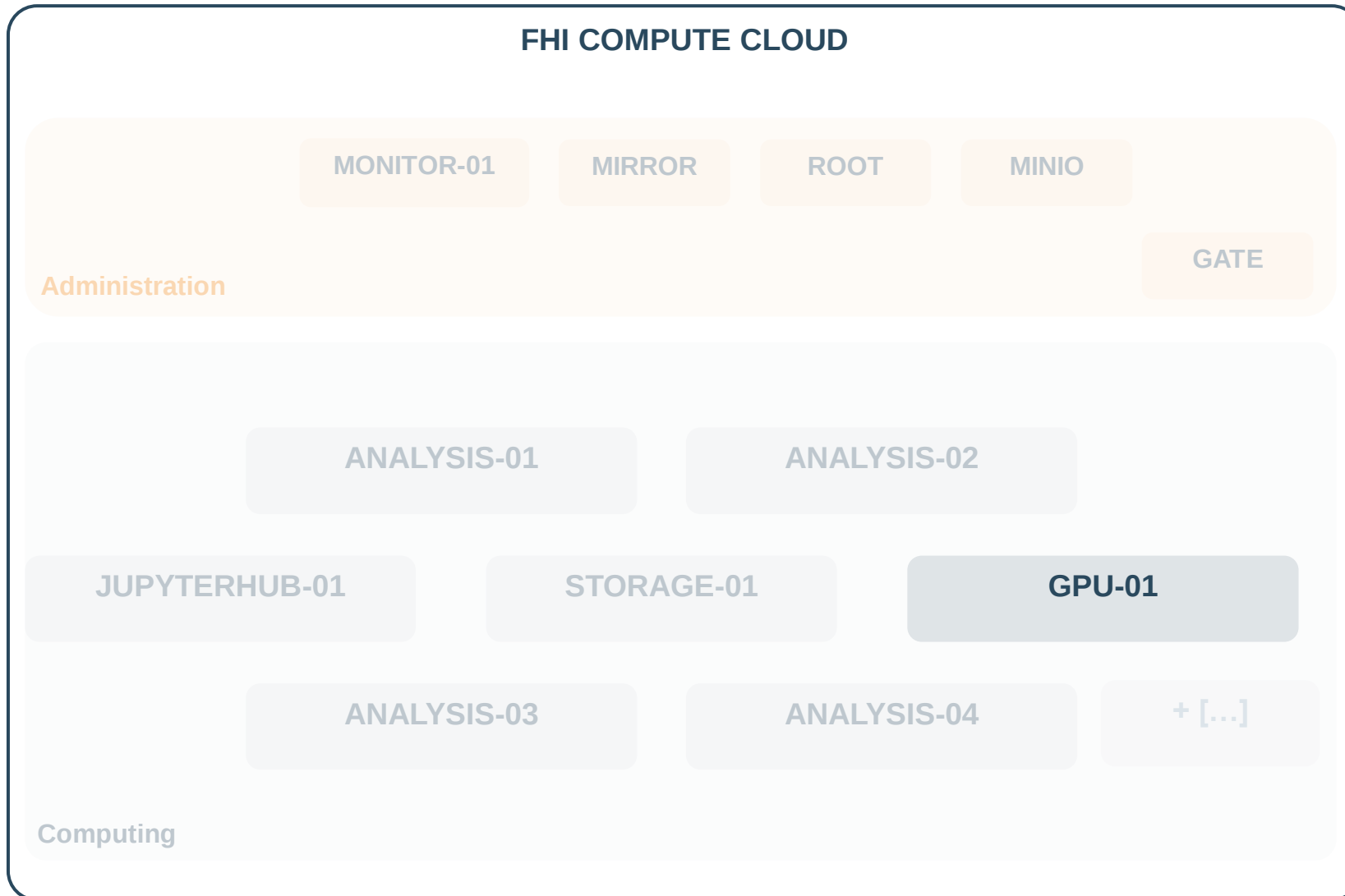


## Requires:

- Compute Account + Desktop Session
- Browser (*Chrome recommended*)  
or SSH (*Powershell / Terminal recommended*)  
+ VNC Viewer (*tigervnc, tightvnc, realvnc*)

## Documentation:

- PPBwiki (small): <https://ppbwiki.fhi.mpg.de/Main/Compute-Analysis>
- GIT Wiki (large): <https://gitlab.fhi.mpg.de/fhi-hpc/analysis-compute/-/wikis/HOME>





# GPU-01 – VIRTUAL DESKTOP + NVIDIA TESLA A100 GPU

## Use case

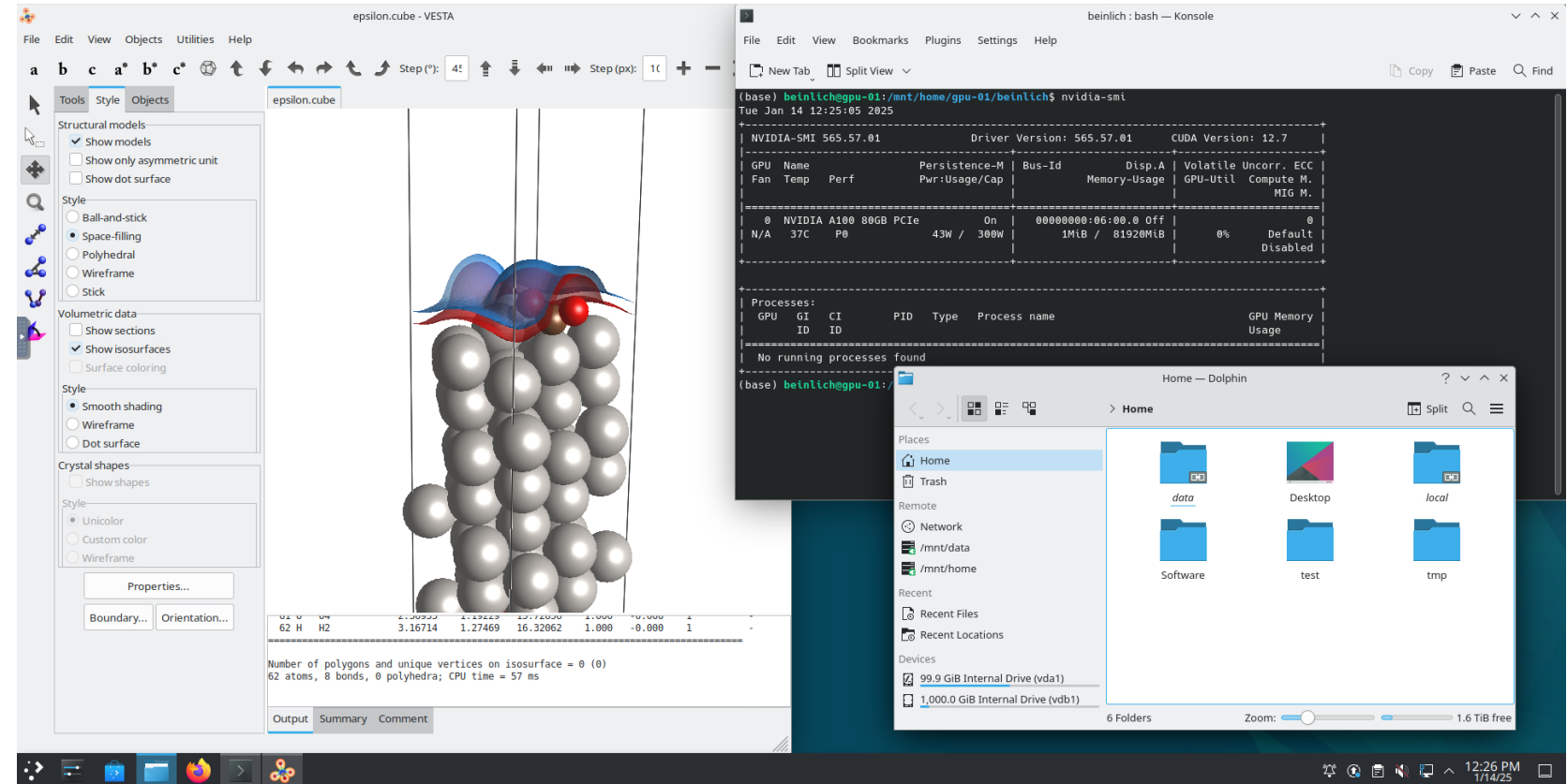
- see Analysis-0X
- + Machine Learning / AI
- + Accelerated Computing (GPU-accel.)

## Resources

- 32 cores, 128GB RAM, NVIDIA Tesla A100 GPU (80GB)
- Central, fhi accessible storage
  - 2TB Home
  - 10TB Data
- Fast local Storage (1TB)

## Advantages

- see Analysis-0X
- + AI / ML / GPU acceleration

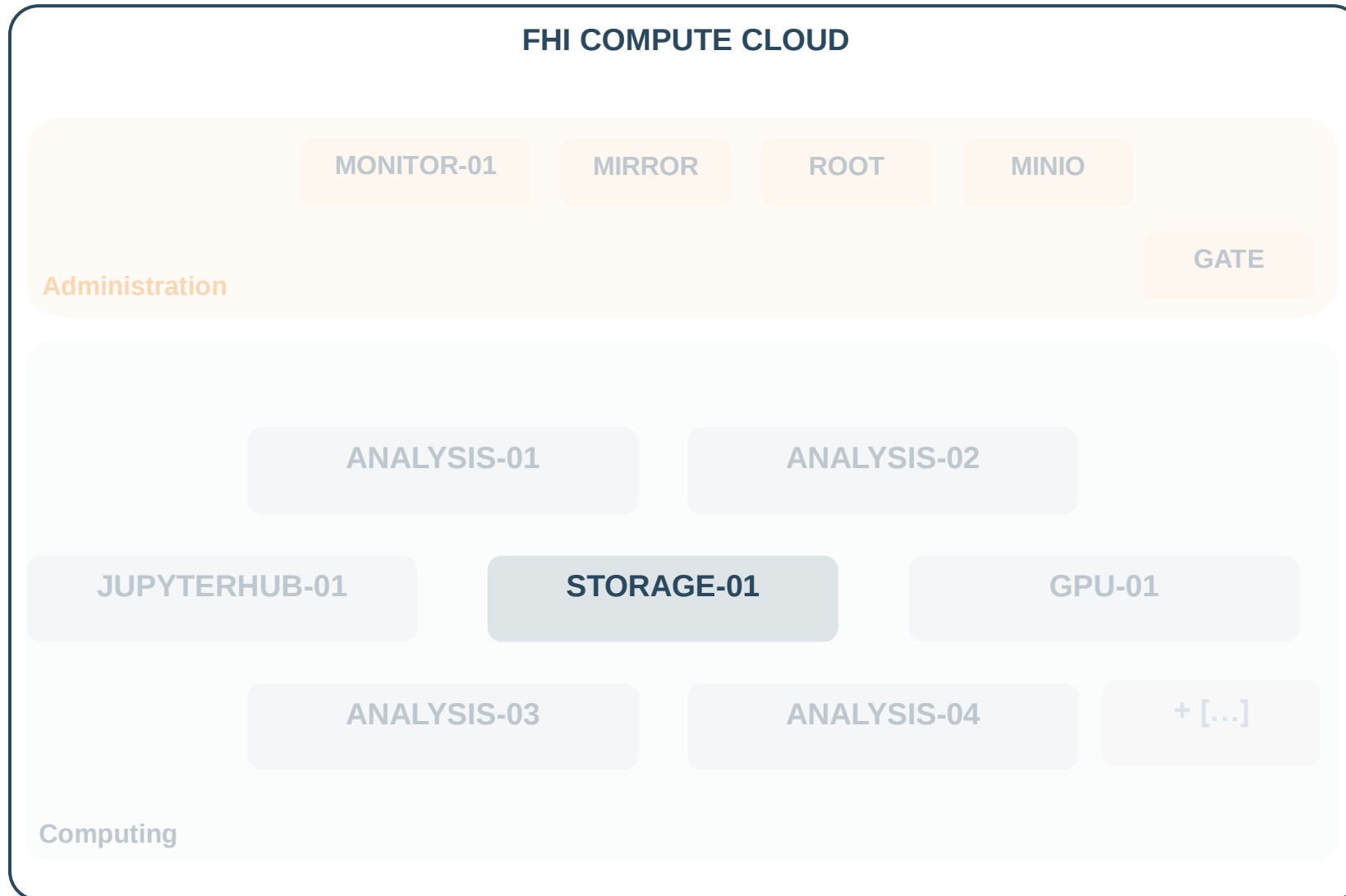


## Requires:

- Compute Account + Desktop Session
- Browser (*Chrome recommended*)  
or SSH (*Powershell / Terminal recommended*)  
+ VNC Viewer (*tigervnc, tightvnc, realvnc*)

## Documentation:

- PPBwiki (small): <https://ppbwiki.fhi.mpg.de/Main/Compute-Analysis>
- GIT Wiki (large): <https://gitlab.fhi.mpg.de/fhi-hpc/analysis-compute/-/wikis/HOME>





# STORAGE-01 – CENTRAL STORAGE FOR COMPUTE SERVERS

## Use case

- Transfer files to/from the Compute Servers
- Manage all your compute server files in one place

## Resources

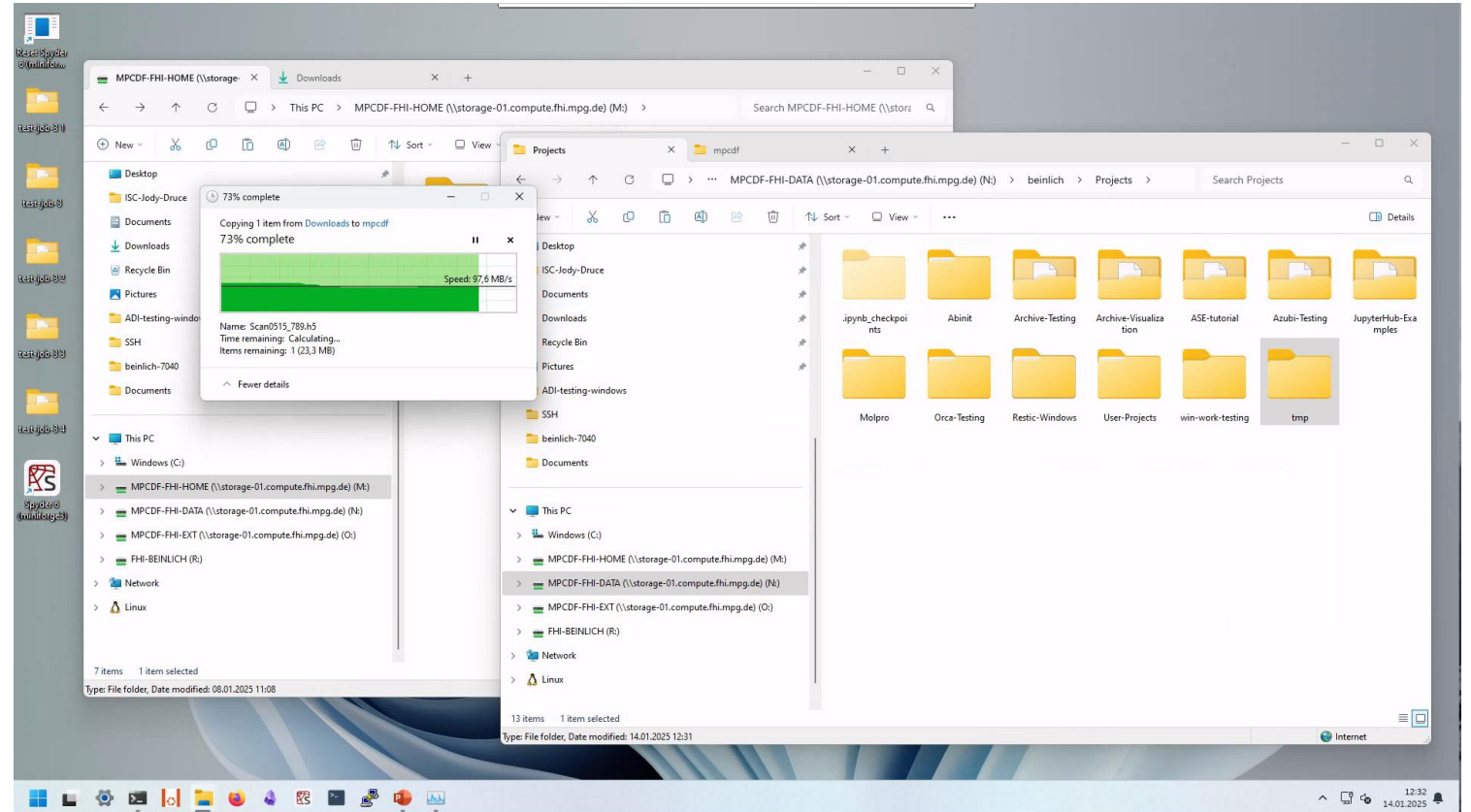
Shared filesystems (snapshotted, backed-up)

- 2TB Home
- 10TB Data
- X TB Storage Extensions (>10TB, on demand)

SMB/SFTP/SSH access from FHI

## Advantages

- Simple File/Data Management
- Backed up data
- Integration of large data buckets possible on demand (>100TB)

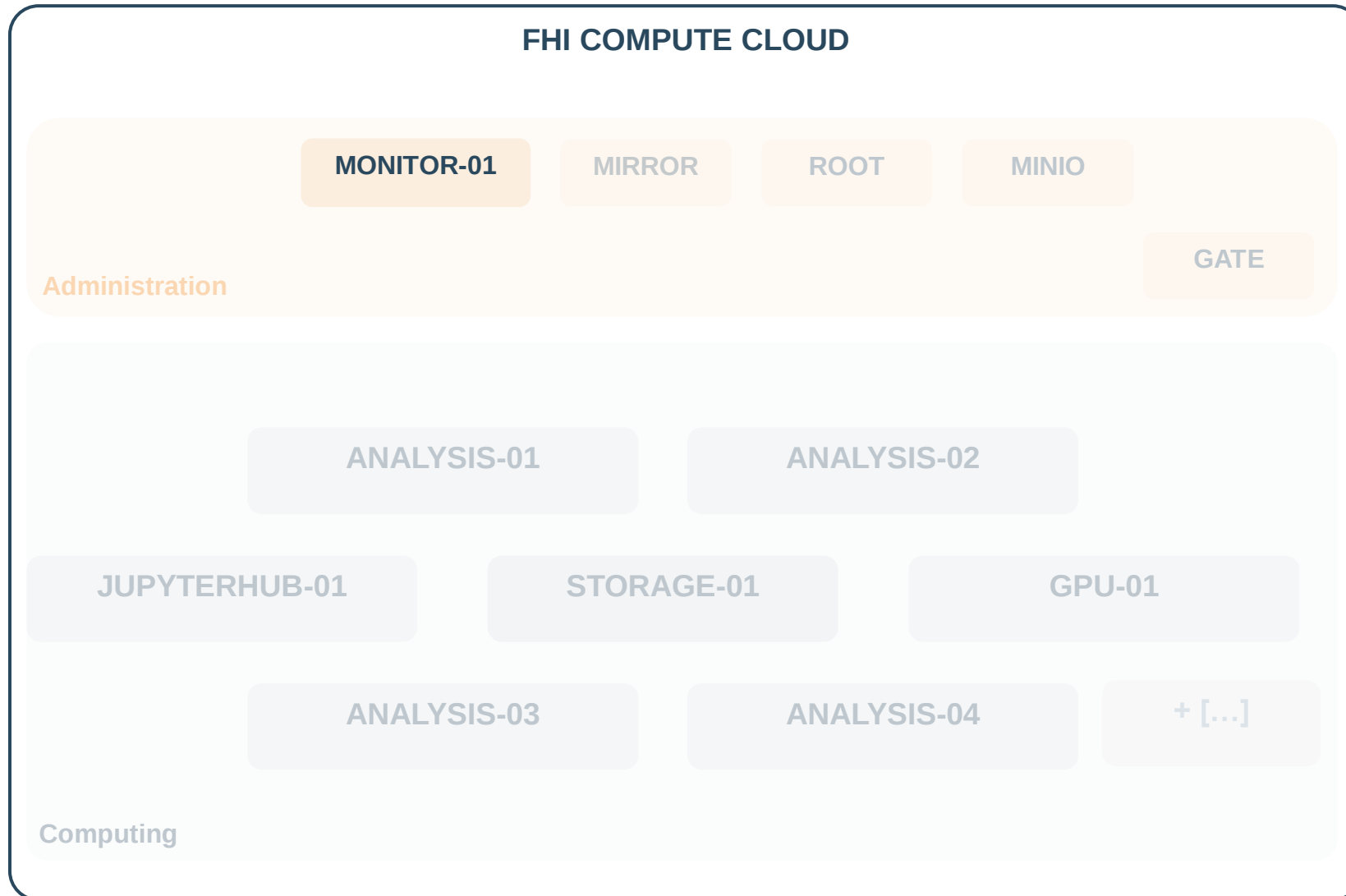


## Requires:

- Compute Account
- File Explorer (Windows Explorer or SMB/SFTP-enabled file explorer)

## Documentation:

- PPBwiki (small): <https://ppbwiki.fhi.mpg.de/Main/Compute-Storage>
- GIT Wiki (large): <https://gitlab.fhi.mpg.de/fhi-hpc/storage-compute/-/wikis/HOME>





# MONITOR-01 – CENTRAL RESOURCE MONITORING

## Use case

- Monitor current state and usage of compute servers and storage
- Monitor current usage of your sessions
- Monitor Viper Compute Budgets

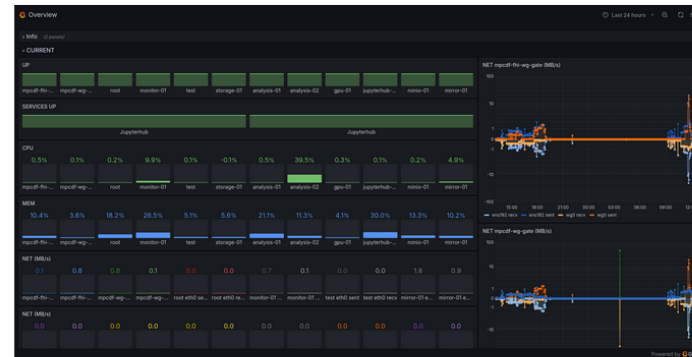
## Resources

- Compute Server Monitoring:
- UP state, CPU & Memory usage, Network traffic, Storage Space
- Viper Usage Monitoring

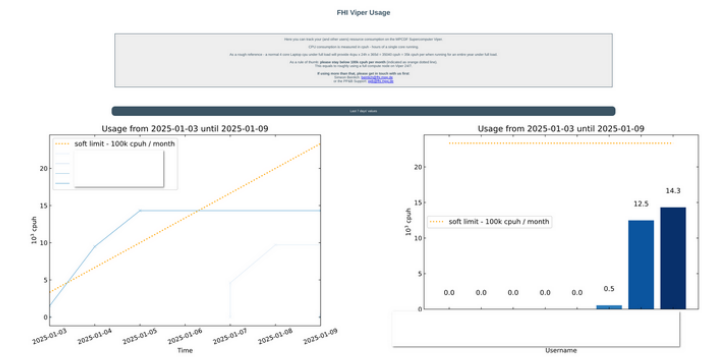
## Advantages

- Easy monitoring of cpu and memory consumption to prevent issues or optimize your processes
- Check that you stay within the Viper budget

## MONITOR-01 - Compute Server and Viper Usage Monitoring



Compute Server Monitoring



Viper Usage Monitoring

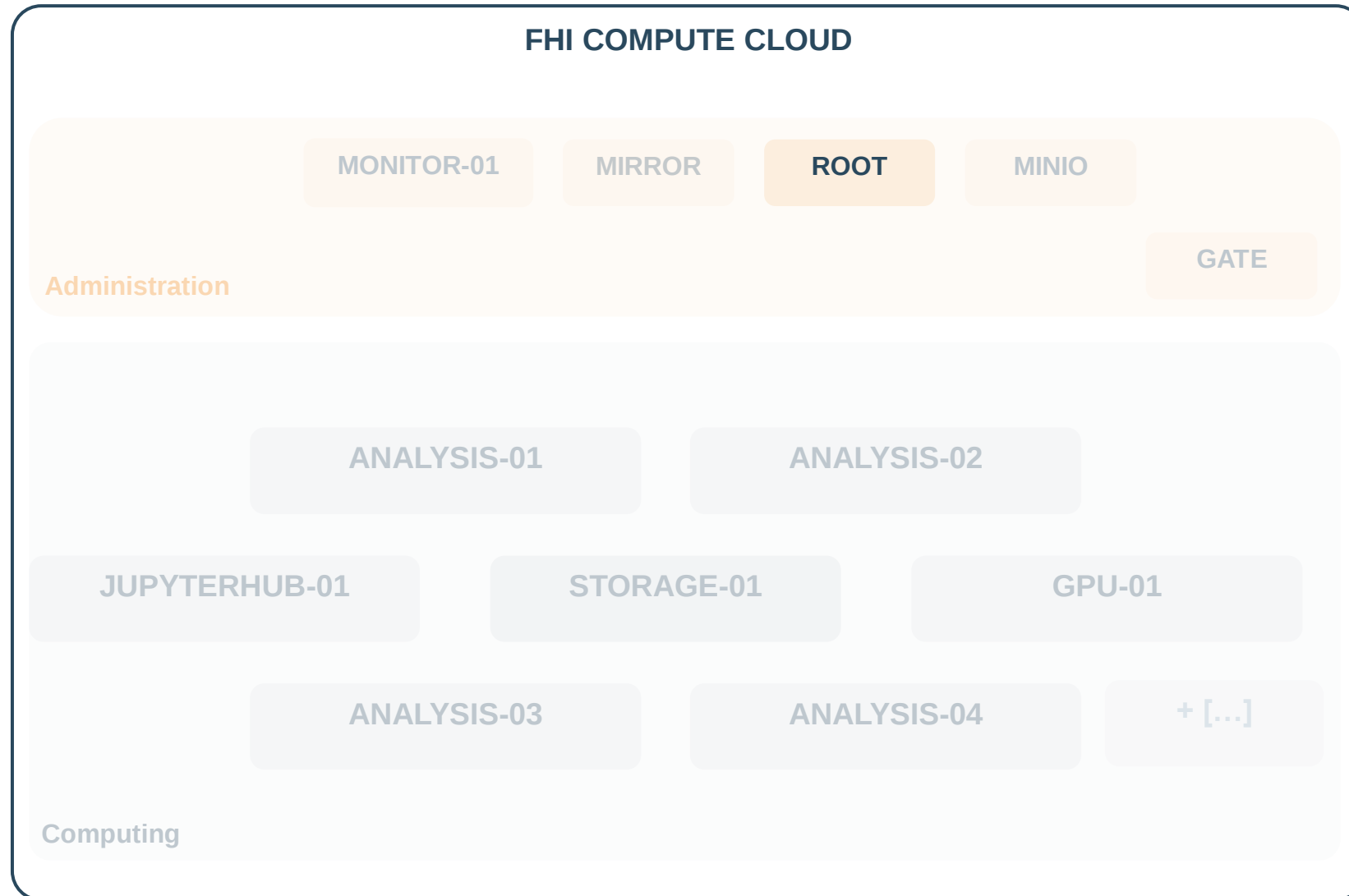
Administration  
Grafana admin interface  
Prometheus admin interface

## Requires:

- Browser (*Chrome recommended*)

## Documentation:

- PPBwiki (small): <https://ppbwiki.fhi.mpg.de/Main/HPC>





# ROOT – CENTRAL ACCOUNT MANAGER

## Use case

- Account Creation
- Password Change / Reset
- Desktop Session (Re-)Start

## Advantages

- Simple restarting of VNC servers
- Simple Password Change for all compute services

### FHI Compute Manager

Welcome to the FHI Compute Manager!



#### Account Management

- Create Account
- Change Password
- Reset Password
- Start / restart VNC Session
- Start / restart all VNC Sessions

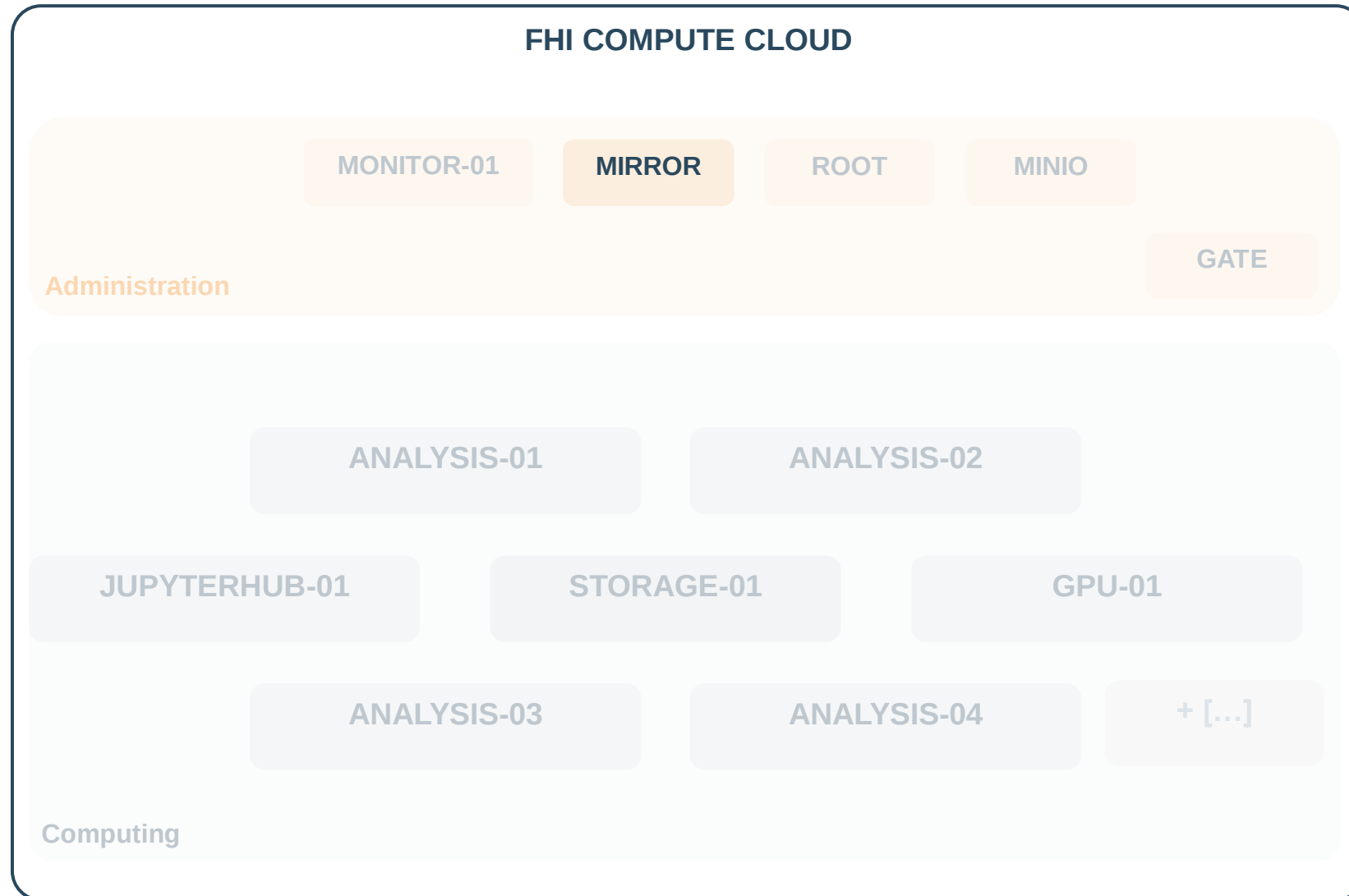
For more information, please visit the [ppbwiki.fhi.mpg.de](https://ppbwiki.fhi.mpg.de).

### Requires:

- Browser (*Chrome recommended*)

### Documentation:

- PPBwiki (small): <https://ppbwiki.fhi.mpg.de/Main/HPC>





# MIRROR – LARGE-SCALE FILE TRANSFER (ASYNCHRONOUS S3 SYNC)

## Use case

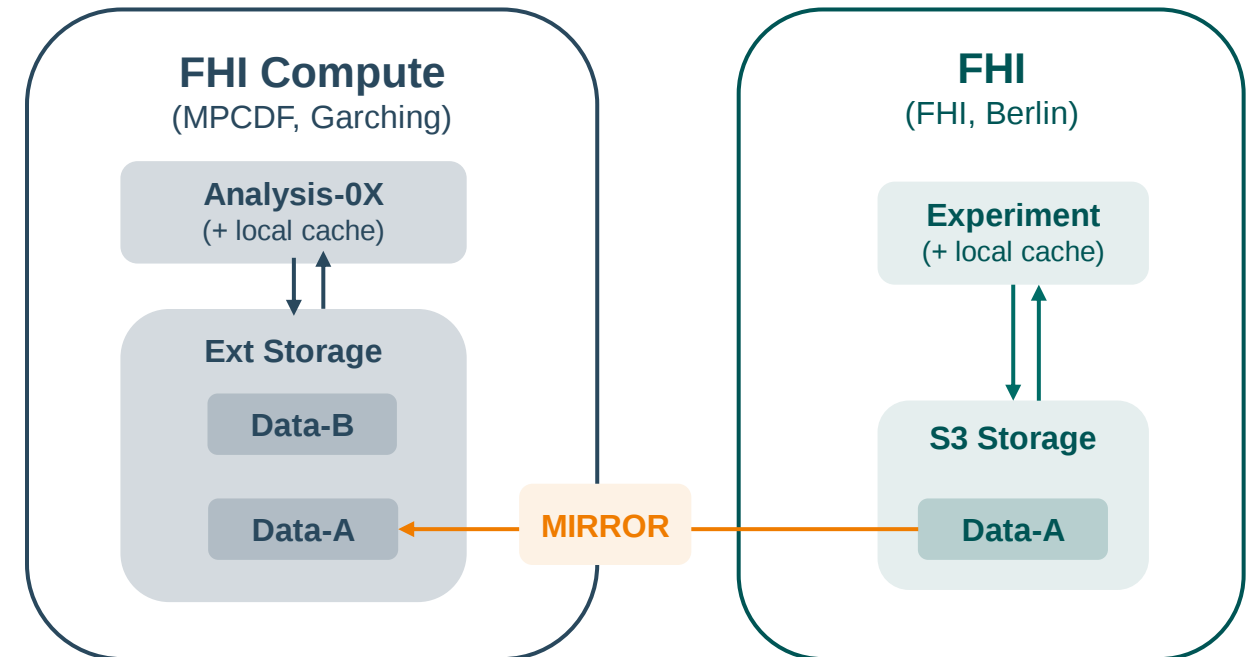
- Near real-time, fast synchronization of FHI S3 buckets to FHI Compute Storage (one-way, ~2.5-8GBit/s)
- E.g. for experiments that produce large data sets that should be processed at Compute Servers
- Large Data Stores (> 100TB), Large Files (>10MB)
- Full Data Access at Compute Servers
- Single- or multi-user access on entire data set or subfolder
- FHI → MPCDF or MPCDF → FHI

## Resources

- Based on S3 Storage (200TB+)
- Local SSD Cache possible (r/w, soon)
- Filesystem-like access from Compute Servers
- SMB (Windows network storage) convenience access from FHI (<1GBit/s)

## Advantages

- Fast, large data store



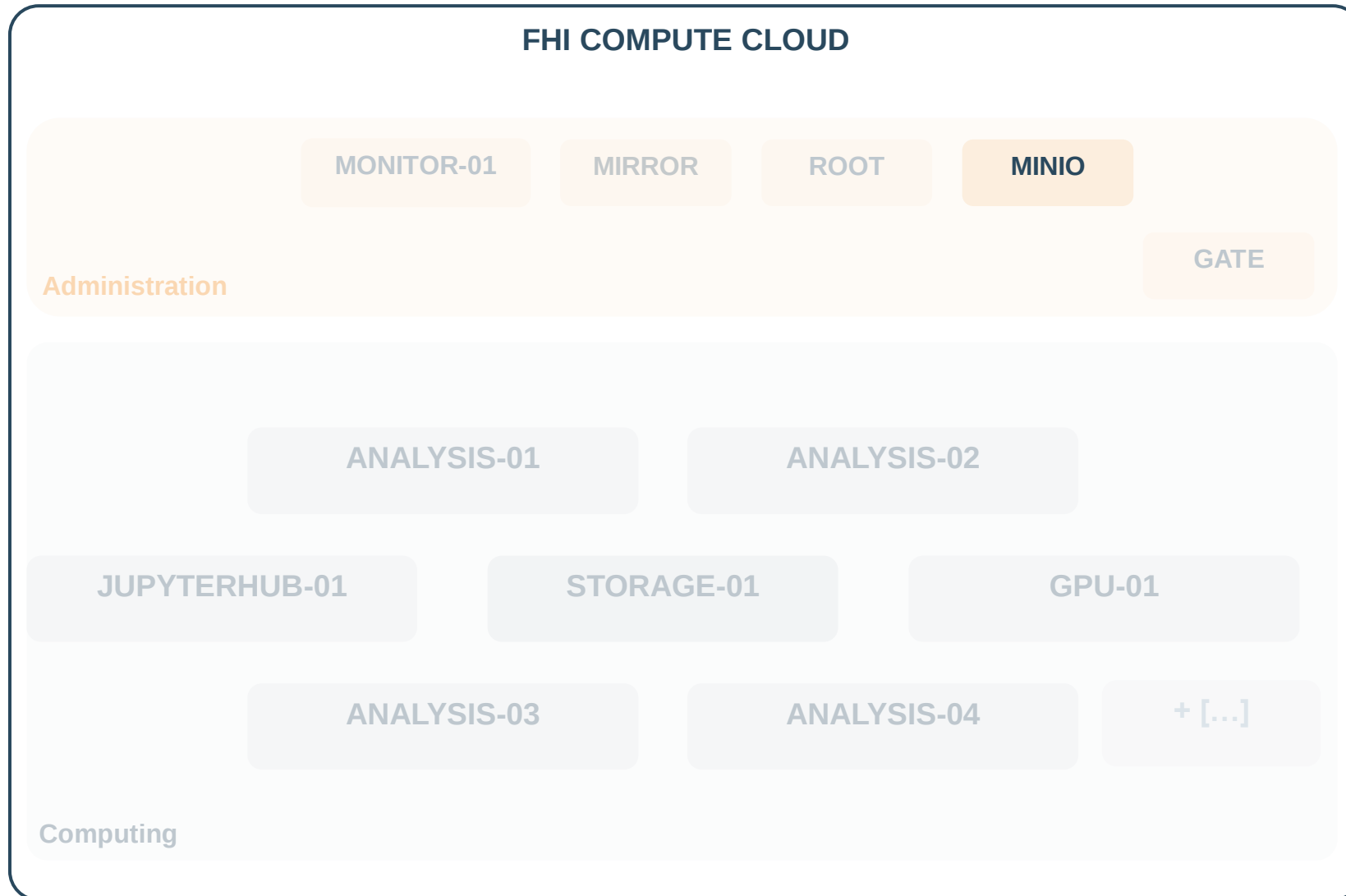
### Requires:

- FHI S3 Bucket (*on request*)
- MPCDF Ext Storage Access (*on request*)

### Documentation:

- TBA

**TESTING PHASE !**





# MINIO-01 – S3 USER BUCKETS

## Use case

- S3 User/Project buckets similar to FHI S3 buckets

## Resources

- 80TB (more on demand)

## Advantages

- Known S3 interface as with the FHI S3 Storage
- No File Size restrictions
- Direct S3 API and Webinterface access with per-user credentials

The screenshot displays the Minio Object Store web interface. On the left is a dark sidebar with navigation options: User (Object Browser, Access Keys, Documentation), Administrator (Buckets, Policies, Identity, Monitoring, Events, Tiering, Site Replication), and a top section for MINIO OBJECT STORE with an AGPL LICENSE. The main content area shows the 'Object Browser' for the bucket 'beinlichtests3'. It includes a search bar, a breadcrumb 'beinlichtests3 / subset', and a table of objects. The table has columns for Name, Last Modified, and Size. The objects listed are Scan0515\_7.h5 through Scan0515\_705.h5, all created on 'Thu, Oct 31 2024 16:11 (GMT+1)' with sizes ranging from 115.3 MiB to 115.4 MiB. Action buttons like 'Rewind', 'Refresh', and 'Upload' are visible at the top right of the bucket view.

| Name            | Last Modified                  | Size      |
|-----------------|--------------------------------|-----------|
| Scan0515_7.h5   | Thu, Oct 31 2024 16:11 (GMT+1) | 115.3 MiB |
| Scan0515_70.h5  | Thu, Oct 31 2024 16:11 (GMT+1) | 115.3 MiB |
| Scan0515_700.h5 | Thu, Oct 31 2024 16:11 (GMT+1) | 115.4 MiB |
| Scan0515_701.h5 | Thu, Oct 31 2024 16:11 (GMT+1) | 115.4 MiB |
| Scan0515_702.h5 | Thu, Oct 31 2024 16:11 (GMT+1) | 115.4 MiB |
| Scan0515_703.h5 | Thu, Oct 31 2024 16:11 (GMT+1) | 115.4 MiB |
| Scan0515_704.h5 | Thu, Oct 31 2024 16:11 (GMT+1) | 115.4 MiB |
| Scan0515_705.h5 | Thu, Oct 31 2024 16:11 (GMT+1) | 115.4 MiB |

## Requires:

- Browser / Minio Client / S3 client / rclone / restic / ...
- FHI MINIO-01 Bucket (*on request*)

## Documentation:

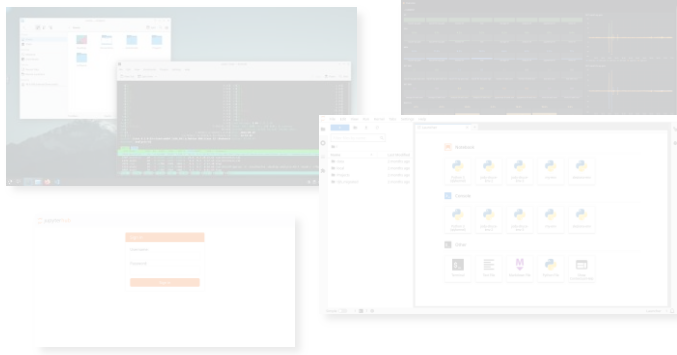
- TBA



# VIPER SUPERCOMPUTER



# HPC INFRASTRUCTURE AT FHI (& MPG)



FHI  
(PPB)

## FHI COMPUTE Servers

- Large Compute Servers (6x):  
Large CPU + Memory + GPU
- **Virtual Desktops** (Graphical Applications)
- **Jupyterhub** (Python Server)
- Terminal Access
- **Direct Execution** (no Scheduling)
- User friendly



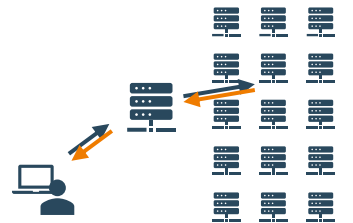
*Anything that requires more than a desktop PC*



MPG  
(MPCDF)

## VIPER Supercomputer

- Huge Compute Servers (768x):  
Huge CPU + Memory + GPU
- Terminal Access
- **Job Submission** (Scheduling)
- Advanced



*+ RAVEN, SuperMUC, JUWELS, ...*  
*(department-level / grant-based)*

*Large Scale Analyses / Processing  
& Ab-Initio Simulations*

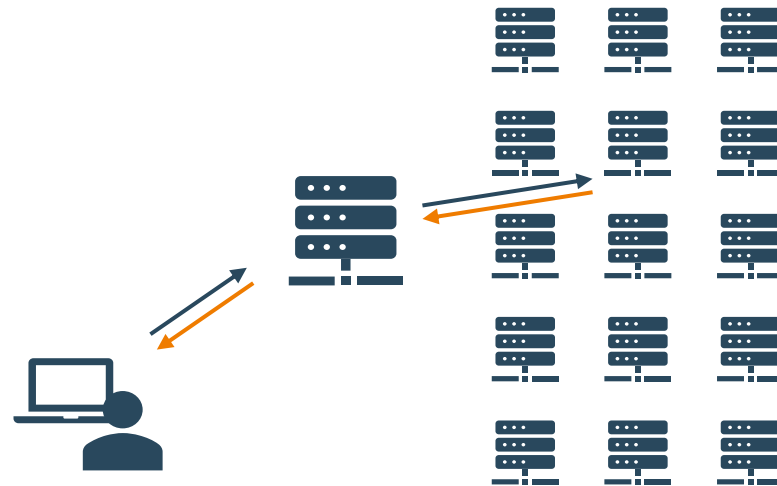
# VIPER SUPERCOMPUTER – MPCDF

## Use case

- Large-Scale Data Analysis & Processing
- Machine Learning / AI
- Ab-initio Simulations (DFT etc.)
- Anything massively parallelizable

## Resources

- SLURM scheduler
- Massive module system  
(Compiler, Software, ...)
- 768 CPU compute nodes:
  - 609 x 128c/512g
  - 90 x 128c/768g
  - 66 x 128c/1024g
  - 3 x 128c/2304g
- 340 APU compute nodes:  
(GPU accelerated CPUs)
  - 2 x AMD MI300A + 2 x 128g HBM3



## Documentation:

- PPBwiki (small): <https://ppbwiki.fhi.mpg.de/Main/HPC>
- GIT Wiki (large): <https://gitlab.fhi.mpg.de/fhi-hpc/viper/-/wikis/home>



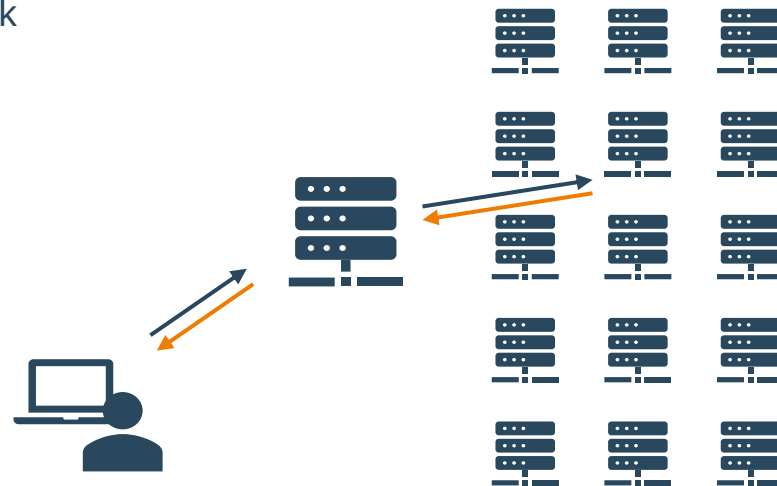
# VIPER SUPERCOMPUTER – MPCDF

## Advantages

- Huge computing power
- State-of-the art software & hardware stack

## Common Workflow

- Transfer data to Viper  
e.g. *WinSCP, Rsync, ...*
- Log in via SSH  
e.g. *Powershell, Putty, ...*
- Create a job script  
e.g. *job.sh*
- Submit the job  
e.g. *sbatch job.sh*
- Check the Queue until the job is done  
e.g. *squeue -u myusername*
- Transfer the results back to your PC  
e.g. *WinSCP, Rsync, ...*



```
#!/bin/bash

#SBATCH --job-name=my-job
#SBATCH --no-requeue
#SBATCH --nodes=4
#SBATCH --ntasks-per-node=128
#SBATCH --time=02:00:00
#SBATCH --output=std.out
#SBATCH --error=std.err

QE_root="/u/sbeinlich/Software/QE/qe-7.3.1/intel_standard/"
inputfile="struc.pwt"
outputfile="struc.pwo"

#QE_flags=""
QE_flags="-nk 4"

srun_flags=""
#srun_flags="-n 512"

module purge

unset LD_LIBRARY_PATH

module load intel/2024.0
module load impi/2021.11
module load mkl/2024.0

source "$MKLR00T/env/vars.sh"

# This should ideally not be needed, but QE had trouble linking these libraries correctly...
export LIBS="-lmkl_intel_lp64 -lmkl_sequential -lmkl_core"

# magic environment variable settings similar to raven (might not be needed but unlikely to do harm)

srun $srun_flags $QE_root/bin/pw.x $QE_flags -in $inputfile > $outputfile

"job.sh" 38L, 828B written
```

*QuantumESPRESSO job.sh*

## Documentation:

- PPBwiki (small): <https://ppbwiki.fhi.mpg.de/Main/HPC>
- GIT Wiki (large): <https://gitlab.fhi.mpg.de/fhi-hpc/viper/-/wikis/home>



# VIPER SUPERCOMPUTER – MPCDF



Laptop / Desktop



Computing Infrastructure

1 node on Viper  
~ 4 FHI Compute Servers  
~ 32 Laptops

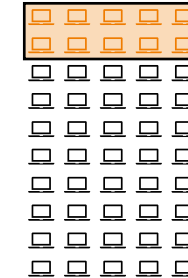
5 jobs x 10 nodes on Viper  
~ 200 FHI Compute Servers  
~ 1600 Laptops



Laptop

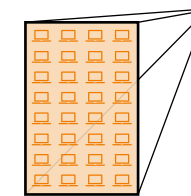
Run 24/7 for 1 month

1 Server

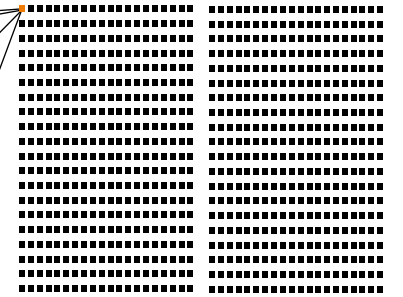


FHI COMPUTE  
x 50

Run 3 days (1 server)



1 Server



VIPER  
x  
25000

Run 22 hours @ 1 server  
or 2 hours @ 10 servers



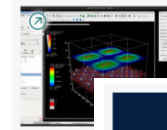
## OTHER MPCDF SERVICES

# OTHER MPCDF SERVICES

*user-based, requires MPCDF account*

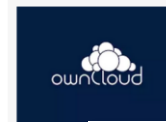
- **Remote Visualization Services**  
(small, shared compute cloud)
  - Remote Desktop (VNC): 12c/256g, shared
  - Remote Visualization (VNC+GPU): 12c/256g/A30, shared
  - Jupyter Lab / Jupyter Notebook: 12c/64g, shared
  - Jupyter Lab + GPU: 12c/64g/A30, shared
  - R Studio
- **DataShare: OwnCloud 200GB**
- **Nexus-S3: S3 Bucket 1TB**
- **Gitlab**
- **Archiving (Tape Backup)**
- **Globus Data Transfer**
- ...

## Remote visualization service



### Compute infrastructure for remote visualization and data analysis

Jupyter notebooks, remote visualization, remote desktops enable convenient web-



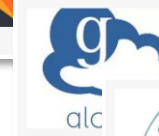
### Sync & Share

With the increasing prevalence of mobile and hand-held devices cloud-based sync & share solutions are getting more and more popular. People appreciate the convenience of services like Dropbox and expect similar offerings in their work environment. To this end the MPCDF has set up an in-house private storage cloud using the OwnCloud



### Gitlab

GitLab is a complete DevOps platform, delivered as a single application. The MPCDF GitLab service offers git repository management, code reviews, issue tracking, activity feeds, wikis and more.



### Data Transfer

**DataHub** - Globus Online based service for reliable large scale data transfers (ranging into the 10s of TBs). Globus online provides a web portal where users can schedule



### Data Publication and Metadata Management

The publication of research data is a complex task. This is due to the heterogeneous landscape of storage systems, different sizes of datasets, permission management and further constraints. The MPCDF supports its users by providing tools for metadata



### Backup

A **backup** is a copy of data that you have, intended for disaster recovery: if you lose your data because of a hardware or software failure or a user mistake, you can recover the lost data from the backup. The backup is typically done automatically and periodically so that you always have a current copy of your data. In addition, a good backup system

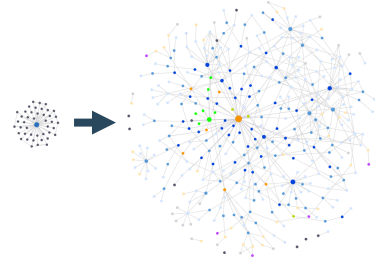


### Archive

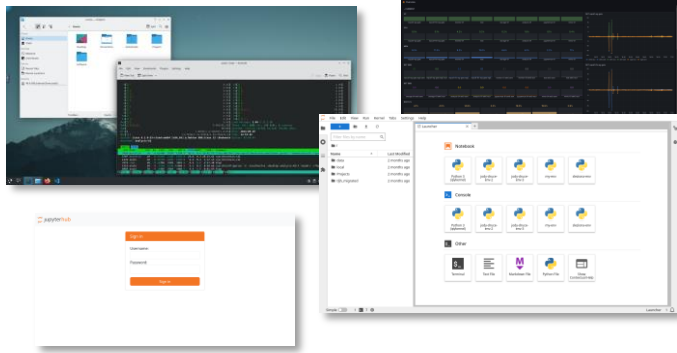
An **archive** is a collection of data that you want to store somewhere other than your local disk because you currently don't need the data anymore, but you might need it again in the future. Or you might have a legal or contractual requirement to keep certain data for very long periods of time, even if the data belongs to projects which are already finished. An archive is usually done by hand by the user, who must decide what data to archive. For safety, a good archive system will automatically store at least two copies of the archived data. [\[documentation\]](#)



# SUMMARY



THANK YOU!



FHI  
(PPB)



MPG  
(MPCDF)

## FHI COMPUTE Servers

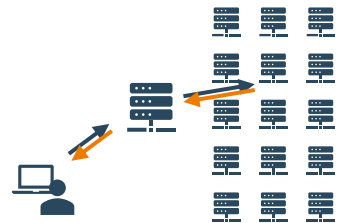
- Large Compute Servers (6x):  
Large CPU + Memory + GPU
- **Virtual Desktops** (Graphical Applications)
- **Jupyterhub** (Python & Matlab Server)
- Terminal Access
- **Direct Execution** (no Scheduling)
- User friendly



*Anything that requires more than a desktop PC*

## VIPER Supercomputer

- Huge Compute Servers (768x):  
Huge CPU + Memory + GPU
- Terminal Access
- **Job Submission** (Scheduling)
- Advanced



*+ RAVEN, SuperMUC, JUWELS, ...*  
*(department-level / grant-based)*

*Large Scale Analyses / Processing  
& Ab-Initio Simulations*