New Open Ondemand Features

Wouter Lampaert

Ehsan Moravveji

VSC User Day - 13 June 2024

Overview

OnDemand Upstream

- OnDemand at KU Leuven
- Globus integration

Updated Apps

- RStudio Server
- JupyterLab
- NVidia RAPIDS

New Apps

- VNC-based apps
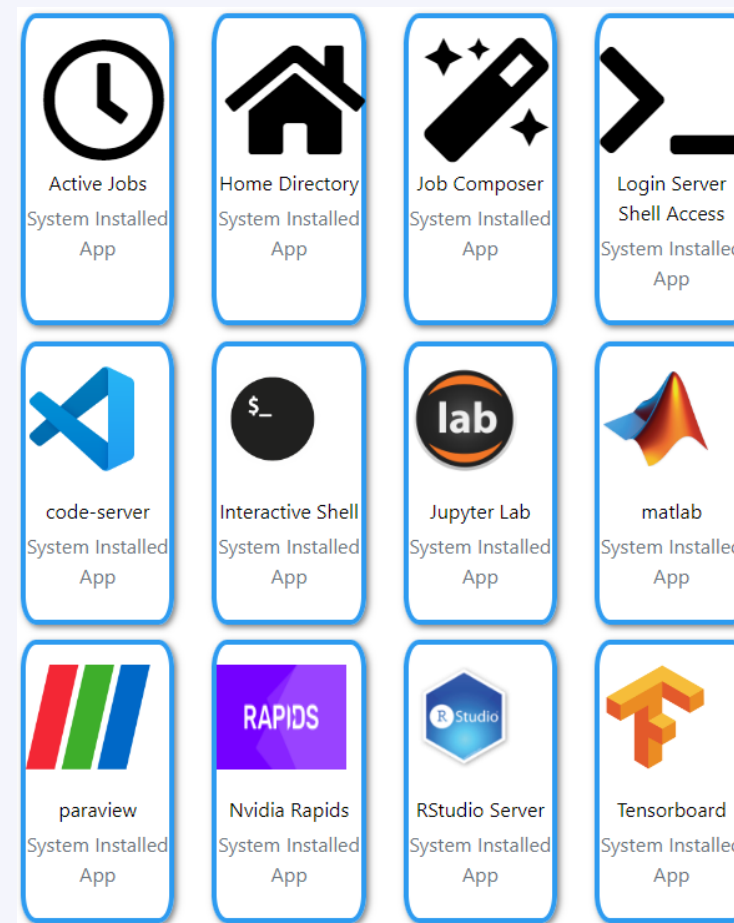 - Matlab
  - ParaView

Beyond OOD

- Scaling up

VLAAMS
SUPERCOMPUTER
CENTRUM

Overview

OnDemand Upstream

- o OnDemand at KU Leuven
- o Globus integration

Updated Apps

- o RStudio Server
- o JupyterLab
- o NVidia RAPIDS

New Apps

- o VNC-based apps
  - o Matlab
  - o ParaView

Beyond OOD

- o Scaling up

VLAAMS
SUPERCOMPUTER
CENTRUM

What is Open OnDemand?

o Web-based Access to HPC facilities
  More info: OOD website, OOD documentation
o Latest release: 3.1.1
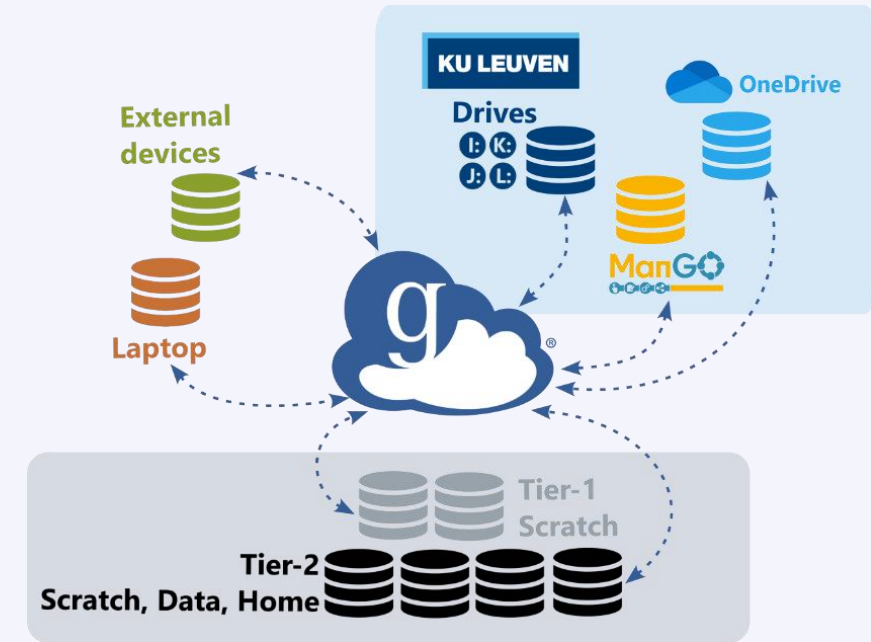
OOD@KULeuven

o Login via https://ondemand.hpc.kuleuven.be
o Authentication: MFA (no SSH keys needed!)
o Basics:
  File/Folder browser, File editor, File transfer (see later)
o Native Apps:Active Jobs, Job Composer, Login &
  Interactive Shell
o Extra Apps (running as jobs):
  Code Server, RStudio, JupyterLab, RAPIDS, Tensorboard,
 MATLAB, ParaView
o Last release: Yesterday!

OPEN OnDemand

| Active Jobs | Home Directory | Job Composer | Login Server Shell Access |
| System Installed App | System Installed App | System Installed App | System Installed App |

| code-server | Interactive Shell | Jupyter Lab | matlab |
| System Installed App | System Installed App | System Installed App | System Installed App |

| paraview | Nvidia Rapids | RStudio Server | Tensorboard |
| System Installed App | System Installed App | System Installed App | System Installed App |

VLAAMS
SUPERCOMPUTER
CENTRUM

Globus

- o Interface to easily transfer (large) files and directories between:
  - o VSC clusters
  - o Remote storage (e.g. K- and L-drive)
  - o Tier-1 data/ManGO
  - o Remote servers
  - o Personal computer
- o Stable and fast transfers with restart options
- o Accessible through globus.org or via the Ondemand File Browser
- o Extensive user documentation available on our Globus VSC-Docs pages!
- o Do you still need FileZilla/WinSCP?

VLAAMS
SUPERCOMPUTER
CENTRUM

Globus: using it from Open Ondemand
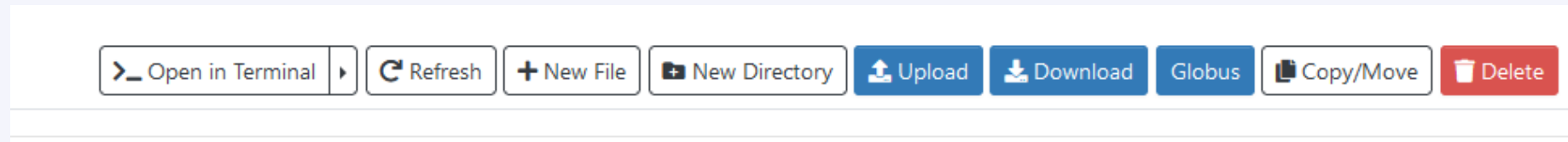
How-to

- Go to 'Files'
- Choose 'Data Directory' (or any sub-directory)
- Click on the 'Globus' button
- Log in to Globus with your KU Leuven ID (if asked)
- Sub-directory will be the starting location

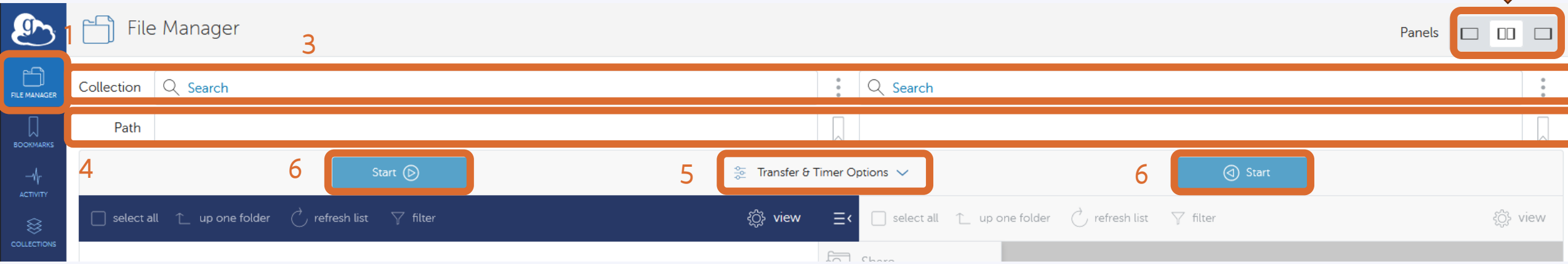Log in to use Globus Web App

Use your existing organizational login

e.g., university, national lab, facility, project

KU Leuven / KU Leuven Association

By selecting Continue, you agree to Globus terms of service and privacy policy.

Continue

VLAAMS
SUPERCOMPUTER
CENTRUM

Globus: how to use it

Steps

1. Select 'File Manager'
2. Choose the two-pane view
3. Select two collections to transfer between.
   o Have a look at our list of managed collections on our Globus VSC-Docs pages
4. Choose the two directories to transfer between
5. Set your transfer and timer options
   o Select syncing if you do not want to overwrite existing files
   o Select quality check protocol
   o Set name for transfer
6. Start the transfer

Globus: local endpoints

o Easily transfer to/from a machine without a managed endpoint
- o Personal computer
- o Local server in department
- o Other HPC clusters without Globus endpoints
- o ...

o While the machine/connection to a remote machine is live, the endpoint is available

o Install Globus Connect Personal to use on any machine

VLAAMS
SUPERCOMPUTER
CENTRUM

Overview

OnDemand Upstream

- OnDemand at KU Leuven
- Globus integration

Updated Apps

- RStudio Server
- JupyterLab
- NVidia RAPIDS

New Apps

- VNC-based apps
  - Matlab
  - ParaView

Beyond OOD

- Scaling up

VLAAMS
SUPERCOMPUTER
CENTRUM

RStudio Server

o RStudio installed on the OS	**Now**
o Latest R was automatically pulled (now v. 4.4.0)	
o But: no extra R packages	
o User package management: difficult and fragile	
o Possible inconsistencies w.r.t. different archs	

o RStudio as a module	**Soon**
o Using R/4.2.2-foss-2022b module	
o Extra (>1,100) packages from CRAN repo	
o Extra (>400) packages from Bioconductor repo	
o Consistent package installation w.r.t. different archs	

o Latest RStudio with R/4.3.3-gfbf-2023b	**Extra**
o Latest R bundle from CRAN Bioconductor	

Toolchain year and R version

2022b: Using R/4.2.2-foss-2022b

☑ Load R-bundle-CRAN module for additional packages

☑ Load R-bundle-Bioconductor module for additional packages

```
$ module load R-bundle-CRAN/2023.12-foss-2022b
$ module load R-bundle-Bioconductor/3.16-foss-2022b-R-4.2.2
$ R
> library(DESeq2)
...
> library(Seurat)
Attaching SeuratObject

Attaching package: 'Seurat'
...
```

VLAAMS
SUPERCOMPUTER
CENTRUM

JupyterLab

o JupyterLab installed on the OS   **Before**
o Python (pre-built) from a container
   Python version: 3.9
o Too dry: no extra Python packages
o User package management: possible with pip
   Fiddling with "–prefix" and "Pre-run scriptlet"
   needed
o User packages are architecture-dependent

o JupyterLab as a module   **Now**
o Choose from 2021a, 2022a, 2023a toolchains
o Extra SciPy-bundle and Matplotlib modules
o Consistent package installation w.r.t. different archs

o Python v3.11.5 from 2023b toolchain   **Soon**
o PyTorch and TensorFlow modules
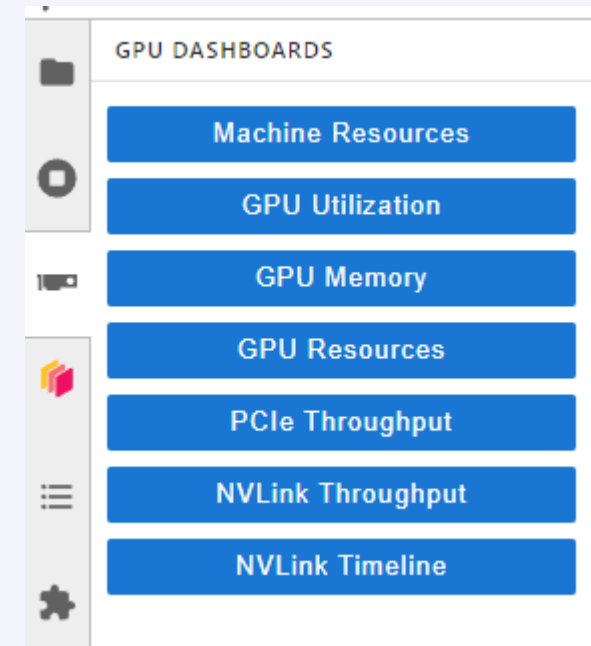o Resource Utilization Dashboard

Toolchain year

2023a

☐ Load the SciPy-bundle module with FOSS

☐ Load the matplotlib module with FOSS

GPU DASHBOARDS

Machine Resources

GPU Utilization

GPU Memory

GPU Resources

PCIe Throughput

NVLink Throughput

NVLink Timeline

Nvidia RAPIDS

- See https://rapids.ai/
- GPU-accelerated data structures and algorithms
- Drop-in replacement for e.g. Pandas, NetworkX, scikit-learn, scikit-image, and more
- For Data Science, Machine Learning and more

- RAPIDS launched from container
- Converted from Docker to Apptainer
- Choosing RAPIDS version
  Container names follow RAPIDSAI convention
- Stable releases are included
- Now: Using JupyterLab
- Kernel: `Python 3 (ipykernel)`

- RAPIDS v.24.04 is incompatible with P100 GPUs

RAPIDS module

```
RAPIDS/24.04-cuda12.2-py3.11                                    ∨
  RAPIDS/22.10-CUDA-11.5-runtime-Ubuntu-18.04-Python-3.9-singularity
  RAPIDS/23.06-cuda11.8-runtime-ubuntu22.04-py3.10-apptainer-jupyter-1.24.0
  RAPIDS/24.04-cuda12.2-py3.11
```

```python
import numpy
import pandas
import cudf

# instantiate data directly in GPU memory
cdf = cudf.DataFrame({'indx': numpy.arange(10),
                      'vals': numpy.random.random(10)})

arr = numpy.ndarray((101, 101, 21), ...)        # in CPU memory
garr = cudf.DataFrame.from_records(arr, ...)    # in GPU memory

pdf = pandas.DataFrame(...)                      # native pandas
gdf = cudf.from_pandas(pdf)                      # convert from Pandas
```

VLAAMS
SUPERCOMPUTER
CENTRUM

Nvidia RAPIDS

o [cudf benchmark](#) on github
o GPU: 1x Nvidia V100
   CUDA v. 12.0
o CPU: 4x CascadeLake core
   Frequency: 2.6GHz
   Memory: 20GB/core
o RAPIDS v. 24.04

Metrics

o Measuring runtime
o Only for function calls
o Data movement (copy, allocation) on GPU is not measured
o Unfair comparison
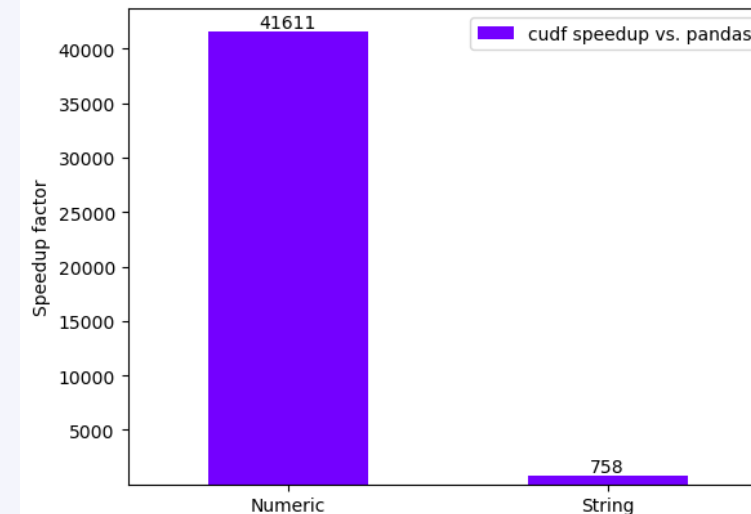
o Runtime_CPU / Runtime_GPU   Speedup

VLAAMS
SUPERCOMPUTER
CENTRUM
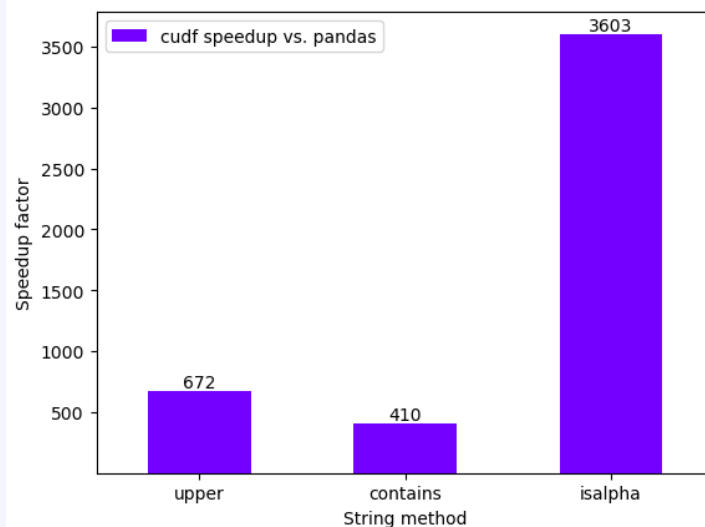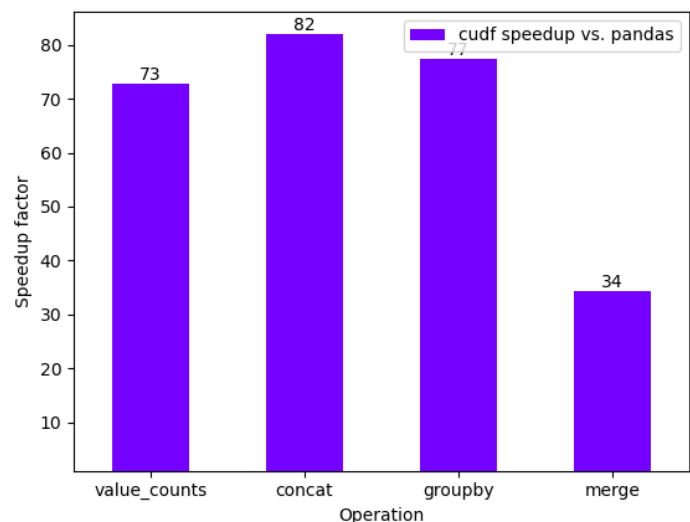
Nvidia RAPIDS

Numerics

- Counting
- Concatenating
- Group-by comparison
- Merge datasets

Strings

- Uppercase
- Contains
- Isalpha

UDF

- UDF: User-defined Function
- Numerics and Strings
- Without JIT overhad (=after warmup)

Overview

OnDemand Upstream

- OnDemand at KU Leuven
- Globus integration

Updated Apps

- RStudio Server
- JupyterLab
- NVidia RAPIDS

New Apps

- VNC-based apps
  - Matlab
  - ParaView

Beyond OOD

- Scaling up

VLAAMS
SUPERCOMPUTER
CENTRUM

VNC-based apps

o   Running interactive apps on a remote desktop in the background

o   (Faster) integration of interactive apps into Open Ondemand

    o   Taking user requests, but... this will need review from our side

o   VNC-hosted apps are minimalistic and contain a single app

    o   Easier maintenance of each app

    o   Problem with one app would not block access to other apps

o   New apps:

    o   MATLAB

    o   ParaView

VNC-based apps: new options

Resolution

- o Available on the form
- o Pre-defined options (Full HD, 2K, 4K)
- o Fine-tune with the sliders at startup, but less recommended

Sharing

- Option to share your session with a link (view-only)

matlab (56265758)   1 node | 4 cores | Running

Host: `r25i13n17.genius.hpc.kuleuven.be`   ⊗ Cancel

Created at: 2024-06-07 14:00:48 CEST

Time Remaining: 59 minutes

Session ID: 7edccf33-e769-417a-97b3-481ac4f197ff
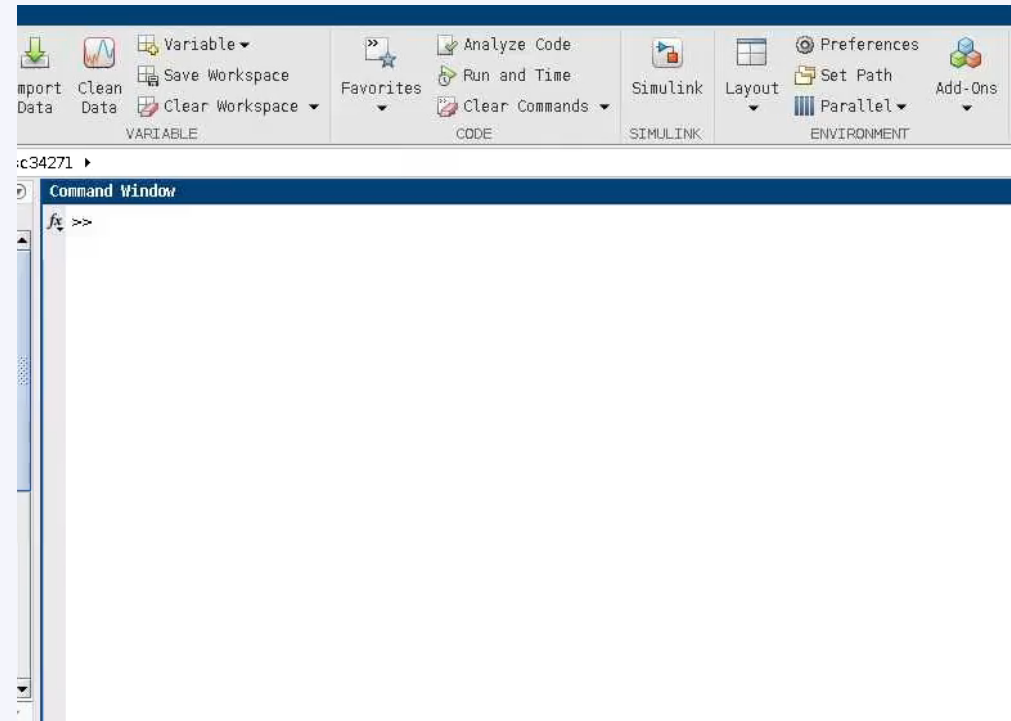
Compression   Image Quality

0 (low) to 9 (high)   0 (low) to 9 (high)

Launch matlab   View Only (Share-able Link)

VLAAMS
SUPERCOMPUTER
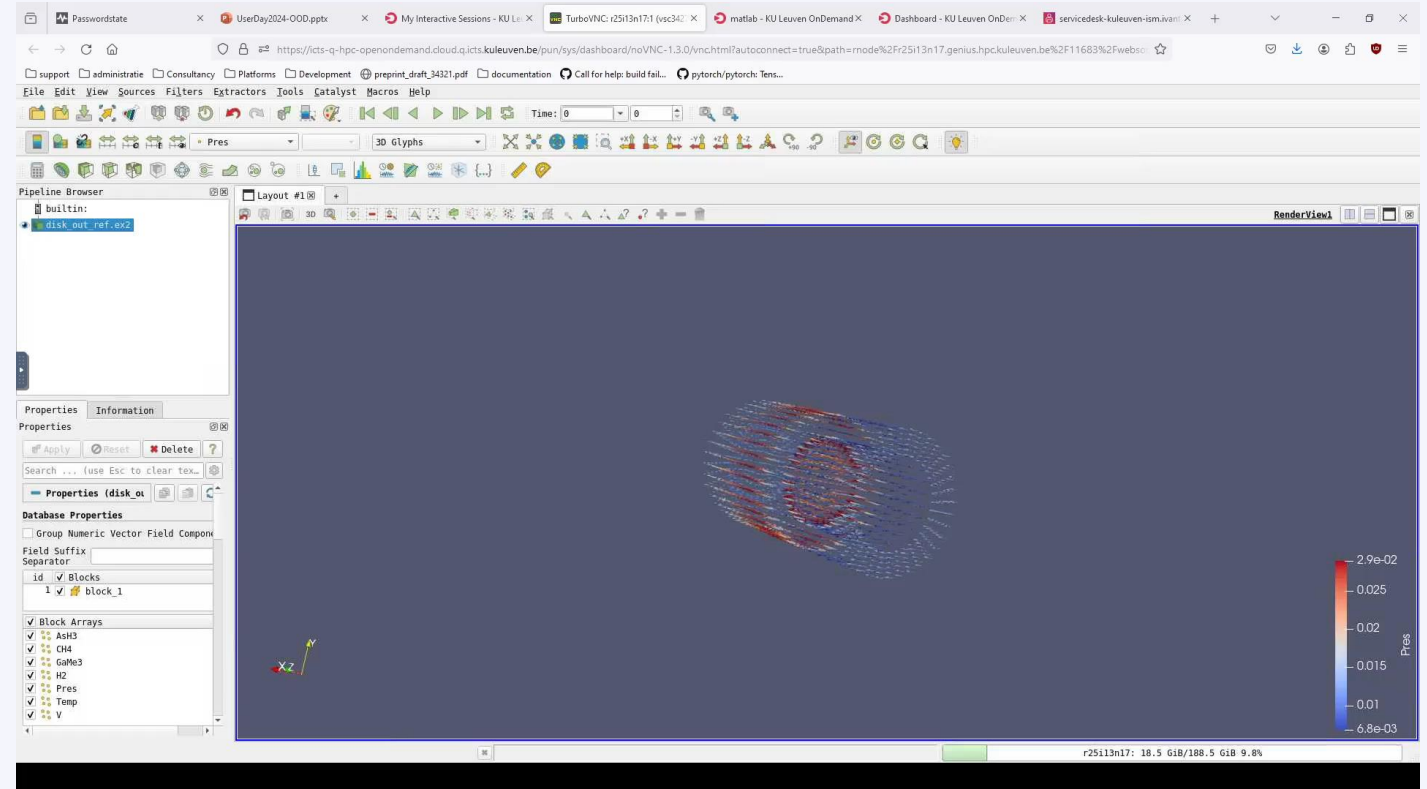CENTRUM

VNC-based apps: MATLAB

- o Current supported versions:
  - o MATLAB/2022a
  - o MATLAB/2022b
  - o Available on Genius and wICE
  - o Access: join `lli_matlab` group
- o Allows running the GUI on compute nodes
- o Improved performance compared to NX session (login node)
- o GPU support

VLAAMS
SUPERCOMPUTER
CENTRUM

VNC-based apps: ParaView

o Current supported versions:

- o 5.6.1
- o 5.9.1-FOSS
- o 5.9.1-Intel
- o 5.10.1-FOSS
- o On both Genius and wICE

VLAAMS
SUPERCOMPUTER
CENTRUM

Overview

OnDemand Upstream

- OnDemand at KU Leuven
- Globus integration

Updated Apps

- RStudio Server
- JupyterLab
- NVidia RAPIDS

New Apps

- VNC-based apps
  - Matlab
  - ParaView

Beyond OOD

- Scaling up

VLAAMS
SUPERCOMPUTER
CENTRUM

Growing out of OOD

- ✓ your dataset is ready
- ✓ your workflow is well tested
- ✓ your scripts are mature
- ✓ Longer walltime needed

THEN

Go to "batch" mode

- o Your input data stays the same
- o Your scripts stay the same

Python

```
#!/bin/bash –l
#SBATCH --account=lp_science
#SBATCH --cluster=wice
#SBATCH --nodes=1 --ntasks=1
#SBATCH --time=2-00:00:00

module load JupyterLab/4.0.5-GCCcore-12.3.0

jupyter execute prepare.ipynb analyse.ipynb postproc.ipynb
```

Load extra modules

Run multiple notebooks

VLAAMS
SUPERCOMPUTER
CENTRUM

Growing out of OOD

IF

- ✓ your dataset is ready
- ✓ your workflow is well tested
- ✓ your scripts are mature
- ✓ Longer walltime needed

THEN

Go to "batch" mode

- o Your input data stays the same
- o Your scripts stay the same

R

```bash
#!/bin/bash –l
#SBATCH --account=lp_science
#SBATCH --cluster=wice
#SBATCH --nodes=1 --ntasks=1
#SBATCH --time=2-00:00:00

module load R/4.2.2-foss-2022b
module load R-bundle-CRAN/2023.12-foss-2022b
Rscript clustering.R
```

VLAAMS
SUPERCOMPUTER
CENTRUM

Growing out of OOD

IF

- ✓ your dataset is ready
- ✓ your workflow is well tested
- ✓ your scripts are mature
- ✓ Longer walltime needed

THEN

Go to "batch" mode

- o Your input data stays the same
- o Your scripts stay the same

RAPIDS

```
#!/bin/bash -l
#SBATCH --account=lp_science
#SBATCH --cluster=wice --partition=gpu_a100
#SBATCH --nodes=1 --ntasks=1 --gpus-per-node=1
#SBATCH --time=2-00:00:00


module load RAPIDS/24.04-cuda12.2-py3.11


apptainer exec --nv $RAPIDSIMAGE python3 data_science.py
```
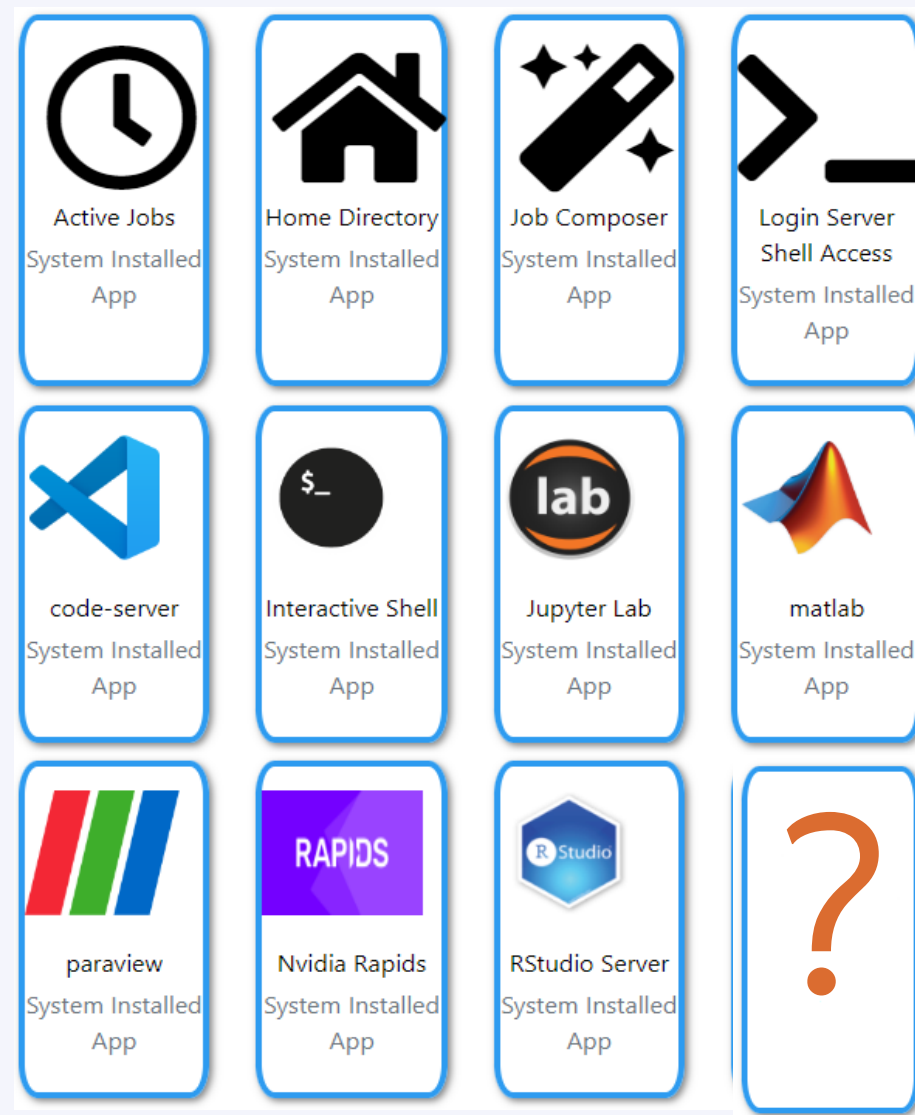
Contains RAPIDS conda env.

Python script must be executable

VLAAMS
SUPERCOMPUTER
CENTRUM

Finally

Work in progress

o You need a new app or feature?
o Contact us via hpcinfo@kuleuven.be
o Development takes time
  testing -> quality -> production
o Next candidates:
  Ansys FLUENT
  Stata
  BAND

VLAAMS
SUPERCOMPUTER
CENTRUM

Stay Connected to VSC

Linked in®

SCAN