

JupyterHub.NRW

Dr. Markus Blank-Burian
WWU Münster



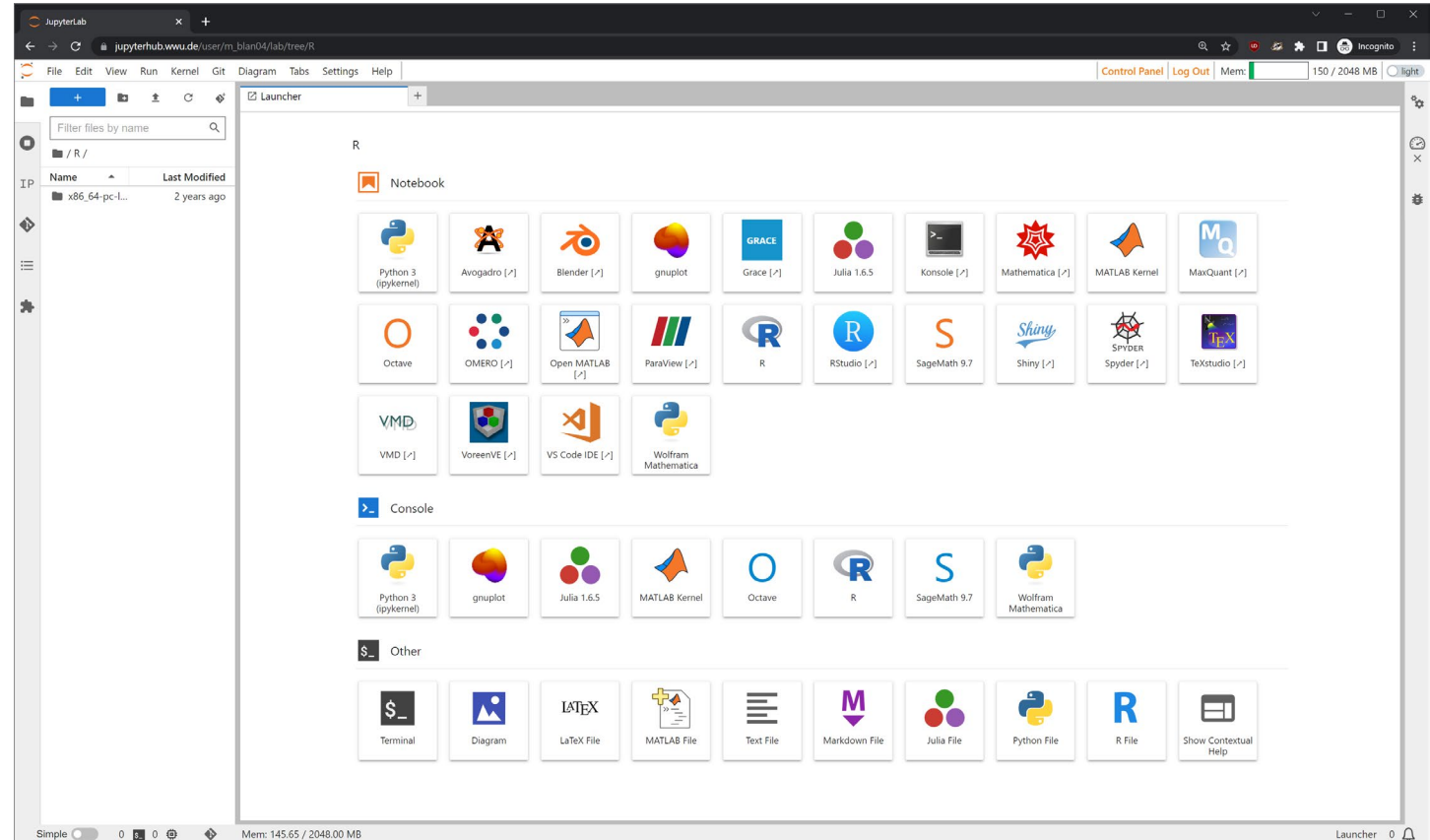
JupyterHub an der Universität Münster



- Im Produktivbetrieb seit 2019
- Ausführung im Kubernetes/OpenStack über KubeSpawner
- Auswahl von benötigten Ressourcen im Spawner
- 2 Images: eScience und Development; Fokus auf Datenanalyse bzw. reine Software-Entwicklung
- Ziel: Virtuelle Arbeitsumgebung
 - Notebooks, IDEs, X11 Anwendungen, Kommandozeile
 - Git-Unterstützung zum verteilten Arbeiten an Programmen
- GPU Unterstützung: CUDA und Remote-Visualization
- Sciebo-Support und Zugriff auf HPC-Dateisysteme sowie Netzwerk-Freigaben im OpenStack

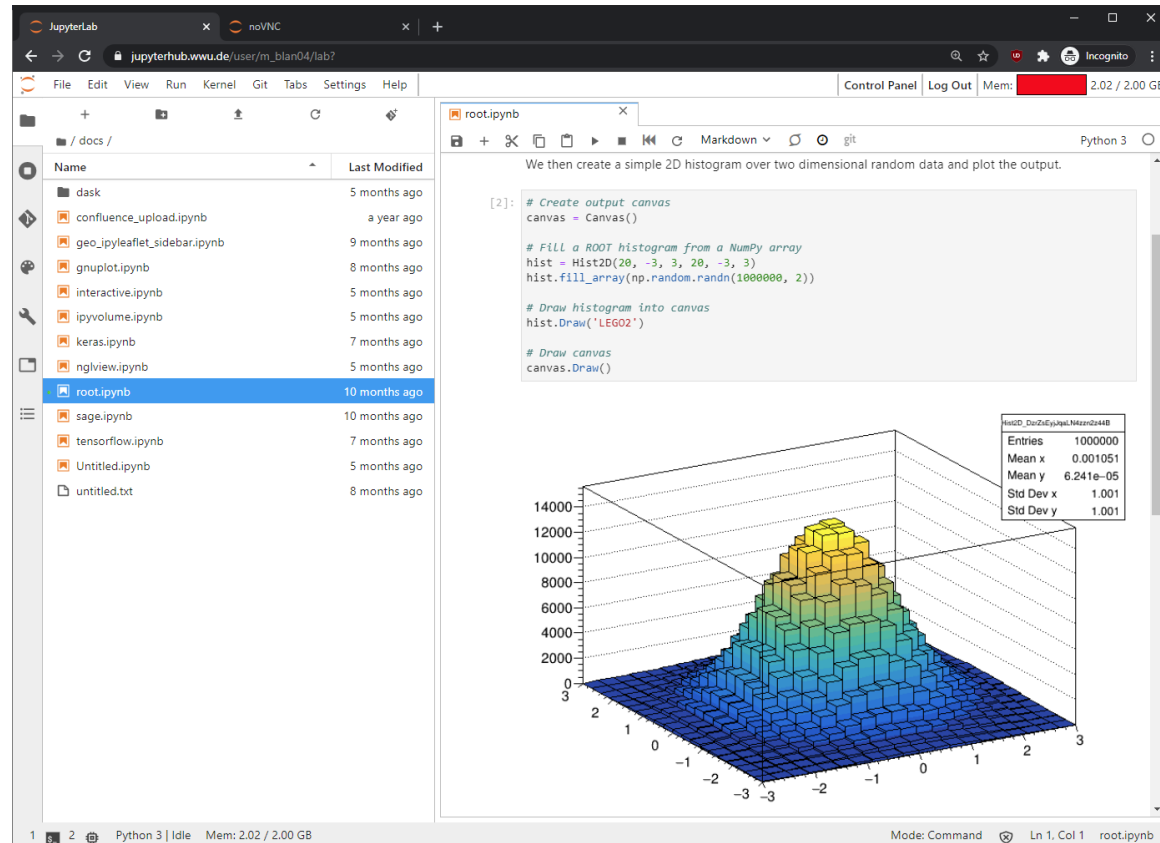
JupyterLab: Viele Anwendungen zur Auswahl

- Launcher
- Treeview
- Tabbed UI
- Git-Unterstützung



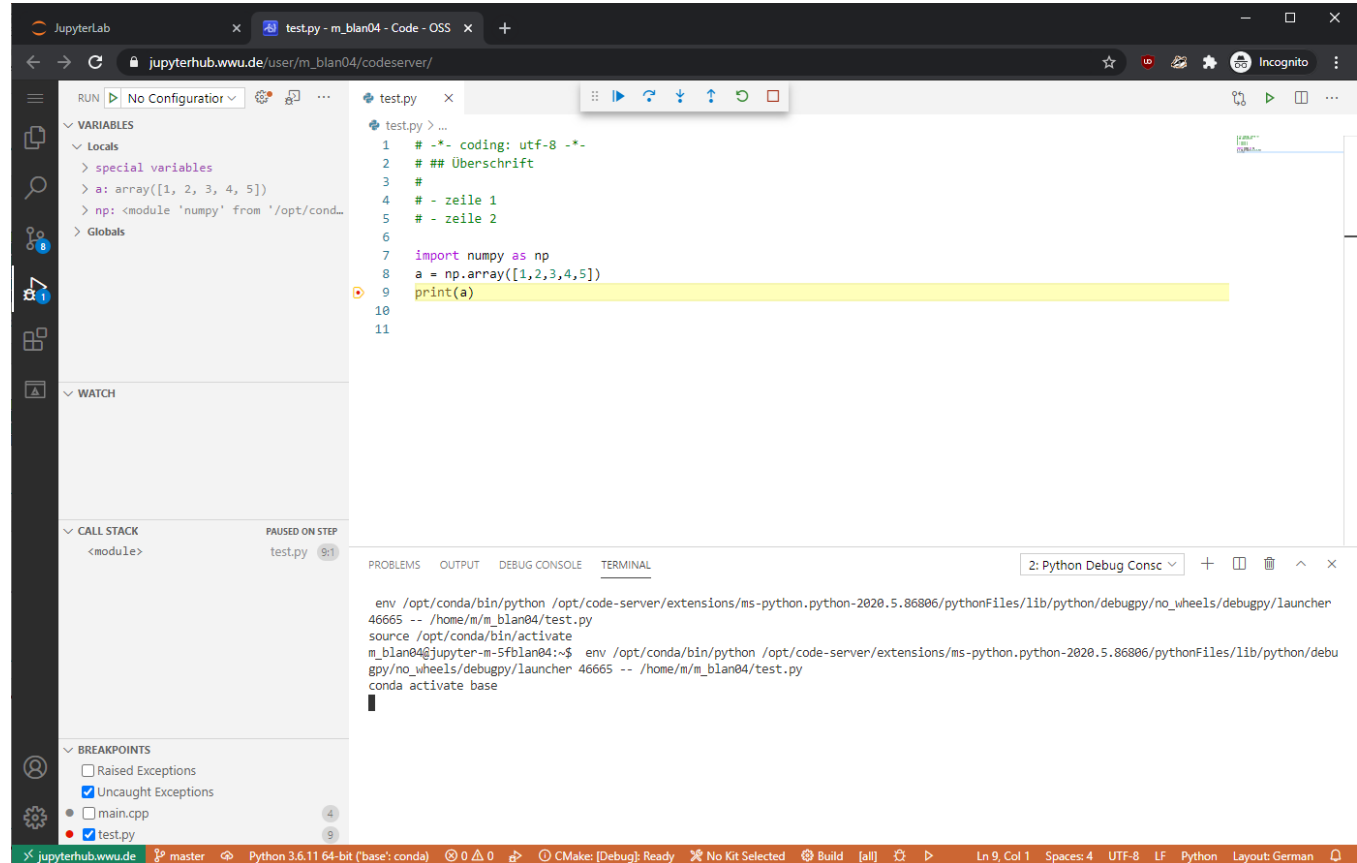
Traditionelle Notebooks

- Julia
- Python
- R
- uvm.



IDE im Browser: Debugging mit Visual Studio Code

- VS Code
 - Python mit Debugging
 - C++ mit Debugging
 - LaTeX Workshop
- RStudio

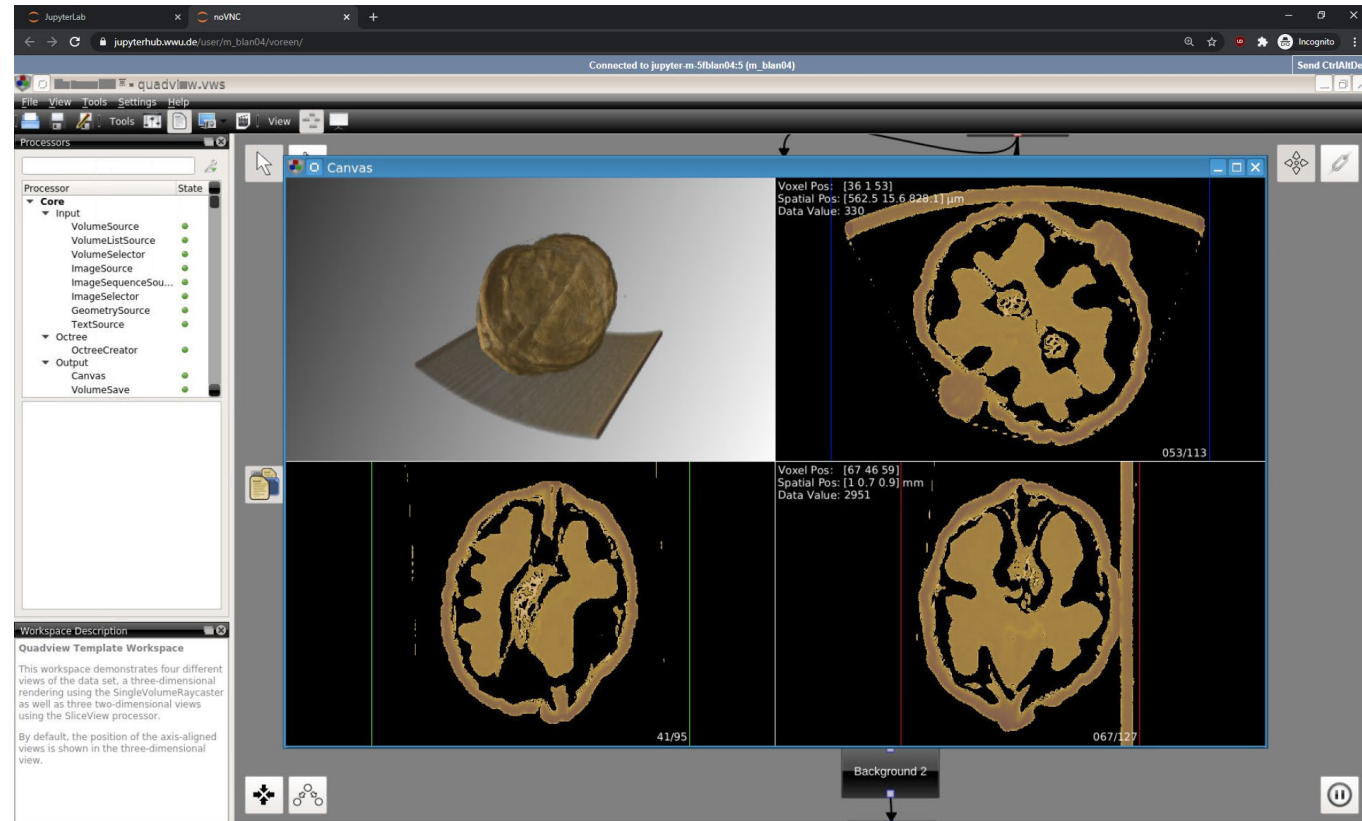


```
test.py > ...
1  # -*- coding: utf-8 -*-
2  ## Überschrift
3  #
4  # - zeile 1
5  # - zeile 2
6
7  import numpy as np
8  a = np.array([1,2,3,4,5])
9  print(a)
10
11
```

env /opt/conda/bin/python /opt/code-server/extensions/ms-python.python-2020.5.86806/pythonFiles/lib/python/debugpy/no_wheels/debugpy/launcher 46665 -- /home/m/m_blan04/test.py
source /opt/conda/bin/activate
m_blan04@jupyter-m-5fblan04:~\$ env /opt/conda/bin/python /opt/code-server/extensions/ms-python.python-2020.5.86806/pythonFiles/lib/python/debugpy/no_wheels/debugpy/launcher 46665 -- /home/m/m_blan04/test.py
conda activate base

X11 Anwendungen mit GPU Unterstützung

- X11 Anwendungen
- CUDA
- Remote Rendering



JupyterHub.NRW: Use-Cases



- Studieren + Forschen:
 - Niederschwellige Arbeitsumgebung zum Erlernen von Werkzeugen und Programmier-/Scriptsprachen
 - Einstiegshürde niedrig durch vorgegebene Umgebungen mit jeweils festem Toolset
 - Keine Konkurrenz zu HPC: Forschende mit höheren Anforderungen an CPU/RAM/GPU sollen HPC nutzen
- Lehren: Verteilen von Notebooks als Übungsaufgaben über öffentliche Git-Repositories
- Bewerten: Kursverwaltung für Verteilung, Abgabe und Bewertung von Notebooks als Übungsaufgaben
- Prüfen: Zunächst **keine** Unterstützung für Prüfungen (Stabilität und Rechtliches)
- LMS-Integration: Vorbereitung von APIs zur Kursverwaltung

JupyterHub.NRW



- Zentrales Angebot von Jupyter Notebook-Images; dezentrale Anpassungen möglich
- Größe der wählbaren Notebooks/Ressourcen/GPUs per Attribute durch IDM.NRW oder Kursverwaltung
- Kursbuchungssystem für große Kurse / Kurse mit GPUs => Gleichmäßigere Auslastung der Ressourcen
- Support:
 - Dezentraler First-Level Support durch Hochschulen (insbesondere durch Kursleiter)
 - Zentraler Support für Kursleiter und zentrale Ansprechpartner der Hochschulen
- Regelmäßige Austauschtreffen mit zentralen Ansprechpartnern der Hochschulen
- Regelmäßige Evaluation und kontinuierliche Weiterentwicklung

JupyterHub.NRW: Meilensteine



- Projektplan vorbehaltlich zeitnaher Stellenbesetzung; Verzögerungen um 6 Monate zu erwarten
- M1 Hardwareintegration: Q2 2024
- M2 Pilotphase: Q3 2024
- M3 Allgemeine Verfügbarkeit: Q1 2025
- M4 Lehrintegration: Q1 2026

JupyterHub.NRW: Hardware

- 10 Compute-Knoten: z.B. 640 Forschende und 2560 Studierende
- 10 GPU-Knoten: z.B. 160 GPU-Sessions mit 12 GB
- 3 Storage-Knoten: z.B. 65.800 Home-Verzeichnisse mit 10 GB und 640x 128 GB Scratch-Space für Forschende

- Integration der Hardware in WWU Cloud / DataInfrastructure.nrw (nächste Folie)
 - => Transparente Mitnutzung der vorhandenen Cloud-Hardware für Storage-Redundanzen und Cloud-Basisdienste
- Erweiterung der Hardware über Folgeantrag nach Bedarf

WWU Cloud / DataInfrastructure.nrw

- Hervorgegangen aus NRW Speicherkonsortium “RDI-NRW”
- OpenSource Cloud Architektur auf Basis von OpenStack, Ceph, Kubernetes
- OpenStack: IaaS Umgebung für virtuelle Maschinen
- Kubernetes: Plattform zum Betrieb von containerisierbaren Diensten (Industriestandard, CNCF)
- Grafana Monitoring Stack (Logs, Metriken, Traces, Alerting)
- Hardware: Storage- und GPU-Knoten; mit JupyterHub.NRW auch Compute-Knoten
- Einfache Skalierbarkeit der Plattform durch vollautomatisierten Softwarestack
- Cloud-Plattform als mögliche Basis für zukünftige NRW Dienste

DataInfrastructure.nrw

JupyterHub.NRW

 **kubernetes**

 **openstack.**

**Vielen Dank für Ihre
Aufmerksamkeit!**