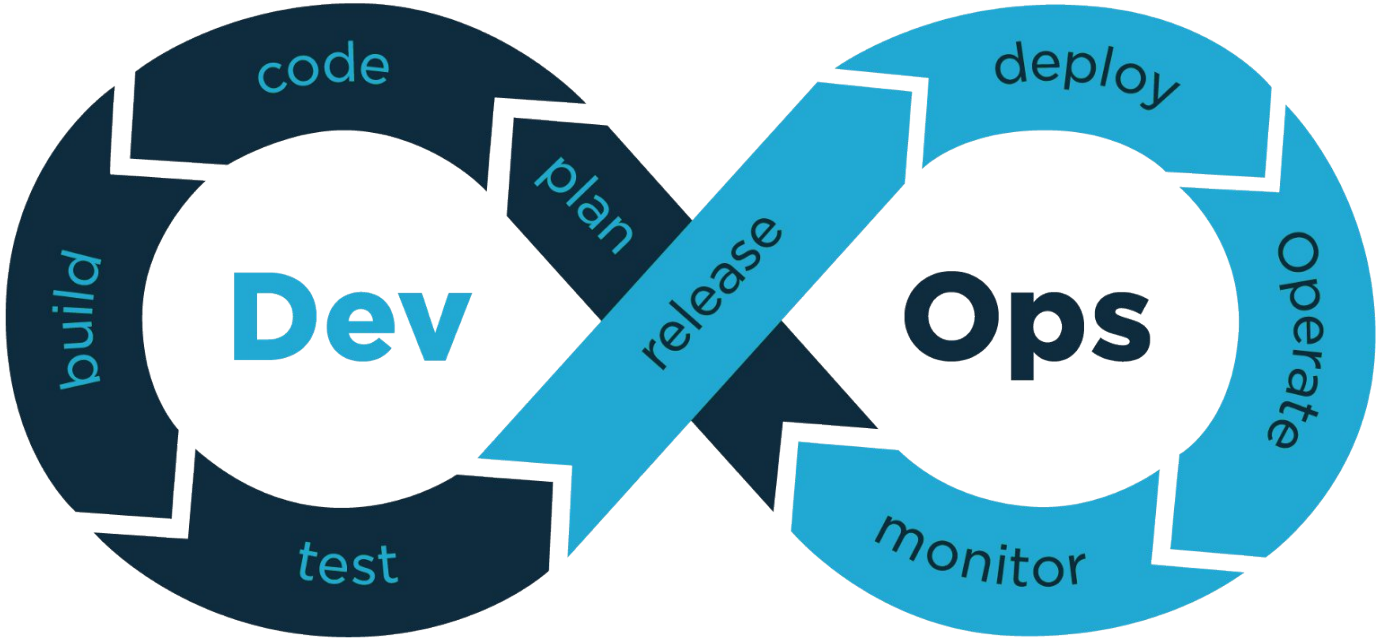
GFZ Research Software Meet-up

GFZ IT infrastructure and service offers for software engineering and computing

From development to operation



Collaborative Coding

- process of working on the code with a team or with another developer
- each team member helps building the code and checks it for bugs or errors
  - ✓ fewer mistakes and bugs
  - ✓ better code quality
  - ✓ projects being completed more quickly
  - ✓ faster debugging and greater project resiliency

➢ GitLab

GitLab

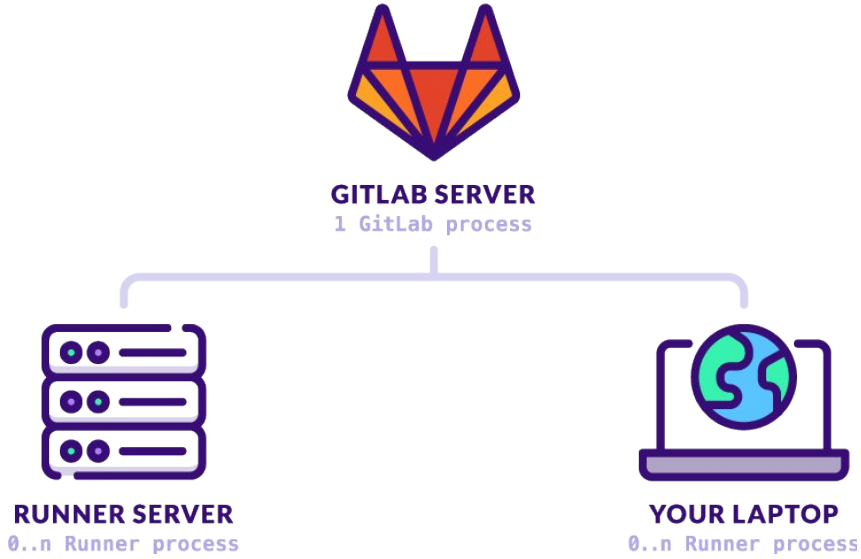
- Community Edition
- 2 instances for different use cases
  - https://git.gfz-potsdam.de
    - login via Helmholtz AAI & GitHub Login
    - ➢ collaboration with the whole world
  - https://git-int.gfz-potsdam.de
    - only for GFZ internal usage
    - ➢ code with high protection needs

Build, Test, Release & Deploy

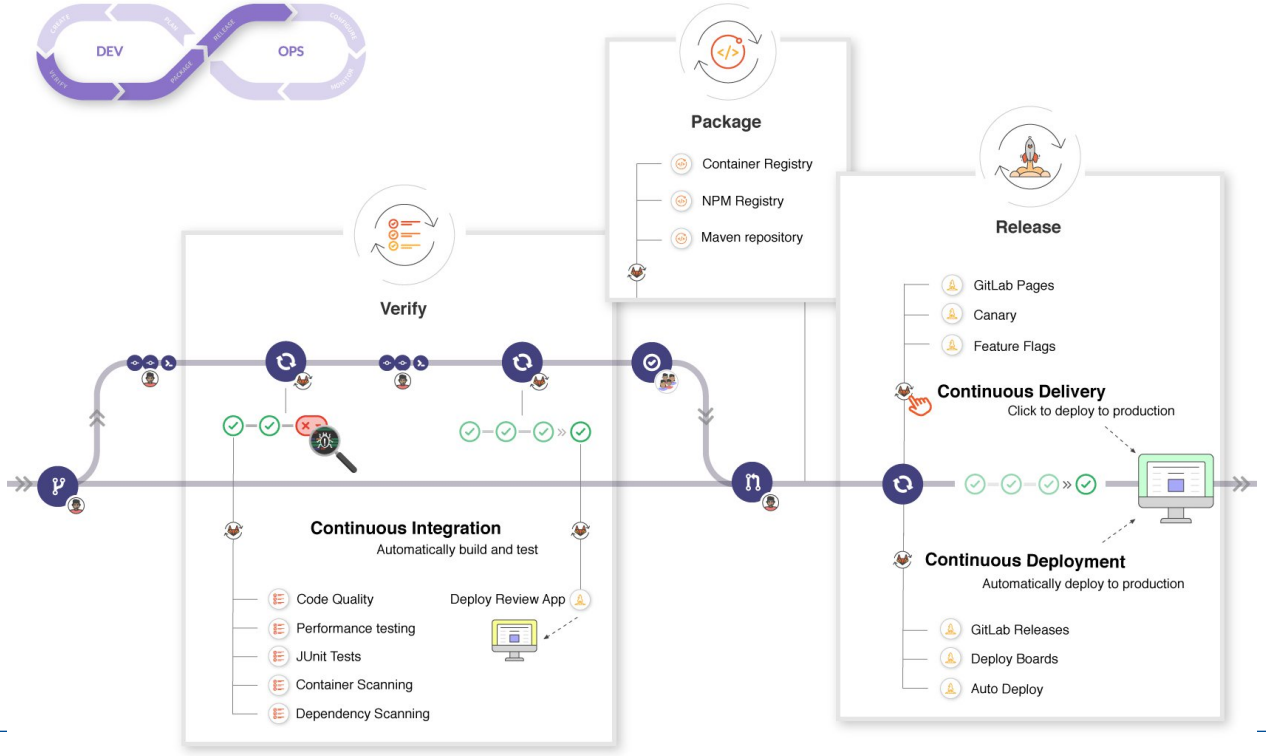
- Continuous Integration
  - process of continuously assembling components into an application
- Continuous Delivery
  - collection of techniques, processes, and tools that improve the software delivery process
- Continuous Deployment
  - software engineering approach in which software functionalities are delivered frequently and through automated deployments

GitLab CI/CD

- GitLab Runner
  - could be any VM, Server, Laptop
  - after pushing commit(s) to GitLab
    1. runner clones fresh copy of repository
    2. runner starts build jobs described in gitlab-ci.yml (parallel jobs possible)
    3. depending on your job description and code language
       1. tests are run (e.g. pytest, junit, boost, ...)
       2. binary packages are build
       3. documentation is created (and deployed via git-pages)
       4. everything that is done locally can also be automated in the runner



DevOps with GitLab

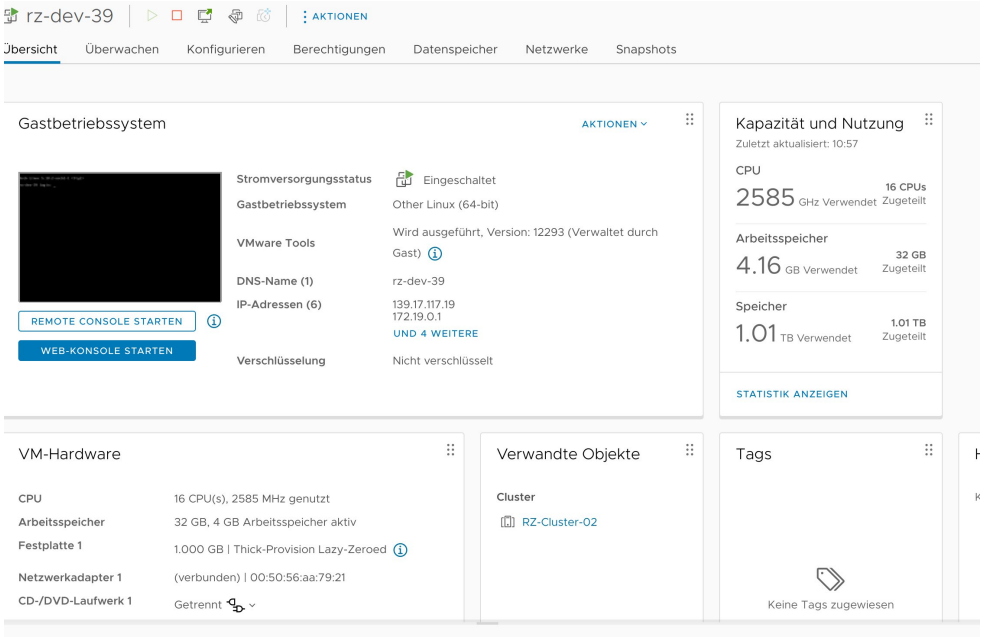


Operating

- Server virtualization via VMWare vSphere
  - ✓ existing resources are better utilized
  - ✓ reduction of costs in aspects of energy efficiency and space requirements
  - ✓ high availability

  - currently 33 Server with 3,14THz, 24TB RAM, 166TB HDD hosting 509 virtual machines
  - default VM upon request via intranet: 1 CPU, 4GB RAM, 50GB HDD
    - http://intranet.gfz-potsdam.de/en/topics/it/server-and-services/virtualization/
  - work via remote console, as if you were at a physical PC

vSphere


rz-dev-39
AKTIONEN
Übersicht
Überwachen
Konfigurieren
Berechtigungen
Datenspeicher
Netzwerke
Snapshots
Gastbetriebssystem
AKTIONEN
REMOTE CONSOLE STARTEN
WEB-KONSOLE STARTEN
Stromversorgungsstatus
Eingeschaltet
Gastbetriebssystem
Other Linux (64-bit)
VMware Tools
Wird ausgeführt, Version: 12293 (Verwaltet durch Gast)
DNS-Name (1)
rz-dev-39
IP-Adressen (6)
139.17.117.19
172.19.0.1
UND 4 WEITERE
Verschlüsselung
Nicht verschlüsselt
Kapazität und Nutzung
Zuletzt aktualisiert: 10:57
CPU
2585 GHz Verwendet
16 CPUs Zugeteilt
Arbeitsspeicher
4.16 GB Verwendet
32 GB Zugeteilt
Speicher
1.01 TB Verwendet
1.01 TB Zugeteilt
STATISTIK ANZEIGEN
VM-Hardware
CPU
16 CPU(s), 2585 MHz genutzt
Arbeitsspeicher
32 GB, 4 GB Arbeitsspeicher aktiv
Festplatte 1
1.000 GB | Thick-Provision Lazy-Zeroed
Netzwerkadapter 1
(verbunden) | 00:50:56:aa:79:21
CD-/DVD-Laufwerk 1
Getrennt
Verwandte Objekte
Cluster
RZ-Cluster-02
Tags
Keine Tags zugewiesen

HPC at GFZ

Overview of available resources and possible workloads

Marco Ende & Max Rutkowski
rz-support@gfz-potsdam.de

Resources

- ~180 Servers
  - up to 1 TB RAM/Server
  - up to 72 CPU Cores/Server
  - Infiniband Interconnect (up to 100 Gbit/s)
  - Slurm as workload manager
- GPUs
  - currently: 48 (A100, A40)
  - soon: +32 (A100 SXM, A30, A40)

- Access
  - https://forms.gfz-potsdam.de/en/it-sb/cluster-usage
  - 10 TB storage/user (more on request, no backup!)
- MATLAB Parallel Server
  - 64 Workers
  - 2 Servers with 750 GB RAM each
  - 2023: 2 Servers with 4x A100 and 2 TB RAM each

Workloads – First View

- HPC Candidates
  - Workstation under your desk?
  - You’re waiting a couple hours for results?
  - Dedicated servers running for your once a year workload?
- Linux compatible software that needs CPU/GPU/RAM/Storage/Time
- Parallel Computations („Workers“, MPI)
- Serial Computations
- MATLAB Processing

Workloads – Second View

- gcc, intel oneAPI, clang, Nvidia HPC SDK (formerly PGI)
- Python, R, Julia
- Jupyter, ParaView
- cuDNN, PyTorch, Tensorflow, Keras, Horovod, Ncview, CDO
- Package managers and environments like
  - pip & venv
  - Spack, easybuild
  - miniconda
  - in future: docker-based containers

- More here: https://git-int.gfz-potsdam.de/hpc/user-manual/-/wikis/home

Happy Computing!