



WHAT CAN HPC CLOUD DO FOR YOU?

What can you do for the HPC cloud?

Frank Berghaus
for the HPC Cloud Team

- Features of the HPC Cloud

- Compute

- *Images*
- *Flavors*
- *Server Groups*

- Volumes, i.e. Block storage

- Network

- *Networks*
- *Security Groups*
- *Load Balancers (not discussed)*
- *Floating IPs (not discussed)*

- Orchestration

- Object storage

- Shares, i.e. shared file systems

- Recipes for the HPC Cloud

- Kubernetes (MoM)

- SLURM (JADE)

- Workstations

- VPN (?)

- Services based on HPC Cloud

- Notebooks:

- *BinderHub & JupyterHub*

- GitLab runners

- Virtual cluster

- *Example: remote visualization service (RVS)*

- Open Build Service (OBS)

- Features of the HPC Cloud

- Compute

- *Images*
- *Flavors*
- *Server Groups*

- Volumes, i.e. Block storage

- Network

- *Networks*
- *Security Groups*
- *Load Balancers (not discussed)*
- *Floating IPs (not discussed)*

- Orchestration

- Object storage

- Shares, i.e. shared filesystems

- Recipes for the HPC Cloud

- Kubernetes (MoM)
- SLURM (JADE)
- Workstations
- VPN (?)

- Services based on HPC Cloud

- Notebooks:
 - *BinderHub & JupyterHub*
- GitLab runners
- Virtual cluster
 - *Example: remote visualization service (RVS)*
- Open Build Service (OBS)

FEATURES → COMPUTE → IMAGES



- **EL-Like:**
 - Alma
 - Centos
 - RHEL (BYOL)
- **Debian-like**
 - Debian
 - Ubuntu
- **SUSE-like**
 - OpenSUSE
 - SLES (BYOL)
- **Wishlist?**
 - OS's: Arch, Container OS's, Windows
 - Local repo mirrors / regular updates
 - Pre-installed "appliances" (e.g. Prometheus)

Name	Type	Status	Visibility	Protected	Disk Format	Size	Launch
AlmaLinux 8	Image	Active	Public	Yes	RAW	10.00 GB	Launch
AlmaLinux 9	Image	Active	Public	Yes	RAW	10.00 GB	Launch
CentOS Stream 8	Image	Active	Public	Yes	RAW	10.00 GB	Launch
CentOS Stream 9	Image	Active	Public	Yes	RAW	10.00 GB	Launch
Debian 11	Image	Active	Public	Yes	RAW	2.00 GB	Launch
Debian 12	Image	Active	Public	Yes	RAW	2.00 GB	Launch
openSUSE Leap 15.4	Image	Active	Public	Yes	RAW	851.00 MB	Launch
openSUSE Leap 15.5	Image	Active	Public	Yes	RAW	809.00 MB	Launch
openSUSE Leap 15.6	Image	Active	Public	Yes	RAW	839.00 MB	Launch
Ubuntu 20.04	Image	Active	Public	Yes	RAW	2.20 GB	Launch
Ubuntu 22.04	Image	Active	Public	Yes	RAW	2.20 GB	Launch



- OS versions removed at EOL
- Updates are manual / irregular



FEATURES → COMPUTE → FLAVORS



- Defined by mapping to hardware
- Shared (default):
 - Two VCPUs share a physical CPU
- Dedicated (on request):
 - VCPUs are pinned to physical CPUs
- Devices
 - GPUs: A30/A40, A100, H100
 - NVMe: 1.8 TB & 3.6 TB
- Wishlist:
 - GPU flavor sizing??
 - Hyper-threaded flavors

RAM per VCPU [GiB]

	shared	dedicated
standard	4	6.7
high	8	---
extreme	16	30

Maximum size

	shared	dedicated	whole-node
standard	64c256g	36c240g	72c480g
high	64c512g	---	---
extreme	32c512g	32c960g	64c1920g



- Hint to the OpenStack scheduler

```
o server group create demo --policy soft-affinity
o server create demo --server-group demo ...
```

policy	effect
soft-affinity	Attempt to schedule instances near each other, for example on the same hypervisor
soft-anti-affinity	Attempt to schedule instances on different hardware, for example separate racks



- Block storage provided by Ceph
- Three variants:
 - 1) Ceph
 - 2) CephSSD
 - 3) CephIntensive
- Little difference in performance
- Faster block devices
 - Flavors with NVMe
- Wishlist
 - NVMeOF (?)

- Volume as root disk
 - Create a volume from an image
 - Mark the volume “bootable”
 - Create server from the volume
- Multi-attach volumes
 - Available on request
 - Must set up parallel file system



- Networks shared by all projects
 - cloud-local-1
 - *very full*
 - cloud-local-2
 - *full*
 - cloud-local-3
- Public IPs
 - cloud-public
- You can define private networks
 - Network
 - Subnet
 - Router



- Special networks for shares
 - Access given on request

```
o network create NETWORK
o subnet create SUBNET \
  --network NETWORK \
  --subnet-range 192.168.0.0/24 \
  --dns-nameserver 130.183.9.32 \
  --dns-nameserver 130.183.1.21
o router create ROUTER
o router set ROUTER --external-gateway cloud-public
o router add subnet ROUTER SUBNET
```




- Firewall rules
- We setup up three for you
 - default
 - web
 - web-campusonly
- Notes on default group
 - Allows ingress from your hosts
 - Blocks ingress from other projects
 - Allows SSH from campus

- AutoDNS:
 - DNS entry for FIPs and Servers
 - Runs every 5 minutes
 - Servers:

VM_NAME.PROJECT_NAME.hpcloud.mpg.de

- FIPs:

FIP_DESCRIPTION.PROJECT_NAME.hpcloud.mpg.de

- Certificates
 - CertBot & Let's Encrypt
 - CA of your choice



- Addressing the storage
 - S3: Amazon's protocol
 - *de facto standard*
 - SWIFT: OpenStack protocol
- Credentials issues by OpenStack
 - Provide full access to project space
- Globally accessible
- Wishlist:
 - How to make this useful to you?

Paul Müller, Wed @ 9:00

Storage in the Cloud, Wed @ 15:00

- Share content
 - Handle access control in separate application
 - Generate links with temporary access tokens



- Shared file system as a service
- Three flavors:
 - NexusPOSIX
 - *Native on HPCs*
 - *Configured for large data files*
 - *Uses MPCDF user & group IDs*
 - *Projects can be exported to cloud via NFS*
 - Manila & NFS
 - *Easy to set up*
 - *Widely compatible*
 - *Low performance (bottleneck)*
 - Manila & CephFS
 - *Requires compatible client*
 - *More configuration*
 - *Good performance*

- Wishlist:
 - Recipes / examples (?)

Storage in the Cloud, Wed @ 15:00





HEAT

- Defined in yaml
- State tracked in OpenStack
- *Example:* Kubernetes (MoM)

Terraform

- Defined in hcl
- State local to client
 - Shared state via GitLab
- *Example:* SLURM (JADE)

- Define compute, storage, and network infrastructure as code
- Collaborate, reproduce, or share systems of compute resources

Automation & Infrastructure as code, Wed @ 13:30

- Features of the HPC Cloud
 - Compute
 - *Images*
 - *Flavors*
 - *Server Groups*
 - Volumes, i.e. Block storage
 - Network
 - *Networks*
 - *Security Groups*
 - *Load Balancers (not discussed)*
 - *Floating IPs (not discussed)*
 - Orchestration
 - Object storage
 - Shares, i.e. shared filesystems

• Recipes for the HPC Cloud

- Kubernetes (MoM)
- SLURM (JADE)
- Workstations
- VPN (?)

• Services based on HPC Cloud

- Notebooks:
 - *BinderHub & JupyterHub*
- GitLab runners
- Virtual cluster
 - *Example: remote visualization service (RVS)*
- Open Build Service (OBS)



- **MoM Recipe:**

 - gitlab.mpcdf.mpg.de/mpcdf/cloud/kubernetes

- **Two branches:**

 - **production:**

 - *Meant to host productive services*
 - *Three control plane nodes*
 - *Problem: warnings from etcd*

 - **test:**

 - *Meant to try out Kubernetes*
 - *Single control plane*

Kubernetes, Containers etc., Wed @ 13:00

- **Use:**

 - Notebooks

 - Many projects, at the workshop:

 - *Flaski*
 - *Nomad*
 - *SciServer*

- **Wishlist**

 - Improve existing recipe

 - Other deployment options (?)

 - *Magnum*
 - *Rancher*
 - *ClusterAPI*



• **JADE Recipe:**

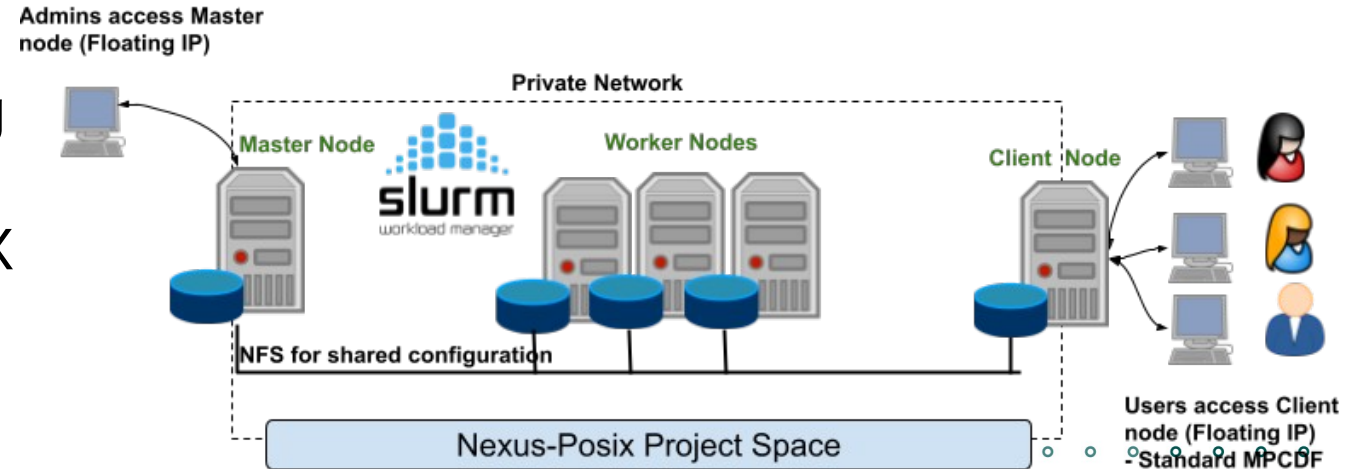
→ gitlab.mpcdf.mpg.de/mpcdf/cloud/jade

• **Slurm Cluster in the Cloud**

- Full control of your cluster config and software environment
- Shared Storage on NexusPOSIX
 - *Also mounted on Raven or Cloud VMs*
- Good for single-node workloads
- Terraform based
 - *Scale out/in easily*
 - *Cleanly spin up or tear down clusters*

Clusters in the Cloud, Wed @ 15:00

JADE: High-level architecture



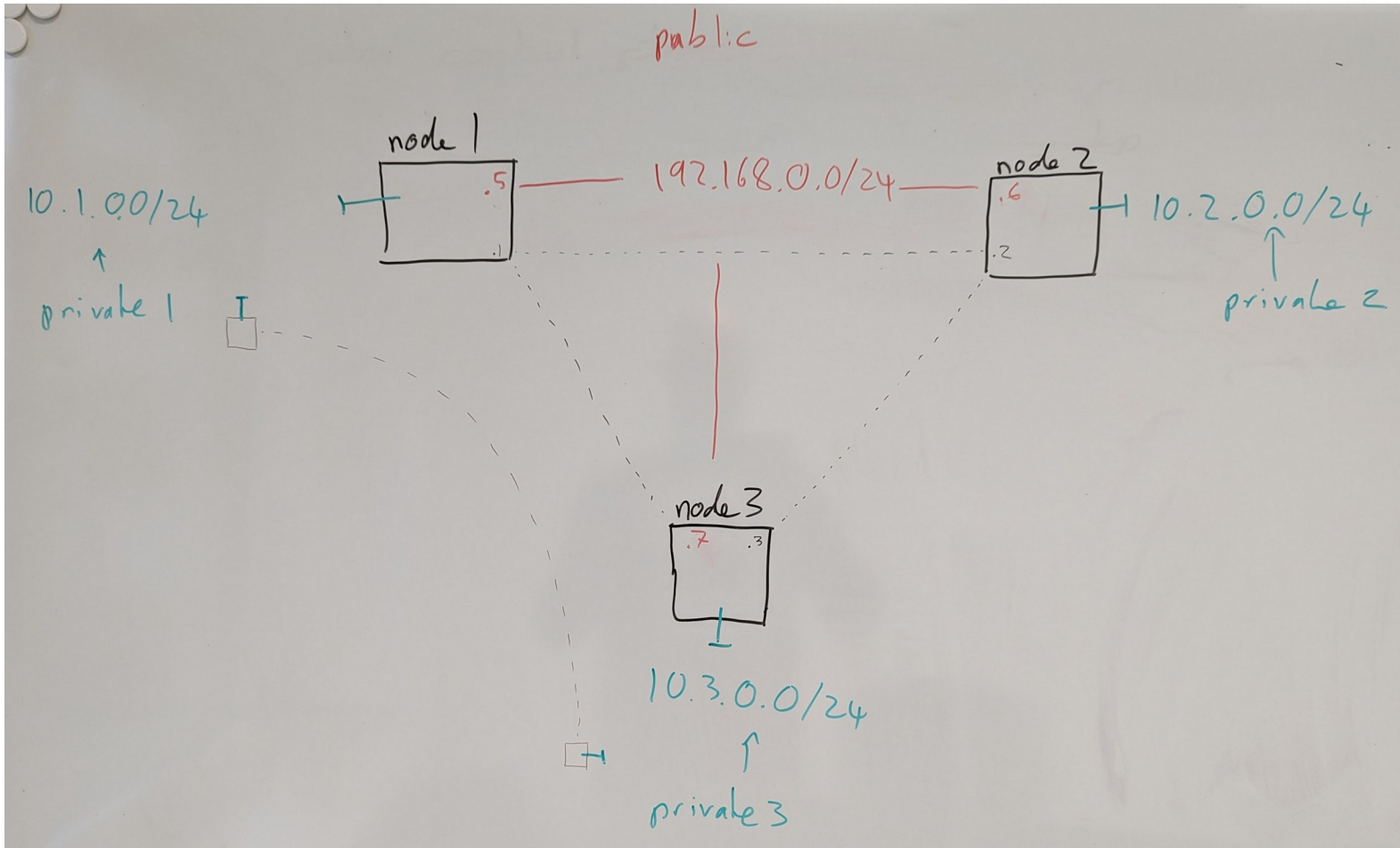
Core features:

- All nodes in Private Network
 - Master/Client have public IP
- Standard MPCDF user accounts (LDAP, Kerberos)
- Nexus-Posix mounted on all nodes (possible use from RAVEN too)
- Can scale out/up (increase Workers number or size)



- Workstations in the cloud
 - High degree of control over infrastructure
 - *Size of VM*
 - *Storage*
 - Local or MPCDF Accounts (LDAP)
 - NexusPOSIX mount possible
 - Ideal for long running post processing jobs
 - *Production on Raven, data store in NexusPOSIX, post processing on workstation*
 - Possible mount of block storage volumes
 - Full control over software environment
 - Remote Desktop (VNC)
 - Access via SSH through Gateways





- WIP Example:
[gitlab.../fberg/vpn-demo](https://gitlab.com/fberg/vpn-demo)
- Wishlist:
 - Recipe (WG mesh)
 - OpenStack VPNaaS
- Interest (?)

Maurits & Simeon, Tue @ 16:30

- Features of the HPC Cloud

- Compute

- *Images*
- *Flavors*
- *Server Groups*

- Volumes, i.e. Block storage

- Network

- *Networks*
- *Security Groups*
- *Load Balancers (not discussed)*
- *Floating IPs (not discussed)*

- Orchestration

- Object storage

- Shares, i.e. shared file systems

- Recipes for the HPC Cloud

- Kubernetes (MoM)

- SLURM (JADE)

- Workstations

- VPN (?)

- Services based on HPC Cloud

- Notebooks:

- *BinderHub & JupyterHub*

- GitLab runners


- Virtual cluster

- *Example: remote visualization service (RVS)*

- Open Build Service (OBS)



- Service:
 - notebooks.mpcdf.mpg.de/binder
- Binder:
 - Run notebooks from git repo
 - Restricted to MPCDF GitLab
- JupyterHub
 - Create JupyterHubs for projects
 - e.g. Workshops
- Wishlist
 - Authentication
 - *EduGAIN for Binder*
 - Plugin to Storage (?)



Turn a Git repository into a
Have a repository full of
executable environments
New to Binder?

Build and launch a repository

MPCDF GitLab repository URL (https://gitlab.mpcdf.mpg.de)

MPCDF GitLab

Git ref (branch, tag, or commit)

HEAD

Copy the URL below and share your repository

Fill in the fields to see the repository

Expand to see the text below, page 4 of 4

4.4.2. BioEn result for full calculated intensities

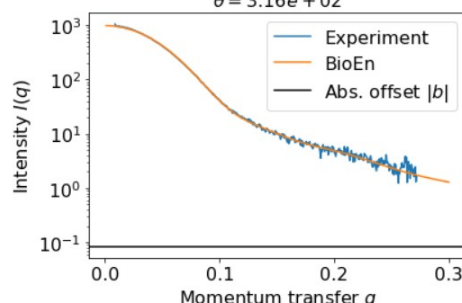
We next show our optimal BioEn refinement for the calculated intensities over their full q-range. The full intensity contains more information and can, for example, be used to determine the radius of gyration. Additionally, we show the absolute value of the offset as a horizontal line as a consistency check.

```
M_full = intensities[0].shape[0]
q_list = intensities[0][:,0]
y_full=np.zeros((M_full, N))
for i in range(N):
    y_full[:,i] = intensities[i][:, 1]

y_full_fit, dummy = saxes.update_y(a, b, y_full, yTilde)
Iq_ave = (wopts[idx][:,np.newaxis]*np.asarray(y_full_fit.T)).sum(axis=0)

plt.plot(Iq_exp[:,0], Iq_exp[:,1], label="Experiment")
plt.plot(q_list, Iq_ave, label="BioEn")
plt.axhline(np.abs(b), color='k', label="Abs. offset |b|$")
plt.xlabel("Momentum transfer $q$")
plt.ylabel("Intensity $I(q)$")
plt.yscale('log')
plt.legend(fontsize=16)
plt.title(r"$\theta = %3.2e$" % results[idx,0], size=16)
```

Text(0.5, 1.0, '\$\theta = 3.16e+02\$')

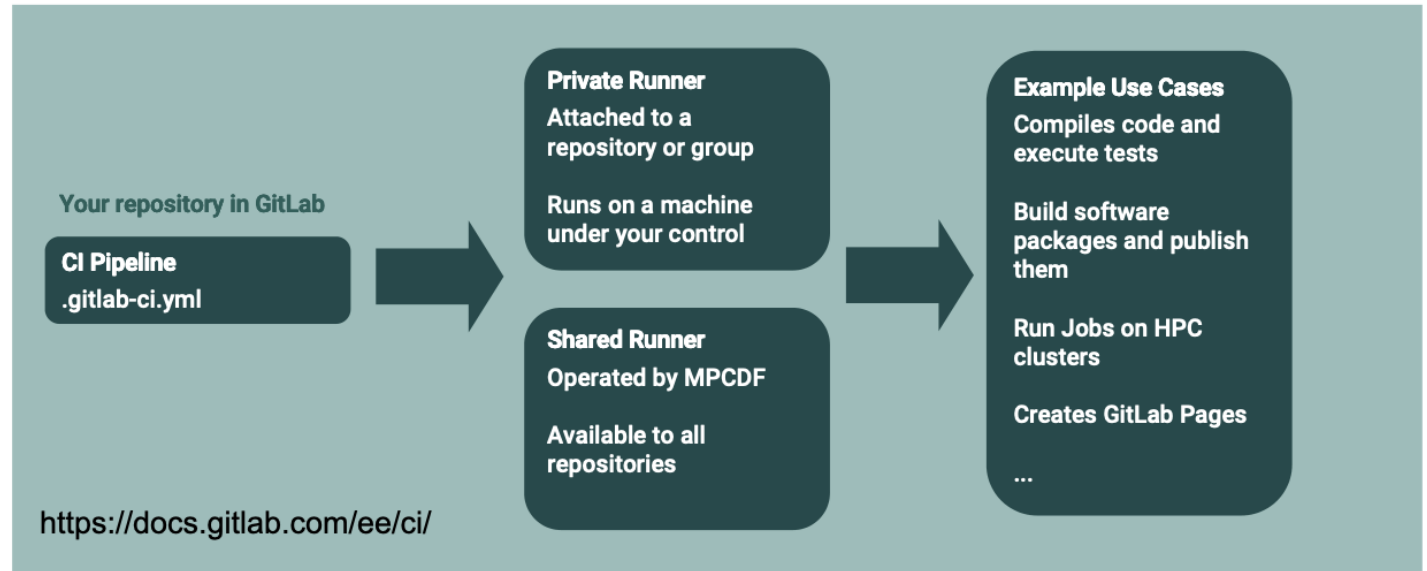




- Service:
 - gitlab.mpcdf.mpg.de
- Run CI/CD jobs
 - Shared runners
 - Group/projects runners

GITLAB CONTINUOUS INTEGRATION

Automatic integration and processing of code changes





- **Managed SLURM cluster**
 - For large projects
 - *Solution provided by MPCDF Cluster team on demand when fitting*
 - Caveats:
 - *Bound to MPCDF accounts*
 - *Limited customization*
- **Examples:**
 - FHI General
 - Pirol
 - Robin / RVS
- **Alternative:**
 - SLURM (JADE) recipe
 - Build your own from scratch

Clusters in the Cloud, Wed @ 15:00





- Internal project
- Builds software for HPC systems



- Features of the HPC Cloud

- Compute

- *Images*
- *Flavors*
- *Server Groups*

- Volumes, i.e. Block storage

- Network

- *Networks*
- *Security Groups*
- *Load Balancers (not discussed)*
- *Floating IPs (not discussed)*

- Orchestration

- Object storage

- Shares, i.e. shared file systems

- Recipes for the HPC Cloud

- Kubernetes (MoM)

- SLURM (JADE)

- Workstations

- VPN (?)

- Services based on HPC Cloud

- Notebooks:

- *BinderHub & JupyterHub*

- GitLab runners

- Virtual cluster

- *Example: remote visualization service (RVS)*

- Open Build Service (OBS)



THANK YOU

