

# Bits & Bytes

No.216, August 2024

Computer Bulletin of the Max Planck Computing and Data Facility (MPCDF)\*  
<https://docs.mpcdf.mpg.de/bnb>

## High-performance Computing

### New HPC system *Viper* (Phase-1: CPU)



Figure 1: Deployment of *Viper*-CPU (on premises of the LRZ). Image credit: H. Huber (LRZ)

The deployment of the first phase of the new HPC system of the Max Planck Society, *Viper*, has been completed

by Eviden and MPCDF and the CPU-based part of the machine is operational (still in “user-risk” mode, and with some hardware repairs ongoing) since June 2024. The machine comprises 768 compute nodes, each with two AMD EPYC 9554 “Genoa” CPUs<sup>1</sup>, providing 128 Zen4 cores per node and 512 GiB (609 nodes), 768 GiB (90 nodes), 1024 GiB (66 nodes), or 2048 GiB (3 nodes) of DDR5 memory per node, respectively. The nodes are interconnected by an Nvidia/Mellanox InfiniBand (NDR 200 Gb/s) network with a non-blocking fat-tree topology, and are connected to online storage based on IBM SpectrumScale (aka GPFS) parallel file systems with a total capacity of ca. 11 PB. The new machine replaces *Cobra* which was decommissioned in July after more than 6 years of operation, and is already productively used by a large number of users. A high-level overview of *Viper* was given in a MeetMPCDF seminar<sup>2</sup> on July 4. Users can find all relevant technical details on the hardware and software and its usage, including example batch submission scripts, in a comprehensive user guide<sup>3</sup>.

Among the most notable changes, users transitioning from *Cobra* or *Raven* should be aware that compiler optimization settings<sup>4</sup> need to be adapted from Intel CPUs to the specifics of an AMD x86 CPU, and that some resource limits have been introduced on the login and interactive nodes<sup>5</sup> in order to keep the “front-end” nodes responsive for a large number of users and different (interactive) use cases.

The shipment of the second, GPU-accelerated, phase of *Viper* with more than 600 AMD MI300A APUs<sup>6</sup> (Accelerated Processing Units) has just started. This part of the machine is expected to be operational in autumn 2024.

*Markus Rampp*

\*Editors: Dr. Renate Dohmen & Dr. Markus Rampp, MPCDF

<sup>1</sup><https://www.amd.com/en/products/cpu/amd-epyc-9554>

<sup>2</sup><https://www.mpcdf.mpg.de/training/meetmpcdf>

<sup>3</sup><https://docs.mpcdf.mpg.de/doc/computing/viper-user-guide.html>

<sup>4</sup><https://docs.mpcdf.mpg.de/doc/computing/viper-user-guide.html#migration-guide-for-users-coming-from-intel-based-hpc-systems>

<sup>5</sup><https://docs.mpcdf.mpg.de/doc/computing/viper-user-guide.html#login>

<sup>6</sup><https://www.amd.com/en/products/accelerators/instinct/mi300/mi300a.html>

## HPC Software News

### MPCDF GitLab `module-image` to be discontinued on October 31

Users can run their continuous integration (CI) pipelines on the shared runners of the MPCDF GitLab instance using the same software stack as present on the HPC clusters. Last year, new CI module images<sup>7</sup> were introduced, replacing the deprecated `module-image` that has been available for a long time. The latter will now finally be discontinued and cannot be used anymore after October 31, 2024. All users still referencing the `module-image` (i.e., the `.gitlab-ci.yml` file contains the line `image: gitlab-registry.mpcdf.mpg.de/mpcdf/module-image`) need to adapt their CI pipelines to use the new images by then. Further information can be found in the documentation<sup>8</sup>.

### New module images available in MPCDF GitLab

Additional CI module images<sup>9</sup> (see also previous section) were deployed based on the `intel/2024.0` software stack. The tags `intel:latest`, `intel-mpi:latest`, and `intel-openmpi:latest` point to `intel/2024.0` now.

### Nvidia HPC SDK version 24.3 available on *Raven*

The Nvidia HPC SDK version 24.3 was installed on *Raven* and is available by loading the module `nvhpcsdk/24`. It ships its own CUDA versions 11.8 and 12.3, similar to the previous `nvhpcsdk/23`. An update of the `nvhpcsdk/24` module from the currently installed version 24.3 to a more recent minor version is in preparation.

*Tobias Melson*

## Major Change in the Python Infrastructure on the HPC Clusters

For nearly a decade, MPCDF has been providing the Anaconda Python Distribution as the foundation for Python user applications, offering a large selection of important packages such as NumPy, SciPy, matplotlib, etc. in compatible versions and using optimized builds and libraries.

However, earlier this year, Anaconda Inc. has changed its software licensing model such that MPCDF is not allowed to install *new versions* of Anaconda Python any more. Moreover, the package channels ‘default’ and ‘anaconda’ had to be disabled in the global `.condarc` config files for the existing installations, in order to prevent newly created conda environments to download from these channels which would require licensing.

From now on, MPCDF will deploy its own comprehensive Python stack entirely based on software from conda-forge<sup>10</sup>, a community-driven initiative that devel-

ops conda packages which do not fall under the strict licensing of Anaconda Inc. and can therefore be used freely. For each release, the versions of the key packages such as Python, NumPy, SciPy, matplotlib, Numba, pandas will be the same as the versions in the respective release of the commercial Anaconda distribution, whereas the versions of less important dependencies may vary for dependency-resolution reasons.

Along these lines, we start with *Water Boa Python 2024.06* (environment module `python-waterboa/2024.06`) which has been rolled out on the HPC clusters recently and might still experience minor modifications or extensions. Existing Anaconda installations, including 2023.03, stay available unchanged.

*Klaus Reuter*

## Introducing containerized Applications as Part of the MPCDF Software Stack

Software containers are ubiquitous, e.g. in cloud computing, and are also gaining popularity on HPC systems for various reasons and with certain advantages. On MPCDF systems, the Apptainer container platform is provided to support containers in user space. A typical use case is, for

example, the installation of a complex software package that expects a certain operating system version together with the respective libraries (e.g. Ubuntu 22.04), and is therefore incompatible with the host operating system of the HPC cluster (e.g. SLES 15). Once installed success-

<sup>7</sup><https://docs.mpcdf.mpg.de/bnb/214.html#announcing-legacy-status-and-later-discontinuation-of-the-module-image>

<sup>8</sup><https://docs.mpcdf.mpg.de/doc/data/gitlab/gitlabrunners.html#docker-images-for-ci-with-mpcdf-environment-modules>

<sup>9</sup><https://docs.mpcdf.mpg.de/doc/data/gitlab/gitlabrunners.html#docker-images-for-ci-with-mpcdf-environment-modules>

<sup>10</sup><https://conda-forge.org>

fully into a container image, the execution of the software is then largely portable between host Linux operating systems.

From a system administration and software deployment point of view, another advantage of containerization is that the application and its complete set of dependencies are contained within a single compressed image file (e.g. Apptainer `.sif`). This saves disk space, inodes, and installation time, compared to a regular installation of the application directly in the cluster file system. Large software packages can easily consume  $O(100,000)$  inodes, in the case of MATLAB R2023bU5 the inode count even amounts to about 680,000, for example.

To mitigate the associated negative impacts on the file

system, MPCDF is going to provide certain large software packages in containerized form in the future, starting with MATLAB R2024aU2 which is available already now as an environment module. The fact that the software is running within a container is largely hidden from the user by providing executable wrappers for the containerized executables (e.g. `matlab`, `mex`, `mcc`). These wrappers mount the cluster file systems into the container as they would appear on the host operating system.

The command `module help matlab/R2024aU2` gives further hints, advanced users may explicitly launch e.g. a containerized shell and adapt the execution command line of the containerized application, if necessary.

*Klaus Reuter*

## Nexus-S3: Object Storage in the HPC-Cloud and beyond

Nexus-S3 is a new, scalable object storage service by MPCDF, compatible with the Amazon S3 protocol. MPCDF users can opt-in to Nexus-S3 via the SelfService portal, which provides a free 1TB (1M objects) quota (see opt-in instructions below). Data can be accessed using standard S3 clients and libraries such as `minio-client`<sup>11</sup>, `s3cmd`<sup>12</sup>, `rclone`<sup>13</sup> and `python-boto3` as well as via Globus (MPCDF GO Nexus S3 Collection) or via a web browser/curl. *Note:* The `minio-client` and `rclone` are both available via the modules system on MPCDF clusters.

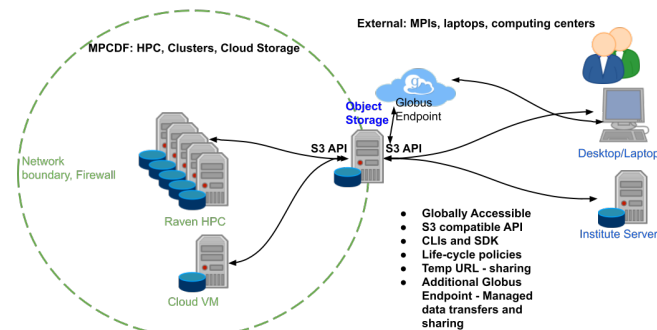


Figure 2: The Object Storage System (including Globus Endpoint)

Nexus-S3 also supports object storage functionality such as versioning, life-cycle policies and temporary URL generation to allow users to download files with an expiry date. Together with the transfer and sharing functionality available via Globus this provides many solutions to use cases such as large-scale data sharing and publishing.

<sup>11</sup><https://min.io>

<sup>12</sup><https://s3tools.org/s3cmd>

<sup>13</sup><https://rclone.org/>

<sup>14</sup><https://selfservice.mpcdf.mpg.de>

<sup>15</sup><https://docs.mpcdf.mpg.de/doc/cloud/renting/index.html>

### Example use case

An example use case would be the generation of data via computational jobs in a batch system with the following consolidation and sharing of the data via Object storage and Globus.

1. Submit batch jobs to produce data
2. Each job uses S3 cli tools, such as `minio-client`, to upload data to Object Storage
3. Data in Object Storage is shared with collaborators via Globus. Benefiting from the use of groups and sharing provided by Globus.

### Opt-in via SelfService

Access to Object Storage is possible via the MPCDF SelfService<sup>14</sup>. Log in with your MPCDF account and go to “My account / Services” to opt-in for Nexus-S3. *Note:* After opt-in, it can take up to 60 minutes until the accounts are created and for the access/secret keys to become available. Once the account has been created in the S3 service you can access your access/secret keys by clicking “View Access Token”. These access/secret keys are used by S3 clients and Globus to access your S3 storage, please keep them secure and treat them as you would do with a password. If you feel these keys may have been exposed, please create a helpdesk ticket and request that new keys be generated.

### Object storage for larger projects

For larger projects it is possible to, cost-effectively, rent object storage in the 10s-100s TiB region via the HPC-Cloud. For more information please see our documentation<sup>15</sup>

*John Kennedy, Robert Hish, Kathrin Beck, Bolarinwa Adeoye*

## The MetaStore Research Data Publication Platform

MetaStore is the catch-all data publishing platform of the MPCDF. It is meant as a place to create and publish metadata which describes already existing data stored in the various storage systems at MPCDF (e.g. Nexus-S3). MetaStore provides a landing page and a Digital Object Identifier (DOI) for the linked data, irrespective of where exactly the data is stored or how it is accessible. The data set on MetaStore will be then used as a landing page for the DOI. DOIs can also be created later in time, and not necessarily directly at the time when the data set is uploaded. Please be aware that DOI creation is irreversible.

### Managing data sets and resources in MetaStore

Data sets and resources can be managed via the Web UI or a REST-style API. MetaStore's default<sup>16</sup> metadata

schema is the DataCite metadata schema in a simple key-value form. In addition, the extended<sup>17</sup> metadata schema supports the full functionality of DataCite's metadata schema, including controlled vocabularies.

Each data set can contain one or more resources. A resource can be either an URI or an uploaded file, typically corresponding to supplementary data like papers or visualizations. Uploads larger than 1GB are not supported.

### Who can use it

MetaStore is not meant to be used by individual users, but by Max Planck Institutes. If your institute has already an account on MetaStore, you can use it with this account. For more details, please look at the user documentation<sup>18</sup>. You can contact us at support@mpcdf.mpg.de.

*Nicolas Fabas, Thomas Zastrow*

## News

### Password policy

#### MPG regulations

New password regulations have been defined in the MPG, which can be found in the OHB (XIX.04)<sup>19</sup>. According to chapter 2.1 of the password-policy document the minimal length of a password is now 12 characters (for privileged accounts 14), and it must consist of at least three different character sets out of these four categories: lower case, upper case, digits, and special characters. Beside that, a minimal complexity is required, and the use of other information linked to the account, like names, telephone numbers, birthdays or similar are not allowed to be used in passwords. Some of these requirements can be enforced by checking passwords when setting them. Furthermore, passwords can be checked against databases containing exposed passwords, like those found in HavelbeenPwnd.

#### No expiration, but checking

Following the latest NIST recommendations there is no longer the need to regularly change the password. Therefore it is wise to go for a sufficiently complex password. MPCDF will no longer enforce a yearly change of the passwords. In order to assure that the password still fulfills the MPG rules and is not found in a database of broken passwords, a yearly check of the passwords will be performed instead.

This means, that MPCDF still will set yearly expiration dates on passwords, but there is no need to change them,

if the check turns out to be OK. This check has to be done in the SelfService and if it succeeds, the expiration date will simply be shifted one year into the future. In case the check fails, either because the password does not comply with the latest MPG rules or because it is found in the database of broken passwords, a password change is enforced. This change to the previous password policy of MPCDF should allow all users to use complex passwords, which then will be valid forever. Therefore, already sufficiently complex passwords which were set up in the past, will also no longer require a change, but only a check via SelfService.

*Andreas Schott*

### 2FA for DataShare and GitLab

MPCDF will introduce two-factor authentication (2FA) for the two externally exposed services DataShare and GitLab later this month. In the beginning there will be an additional login option which enforces the 2FA with the MPCDF-2FA server. This will be available for the regular accounts as well as for guest accounts.

Later this year the standard login method for DataShare and GitLab will be switched over to this method as the only way to login. We therefore ask everybody to try out this new login method, which as a bonus provides a single-sign-on for the two services. So logging into DataShare will automatically log you in into GitLab and vice versa.

*Andreas Schott*

<sup>16</sup><https://docs.mpcdf.mpg.de/doc/data/publication/metastore/docs/datacite-standard-format.html>

<sup>17</sup><https://docs.mpcdf.mpg.de/doc/data/publication/metastore/docs/datacite-extended-format.html>

<sup>18</sup><https://docs.mpcdf.mpg.de/doc/data/publication/metastore-documentation.html>

<sup>19</sup><https://ohb.mpg.de>



## Events

### AMD GPU workshop & hackathon (November 5-7)

In order to prepare for the second phase of the new HPC system *Viper* of the MPG with a large number of AMD MI300A APUs, the MPCDF in collaboration with AMD offers an online course with hands-on for this new architecture. The workshop comprises two and a half days, starting on November 5th with an afternoon of lectures by experts from AMD, followed by two full days of expert-guided hands-on work on individual codes participants are invited to bring in. The workshop targets intermediate to advanced developers who can start out with a code that already uses GPUs (for example on the *Raven* GPU partition with Nvidia A100 GPUs), and who would like to also leverage the new *Viper* system with MI300A APUs. For registration and further details please visit the workshop registration page<sup>20</sup>.

*Markus Rampp, Tilman Dannert*

### Introduction to MPCDF services (October 24)

The next issue of our semi-annual online course “Introduction to MPCDF services” is scheduled for October 24th, 14:00-16:30 via Zoom. Topics comprise login, file systems, HPC systems, the Slurm batch system, and the MPCDF services remote visualization, Jupyter notebooks and DataShare, together with a concluding question & answer session. No registration is required, just connect at the time of the workshop via the zoom link<sup>21</sup>.

*Tilman Dannert*

### Meet MPCDF

The next editions of our monthly online-seminar series “Meet MPCDF” are scheduled for

- September 5th, 15:30 “Basic profiling of HPC applications” by Sebastian Ohlmann (MPCDF)
- October 10th, 15:30 (Topic to be announced)
- November 7th, 15:30 “The *Viper* GPU system”

All announcements and material can be found on our training webpage<sup>22</sup>.

We encourage our users to propose further topics of their interest, e.g. in the domains of high-performance computing, data management, artificial intelligence or high-performance data analytics. Please send an E-mail to

training@mpcdf.mpg.de<sup>23</sup>.

*Tilman Dannert*

### MPCDF at Garching Campus Open Doors (October 3)

For the first time MPCDF will open its doors for the general public at the *Open Day Campus Garching*<sup>24</sup> at October 3, 10:00-17:00. We will provide short talks, posters about scientific high-performance computing, data science, and artificial intelligence and offer the opportunity to have a peek into the machine hall. The program<sup>25</sup> is targeted at the general public, but we, at MPCDF, always appreciate exchange with our friends and expert users who might take the opportunity of the campus event to meet in person with MPCDF staff.

*Friederike Neu, Markus Rampp*

### HPC-Cloud workshop (September 10-12)

Following the growing popularity of the HPC-Cloud the MPCDF will host a cloud workshop which aims to establish a community of cloud users from the scientific projects. Projects will share ideas and real-world experience gained in the past months and years. The cloud experts from MPCDF will be present to facilitate the exchange and discuss future directions for the HPC-Cloud.

For more information and to register please visit Agenda and Registration at <https://plan.events.mpg.de/e/hpccloudws>.

*Raphael Ritz*

### Course on “Python for HPC” (November 26-28)

The next iteration of our popular course on “Python for High-performance Computing” is scheduled for November 26th to 28th, 2024. The event takes place online via Zoom and teaches how to use the Python ecosystem efficiently on HPC systems. We detail on topics such as using NumPy, SciPy, Cython, Numba, JAX, writing compiled extensions in C, C++, Fortran, making use of multithreading, GPU programming, and leveraging distributed memory parallelization using mpi4py and Dask. The lectures in the morning are complemented by exercises and Q&A sessions in the afternoon. Registration is now open<sup>26</sup>.

*Sebastian Kehl, Sebastian Ohlmann, Klaus Reuter*

<sup>20</sup><https://plan.events.mpg.de/e/amd-gpu-workshop-24>

<sup>21</sup><https://mpcdf-mpg-de.zoom-x.de/j/69488919156?pwd=T256S1dXSjhXcU1hNGhZNEExkeVVvQT09>

<sup>22</sup><https://www.mpcdf.mpg.de/services/training>

<sup>23</sup><mailto:training@mpcdf.mpg.de>

<sup>24</sup><https://forschungscampus-garching.de/>

<sup>25</sup><https://www.mpcdf.mpg.de/opendoors2024>

<sup>26</sup><https://plan.events.mpg.de/e/mpcdf-python-for-hpc-2024>