Cluster Computing Resource Basic User Guidelines

MRI Center, NYSPI

Juan Sanchez and David Semanek

v1 5/2022

v2 1/2023
v3 9/2023

The purpose of this document is to outline user guidelines for the MRI Center computing cluster resources, including MRImaker and Docker Swarm.

The MRI Center computing resources are provided for the purpose of processing (“pre-processing”) approved MRI and related data through pipelines managed by the MRI Center cluster computing staff and PsyIT. Regular users of the systems should never execute their own scripts or create or call Docker containers outside of the managed pipelines following the instructions of the cluster computing staff and documentation. For experienced shell users, running user generated scripts created for the purpose of gathering information about or re-organizing data within a specific project is allowed, provided the script does not change the names or attributes of folders or files that would break the pipelines enabled for your projects; such scripts are run at your own risk regarding the integrity of your data.

**Users should never use their typical Linux home folder on MRImaker (found under *~/.* or */home/[username]/.* ) as a workspace.** Please request from us a Home folder on our cloud storage in this location: */MRI\_DATA/home/[username]*. When making the request, let us know how you intend to use this space, as we prohibit users from storing imaging data or outputs in any Home folder. All image processing and pipeline testing should be done under */MRI\_DATA/mritesting*, and if you need testing space you can request access to the “mritesting” group to gain access to this location. Please store and maintain the minimum amount of data necessary for testing and be sure to remove unnecessary processing outputs as work progresses. NB: due to the permissions used under */MRI\_DATA/mritesting/*, all testing users can change and remove all data under this path, so please be mindful of which top level folder you are working in to avoid interfering with other users! It’s very helpful if all users are mindful of being efficient in using space such as not making unnecessary copies, removing stale or malformed data, and eliminating the need for DICOM file usage and handling wherever possible.

Do not under any circumstances run commands on MRImaker that query or issue instructions directly to the Docker daemon or Docker Swarm cli. Only cluster users with prior administrative approval may run commands directly through the Docker cli, including starting and stopping containers or services, or building or altering images. All other users access the Docker Swarm resources seamlessly through pipeline commands with provided documentation.

The MRI computing cluster is a work in progress, and all users have a responsibility to maintain the stability and integrity of the system by being conservative about the commands run on the system and reporting any issues or bugs immediately to the cluster computing staff. When in doubt, please check with us!

Juan.Sanchez@nyspi.columbia.edu

David.Semanek@nyspi.columbia.edu

Waseem.Ijaz@nyspi.columbia.edu