Rapporti Tecnici INAF
INAF Technical Reports

Number	390
Publication Year	2026-05-19
Acceptance in OA@INAF	2026-05-29T10:44:41Z
Title	TRG COMPUTING INFRASTRUCTURE FOR ASTROPHYSICAL DATA ANALYSIS AND SIMULATIONS: USER MANUAL
Authors	MUSCAS, Mariano
Publisher's version (DOI)	https://doi.org/10.20371/INAF/TechRep/390
Handle	http://hdl.handle.net/20.500.12386/49874

TRG COMPUTING INFRASTRUCTURE FOR ASTROPHYSICAL DATA ANALYSIS AND SIMULATIONS: USER MANUAL

Author: Mariano Muscas[1]

Referee: Alessandro Cabras[1]

[1]INAF – Osservatorio Astronomico di Cagliari

1. Introduction

The trg1 and trg2 are two workstations that have been operational since 2018, trg3 and trg4 nodes were recently purchased with the goal of improving performance compared to the existing nodes, mainly for use in the Pulsar group's data analysis. trg3 and trg4 were configured to match the configuration of trg1 and trg2: the same operating system (Ubuntu 22), the same usernames, IDs, and GIDs.
Access to **trg1, trg2, trg3, and trg4** is provided through the **feric** node.

2. Purpose and Scope of this Document

This document provides an overview and usage guide for the trg computing environment used within the Pulsar group at the Cagliari Astronomical Observatory (INAF–OAC).

The manual describes the available computing nodes, their hardware characteristics, the procedures required to access them, the storage configuration, the available software environment, and the rules for using the shared resources.

The goal of this document is to provide both new and experienced users with a practical reference for working within the trg environment.

3. What is trg?

trg is the local computing environment used by the Pulsar group for scientific data analysis, software development, and computational processing.

It consists of a small set of Linux workstations designed to support pulsar data processing pipelines, astronomical analysis software, and general scientific workloads.

The trg environment currently includes four compute nodes:

- trg1
- trg2
- trg3
- trg4

These machines provide CPU resources, GPUs, and local storage that can be used interactively by group members. Access to the nodes is managed through a frontend machine called **feric**.

The nodes differ in hardware generation and performance characteristics, allowing users to select the most appropriate system depending on the computational requirements of their tasks.

4. trg System Architecture Overview

The trg environment is composed of four compute nodes divided into two hardware generations:

Older generation nodes:

- trg1
- trg2

These systems are based on Intel Xeon CPUs and use traditional mechanical disks. They are suitable for legacy software environments and standard CPU-based workloads.

New generation nodes:

- trg3
- trg4

These nodes use modern AMD EPYC processors, DDR5 memory, NVMe storage, and NVIDIA RTX GPUs. They provide significantly higher compute performance and are intended for demanding workloads, GPU-accelerated applications, and I/O-intensive data processing.

Access to all trg nodes is performed through the **feric** frontend node using SSH.

The storage architecture differs between node generations:

- **trg1 and trg2** share several storage volumes using NFS, allowing both machines to access the same disks.
- **trg3 and trg4** use high-speed local NVMe storage configured in RAID0 to maximize I/O performance.

At present, the trg environment does **not use a job scheduler or disk quota system**, and resources are shared cooperatively between users.

5. trg Nodes Summary

The following table provides a quick overview of the main characteristics of each node.

Node	CPU	RAM (GB)	Storage	GPU	Recommended Usage
trg1	2× Intel Xeon E5-2640 v4	112	Mechanical HDD	GTX 1080 Ti	Legacy software, standard CPU workloads
trg2	2× Intel Xeon E5-2640 v4	128	Mechanical HDD	GTX 1080 Ti	Legacy software, standard CPU workloads
trg3	AMD EPYC 9224	384	NVMe RAID0	4× RTX 4000 Ada	GPU computing, fast I/O workloads
trg4	AMD EPYC 9224	384	NVMe RAID0	4× RTX 4000 Ada	GPU computing, fast I/O workloads

Users should choose the node that best matches the requirements of their analysis tasks.

6. Which Node Should I Use?

The choice of node depends on the type of workload and software environment required.

Use **trg1 or trg2** when:

- running legacy pulsar software environments
- using software that is already configured on those systems
- performing moderate CPU-based processing

Use **trg3 or trg4** when:

- performing computationally intensive analyses
- using GPU-accelerated software
- working with datasets that benefit from high-speed NVMe storage
- using software provided through the module system

Because trg3 and trg4 offer significantly higher compute performance, they are generally recommended for new analyses and heavy workloads.

7. Access Requirements

To use the trg computing environment, the following prerequisites are required:

- an active user account on the feric frontend node
- an account on the trg nodes (trg1, trg2, trg3, trg4)
- SSH access from your local machine
- access to the Observatory network, or connection through the Observatory VPN when working remotely

If you do not yet have an account on feric, contact:

helpdesk.oaca@inaf.it

The system administrator of feric will create the required accounts.

If you do not have an account on **trg1, trg2, trg3, or trg4**, please contact **Mariano Muscas** (*mariano.muscas@inaf.it*) to request the creation of your account on these systems.

8. Typical User Workflow

A typical workflow when using the trg environment consists of the following steps:

1. Connect to the frontend node:

   ssh <username>@feric.oa-cagliari.inaf.it

2. Log into the desired compute node:

   ssh <username>@trg3

3. Move to the appropriate working directory (for example a scratch or data directory).
4. Check available software modules if needed:

   module avail

5. Load the required module:

   module load <module_name>

6. Run the analysis or processing task.
7. Once the computation is completed, move important results to a safer storage location if needed.

9. Monitoring System Resources

Since trg resources are shared between users and there is currently no job scheduling system, users are encouraged to monitor resource usage before launching heavy workloads.

The following commands can be useful:

CPU usage:

```
top
```

```
htop
```

Memory usage:

```
free -h
```

Disk usage:

```
df -h
```

```
du -sh *
```

GPU usage (on trg3 and trg4):

```
nvidia-smi
```

Checking these resources before starting large jobs helps avoid conflicts with other users.

10. Basic GPU Usage (trg3 / trg4)

Nodes trg3 and trg4 are equipped with NVIDIA RTX 4000 Ada GPUs.

Before running GPU-based workloads, users should check whether GPUs are already in use:

```
nvidia-smi
```

If multiple GPUs are available, it is possible to restrict an application to a specific GPU using the environment variable:

```
CUDA_VISIBLE_DEVICES
```

For example:

```
CUDA_VISIBLE_DEVICES=0 python script.py
```

Users are encouraged to coordinate with others when running long GPU-intensive jobs.

11. Installing Additional Software

Users may sometimes need software that is not already installed on the system.

Possible approaches include:

- installing software in the user's home directory
- using Python virtual environments
- using Docker containers where appropriate

For software that needs to be installed system-wide or provided as a module on the trg workstations, users should contact the system maintainer (Mariano Muscas).

If you have any difficulties installing astronomical software on the trg nodes, please contact Mariano Muscas (mariano.muscas@inaf.it).

12. Known Limitations of the trg Environment

Users should be aware of several current limitations:

- There is **no job scheduler** (such as SLURM).
- There are **no enforced disk quotas**.
- Storage on **trg3 and trg4 uses RAID0**, which provides no redundancy.
- The NFS storage shared between **trg1 and trg2 depends on the availability of the node that physically hosts the disks**.

Users are therefore expected to use the system responsibly and cooperate with other users.

13. Hardware Features

The trg computing environment is composed of two generations of hardware with significantly difference performance characteristics.

The older nodes, **trg1** and **trg2**, are based on Intel Xeon **E5-2640 v4** CPUs (2016 generation) and use traditional mechanical storage. These systems are suitable for standard workloads and legacy software environments.

By contrast, the newer nodes, **trg3** and **trg4**, are equipped with **AMD EPYC 9224** processors (2022 generation), high-speed **NVMe** storage, **DDR5** memory, and **NVIDIA RTX 4000 Ada** GPUs. These nodes deliver substantially higher computational performance, memory bandwidth, and I/O throughput, and are therefore recommended for demanding workloads, GPU-accelerated applications, and data-intensive analyses.

Users are encouraged to select the appropriate node type based on the computational requirements of their tasks to achieve optimal performance.

13.1. trg1 / trg2

CPU

- **Model:** 2 x Intel® Xeon® E5-2640 v4 per node @ 2.40 GHz 3.4 GHz boost clock, 10 cores / 20 threads per CPU; 20 cores / 40 threads total per node
- **Threads per core:** 2
- **Total cores: 20** (40 threads).

Disk

- **Mechanical Hard Drives:** trg1: (17 TB × 7) = 119 TB trg2: (17 TB × 7) = 119 TB. Six out of seven disks on each trg1 and trg2 node are shared with each other via NFS.

RAM

- **trg1: Total:** 112 GiB (7 × 16 GiB) **trg2: Total:** 128 GB (8 x 16 GiB)
- **Type:** DDR4
- **Speed:** 2400 MT/s
- **Manufacturer:** Hynix Semiconductor

GPU - Each node (trg1 and trg2) is equipped with:

- **ASPEED Graphics Family** (on-board)
- **NVIDIA GeForce GT 1030:** 2 GB GDDR5
- **NVIDIA GeForce GTX 1080 Ti:** 11 GB GDDR5X - 3584 CUDA cores

Chassis / Node

- Supermicro

13.2. trg3 / trg4

CPU

- **Model:** AMD EPYC 9224
- **Cache:** 64 MB
- **Cores:** 24
- **Threads:** 48
- **TDP:** 200 W

Disk

- **System Disk:** 2 x 480 GB SATA SSD (Software RAID for the operating system)
  - Form factor: 2.5"
  - Interface: SATA 600 (6 Gb/s)
  - Transfer speed: 600 MB/s
- **NVMe Storage:** 3 × 15.3 TB
  - Model: Kioxia CD8-R RI 15.3 TB
  - Form factor: U.2 PCIe
  - DWPD: 1

RAM

- **Type:** DDR5-4800
- **Total:** 24 × 16 GB modules = Total: 384 GB

GPU

- **Each node has 4 × NVIDIA RTX 4000 Ada Generation GPUs. The following specifications are per GPU:**
  - Memory: 20 GB GDDR6 (ECC)
  - Memory bandwidth: up to 360 GB/s
  - FP32 performance: 26.7 TFLOPS
  - RT core performance: 61.8 TFLOPS
  - Tensor performance: 427.6 TFLOPS
  - Max power consumption: 130 W
  - Bus: PCIe Gen4 ×16
  - CUDA cores: 6144
  - Tensor cores: 192
  - RT cores: 48

  Total GPU memory per node: 4 × 20 GB = 80 GB

13.3. trg1 / trg2 / trg3 / trg4

OPERATING SYSTEM: Ubuntu 22.04.5 LTS
Kernel: 6.8.0-106-generic

14. SSH access to the trg nodes

Compared to the configuration in place until December 2025, **trg1, trg2, trg3, and trg4 now use feric as their frontend node.**
Access to feric from a local PC can be performed as follows:

[mylaptop_username@mylaptop ~]$ ssh <feric_username>@feric.oa-cagliari.inaf.it

Once you are logged into **feric**, you can then access:

<feric_username>@feric:~$ ssh <trg1_username>@trg1
or:
<feric_username>@feric:~$ ssh <trg2_username>@trg2

As you may have noticed, *trg2* is no longer under *trg1*. This means that if *trg1* goes down, you can still access *trg2* from **feric**.
All of this applies when you are **inside the Observatory network**. If you are **outside the Observatory**, you need to activate the VPN provided by: *helpdesk.oaca@inaf.it*

If you do not yet have an account on feric, you need to ask helpdesk.oaca@inaf.it to create one for you (contact the system administrator of **feric** to perform this operation, as administrative privileges are required). Once you have an account, you will be able to access the *trg* nodes from feric.

Finally, if you are on the Observatory network but still cannot connect directly to **feric**, and you want to enable direct access from your laptop, you must add the following line to your local /etc/hosts file:
192.168.140.55 *feric.oa-cagliari.inaf.it*
Remember that you need to open a **new shell** after saving the file for the changes to take effect.
The trg machines are no longer part of the "research" network, and therefore they no longer use the fully qualified domain name *trg1.oa-cagliari.inaf.it*. Because of this change, SSH connections that include the full domain name will fail.
To connect to trg1, you must use the short hostname only, without any domain extension.
ssh <trg1_username>@trg1 (without domain extension)

15. Disk configuration of trg3 and trg4

Nodes trg3 and trg4 are configured with a RAID0 array on the NVMe drives.

This configuration was chosen to provide maximum performance, offering very high read and write speeds for demanding computational workloads. Please note that RAID0 does not provide any data redundancy: if one of the drives fails, all data stored on the node may be lost.
RAID0 was chosen to maximize capacity and I/O performance. Redundant RAID levels were not selected because this storage is intended as high-performance scratch/local working space rather than long-term reliable storage. For this reason, each user is responsible for maintaining their own data backups, and it is strongly recommended to store important files on external or redundant storage systems after computations are completed.

User best practices

- Use trg3 and trg4 for workloads that benefit from high-speed local storage (e.g., large dataset processing, I/O-intensive tasks).
- Do not use these nodes as permanent long-term storage.
- Move important results to a safer or redundant storage area after processing.
- Remove unnecessary files to keep workspace available.

16. Environment Modules

Environment Modules provide an easy and flexible way to use pre-installed software by automatically setting the appropriate environment variables. They are the recommended solution for managing multiple compilers, libraries, or different versions of the same application that users may need.

Modules are particularly useful when software packages install binaries with common names that could otherwise conflict with system executables or clutter directories such as /usr/bin.

To view the list of modules available on the trg3 and trg4, use:

```
$ module avail
```

To dynamically load a module and update your shell environment variables:

```
$ module load gcc11.4/fftw/3.3.10
```

To display the list of modules you currently have loaded:

```
$ module list
Currently Loaded Modulefiles:
  1) cuda/11.8   2) gcc11.4/fftw/3.3.10
```

17. How to Use TigerVNC on trg1, trg2, trg3, trg4

To access the graphical desktop on the trg nodes use TigerVNC from your laptop. TigerVNC is installed and configured on all nodes: trg1, trg2, trg3, trg4.

1. Start your VNC session on the target node
   First, connect to the node where you want to work (for example trg3), and start your VNC server:
   vncserver :1
   This creates your remote desktop on display :1, which corresponds to port 5901.
   (You only need to set the VNC password the very first time.)

2. From your laptop, create the SSH tunnel
   Before opening TigerVNC Viewer, you must create a secure SSH tunnel from your laptop to the target node.
   For trg3
   mylaptop_username@mylaptop:~$ ssh -J <feric_username>@feric.oa-cagliari.inaf.it -L 5901:localhost:5901 <trg3_username>@trg3

3. Open TigerVNC Viewer
   Now start **TigerVNC Viewer** on your laptop.
   In the "VNC Server" field, if you use port 590, 1enter:
   localhost:5901
   Click **Connect**, enter your VNC password, and your remote desktop will appear.

4. Closing the session
   When you are done, close TigerVNC Viewer and then close the SSH terminal.
   If you want to stop your VNC session:
   vncserver -kill :1

18. Storage configuration between trg1 and trg2.

Both machines are equipped with their own local hard disks, but they are also configured so that each node can directly access the disks physically attached to the other node. In practice, users working on trg1 can see not only the trg1 local disks but also all the storage devices that belong to trg2, and vice-versa. This makes the two systems behave as if they were a small two-node storage cluster, with all data accessible from both sides.

Below the shared storage volumes and their sizes are listed:

Disks physically attached to trg1 (also visible from trg2):

- `/data1bck` – 16 TB
- `/backup` – 17 TB
- `/data2` – 17 TB
- `/data3` – 17 TB

- `/scratch1` – 17 TB
- `/scratch2` – 17 TB
- `/scratch3` – 17 TB

Disks physically attached to trg2 (also visible from trg1):

- `/data4bck` – 16 TB
- `/backup` – 17 TB
- `/data5` – 17 TB
- `/data6` – 17 TB
- `/scratch4` – 17 TB
- `/scratch5` – 17 TB
- `/scratch6` – 17 TB

This configuration allows users on either node to access the complete set of storage resources transparently, regardless of whether the data is physically hosted on trg1 or trg2.

In trg1 and trg2 all pulsar software now directly available at login are:

TEMPO,
python3-tk,
python-3.6,
python-3.7,
nmon,
pandas,
gVNC,
PSRCAT version 1.61,
SEADAS-SETUP-EDITOR version 0.4,
stimela Python library,
Docker,
tmux,
BOINC client,
h5py Python library,
QT4 libraries,
“PulsarSearch” pipeline by Vishal Gajjar,
SPLICE_DADA,
PFITS (Andrew Jameson’s fork),
COAST_GUARD,
chromium-browser,
vnc4server,
xfce4,
SAOImage DS9,
fv,
PSRCAT version 1.60,
PSRFITS_UTILS (Scott Ransom’s fork),
PYPULSAR python libraries.

19. Rules and sharing of resources

At present, there are no disk quotas or job queues. You are free to use all the resources (disk space and processing power) that you need at any time.

However, please keep an eye on how much disk space you are using and whether others may also need access to processing cores (the command top can be helpful).

Please also follow these rules:

- **Do not use your home directory to store or process data.** Your home directory is intended for personal documents, software installations, and configuration files. The partition hosting home directories is very small and also contains the operating system and other users' home directories. Filling it up can cause serious problems for the entire trg1 system.

- **Store your data only in the partitions** /data1bck, /data2, /data3, and their subdirectories. These partitions are meant for data storage and should not be used for running scientific analyses.

- **Perform all processing and analyses in the partitions** /scratch1, /scratch2, /scratch3. Avoid storing data permanently in these partitions. If you need the data in the same directory where you process them, consider using symbolic links instead. This will also improve performance, since one disk will be used for reading the data and another for writing the results, while helping keep the system organized and clean.

20. Software in trg3 and trg4

On **trg3** and **trg4**, however, the astronomical software is not loaded by default. It must be loaded when needed using modules. This allows each user to choose among multiple versions of the same software.

The modules available on **trg3** and **trg4** are:

cuda/11.8
gcc/7.5.0
gcc11.4/epsic
gcc11.4/presto/5.0.2
gcc11.4/psrdada/1.1
gcc11.4/tempo/13.103
cuda/12.2
gcc11.4/cfitsio/4.6.2
gcc11.4/fftw/3.3.10

gcc11.4/presto2_on_gpu/2
gcc11.4/python/2.7.18
gcc11.4/tempo2/25-01-02
cuda/13.0.2
gcc11.4/dedisp/1.1
gcc11.4/heimdall/3
gcc11.4/psrcat/1.54
gcc11.4/sigproc/2025-03-29
isl/0.18
gcc11.4/dspsr/2025-07-16
gcc11.4/pgplot/5.2
gcc11.4/psrchive/2025-08-21
gcc11.4/sigpyproc/2