

Instructor Guide

The main objective of this module is to introduce students to remote execution of programs in a cluster. As a starting point, the jobs will be executed on TACC Stampede2. In addition, to familiarize the students with the different types of nodes present in a supercomputer, namely login and compute nodes; different types of file systems. A few points the instructor should try to emphasize:

1. Remind students that there are several options to transfer code from/to the supercomputer. Using GLOBUS is generally a good option for new users.
2. Familiarize students with the authentication systems present in highly secure environments. Remind students to not share their credentials and to be responsible where they store the OTP devices.
3. The instructor is encouraged to stress the importance of scratch space but also make sure that students are aware of the purging policy.
4. The instructor is also encouraged to make use of the workspace for staging.
5. The instructor is encouraged to motivate the students to start using revision control systems or at least start considering these systems for the synchronous development of their code.
6. The instructor is encouraged to spend some time explaining NUMA control and CPU-affinity control at the submission job level.

Code package:

1. Before the guided exercise make sure that all scripts are functional in Stampede2. Although the developers have made every effort to make sure that the scripts work in the system as of 07/24/2020; several system updates may compromise the ability of the scripts to work.
2. There is a readme file in the tarball as well as Makefiles and submission scripts ready to use.

Common pitfalls for students and instructors

1. Executing jobs in the login nodes versus the compute nodes.
2. It is a common mistake to allocate all MPI ranks to a single node.
3. Misuse of the NUMA commands resulting in all jobs running on a single socket.
4. Misunderstanding of the rank, task and job concepts.