OpenMPI

The Open MPI Project is an open source MPI-2 implementation that is developed and maintained by a consortium of academic, research, and industry partners. Open MPI is therefore able to combine the expertise, technologies, and resources from all across the High Performance Computing community in order to build the best MPI library available. Open MPI offers advantages for system and software vendors, application developers and computer science researchers.

- OpenMPI installations on LRZ systems
- Specific OpenMPI usage scenarios
  - Running OpenMPI programs on SuperMUC
  - Running OpenMPI programs on the Linux Cluster
- Documentation

OpenMPI installations on LRZ systems

On LRZ HPC Systems, OpenMPI is used for research and experimental purposes. This means that while LRZ support staff will do its best to help you with problems, we do not have any commercial-level support and hence cannot give any functional or reliability guarantees for using this software. Also, programs built with OpenMPI will often not scale as well as the proprietary MPI implementations on the high-end systems.

The following table gives an overview over the available OpenMPI installations on the LRZ HPC systems:

<table>
<thead>
<tr>
<th>Platform</th>
<th>Environment module</th>
<th>Supported compiler module</th>
</tr>
</thead>
<tbody>
<tr>
<td>SuperMUC, CooLMUC2</td>
<td>mpi.ompi/1.8/intel</td>
<td>intel/15.0</td>
</tr>
<tr>
<td></td>
<td>mpi.ompi/1.10/intel</td>
<td>intel/16.0</td>
</tr>
<tr>
<td></td>
<td>mpi.ompi/2.0/intel</td>
<td>intel/16.0 and intel/17.0</td>
</tr>
<tr>
<td>SuperMUC, CooLMUC2</td>
<td>mpi.ompi/1.8/pgi</td>
<td>pg/14</td>
</tr>
<tr>
<td></td>
<td>mpi.ompi/1.10/pgi</td>
<td>pg/15</td>
</tr>
<tr>
<td></td>
<td>mpi.ompi/2.0/pgi</td>
<td>pg/17</td>
</tr>
<tr>
<td>SuperMUC, CooLMUC2</td>
<td>mpi.ompi/1.8/gcc</td>
<td>gcc/4.7</td>
</tr>
<tr>
<td></td>
<td>mpi.ompi/1.10/gcc</td>
<td>gcc/4.9</td>
</tr>
<tr>
<td></td>
<td>mpi.ompi/2.0/gcc</td>
<td>gcc/4.9 and gcc/7</td>
</tr>
<tr>
<td>SuperMUC, CooLMUC2</td>
<td>mpi.ompi/1.10/nag</td>
<td>nag/6.1</td>
</tr>
<tr>
<td></td>
<td>mpi.ompi/2.0/nag</td>
<td>nag/6.1</td>
</tr>
<tr>
<td>CooLMUC3</td>
<td>mpi.ompi/2.1/intel</td>
<td>intel/17.0</td>
</tr>
<tr>
<td></td>
<td>mpi.ompi/3.0/intel</td>
<td>intel/17.0</td>
</tr>
<tr>
<td>CooLMUC3</td>
<td>mpi.ompi/2.1/gcc</td>
<td>gcc/6.3 and gcc/7</td>
</tr>
</tbody>
</table>

In order to access the OpenMPI installation, the appropriate environment module must be loaded after unloading the default MPI environment, and after loading a suitable compiler. For example, the command sequence

```
module unload mpi.ibm mpi.intel
module load mpi.ompi/2.0/intel
```

can be used, to select one of the OpenMPI installations from the table above. To compile and link the program, the mpicc / mpiCC / mpifort commands are available.

Usage of OpenMPI is typical for the LRZ MPI installations with respect to compiler wrappers and startup commands; special variants are described below.

The OpenMPI builds are all done with MPI_THREAD_MULTIPLE enabled.

Specific OpenMPI usage scenarios

Running OpenMPI programs on SuperMUC
Please consult the LoadLeveler page for example job scripts; starting out from a job script for Intel MPI (i.e. job_type=MPICH), you only need to replace the command "module load mpi.intel" by "module load mpi.ompi/1.6/intel" (assuming the Intel compiler flavor should be used).

For starting multi-node jobs on SuperMUC (or the C2PAP cluster), please also specify additional variables as follows:

```bash
mpiexec ... -x MXM_OOB_FIRST_SL=0 -x MXM_LOG_LEVEL=error ./my_mpi_prog.exe
```

The first one avoids a hang on the slave nodes and is necessary. The second one is optional and suppresses an error message due to a missing software component. For execution of hybrid programs, use e.g.

```bash
mpiexec ... -x MXM_OOB_FIRST_SL=0 -x MXM_LOG_LEVEL=error -x OMP_NUM_THREADS=4 -x OMP_PROC_BIND=true ./my_mpi_prog.exe
```

to assure that the OpenMP settings are propagated to all tasks.

### Running OpenMPI programs on the Linux Cluster

Please consult the parallel SLURM example job script page; starting out from a job script for Intel MPI, it should be sufficient to make the following changes:

1. Add the lines

   ```bash
   module unload mpi.intel
   module load mpi.ompi/2.0/intel
   ```

   immediately after the "source /etc/profile.d/modules.sh"

2. Start the program with the `mpirun` or `mpiexec` command analogous to the command line specified for SuperMUC above.

### Documentation

Documentation as well as Frequently asked Questions pages are available on the OpenMPI web site.

Also, man pages are available for the various OpenMPI commands, especially `mpi(3)` and `mpirun(1)`.