Resource limits for serial jobs on Linux Cluster

This subdocument contains a description of constraints under which serial jobs execute on the cluster systems: maximum run times, maximum memory and other SLURM-imposed parameters.

- Resource Limits
  - Resource Limits on housed clusters
- Policies for interactive jobs
- Policies for queued batch jobs
  - General restrictions
  - Jobs in Hold
  - Job Submissions
  - Memory use

Resource Limits

The following is an overview of the resource limits imposed for various classes of jobs. These are comprised of run time limits, limits on core counts for shared memory jobs, and memory limits. Please consult the SLURM specifications subdocument for more details, in particular how to correctly specify memory requirements.

- The designation "shared memory" assumes that a number of cores assigned by SLURM will be used by threads; typically a command like export OMP_NUM_THREADS=<number> should be issued to achieve this.
- If a job appears to not use resources properly, it may get deleted at LRZ staff's or surveillance system's discretion.

<table>
<thead>
<tr>
<th>Cluster / Partition</th>
<th>Architecture</th>
<th>Core counts and remarks</th>
<th>Run time limit (hours)</th>
<th>Memory limit (GByte)</th>
</tr>
</thead>
<tbody>
<tr>
<td>--clusters=serial -- partition=serial_mpp2</td>
<td>28-way Haswell-EP node</td>
<td>1 core (effectively more if large memory is specified)</td>
<td>96</td>
<td>2.2 GByte (for 1 core)</td>
</tr>
<tr>
<td>--clusters=inter -- partition=terame_inter</td>
<td>192-way HP DL580 shared memory node</td>
<td>up to 96 logical cores can be specified. Generally, a memory specification should be provided as well using the --mem submission option.</td>
<td>96 (default 8)</td>
<td>6000 GByte</td>
</tr>
<tr>
<td>--clusters=serial -- partition=serial_large_mem</td>
<td>8-way AMD Opteron node</td>
<td>1 core (effectively more if large memory is specified)</td>
<td>240</td>
<td>4 GByte (for 1 core)</td>
</tr>
<tr>
<td>--clusters=serial -- partition=serial_long</td>
<td>4-way AMD Opteron node</td>
<td>1 core (effectively more if large memory is specified)</td>
<td>480</td>
<td>2 GByte (for 1 core)</td>
</tr>
<tr>
<td>--clusters=serial -- partition=serial_shm4</td>
<td>4-way AMD Opteron node</td>
<td>Up to 4 cores</td>
<td>120</td>
<td>8 GByte (complete node)</td>
</tr>
</tbody>
</table>

Resource Limits on housed clusters

The clusters and partitions described in this section are only available to institutes that have a housing contract with LRZ.

<table>
<thead>
<tr>
<th>Cluster / Partition</th>
<th>Architecture</th>
<th>Core counts and remarks</th>
<th>Run time limit (hours)</th>
<th>Memory limit (GByte)</th>
</tr>
</thead>
<tbody>
<tr>
<td>--clusters=lmu_asc -- partition=lmu_asc_shm16</td>
<td>16-way AMD Opteron node</td>
<td>Up to 16 cores. Access is limited to users from the Arnold Sommerfeld Center.</td>
<td>240</td>
<td>128 GByte (complete node)</td>
</tr>
</tbody>
</table>
Policies for interactive jobs

- serial interactive program runs are not started via SLURM. Such runs should be kept short; anything running longer than 30 minutes should be submitted as a scripted batch job
- Usage of the cron system (e.g. via /usr/bin/crontab) as well as the /usr/bin/at or /usr/bin/batch commands is not allowed. You are required to use the documented SLURM commands.
- Submission of serial jobs is supported on the login nodes lxlogin5, lxlogin6 and lxlogin7

Policies for queued batch jobs

General restrictions

- The job name should not exceed 10 characters. If no job name is specified, please do not use excessively long script names.
- Do not use the xargs command to generate command line arguments at submission time. Instead, generate any necessary arguments inside your script.

Jobs in Hold

Jobs in user hold will be removed at the LRZ administrators’ discretion if older than 8 weeks.

Job Submissions

- There exists a maximum number of jobs that can be submitted by a user into a serial queue. This limit may change over time, dependent on the cluster load.
- Submission of large numbers of jobs (>100, including array jobs) with very short run time (< 1min) is considered a misuse of resources. It causes both waste of computational resources and - if mail notifications are used - disruption of the notification system. Users that submit such jobs will be banned from further use of the batch system. Bundle the individual jobs into a much bigger one!

Memory use

Jobs exceeding the physical memory available on the selected node(s) will be removed, either by SLURM itself, or the OOM ("out of memory") killer in the operating system kernel, or at LRZ's discretion since such a usage typically has a negative impact on system stability.