Table of Contents

System usage

The DEEP-EST Extreme Scale Booster (ESB) can be used through the SLURM based batch system that is also used for (most of) the Software Development Vehicles (SDV). You can request ESB cluster nodes (`dp-esb[01-25]` with an interactive session like this:

```
srun -A deep --partition=dp-esb -N 4 -n 2 --pty /bin/bash -i
srun ./hello_cluster
Hello World from rank 3 of 8 on dp-esb02
Hello World from rank 2 of 8 on dp-esb02
Hello World from rank 7 of 8 on dp-esb04
Hello World from rank 6 of 8 on dp-esb04
Hello World from rank 4 of 8 on dp-esb03
Hello World from rank 0 of 8 on dp-esb01
Hello World from rank 5 of 8 on dp-esb03
Hello World from rank 1 of 8 on dp-esb01
```

Attention: The remaining two ESB racks with nodes `dp-esb[26-75` are planned to be installed in April 2020.

When using a batch script, you have to adapt the partition option within your script: `--partition=dp-esb`

Filesystems and local storage

The home filesystem on the DEEP-EST ESB Module is provided via GPFS/NFS and hence the same as on (most of) the remaining compute nodes. The local storage system of the ESB running BeeGFS is available at

```
/work
```

There is a gateway being used to bridge between the Infiniband EDR used for the CM and the 40 GbE network the file servers are connected to.

This is NOT the same storage being used on the DEEP-ER SDV system. Both, the DEEP-EST prototype system and the DEEP-ER SDV have their own local storage.

It's possible to access the local storage of the DEEP-ER SDV (`/sdv-work`), but you have to keep in mind that the file servers of that storage can just be accessed through 1 GbE ! Hence, it should not be used for performance relevant applications since it is much slower than the DEEP-EST local storages mounted to `/work`.

There is also some node local storage available for the DEEP-EST Cluster nodes mounted to `/scratch` on each node (about 380 GB with XFS). Remember that this scratch is not persistent and **will be cleaned after your job has finished** !

Multi-node Jobs

The latest `pscom` version used in ParaStation MPI provides support for the Infiniband interconnect currently used in the DEEP-EST ESB and Cluster Module. Hence, loading the most recent !ParaStationMPI module will be enough to run multi-node MPI jobs over Infiniband within the ESB as well as between ESB and CM nodes. Currently (as of 2020-04-03) the ESB rack is equipped with IB and directly connected to the CM nodes. Hence, no gateways has to be used when running on CM and ESB nodes. This will change once the targeted Extoll Fabri³ solution is implemented for the ESB.

```
module load ParaStationMPI
```

For using ESB nodes in heterogeneous jobs together with DAM nodes, please see info about [heterogeneous jobs](#).

Using GPU direct

For using GPU direct over IB on the ESB when passing GPU Buffers directly to MPI send and receive functions the following modules, currently, have to be loaded (it's planned to later integrate those modules in the production stage of Easybuild as well so that only loading the `Intel` and `ParaStationMPI` modules will be required):

```
module --force purge
module use $OTHERSTAGES
module load Stages/Devel-2019a
module load Intel
```

```
module load ParaStationMPI
```

In addition the following environment settings are required for your job scripts or interactive sessions:

```
PSP_CUDA=1
PSP_UCP=1
```

```
module load ParaStationMPI
```

In addition the following environment settings are required for your job scripts or interactive sessions: