

Why do I seem to be running out of memory with MPI jobs?

Users sometimes find that their MPI job crashes due to the way the Altix systems handle memory.

The default mode is to enhance the performance of MPI jobs running on the shared memory Altix system by mapping their memory across all CPU sets involved in the computation. The downside of this improved performance is that you will need a lot more RAM than you perhaps realise, unless you consciously ask for less.

For example, to run a parallel application like Mr Bayes, you would find that a job like this would not run on gust because it is actually asking for more RAM than is available (32 GB)

```
mpirun -np 10 /HPC/apps/MrBayes-3.1.2/mb-mpi WT_Combo.nex
```

but this one would run and would require a total of 30.6GB

```
mpirun -np 8 /HPC/apps/MrBayes-3.1.2/mb-mpi WT_Combo.nex
```

The total (VIRT) memory required for the job will be
 $\text{MPI_MAPPED_HEAP_SIZE} * \text{NPROCS} * \text{NPROCS}$

By setting the environment variables up as

```
export MPI_MAPPED_HEAP_SIZE=500000000  
export MPI_MAPPED_STACK_SIZE=1000000  
mpirun -np 8 /HPC/apps/MrBayes-3.1.2/mb-mpi WT_Combo.nex
```

would run within 4GB.

You will need to experiment with HEAP and STACK size to suit your job. Also remember that the more memory you ask for the longer you may have to wait in the queue.