

Moab

How can I run two applications in one job request in the Cray environment with one on Haswell compute nodes and the other on KNL nodes?

Question:

How can I run two applications in one job request in the Cray environment with one on Haswell compute nodes and the other on KNL nodes?

Solution:

With the introduction of NUMA this is made possible. The -L multi-req job creates "chucks" or "instruction sets" that then create an ALPS reservation and are in turn used by aprun as -B. Each -B runs in a reservation chunk or "instruction set".

Example:

```
aprun -B ./MPIcheck >>$outputFile -B /MPIcheck2 >>$outputFile
```

Here is some information for users desiring to submit jobs on a heterogeneous Cray system with Xeon- and Xeon Phi-based compute nodes.

1. For the purpose of scheduling multi-req jobs that request Haswell and KNL nodes, administrators must assign feature names to the Intel Xeon (Haswell) and Intel Xeon Phi (KNL) nodes; e.g., "haswell" for Xeon nodes and "knl" for Xeon Phi nodes. This permits users to identify the type of nodes (Xeon or Xeon Phi) on which they want to run their job.
2. For those users running concurrent applications in a single job that require heterogeneous resources (e.g., a Xeon Phi-based simulation application + a Xeon-based in-situ visualization application), users must ask for Xeon Phi nodes first using "**-L tasks=nnn:feature=knl:opsys=<OSname>:...**" for the simulation application and for Xeon nodes second using "**-L tasks=nn:feature=haswell:...**" for the in-situ visualization application.
3. TORQUE creates a single ALPS reservation containing two separate resource allocations; the first for the Xeon Phi nodes and the second for the Xeon nodes.
4. Assuming the simulation and visualization applications are separate program files, the job script must execute the applications using the ALPS command
aprun -B <simPath> : -B <vizPath>
to identify the separate program files, which uses the feature Cray calls MPMD (Multiple Program / Multiple Data).
5. The **-B** options tells the **aprun** command to use the resource allocations defined in the ALPS reservation instead of requiring the user to specify node quantity, tasks per node, etc, on the **aprun** command's command line. For a multi-req job, the **-B** parameter indicate "use a defined set of resources"

Page 1 / 2

Moab

within the ALPS reservation, which resource "sets" the **-L** parameters define.

Use Case:

Site that have both Haswell and KNL compute nodes have discovered that visulation applications and statistic applications runn better on different systems. By specifing two applications, one to run on the Haswell systems and one set to run on the KNL systems, job complete more quickly and make better use of the network and disk.

Unique solution ID: #1136

Author: Jason Booth

Last update: 2016-08-22 22:58