

PROJECT INTRODUCTION

Objectives

To execute MPI-based programs over the National Grid Pilot Platform (NGPP).

Project Investigator / Manager

Jason Oh
Hewlett Packard Singapore
Jason.oh@hp.com

Website

<http://www.ngpp.ngp.org.sg/operations/initiatives.html>

Abstract

MPI-based programs, which have been ported to MPICH-G2, have difficulty to run across two or more clusters whose compute nodes have their respective private IP address spaces. This difficulty has been addressed by solutions that use Realm Specific IP (RSIP) framework and protocol proposed and deployed by Hewlett-Packard.

PROJECT DETAILS

Description

For programs that are based on the MPI (Message Passing Interface) standard, which make use of a common interface for message-based communication and synchronization among its processes running on distributed-memory parallel computers and networks of workstations, the use of a private IP address space does not pose a problem. Many such programs are developed using an open source implementation of the MPI standard known as MPICH.

In order for an MPICH program to run over a grid, it must first be ported to MPICH-G2. The latter is an MPICH implementation using Globus Toolkit 2 (or GT2) as its underlying communication device.

A characteristic of MPICH-G2 is its requirement for all nodes involved in running an MPI program to establish communication with all other nodes. This, in turn, means that all the compute nodes of a Beowulf cluster should be in the global or public IP address space (relative to the IP address of the machine from which the program is launched).

Moving the compute nodes to a public IP address space is often not possible for a variety of reasons, which include the following:

- Unavailability or limited number of public IP addresses;
- Reluctance of compute resource owners to expose the compute nodes; and

- Efforts required to re-configure the compute resources.

Without putting the compute nodes on a public IP address space, an MPICH-G2 program is unable to harness the other compute resources on NGPP beyond a single cluster.

This dilemma was first encountered when the National Grid Competency Centre was assessing a request to grid-enable an MPI-based program to run beyond the Hydra-3, a 16-node quad Itanium cluster at the Singapore - MIT Alliance (SMA).

The initial implementation involved two clusters of Itanium located at the NGPP Hub and the Bioinformatics Institute, totaling 18 CPUs. These two clusters reside at different geographical locations on the NGPP and are connected by a 1 Gbps link.

With RSIP, it is possible to execute MPICH-G2 programs across clusters even when the nodes do not have public IP addresses. Since then, other Itanium clusters on NGPP also support RSIP

Collaborating Organizations:

- HP Labs (India)
- National Grid Office