Notre Dame Center for Research Computing (CRC)
Cluster Partnership Program

This document defines the CRC's Cluster Partnership Program (CPP) for ND Faculty regarding system procurement, administration, and support. Under this program the CRC provides specification/acquisition expertise, space (in an appropriate and secure operating environment), all necessary infrastructure (e.g., power, cooling, networking, rack space, switches, power strips, and cabling), system administration, and server/hardware support (e.g., error diagnosis and warranty work).

Given the CRC's provision of infrastructure and services, our partner researchers are able to leverage their capital resources to purchase approximately 50% more compute nodes while significantly reducing the overhead of internal system administration. In return for infrastructure and services provided by the CRC, when partner compute resources are not being used by the partner they are automatically available to other "trusted" partner program users within the University's computational community.

Partner Program Details

Partner program hardware options are compatible with the CRC's general access hardware resources. Compatible hardware allows a streamlined systems staff to effectively support a large number of systems—since the technical efforts required to manage the compute node clusters increases very little with additional, compatible hardware.

Three compatible high-performance computing (HPC) environments will be available for partners in fall 2009. These are based on the core, general access HPC environments also available in fall 2009, and include:

- Three hundred and sixty nodes with AMD Istanbul processors (a total of 4,320 cores). This is a distributed memory Linux cluster environment for high throughput computing (HTC) based on Sun, HP, or Dell hardware with gigabit Ethernet interconnections.
- Around six hundred Intel Nehalem processor cores with high speed interconnect (ten gigabit Ethernet or InfiniBand). This is a distributed memory Linux cluster for highly parallel computing.
- A multi-processor, shared memory system—each processor with multiple cores and up to 256 gigabytes of RAM.

For the distributed memory, high throughput Linux cluster, partners can choose among specific "tier-1" vendor compute nodes. These devices have two six-core AMD processors and are typically configured with one gigabyte of memory per processing core and 100 gigabytes (or more) of local disk storage. The partner cost for a Linux cluster compute node is the actual cost of a "1U" (rack-mountable) server with three years of maintenance included. The CRC provides rack space and all other necessary infrastructure for a gigabit Ethernet connection. Partners who are able to bring at least

two 1U servers are welcome to participate.

The distributed memory, highly parallel computing Linux cluster partners can choose between IBM and Dell compute nodes. These devices have two four-core Intel Nehalem processors and are typically configured with two gigabytes of memory per processing core and 100 gigabytes (or more) of local disk storage. Here, the partner cost for a Linux cluster compute node is the actual cost of the 1U server with three years of maintenance included. The CRC provides rack space and all other necessary infrastructure for an InfiniBand or ten gigabit Ethernet connection. Partners who are able to bring at least one rack of (32) 1U servers are welcome to participate.

Shared Memory System Partners can add additional nodes to the CRC shared memory system. These are eight-processor nodes typically configured with 64 to 256 gigabytes of memory, two 500 gigabyte SCSI disks, and fiber optic channel ports for connection to disk arrays. The partner cost for a shared memory processor (SMP) node is the actual node cost with three years of maintenance. The CRC provides rack space, fiber optic network cards, and all other necessary infrastructure to include the node within the SMP system environment.

Management of Partner Compute Resources

Both the distributed memory Linux clusters and the shared memory system use the CRC's software stack: Red Hat Linux, Sun Grid Engine 6.2, CRC file systems, and faculty-requested applications installed on the CRC's distributed file systems.

Sun Grid Engine (SGE) share-based priority scheduling is used to provide equitable access to compute resources. Partners have their own dedicated SGE queues to manage their exclusive resources and utilize general compute nodes with a higher allocation (based on a fraction of overall resources owned by the partner). All access to compute node cycles is through SGE; separate, shared login nodes provide shell access to all HPC compute resources.

In return for infrastructure and services provided by the CRC, when partner compute resources are not being utilized they are made available to the "trusted/experienced" users within the University's computational community. The CRC will help with identifying and excluding "untrusted/inexperienced" users. Trusted users will have access to Partner resources through a special, short (two to four hours long) queue. This means that if a Partner will need his/her resource back, their maximum waiting time would be two to four hours, respectively. The CRC assumes responsibility to ensure nodes shared by faculty partners will be returned to use by the faculty without detrimental remnants from prior jobs/simulations.

The CRC will negotiate a memorandum of understanding (MoU) with each program partner. In general, after three years from the initial hardware installation/operation date, this MoU will be revisited. As equipment ages, the CRC will work with all partners in advance to raise funds for hardware replacement.