HyperClustering@CSAIL.MIT.EDU

Need

To provide both a permanent general purpose batch queued cluster for the whole CSAIL community and multiple ephemeral customized cluster environments for subgroups with special hardware and or software needs while optimizing the use of computational hardware resources.

Vision

Customized on demand utility computing.

A 'Hypercluster' of physical compute nodes (hypernodes) running software hypervisors (Xen in this case) which can dedicate hardware resources such as CPU time, memory, and PCI bus devices among multiple running or potentially operating systems (virtual nodes) which can be built into multiple virtual clusters within the hypercluster framework.

The hypercluster will use shared storage and centralized configurations such that any virtual cluster or virtual node can be run transparently on any hypernode and migrated while running from its current hypernode to any other running hypernode with sufficient free resources.

This will allow transient clusters to be built for research groups when needed, typically prior to major conferences or competitions, and dynamically grow or shrink the general purpose cluster to accommodate these uses while optimizing utilization.

Initial Deployment

- 64bit multicore chips with 4G RAM per processor core
- Two gigabit Ethernet ports per hypernode
- Small internal disk sufficient to run the Xen hypervisor
- At least one free PCI-X slot to provide for high speed interconnects (Infiniband, Myrinet, etc.) if requested by a research group
- NAS device for virtual node storage (existing on site)
- Switching for Ethernet interconnect

Initially, provisioning and load balancing of virtual nodes will be done manually.

Future Work

The flexibility provided by this platform allows virtual test environments to be created temporarily in the same manner as other custom clusters. Some areas to be explored are:

- Scheduling and reservation systems to manage booting virtual clusters
- Dynamic load balancing system to migrate running virtual nodes to balance resource allocations
- Load visualization and trending software that is aware of the virtual cluster level of organization and can show a virtual cluster's location and resource use within the hypercluster