

# Building & Maintaining a Cluster of GPUs

by Prof. Tsuyoshi Hamada

This course is based on the experience at the Nagasaki Advanced Computing Center building a GPU cluster to win the Gordon Bell prize in the price/performance category. After several years trying, finally the coveted prize was awarded to Professor Hamada and his co-authors in 2009.

Two main sections in the course will cover the hardware & software aspects, respectively.

## Hardware

The following topics will be covered:

- server and GPU racks
- air conditioning
- power supply
- network
- file servers

Based on his experience, Prof. Hamada will give tips about the choice of components, depending on the size of the cluster. He will describe things that can go wrong and how to avoid them—e.g., proper grounding, using spacers to avoid sparks!

Suggestions about power and air conditioning requirements will be offered, and suggestions about network choices. For example, if the cluster will have 36 nodes or less, choose Ethernet; for more than 36 nodes, a better choice is Infiniband.

Some details that need attention are cable management, and power outlets; solutions to the problems that may arise in this and other hardware issues will be suggested. Prof. Hamada will describe

*What are some tested configurations for building a cluster of commodity hardware, with GPU acceleration? How should the power and air conditioning requirements be determined? What tools can be used to efficiently manage the cluster?*

issues, show photographs of several stages of his cluster, and discuss pricing, sourcing, timing, and labor requirements.

## Software

There are two aspects to be considered in terms of software requirements: the system configuration (operating system, file system, setting accounts) and the cluster management tools. Prof. Hamada will spend time explaining his tested solutions for these two aspects, and also demonstrate his own tools for parallel management. He will not only demonstrate, but make available his programs providing parallel shell scripts for broadcasting instructions to all nodes. He will also share his testing suite, to be used for detecting faulty gaming cards, and discuss strategies for fault-tolerant computing.

Participants in this course will have access to the wealth of knowledge obtained from building several generations of GPU clusters. Starting in March 2007 with a 1 Tflop/s system, and ending with the award-winning GPU cluster *Degima* with 576 GPUs.

### Degima cluster

One of the 144 nodes, equipped with four GT200 chips, and Professor Hamada posing next to the hand-built cluster of GPUs in Nagasaki.

