MPICH on K, PostT2K, and Future Machines

Yutaka Ishikawa
University of Tokyo
RIKEN AICS
MPICH on K

• The first step to porting K computer
  – TCP/IP on Tofu

In K, Open MPI is provided by Fujitsu

• The second step
  – Native Tofu
  – Instead of using just native Tofu interface, we are currently designing lower-level communication layer for K, Infiniband, and etc.

MPICH/K, developed by Masayuki Hatanaka and Toyohisa Kameyama
MPICH on PostT2K

• What is PostT2K machine
  – Designed and will be installed by U. of Tokyo and U. of Tsukuba
  – Based on a commodity manycore architecture
  – O(8K to 10K) nodes

• Software Architecture being considered and developed
  – McKernel/IHK: light-weight micro kernel sitting next Linux kernel using IHK (Interface for Heterogeneous Kernels)
  – DCFA: Direct Communication Facility for Advanced manycore architectures
  – MPICH/DCFA

McKernel and MPICH/DCFA are open source. The USB memory containing those sources is distributed at Booth #1229, PC Cluster Consortium Booth.

DCFA developed by Min Si, a PhD student
MPICH/DCFA developed by Masamichi Takagi, NEC
For workaround of a Sandybridge performance issue, data area is copied to the host machine and that area is sent in the case of larger messages.

For more information, read the following paper:

Important dates:

• Submission deadline of full papers and poster abstracts: April 25th, 2014
• Paper/Poster Acceptance Notification: May 30th, 2014
• Camera Ready: June 20th, 2014
• Early Registration: July 18th, 2014
• Conference: September 9th-12th, 2014
DCFA-MPI, MPI Implementation based on our MPI Implementation

**Host**
- DCFA-MPI CMD Server
- DCFA CMD Server
- DCFA (Host) IBV Extension
- DCFA (Host) MLX4 Library Extension
- IB Verbs Library
- MLX4 Library
- IB uverbs
- DCFA (Host) Modified IB core
- MLX4 Driver

**Xeon Phi**
- DCFA-MPI CMD Client
- DCFA IB IF
- DCFA-MPI CMD Server
- DCFA-IKC
- IHK-IKC
- IHK-lichens

The OFED kernel module is modified to take care of memory area of PCI device, i.e., MIC memory area.