UMBC High Performance Computing Facility

Contact: Matthias K. Gobbert, Department of Mathematics and Statistics, www.umbc.edu/hpcf
Research Assistants: Xuan Huang, Samuel Khuvis, Jonathan Graf, Jack Slettebak, Andrew Raim
Governance Committee: M. K. Gobbert, C. R. Menyuk, M. Olano, L. Sparling, L. Strow, I. Thorpe, C. Welty

HPCF Introduction

The UMBC High Performance Computing Facility (HPCF) is the community-based, interdisciplinary core facility for scientific computing and research on parallel algorithms at UMBC. Started in 2008 by more than 20 researchers from ten academic departments and research centers from all three colleges, it is supported by faculty contributions, federal grants, and the UMBC administration. The facility is open to UMBC researchers at no charge. Researchers can contribute funding for long-term priority access. System administration is provided by the UMBC Division of Information Technology, and users have access to consulting support provided by dedicated full-time graduate assistants.

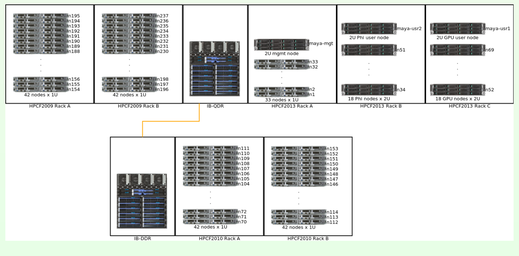
240-Node Cluster maya

- **HPCF2013 = maya (2013) = $540,000:**
  72 nodes, each with two 2.6 GHz eight-core Intel E5-2650v2 Ivy Bridge CPUs:
  - 34 CPU-only nodes
  - 19 hybrid nodes with two NVIDIA K20 GPU
  - 19 hybrid nodes with two Intel Phi 5110P
- **HPCF2010 = maya (2010) = gift from NASA:**
  84 nodes, each with two 2.8 GHz quad-core Intel Nehalem X5560 CPUs
- **HPCF2009 = maya (2009) = $600,000:**
  84 nodes, each with two 2.6 GHz quad-core Intel Nehalem X5550 CPUs

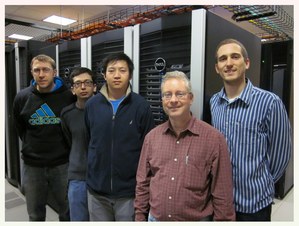
Networks connecting all components:

- quad-data rate (QDR) InfiniBand interconnect for HPCF2013 and HPCF2009
- dual-data rate (DDR) InfiniBand interconnect for HPCF2010

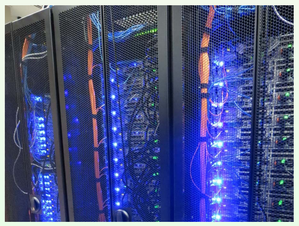
Storage of more than 750 TB connected by IB.



Photos

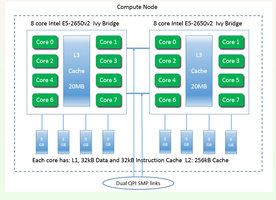

HPCF2013 Front


QDR IB Front


HPCF2013 Back


QDR IB Back

HPCF2013: 2 Eight-Core CPUs



Each node contains two eight-core Intel E5-2650v2 Ivy Bridge CPUs. The 64 GB of the node's memory are connected to CPUs via 8 memory channels. The two CPUs of a node are connected to each other by two QPI (quick path interconnect) links.

Acknowledgments

- REU Site www.umbc.edu/hpcreu
- CIRC www.umbc.edu/circ
- NASA Goddard Space Flight Center
- NSF (MRI, SCREMS), NSA, UMBC, CIRC

References

- HPCF–2014–6, HPCF–2014–7, HPCF–2014–8
- REU Site: HPCF–2013–13, HPCF–2014–14

The Poisson Equation as Challenging Test Problem

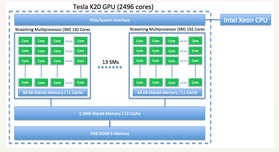
The Poisson equation with homogeneous Dirichlet boundary conditions

$$\begin{aligned} -\triangle u &= f \qquad \text{in } \Omega = (0,1)\times(0,1) \subset \mathbb{R}^2 \\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned}$$

Wall clock time in HH:MM:SS on maya 2013 by number of nodes and processes per node for mesh resolution $N \times N = 16{,}384 \times 16{,}384$:

	1 node	2 nodes	4 nodes	8 nodes	16 nodes	32 nodes	64 nodes
1 ppn	14:08:26	06:57:26	03:29:58	01:45:31	00:53:29	00:27:13	00:13:57
2 ppn	07:01:30	03:31:53	01:46:36	00:54:13	00:27:20	00:14:12	00:07:15
4 ppn	03:55:38	01:58:32	01:00:05	00:30:26	00:15:32	00:08:15	00:04:20
8 ppn	02:55:26	01:28:52	00:44:32	00:22:30	00:11:26	00:06:23	00:03:22
16 ppn	02:49:13	01:25:16	00:43:08	00:22:10	00:11:33	00:06:31	00:06:34

Hybrid Nodes with 2 GPUs



19 hybrid nodes contain two NVIDIA K20 GPUs with 2496 compu–tational cores with 5 GB of global memory.

We use NVIDIA's CUDA (Compute Unified Device Architecture) to program the GPUs and MPI (Message Passing Interface) to utilize resources on multiple nodes. Maya offers the opportunity for research on multiple nodes with GPUs. The most challenging part of GPU programming is to design algorithms with minimal data transfer between CPU and GPU.

CPU/GPU Poisson on $16{,}384 \times 16{,}384$ mesh:

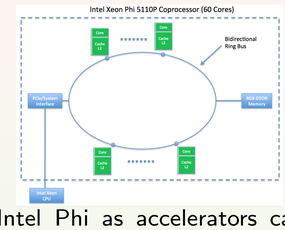
Per node:	4 nodes	8 nodes	16 nodes
1 GPU	01:41:29	00:51:09	00:25:17
2 GPUs	01:14:32	00:37:29	00:18:50

The 3-D HPCG Benchmark

The High Performance Conjugate Gradient (HPCG) Benchmark, developed by Sandia National Laboratories (http://software.sandia.gov/hpcg), is a realistic test of parallel clusters using a Poisson problem in three dimensions.

The MPI–OpenMP reference implementation of HPCG Revision 2.4 on maya achieves at least 344.512 GFLOP/s on a $1{,}280 \times 1{,}280 \times 1{,}280$ mesh using 64 nodes, 8 MPI processes per node, and 2 threads per MPI process.

Hybrid Nodes with 2 Intel Phi



19 hybrid nodes contain two 60-core Intel Phi 5110P. The 8 GB of memory are connected through a bidirectional ring bus.

Intel Phi as accelerators can be used to increase throughput, and the x86 architecture can minimize the programming effort. There are three ways to access Intel Phi: native execution on the Phi, Offloading to the Phi, and symmmetric mode using both CPU and Phi heterogeneously. The most challenging part of Phi programming is to design algorithms with minimal data access across the ring bus.

CPU/Phi Poisson on $16{,}384 \times 16{,}384$ mesh:

Per node:	4 nodes	8 nodes	16 nodes
1 Phi	01:49:22	00:56:46	00:31:01
2 Phis	00:56:15	00:31:00	00:18:27

InfiniBand Stresstest

The low-latency InfiniBand interconnect overcomes network contention of an MPI_Alltoall command for constant global problem size of 6 GB distributed across all cores of N nodes:

N nodes	1	3	9	18	36
time (sec.)	1.14	0.57	0.25	0.15	0.11