Open MPI Community Meeting

SC06, November 14, 2006

Jeff Squyres, Cisco Systems
Brian Barrett, Los Alamos National Laboratory

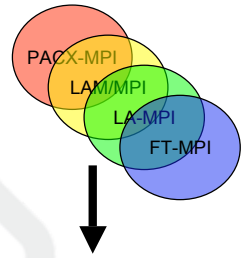
LA-UR 06-8038

Open MPI Sponsors

- DoE
  - ASC
  - LANL CCS-1
  - NNSA
- HLRS
- Lilly Endowment
- Microsoft
- NSF

Open Source High Performance Computing

- Open source implementation of MPI-2
- Combined expertise from 4+ previous MPIs
- High performance & robust
- Works with most interconnects
- Modular Component Architecture
  - Combinatorial capabilities
  - Function pointers faster than shared library calls



OPEN MPI

Current Members

Academia / Research

- HLRS
- Indiana U.
- Sandia National Lab
- Los Alamos National Lab
- U. of Dresden
- U. of Houston
- U. of Tennessee

Industry

- Cisco
- IBM
- Mellanox
- Myricom
- QLogic
- Sun
- Voltaire

Current Status

- Stable release: v1.1.2
  - v1.1.3 expected “soon”
- Upcoming release: v1.2
  - Stability and scalability improvements
  - Sun / Solaris / N1GE / uDAPL support
  - Better MX support
  - InfiniPath support
  - TotalView message queue support
  - …and more

Top 500

- #6: Sandia Thunderbird cluster
  - Dell PowerEdge 1850
  - InfiniBand
- Linpack result
  - 4347 dual processor nodes
  - 53 teraflops
  - 84.66% network efficiency
- Powered by Open Fabrics / Open MPI

Future Directions in Open MPI

Brian Barrett
Los Alamos National Laboratory /
Indiana University

Future Development Plans

- Improved collectives support
  - Better performance
  - Utilize hardware collectives
  - Non-blocking interface
- Fault Tolerance
  - Process checkpoint / migration
  - Network reliability / failover

Future Development Plans

- Scalability improvements
  - Faster, more reliable launching
  - OpenIB Unreliable Datagram (UD) support
- IPv6 support
- Heterogeneous architecture support
- Microsoft Windows support

Open MPI as a Research Platform

- Highly portable source code
- Limited requirements on architecture
  - C / C++ compiler
  - TCP stack or pre-wired network
  - Unix-like environment (Windows coming soon)
- Straight forward interconnect interfaces
  - Reliable RDMA
  - Unreliable send/receive
  - Matching send/receive

Getting Involved

- Source code repository publicly available
- Active, open developer mailing lists
- Straight-forward process for pushing code into main repository
- Always looking to add new members

Questions for Audience

What do You Want From MPI?

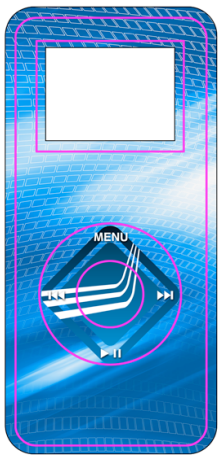
(audience -- you talk now)

How Important Is…

- Thread safety
  - Multiple threads making simultaneous MPI calls
- Parallel I/O
  - Working with parallel file systems
- Dynamic processes
  - Spawn, connect / accept
- One-sided operations
  - Put, get, accumulate

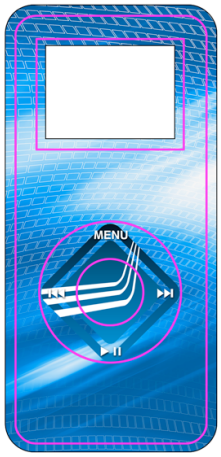
Open OMPI SC06 Events

- iPod giveaways!
- Booth talks today
  - 1:30 IU/Past, present future
  - 3:00 Sun/Open source HPC
- Tomorrow
  - 11:00 Mellanox/OMPI and IB
  - 11:30 ORNL/Overview
  - 1:40 Microsoft/Windows OMPI
  - 3:30 ASC/Fault Tolerance
  - 4:00 ORNL/Overview
  - 4:00 IU/OMPI Research
  - 4:00 AIST/Heterogeneity



Open OMPI SC06 Events

- Thursday
  - 11:00 IU/Wide-reaching fault tolerance
  - 12:40 Microsoft/Windows OMPI
  - 1:00 Mellanox/OMPI and IB
- Full schedule
  - http://www.open-mpi.org/sc06/



Come Join Us!

http://www.open-mpi.org/