Open MPI
Join the Revolution

Jeff Squyres
Indiana University

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

Technical Contributors

• Indiana University
• The University of Tennessee
• Los Alamos National Laboratory
• High Performance Computing Center, Stuttgart
• Sandia National Laboratory - Livermore

MPI From Scratch!

• Developers of FT-MPI, LA-MPI, LAM/MPI
  ▪ Kept meeting at conferences in 2003
  ▪ Culminated at SC 2003: Let’s start over
  ▪ Open MPI was born

Jan 2004  SC 2004  Today  Tomorrow
Started work  Demonstrated  Released v1.0  World peace

MPI From Scratch: Why?

• Each prior project had different strong points
  ▪ Could not easily combine into one code base
  ▪ New concepts could not easily be accommodated in old code bases
  ▪ Easier to start over
    ▪ Start with a blank sheet of paper
    ▪ Decades of combined MPI implementation experience

MPI From Scratch: Why?

• Merger of ideas from
  ▪ FT-MPI (U. of Tennessee)
  ▪ LA-MPI (Los Alamos)
  ▪ LAM/MPI (Indiana U.)
  ▪ PACX-MPI (HLRS, U. Stuttgart)

Open MPI Project Goals

• All of MPI-2
• Open source
  ▪ Vendor-friendly license (modified BSD)
• Prevent “forking” problem
  ▪ Community / 3rd party involvement
  ▪ Production-quality research platform (targeted)
  ▪ Rapid deployment for new platforms
• Shared development effort
Open MPI Project Goals

• Actively engage the HPC community
  • Users
  • Researchers
  • System administrators
  • Vendors
• Solicit feedback and contributions

True open source model

Design Goals

• Extend / enhance previous ideas
  • Component architecture
  • Message fragmentation / reassembly
  • Design for heterogeneous environments
    • Multiple networks (run-time selection and striping)
    • Node architecture (data type representation)
    • Automatic error detection / retransmission
    • Process fault tolerance
    • Thread safety / concurrency

Design Goals

• Design for a changing environment
  • Hardware failure
  • Resource changes
  • Application demand (dynamic processes)
• Portable efficiency on any parallel resource
  • Small cluster
  • “Big iron” hardware
  • “Grid” (everyone a different definition)
  • …

Plugins for HPC (!)

Networks
  • Shmem
  • TCP
  • OpenIB
  • mVAPI
  • GM
  • MX

Run-time environments
  • rsh/ssh
  • SLURM
  • PBS
  • BProc
  • Xgrid

Your MPI application

Plugins for HPC (!)
Plugins for HPC (!)

Networks
- Shmem
- TCP
- OpenIB
- mVAPI
- GM
- MX

Your MPI application
- Shmem
- TCP
- GM

Run-time environments
- rsh/ssh
- SLURM
- PBS
- BProc
- Xgrid

Current Status

- v1.0 released immanently (see web site)
- Much work still to be done
  - Data and process fault tolerance
  - Support more run-time environments (Grid!)
  - Interoperable MPI (IMPI) functionality
  - More external tools
  - …
- Come join the revolution!