Open MPI: $10^{15}$ Flops Can't Be Wrong

Open Source High Performance Computing

Jeff Squyres
Open MPI Architect

Petaflop!!

- Los Alamos Road Runner
- #1 on Nov. 2008 Top500
  1.1 petaflops
- Powered by Open MPI
  Significant community achievement
Open MPI Is…

- Evolution of several prior MPI implementations
- Open source project and community
  - Production quality
  - Vendor-friendly
  - Research- and academic-friendly
- All of MPI-1 and MPI-2

15 Members, 9 Contributors, 2 Partners
Why Does Open MPI Exist?

- Maximize all MPI expertise
  - Research / academia
  - Vendors
  - Customers, enterprise
  - ...elsewhere
- Capitalize on years of MPI research and implementation experience
- The sum is greater than the parts

“Great discoveries and improvements invariably involve the cooperation of many minds.”

Alexander Graham Bell, 1877
Cisco: Why Open MPI?

- It seems obvious to us!
  - Why re-invent the wheel?
  - Who would want “Cisco MPI”?
  - Combined community resources

- Meshes with Cisco values
  - Standards-based
  - Open architectures
  - Consensus driven
  - Collaborate to innovate

Cisco votes “yes” for community MPI

Cisco’s Community Role

- Active development
  - Design, code

- [Very] Extensive testing
  - 300-500k regression tests/night
  - Data fed back to community

- Logistics support
  - Face-to-face engineering meetings

- Member, MPI Forum
Cisco’s Open MPI Goals

**Technical**
- Promote standards
  - Ethernet-based technologies
  - Commodity clusters
- Integrate with tools
  - Make parallel programming [a little] easier
- Understand and accelerate applications
  - RAB and DAL two (pseudo-HPC) examples

**Non technical**
- Promote community
  - Conferences, tradeshows
  - Contribute on open mailing lists
- Partner with academics and researchers
  - Foster cutting-edge research
- Perform “community service”
  - Example: Fortran API maintenance

---

...But How Does That Equal a Petaflop?

- Robustness through QA
- Ease of deployment
- Scalable connectivity scheme
- High quality error reporting (MTBF)
- Small MPI memory footprint
- Scalable job launch
- Scalable MPI collectives

OpenFabrics performance

Optimized Fortran API

PETAFLOP
“Open source is decided by those who show up.”

Cisco is there. Come join us.