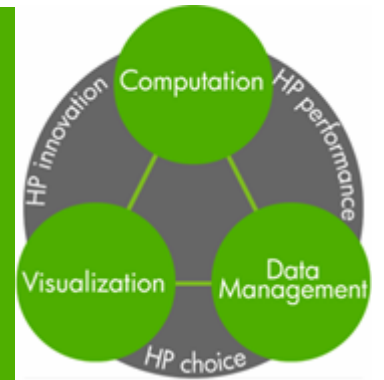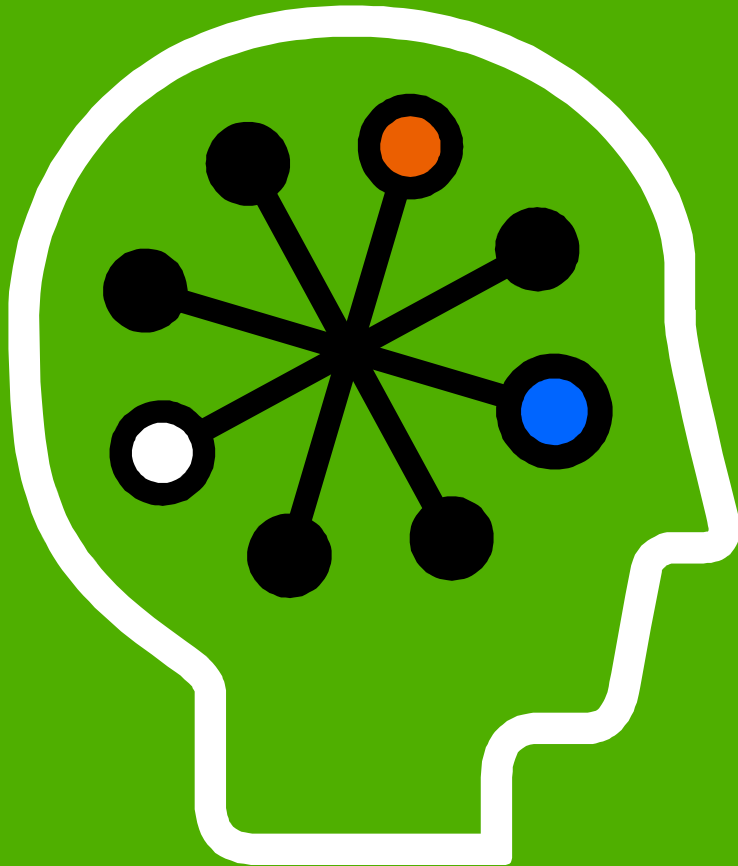HP-MPI™

Industry's Standard Commercial High Performance MPI

Kannan Narasimhan

HP High Performance Computing Division

November, 2007

Technology for better business outcomes

Introduction to HP-MPI

- Universal MPI library supporting full choice of industry-standard processors, interconnects, OSs and HP turnkey systems
  - Linux, Windows CCS and HP-UX complex computing environments
  - Key component of HP HPC solutions
- Enables a single executable and qualification for each platform
  - where Platform = Processor + Operating Environment
  - transparently supports all leading industry-standard switches and interconnects
- Similar interface and run-time environment across all platforms
- Object compatible with MPICH V1.2.5 and later
- Easily deployed when built as a shared, dynamically linked library

HP-MPI: Industry's Standard Commercial High Performance MPI

Portable, Robust, Fast, Supported, Available

- Optimized for all leading architectures, OSs, switches, applications (see below)
- Distributed by >25 ISVs to Linux customers on both HP & non-HP clusters

Supported Operating Systems

ISVs Distributing HP-MPI on Industry-Standard Linux Clusters

Supported Architectures

Supported Switches

HP-MPI Portability for users

OS:	Linux						Windows CCS		HP-UX 11i & 11i V2		
Switch:	Quadrics		Myrinet		Infini-Band *	Gigabit Ethernet		Infini-Band	GigE	Infini-Band	Hyper-Fabric 2	GigE **
Protocol:	Elan3	Elan4	GM2.1	MX	uDAPL VAPI PSM OFED	Level 5 RDMA	TCP/IP	IBAL WSD	TCP/IP	IT-API	HP	TCP/IP
Integrity (Itanium)	V2.2	V2.2.5 XC	V2.2.5.1 XC	V2.2.5.1	V2.2.5.1 XC		V2.2.5.1 XC			V2.2	V2.1	V2.2
ProLiant (Opteron)		V2.2.5 XC	V2.2.5.1 XC	V2.2.5.1	V2.2.5.1 XC	V2.2.5.1	V2.2.5.1 XC	V1.1	V1.1			
ProLiant (Xeon32 & EM64T)			V2.2.5.1 XC	V2.2	V2.2.5.1 XC	V2.2.5.1	V2.2.5.1 XC	V1.1 X86 only	V1.1 X86 only			
HP9000 (PA-RISC)											V2.1	V2.2

*   v2.2.5 supports all 4 IB vendor's hardware: Voltaire, Cisco, SilverStorm, Mellanox (including support for InfiniPath)
**   v2.2 supports 10GigE switches for Itanium/HP-UX

Current Version: v2.2.5.1

HP-MPI v2.2.5 release (Linux) Interconnect Support Enhancements

- Enhanced usability of interconnect environment variables
- IB partitioning
- QLogic's InfiniPath™ support
- Myrinet MX support
- Expansion of Signal Propagation
- New mpirun option for intra-host performance tuning
- Fast one-sided lock/unlock under InfiniBand VAPI
- OpenFabrics Verbs (IBV) support

HP-MPI v2.2.5.1 release (Linux) Highlights

- Adds OpenFabrics 1.2, 1.2.5 support
- Additional cpu-binding features

"HP-MPI is an absolute godsend," notes Keith Glassford, director of the Materials Science division at San Diego, CA-based Accelrys Software Inc. "It allows us to focus our energy and resources on doing what we're good at, which is developing scientific and engineering software to solve customer problems."

HP-MPI Coming Attractions in 2008

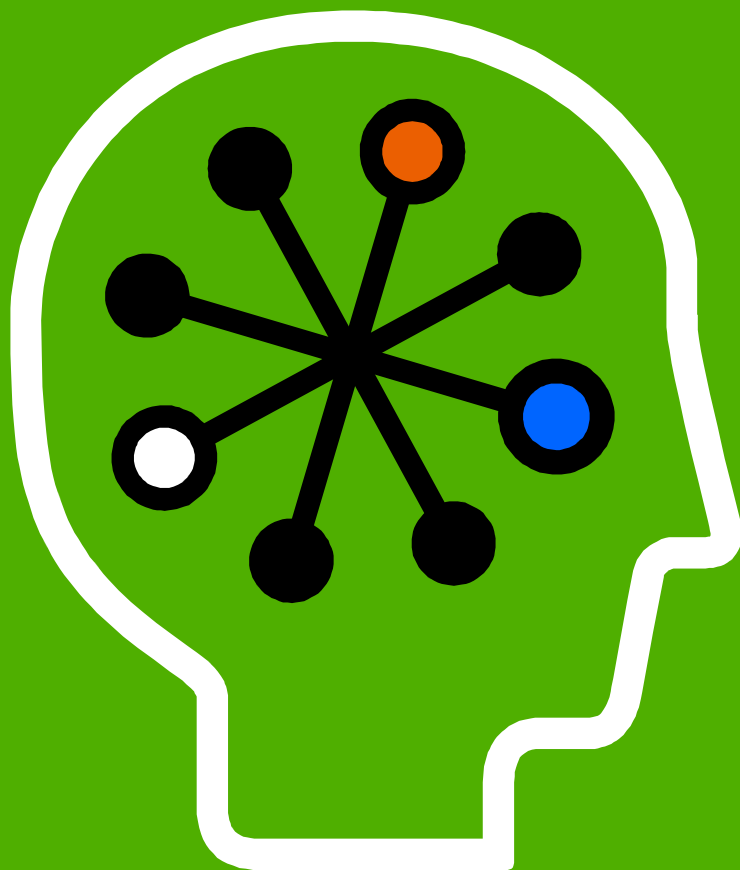| Improved Performance	High Availability	Interoperability
• XRC support • Improved message rate • Improved performance of MPI Collectives • Shared Memory and Network utilization strategies for multiple cores • Topology aware CPU/rank mapping	• Network HA • Card fail-over • Port fail-over • Infrastructure HA • Eliminate single point of failure • Dynamic failure-free subsets of communicators • Completion of collectives across rank failures	• More Interconnects (iWARP) • More tools (PGI Debuggers, MPICH2 compatibility, Intel Debuggers) • More protocols (IPV6)

Observations on the impact of OFA

- OFED based implementations gaining popularity
- Deployed in ISV applications and production clusters
  - Backward compatibility is key
  - Clearly articulate failure scenarios
    - Clarification on corner cases for fork() failure
  - Exposure to limits that show up only at high scale
  - Think about new Verbs API and driver enhancements that can improve performance in large scale clusters:
    - Pin/unpin multiple regions in one call
    - Reduce blocking at the driver layer where possible
    - Increase default per-card QP limits
    - Memory optimizations at the driver
    - Better error messages from driver where possible

Thank you!

http://www.hp.com/go/mpi

Technology for better business outcomes