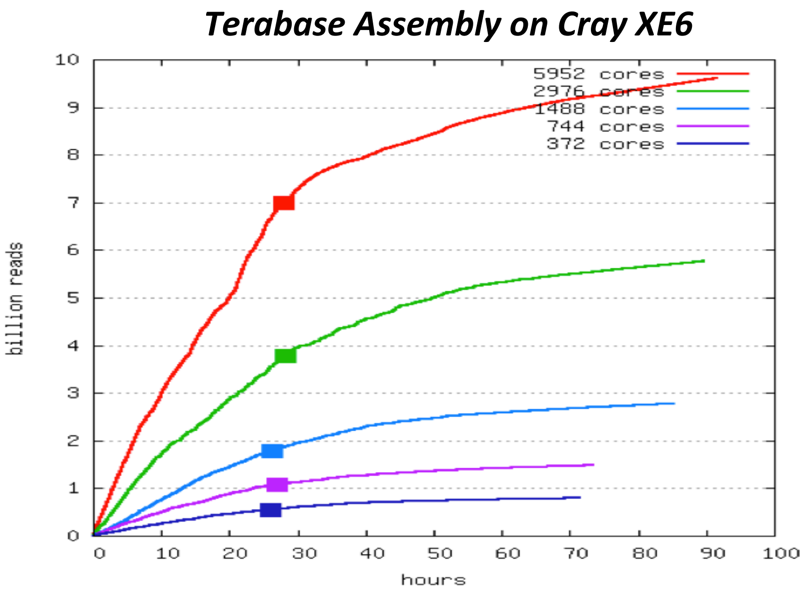
Evolving MPI to Address the Challenges of Exascale Systems

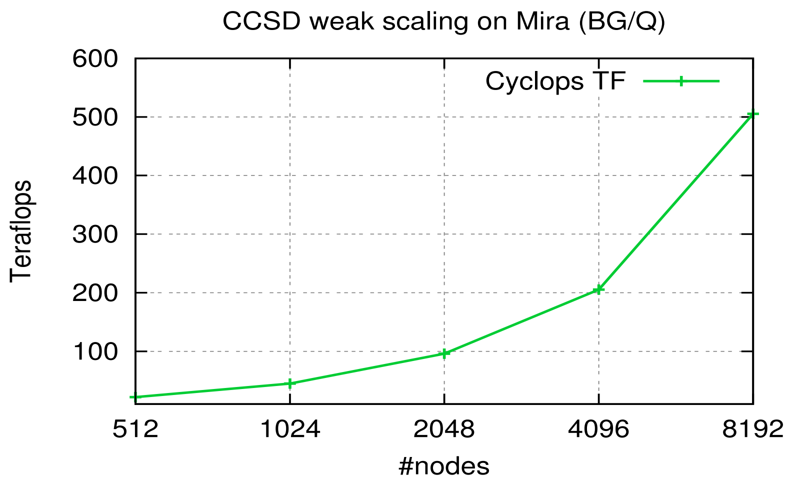
- **Myths**
  - Several myths surrounding MPI: it cannot support dynamic execution, it is not fault tolerant, it cannot take advantage of new architectures, …
- **Project Goals and Core Problems Tackled**
  - Efficient Runtime for High-level Programming
    - Dynamic execution environments (e.g., Charm++, ADLB)
    - Global communication models (e.g., PGAS models, Global Arrays, GVR)
  - Interoperability for Hybrid Programming
    - Interoperability of MPI with lightweight threading and task models
    - Interoperability of MPI with heterogeneous computing models
  - Performance, Scalability, and Resilience
    - Techniques to address challenges posed by new architectures and systems with millions of cores (performance and memory scalability, resilience, …)
- **Impact**
  - Vendors and Supercomputers
    - MPICH is used on virtually every supercomputer in the world (IBM, Cray, Intel, Microsoft, OSU MVAPICH) and on the largest systems in the Top500 list
    - 9 of the top 10 systems in the June 2014 Top500 list use MPICH
    - This project will enable the deployment of a high-performance implementation of the newest version of the MPI standard (MPI-3) on these systems
  - Applications
    - Close interactions with applications to enable them to use advanced functionality in MPI (either directly or through domain-specific models)
    - Recently demonstrated impact with massive-scale computations in various domains including Chemistry, Biology and Nuclear Physics
    - Ongoing effort in solving large computations in other domains

Rajeev Thakur, Pavan Balaji, Marc Snir, Rusty Lusk

Some Success Stories with Applications

Argonne NATIONAL LABORATORY





- **Terascale Genome Assembly**
  - Graph assembly problem that deals with finding an Hamiltonian path in a fuzzy graph with erroneous edges (or non-edges)
  - Highly communication intensive, with (seemingly) random global communication
  - Genome assembly at this scale (2.3TB on Cray XE6) has, for the first time, allowed scientists to study multiorganism genome colonies of completely or partially unknown species
- **Cyclops Tensor Framework (Chemistry)**
  - Fundamental component of quantum chemistry for coupled-cluster methods
  - Supersedes existing algorithms and software for parallel tensor contractions
  - Enabled quantum simulations of 250 electrons in 1000 orbitals (no point-group symmetry) on Argonne Mira
    - Order of magnitude larger scale problem than anything that has been previously done

Recent Accomplishments

- **Publications**
  - **SC14:** "Nonblocking Epochs in MPI One-Sided Communication," Judicael Zounmevo, Xin Zhao, Pavan Balaji, William Gropp, Ahmad Afsahi. (Selected as **Finalist for Best Paper Award**)
  - **SC14:** "MC-Checker: Detecting Memory Consistency Errors in MPI One-Sided Applications," Zhezhe Chen, James Dinan, Zhen Tang, Pavan Balaji, Hua Zhong, Jun Wei, Tao Huang, Feng Qin
  - **ICS'14:** "MT-MPI: Multithreaded MPI for Many-Core Environments," Min Si, Antonio Peña, Pavan Balaji, Masamichi Takagi, Yutaka Ishikawa
  - **PPoPP'14:** "Portable, MPI-Interoperable Coarray Fortran," Chaoran Yang, Wesley Bland, Pavan Balaji, John Mellor-Crummey
  - **ScalComm'13: "**Optimization Strategies for MPI-Interoperable Active Messages," Xin Zhao, Pavan Balaji, William Gropp, Rajeev Thakur **(Best Paper Award)**
- **Tutorials**
  - **SC14:** "Advanced MPI Programming," Pavan Balaji, William Gropp, Torsten Hoefler, Rajeev Thakur. Accepted as full-day tutorial.
  - Gave a full-day tutorial on June 6th at Argonne to around 180 attendees from local area institutions (Argonne, Northwestern, UChicago, UIUC, UIC, NIU, Loyola)