FUJITSU

shaping tomorrow with you

FUJITSU “PHI” Turnkey Solution

Integrated ready to use XEON-PHI based platform

Dr. Pierre Lagier

ISC2014 - Leipzig

PHI Turnkey Solution challenges

FUJITSU

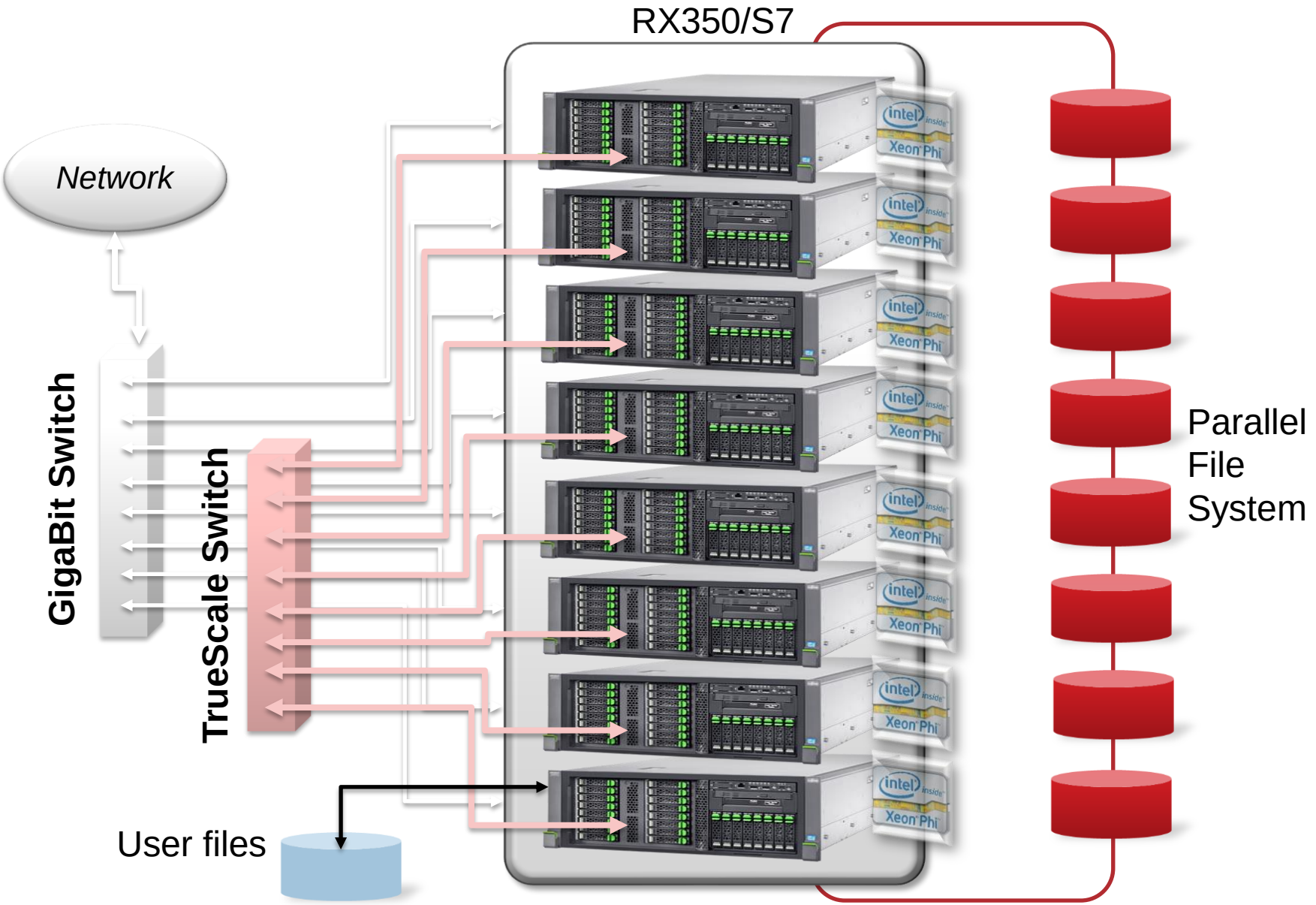
- System performance challenges
  - Parallel IO best architecture design and fine system tuning, includes the integration of SSDs technology
  - Known bottlenecks on application performance, like PCI bus and host sockets relationship
- Application challenges
  - Hybrid programming model and related performance issues (latency, synchronization overhead, MPI sustained bandwidth between PHI boards) must be addressed in parallel with system performance challenges
  - How end users will benefit from using a Web portal (PRIMERGY Gateway) to hide the heterogeneity and related issues ?
- Environment challenges
  - Full integration of all software components with the Cluster Deployment Manager tool.

FUJITSU

FUJITSU “PHI” CLUSTER

The “PHI” Cluster

FUJITSU



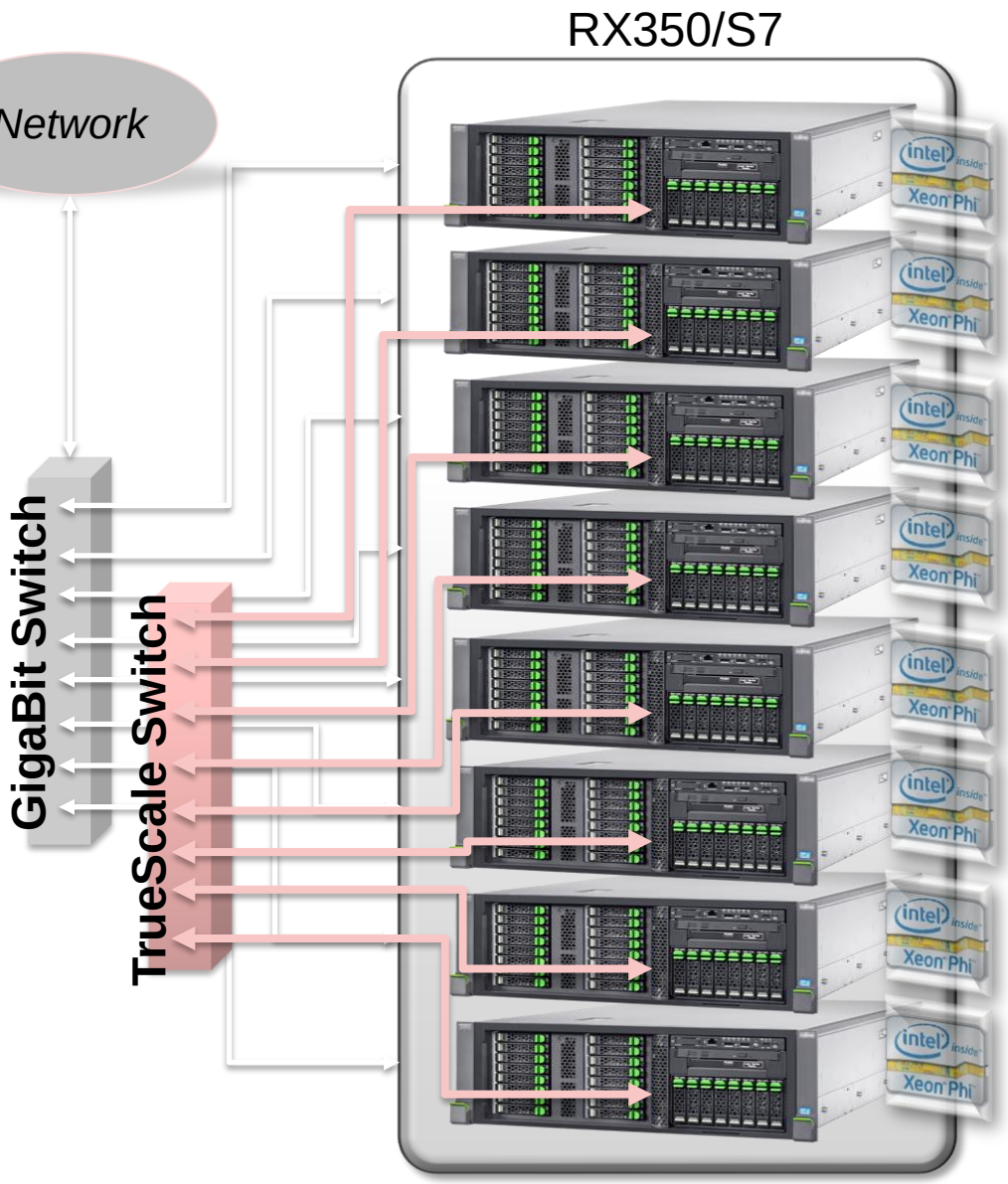
PHI Cluster: Network

FUJITSU

Ethernet and IB Network

- Single GigaBit switch
- Same shared subnet bridging XEON-PHI and all compute nodes
- Only one cluster of heterogeneous compute nodes, XEON-PHI and Dual Ivy-Bridge

- TrueScale IB switch
- Dual rail IB per compute node

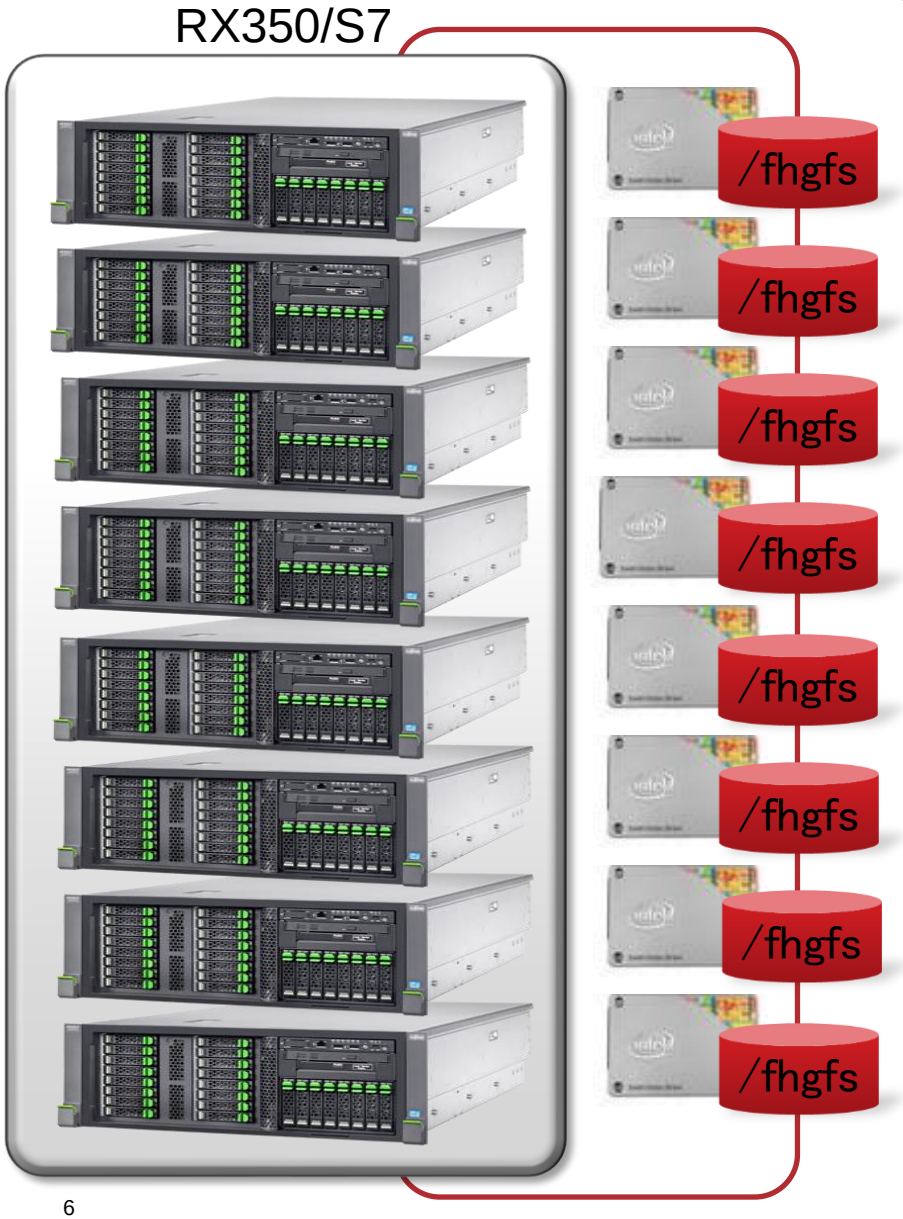


PHI Cluster: File Systems

FUJITSU

Efficiency Driven IOs

- **HOME file system**
  - On login node
  - NFS mounted on the compute nodes
- **Local scratch**
  - On each SB node
  - Mounted on connected XP node
- **Parallel file system (FHGFS)**
  - Integrated to SB nodes with one Intel SSD per node (3TB global storage capacity over 8 nodes minimal configuration)
  - Each SB node is MDS/DS/Client
  - Each XP node is client
  - Simple policies: local prefered MDS and DS from client, striping factor of 8 over all SB nodes

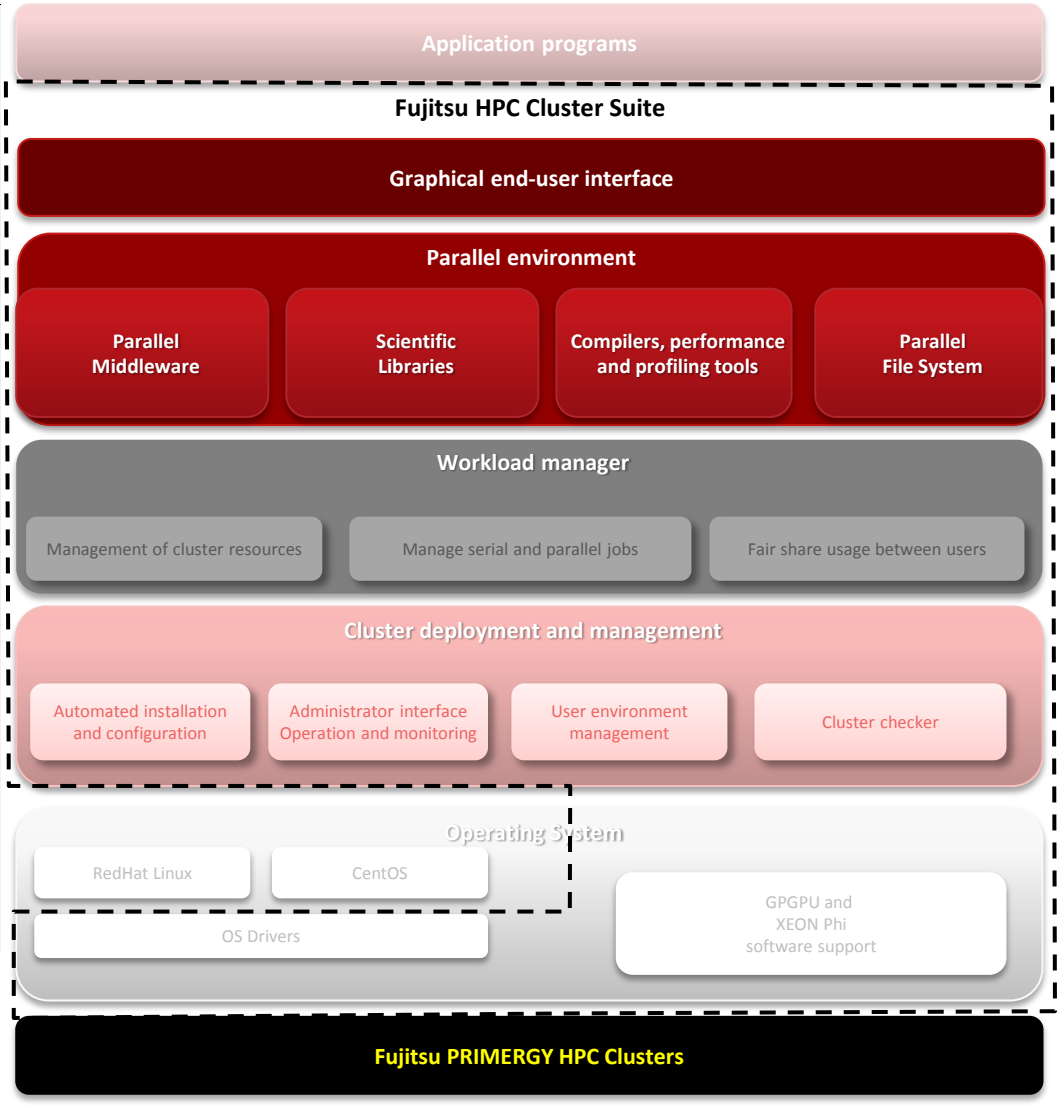


FUJITSU

FUJTISU SOFTWARE STACK

FUJITSU HPC SW Stack

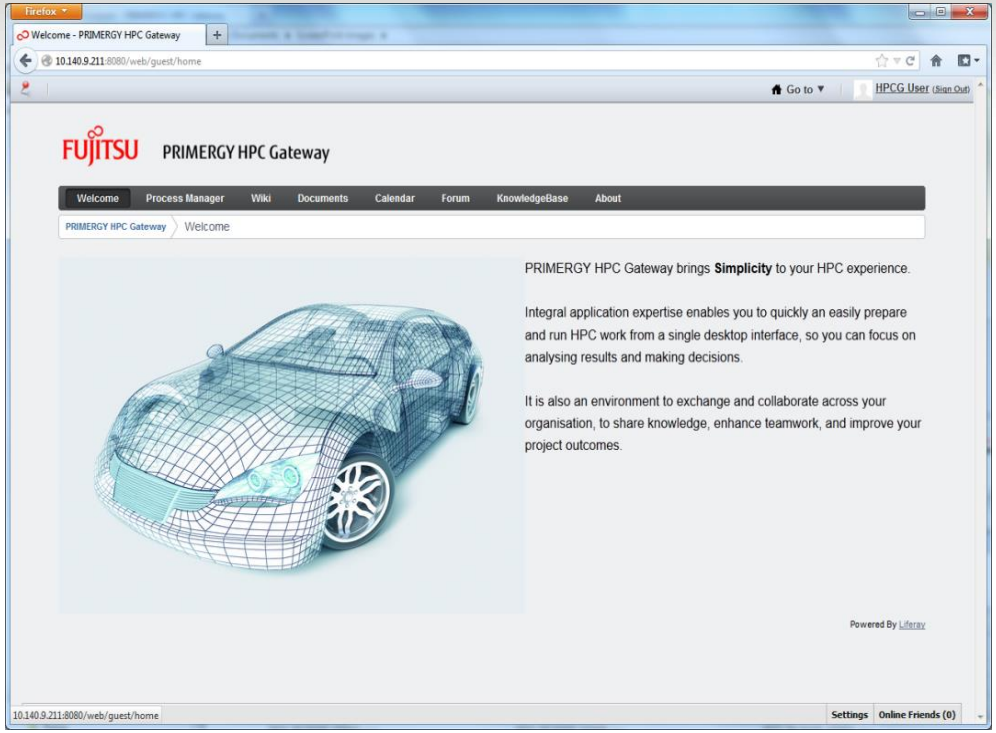
- A mature software stack includes specific software for:
  - Deploying nodes and managing software packages
  - A workload manager for job and resource management
  - Parallel execution environment with libraries
  - Tools for application development (as needed)
  - Storage options (NFS, PFS)
- These HPC software layers are always the same
  - Variety exists only in the actual components used



Fujitsu SW Stack coverage

PRIMERGY HPC Gateway

Fondation for application **solution** development



- PRIMERGY HPC Gateway is the user interface component of the FUJITSU Software HPC Cluster Suite
- Intuitive web environment incorporating application workflows, direct simulation monitoring, data access and collaboration
- Value proposition based on **simplifying** HPC end-use and integrating application **expertise,** to tune business processes and better manage projects

The Gateway delivers additional value by simplifying HPC usage – shipping since 05/2013

FUJITSU

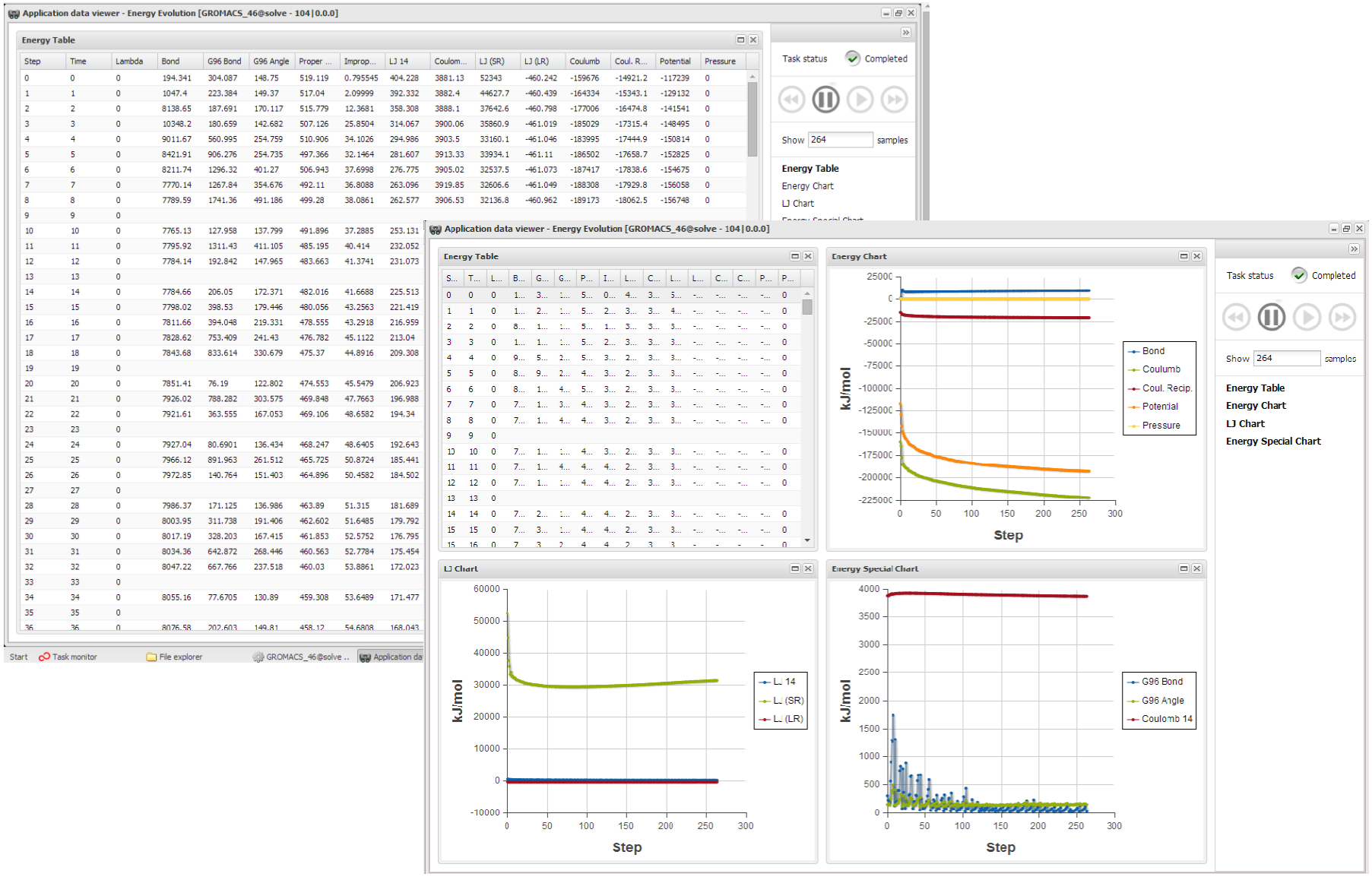
THE “GROMACS” PLATFORM

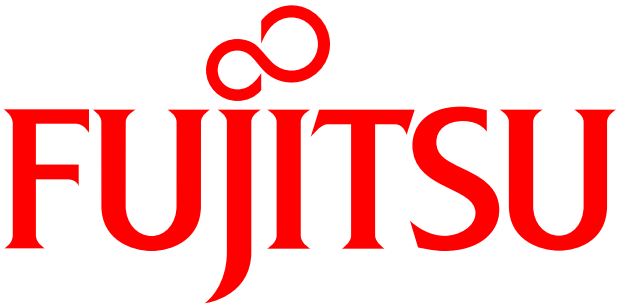
The GROMACS Environment

FUJITSU

- Based on GROMACS 5.0
  - Verlet cutoff scheme tuned for the xeon-phi
  - Running native on XEON-PHI with Fujitsu OpenMP tuning
  - MPI tuning still progressing
- Fully integrated to the HPC Gateway
  - Transparent run on front node or XEON-PHI depending of the tools used as wel as the cutoff scheme (verlet on xeon-phi)
  - Integration of key components of GROMACS (grompp, mdrun,…) with form based parameter control
  - Real time display of energies at run time
- True improvement
  - Protein-Water test case (dhfr)
  - 1.6 times faster on XEON-PHI than dual socket Sandy-Bridge

GROMACS GUI





shaping tomorrow with you