

# The “PHI” solution

Fujitsu Industry Ready Intel® XEON-PHI™ based solution

Dr. Pierre Lagier  
Fujitsu Systems Europe

ISC-2013



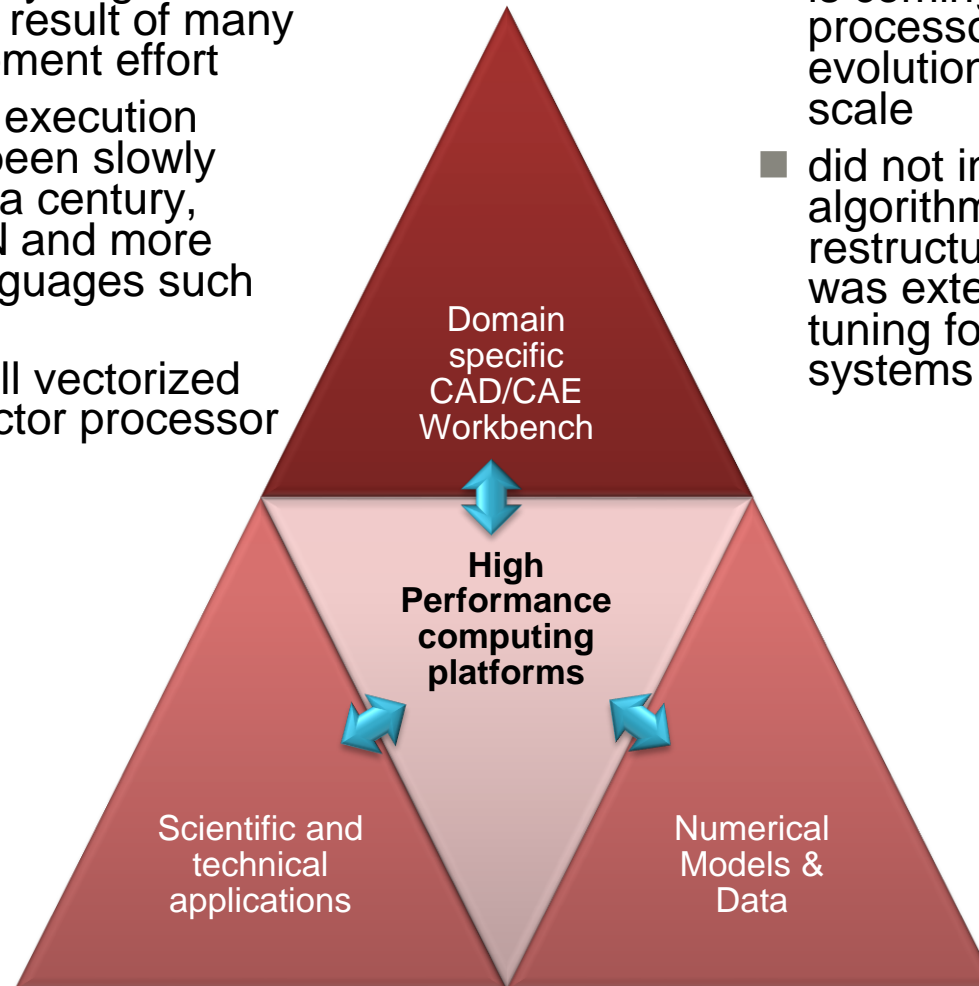
# Industrial Application Challenges

## ■ Most of existing scientific and technical applications...

- Are written for legacy execution model, mostly very large programs as the result of many years of development effort
- Are following an execution model that has been slowly evolving for half a century, using FORTRAN and more recently new languages such as C++
- Are still often well vectorized for traditional vector processor

## ■ Performance increase during the last two decades..

- is coming from single processor performance evolution and larger system scale
- did not involve major algorithmic changes or code restructuring even if there was extensive performance tuning for some specific systems



## ■ Budget side

- ◆ operational cost management is often delegated from central level to departmental/division level
- ◆ Industry projects generating revenue can benefit from significant investments to increase productivity, while less profitable non competitive enough activities can be quickly discarded
- ◆ focusing on TCO and ROI at departmental / division level, on manageable project basis toward most cost effective operation. Global investment is becoming less relevant to preserve competitiveness
- ◆ manage companies in very flexible way, optimizing the balance between cost centers and profit centers, indeed trying to decrease the first ones and increase the second ones.

## ■ Production side

- ◆ Simple environment, easy to deploy and maintain without deep technical knowledge
- ◆ Multi-user, throughput oriented usage to maximize the efficiency of the solution with simple batch system
- ◆ Reliable storage and archiving systems, homogeneous with the compute cluster (ie. Not another kind of platform with different operating system, etc...)
- ◆ Guaranteed support from ISVs for all major applications, efficient support from the manufacturer for proprietary applications porting and tuning
- ◆ Easy to learn graphical tools making more comfortable and as such more efficient day to day life of system administrators, application programmers and end users

- Real World Application from Industrial Companies
- Industry based test suite
- XEON-PHI standard compute node

## Proof of Concept

- Proven vector performance
- XEON PHI efficiency makes CFD hybrid application fast enough for competitive HPC solution

*Fujitsu "PHI"*

*European Expert team*

## Solution Design

- End user « ready to use » design
- Simple architecture with basic building blocks
- Integration of XEON PHI with Business Process engine

- Simplicity of native programming model
- Full smooth integration with existing computer resource
- Native application execution

*Fujitsu Sales Force*

*Worldwide Initiative*

## Fujitsu PHI Product Line

- Better ROI for ISVs with limited effort to port application native on XEON-PHI
- Better TCO control for end users
- Very competitive Fujitsu solution

## ■ Single rack based solution

- No extra hardware outside the rack
- No additional external storage
- Optimal for small to medium range configuration with “reasonable” storage requirements (ie. few hundreds of TB)

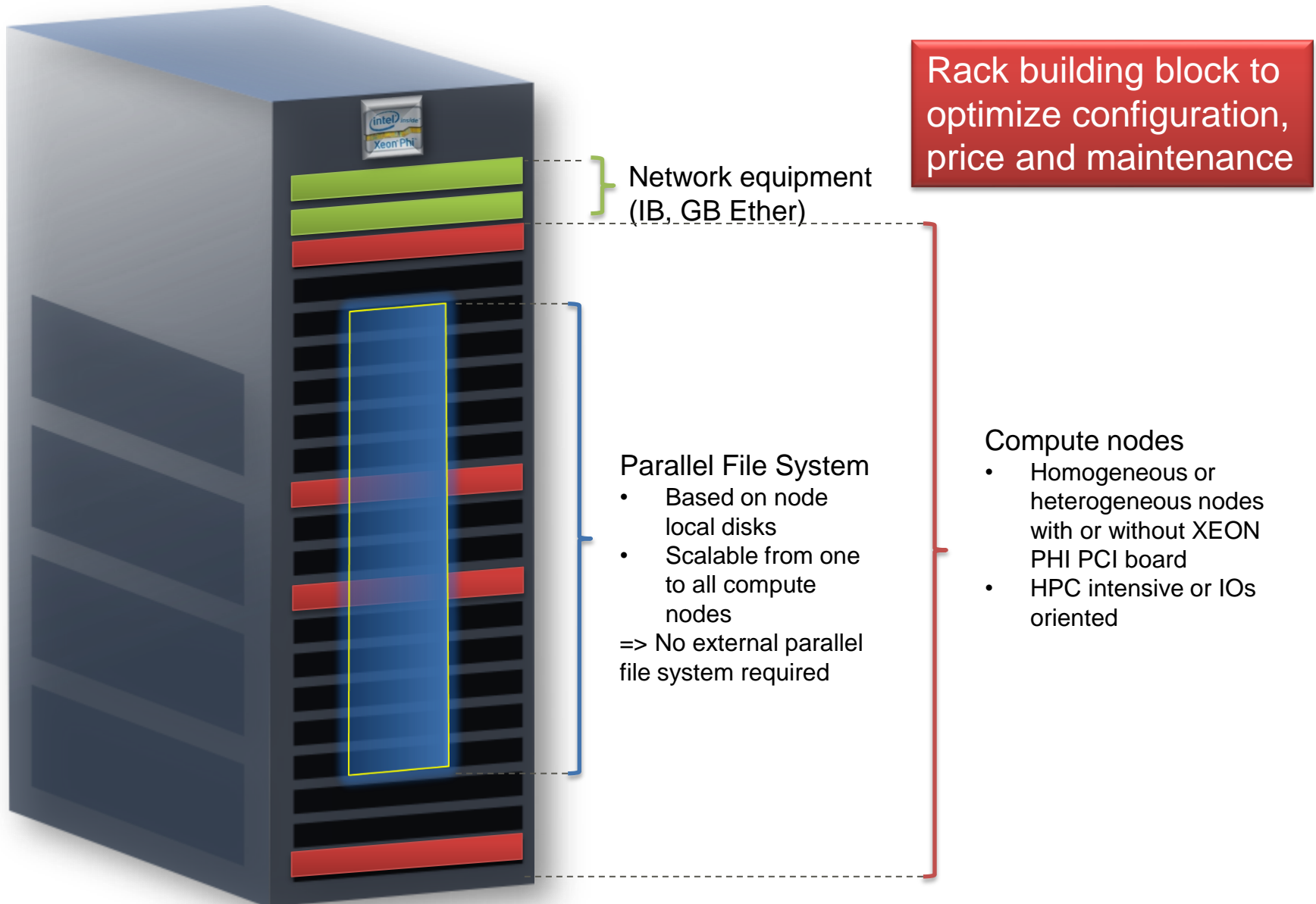
## ■ Compute nodes

- Host: Single or Dual Sandy-Bridge or Ivy-Bridge
- IB/PCI board for IOs and data exchange
- XEON PHI: single PCI board

## ■ Integrated parallel file system

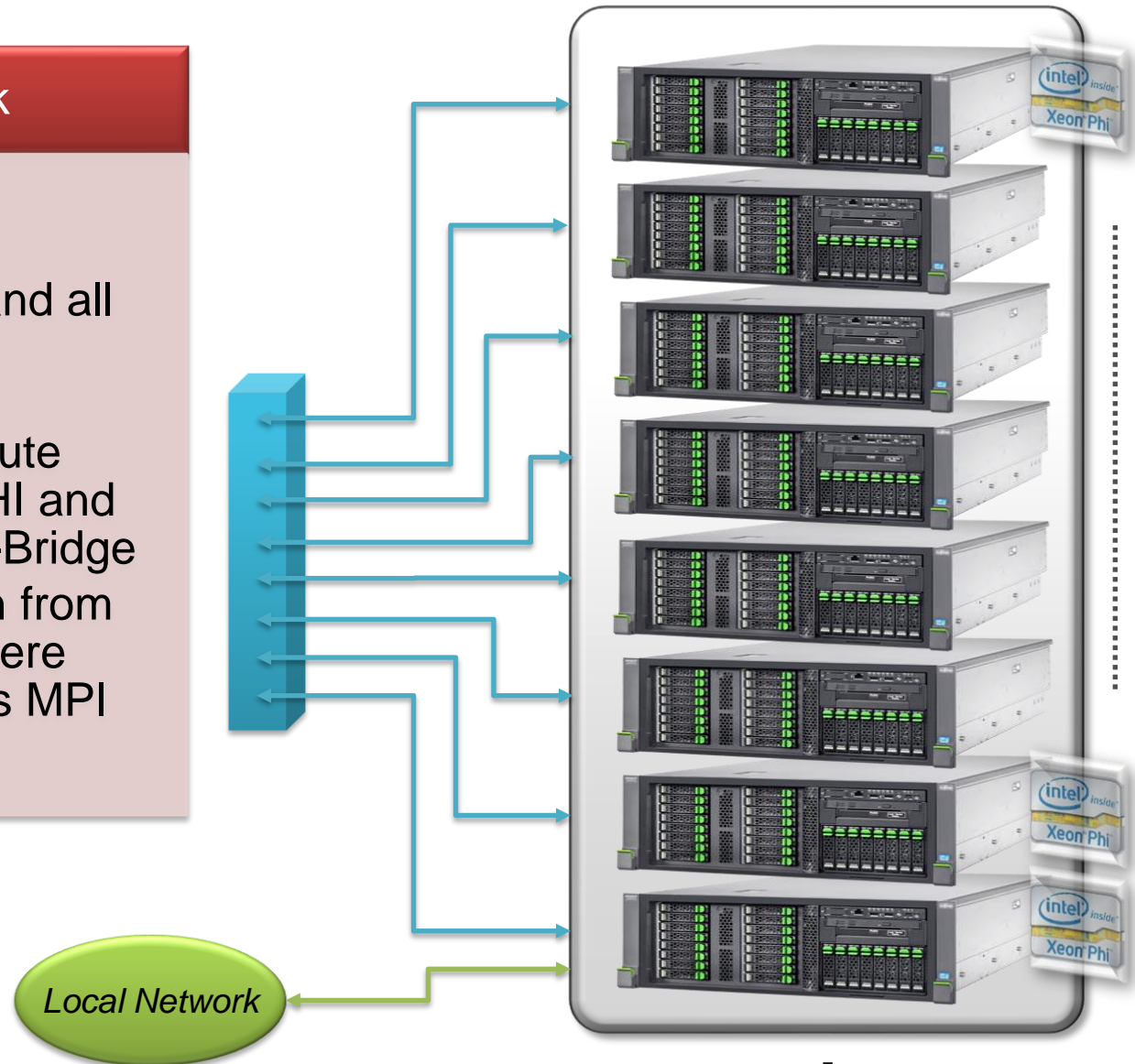
- Using integrated disk from compute nodes
- Direct integration of parallel file system in the rack
- Parallel file system accessible from XEON-PHI

# Fully integrated solution (II)



## Simple Network

- Single switch
- Same shared subnet bridging XEON-PHI and all compute nodes
- Only one cluster of heterogeneous compute nodes, half XEON-PHI and half Dual sockets Ivy-Bridge
- Direct communication from anywhere to everywhere with full heterogeneous MPI support

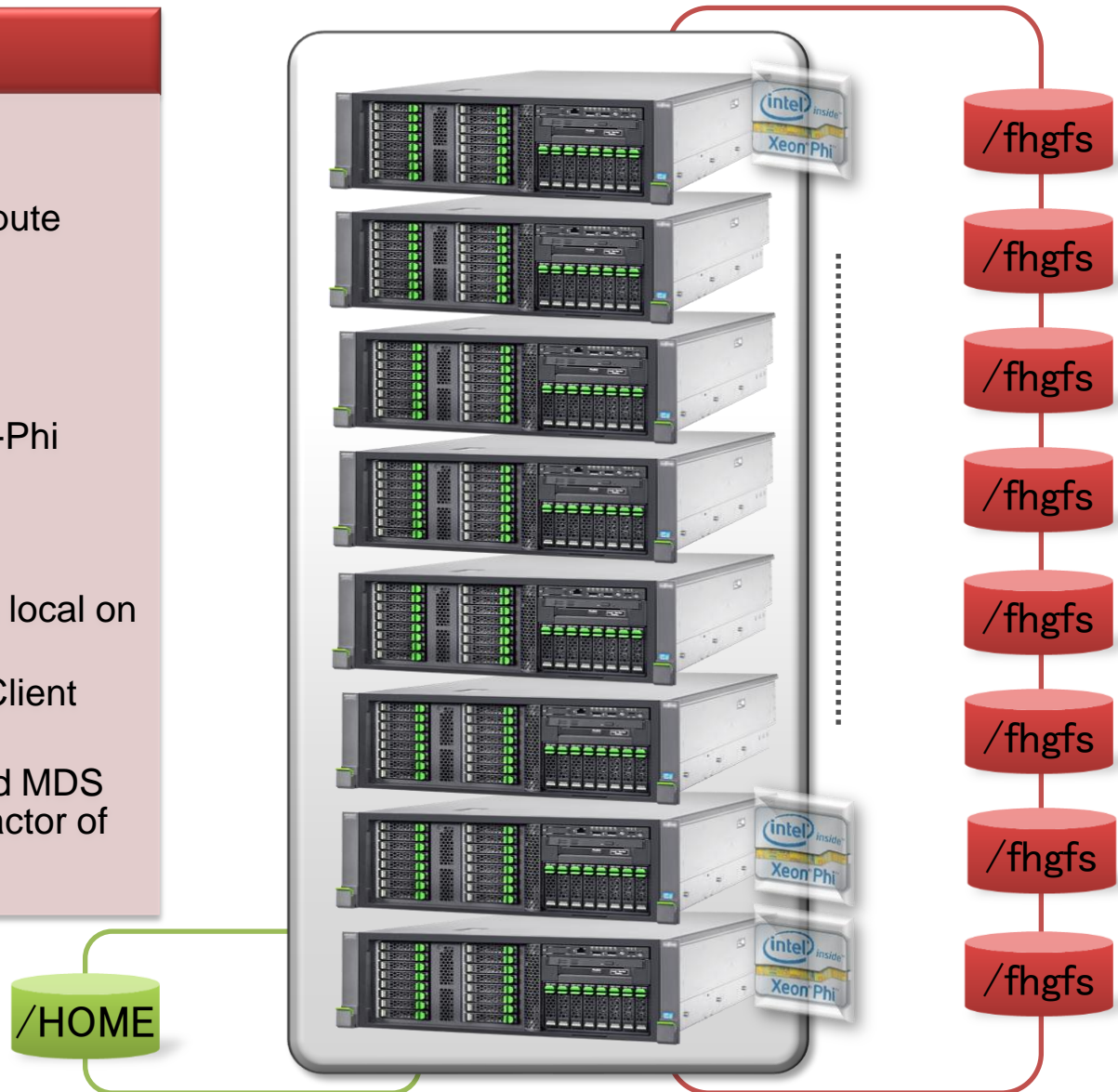




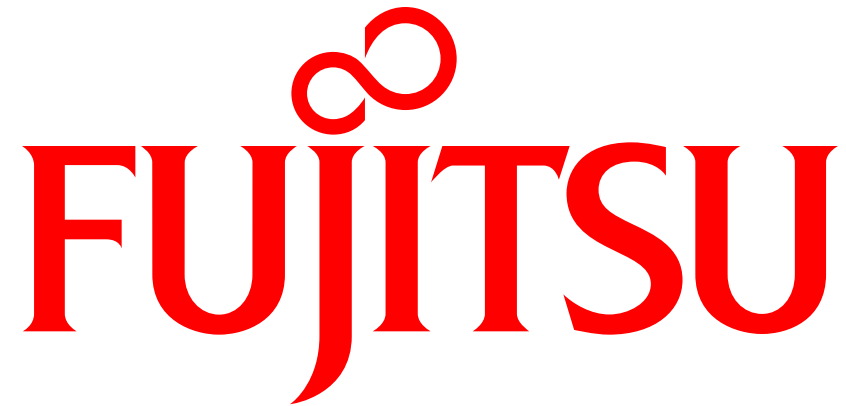
# PHI Cluster: File Systems

## Efficiency Driven IOs

- **HOME file system**
  - On login node
  - NFS mounted on the all compute nodes
- **Local scratch**
  - On each front node
  - Mounted on connected Xeon-Phi node
- **Parallel file system (FHGFS)**
  - Integrated to front nodes with local on board RAID controller
  - Each front node is MDS/DS/Client
  - Each Xeon-Phi node is client
  - Simple policies: local preferred MDS and DS from client, striping factor of 8 over all SB nodes







shaping tomorrow with you