

# ECLIPSE Performance Benchmark and Profiling

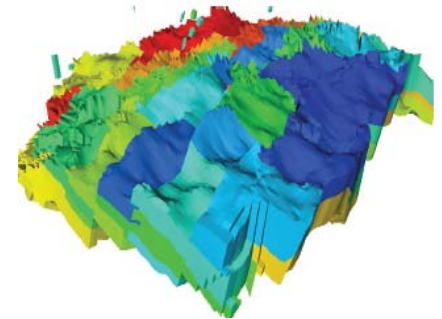
July 2010

**Schlumberger**



- **The following research was performed under the HPC Advisory Council activities**
  - Participating vendors: Schlumberger, HP, Mellanox
  - Compute resource - HPC Advisory Council Cluster Center
  
- **For more info please refer to**
  - [www.mellanox.com](http://www.mellanox.com), <http://www.hp.com/go/hpc>,  
<http://www.slb.com>

- **Oil and gas reservoir simulation software**
  - Developed by Schlumberger
- **Offers multiple choices of numerical simulation techniques for accurate and fast simulation for**
  - Black-oil
  - Compositional
  - Thermal
  - Streamline
  - Others
- **ECLIPSE support MPI to achieve high performance and scalability**

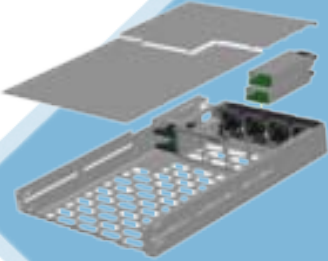


- **The presented research was done to provide best practices**
  - ECLIPSE performance benchmarking
    - Interconnect performance comparisons
  - Ways to increase ECLIPSE productivity
  - Power-efficient simulations
- **The presented results will demonstrate**
  - The scalability of the compute environment
  - Considerations for power saving through balanced system configuration

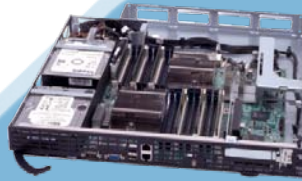
- **HP ProLiant SL2x170z G6 16-node cluster**
  - Six-Core Intel X5670 @ 2.93 GHz CPUs
  - Memory: 24GB per node
  - OS: CentOS5U4, OFED 1.5.1 InfiniBand SW stack
- **Mellanox ConnectX2 InfiniBand adapters and switches**
- **MPI: Platform MPI 5.6.7**
- **Application: Schlumberger ECLIPSE 2009.1**
- **Benchmark Workload**
  - Four million cell model (FOURMILL.DATA)

# About HP ProLiant SL6000 Scalable System

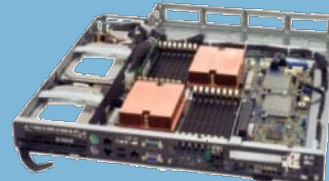
- **Solution-optimized for extreme scale out**



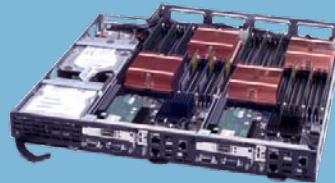
ProLiant z6000 chassis  
Shared infrastructure  
– fans, chassis, power



ProLiant SL160z G6    ProLiant SL165z G7  
Large memory  
-memory-cache apps



ProLiant SL170z G6  
Large storage  
-Web search and database apps



ProLiant SL2x170z G6  
Highly dense  
- HPC compute and  
web front-end apps

Save on cost and energy -- per node, rack and data center

Mix and match configurations

Deploy with confidence

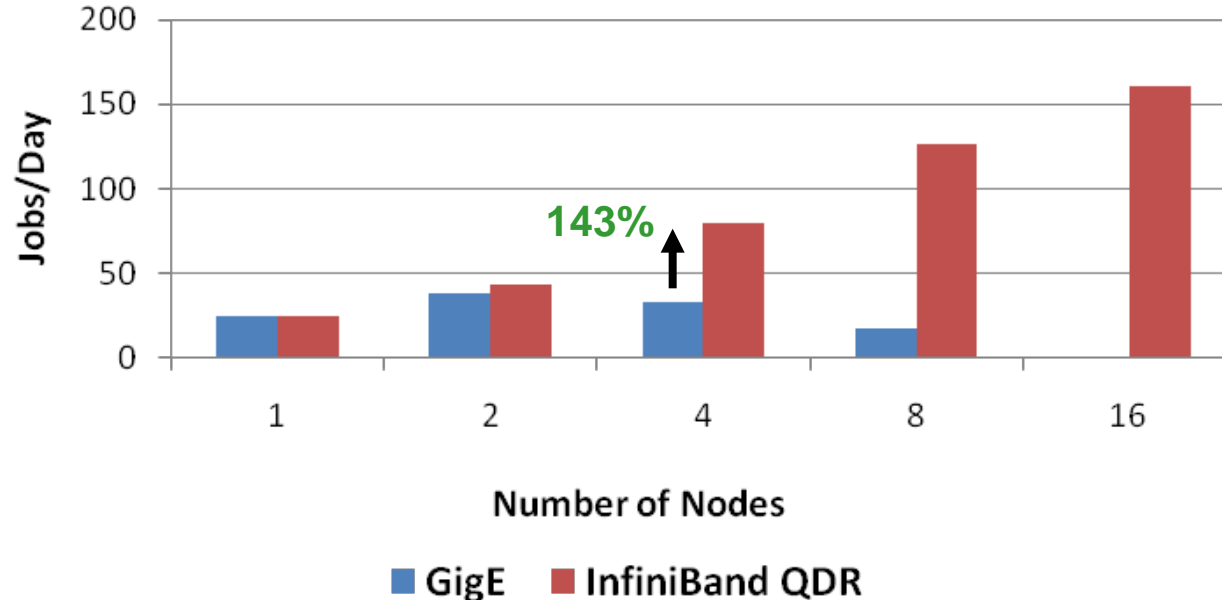


#1  
Power  
Efficiency\*

\* SPECpower\_ssj2008  
[www.spec.org](http://www.spec.org)  
17 June 2010, 13:28

- **Input Dataset: FOURMILL (Four million cell model)**
- **Interconnect performance comparison**
  - InfiniBand reduces electrical energy/job by 60% or more compared to GigE
  - Time to run high priority job is 4x faster with InfiniBand
  - Up to 143% higher performance than GigE at 4 nodes
  - GigE stops scaling beyond 2 nodes

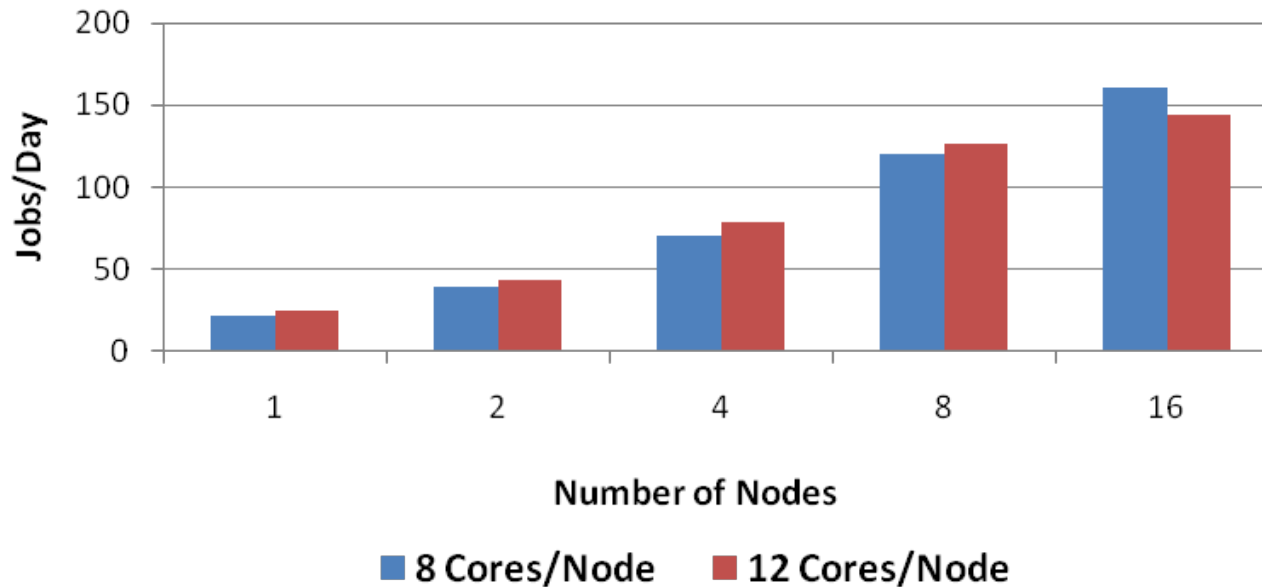
## ECLIPSE Benchmark Result



*Higher is better*

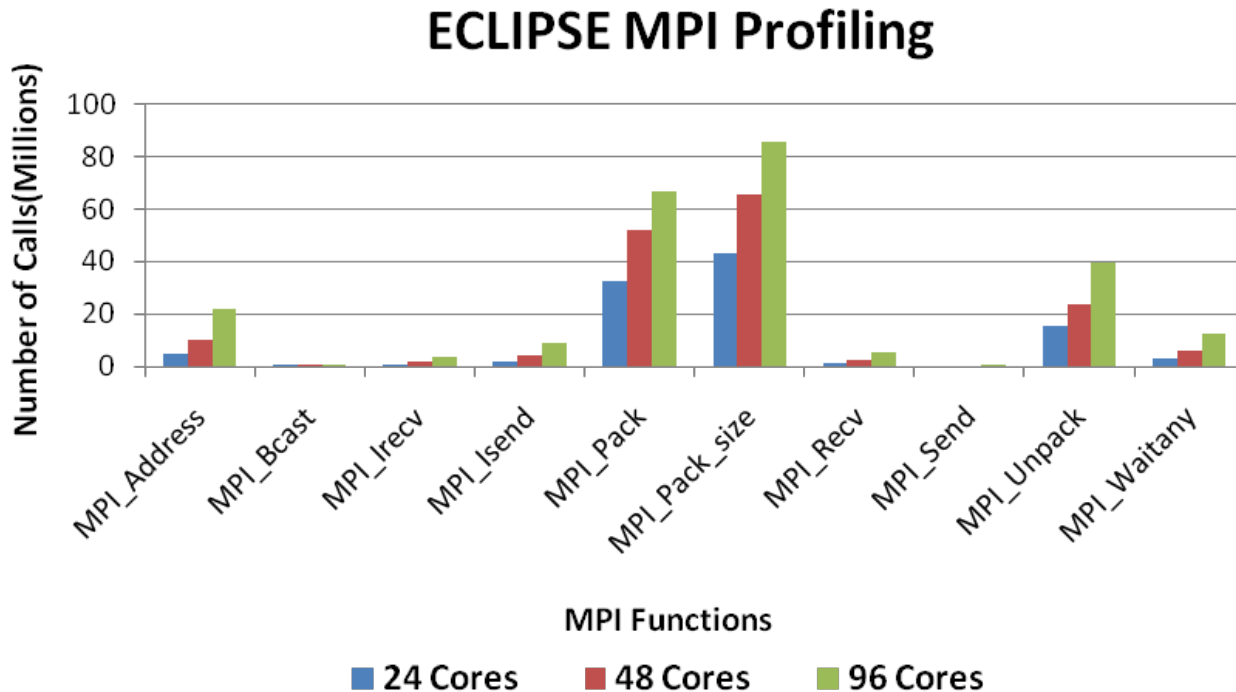
- **Input Dataset: FOURMILL (Four million cell model)**
- **Core/socket comparison**
  - Using 6 cores/socket (12 cores/node) provide higher performance up to 16 nodes
  - At 16 nodes, 4 cores/socket (8 cores/node) provide higher performance
    - Tradeoff between more compute power and higher cost to accomplish parallelism affects application scalability

## ECLIPSE Benchmark Result



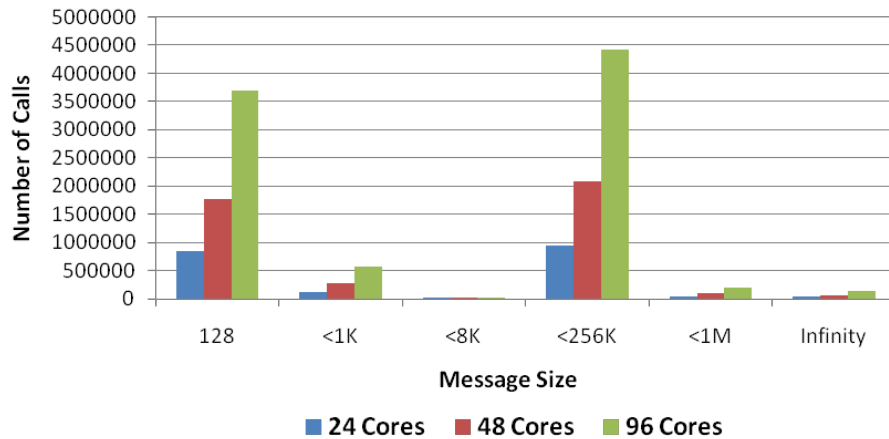
*Higher is better*

- **Majority MPI Functions are point-to-point functions**
  - MPI\_Pack, MPI\_Pack\_Size, and MPI\_Unpack
    - Packs data of a given datatype into contiguous memory
    - All three functions increase as cluster size scales
  - MPI\_Isend and MPI\_Recv

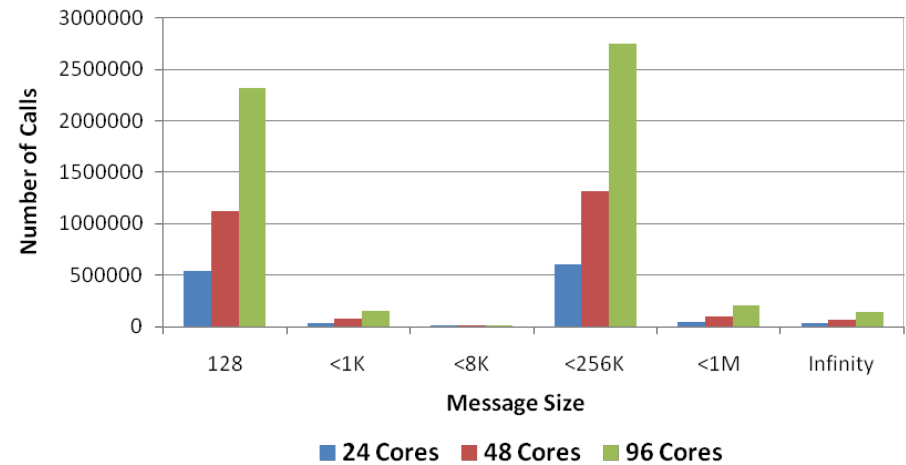


- Both small and large messages are being used

### ECLIPSE MPI Profiling (Isend)



### ECLIPSE MPI Profiling (Recv)



- **ECLIPSE was profiled to identify its communication patterns**
  - MPI Functions
    - MPI\_Pack/Unpack and MPI\_Isend/recv generate most overhead
  - MPI Message Size
    - Both small and large messages are being used
  - Interconnect performance impacts application performance
    - MPI\_Pack/Unpack performance receives big benefit from InfiniBand architecture
    - Both Interconnect latency and bandwidth are critical to ECLIPSE performance
- **InfiniBand enables higher ECLIPSE performance and scalability**
  - Up to 143% higher performance than GigE at 4 nodes
  - GigE stops scaling beyond 2 nodes

# Thank You

## HPC Advisory Council

**Schlumberger**



All trademarks are property of their respective owners. All information is provided "As-Is" without any kind of warranty. The HPC Advisory Council makes no representation to the accuracy and completeness of the information contained herein. HPC Advisory Council Mellanox undertakes no duty and assumes no obligation to update or correct any information presented herein