



# tcc

harnessing boundless capacity across client, cluster & cloud

## PERFORMANCE AND SCALABILITY OF WINDOWS HPC SERVER 2008 R2

Wenhao Wu, Sr. Technical Product Manager, Microsoft Technical computing



# Agenda

- Microsoft' s Vision for HPC
- Windows HPC focuses on performance
- Windows Rivals Linux Performance on key HPC benchmarks
  - ❖ State of the Art Windows HPC Benchmarking System
  - ❖ Windows delivers high performance on standard benchmarks
  - ❖ Windows Rivals Linux on major CAE applications
  - ❖ Windows shows great scalability on Finance applications

# Microsoft's Vision for HPC

"Provide the platform, tools and broad ecosystem to reduce the complexity of HPC by making parallelism more accessible to address future computational needs."

## Reduced Complexity

Ease deployment for larger scale clusters

Simplify management for clusters of all scale

Integrate with existing infrastructure

## Mainstream HPC

Address the performance needs of traditional supercomputing

Address emerging cross-industry computation trends

Enable non-technical users to harness the power of HPC

## Broad Ecosystem

Increase number of parallel applications and codes

Offer choice of parallel development tools, languages and libraries

Drive larger universe of end-users, developers, and system administrators

*Performance!*

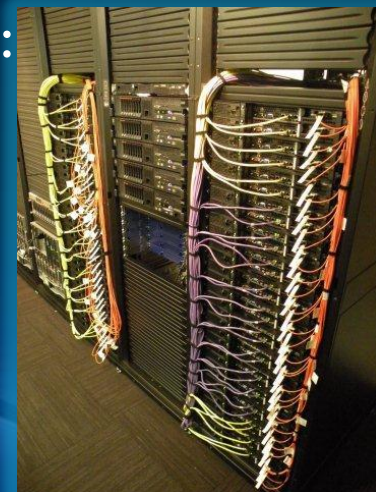
# Windows HPC focuses on performance

Windows Computer Cluster Server 2003	Windows HPC Server 2008	Windows HPC Server 2008 R2
<ul style="list-style-type: none"><li>•Windows Server 2003</li><li>•Scale to 64 nodes</li><li>•MPI over Windows Socket Direct</li><li>•File system: SMB 1 (Server Message Block)</li></ul>	<ul style="list-style-type: none"><li>•Windows Server 2008</li><li>•Scale to 256 nodes and beyond</li><li>•MPI over NetworkDirect (RDMA on Windows) offer comparable performance to native API with multi Fabric support: InfiniBand, iWarp, Myrinet</li><li>•Scalability on a single node goes up to 64 cores</li><li>•SMB 2 improves I/O performance</li></ul>	<ul style="list-style-type: none"><li>•Windows Server 2008 R2</li><li>•Scale to 1000 nodes and beyond</li><li>•Further improvements in MPI shared memory performance</li><li>•Timer coalescing reduces OS overhead</li><li>•Scalability on a single node goes up to 256 cores</li><li>•SMB 2.1 reduce network chattiness</li><li>•User mode threads (aka Light weight threads)</li></ul>

# State of the Art Windows HPC Benchmarking System

## • Rhiannon at Microsoft Enterprise Engineering Center:

- *The IBM System x® iDataPlex™ DX360 cluster*
- Windows HPC Server 2008 R2
- Total 512 cores with 64 compute nodes.
- CPU: Intel® Xeon® Processor X5550, 2.66 GHz
- Memory: 24GB DDR3 RAM (1333MHz) per node.
- Interconnect: Mellanox Connect-X® QDR InfiniBand.
- Storage: IBM GPFS® on 4 IBM x3650 M2 nodes.



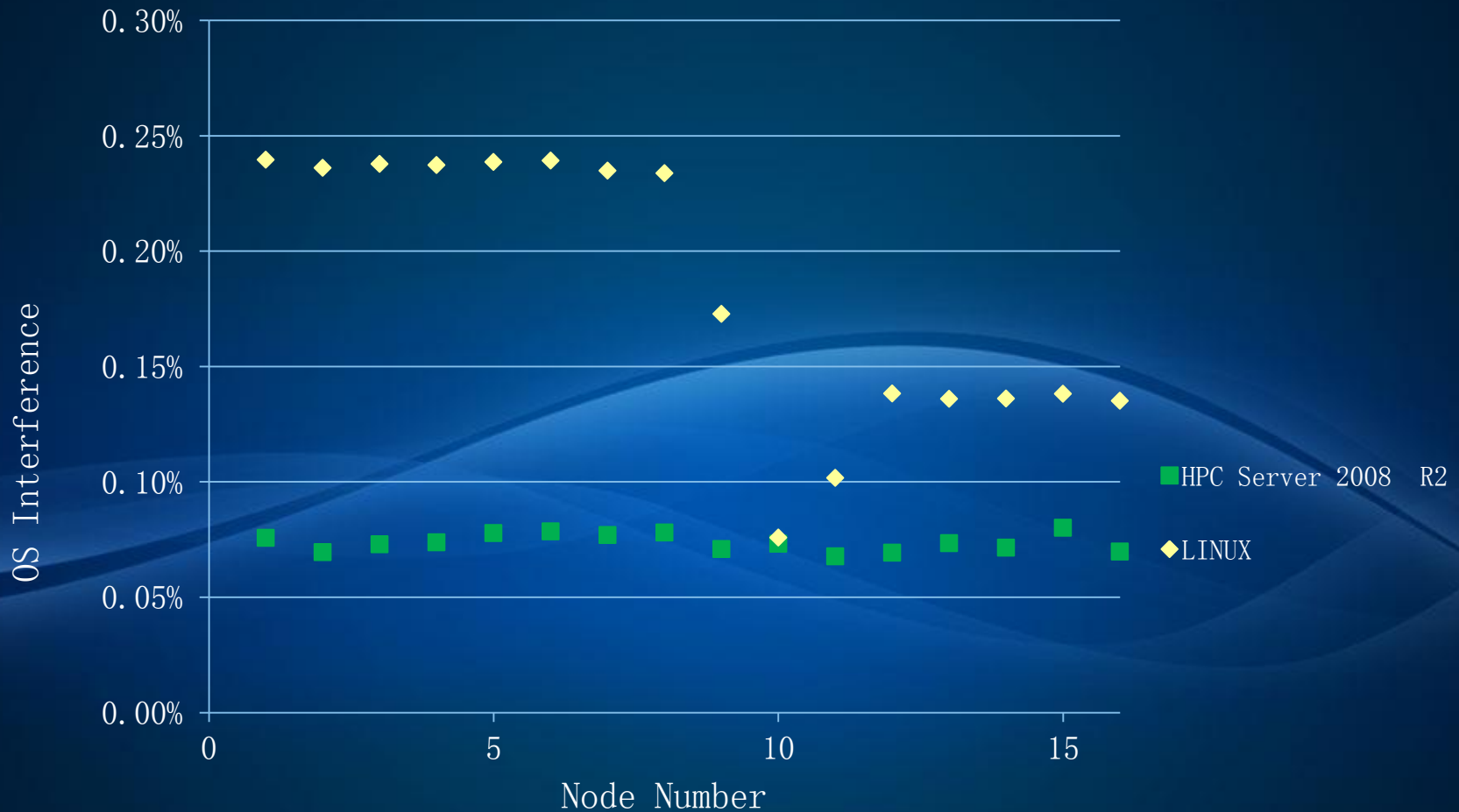
## • WESTMERE at Microsoft Enterprise Engineering Center:

- *The HP Blade System c7000*
- Windows HPC Server 2008 R2
- CPU: Intel X5650 2.67GHz (2 Processors)
- Interconnect: Mellanox QDR IB
- Memory: 24GB RAM

## • Gambrinus at Fraunhofer Germany

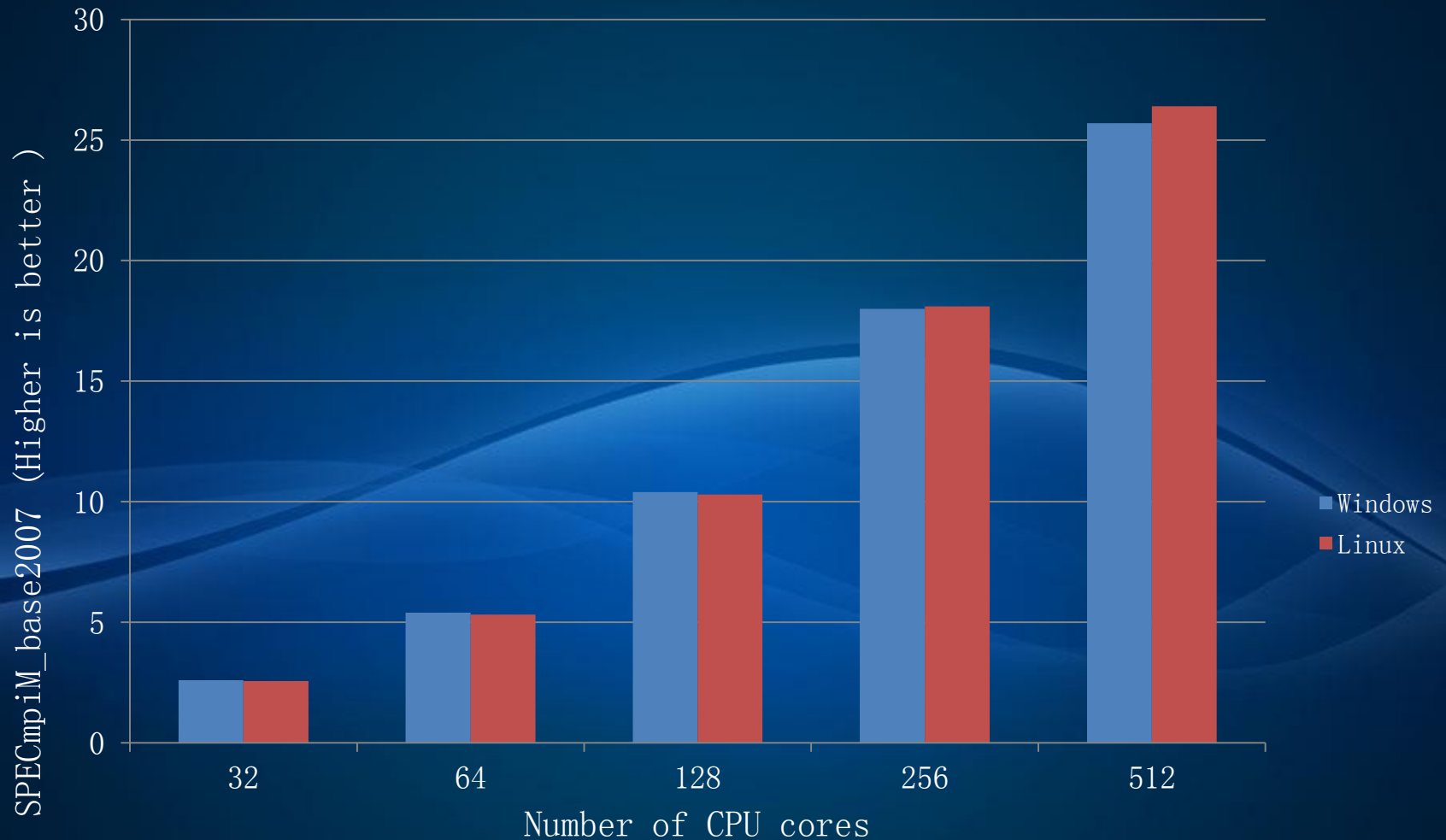
- *16 nodes of Acer Blades*
- Windows HPC Server 2008 R2
- Intel Xeon 5570 and Mellanox QDR InfiniBand





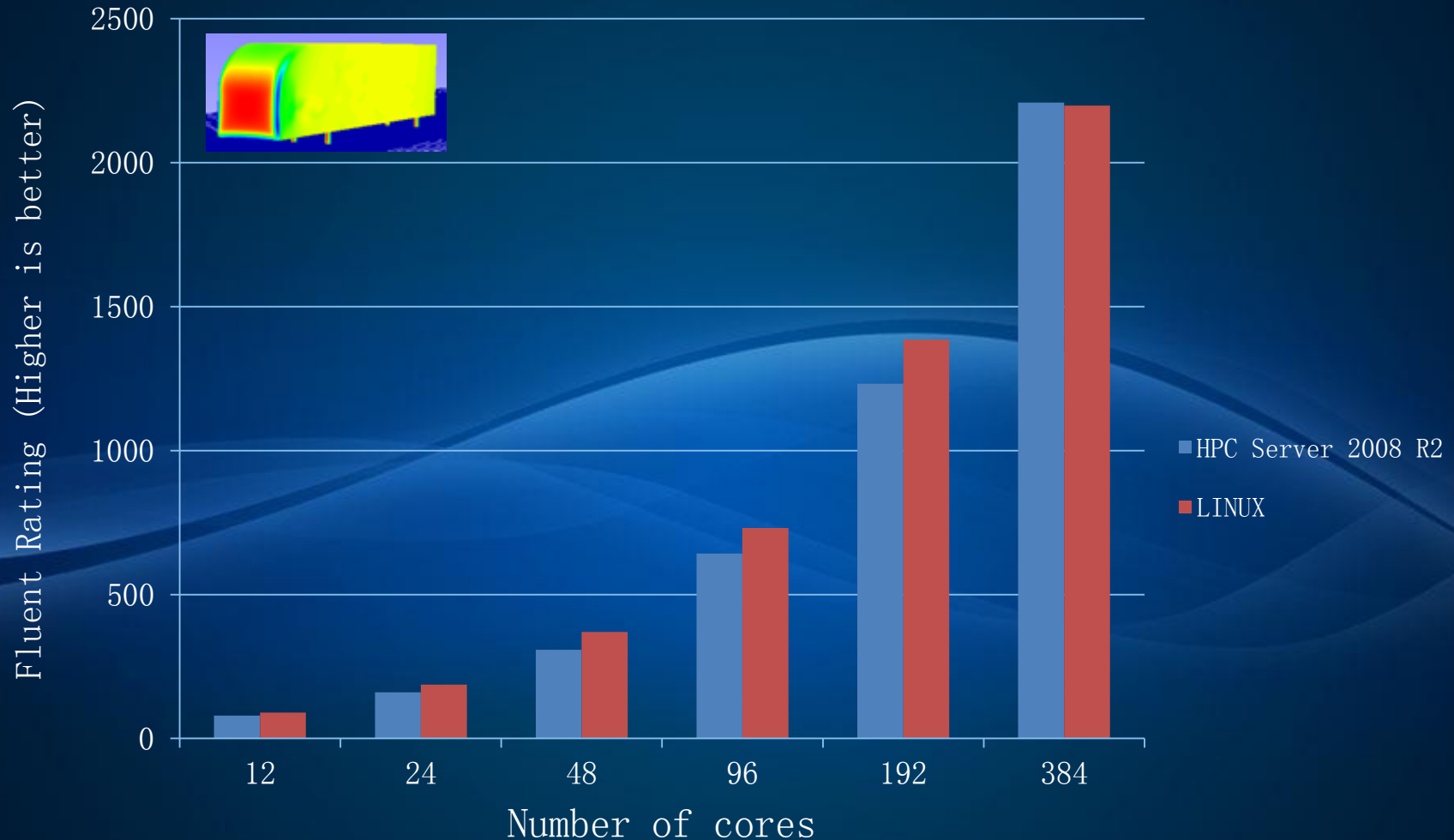
Reference:

- PSNAP 1.2 is an Open Source test to measure OS Interference and is available at the public website: <http://www.ccs3.lanl.gov/PAL/software/psnap/>
- OS Interference in HPC Server 2008 R2 is less than what is seen in a LINUX environment run on an identical HW Configuration.
- Both Windows HPC Server 2008 R2 and LINUX benchmarks were run on the 16 Node Acer Blade System (Intel Nehalem CPU 5570 2.93 GHz (Dual-Processor, Quad Core), 24GB (1333 MHz), Mellanox Infiniband) - GAMBRINUS (Fraunhofer SCAI, Germany).



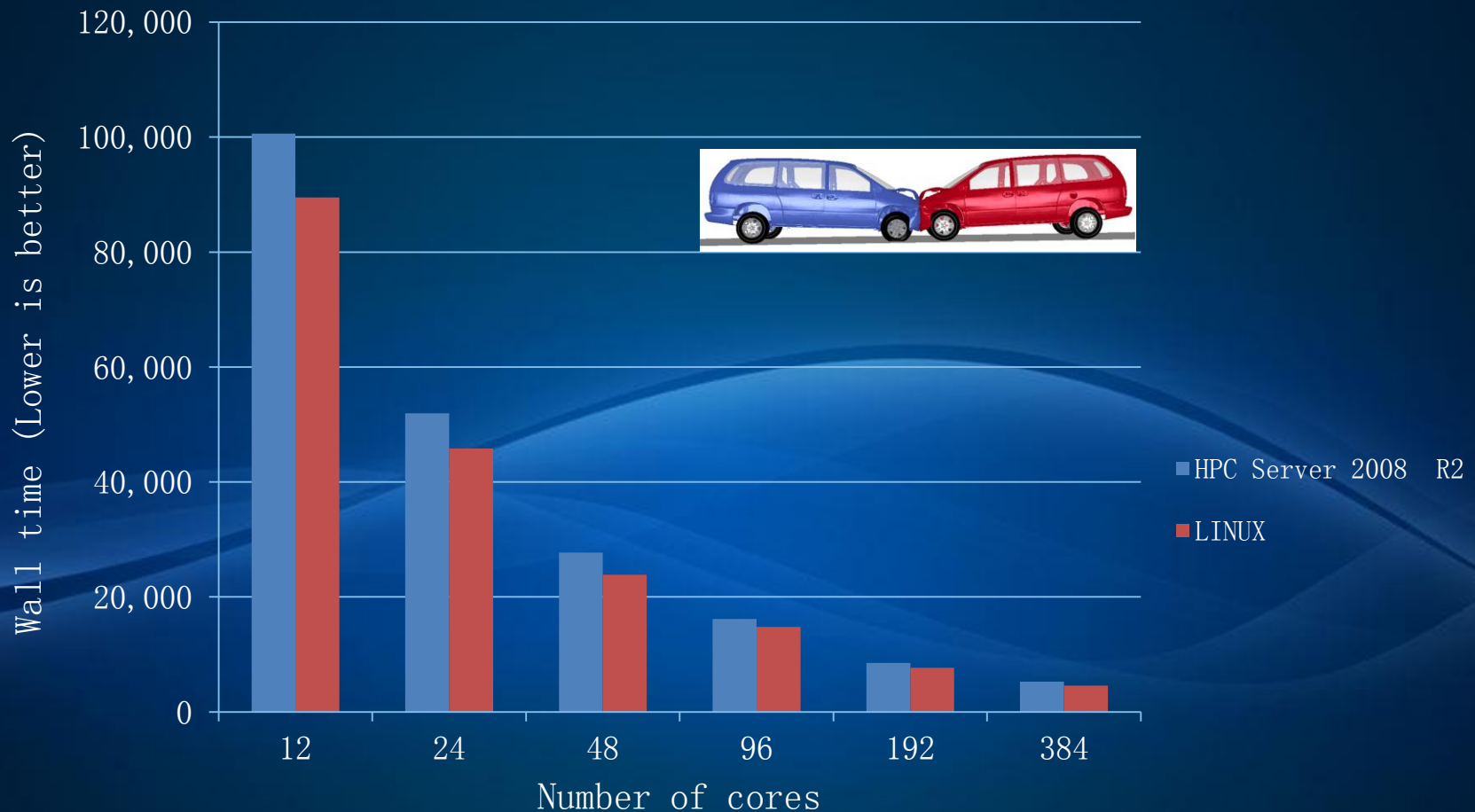
## Reference:

- SPEC MPI2007 version 2.0 was used to benchmark both Windows and LINUX systems.
- Excellent scalability to 256 cores.
- Comparable Performance for HPC Server 2008 R2 and LINUX.
- Both WINDOWS and LINUX benchmarks were run on IBM System x iDataPlex dx340 with 2.5GHz Intel® Xeon® Processor L5420 at Indiana University.
- Both WINDOWS and LINUX data was obtained from the public website:
- <http://www.spec.org/mpi2007/results/res2010q1/#SPECmpiM>.



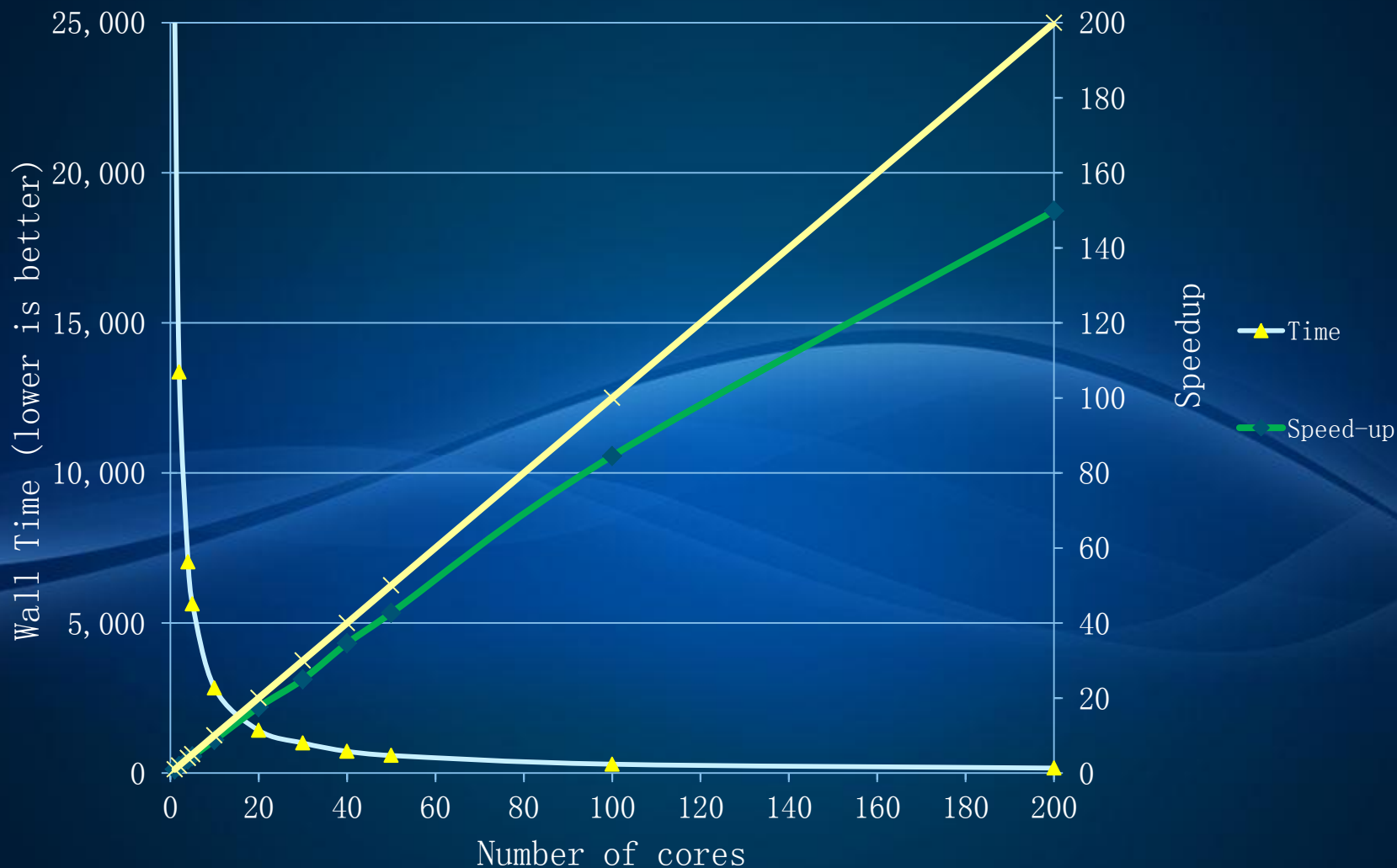
## Reference:

- FLUENT version 12.1. Dataset is ANSYS FLUENT truck\_14m, 14M cells, external aero, public benchmark.
- Near linear scaling to 384 cores.
- Comparable Performance for HPC Server 2008 R2 and LINUX.
- WINDOWS benchmarks were run on HP Blade System c7000™ with 2.66GHz Intel® Xeon® Processor 5650.
- LINUX benchmarks were run on SGI® (ALTIX XE1300C with 2.93GHz Intel Xeon Processor 5670, SLES11, IB).
- LINUX data was obtained from the public website: [http://www.fluent.com/software/fluent/fl6bench/fl6bench\\_12.0/](http://www.fluent.com/software/fluent/fl6bench/fl6bench_12.0/).



## Reference:

- LS-DYNA version 971 R3.21. Dataset is LS-DYNA car2car, public benchmark available at [www.topcrunch.org](http://www.topcrunch.org).
- Excellent scalability to 192 cores.
- Comparable Performance for HPC Server 2008 R2 and LINUX
- WINDOWS benchmarks were run on HP Blade System c7000™ with 2.66GHz Intel® Xeon® Processor 5650, Mellanox IB
- LINUX benchmarks were run on Intel® SR1600UR system with 2.93GHz Intel Xeon Processor 5670, IB
- LINUX data was obtained from the public website: <http://www.topcrunch.org>.



Reference:

- Dataset is Test Case B an MG-ALFA benchmark.
- MG-ALFA achieves near linear scalability using test case B.
- MG-ALFA version 7.1 was used on HPC Server 2008 R2.
- WINDOWS benchmarks were run on IBM System x® iDataPlex, with 2.66GHz Intel® Xeon® Processor 5550.

# More Resources

Windows HPC Server 2008 R2 Suite

<http://www.microsoft.com/hpc/en/us/default.aspx>

Windows HPC Server 2008 Benchmarks

<http://www.microsoft.com/hpc/en/us/product/performance-benchmarking-hpc.aspx>

Windows Performance Analysis Development Center

<http://msdn.microsoft.com/en-us/performance/default.aspx>

Microsoft Windows Performance Analyzer - Quick Start and Overview

<http://msdn.microsoft.com/en-us/library/ff190971%28VS.85%29.aspx>