HPC Strategy and direction for Meteorological Modelling

Hans Joraandstad

Hans.Joraandstad@Sun.COM

HPC – A Sun Priority

"HPC represents an enormous opportunity for Sun and Sun's partners. We have products today as well as on our future roadmap which are uniquely positioned to gain market share in HPC. I am personally leading the cross-Sun team to grow our position in this opportunity area and am looking forward to rapid success."

John Fowler - Executive Vice President

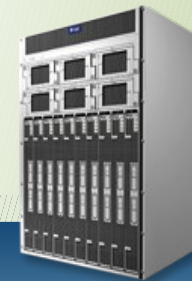# Sun's HPC Technology Strategy

Standards based components

Power, Space, Performance Balance, Efficiency, Density

Open Source, Integratable

Low risk solutions

Sun Fire Systems for Every Application

Delivering Real-World Application Performance

Scale Up

- Large databases
- Enterprise apps—
  CRM, ERP, SCM
- Data warehousing,
  business intelligence
- Server consolidation/
  migration
- Mainframe rehosting

Scale Out

- Web services, mail, messaging, security, firewall
- Applications server,database, ERP, CRM
- HPC, Compute Grid solutions
- Network-facing, I/O intensive
- Load balancing, business logic
- Distributed databases
- Server consolidation

Your Choice of Operating Systems

Sun Joins Open MPI

Sun has joined Open MPI with the firm belief that HPC ultra-scale MPI computing requirements are best met through a concerted and collaborative effort for the good of the community as a whole.

- Open MPI brings together world-class expertise to deliver ultra-scale MPI capabilities
- Sun brings nine years of MPI implementation experience and expertise to the community
- Sun engineers will participate as active developers of Open MPI
- Sun will ship and support Open MPI for Solaris x64 and SPARC platforms

Innovation at Sun

System

Innovate at the system level with industry-standard chip.

AMD 64 Opteron

Chip

Innovate at the system AND the chip level

ULTRASPARC

Sun and US Government Partnership
Basic Computer Systems Research

2002	**1 Year**	I: Industry Concept Study	**$ ~3M**	IBM sgi CRAY hp Sun microsystems
	3 Years	II: Research	**$ ~50M** (July 2003)	IBM CRAY Sun microsystems
2010	**4 Years**	III: Product Development	**$~250M** (Mid 2006)	**???** 1 or 2 to Be Selected

DARPA's "Ultrascale" Goals

2+ PetaFlops on Linpack Top500 benchmark
- Solves large linear system using Gaussian Elimination

6.5 PetaBytes/sec data streams bandwidth
- Copy, add, scale large vectors locally

3.2 PetaBytes/sec bisection bandwidth
- Transpose a large matrix

64,000 GigaUpdate /sec
- Random updates to memory

*"May require up to 100,000 processor cores,
but must be easier to use than current 1,000 processor systems."*

From Strategy to Products and Solutions

- ## Fast
  - > Sun's fast HPC solutions allow customers to solve more complex problems, run more simulations, perform intricate business analysis, and bring new products to market quicker than their competition.

- ## Affordable & Efficient
  - > Sun HPC solutions increase system utilization, lower TCO and reduce energy utilization/minimize power and cooling costs, resulting in big savings to a customers bottom line.

- ## Open & Scalable
  - > Giving customers a choice in running Solaris, Linux or Windows enables customers to choose the best solution and components for their infrastructure, and to scale when needed simply and cost effectively.

- ## Easy to Manage & Deploy
  - > All Sun HPC solutions include sophisticated management and workload software that allows customers to manage very large clusters consisting of thousands of compute nodes from a single console. The Sun Customer Ready Systems program offers custom built and tested solutions, such as the Sun Grid Rack System, delivered directly from Sun's factory to speed and simplify deployment.

Adding to the Galaxy Capabilities ...

More than
45 world records...
38.18 TFlops!!!

Sun x64
Servers

Fast. HPC Is More Than Computations
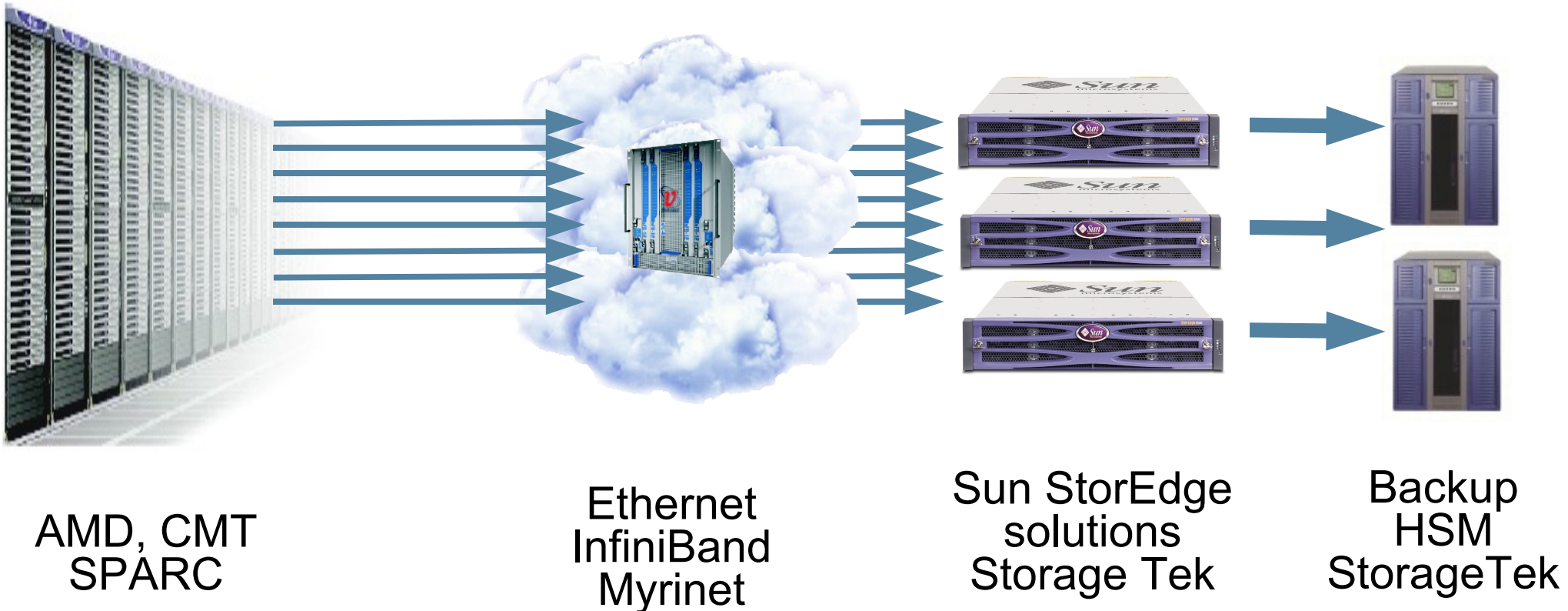# Balance Across the Workflow is Critical

Grid Software: compile, run, manage, store, transfer, provision, visualise, ...

Cluster Nodes	Interconnect	Storage	Tape

AMD, CMT
SPARC

Ethernet
InfiniBand
Myrinet

Sun StorEdge
solutions
Storage Tek

Backup
HSM
StorageTek

A Complete HPC Portfolio From Sun

Sun CRS, Support, Architectural, Professional Services

Applications

Custom or ISV Applications

Sun HPC Cluster Tools/Development Tools — *Open, Free*

Management
Workload Management
Cluster Management

Sun N1™ Grid Engine Software — *Open, Free*

Sun N1™ System Manager Software — *Open, Free*

Operating System

redhat. SUSE solaris — *Open, Free*

Node
Processor

64 Bit AMD 64 Opteron ULTRASPARC

Interconnect

Gigabit Ethernet, Myrinet, Infiniband

Most Complete, Scalable Server Line

Sun Fire™ x64 and SPARC® Server Family

Sun Fire x64 Servers (1-16 way)

- Sun Fire X2100 Single/Dual Core
- Sun Fire X4100 Single/Dual Core
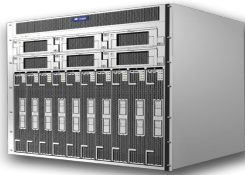- Sun Fire X4200 Single/Dual Core
- Sun Fire V40z Single/Dual Core
- Sun Fire X4600 4 way/Dual Core
- Sun Grid Rack Systems

Other x64 Sun Fire Servers (Blades & Data Storage)

- Sun Fire Blade 8000 Modular System
- Sun Fire X4500
- **New Sun Fire CoolThreads™ Servers** (up to 32 threads)
- Sun Fire T1000
- Sun Fire T2000

Sun Fire UltraSPARC IV+ Servers (4-72 processors)

- Sun Fire V490
- Sun Fire V890
- Sun Fire E2900
- Sun Fire E4900
- Sun Fire E6900
- Sun Fire E20K/E25K

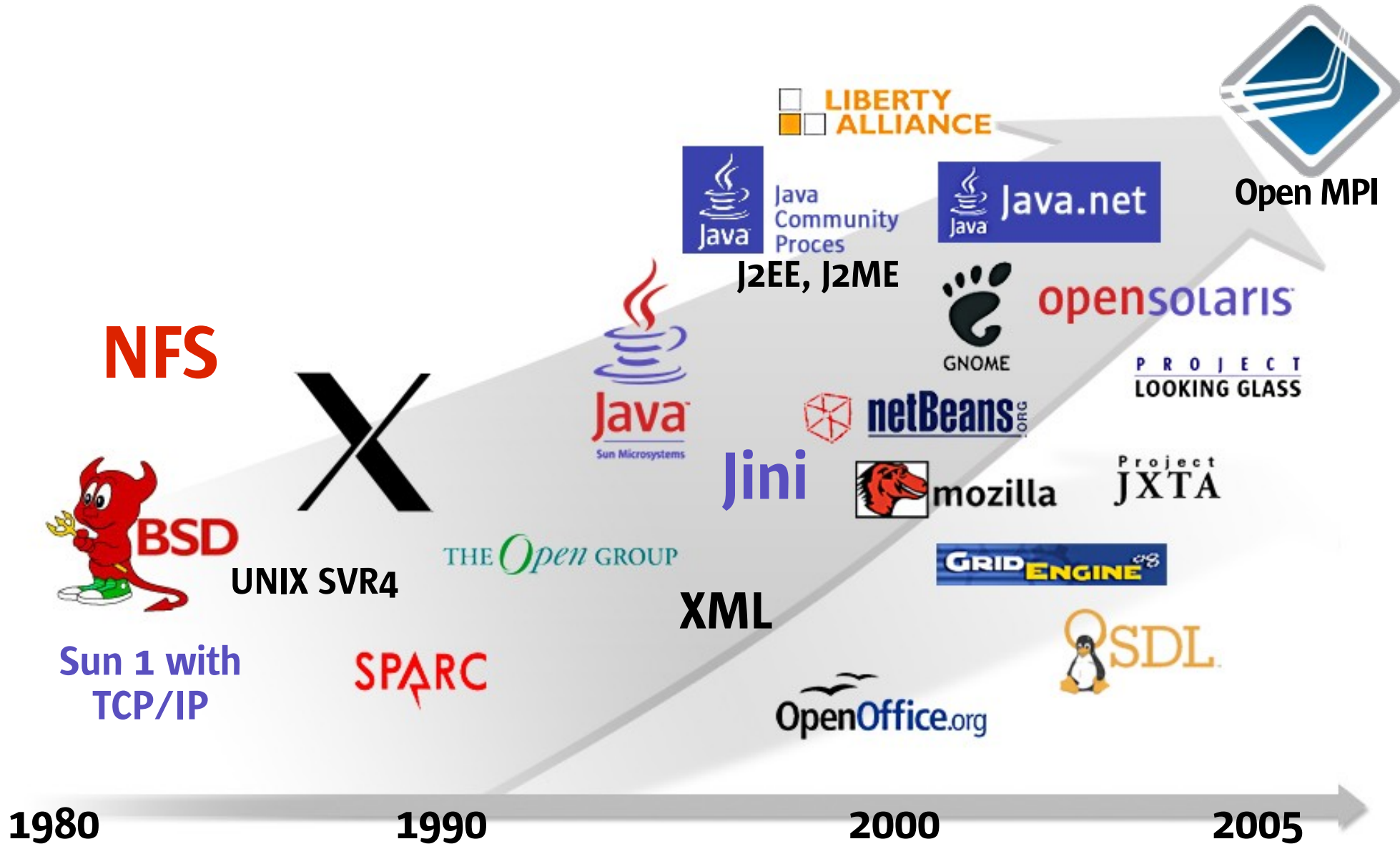**Sun Fire UltraSPARC IIIi Servers** (1-4 processors)
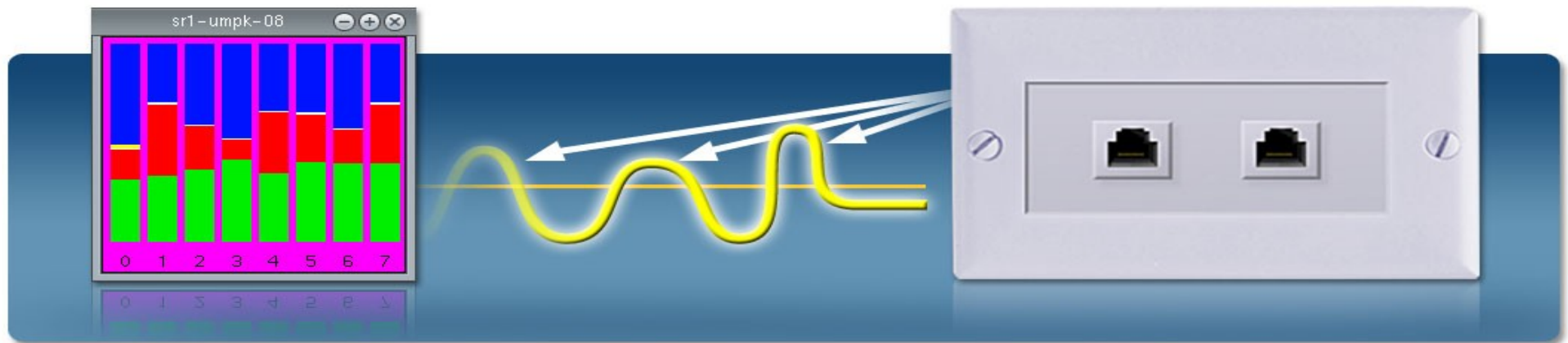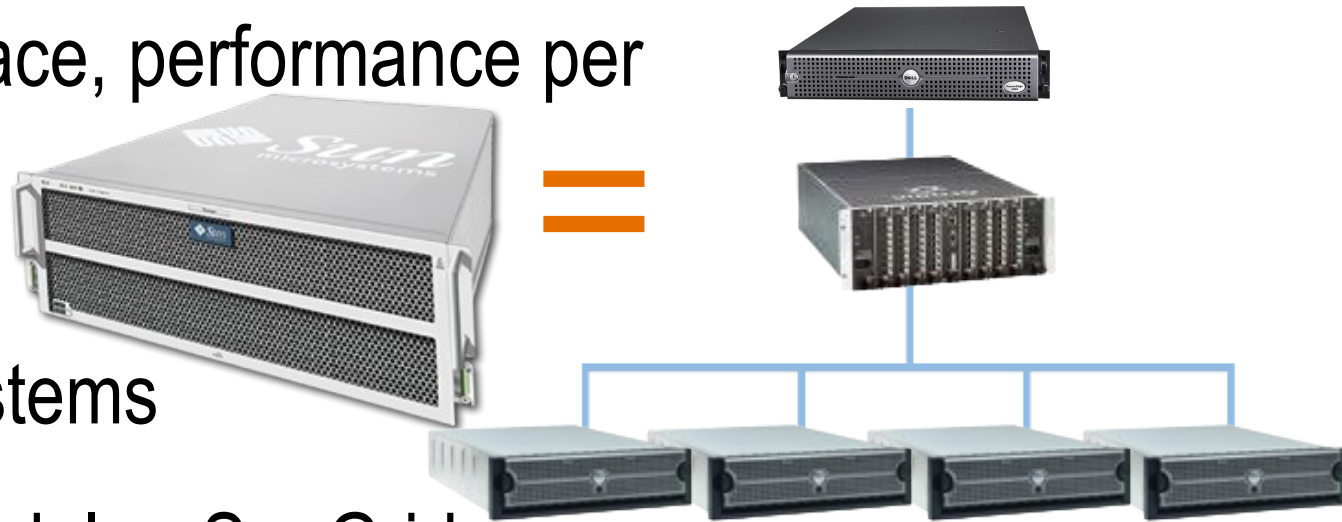
- Sun Fire V210
- Sun Fire V240
- Sun Fire V440

Affordable and Efficient.

- Less power, less space, performance per watt

- Higher densities

- General purpose systems
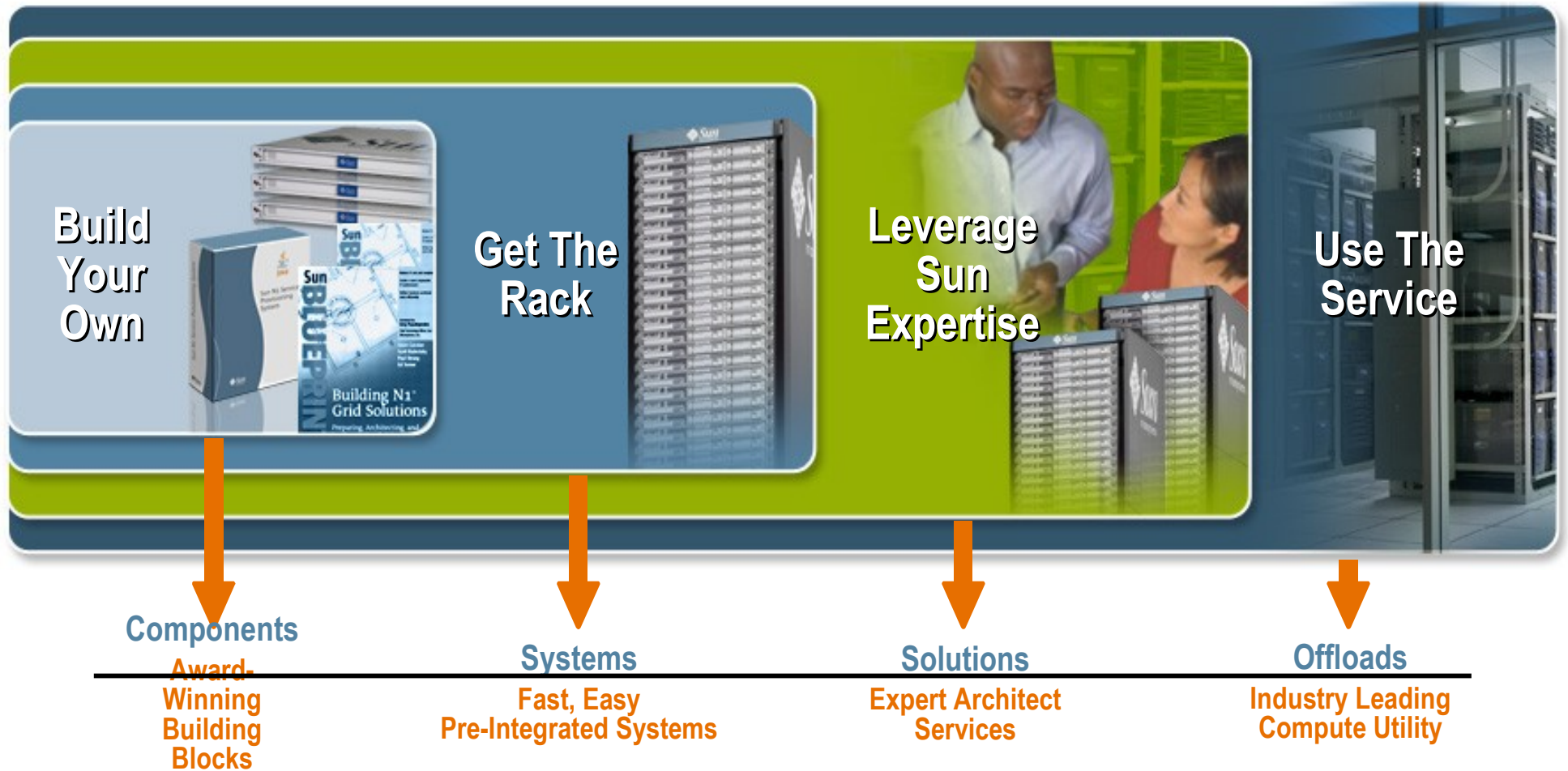
- Utility Computing Model or SunGrid

  $/CPU/Hour

Sun Fire X4500 Data Server

NEW

- ✓ Integrated Compute and Storage Server
- ✓ 24TB in 4RU
- ✓ $2/GB

Sun Gives You The Power of Choice and Flexibility

Build Your Own — **Get The Rack** — **Leverage Sun Expertise** — **Use The Service**

Components	Systems	Solutions	Offloads
Award-Winning Building Blocks	Fast, Easy Pre-Integrated Systems	Expert Architect Services	Industry Leading Compute Utility

Sun's ability to innovate with both technology and business models provides a unique advantage to our customers

Large HPC Clusters Made Faster and Simple

Only Sun Can Do This

Tokyo Institute of Technology with Sun Fire X4600 Server

7th fastest Super-Computer in the World

- Others would need 1,300+ servers
- Sun delivered 650+ servers
- 4 weeks to 37 Tflops vs 3-months
- Customer much better off:
  - > Quicker bring-up of grid
  - > Reduced network complexity
  - > Reduced costs on switches and network infrastructure
  - > Reduced TCO

655 Sun Fire x4600's and 42 Sun Fire x4500's deployed at Tokyo Tech in Tokyo, Japan, deployed March 2006

85 TFlops. IB network, 10,000 cores, 21 TB memory, 1 PB storage

Project Blackbox:
The Virtualized Datacenter
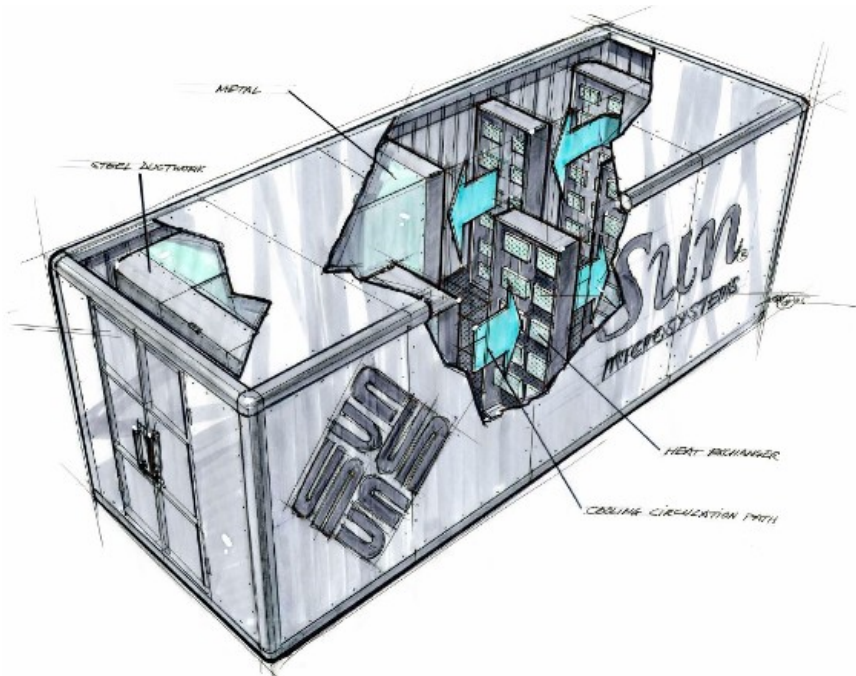
- Standard shipping container packaged with eight standard racks
- Integrated, high-efficiency power and cooling
- Supports a wide range of compute, storage and network infrastructure – build once, deploy anywhere when fully configured

Tsubame: By the Numbers

- 655 16-core Compute Nodes
- 10480 Opteron Cores
- 21.4 TB main memory
- 1.1 PB Storage
- 8 Voltaire ISR9288 InfiniBand Switches
- 360 Clearspeed Advance Boards

Running Tsubame

- N1GE and N1SM on Solaris management nodes
- Sun Studio 11, Pathscale, & PGI compilers
- SuSe 9 SP3, Solaris 10*
- Voltaire MPI with Dual-Rail HCAs
- Lustre, ZFS file systems
- Everybody's Supercomputer, 10,000+ users
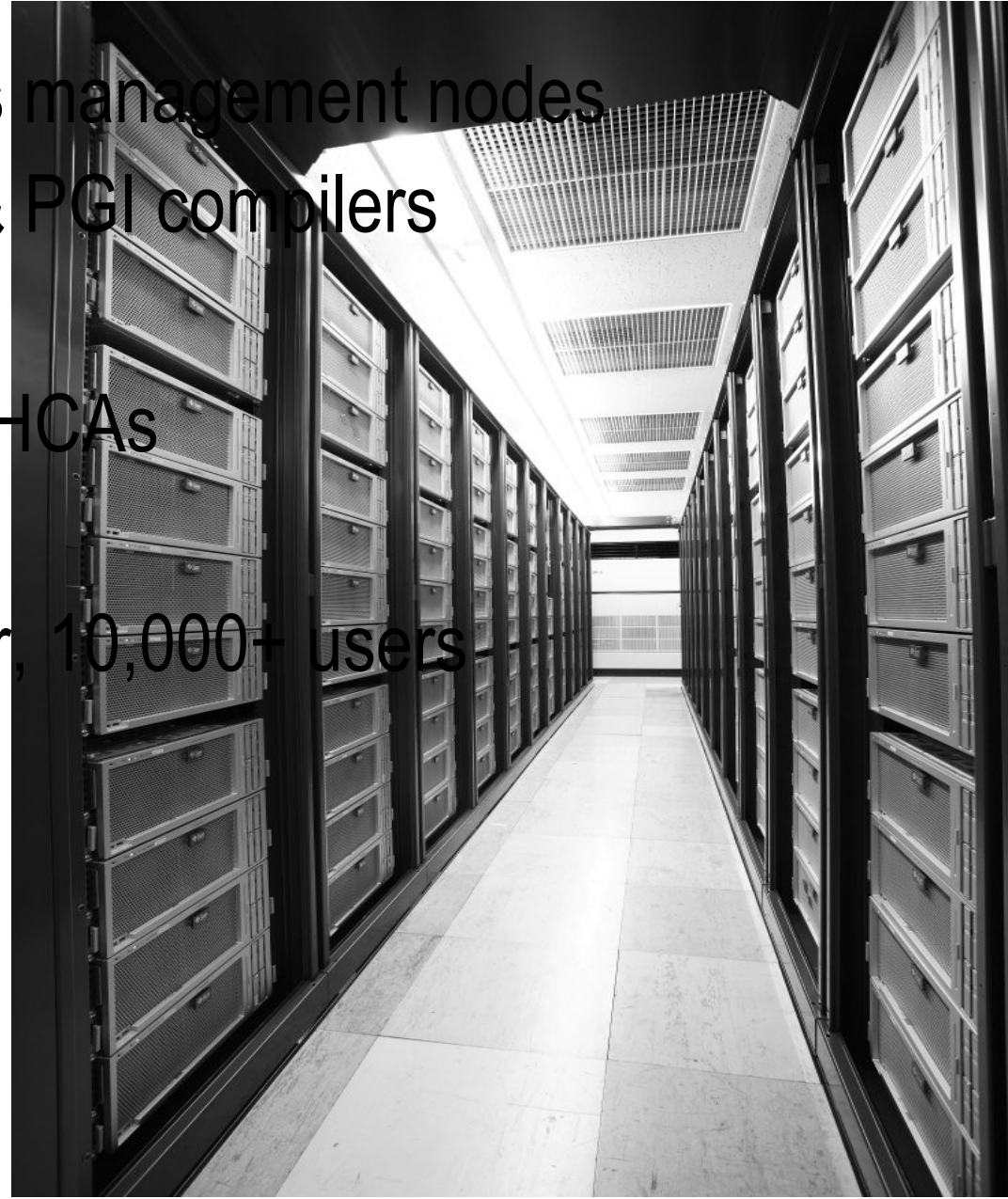
Recent news - TACC

National Science Foundation Awards Texas Advanced Computing Center $59 Million for High-Performance Computing

TACC is partnering with **Sun Microsystems** to deploy a supercomputer system specifically developed to support very large science and engineering computing requirements. In its final configuration in 2007, the supercomputer will have a peak performance in excess of **420 trillion floating point operations** per second (teraflops), making it one of the most powerful supercomputer systems in the world. It will also provide over **100 trillion bytes (terabytes) of memory** and 1.7 quadrillion bytes (petabytes) of disk storage. The system is based on Sun Fire(TM) x64 (x86, 64-bit) servers and Sun StorageTek(TM) disk and tape storage technologies, and will use AMD's forthcoming **quad-core** processors.

Sun Solution Center for HPC
Located in Hillsboro, Oregon, USA

- Over 10 Teraflops deployed
  - > More than 600 x64 and UltraSPARC nodes
  - > Continually refreshed (located next to the factory)
  - > Built with Sun Grid Rack Systems

- Available for:
  - > Proofs of Concept
  - > Benchmarks
  - > Scalability testing
  - > Risk mitigation

- Leverages years of HPC expertise from specialists all over Sun
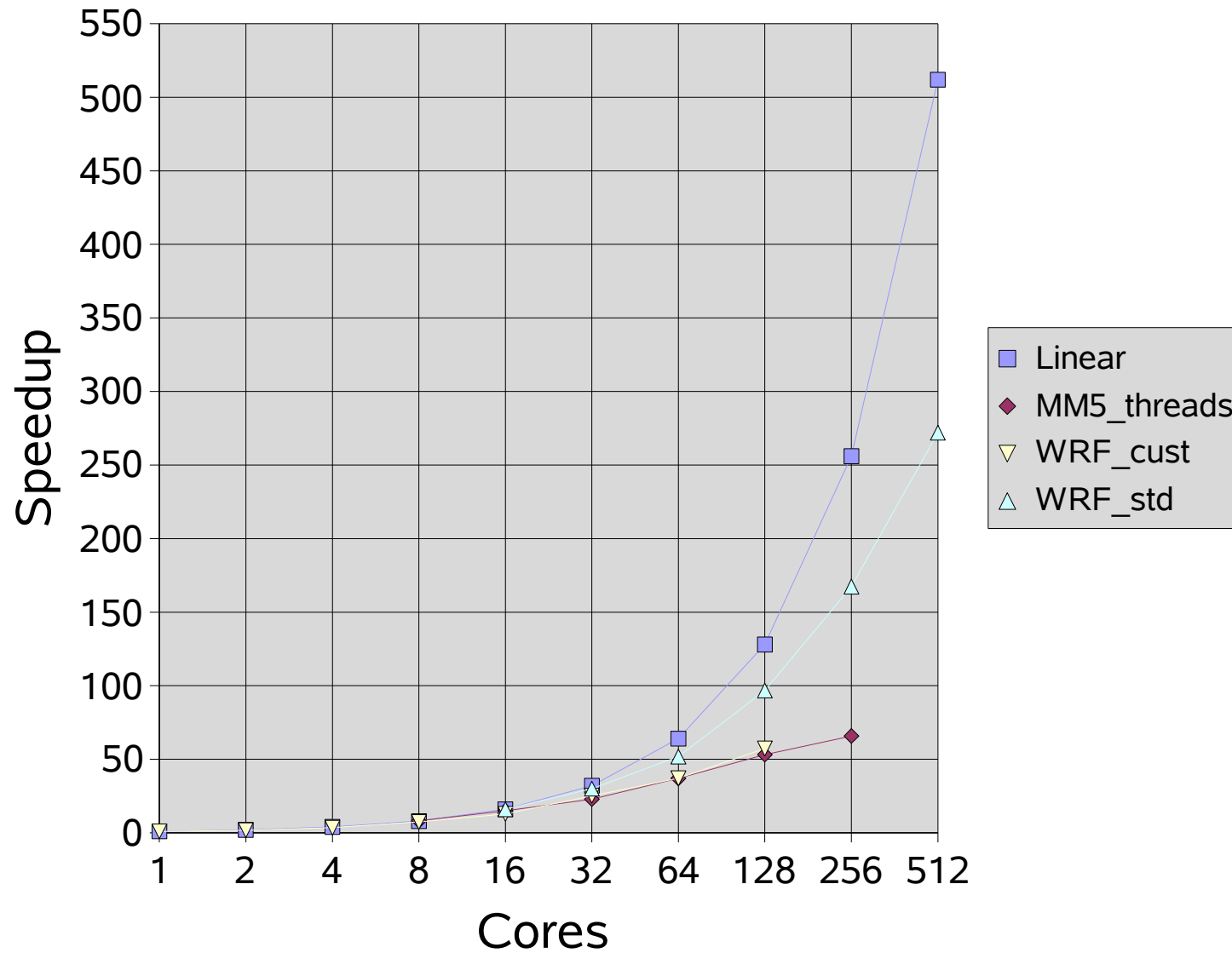
Speedup of meteorological codes

- Codes are MM5, WRF, Hirlam, Aladin and LM RAPS
- Different everything!
  - > Different nodes: 2-16 cores
  - > Different compilers: PGI, PathScale, Studio
  - > Different model sizes, time steps, forecast length
  - > Different versions (even for same code!)
  - > Different builds: MPI with and without OpenMP
- Consider speedup with number of cores used
  - > All runs with SilverStorm Infiniband and ScaliMPI
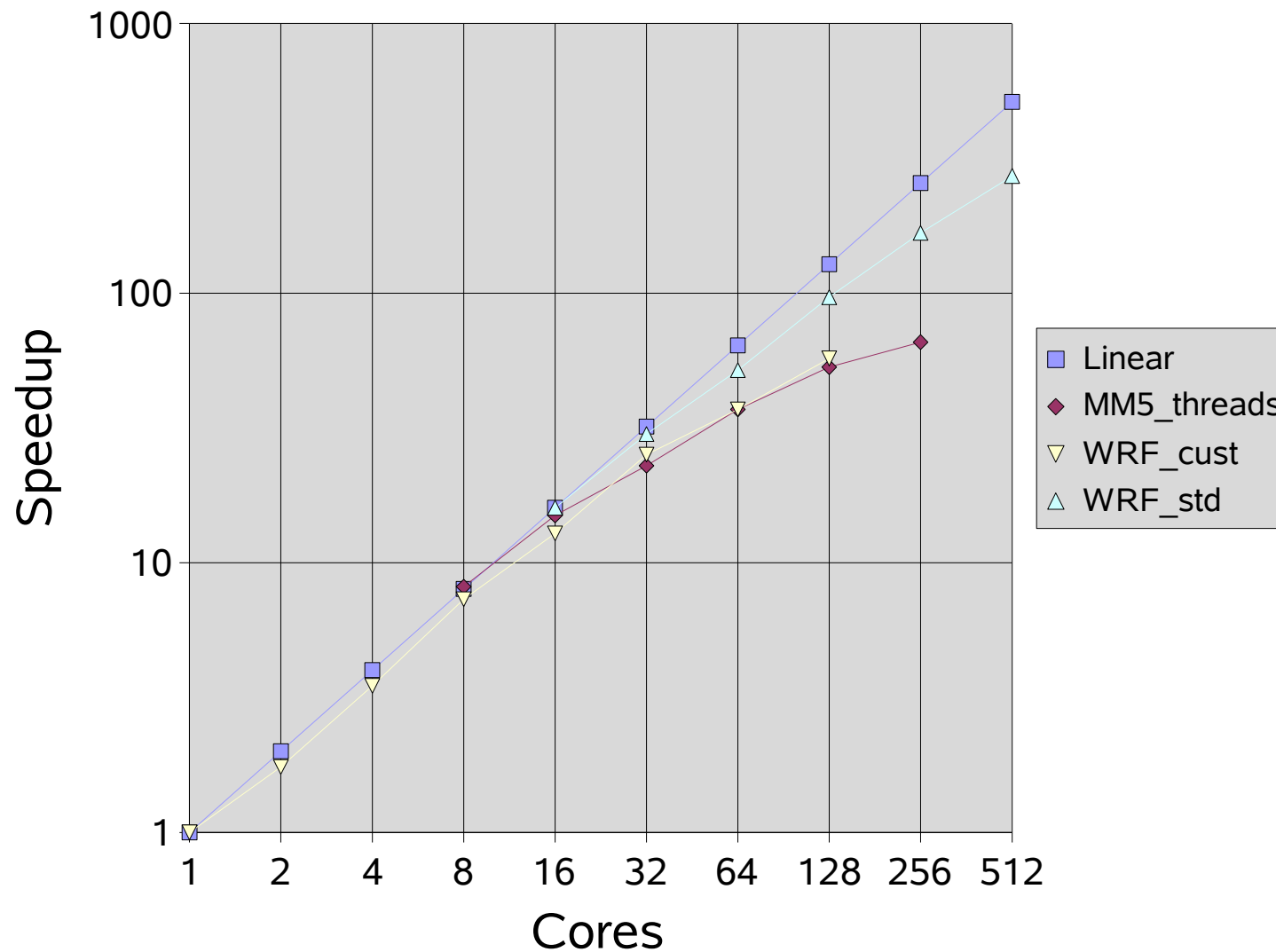 - > Based on time per time-step (different FS used)
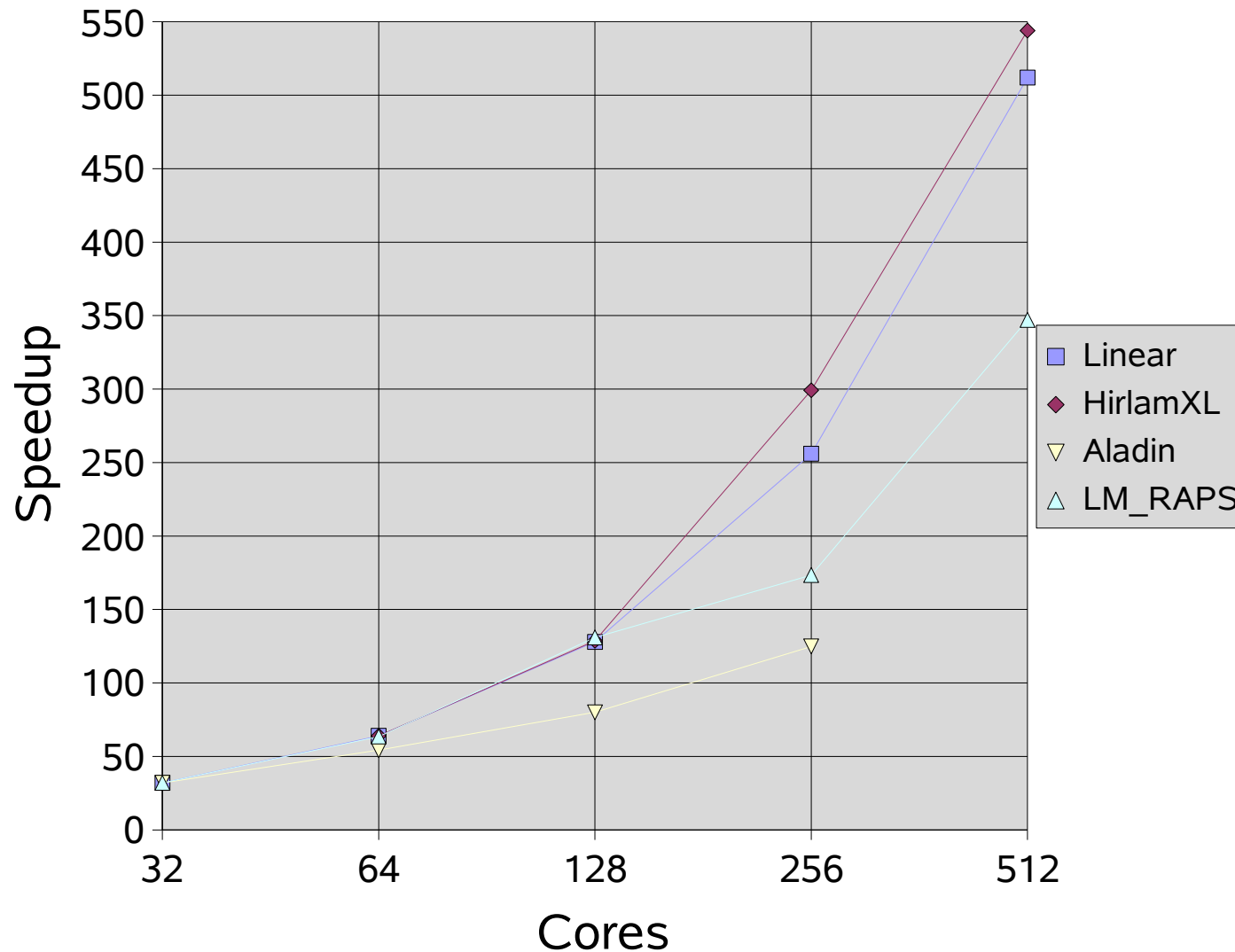
Speedup of Meteorological codes
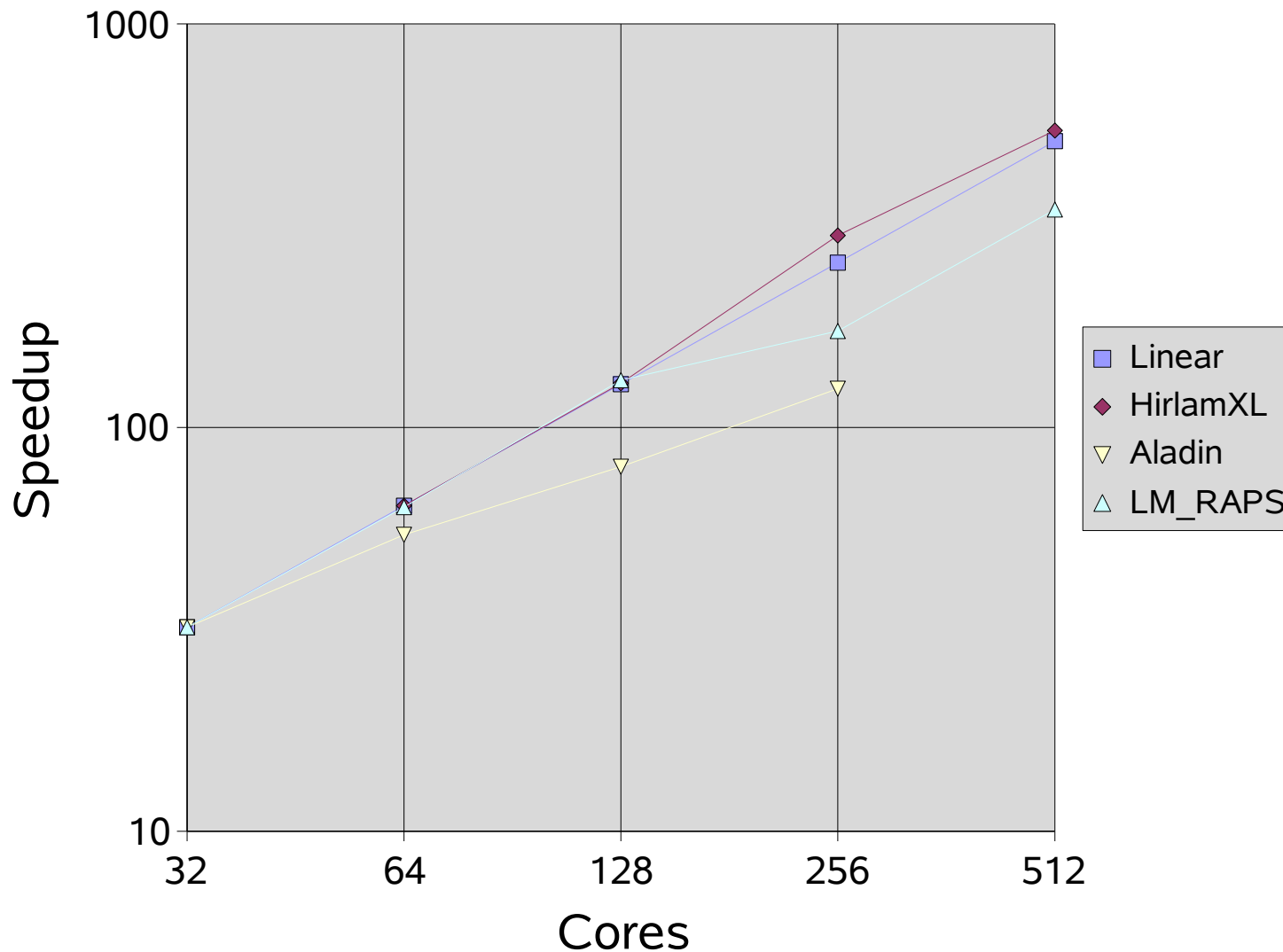## WRF and MM5

Speedup of Meteorological codes
WRF and MM5

Speedup of Meteorological codes

Hirlam, Aladin and LM RAPS

Speedup of Meteorological codes

Hirlam, Aladin and LM RAPS

Summary

- The HPC market is very important for Sun
- Main direction is still
  - > General Purpose Systems
  - > Open Software
  - > R&D - Innovation
- Meteorological codes does scale on distributed systems
- Exciting new products to come
- Visit Sun at SC2006 for NDA presentations
- Visit Sun at www.sun.com or www.sun.com/hpc

Hans Joraandstad
Hans.Joraandstad@sun.com