



**High Performance Computing Clusters**

**Ibergrid**

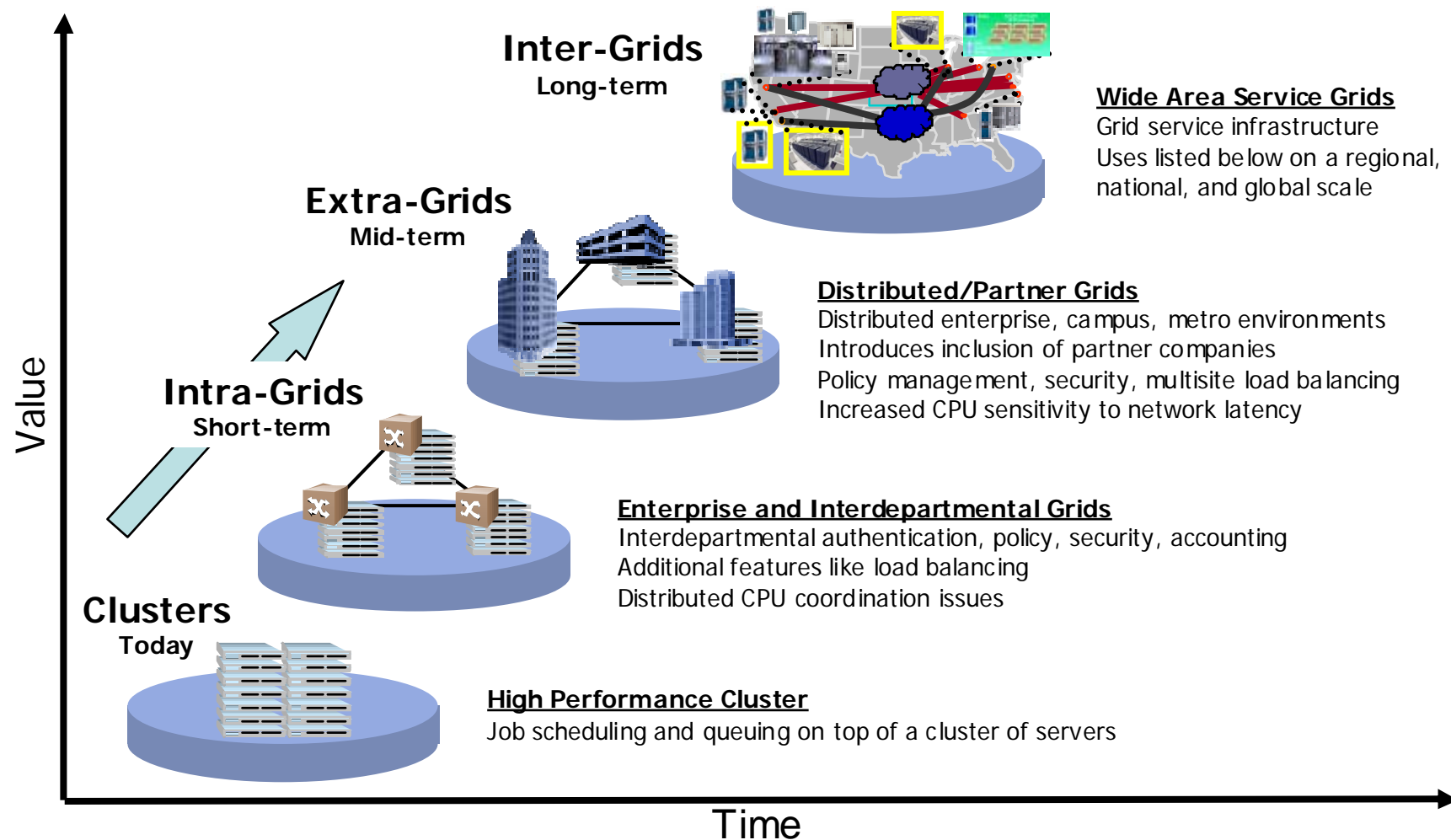
**Miguel Lobato**  
**Solutions Spain and Portugal**



**Santiago de Compostela, May. 14-16, 2007**

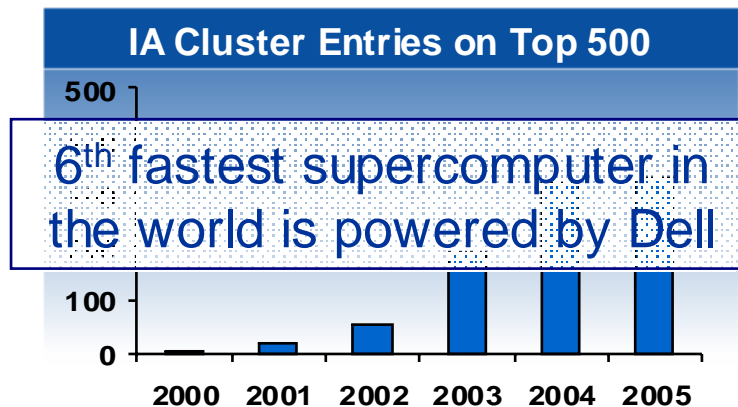
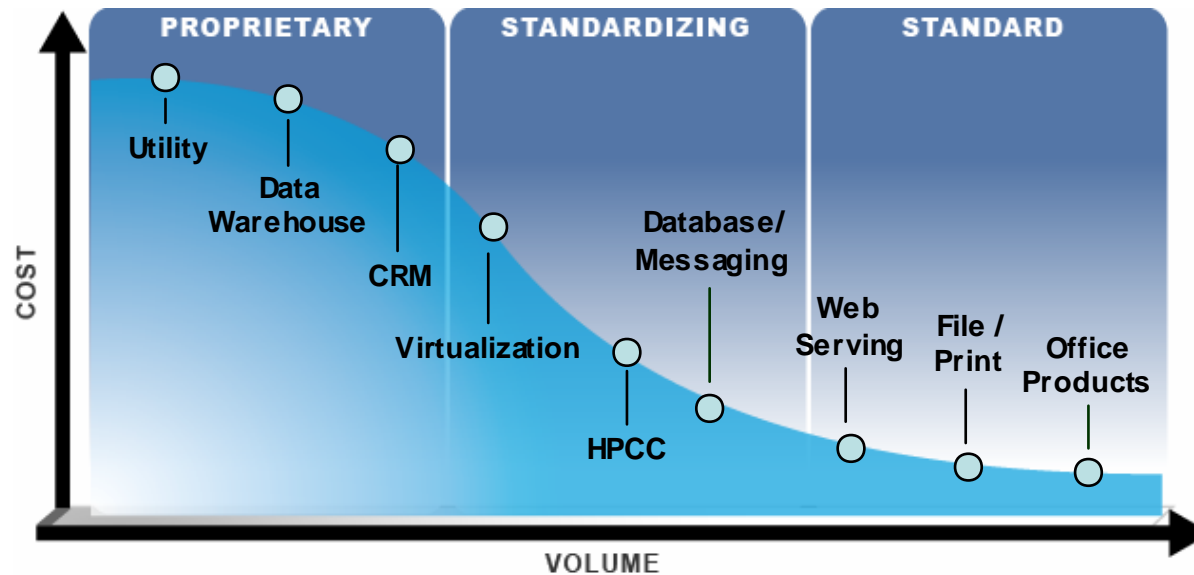
# Evolution of Clustering: Grid Computing

Grid Defined: Aggregation of geographically dispersed computing, storage, and network resources, coordinated to deliver IT services when and where needed.



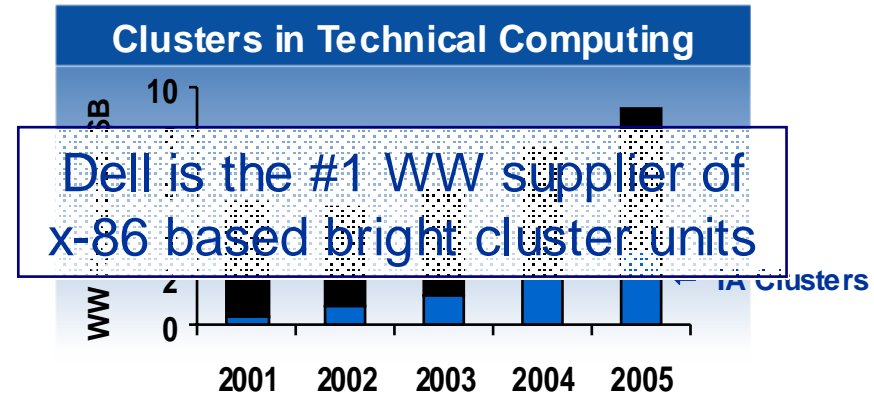
# IA HPC Clusters deliver performance & superior value.

## Rapid market adoption of industry-standard HPC Clusters



6<sup>th</sup> fastest supercomputer in the world is powered by Dell

Source: [www.top500.org](http://www.top500.org), November 2005

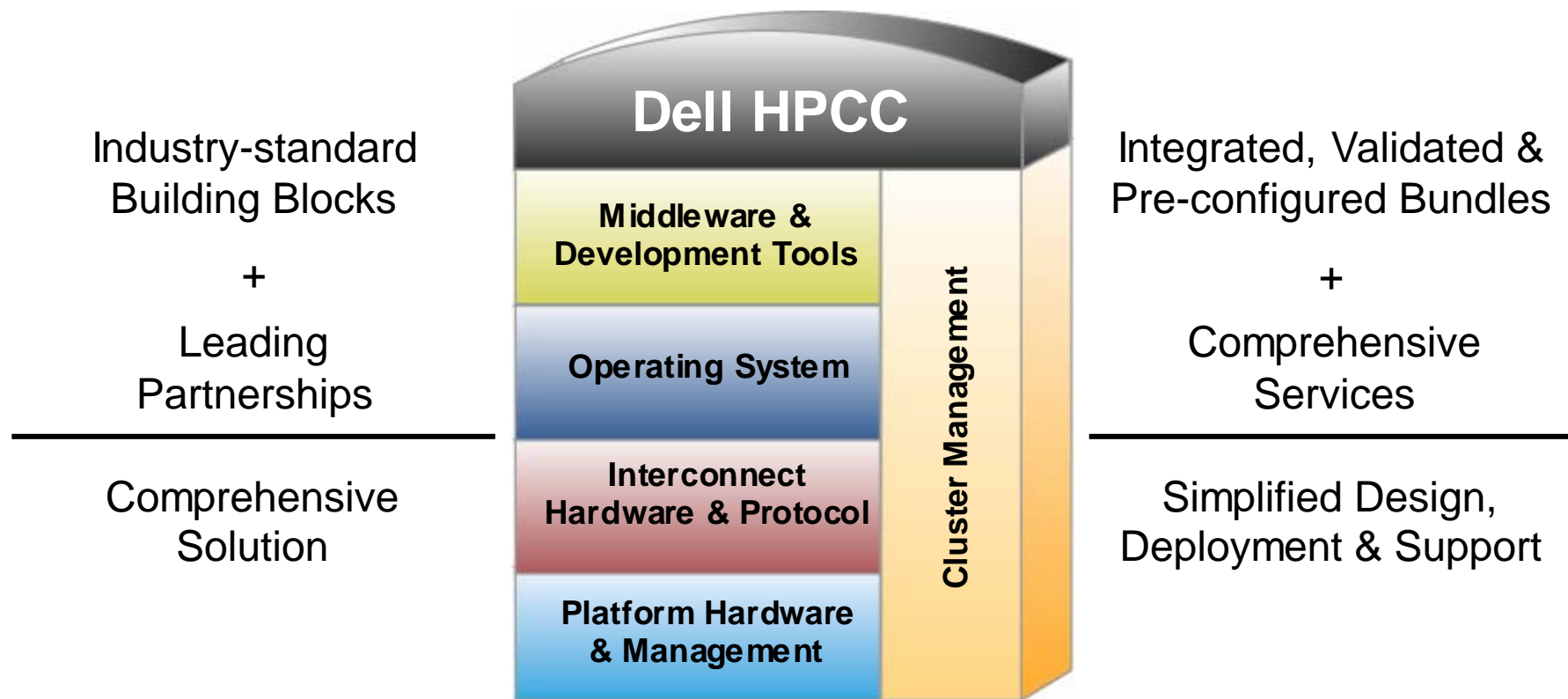


Dell is the #1 WW supplier of x-86 based bright cluster units

Source: IDC High Performance Technical Computer Tracker, Q4 '05

# Dell HPCC Solution Stack

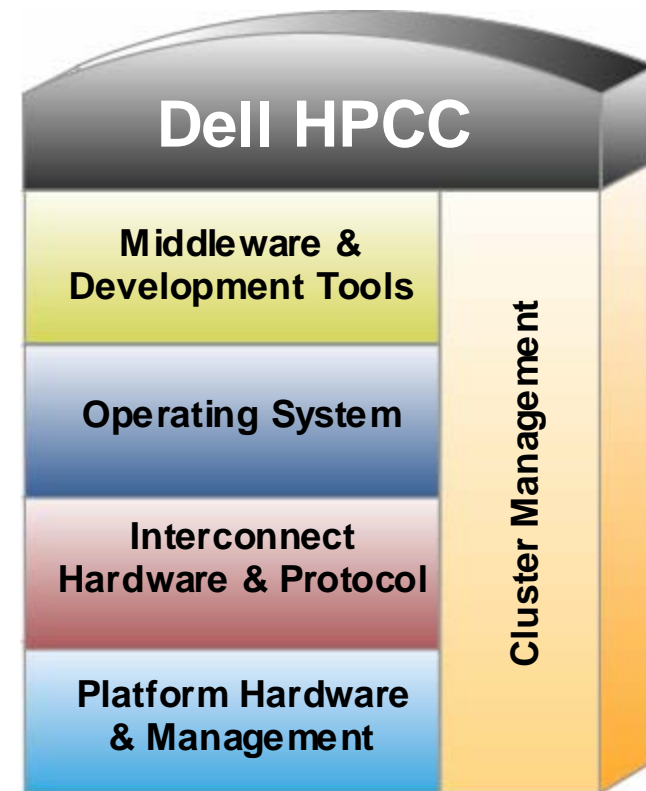
**Maximize Performance – Minimize Cost & Complexity**



# Comprehensive Solutions

## Industry Standard Building Blocks - Leading Partnerships

- **PowerEdge™ Servers:** Opteron™ & Xeon™
- **PowerVault™, Dell/EMC and DataDirect Network®**
- **Red Hat®** - Enterprise Linux® AS & WS
- **Suse Novell® Linux**
- **Gigabit Ethernet, Infiniband & Myrinet**  
Interconnects
- **Cluster Management**
  - **Platform Computing** – Cluster management tools, Job Scheduler & Grid-enabled Middleware
  - **IBRIX®** - Scalable Cluster File System

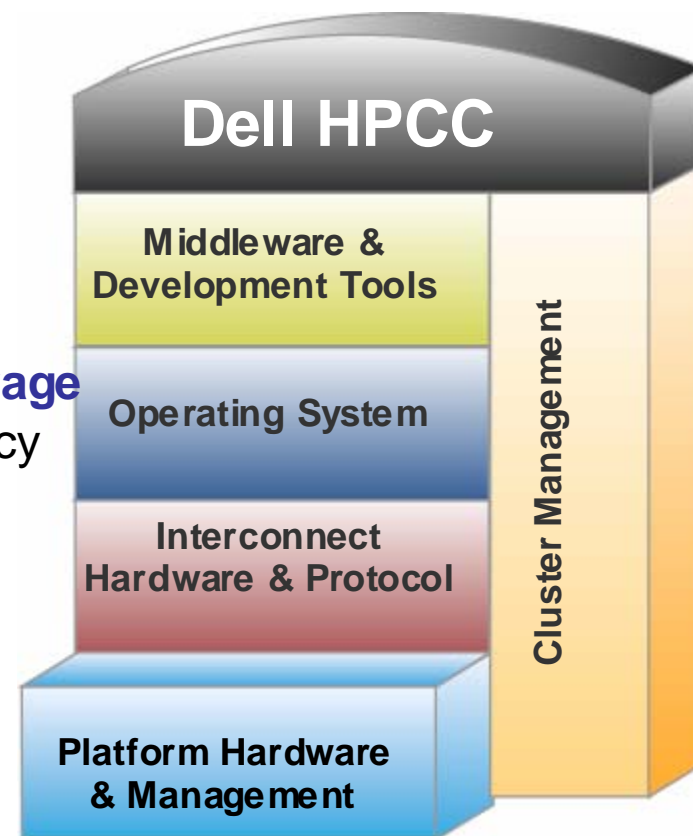


# Industry Standard Building Blocks: Servers

**PowerEdge Servers** leading price / performance & density

- 1U rack and blade form factor
- Intel® Quad-core and AMD Opteron processors
- 1333MHz FSB, Full Buffered and DDR2 memory
- Standards based hardware management IPMI / BMC
- PCI-Express

**PowerVault, Dell / EMC & DataDirect Networks Storage**  
delivers leading performance, scalability and redundancy



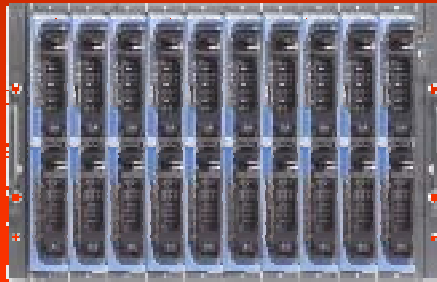


# Blade Compared to Standard Servers

## 10 x PowerEdge 1955 Blades

### Occupies 7U Rack Space

- Typical Power Draw = 1681W
- Typical BTU/hr = **5737**
- Number of AC cords = **4** with redundancy



## 10 x PowerEdge 1950 Server

### Energy Smart

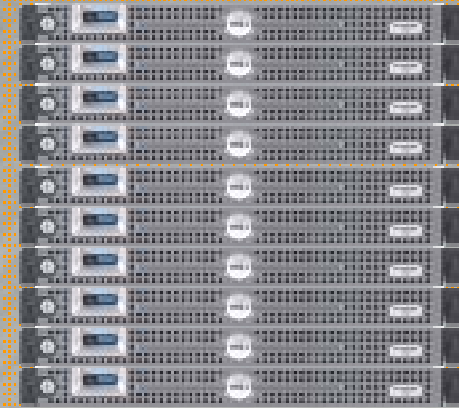
- Typical Power Draw = 2165W
- Typical BTU/hr = **7392**
- Number of AC cords = 20 with redundancy
- Occupies 10U Rack Space



## 10 x PowerEdge 1950 Server

### Standard

- Typical Power Draw = 2388W
- Typical BTU/hr = **8152**
- Number of AC cords = 20 with redundancy
- Occupies 10U Rack Space



## Power Advantages of PowerEdge 1955 Blade Server to PowerEdge 1950 1U Server

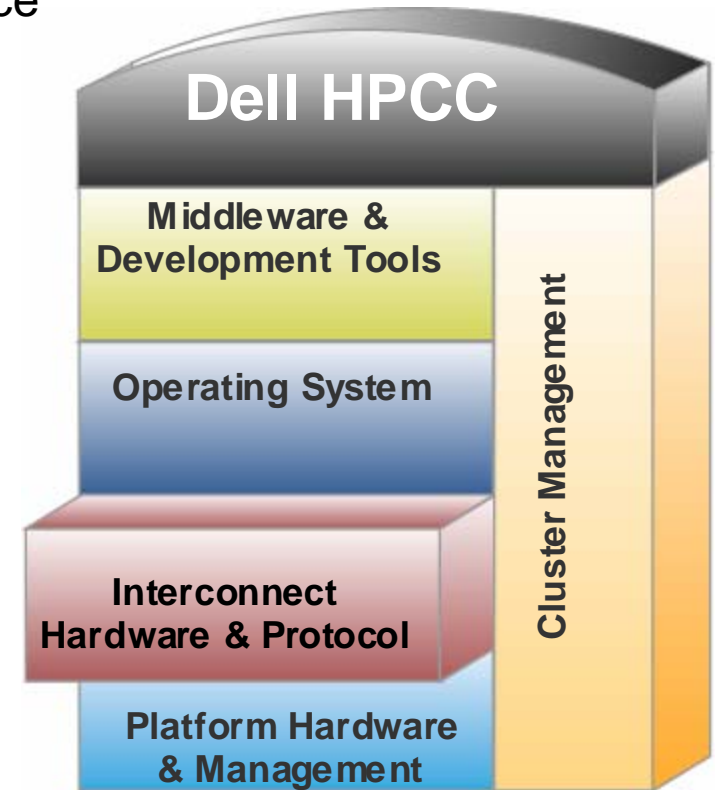
- PowerEdge 1955 Blade Server consumes 22% less power than similarly configured 1U Energy Smart 1950
- PowerEdge 1955 Blade Server consumes 29% less power than similarly configured 1U Standard 1950

Dell Power and thermal results from Dell's Datacenter Capacity Planner (www.dell.com/calculator) January 31, 2007, which is based on SPECint005 benchmark test results. Based on a sample server configuration: Dual Intel Xeon 5148 Dual-Core low voltage 2.33Ghz processors 4x1GB 667Mhz RAM 2x SAS 10k hard drive, redundant power; Blade Chassis configuration: 2xEthernet pass-through, redundant management modules, redundant power. Actual performance will vary based on configuration, usage, and manufacturing variability.

# Best-of-Breed Partners: Interconnects

**Choice of interconnects** balances cost & performance\*

- Infiniband by Cisco
- Myrinet by Myricom
- Gigabit Ethernet by Dell and Extreme Networks





# Interconnect Performance

**Throughput:** Amount of data that can be sent in one unit of time

- MPI point to point bandwidth (approximate MB/s):

<i>Quadrics QsNet</i>	Infiniband			<i>Myricom</i>	Ethernet			
Elan4	1x	4x	12x	Myrinet	Fast	Gig	10Gig	40Gig
911	169	963/893	2025	250	11	110	1100	4400

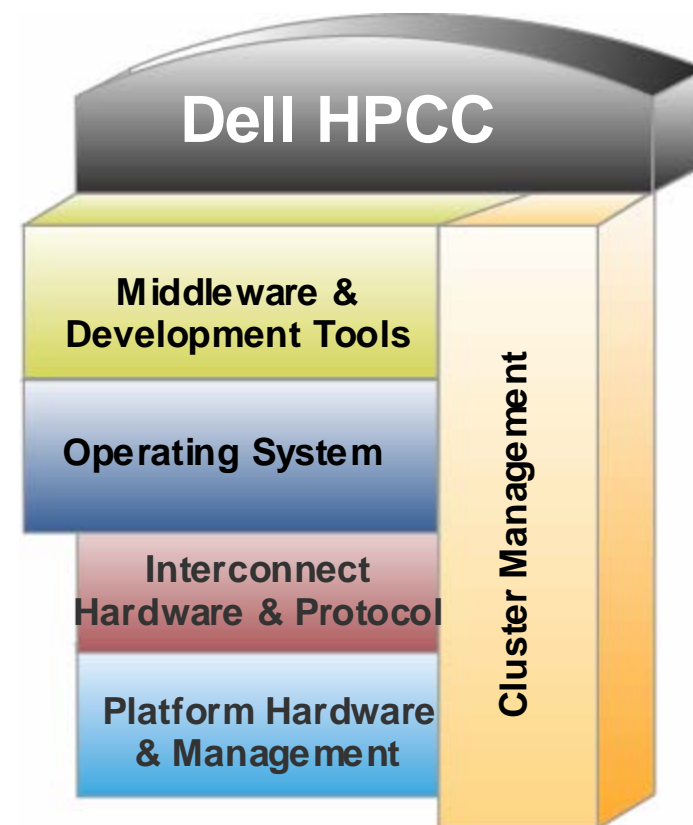
**Latency:** Time required to send a small message from application to a remote application

<i>Quadrics QsNet</i> Elan4	Infiniband	<i>Myricom</i> Myrinet	Ethernet without TOE / with TOE
1.1µs	4.5-5 µs	5.5 µs	50 µs / 17 µs

# Best-of-Breed Partners: Software Stack

**Comprehensive software & management stack**  
simplifies deployment & management.

- **Red Hat and Suse Linux**
- **Platform Computing**
  - Ganglia & Clumon: management & monitoring
  - Parallel Virtual File System 2
  - Platform Lava and Platform LSF HPC for job scheduling & workload management
  - Message Passing libraries
    - Platform LSF HPC grid-enabled middleware
- **Scali Clusters.** Management and monitoring
- **Intel** - Compilers & performance monitoring tools
- **IBRIX** - Scalable cluster file system



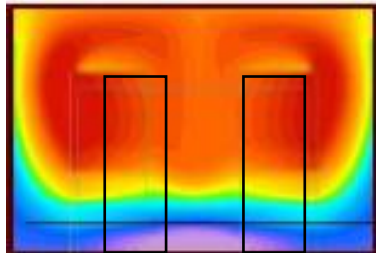
\* Platform Rocks is based on the NPA CI ROCKS cluster software package developed by the San Diego Supercomputing Center.

# Comprehensive Services Portfolio

## Speed Time to Production - Streamline Resolution

### Plan

- Design & Solution Validation
- Benchmark Testing  
Dell & University of Texas  
Advanced Computing  
Center (TACC)
- Environmental Assessment



**Datacenter  
Assessment**

### Implement

- Custom Integration
- Pre-configure, Rack & Ship
- Onsite Deployment & Acceptance Testing
- Technical Training



**Custom Factory  
Integration**

### Maintain

- Professional Support Services
- Partner collaboration
- Technical Account Management
- Onsite service options



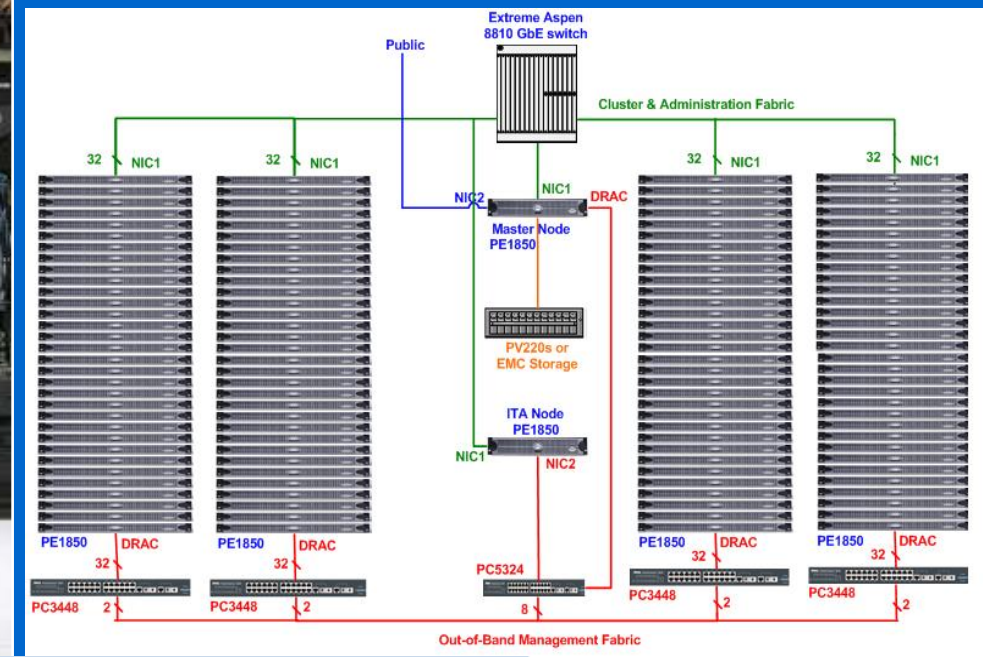
**Enterprise Command  
Center**

# Design, Deployment & Support Simplified

## Integrated, Validated & Pre-configured Bundles



### 8 – 256 node bundles



### HPCC Design and Test Lab

- State of the art engineering lab
- Collaborative Partner design & testing
- Standardized architectures
- Technical papers & Best Practices

# Dell Top 500 Results

## Dell Clusters on Top500.org\*

### #6 Sandia National Laboratories (NNSA)

11 Maui High Performance Computing Center (MHPCC)  
12 Texas Advanced Computing Center (TACC)  
27 NCSA  
45 BYU  
59 NCSA  
78 CalTech  
87 SUNY  
96 University of Sherbrooke  
104 University of North Carolina  
108 NCSA  
123 University of Oklahoma  
128 KTH Royal Institute of Technology  
166 Imperial College ICT HPC  
186 UT SimCenter, Chattanooga  
188 CalTech/JPL  
406 University of Sherbrooke



## Highlights

- **17 systems** on the Top 500 list
- **6<sup>th</sup> fastest supercomputer** in the world
- **Industry Standard leadership**
  - Intel EM64T
  - Gigabit Ethernet
  - Infiniband
  - Myrinet

\*Source: Top500 Supercomputer Sites, Nov 2006 [www.top500.org](http://www.top500.org)



# Brigham Young University

## *expands supercomputing resources with MaryLou 4*

### Challenge:

- Increase computational capacity
- Strict heat, space & power restrictions
- Limited technology staff

### Solution:

- 630 PowerEdge™ 1855 blade servers, each with dual Intel® Xeon™ EM64T 3.6GHz processors
- Red Hat® Linux® operating system
- Platform Computing LSF Rock
- Cisco InfiniBand® interconnect, 10 Gigabit Ethernet interface & Force 10 network switches
- Data Direct Networks storage

### Benefits:

- Exceptional computational performance within its heat, space and power parameters
- Enables students and faculty in both the graduate and undergraduate programs to forge ground-breaking research

*“We needed as much computing power as possible with the lowest heat profile. Our primary goal was the biggest bang for BTU of heat generation.”*

**-David Stirling**

Supercomputing Administrator  
BYU Supercomputing Laboratory  
August 2005

*45th fastest supercomputer worldwide\**  
*Top 500- Nov 2006*

\*Source: Top500 Supercomputer Sites, Nov 2006 [www.top500.org](http://www.top500.org)

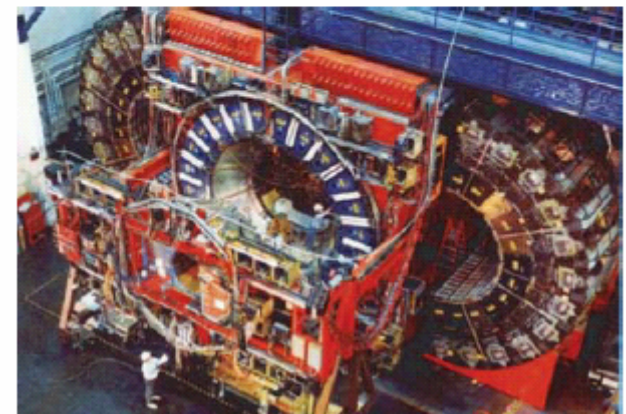


# Oxford University: Scientific Research

- Oxford University is one the UK's and the world's leading education and research establishments.
- Uses Dell cluster for high energy physics research
- Oxford uses HPCC to speed UK-based research on data obtained from experiments conducted on the Fermilab accelerator in the US,
- Serviced by Dell/Scali partnership
- Research centre deploys four HPCC clusters across the UK, each consisting of Dell PowerEdge 2650 servers and PowerVault 220S storage



Collider Detector Facility (CDF) located at the collision point between protons and antiprotons in Fermi National Accelerator Laboratory's Tevatron accelerator.











## Spain Case Studies: CESGA

### SVG – Superordenador Virtual Galego

- 80 servers 1U DELL 750 P4 3.2GHz, 112Gb RAM. **512 GFlops**
- 36 servers Blade DELL 1955 Quad Core Intel, 148Gb RAM. **2.227GFlops**
- **15,3 Tb Scratch Parallel Storage**

#### MONITOR SISTEMAS DE CÁLCULO

	Nome	Sistema	CPUs	GFlops Pico	Memoria Gbytes	Disco Gbytes	Carga	Traballos Execución
	SC	Compaq HPC320	32	64	80	2000	 93 %	25
	SUPERDOME	Cluster SUPERDOME	128	768	384	7168	 100 %	100
	SVGD	Cluster SVG	391	2371	273	16032	 43 %	123

## Other references





**Gracias  
Obrigado  
Thank you**

**miguel\_lobato@dell.com**

[www.dell.com/hpcc](http://www.dell.com/hpcc)



**Santiago de Compostela, May. 14-16, 2007**