



Cisco InfiniBand Switching for High Performance Applications Defining a Unified Compute Fabric for The Enterprise

**Amir Sharif, Product Manager
InfiniBand Infrastructure**

June 2006

Agenda

Cisco.com

- **Unified Compute Fabric and Distributed Systems**
- **Cisco leadership: Open fabrics & Open MPI**
- **SFS-OS 2.7 update & CiscoWorks Integration**

InfiniBand Adoption **Across Customer Markets** Example 2005 Cisco SFS Deployments

Cisco.com

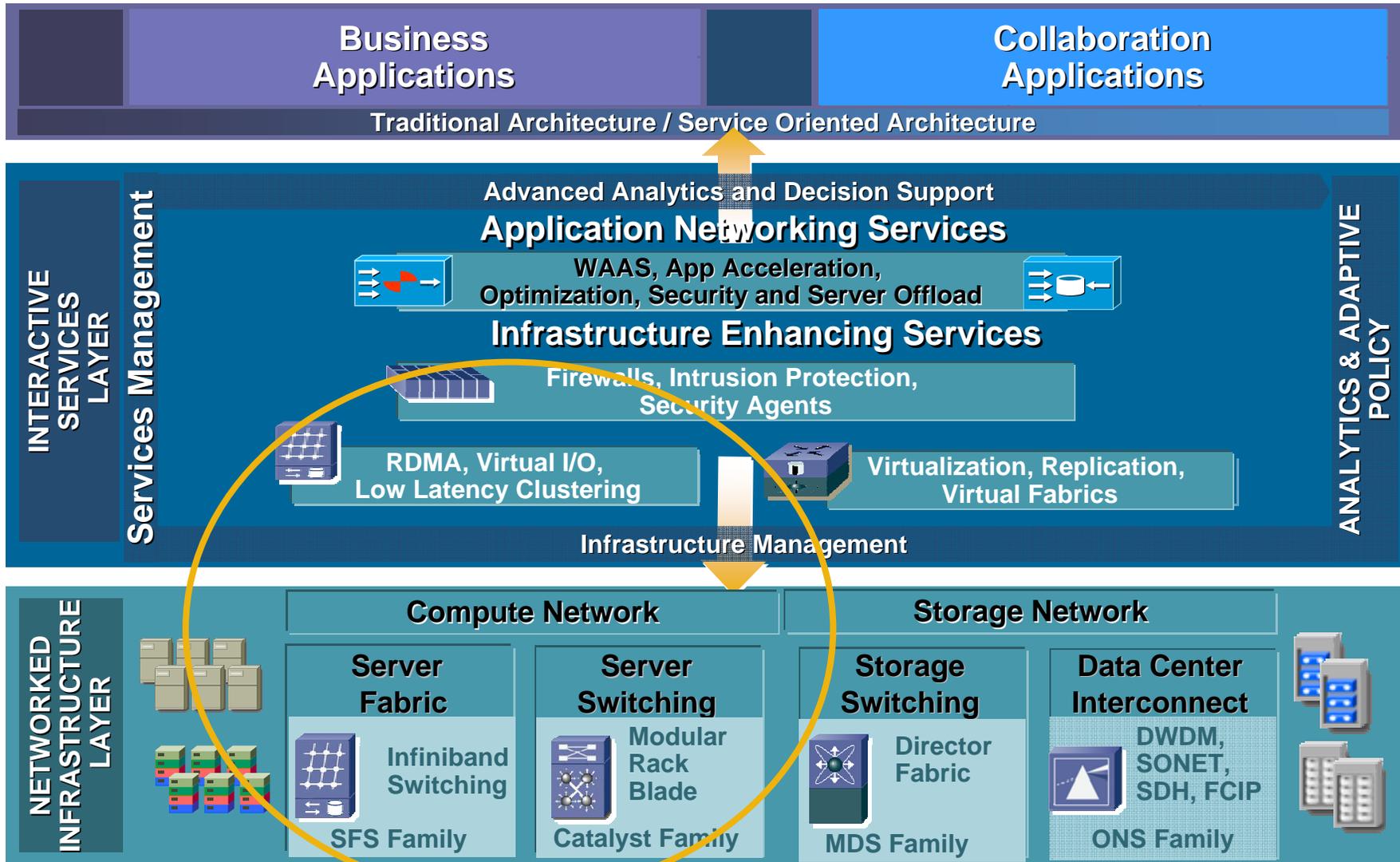
| | Vertical | InfiniBand-attached servers | Top 500 Ranking |
|-----------------------------------|----------------------|-----------------------------|-----------------|
| Sandia National Labs | Government/Research | 4400 | 5 |
| Bio-Infomatics Cluster* | Research | 1066 | |
| Tier 1 Systems Company* | Service Provider | 1024 | |
| University of Sherbrooke (Canada) | Education/Research | 576 | 51 |
| NCSA | Education/Commercial | 512 | 59 |
| University of Oklahoma | Education | 512 | 67 |
| University* | Education | 386 | |
| Wall Street Bank* | Financial/Grid | 370 | |
| Japanese Car Company* | Automotive/CAE | 300 | |
| MCNC | Government/Education | 284 | |
| SARA (Netherlands) | Government/Education | 272 | 277 |
| Tier 1 Systems Company* | Hosted Services | 256 | |
| Wall Street Bank* | Finance/Grid | 256 | |
| TACC | Education/Commercial | 200 | 301 |
| Arizona State | Education | 200 | 328 |

* Non-Referenceable Customer Name

Cisco Data Center Network Framework

Defining a Unified Compute Fabric

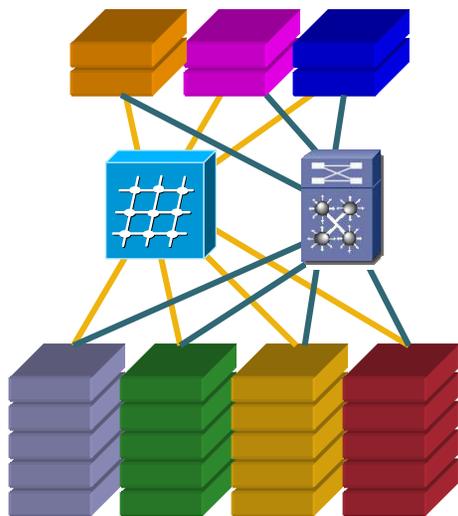
Cisco.com



Distributed Applications & Networking

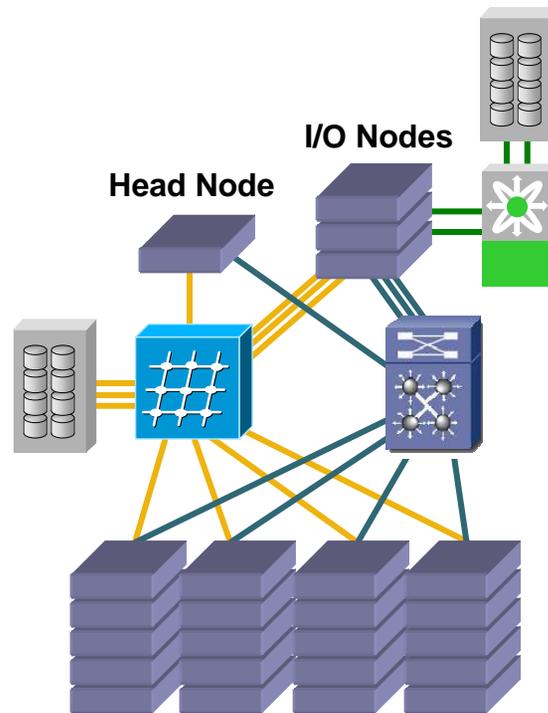
Cisco.com

Distributed Systems



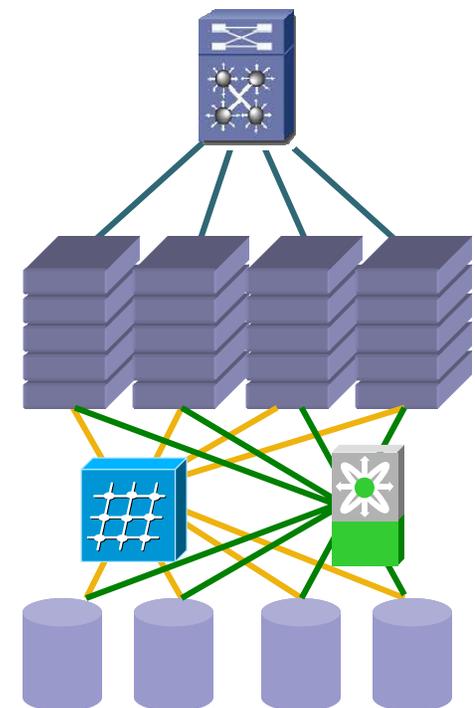
TIBco, Wombat, 29West, etc

High-Performance Computing



Fluent, Ansys, Charmm, Amber, LS-Dyna, etc

High-Performance Database and Storage



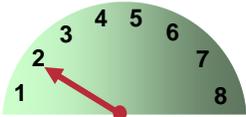
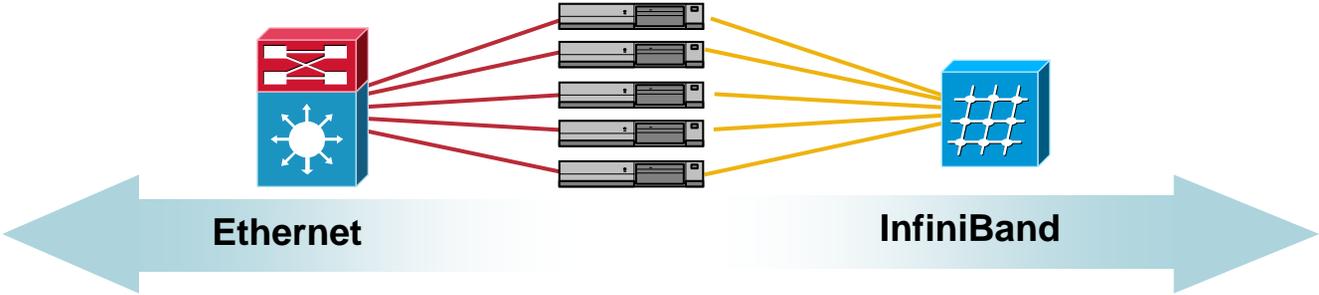
Oracle, IBM DB2
Luster, IBM GPFS, Panasas,

Are feeds and Speeds enough?

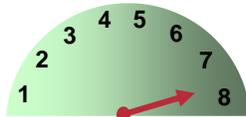
Cisco.com



Fragmented Landscape Inhibits Distributed Systems Adoption



- **Multiple MPI implementations – no standards to build off of, more options to certify**



- **Extensive Configuration and management tools**

- **Limited Configuration and management tools**



- **Broad support for standard protocol stacks - Interoperable**

- **Loosely defined protocol stacks – limited interoperability**



- **Simple “Plug & Play” Operation**
- **10/100/1000 + 10GE**

- **Do-it-yourself testing, integration and certification**
- **4X 10Gbps Ports**



Multiple Fabric Options – No uniform integration of Capabilities

What is a Unified Compute Fabric?

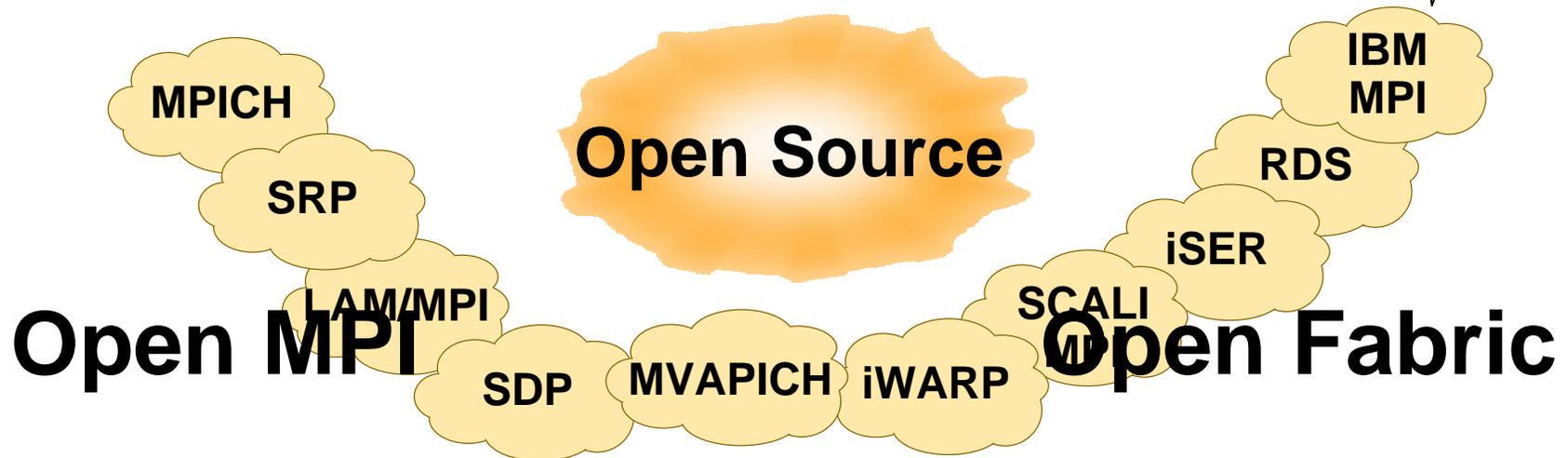
Cisco.com

A unified compute fabric integrates network management and troubleshooting, application protocols and APIs for Ethernet and InfiniBand fabrics, offering IT cluster administrators a single operational model for greater ease-of-use, rapid deployment, higher reliability and integrated security

Driving Open Standards Open Fabric and Open MPI

New

cisco.com



- **Problem:** Proprietary protocol stacks multiply options for application development and complicate certifications
- **Solution:** Cisco will move strategic focus to Open Fabrics and Open MPI
- **Benefits:**
 - Single stack, fully interoperable between vendors
 - Eliminates fragmented infrastructure that reduce performance
 - Accelerates application development for the ISV and deployment and certification for the end-user
 - Consistent operation across Ethernet & InfiniBand

Bringing Ethernet-Like Maturity to Infiniband

Cisco.com

Ethernet

Standard Manageability

Standard MIBs and Manageability integrated into Ethernet

Standard Drivers and Application Interface

IEEE 802.3, 802.1, TCP/IP have unified the world around a common capability

Evolutionary Performance

Incremental performance capabilities provide right-sized price per port

Unifying Ethernet and Infiniband for High Performance Applications

Infiniband

CiscoWorks LMS Integration and TopSpin OS 2.7

Common platform for multi-fabric management and new management and security capability

Open Fabrics and Open MPI

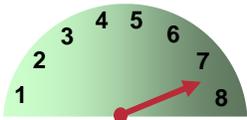
Standard development platform for drivers and app integration

Double-Data Rate (DDR)

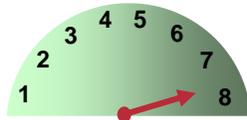
Higher bandwidth option for incremental bandwidth requirements

Delivering the Unified Compute Fabric

Cisco.com



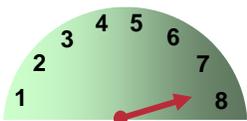
- Drive commercial Open MPI middleware to deliver standard features and performance, and reduce implementation and support overhead



- InfiniBand Integration into CiscoWorks Management tools



- Drive standardization of Protocols within Open Fabrics Consortium to deliver standard features and reduce time to deploy



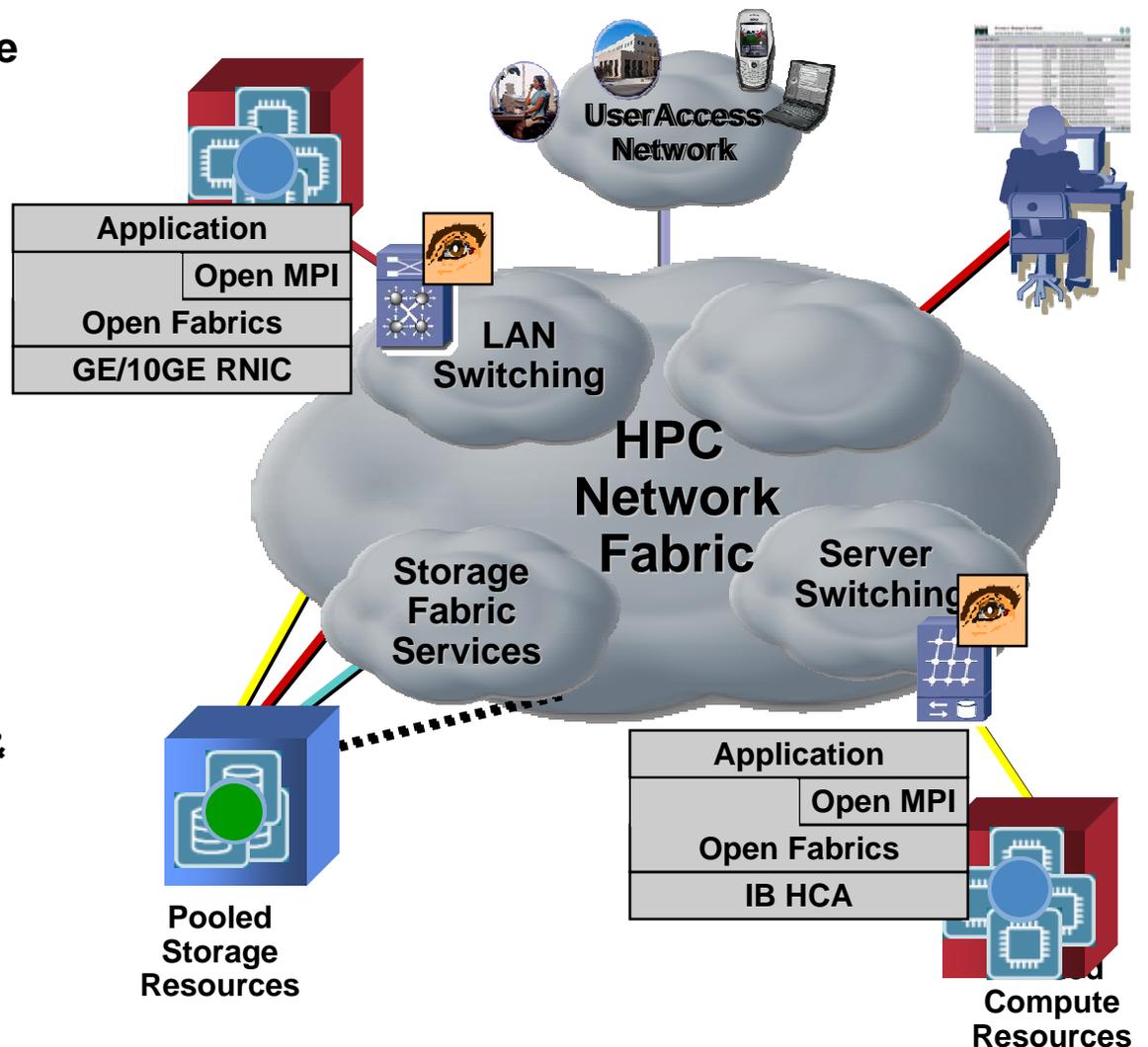
- Extend Ethernet “Plug & Play” functionality to InfiniBand
- Deliver Next-generation technologies – InfiniBand dual speed DDR/SDR

Uniform Cross-Fabric Capabilities

Cisco Open Fabric Strategy

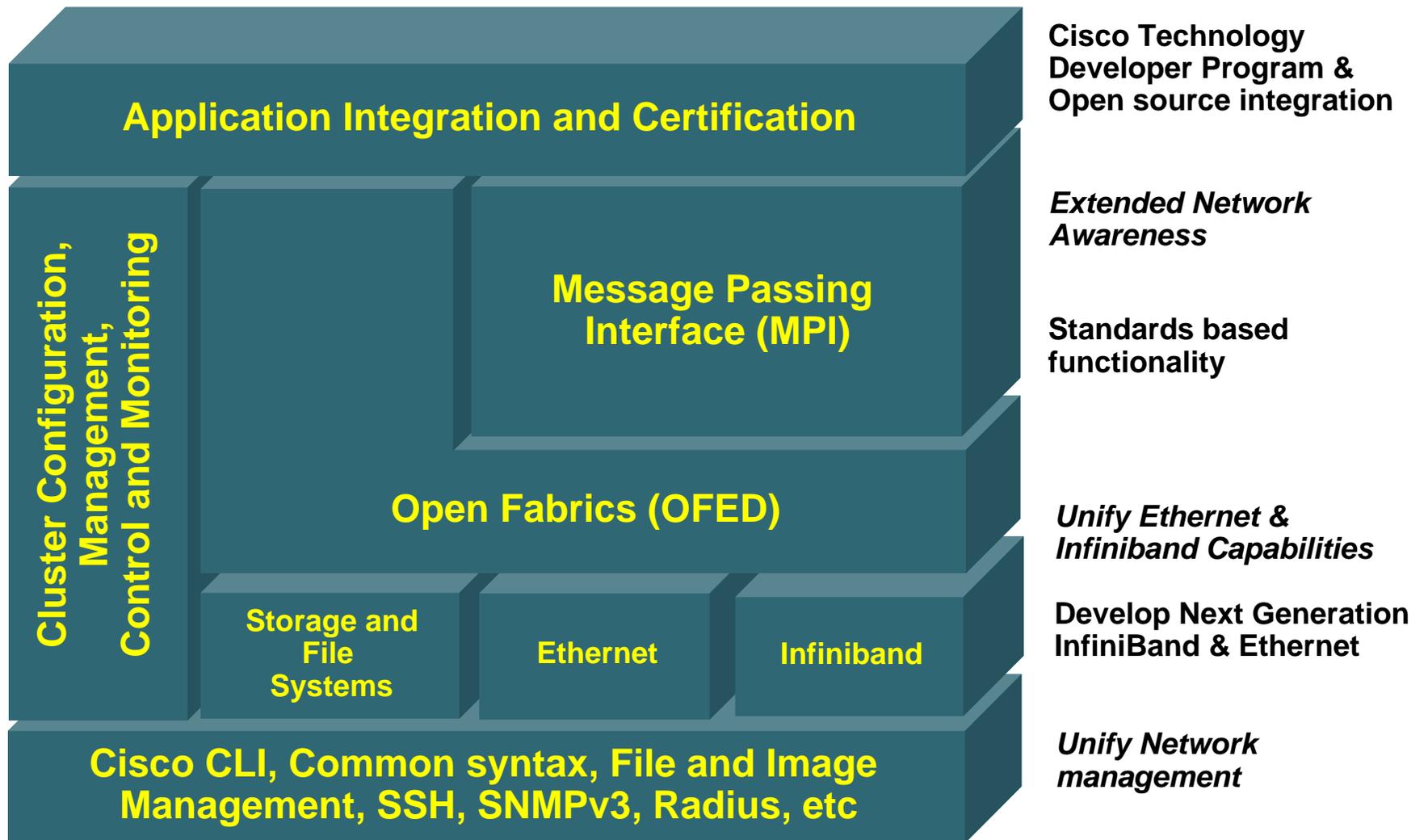
Cisco.com

- **The Enabler for true Enterprise High-performance Computing**
 - HPC/GRID
 - Applications Systems
- **Only vendor to offer complete multifabric solution**
- **Multi-fabric Management**
 - Common tools for *Ethernet & InfiniBand*
 - Comprehensive APIs for 3rd Party Integration
- **Fabric Agnostic Protocols**
 - Unify MPI across *Ethernet & InfiniBand* with Open MPI
 - Unify RDMA across *Ethernet & InfiniBand* with Open Fabrics.
- **Cisco ISV Certification**
 - InfiniBand & Ethernet*
 - Certification for *InfiniBand & Fiber Channel Storage*



Cisco Delivering Unification to Distributed Computing

Cisco.com



Cisco SFS-OS 2.7

Consistent Configuration & Management



- **Common CLI across all products**
Command Syntax, scripting, etc.
- **Consistent Security model**
TACACS and RADIUS for Centralized Auth/ACS Integration
SSH/SSL/SNMPv3 for full management security
Multiple authorization levels
- **File & Image Management**
System image and configuration file libraries.
- **Consistent Management Notification**
Full SNMP v1/v2/v3 Support across all fabrics
Streaming Syslog: Integrates with Syslog Analyzer



CiscoWorks LMS Support for InfiniBand

New

- Single network management application for Ethernet and InfiniBand networks
- Delivers Centralized Device, Software and Configuration Inventory Manager
- Diagnostic Tools and Syslog Analyzer
- Centralized Reporting
- Device level fault analysis for network fabric, including high availability monitoring, pager/email/trap notification
- **Benefit: Eliminates administrative and usage barriers; identify and fix problems -> increased performance**

The screenshot displays the CiscoWorks LMS interface with several overlapping windows. The top window is 'Archive Mgmt' showing configuration files. Below it is a 'Hardware Report' window. The central window is 'Resource Manager Essentials' displaying a 'Syslog Analyzer Standard Report' with a table of log entries. The bottom window is a 'Detailed Device Report' for device 10.64.158.162, showing system and chassis information.

| Interface | Timestamp | Facility/Sub-facility | Severity | Mnemonic | Description |
|--------------------|----------------------|-----------------------|--------------|--|-------------|
| 1. 192.168.158.106 | Jan 25 2005 00:27:20 | SYS | 4 SNMP_WRITE | SNMP WriteNet request. Writing current configuration to 192.168.138.24 | |
| 2. 192.168.158.106 | Jan 25 2005 01:27:27 | SYS | 4 SNMP_WRITE | SNMP WriteNet request. Writing current configuration to 192.168.138.24 | |

| Update Time | Device Name | Domain Name | User defined Serial No. | System Name | Description | Location | Contact |
|--------------------------|---------------|-------------|-------------------------|-------------|--|----------|---------|
| 13 Aug 2001 03:00:35 GMT | 10.64.158.162 | | 69029236 | CM5500 | Cisco Systems WS-C5500 Cisco Catalyst Operating System Software Version 4.5(10) Copyright (c) 1995-2000 by Cisco Systems | chennai | rmuthua |

| Available Slots | Slot Capacity | Model | Power Supply1 | Power Supply2 | Backplane Type | Chassis Type | Mgmt Type | Network Mask | Broadcast Address |
|-----------------|---------------|----------|---------------|---------------|----------------|--------------|-----------|-----------------|-------------------|
| 10 | 13 | WS-C5500 | none | wsc5508 | giga3E | wsc5500 | snmpV2cV1 | 255.255.255.240 | 10.64.158.175 |

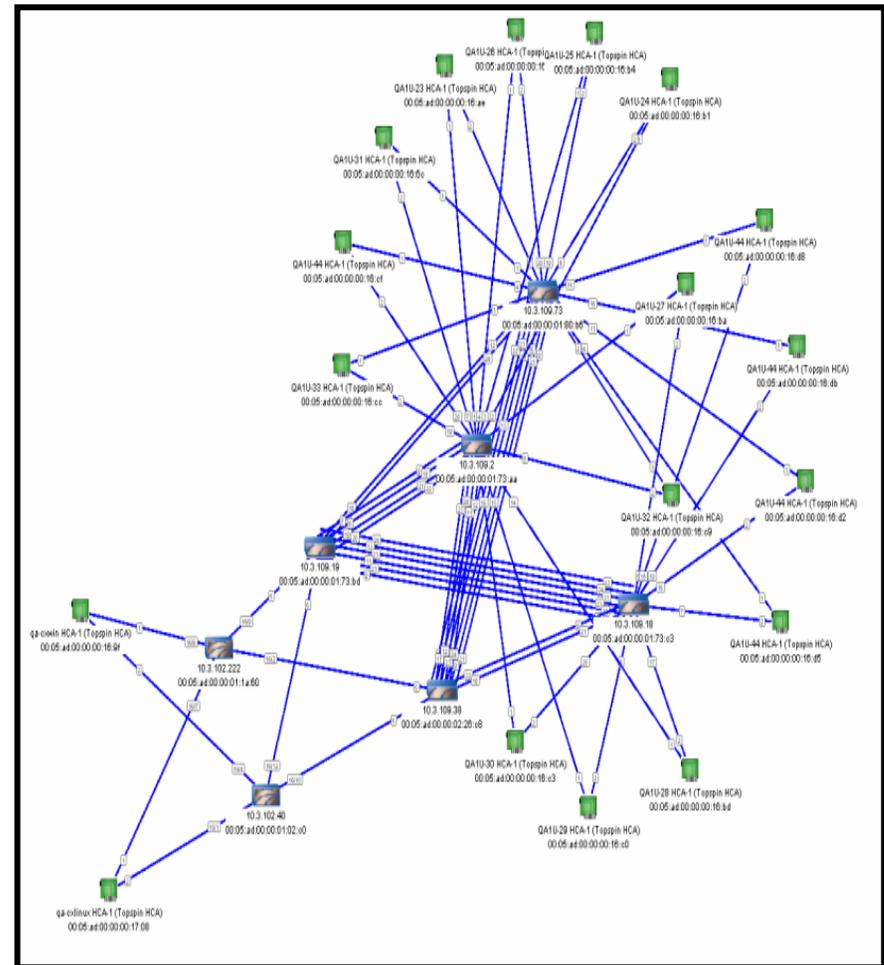
| Slot No | Description | Type | Speed (bps) | Physical Address | Network Address | Status |
|---------|-------------|----------------|-------------|-------------------|-----------------|--------|
| N/A | sc0 | ethernetCsmacd | 10000000 | 00:90:5f:8b:2d:1f | 10.64.158.162 | up |
| N/A | sl0 | slip | 9600 | 00:00:00:00:00:00 | | down |
| N/A | VLAN 11 | prop/Virtual | 0 | 00:90:5f:8b:2a:0a | | up |
| N/A | VLAN 12 | prop/Virtual | 0 | 00:90:5f:8b:2a:0b | | up |

Available: Q3, 2006

High Performance Subnet Manager

Cisco.com

- Overall Subnet Manager and framework tested on Sandia Thunderbird -- World's largest standard HPC server cluster (4500 IB-attached servers)
- SM brings multi-thousand server cluster up in less than one minute
- Includes database synchronization between redundant SM's for HA
- Includes performance and statistics tools capable of reporting on tens of thousands of network ports in half a minute
- Multi-Vendor support



Cisco InfiniBand Hardware Roadmap

The diagram illustrates the Cisco InfiniBand hardware roadmap, featuring several product categories and their status:

- New Products (marked with yellow starburst icons):**
 - Cisco PCI-X & PCI-EX 1*4X HCA
 - IBM Bladecenter InfiniBand 4X Switch & HCA
 - SFS-3012P
 - 4X DDR PCI-Ex HCA
- Other Products:**
 - IBM Bladecenter InfiniBand 1X Switch & HCA
 - Dell 1855 Bladecenter InfiniBand 4X Module & HCA
 - SFS-3012
 - SFS-3012 Gateways
 - SFS-7008
 - SFS-7000
 - SFS-7000P
 - SFS-7000P
 - SFS-7012P
 - SFS-7024P
 - SFS-7000D Family
 - Cisco PCI-X & PCI-Ex 2*4X HCA

Cisco InfiniBand Software Roadmap

Cisco.com

OpenFabrics
iWARP integration

HPSM 1.3

New
OpenFabrics
Enterprise
Distribution v2.0 –
RDS and iSER
support

New
Resource Manager
Essentials SFS
integration

SFS-OS

Advanced QoS (SL to VL
mappings)

InfiniBand port mirroring
InfiniBand packet sniffing

High-Performance
InfiniBand Subnet
Manager

Cluster Readiness Tool Kit

TopspinOS Switch Management v3.2

OpenFabrics Enterprise
Distribution
(iPoIB, SDP, SRP, OpenMPI,
MVAPICH)

Embedded Subnet Manager <1100 nodes

New
SFS-OS v2.7.0

New
High-Performance
InfiniBand Subnet
Manager 1.2
(multiple vendors)

CiscoWorks Device fault
Manager

- Advanced Fabric Management Tools
- Additional topology-aware logic/variables

Q and A



