Euro PVM/MPI 2007

Jeff Squyres

Open MPI Architect, Cisco

 Cisco Public

Questions Jeff gets asked

- Aren't you…

  Yes, we are the group that used to be Topspin

  Cisco acquired Topspin ~two years ago

  InfiniBand is now part of Cisco's datacenter portfolio

Questions Jeff gets asked

- But Cisco = Ethernet. Did Cisco acquire Topspin to kill InfiniBand?

  ### *Absolutely not!*

  InfiniBand *complements* Ethernet in both HPC and other datacenter environments

  Cisco is aggressively marketing and selling InfiniBand products

  Cisco has an extended, long-term roadmap for InfiniBand

  Strategic partnerships grow the entire InfiniBand ecosystem

Questions Jeff gets asked

- Why does Cisco care about MPI?

  HPC customers want a reliable, scalable MPI

  Cisco spends a lot of effort on QA resources for many MPI's (HP, Intel, Open MPI, MVAPICH)

  Many customers _require_ an open source MPI

  Who would want "Cisco MPI"? No one!

  It made obvious sense to join a larger industry-supported, research-driven community MPI project (Open MPI)
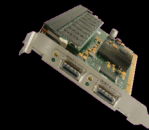
Cisco SFS InfiniBand Products

- Host-side products

- InfiniBand switches

- I/O gateways

                                                                5

Cisco InfiniBand Host-Side Products

- ## HCAs ("InfiniBand NICs")
    PCI-e for individual servers

    Incorporated into IBM / HP blade systems

- ## Software drivers for Linux
    OpenFabrics Enterprise Distribution (OFED) v1.2.5
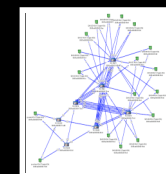
    Open source, standards-based

    Active development of Open MPI

    QA testing of other MPI's: HP, Intel, MVAPICH

- ## Cisco High Performance Subnet Manager
    Tested scalability to 4,600 hosts
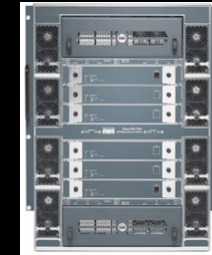
Base InfiniBand Switch: Cisco SFS 7000D

- **24 dual-speed InfiniBand 4X ports**
  - **20-Gbps double data rate (DDR)**
  - **10-Gbps single data rate (SDR)**

- **Non-blocking 480-Gbps bisectional bandwidth**

- **Port-to-Port latency of 140 nanoseconds**

- **Embedded subnet manager** up to 1000 nodes

- Cisco CLI, Web, and Java-based systems management options

- Powered ports for flexible copper and optical interfaces

- Dual, hot swappable, redundant power supplies and fans

Larger InfiniBand Switches:
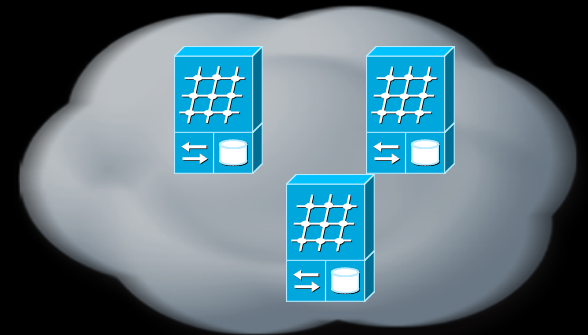# Cisco SFS 7012D and 7024D
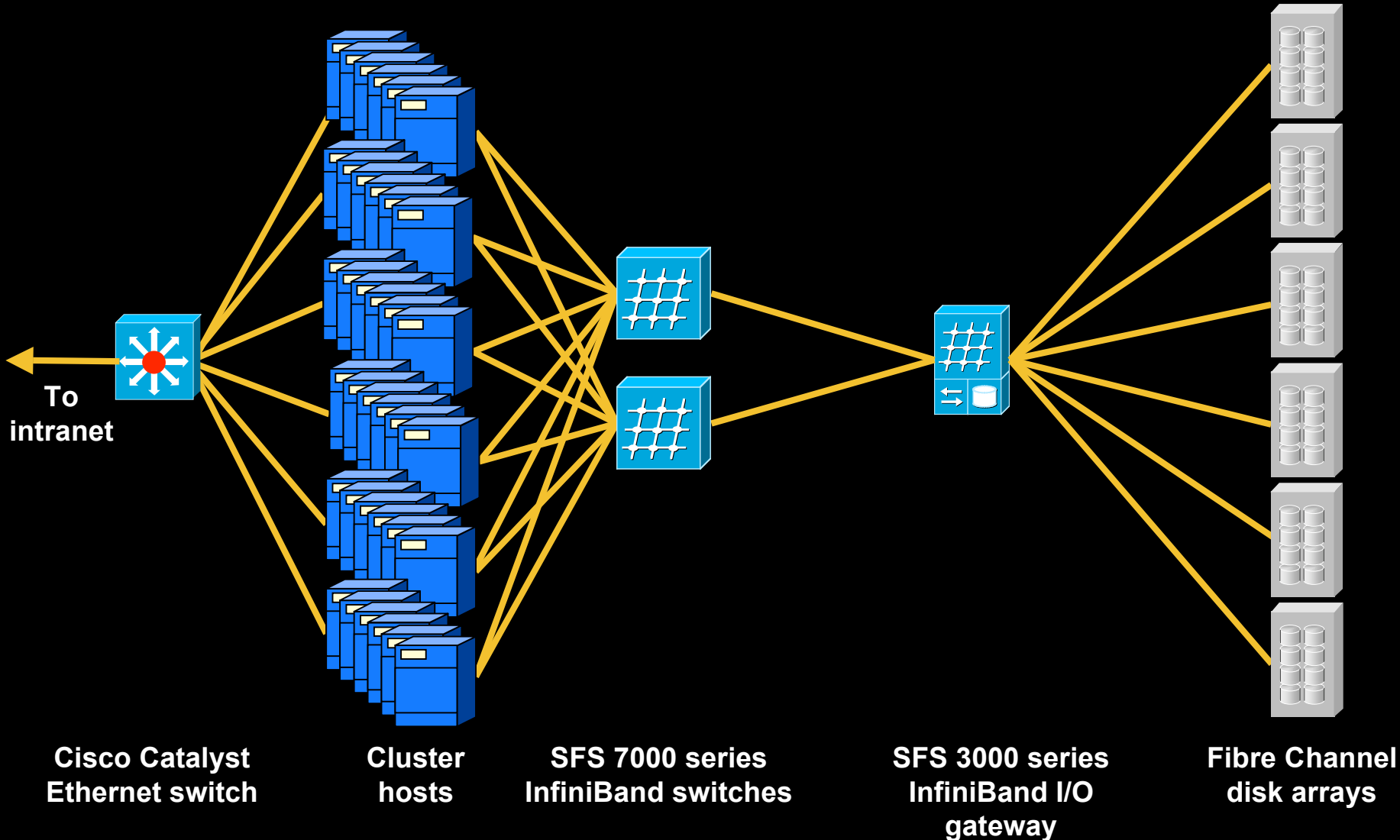
SFS 7012D

SFS 7024D

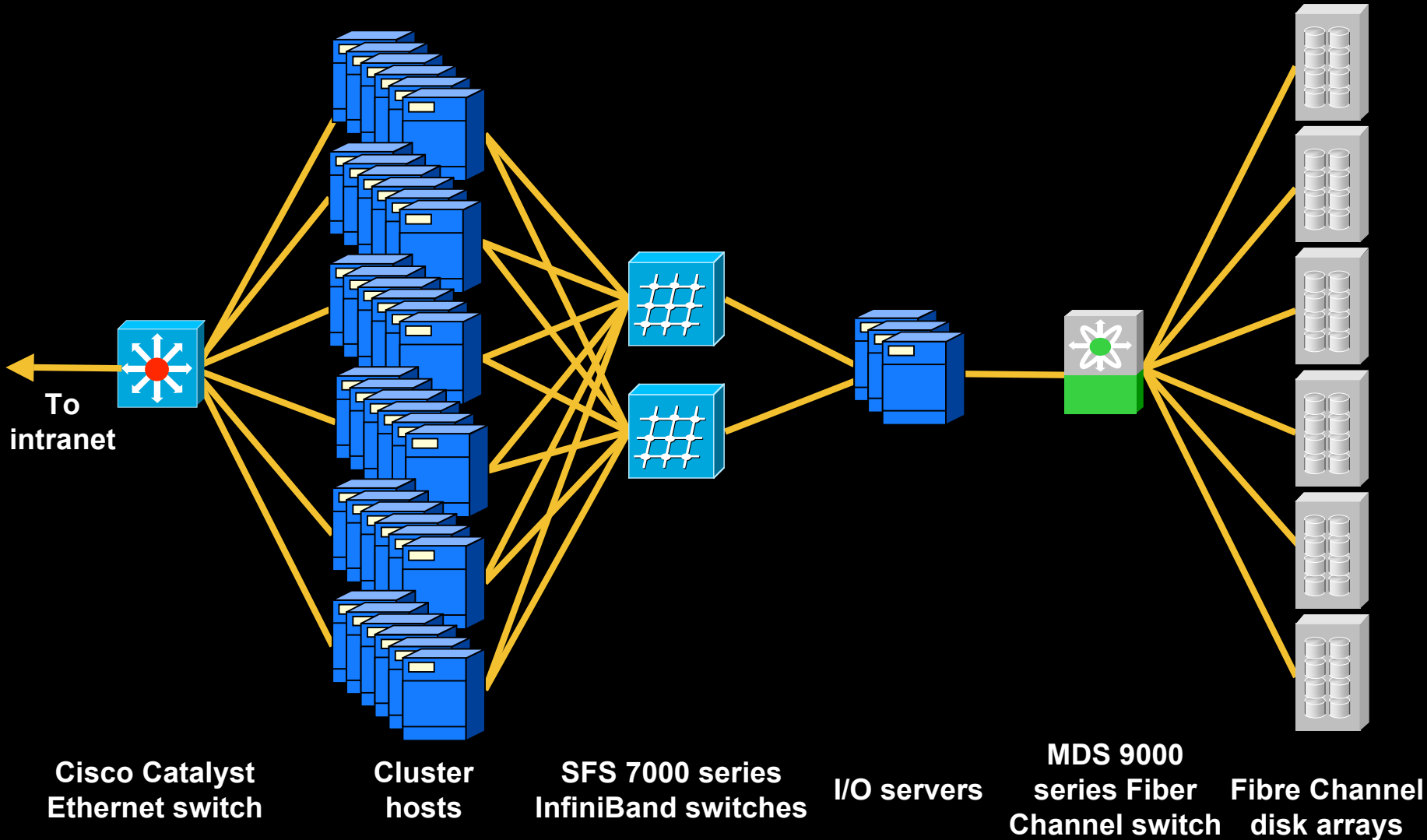| Chassis Type	7U Modular	14U Modular
Max 4X ports	**144 ports**	**288 ports**
Link Speed	**DDR/SDR Auto Sensing**	**DDR/SDR Auto Sensing**
Port Module Options	**12 Side by Side Slots** **12 by 4X LIMs** **Powered Interfaces**	**24 Side by Side Slots** **12 by 4X LIMs** **Powered Interfaces**
High Availability	**Redundant Power/Cooling** **Redundant Management** **Hot Swappable FRUs**	**Redundant Power/Cooling** **Redundant Management** **Hot Swappable FRUs**
Subnet Manager	**High Perf Subnet Mgr**	**High Perf Subnet Mgr**
Best Use	▪ **48-144 node clusters**	▪ **145-288 node clusters** ▪ **Core switch for >384 node clusters**

SFS Subnet Manager

- **Embedded with scalability up to 1000 nodes**
  - **Runs on switch**

- **x86 appliance hosted with validated scalability to 4600 nodes**

- **Interoperable with other vendor's IB switches, HCA's etc**

- **Support for up to four SM instances per fabric**
  - **Transparent failover between SMs**

- **Manages mixed DDR/SDR fabrics route calculations and weights links based on link speed**

- **Can bring up 1000 node network in less than 60 seconds**

- **On-the-fly configuration changes with no fabric disruptions**

Cisco HPC Vision: Possibilities

To intranet

Cisco Catalyst Ethernet switch

Cluster hosts

SFS 7000 series InfiniBand switches

SFS 3000 series InfiniBand I/O gateway

Fibre Channel disk arrays

Cisco HPC Vision: Possibilities

To intranet

Cisco Catalyst Ethernet switch

Cluster hosts

SFS 7000 series InfiniBand switches

I/O servers

MDS 9000 series Fiber Channel switch

Fibre Channel disk arrays

Cisco HPC Partnerships

- **Cisco is aligned with Partners across HPC Ecosystem**

- **Application certification & testing with all major ISVs across**

- **Storage certifications with EMC, NetApp, etc**

- **Support all MPI implementations**

- **Cluster integration and management capability with HP XC, Scali, ROCKS, Warewulf and many more**

- **Parallel file system validation and testing**

Example ISV/Apps Developing on InfiniBand

Example Management and MPI Developers on InfiniBand

Cisco's HPC Open Source Roles

- OpenFabrics core contributor

    "Verbs" library development and support

    Leader in QA testing / bug reporting of all OFED releases

- Open MPI community member

    Active code development, support, and maintenance

    Hundreds of thousands of community and Cisco-specific nightly regression tests (see Ethan's talk later this morning)

- Cisco is fully committed to the HPC community

Cisco at Euro PVM/MPI

- Cisco has a strong product lineup

  …but we want to hear from real customers

- What do you want out of HPC products?

  Switches

  Host-side network drivers

  Network management tools

  MPI

  …?

- Please come to the open forum on Wednesday!