



# Open Source, Open Storage Technical Seminar

## Welcome Keynote



# Red Hat Storage Management

Integrated enterprise storage management based on open source technology

**Red Hat Korea**

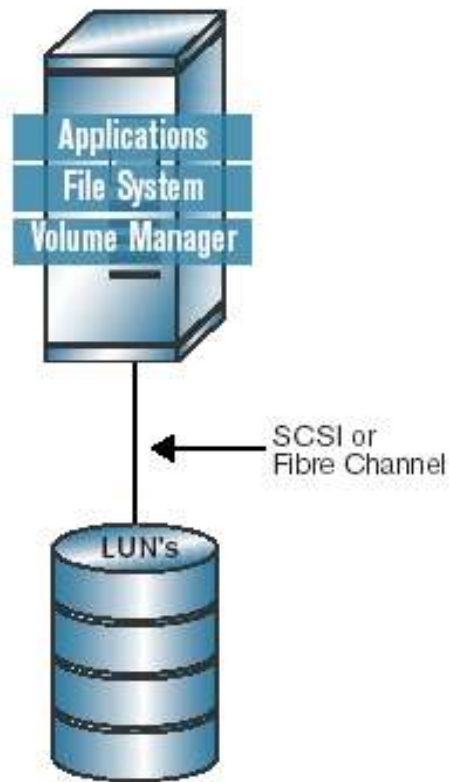
양승도 (syang@redhat.com)

**Sales Engineer**

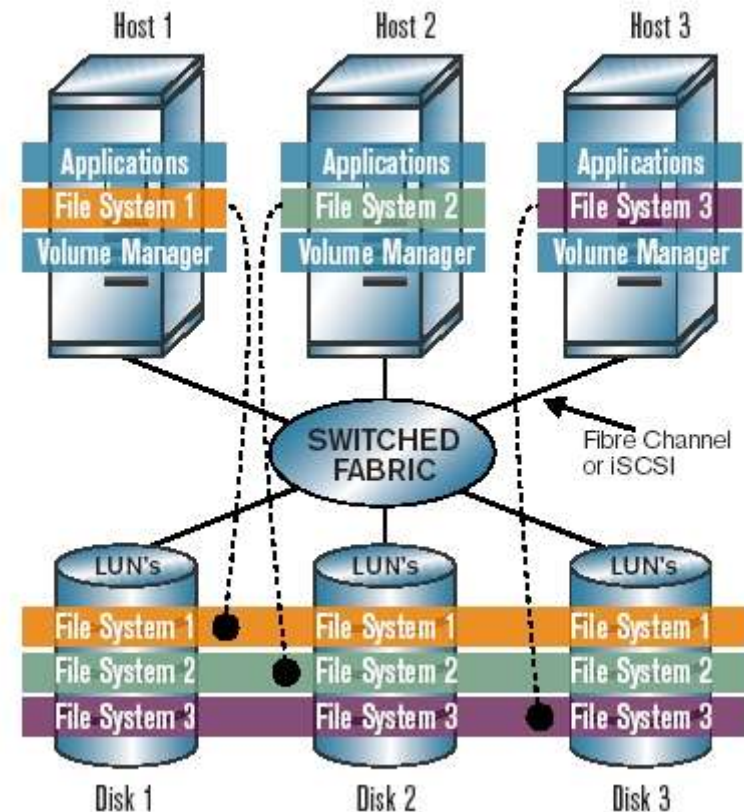
# Data Sharing Architecture

- Storage and Processor Topologies

## 1. Direct-Attached Storage

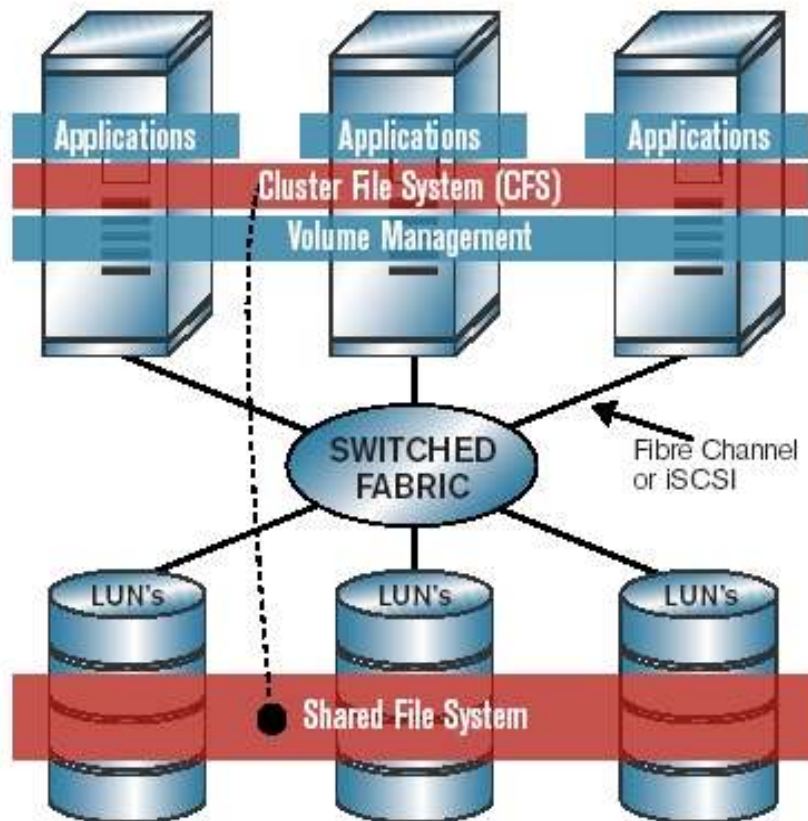


## 2. Storage Area Networks

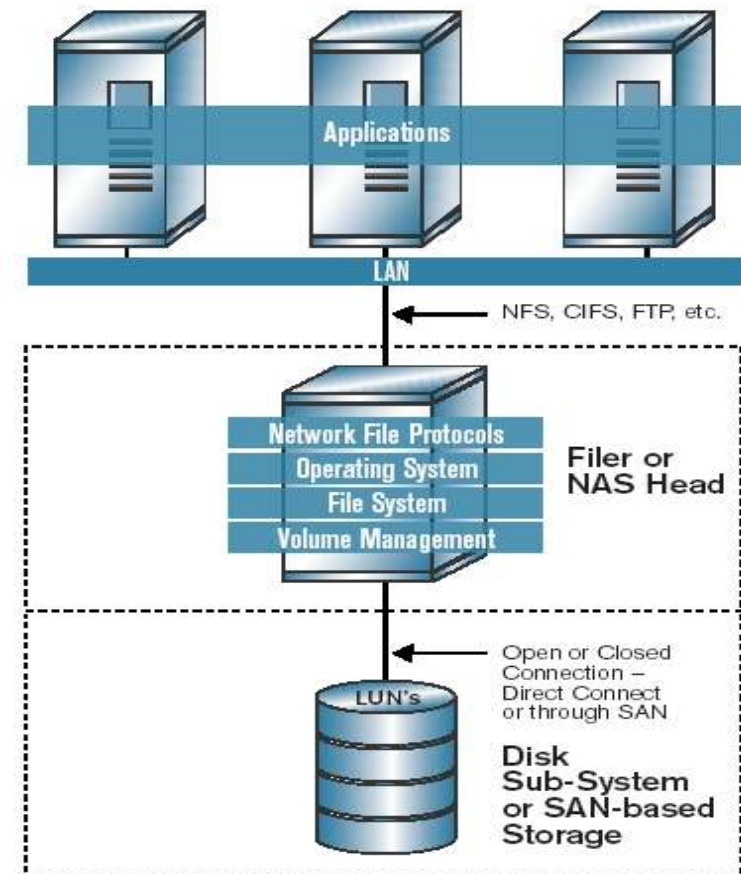


## ■ Storage and Processor Topologies

### 3. Processor Clusters



### 4. Network-Attached Storage





■ Comparison and Application

Requirements	DAS	SAN	Processor Clusters	NAS
General				
<b>High-Availability</b> Fault tolerance and planned downtime	Host availability achieved at high cost	Host availability achieved at high cost	Only requires redundant storage	NAS redundancy achieved at high cost
Price/Performance	Excellent	Good	Excellent	Fair
Manageability	Poor	Good	Good to Excellent	Excellent
<b>Openness and Interoperability</b>	Excellent	Good	Fair to Excellent	Excellent
Incremental Computing	Fair	Excellent	Excellent	Poor
Specific Applications				
Database	Excellent	Excellent	Excellent	Fair to Good
Email	Excellent	Excellent	Excellent	Good
Compute Intensive	Fair	Good	Excellent	Good
Web Serving	Poor	Fair	Excellent	Good

# Red Hat Storage Management

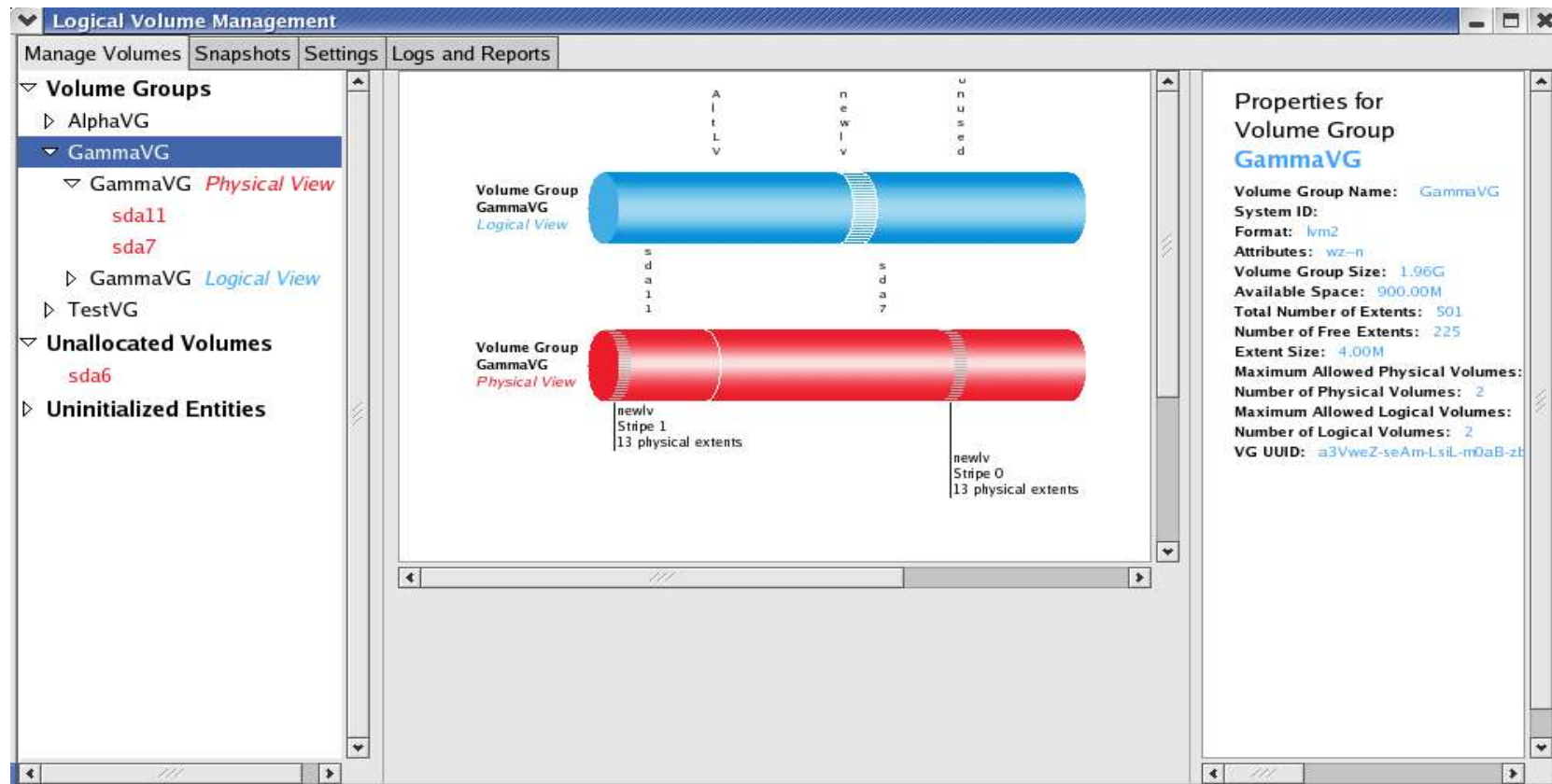
Integrated enterprise storage management based on open source technology

	<i>Best for...</i>	<i>Key technologies..</i>	<i>Key benefits...</i>
<b>Red Hat Enterprise Linux</b>	Single node integrated storage management	LVM, EXT3fs, NFS, AutoFS, Volume Mirroring, IO Multipathing,* iSCSI	Reduce dependency on expensive, complex 3 <sup>rd</sup> party storage management software & HBAs; fully integrated features of RHEL.
<b>Red Hat Cluster Suite</b>	Application failover and IP load balancing	HA clustering, IP Load balancing, Distributed Lock Manager	Low cost, integrated high availability for RHEL; a foundation of cluster services for multiple cluster configurations.
<b>Red Hat Global File System</b>	Enterprise clusters using shared storage	Cluster file system, CLVM, Red Hat Cluster Suite (included)	Share data across the cluster; Decrease storage/data management costs; Increase performance and scalability over NFS.

\*Available in future RHEL Updates

# Storage: Logical Volume Management (LVM2)

- LVM2 provides significantly improved GUI-based storage management capabilities
  - Goal to provide consistent, easy to understand, administrator interface



The screenshot displays the Logical Volume Management (LVM2) GUI interface. The window title is "Logical Volume Management". The interface is divided into several sections:

- Left Panel (Volume Groups):**
  - Volume Groups
    - AlphaVG
    - GammaVG
      - GammaVG *Physical View*
        - sda11
        - sda7
      - GammaVG *Logical View*
    - TestVG
  - Unallocated Volumes
    - sda6
  - Uninitialized Entities

- Center Panel (Physical and Logical Views):**
- Volume Group GammaVG Logical View:** A blue cylinder representing the logical view. Above it, the text "A l t L V" is displayed. Below it, the physical volumes "s d a 1 1" and "s d a 7" are listed.
- Volume Group GammaVG Physical View:** A red cylinder representing the physical view. Below it, the physical stripes are listed: "newlv Stripe 1 13 physical extents" and "newlv Stripe 0 13 physical extents".
- Right Panel (Properties for Volume Group GammaVG):**
- Volume Group Name: GammaVG
- System ID:
- Format: lvm2
- Attributes: wz-n
- Volume Group Size: 1.96G
- Available Space: 900.00M
- Total Number of Extents: 501
- Number of Free Extents: 225
- Extent Size: 4.00M
- Maximum Allowed Physical Volumes: 2
- Number of Physical Volumes: 2
- Maximum Allowed Logical Volumes: 2
- Number of Logical Volumes: 2
- VG UUID: a3VweZ-seAm-LsiL-mDaB-zb

# Red Hat Cluster Suite v4 Overview

- Provides two major technologies
  - High Availability failover – suitable for unmodified applications
  - IP Load Balancing – enables network server farms to load share IP load
- New with Cluster Suite v4
  - Elimination of requirement for shared storage
    - Significantly reduces the cost of high availability clustering
    - **Shared Quorum partition is no longer required**
      - Service state, previously stored in Quorum partition, is now distributed across cluster
  - Online resource management modification
    - Allows services to be updated without shutting down (where possible)
  - Additional fencing agents

# Red Hat Global File System v6.1 Overview

- New version for Red Hat Enterprise Linux v.4
  - Uses new common cluster infrastructure in Red Hat Cluster Suite (included)
- Provides two major technologies
  - GFS cluster file system – concurrent file system access for database, web serving, NFS file serving, HPC, etc. environments
  - CLVM cluster logical volume manager
- Fully POSIX compliant
- Much faster fsck
  - Ported to GFS 6.0 on RHEL 3 Update 5
- Data and meta-data journaling (per-node journals, clusterwide recovery)
- Maximum filesize & file system size: 16TB with 32-bit systems, 8EB with 64-bit systems
- Supports file system expansion
- Supports SAN, iSCSI, GNBD
- Default use of new Distributed Lock Manager (DLM)
  - Client-server locking optional



# The Benefits (and Costs) of Scale-Out

Horizontal scaling (clustering, grid, etc.) is happening...

What are the benefits, and what are the costs?

# The clusters are coming...

- *20% of all servers shipped will be clustered by 2006. - Gartner*
  - *Linux clusters are growing at a cagr of 44% per year. - IDC*
  - *30%+ of Red Hat inquiries are about implementing clusters.*
- **Why? The benefits of horizontal scale-out are clear...**
- Cost effective performance (vs. scale up)
  - Better flexibility to scale and repurpose IT resources (servers & storage)
  - Increased vendor choice
  - Increased availability and uptime of key applications

***But at what cost?***



# Clustering: Three Common Types

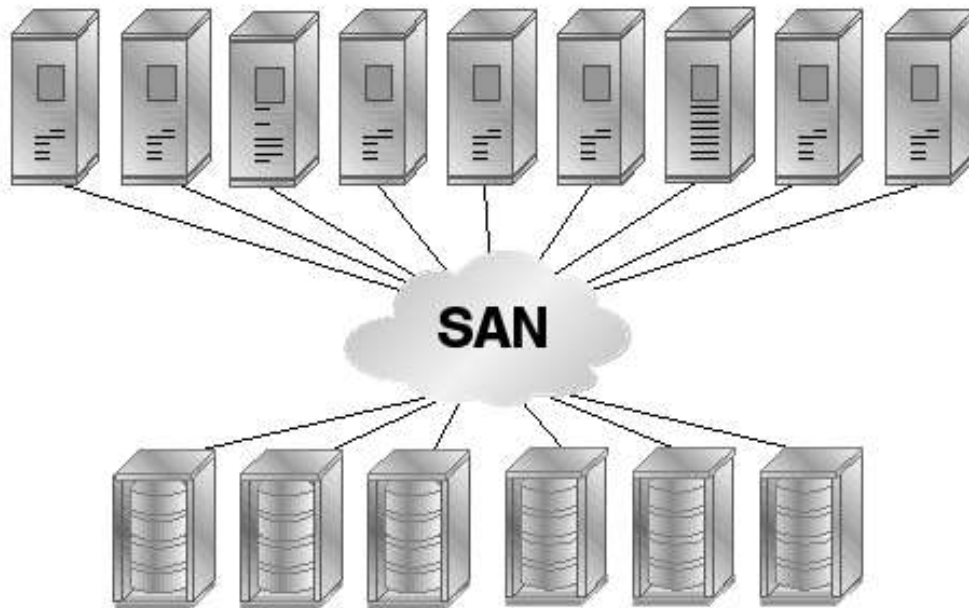
- **High Availability: Small clusters**
  - Goal: Application Uptime, Load balancing
  - Hardware Configuration: Small clusters (2 servers and up)
  - Applications: Any... often web applications, application servers, databases
- **Enterprise: Small-medium clusters**
  - Goal: Unix-Linux migration, scalability of enterprise applications
  - Hardware Configuration: Small-medium clusters (typically 4-32 servers)
  - Applications: Oracle RAC, Application Servers
- **High Performance (HPCC): Medium-large clusters**
  - Goal: High performance, data intensive apps (supercomputing)
  - Hardware Configuration: Medium-large clusters (10s to 100's of nodes)
  - Applications: Scientific, DCC, EDA, Oil & Gas, etc.

# The costs of horizontal scale-out

- **Increased complexity... why?**
  - Proliferation of hardware and software choices
  - Implications of choices (especially software) are far-reaching
  - Multiple potential points of failure
  - IT resource utilization hard to define and manage
  - Requires increased storage capacity (growing at 40-60% annually)
  
- **Increased management costs... why?**
  - Set-up, provisioning and management tasks are multiplied
  - Software deployment and update tasks are multiplied
  - Data synchronization, back-up and recovery tasks are multiplied

# What if you could...

**Manage this...**



**Like this...**



## You would be able to...

- Install and patch applications **once**, for the entire cluster
- **Decrease or eliminate** the need for redundant copies of application data (data replication)
- Greatly **simplify** back-up and recovery tasks
- **Manage** your storage resources as a whole vs. by partition

### *The result?*

- Mitigate the impact of **complexity** in the architecture, control **management effort and costs**, and *maximize the benefits of scaling-out.*

***Red Hat technologies can help you get there today***



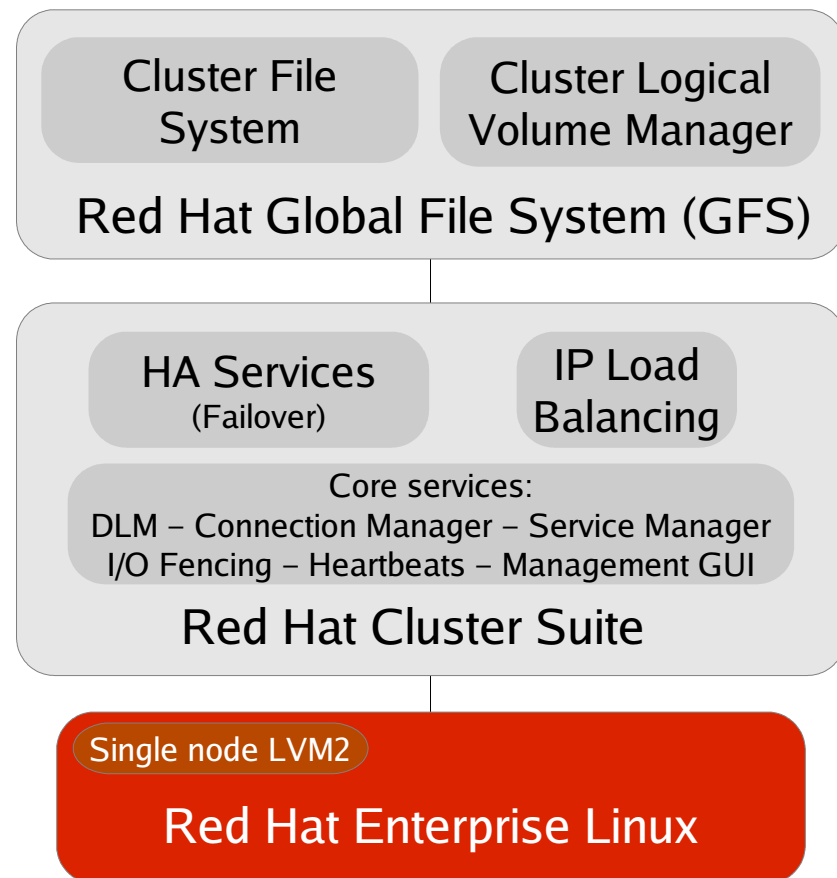
**Red Hat Cluster Suite**

# **Red Hat Global File System (GFS)**

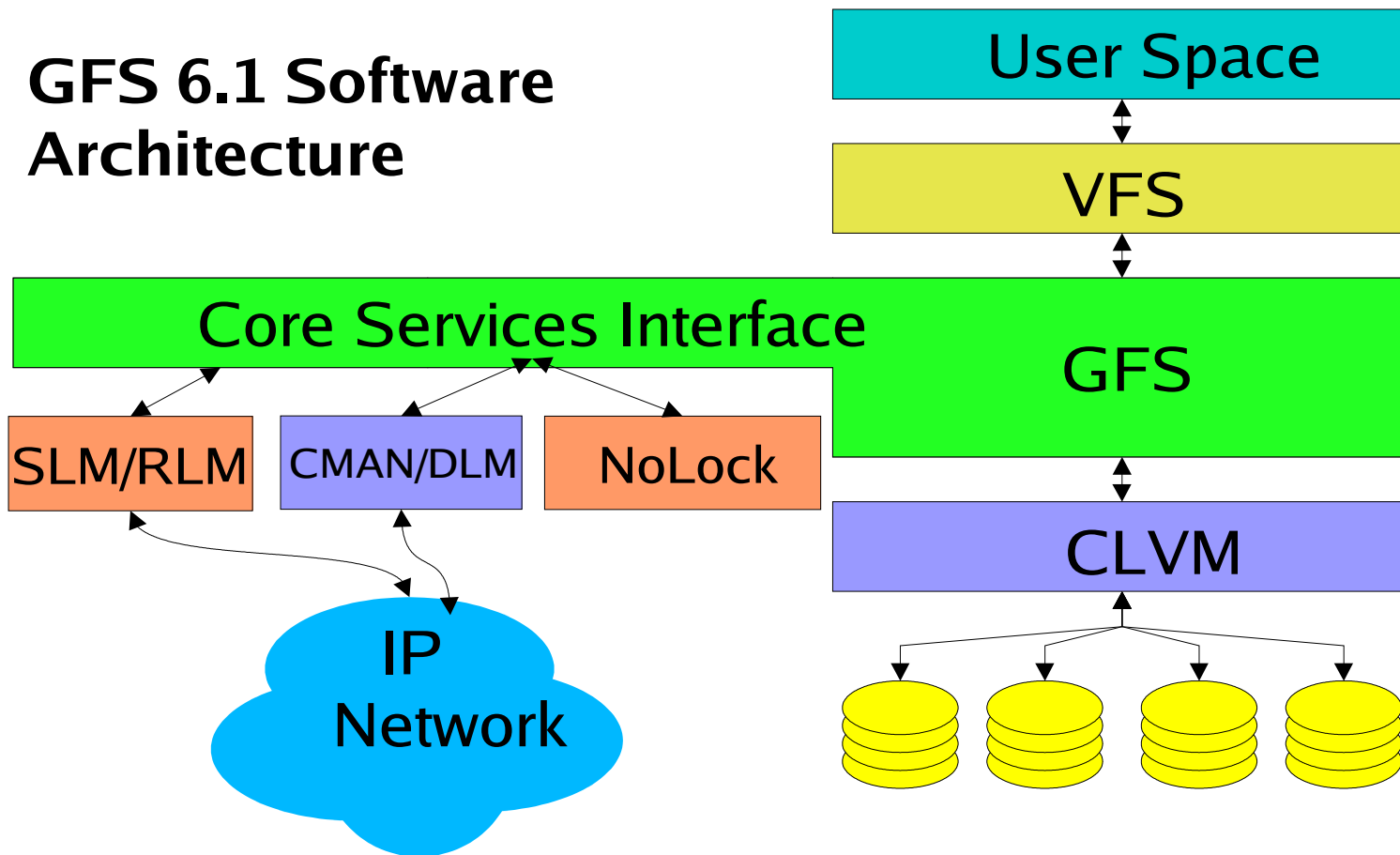
Open source technology for clustering and storage management

# An integrated clustering architecture

- **Red Hat Cluster Suite** provides
  - Application failover & IP load balancing
  - Improves application availability
  - Core services for enterprise cluster configurations
- **Red Hat Global File System (GFS)** provides
  - Cluster-wide file system
  - Concurrent read-write of data
  - Improved cluster availability, scalability and performance
  - Cluster Logical Volume Manager (CLVM)



# Red Hat Global File System v.6.1





# Distributed Lock Manager

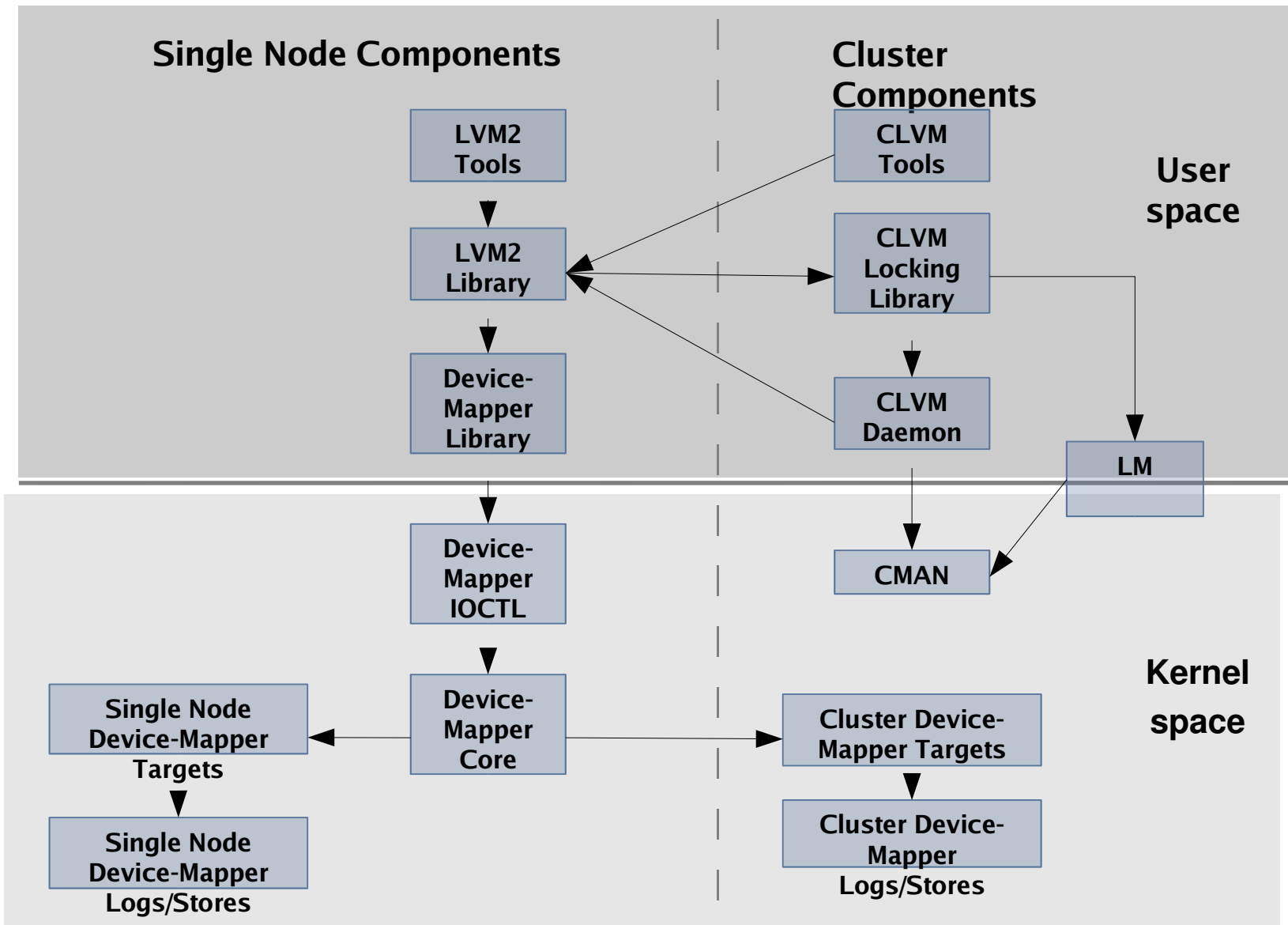
- 32 byte lock value block provides a way to propagate a small amount of context around the cluster as locks are granted and released
  - e.g. sequence/ID numbers
- Provides a set of asynchronous & synchronous APIs
  - `dlm_lock()`
  - `dlm_unlock()`
  - `dlm_query()`
- Locking modes:
  - NL: Null
  - CR: Concurrent read
  - CW: Concurrent write
  - PR: Protected read (“share”)
  - PW: Protected write (“update”)
  - EX: Exclusive
- Designed to use minimal & scalable intra-cluster messaging

Mode of Requested Lock	Mode of Currently Granted Lock					
	NL	CR	CW	PR	PW	EX
NL	Yes	Yes	Yes	Yes	Yes	Yes
CR	Yes	Yes	Yes	Yes	Yes	No
CW	Yes	Yes	Yes	No	No	No
PR	Yes	Yes	No	Yes	No	No
PW	Yes	Yes	No	No	No	No
EX	Yes	No	No	No	No	No

# Cluster Logical Volume Manager (CLVM)

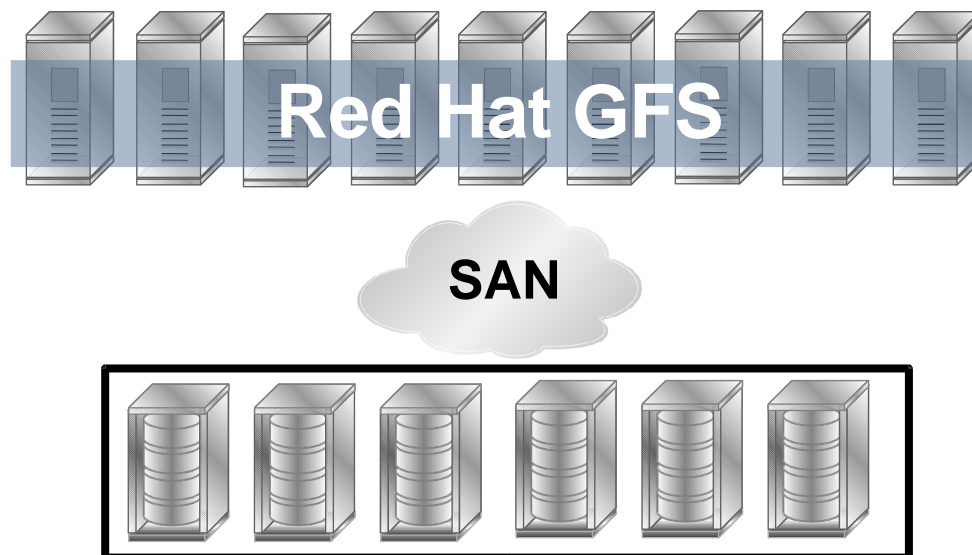
- CLVM builds upon LVM 2.0 and the kernel device mapper component included in 2.5 and 2.6 Linux distributions
- Essentially a cluster-aware version of LVM 2.x
- Commands, features, functions all work just fine in a cluster, any Linux server may mount any volume
- Provides
  - Cluster safe volume operations
  - Cluster-wide concatenation and stripping of volumes
  - Dynamic Volume resizing
  - Cluster-wide snapshots (mid '05)
  - Cluster-wide mirroring (mid '05)
  - Other RAID levels (mid '05)

# CLVM High-Level System Architecture



# Red Hat Global File System (GFS)

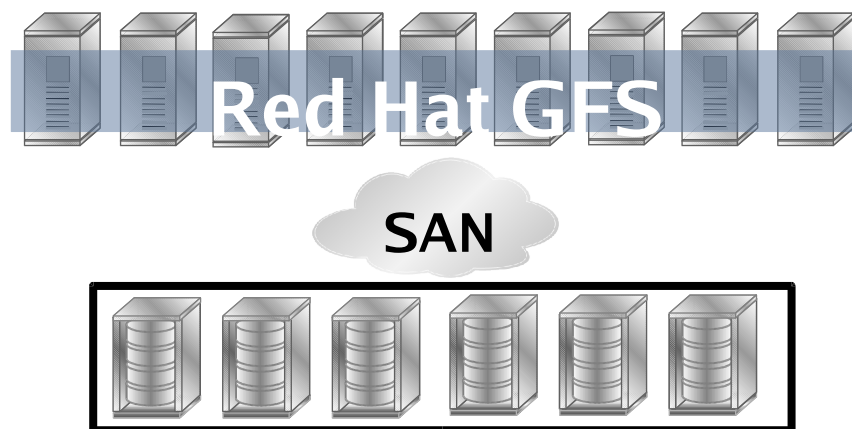
- Allows a **cluster of Linux servers** to **share data** in a **common pool of storage**



*“The main attraction – and, frankly, the original promise – of storage networking is the ability to connect multiple systems to a common pool of storage.” -  
Illuminata, Sept. 2004*

# What is Red Hat GFS?

- Red Hat GFS is an open source, POSIX-compliant **cluster file system**.
- It provides a **consistent file system image across the server nodes** in a cluster, allowing Red Hat Enterprise Linux servers to **simultaneously read and write to a single shared filesystem**.
- Red Hat GFS also includes **integrated application high availability and failover** with the included Red Hat Cluster Suite.



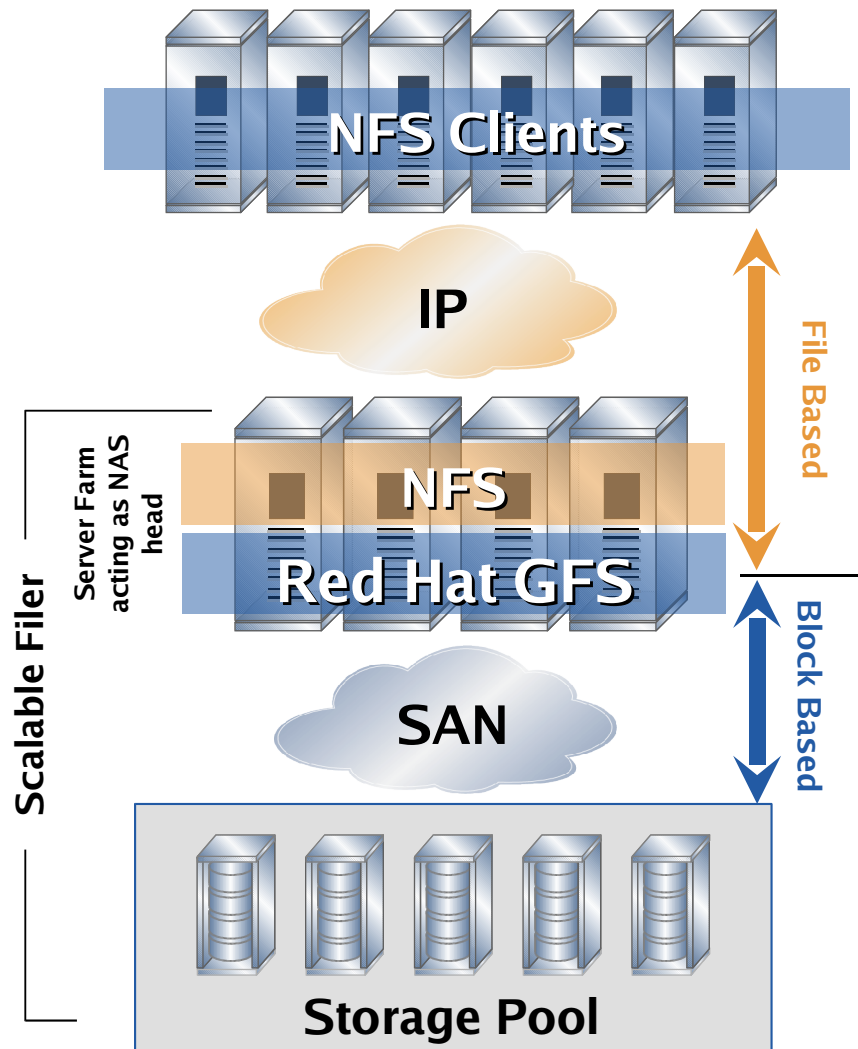
# Red Hat GFS: The best cluster file system for Linux

- The **only native 64-bit** general purpose Cluster File System for Linux – support for x86, IA64, AMD64, and EM64T
- **Most scalable Cluster File System on Linux** – supported up to 256 nodes (more than Veritas (32) or PolyServe (16))
- **Tightly integrated with Red Hat Enterprise Linux** (no patching, no lag from RHEL updates)
- **Proven** in hundreds of production environments with varied applications
- **Only open source (GPL) general purpose cluster file system**, headed for upstream adoption:

*“Significant features will be added [to the kernel] in the next year, including NFS 4 and clustering file support, possibly Red Hat's Global File System technology.”*

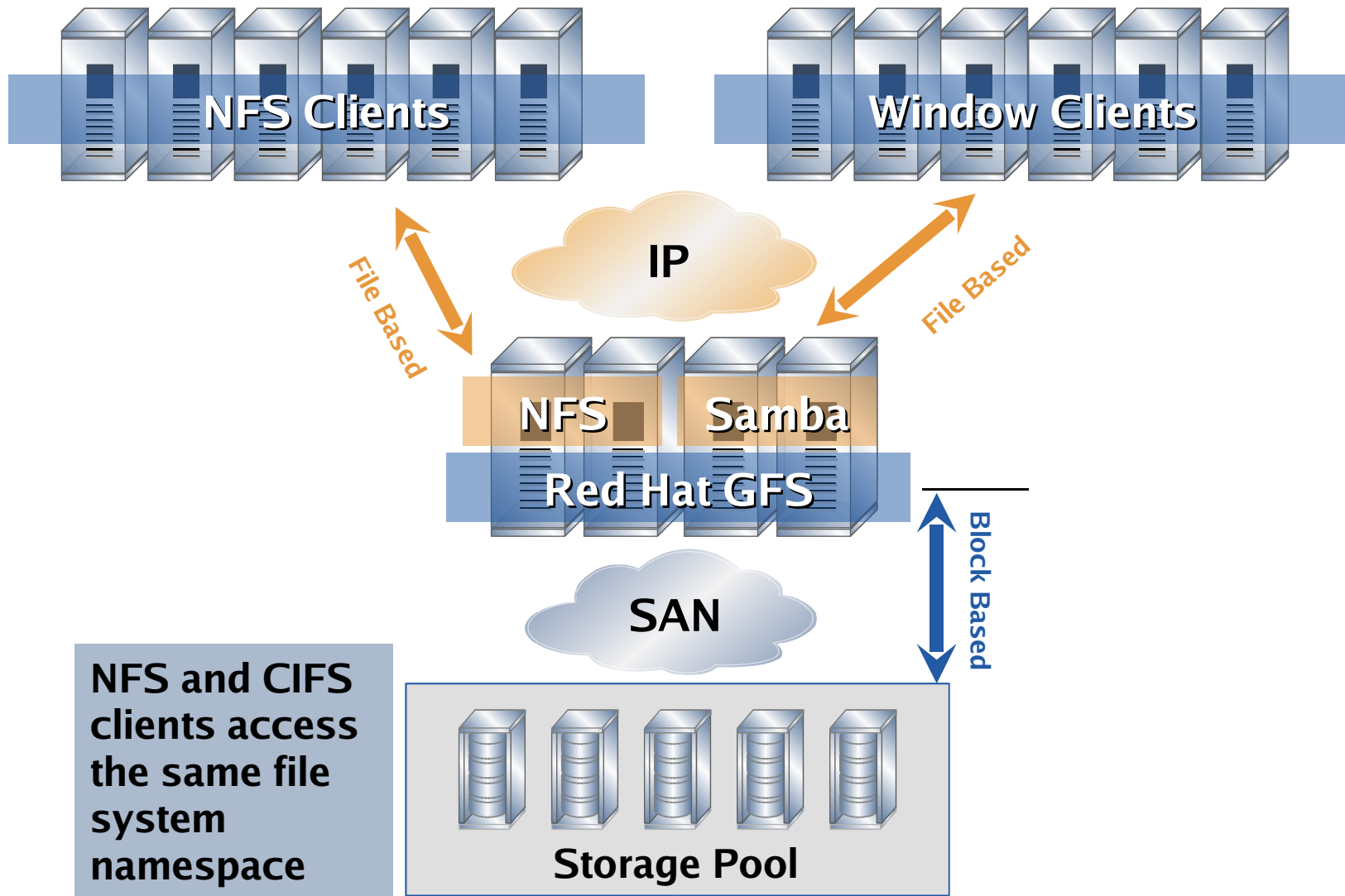
*- Andrew Morton, 2.6 Kernel Maintainer, at Feb. 2005 OSDL Summit*

# Red Hat GFS can make NFS scale & perform



- Eliminate NFS performance and scalability limitations
- Create a scalable NAS-like cluster with no single point of failure
- Dynamically add compute and I/O resources
- Streamline development environments and accelerate build times
- Eliminate need for data duplication





# Common applications for Red Hat GFS

## ■ Enterprise Applications

- *Typical configuration:*
  - Cluster of Red Hat Enterprise Linux servers + Red Hat GFS + a Storage Area Network (SAN)
- *Applications:*
  - Application Servers (BEA WebLogic, IBM WebSphere, Red Hat App Server),
  - Web Servers (Apache)
  - Databases (Oracle, DB2, MySQL, etc.)
- *Benefits:*
  - Provision and update applications once for the entire cluster
  - Share application data (session, user, etc.) throughout the cluster

# Common applications for Red Hat GFS

## ■ Oracle RAC

- *Typical configuration:*
  - Cluster of Red Hat Enterprise Linux servers running Oracle RAC + Red Hat GFS + a Storage Area Network (SAN)
- *Benefits:*
  - Drastically simplify Oracle RAC installation and management
  - Perform one installation, with a single Oracle Shared Home and shared root
  - Expand and manage RAC storage space on the fly
  - Run all POSIX compliant applications on RAC nodes

# Common applications for Red Hat GFS

## ■ **Scientific, Data Analysis, Oil & Gas, EDA, DCC**

- *Typical configuration:*

- HPC Cluster (NFS Clients) + NFS Server Cluster (RHEL servers with Red Hat GFS) + a Storage Area Network (SAN)

- *Benefits:*

- Eliminate NFS performance and scalability limitations
- Create a scalable NFS server cluster with no single point of failure
- Dynamically add compute and I/O resources to the cluster

# Red Hat Storage Training & Services

## ■ Red Hat Global Training

- **Red Hat 436: Red Hat Enterprise Storage Management**
- Intensive hands-on experience with Red Hat Global File System (GFS)
- 4-day course focuses on the implementation of native Red Hat Enterprise Linux technologies included in Cluster Suite and GFS

## ■ Red Hat Professional Services

- **Rapid Solution Service:** Installation, Configuration, Troubleshooting, & Knowledge Transfer – per day, as much or as little as you need
- **Red Hat Professional Consulting:** Specific consulting engagements to accelerate and enhance your adoption of open source, including:
  - HA Clustering
  - HPC Clustering
  - Enterprise Storage Architecture



# Red Hat Storage Management Proof Points

Real-world customer successes with Red Hat Cluster Suite and Red Hat GFS

# Customer Proof Point: Polycom

- *Problem:* Polycom's software development team was experiencing increasingly slower build times as demand for the internal build system increased
- *Solution:* Replaced an existing Unix SMP environment with a cluster of Red Hat Enterprise Linux servers running Red Hat GFS connected to a SAN

- ***Benefits:***

- Significant acceleration in build times
- Able to simultaneously distribute content over Telnet, FTP, SSH and HTTP without impacting development
- Provides reliable and scalable file services via NFS and Samba
- Eases backup with simultaneous read only access of the file system

<i>Process (Compiles)</i>	<i>SMP Unix (Proprietary)</i>	<i>Red Hat GFS</i>
Single Build	14.5 minutes	2.5 minutes
2 Simultaneous Builds	30.0 minutes	2.6 minutes
4 Simultaneous Builds	40.0 minutes	2.7 minutes
8 Simultaneous Builds	60.0 minutes	2.7 minutes
16 Simultaneous Builds	120 minutes	2.8 minutes



# Customer Proof Point: Sprint ATL

- *Problem:* The very large data files collected from the Sprint backbone infrastructure for analysis were extremely slow to access by researchers via NFS (Sun). In addition, researchers had to replicate the data over the network creating a significant burden on administration resources to backup and synchronize the data.
- *Solution:* Replaced an existing Unix SMP environment with a cluster of Red Hat Enterprise Linux servers running Red Hat GFS connected to a SAN
- *Benefits:*
  - Provide reliable and rapidly scalable infrastructure to handle large files
  - Provided high performance application scaling
  - Reduced or eliminated data replication and data movement over the network
  - Addressed NFS limitations of performance, scalability and reliability

# Customer Proof Point: White Cap

## ■ *Problem:*

- Performance of e-commerce website slowed due to data replication issues.
- Build environment for developers required significant management effort to keep all data and sources in sync.

## ■ *Solution:*

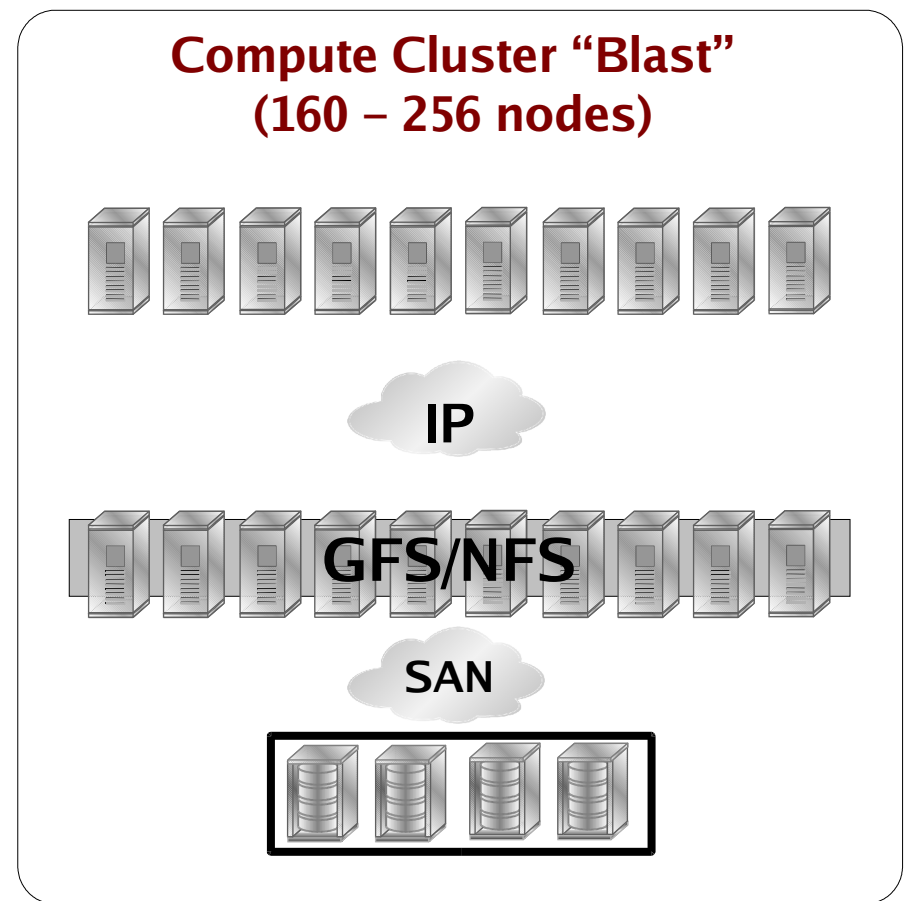
- Deployed a single cluster using RHEL3/GFS 6.0 for managing the e-commerce site and a single source build depository.
- 2 GFS nodes for App Servers, 2 GFS nodes for Web Servers, 8TB storage, Cisco 9509 switch, Dual fibre, HP Servers, Qlogic HBA

## ■ *Benefits:*

- Applications can be developed and tested without data movement, and easily pushed live when ready
- Scalability to address growing web presence, with traffic doubling every year

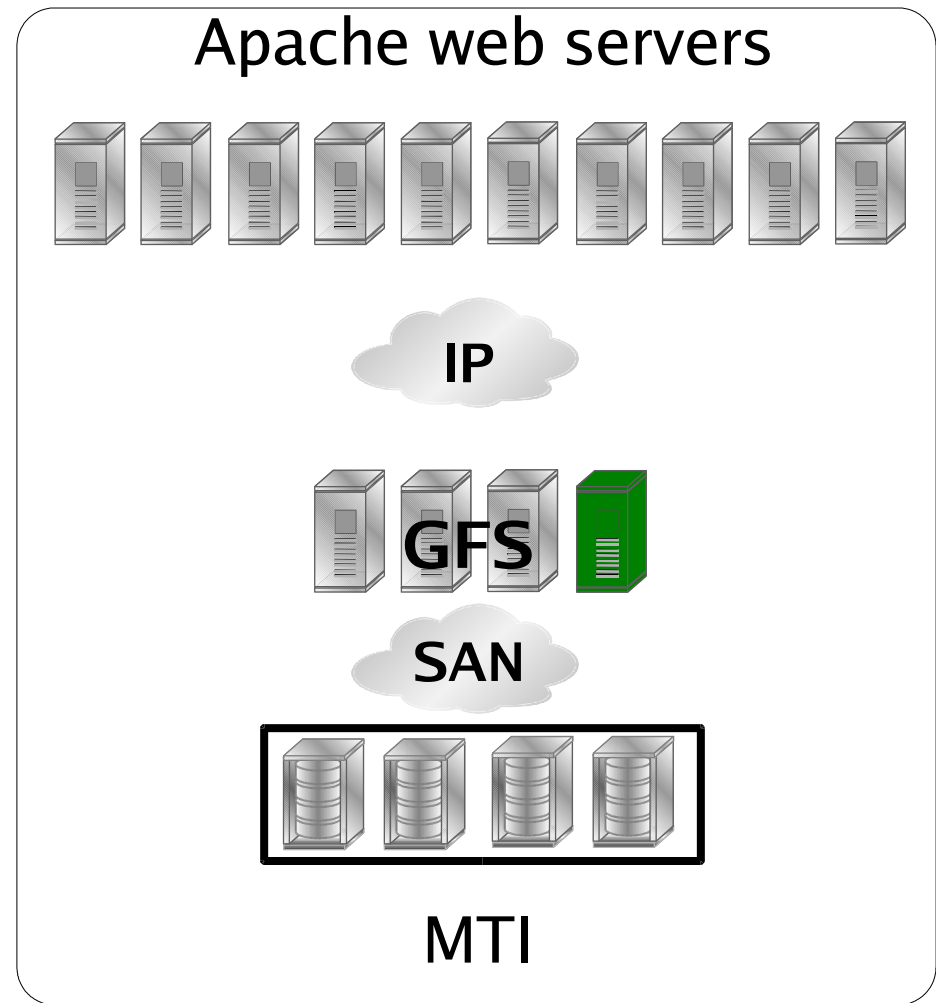
# Customer Proof Point: Cengent Therapeutics

- *Problem:* NFS didn't scale or perform for large compute application “Blast”
- *Solution:* Deployed scalable NFS file serving via a RHEL/GFS cluster
- **Benefits:**
  - 2 years of reliable and scalable file serving farm (NFS)
  - Grew from 5 GFS nodes to 10 GFS nodes serving 128 to 256 peak compute cluster requirements. Currently 160 compute nodes.
  - Reduction in cost using standards based hardware for servers and storage: Dual Pentium CPU, Brocade fabric, IBM FastT storage



# Customer Proof Point: DHDMedia (Web Content)

- **Problem:** Need to address 10 million web transactions a day for content, both static and dynamic (streaming) and transactions, with minimal writes and mostly reads.
- **Solution:** A clustered environment using GFS for content delivery and transactions. Four GFS nodes implemented to support 10 web servers. MTI Storage, Brocade switches, Dual Zeons, Qlogic HBAs (1GB).
- **Benefits:**
  - In deployment for 2 years
  - GFS cluster handles small and large files efficiently (20MB/sec)



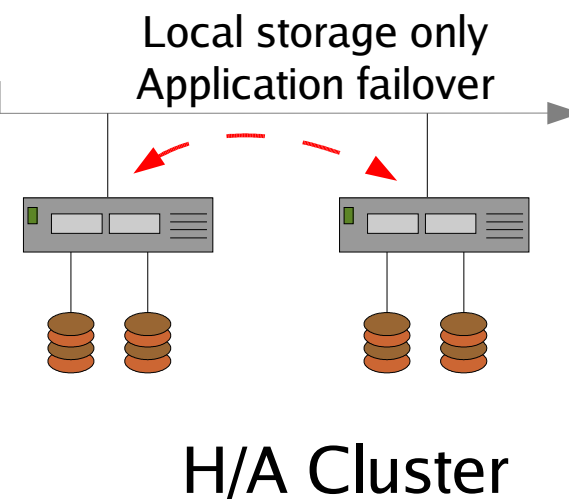


# Red Hat Storage Management Use Cases

Sample storage architectures using Red Hat Storage Management technology

# Use Case: High availability with local data

- Red Hat Enterprise Linux (ext3) + Red Hat Cluster Suite v4
- Requires NO ADDITIONAL HARDWARE
  - No physically shared storage
    - Prior to RHEL4, physically shared storage was required
- Ensures that an application stays running
  - Monitors all cluster nodes
  - Fails-over applications/services from stopped nodes
  - Simple management GUI
  - Service definitions are automatically propagated and synchronized, cluster-wide



## Topology:

- 2 or more nodes
- NIC card
- NO shared storage

## Applications:

- Read-only data
- Small web serving
- NFS/FTP serving
- Edge of network

## Data model:

- Unshared, static data
- Replication w/ rsync
- ext3 file system

# Use Case: Shared file datastore, NAS

- Red Hat Enterprise Linux (NFS)
- Uses an external NFS file server

## Addnl. Cluster features:

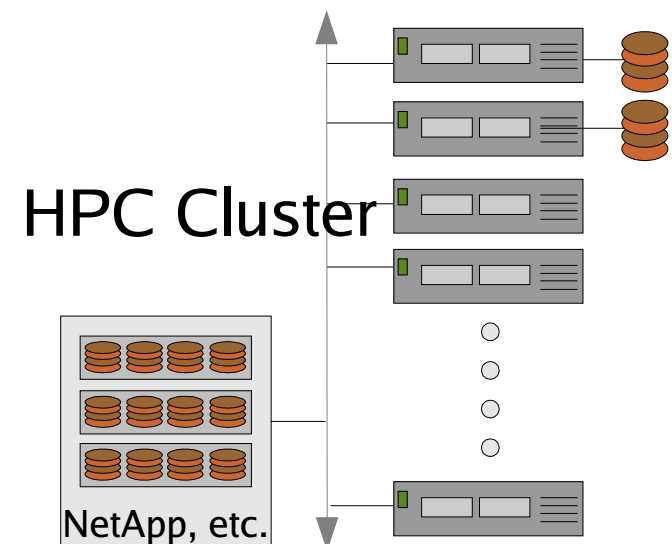
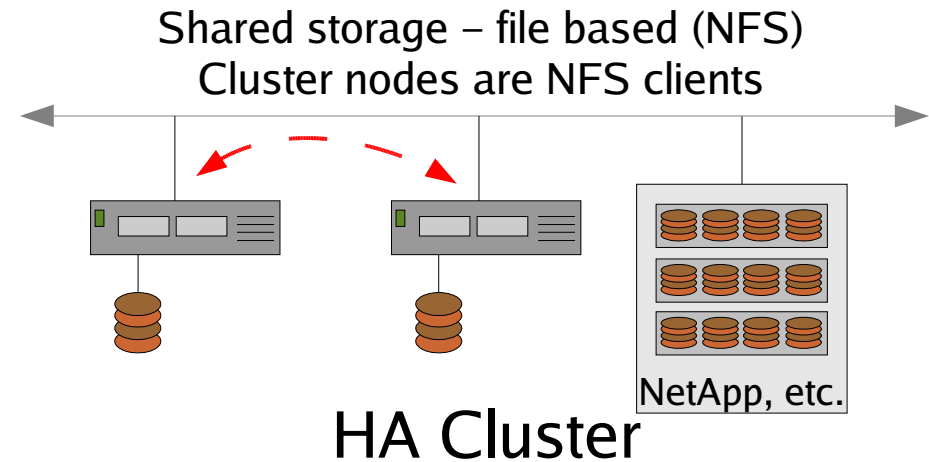
- Increased storage capacity
- Applications can failover and access the same data as before
- NFS sharing model

## Applications:

- Small/medium web serving
- NFS/FTP serving
- Read/write data
- File level access

## Data model:

- Physically shared data
- File based (NFS)



# Use Case: Shared block datastore, NAS

- Red Hat Enterprise Linux (ext3, iSCSI)
- Uses an external iSCSI block server

## Addnl. Cluster features:

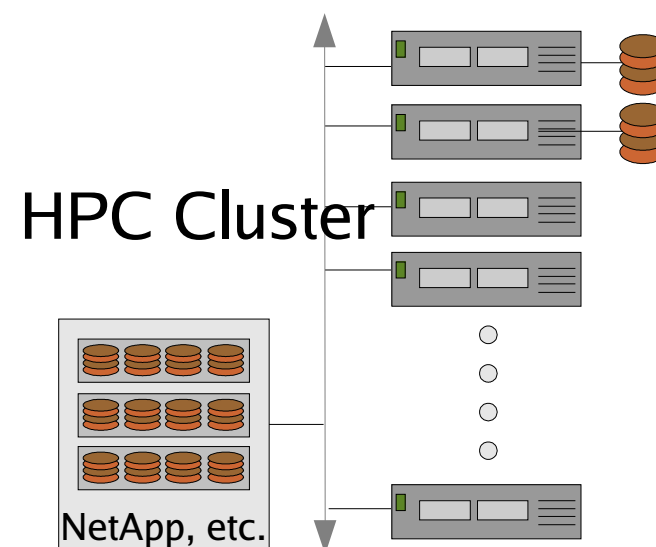
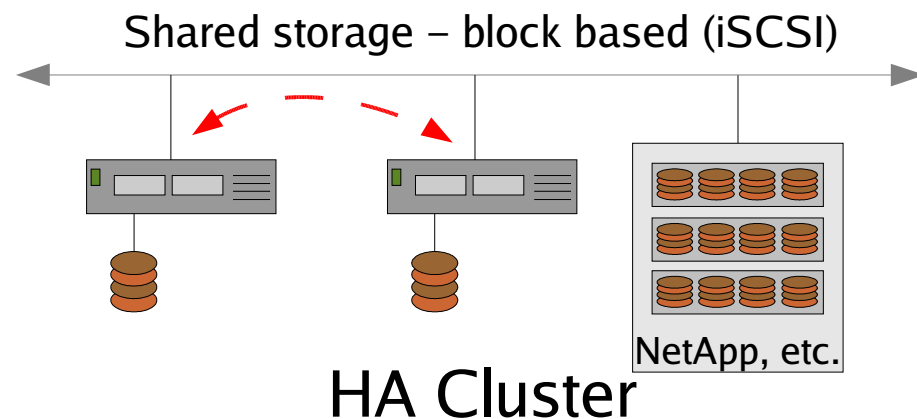
- Block based access provides improved performance
- Replication required if multiple nodes need to access the same data

## Applications:

- Medium web serving
- Small/medium database
- Read/write data
- Block level access

## Data models:

- Physically shared storage
- Block based (higher performance)
- ext3 file system





# Use Case: Shared block datastore, SAN

- Red Hat Enterprise Linux (ext3, integrated HBA drivers)
- Offers higher SAN-based, performance

## Addnl. Cluster features:

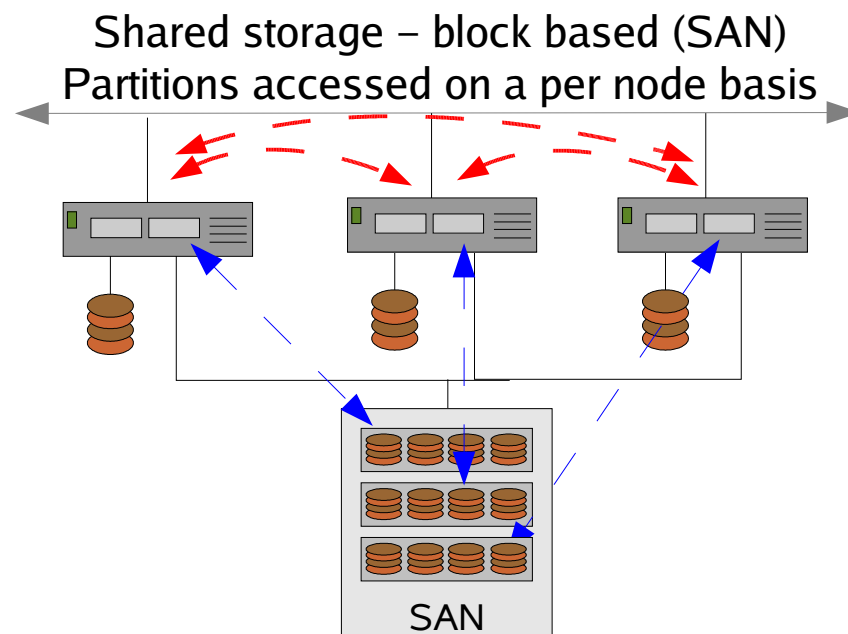
- SAN access provides:
- improved performance
- heterogeneous access by other systems

## Applications:

- Medium/large web serving
- Medium database
- Read/write data
- Block level access

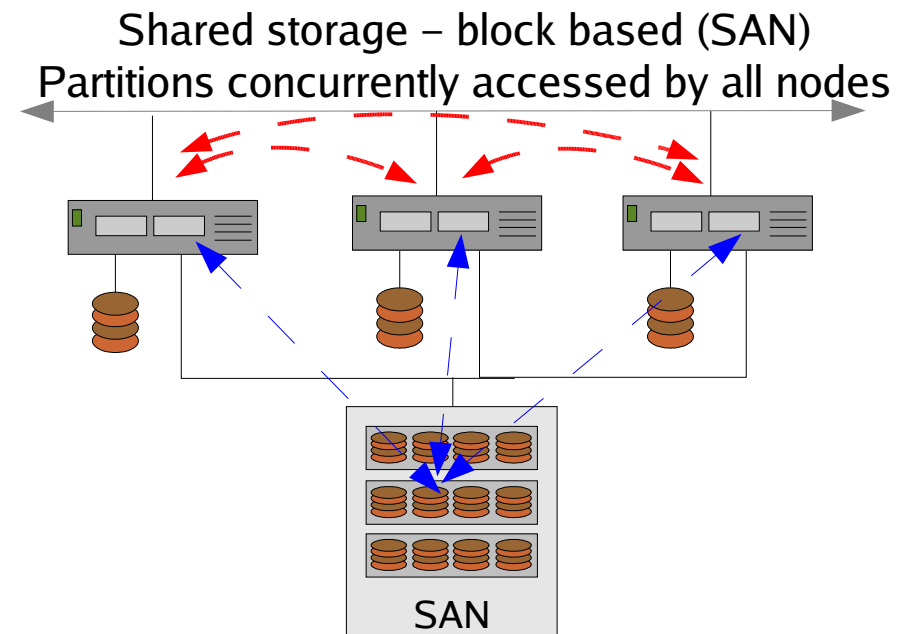
## Data models:

- Storage Area Network
- Physically shared storage
- Block based
- ext3 file system



# Use Case: Concurrent file system access

- Red Hat Enterprise Linux + Red Hat Global File System
  - Includes Red Hat Cluster Suite
- Shared file system access
- Scalable performance
- Same hardware as previous configurations
- Common software infrastructure with previous configuration



## Addnl. Cluster features:

- SAN access provides:
  - improved performance
  - heterogeneous access by other systems

## Applications:

- Large web serving
- Large database
- Read/write data
- Block level access
- Concurrent access

## Data models:

- Storage Area Network
- Shared file system access
- Block based
- GFS file system

# Red Hat GFS Advanced Features

	<i>Helps with...</i>	<i>Details...</i>
<b>256 node support</b>	Scale and performance	Support up to 256 nodes in a single cluster (8x scalability of nearest competitor)
<b>Native 32-bit or 64-bit</b>	Performance	Support natively on x86, Itanium, AMD64 & EM64T (only native 64-bit general purpose cluster file system on Linux)
<b>Quota System for Storage Utilization</b>	Management	Quota system included for cluster-wide storage capacity management
<b>Direct I/O Support</b>	Performance	Allows distributed databases to achieve raw-storage-like performance with the management benefits of a filesystem
<b>Dynamic Multi-Pathing</b>	Data availability	Ensures data availability by routing around switch or HBA failures in the SAN
<b>Online resizing</b>	Management	Add storage capacity while the file system is online and available



# Red Hat Global File System (GFS) Red Hat Cluster Suite

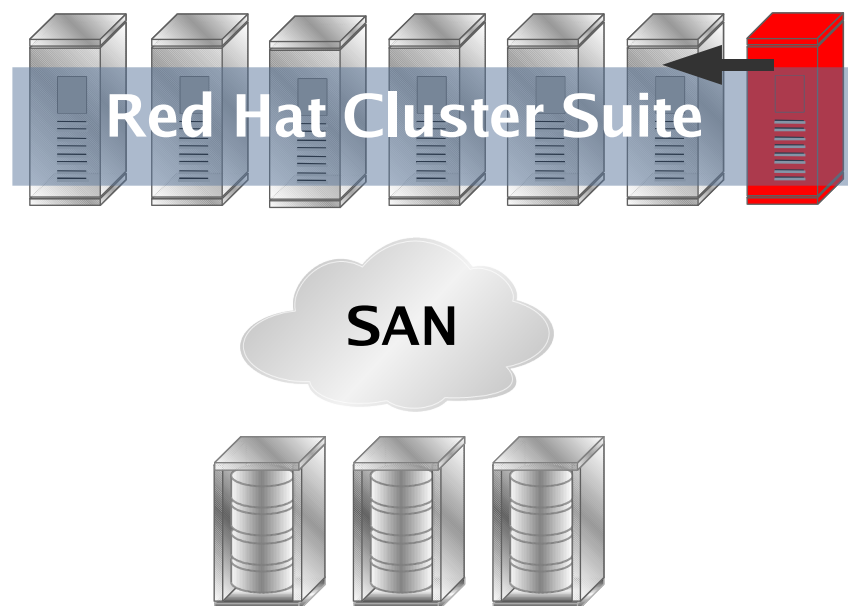
High Availability and Load Balancing

# Red Hat Cluster Suite

- Key Benefits
  - Increased application availability
  - Increased performance & scalability
- Can be configured in NSPF configurations
- Red Hat Cluster Suite
  - Red Hat Cluster Manager - Cold Failover
  - Piranha - Load Balancing

# Red Hat Cluster Suite

- Low-cost **high availability** for applications
  - Create n-node server clusters for desired level of availability
  - In the event of a failure, workload is picked-up by other servers in the cluster
- **Core services** for enterprise cluster configurations (with v4)
  - Distributed Lock Manager, Service Manager, I/O Fencing, Heartbeats, GUI



# History of Red Hat's HA Offerings

- Red Hat High Availability Server 1.0 (2000)
  - Introduced *Piranha* to customers
  - Red Hat's first enterprise-targeted offering
- Red Hat Enterprise Linux v2.1 (2002)
  - Introduced *Cluster Manager* to customers
    - Based heavily on Kimberlite, pioneered by Mission Critical Linux
    - Provided Cold Failover for off the shelf applications
    - Command-line based administration shell (*cluadmin*)
- Red Hat Cluster Suite v3 (2003)
  - For Red Hat Enterprise Linux v3
  - Cluster Manager gains multi-node support
  - New Python / GTK-based configuration utility (*redhat-config-cluster*)

# History of Red Hat's HA Offerings (cont)

- Red Hat Cluster Suite 4 (2005)
  - For Red Hat Enterprise Linux 4, of course
  - Cluster Manager modularized
    - Resource Agent model based on the Open Cluster Framework RA API draft
      - Fairly easy to write new resource scripts
    - Supports on-line reconfiguration of services
  - Cluster Manager integrated with CMAN and DLM
    - GFS and Cluster Manager have the same membership / quorum views
    - Updated UI for configuring cluster infrastructure and Cluster Manager



# On Piranha...

- Typically runs on two computers for redundancy, but can be run on one
- Controls IPVS via the ipvsadm command
  - Piranha is not required to make use of IPVS, but it makes life easier
- Provides
  - Linux IPVS director failover
  - Real Server monitoring via both expect/send scripts and simple port monitoring
  - Increased application availability
  - “Big Computer” abstraction: SSI at the IP level
- Includes a web-based GUI for configuring Virtual IP addresses, Real Servers, directors, etc.




# Piranha – Routing Modes

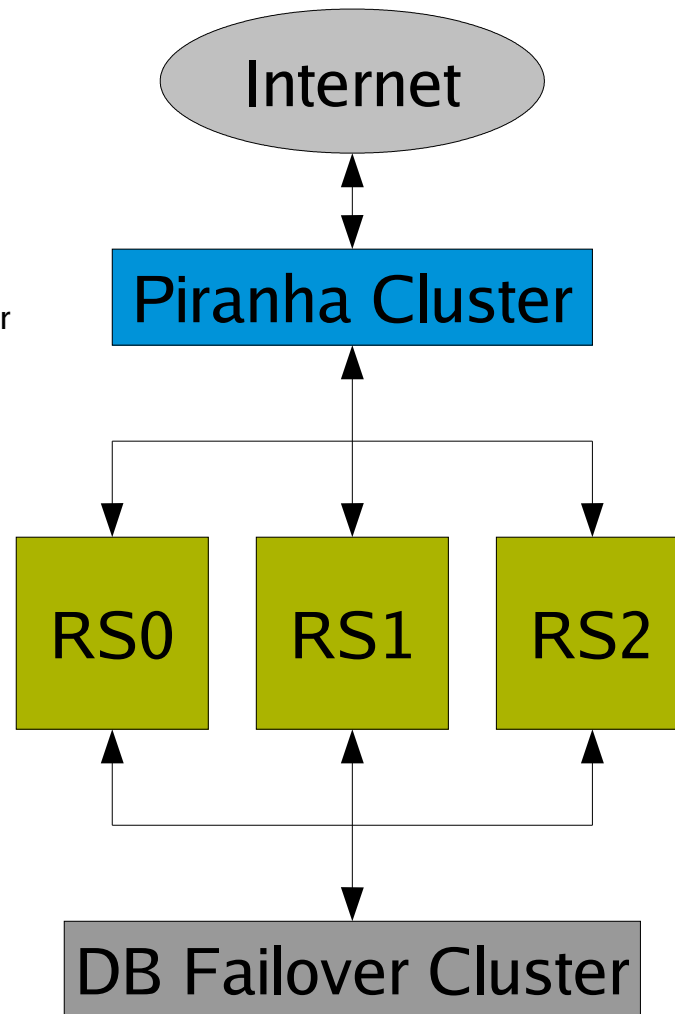
- NAT Routing
  - Exposes only the Piranha director(s) to the outside world
    - Allows real servers to be protected from all traffic except the type expected
  - Easy to configure (no additional steps necessary on real servers)
- Direct Routing
  - High performance
  - Exposes director(s) and real servers to the outside world
  - More difficult to configure (requires steps to be taken on each real server)
- Tunneling
  - TBD. Currently untested.

# Piranha Scheduling Algorithms

- *Round-Robin* – Distributes new requests in sequentially around the real servers
- *Weighted Round-Robin* – Distributes new requests round-robin, but gives more heavily-weighted machines more connections.
- *Least-Connections* – Distributes new requests to the machines with the least amount of active connections
- *Weighted Least Connections* – Distributes new requests to the machine with the least amount of active connections, according to the machines' weights (default)
- More complex scheduling policies exist and are used for different applications and configurations (such as multiple firewalls).
- Server weights are direct ratios: A server with a weight of 2 gets twice as many connections as a server with a weight of 1.

# Piranha – A Typical Example

-  Piranha IPVS Director Cluster
-  Real Server
-  Red Hat Cluster Manager Cluster



# Piranha Configuration Considerations

- Be aware of potential and existing use patterns for your server farm.
  - Performance requirements: Bandwidth, CPU, hits/day, etc.
- Network accessibility: Can you have all of your machines on the front line?
- Security considerations: *Do you want* all of your machines on the front line?
- Application considerations: Database required? How many different applications need balancing?
  - Some applications (e.g. FTP) require the installation of helper modules for IPVS.
- Architectural considerations: What makes the most sense for your network architecture?
- Hardware considerations: Are the nodes of a similar configuration? How many machines will be acting as real servers?

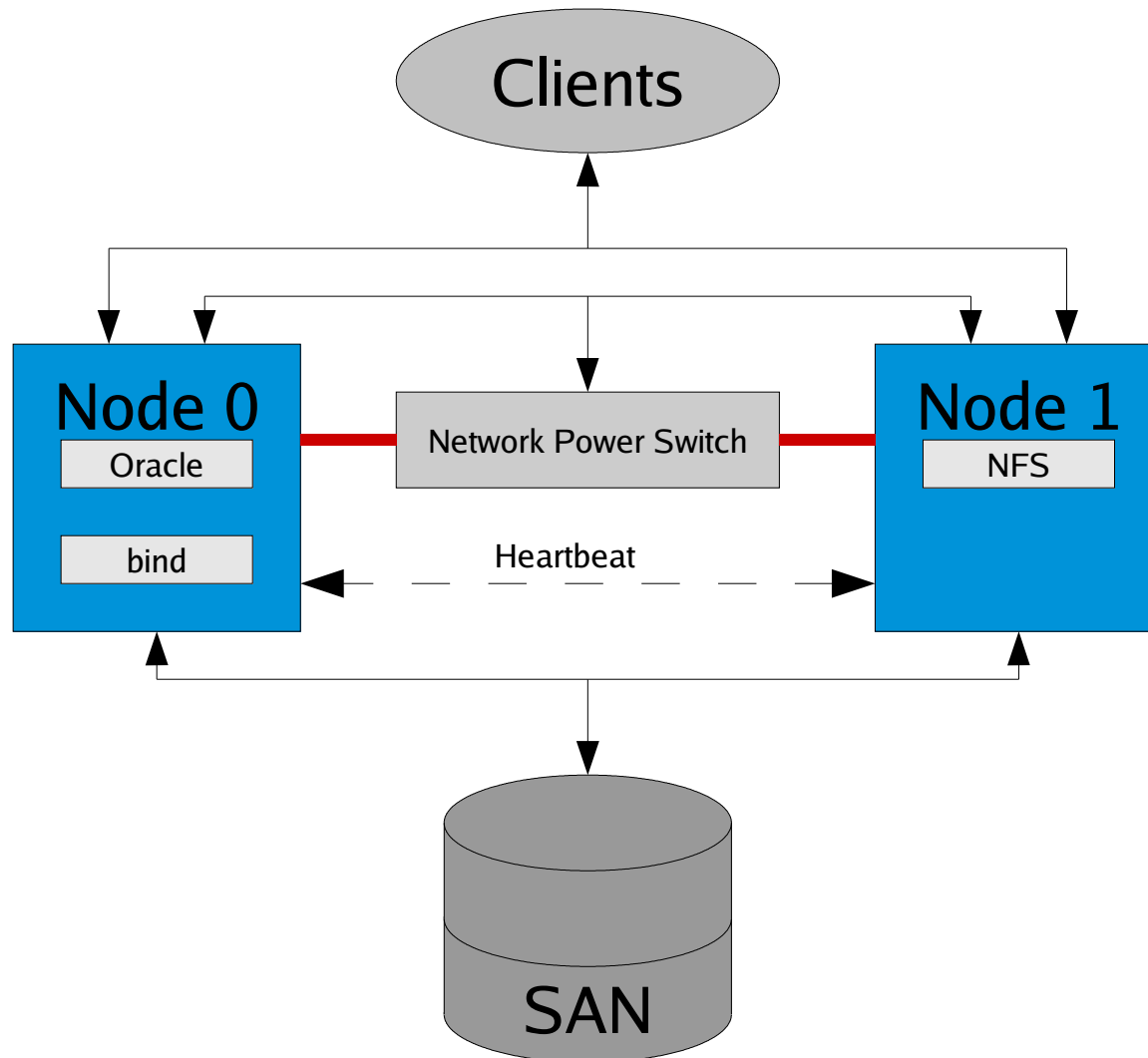
# Red Hat Cluster Manager

- Useful for making off-the-shelf applications highly available
  - Does not require the application to be cluster-aware, but may require some configuration tweaks
- Uses a “virtual service” design.
- Provides methods to describe preferred nodes and/or restricted sets of nodes on which a service should run
- Extensible script-based framework for building new *resource agents*
  - ... or plug in an existing SysV-style init-script
- Simple dependency tree for services: only touch the affected parts
  - Alter any piece of a service; rgmanager will only restart the affected parts of the service
  - If a piece of a service fails, rgmanager will only restart the affected pieces

# Typical Uses

- An Oracle® 10g iAS database (infrastructure mode) cold failover cluster
- NFS Failover cluster for file serving
- Server consolidation
  - Run DNS, DHCP, NFS, and sendmail as 4 distinct services with their own IP addresses on a 2-node failover cluster for increased availability

# Red Hat Cluster Manager Architecture

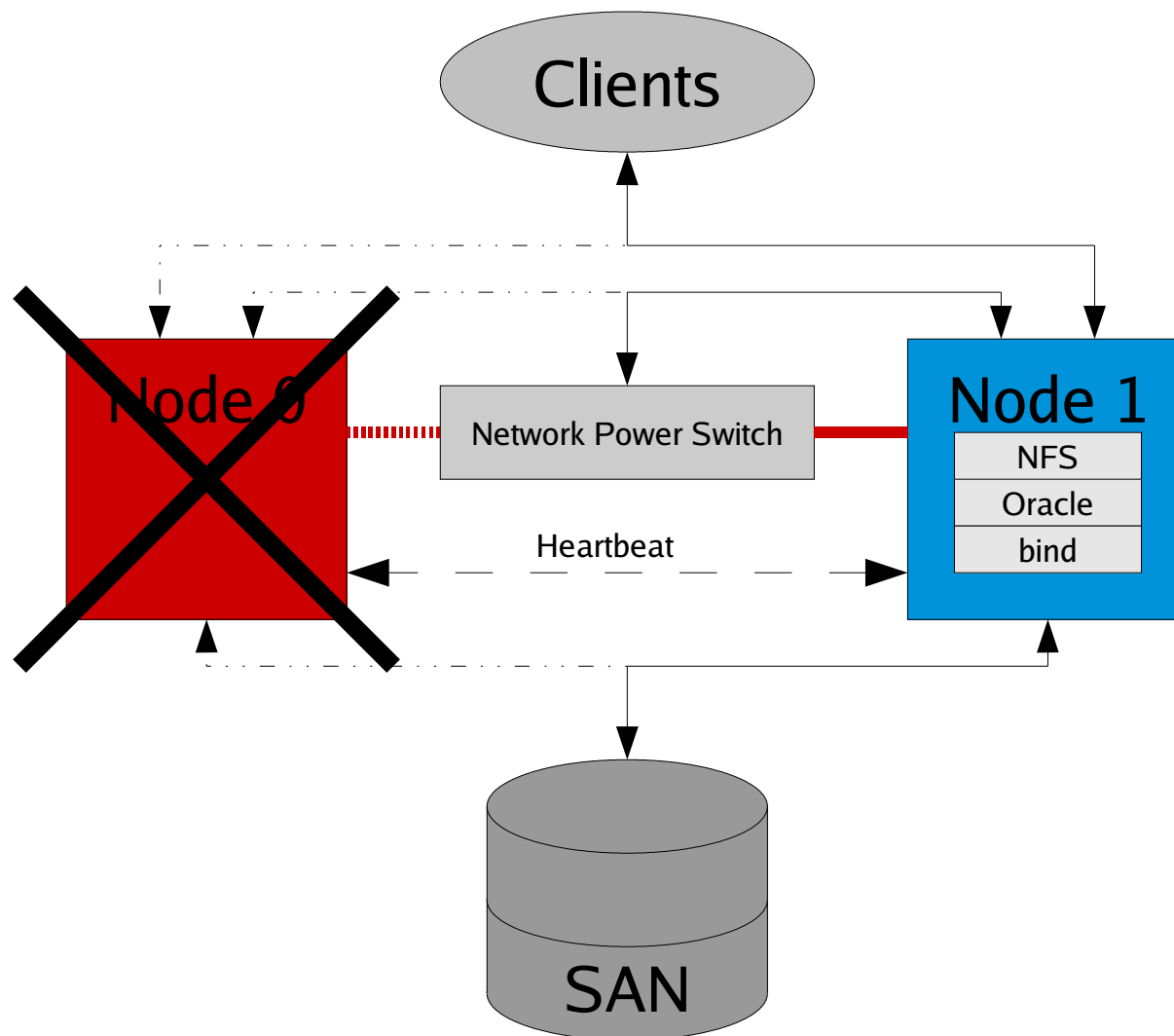




# Handling Failures

- Software Failures
  - Monitoring provided by resource agents handles restarting
  - If a resource fails and is correctly restarted, no other action is taken
  - In the event that the resource fails to restart, RHCM will stop and relocate the entire service to another node
- Hardware, cluster failures
  - If the cluster infrastructure evicts a node or nodes from the cluster; RHCM selects new nodes for the services it was running based on the failover domain if one exists
  - If a NIC fails or a cable is pulled (but the node is still a member of the cluster), the service will be relocated
- *Double Faults* – Usually difficult or impossible to choose a universally correct course of action when one occurs. Ex: Node with iLO losing all power vs. pulling all of its network cables.

# After a node failure...



# RHCM Configuration Considerations

- Some applications are internally highly-available and would receive little benefit from running as part of a Red Hat Cluster Manager service.
- It's good to be proficient in shell scripting when attempting to configure a third party application to run as part of a cluster service
  - *#1 pitfall in the field with respect to service configuration has been improperly written user scripts!*
- Consolidating a bunch of older machines on to one cluster can increase availability while saving on power costs and rack space
- Good shared storage *is not cheap!* Disk failures and faults are one of the highest causes of application outages.
- Larger clusters tend to have more fault tolerance than two-node clusters
- Do your nodes need access to the same data for different services?  
Consider adding Red Hat GFS

- **Founded 1993**
- **Headquarters: Raleigh, NC**
- **27 offices worldwide**
- **IPO in 1999 (NASDAQ: RHAT)**



# Red Hat solutions are sold as subscriptions

An Annual Subscription includes:

- **Technology**
  - Product & Documentation
- **Certifications**
  - A wide choice of certified hardware and software
- **Maintenance**
  - Red Hat Network delivers updates and errata (e.g. security & bug fixes)
- **Upgrades**
  - Customers get new releases at no extra charge
- **Technical Support**
  - Up to 24x7 with 1 hour response
  - Basic, Standard, Premium options available



**Questions?**