

The Samba logo consists of the word "Samba" in a bold, white, sans-serif font. The letter 'S' is stylized with a white arrow pointing to the right, and the letter 'A' has a white arrow pointing to the left. The logo is set against a dark grey rectangular background.

**SAMBA**

opening windows to a wider world

# SambaXP 2010: Clustering Experiences with Samba CTDB

John H Terpstra, CTO  
PrimaStasys Inc.

---

# Agenda

- Overview of cluster deployments
  - Get it right first time!
  - Electronic Search Application
    - First cluster
    - Second cluster (WIP)
-

# Original Cluster Objectives

Application: Electronic Search and Filtering

- 90% Reduction in processing time
    - Achieved: 75%
  - Provide higher processing reliability
    - Mostly achieved
  - Reduced infrastructure cost
    - Costly SAN replaced by low cost infrastructure
-

# Problems - Hardware

- Hardware
    - Get it right or you will waste much time
  - Disk drive failures
    - Drive failures
    - Firmware updates had to be applied to all drives
      - Required 2 total shutdowns
        - Replace drives
        - Update firmware on all drives
  - Network Adaptors (NICs)
    - Ethernet bonding does work ... but ...
-

# RHCS Samba-CTDB Cluster

## Redundant design

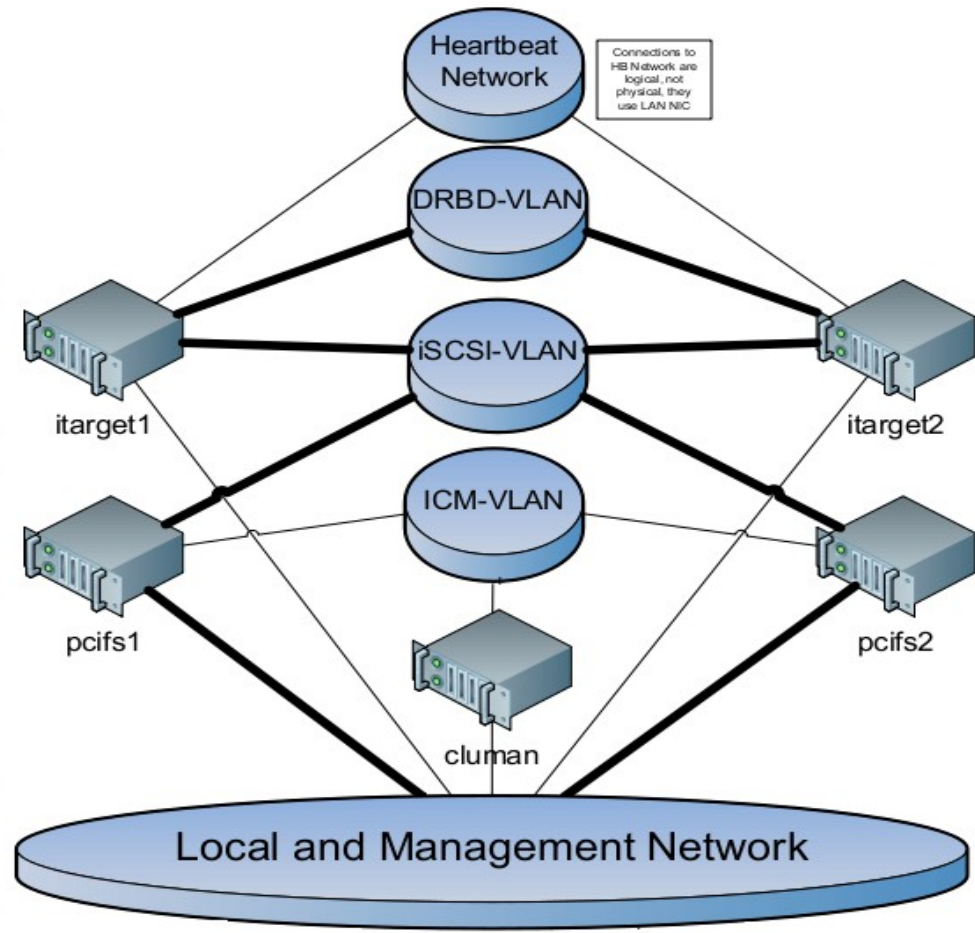
Ideally the heartbeat network would be a simple crossed over network between the two hosts. In this situation, we do not have the spare NICs so we run it logically over LAN

DRBD is a network which optimizes sync traffic by sending write data for SQL one way while sending CIFS data the other. In many ways, it is efficient because it duplexes.

iSCSI is a busy network which connects the pcifs to the iSCSI targets. Devices use bonded NICs to optimize performance

ICM network is a management network between the cluster manager and the pCIFS servers. The pCIFS servers also use this network for SAMBA communication because the LAN connection can potentially fence.

The LAN network has all the fronted server, application servers, and Management devices. This is a routed network using redundant firewalls to the internet.

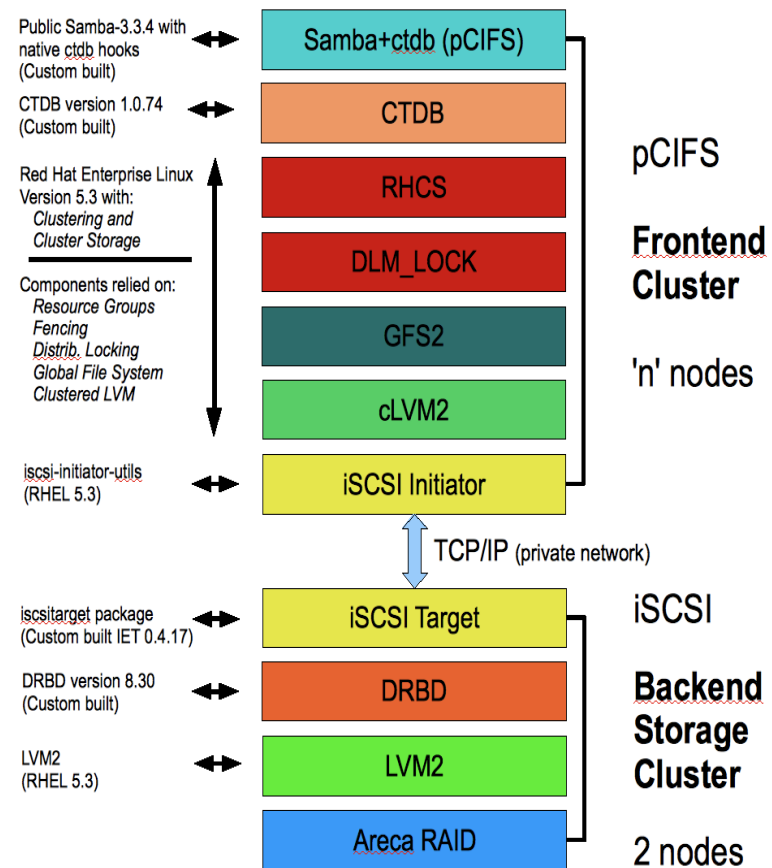


The iTarget servers are large storage server that cluster their drives in a redundant manner for data protection

pCIFS servers are highly available file servers using Samba and CTDB.

# Software Layers (Old Cluster)

- Current Cluster
  - Backend RHEL
    - DRBD
    - IETD
  - Frontend RHCS
    - Open-iscsi
    - CLVM2
    - Samba CTDB



# Software Issues

- File system issues
    - Directories 20 levels deep, 128 char file and directory names, 2-50Kbytes/file, 2-50K files/dir
      - Tried Ext3 and XFS on stand-alone systems
        - Very poor performance
        - <2.5 Mbytes/sec transfer with rsync
      - Using GFS2 in the cluster
        - DLM performance issues
  - Machine password change behavior
    - Not compatible with Samba-CTDB
  - Windows Discovery Software update suddenly led to new file names and POSIX locking problems
    - Solved by disabling Linux Kernel OpLocks
-

# Other issues

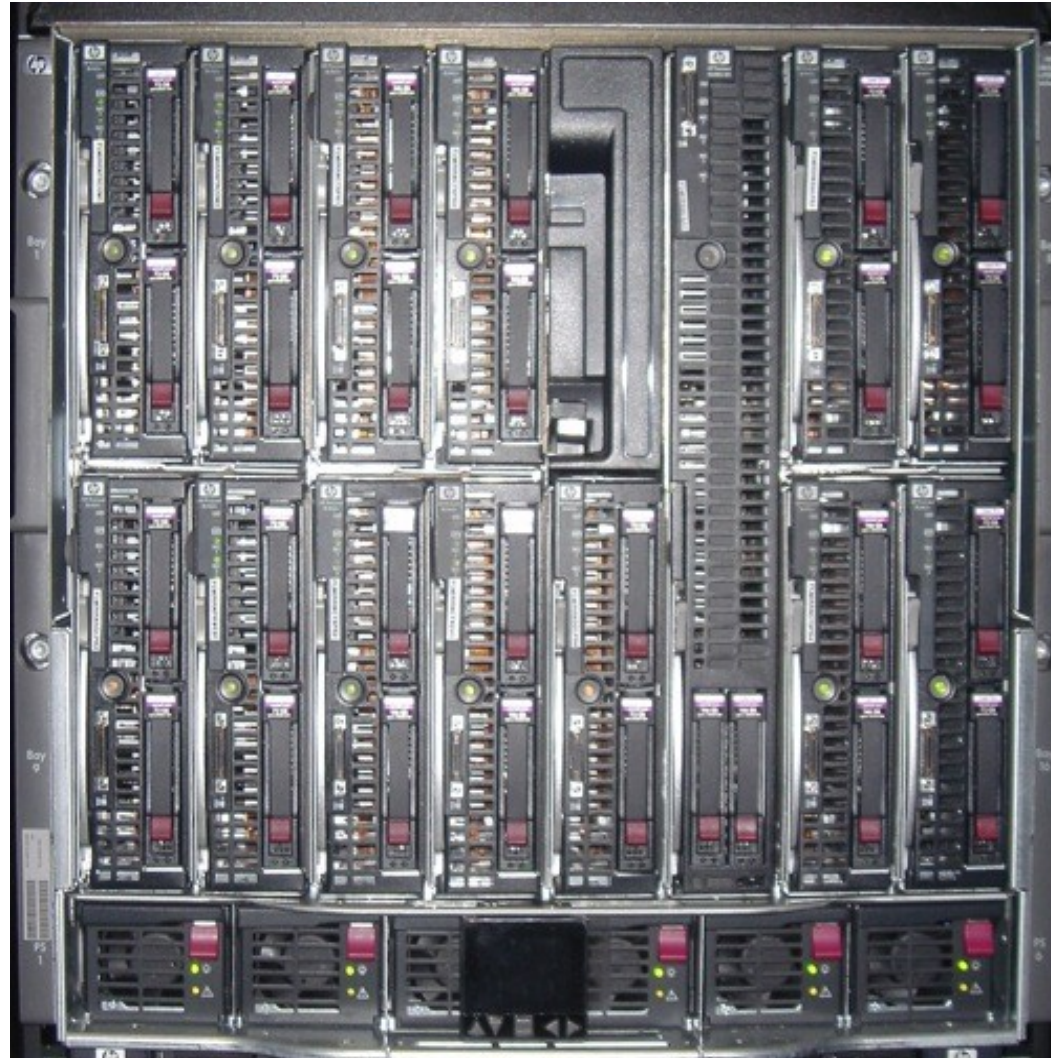
- Corruption of CTDB log files
    - Not found cause, but removal of /var/ctdb transient files solved problem
  - Windows application software required directory listings to be alphabetically ordered
    - Tried the *dirsort* VFS module
      - Failed due to memory exhaustion due to number of files per directory
    - Software vendor fixed software
-



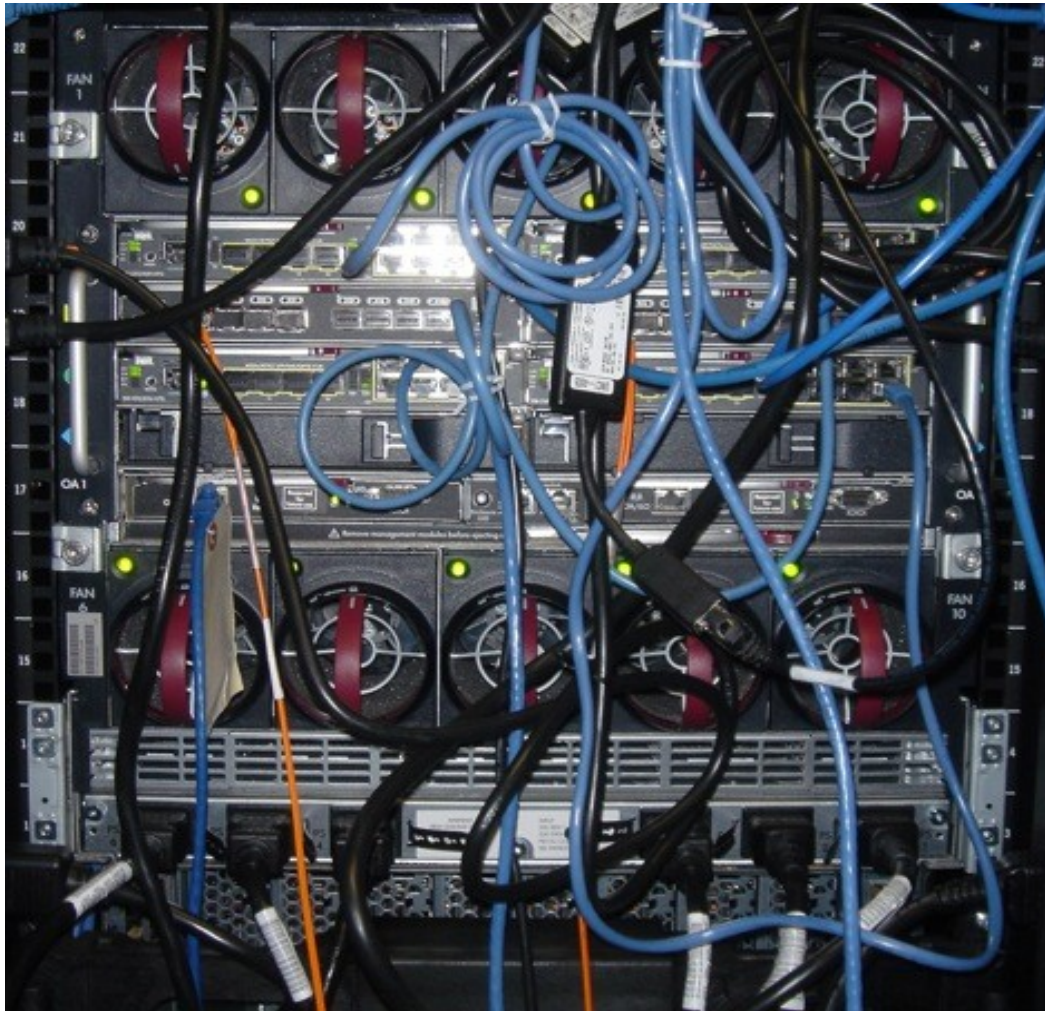
# ClusterNG Objectives

- Reduce processing time by 75%
  - Increased scalability
    - Backend storage several petabytes
    - Frontend – 2-9 nodes initially, maybe up to 30 nodes over 12 months
    - Use VM frontend nodes
  - Reduce power consumption
  - Increased uptime
-

# ClusterNG Rack Front View

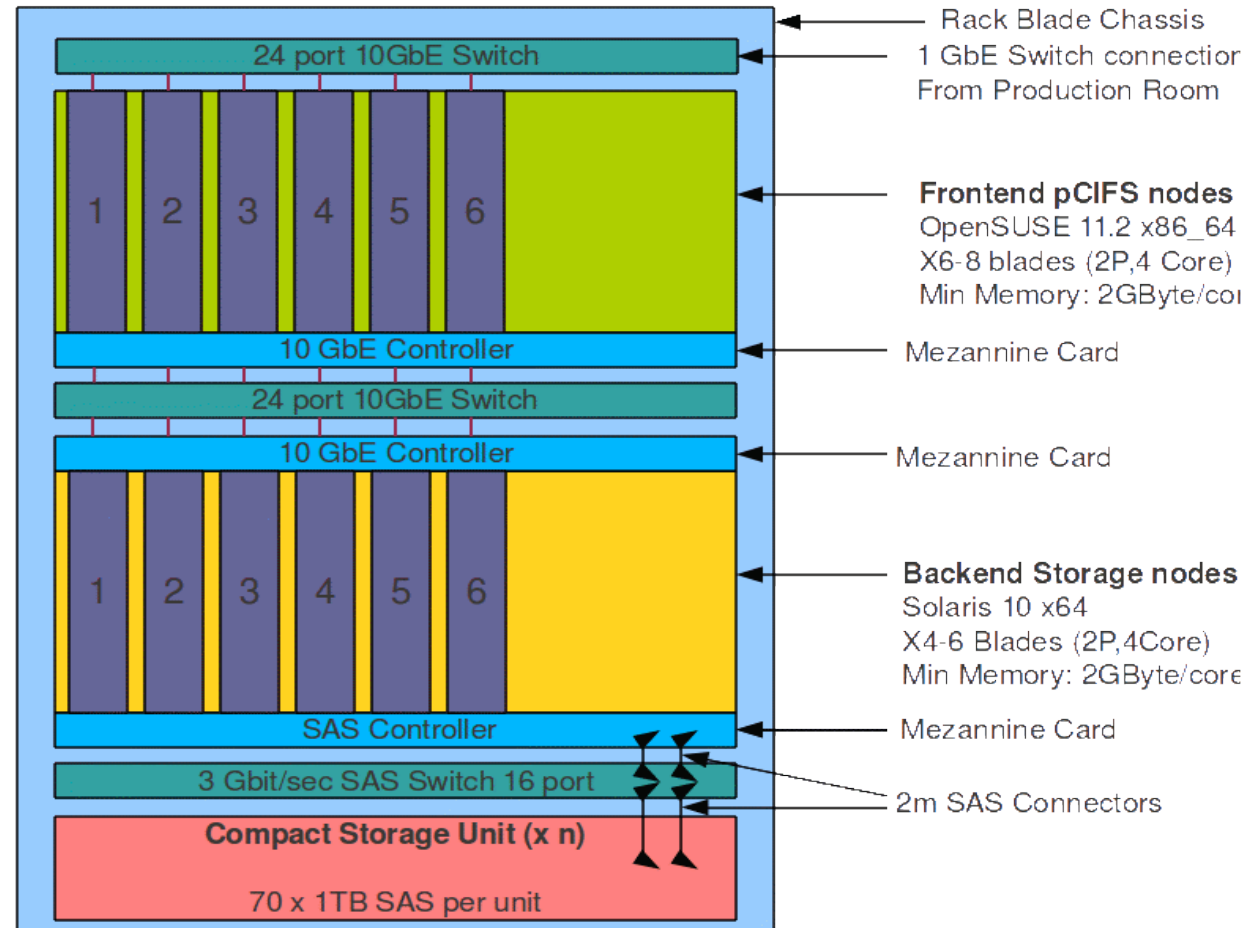


# ClusterNG Rack Rear View



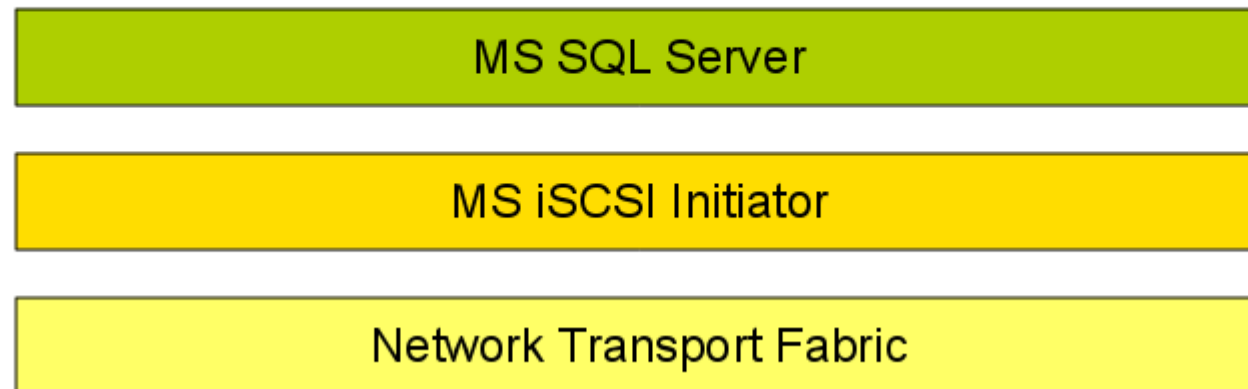
# ClusterNG – Rack Layout

Cluster Rack



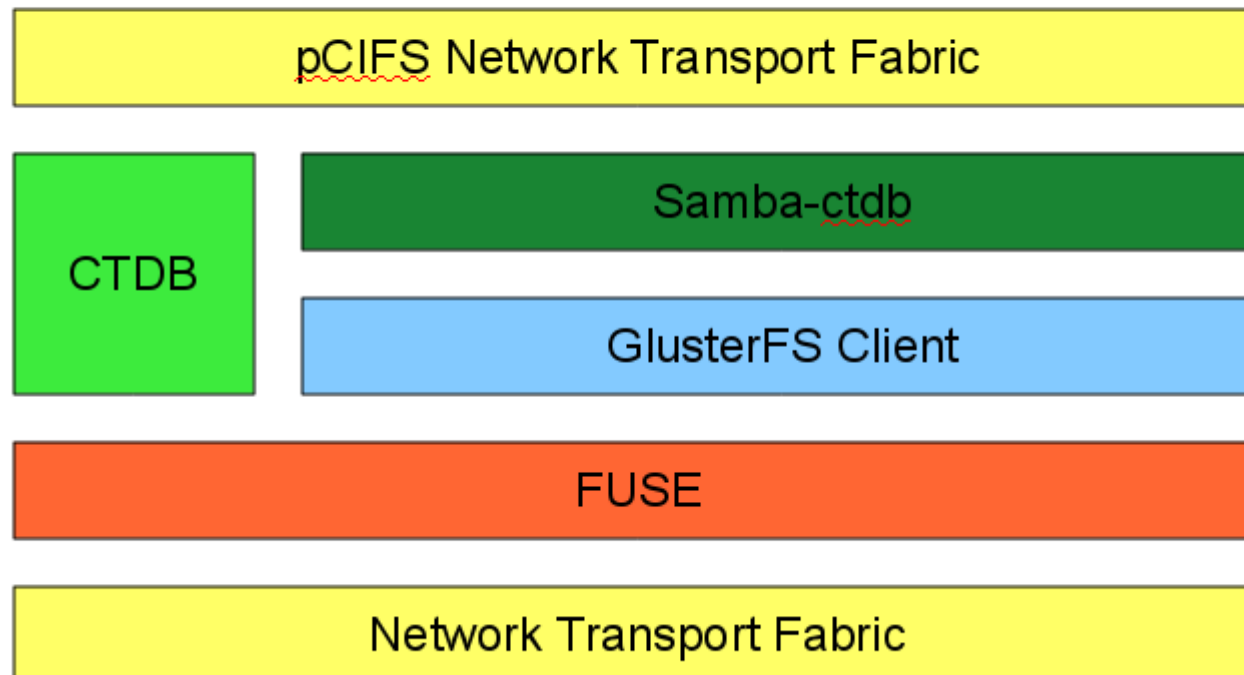
# Microsoft SQL

## MS SQL Software Stack



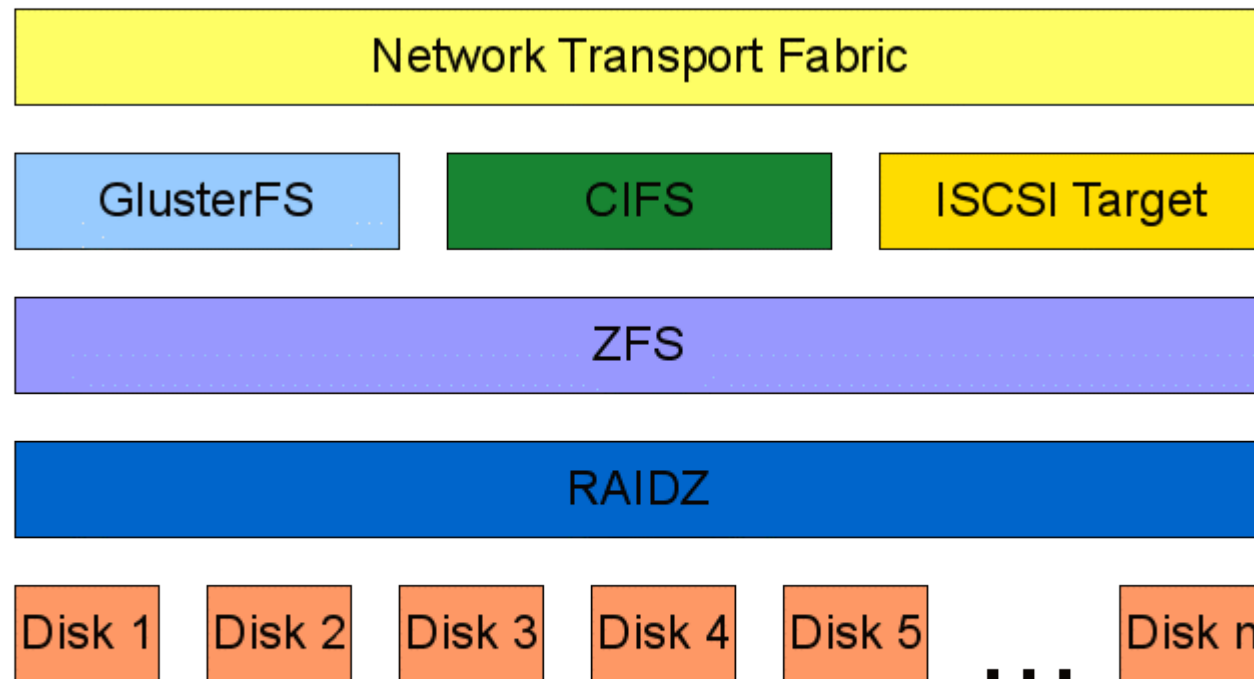
# pCIFS Nodes

## pCIFS Server Software Stack





# ClusterNG Backend Storage Nodes



OpenSolaris Storage Server  
Software Stack

---

# More info ...

Will be provided at SambaXP2011  
Our 10<sup>th</sup> Anniversary Conference

Se you there!

