



Copyright © 2003 VERITAS Software Corporation. All rights reserved. VERITAS, the VERITAS Logo and all other VERITAS product names and slogans are trademarks or registered trademarks of VERITAS Software Corporation. VERITAS, the VERITAS Logo Reg. U.S. Pat. & Tm. Off. Other product names and/or slogans mentioned herein may be trademarks or registered trademarks of their respective companies. Specifications and product offerings subject to change without notice.

# The Ten Essentials of a Successful RAC Implementation



Brian Schwarz  
*VERITAS Software*

# The Essentials

---

## Choosing the Right Architecture

- 3 Designs for Highly Available Oracle

## Architecting a RAC System

- Managing LUNs, Volumes and Files
- Ensuring I/O Performance
- Storage Networking Considerations
- CacheFusion Transport Options
- Clustering Split Brain and Prevention Methods
- Backups with RAC

VERITAS Database Edition / Advanced Cluster

VERITAS – HP Relationship

# **Admissions**

---

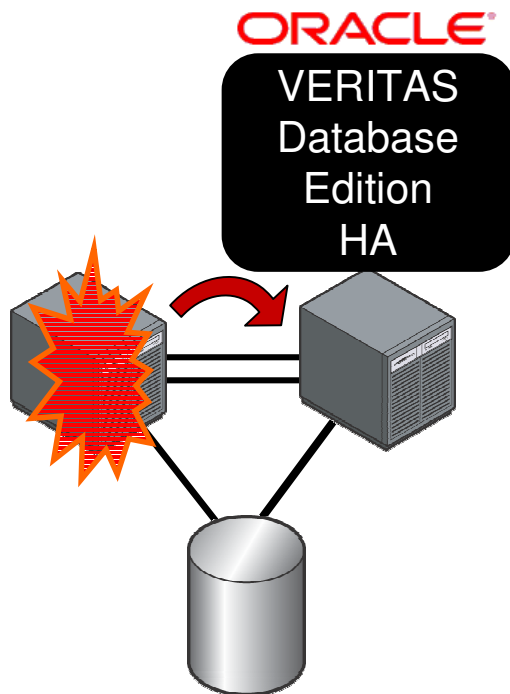
**There are 10+ additional topics that could be included in this session**

**This is a survey course, I recommend follow-on study**

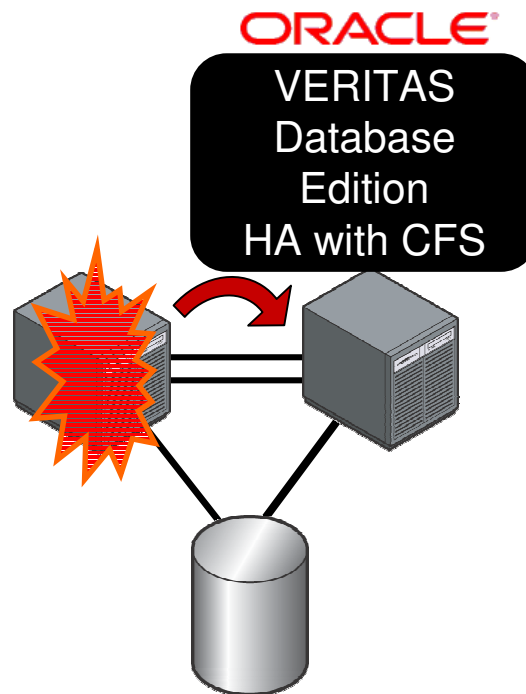
**Tailor this information to your environment**

# Three Architectures for Oracle

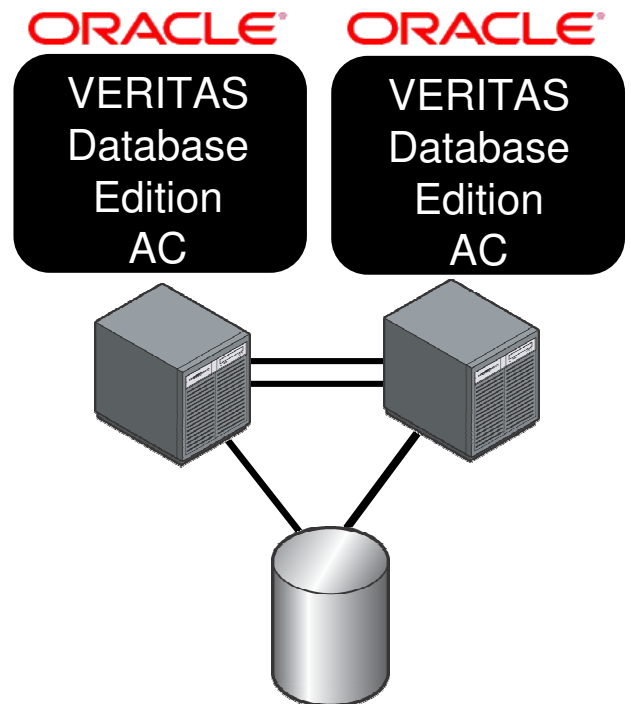
High Availability



Better Availability and  
Storage Manageability



Highest Level of Availability,  
Storage Manageability, and  
Scalability



# Availability Profiles

---

What is the cost of downtime for your application?

	Database Edition HA	Database Edition HA w/ CFS	Database Edition AC
Failover Disks/FS	10 minutes	N/A	N/A
Restart Oracle	5 minutes	5 minutes	N/A
Reconfigure Distributed Locks	N/A	N/A	1 minute
Recover Oracle	1 minute	1 minute	1 minute
Estimated Time	16 minutes	6 minutes	2 minutes

\* All times are estimates only

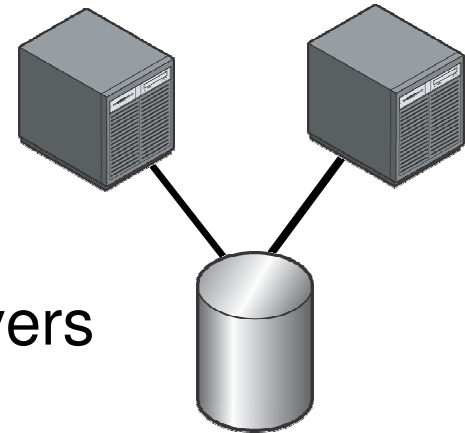
**Ten Essentials of a Successful RAC  
Implementation**

# **Architecting a RAC System**

# Managing LUNs and Volumes

---

- RAC Requires a shared storage architecture
- Manage the storage for cluster, not for each node
- Consistent naming of devices across all nodes
- Device failure and addition
- Striping and Mirroring
- Online Growth and Reconfiguration
- Multi-pathing between storage and servers
- Options
  - No Volume Management
  - Single Host Volume Manager
  - Cluster Volume Manager





# Managing Files in a Cluster

---

- What needs to be stored?
  - Datafiles
  - Oracle files
    - \$ORACLE\_HOME (Binaries), Archive logs, B-files, trace files
  - Other files
    - Admin scripts, 3<sup>rd</sup> party utilities (backup, perf. monitoring, etc..)
- Installation and patching
- How to backup?
- Instance Recovery and access to Archive Redo Logs
- Number of Volumes to Manage
- Options
  - Raw for Datafiles, local file system for rest
  - Raw for Datafiles, cluster file system for rest
  - Cluster file system for all

# Ensuring I/O Performance

---

- Direct Read and Write from all nodes
- Bypassing CFS overhead
  - Oracle Disk Manager API
- Multi-Pathing to take advantage of
  - Several HBAs
  - SAN connections
  - Array controllers
- Good volume layouts
  - RAID 1+0 for redo logs
  - RAID 1 for datafiles

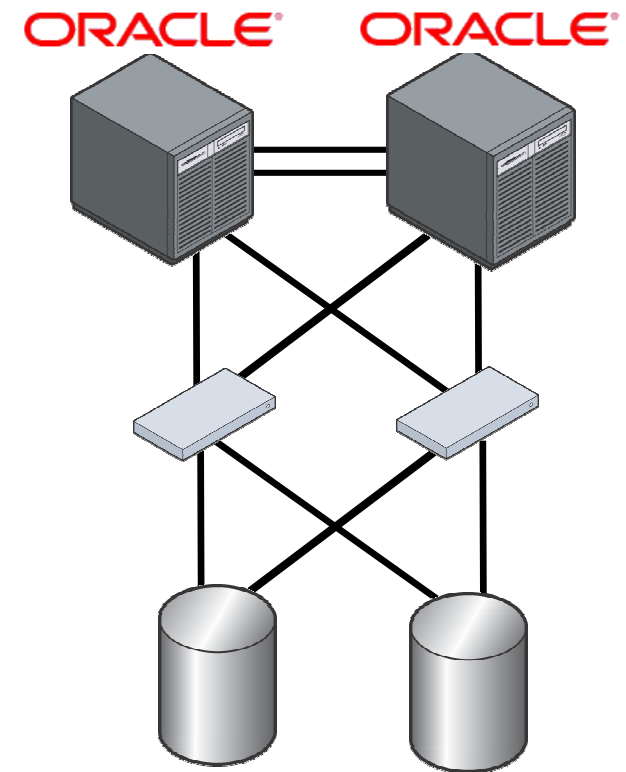
# Storage Networking Considerations

---

- SAN design is a deep topic
- Try not to create SAN islands for each app/DB
- Several network topologies available
  - Star, full mesh, core-edge
  - One fabric or two?
- Minimum Recommended Configuration for RAC
  - 2 HBAs per host
  - 2 switches
  - 2 storage arrays

# Basic Redundant SAN Architecture

- Two Fabrics
- Two HBAs per server
- Multi-pathing SW on the host
  - Array & Multi-pathing SW must be supported configuration
  - Active/Active or Active/Passive
- Two storage arrays
- No SPOF in the architecture



# CacheFusion Transport Options

---

- CacheFusion is a key part of Oracle RAC architecture
- Transport the performance bottleneck? **Not!**
  - Use the Wait Interface
  - Search for “global cache” events
  - Incremental spending, tuning of transport often wasted
- Locality of Data Access drives CacheFusion usage
- Options
  - Ethernet with TCP/IP or UDP/IP
  - Ethernet with proprietary network protocol (Ex. VERITAS LLT)
  - Proprietary network cards, switches, SW

## Transport Options (con't)

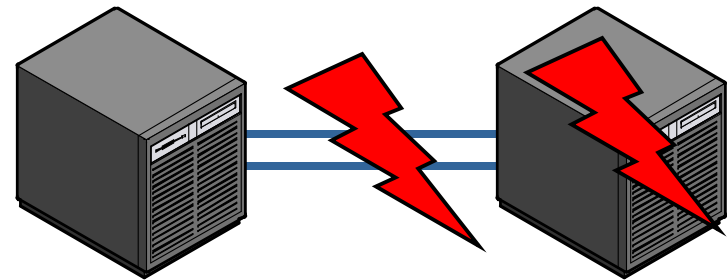
---

- With TCP or UDP, 1 GigE connection at full bandwidth consume ~ 1 GHz of CPU processing
- Shared transport with heartbeats and cluster membership services
  - Not throughput intensive, but need multiple for HA
  - Don't send CacheFusion through disk or public network

# Clustering Split Brain and Prevention Methods

---

- All nodes in cluster need to know who else is a member
- Cluster membership services generally use heartbeats on inter-node connections
- What happens on different failure mechanisms?
  - System Failure
  - Network Failure
  - System Hang



# Clustering Split Brain

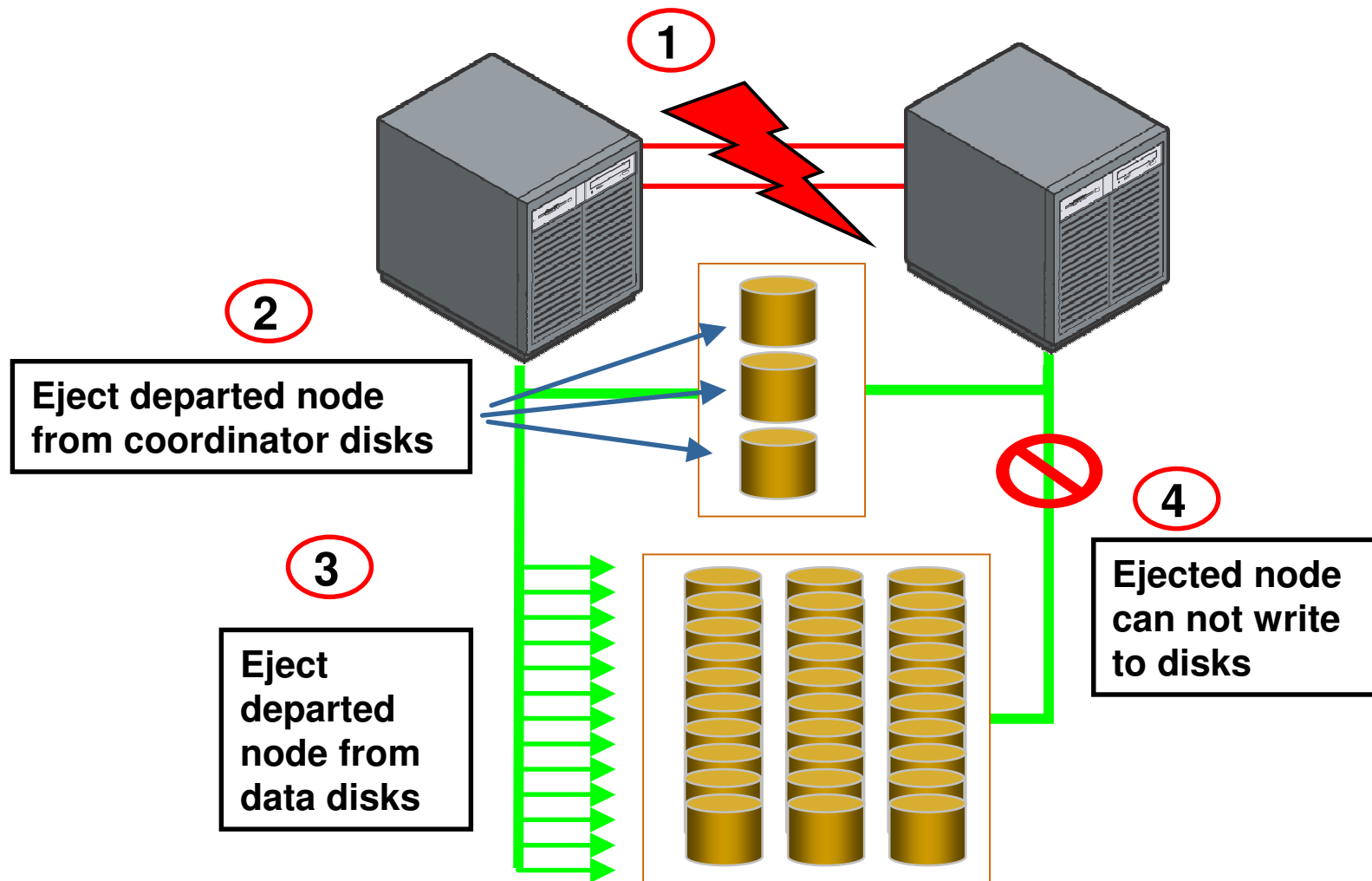
---

- Split Brain can create logical database corruption
- Generic clustering problem
- More important/dangerous in with RAC
  - Shared access to database always enabled
  - Instances up and running on many nodes
- Prevention Methods
  - Heartbeat Timeout
  - Quorum
  - I/O Fencing
    - Disk Level
    - SAN Level





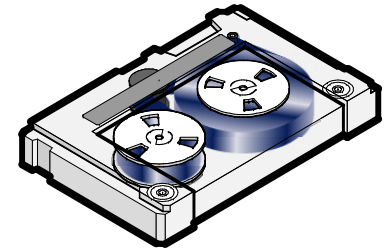
# Example: VERITAS Disk Based I/O Fencing



# Backups and Recovering from Media Failures

---

- What's needed?
  - Access to backup datafiles
  - Access to Redo and Archive Redo logs
    - Each instance in RAC has its own redo thread
  - Control file
- Using NetBackup to Protect Oracle RAC
  - Whitepaper also available
- Can the backup run parallel?
- Use the cluster's virtual host name for backups?



# The Final Three

---

- Know your goals upfront
  - Initial size and growth of data?
  - How many TPM average and peak?
  - What response time is needed?
  - What is the I/O profile?
- Make sure joint installation team is formed
  - System Admins, DBAs, Oracle, VERITAS, HP
- Design for Scalability and Availability
  - Make sure clients are load balanced effectively by Oracle Listener
  - Use Oracle TAF interface

# Where to Get More Information

---

- VERITAS Architects Network
  - All VISION conference presentations
  - [www.veritas.com/van](http://www.veritas.com/van)
- Using NetBackup to Protect Oracle RAC
  - Bill Webster – S150
- Database Downtime is not an Option
  - Subbu Iyer – S121
- Optimize DB Performance for Your Environment
  - Tom Murphy – S120
- I/O Fencing Explained
  - Jim Senicka – 3 part series

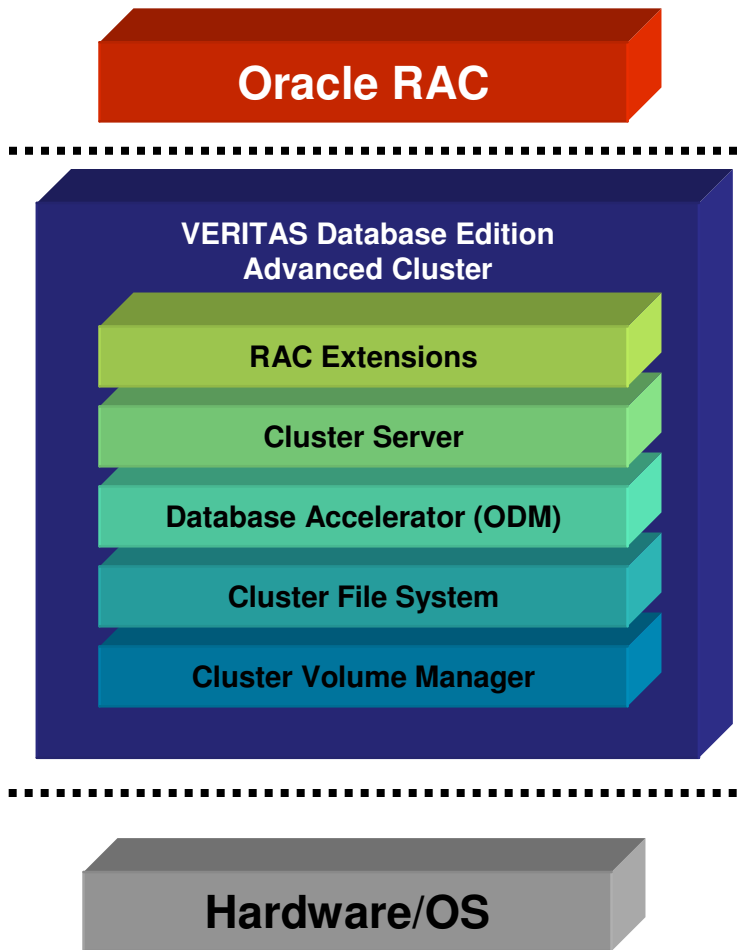
**Ten Essentials of a Successful RAC  
Implementation**

**VERITAS Database  
Edition / Advanced  
Cluster**

# VERITAS

## Database Edition Advanced Cluster

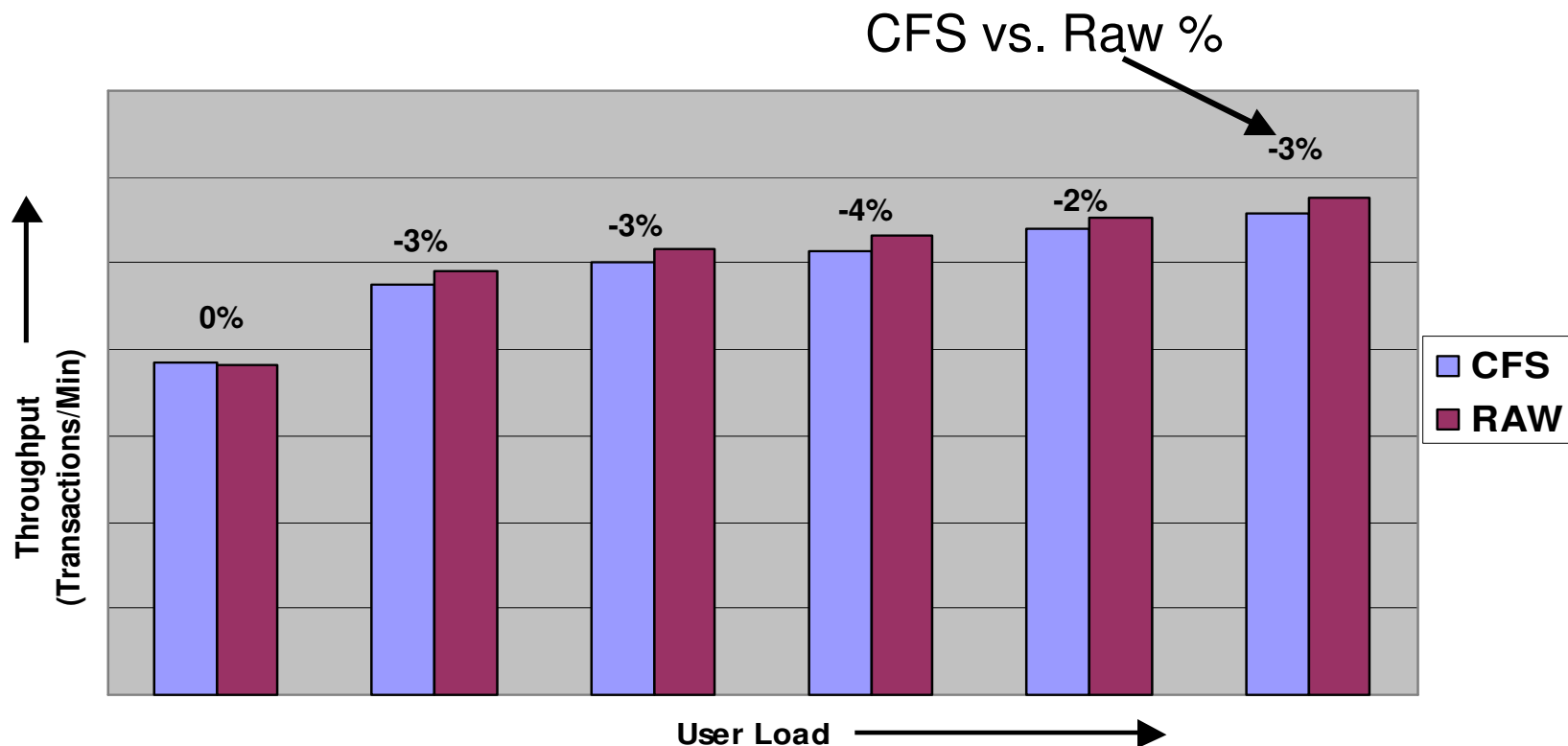
---



- Complete solution for RAC
- Based on core VERITAS technologies
  - Storage Foundation
  - Cluster Server
- Optimized & Tested with Oracle RAC
- Available HP-UX 11.11
  - **Certified by Oracle**
- Also available on Linux

# Performance Test Results

- VERITAS CFS performance on HP with Oracle9i RAC
  - CFS manageability with no performance penalty



# **Ten Essentials of a Successful RAC Implementation**

## **Does Anyone Use It?**



**Ten Essentials of a Successful RAC  
Implementation**

# **HP & VERITAS Relationship**

# **HP and VERITAS**

## ***A Time-Tested Partnership***

---

- 1989**      **HP resells the entire, multi-platform VERITAS product line**
- 1993**      **HP OEMs VERITAS File System as Online JFS**
- 2000**      **Lite versions of VERITAS File System (JFS) and VERITAS Volume Manager included with HP-UX**
  - Full version of VERITAS Volume Manager available**
  - Full version of VERITAS File System included in HP's packaged software bundles**
  - HP and VERITAS announce joint collaboration on Zero Downtime Backup initiative**
- 2001**      **HP OEMs VERITAS Cluster Volume Manager**
- 2002**      **Refreshed OEM agreement covering future releases of HP-UX and VERITAS products**
  - Integrated Services and Support Agreement for VERITAS products**
  - HP and VERITAS collaborate to deliver server-free data protection for HP XP family of arrays**
- 2003**      **HP announces Mezzanine Backup support with VERITAS NetBackup**
  - HP and VERITAS announce API agreement in support of SNIA/CIM initiative**
  - HP & VERITAS collaborate to provide VERITAS Database Edition Advanced Cluster *for Oracle9i Real Application Clusters***

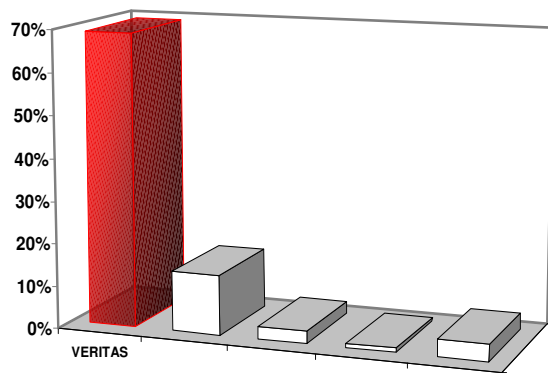
# Storage Management/OEM Partnership

---

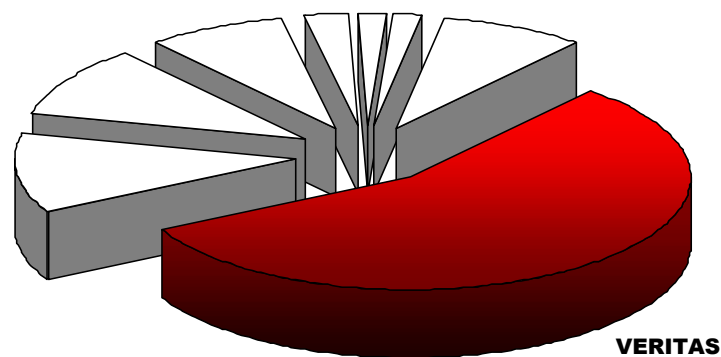
- HP ships VERITAS with every HP-UX box they sell.
- The latest OEM addendum extends this partnership well into 2006 (and beyond).
- 'Lite' versions of VxVM and VxFS included with the HP-UX Operating System
- 'Full' versions of VxVM, CVM and VxFS are available from HP via OEM relationship
- HP|VERITAS DBE/AC Test Drive: New Jersey
- VERITAS' Value Proposition:
  - Technology & Market Share leadership
  - Ease of migration
  - Common storage management look and feel
    - reduction in training costs
    - quicker deployments
  - Flexibility of choice

# VERITAS Market Leadership

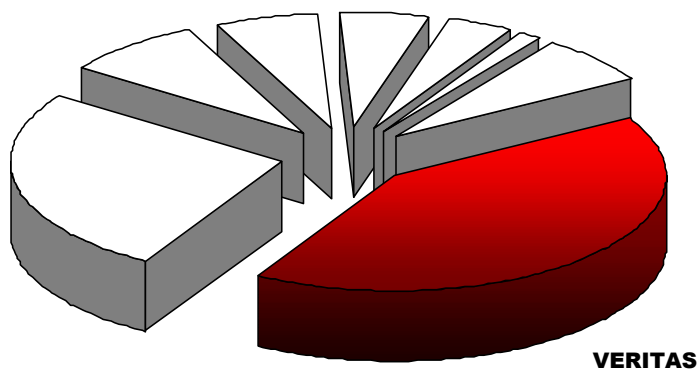
**#1 Storage Virtualization: 69.9%**



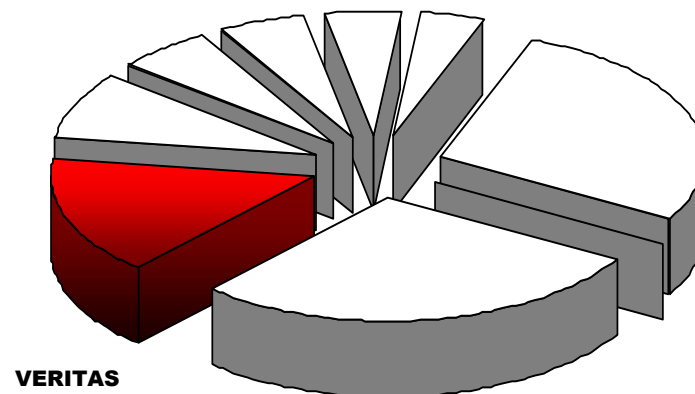
**#1 Backup: Windows – 57.2%**



**#1 Backup: Unix – 39.6%**



**#2 Clustering – 15.6%**



## HP Quote

---

*"HP's work with VERITAS enables enterprise customers to achieve both high performance and manageability of a clustered Oracle database environment."*

*"This [DBE/AC] solution on HP-UX 11i provides additional flexibility for customers to adapt to changing business requirements."*

**Ron Eller, vice president, solutions alliances, HP Enterprise Storage and Servers.**

Quoted in VERITAS press release issued on Aug. 13, 2003 "VERITAS Delivers First Cluster File System for Oracle9i RAC"

# QUESTIONS & ANSWERS

# Conclusions

---

**Don't go into the woods alone**

**Bring along your ten essentials**

**Enjoy HP World 2004 !**





VERITAS™