Planning & Installing a RAC Database

Caleb Small, BSc, ISP
Caleb@Caleb.com
www.Caleb.com/dba

Who is Caleb?

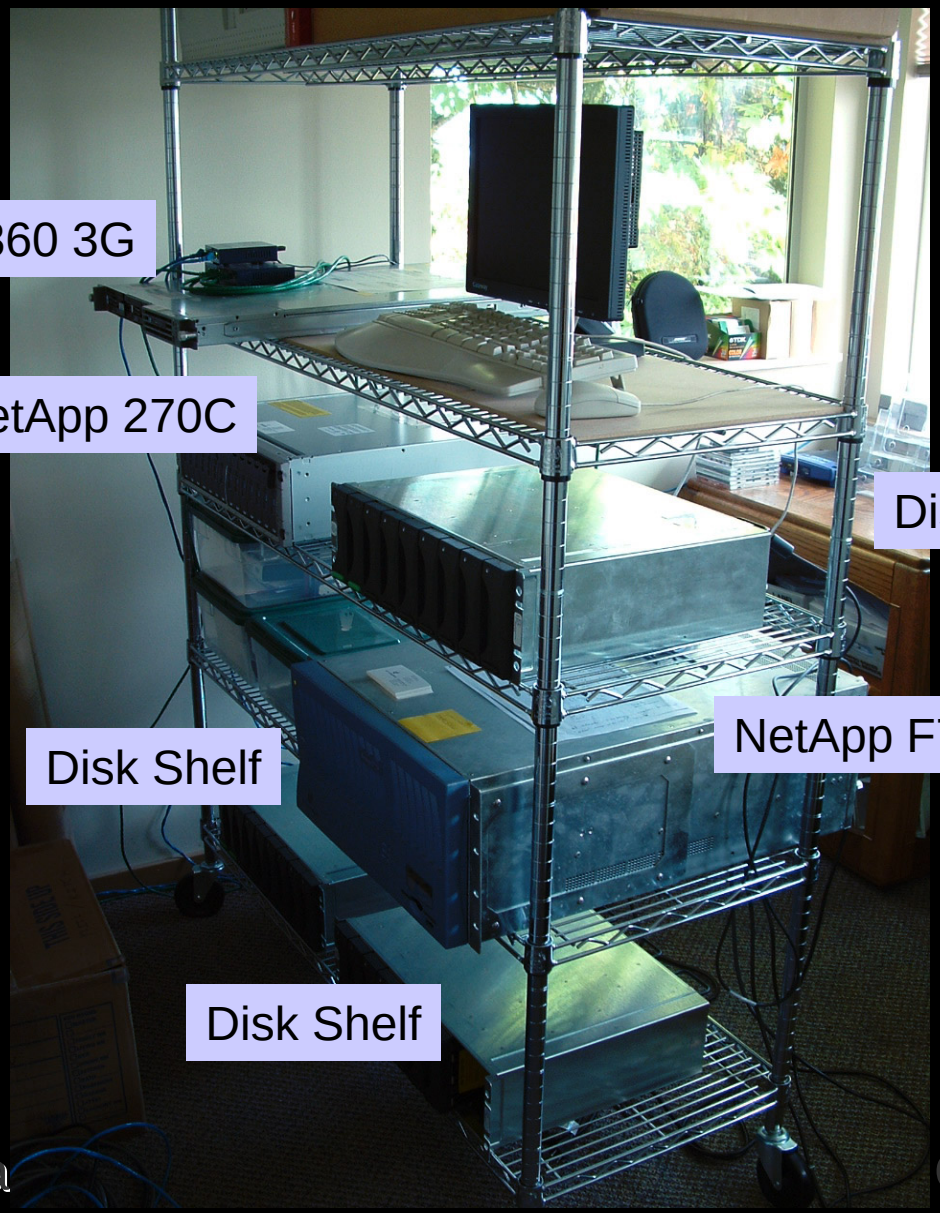
- Lifelong IT career
- Oracle DBA since v7.0
- Former Instructor for Oracle Corp.
- Independent consultant
- Faculty appointment Camosun College
- U of W night courses
- Vice-president, VicOUG
- Director, PSOUG

What does he know about RAC?

- Installing Oracle Databases for 12yrs
- Studied RAC on 9i
- Re-wrote RAC curriculum for 10gR2
- Teaching RAC workshop for PSOUG
- Various RAC installations
- Setup PSOUG and VicOUG RAC labs

PSOUG Training Lab

PSOUG Server Room



Caleb@Caleb.com

So, you're thinking of RAC…

Why?

Caleb@Caleb.com

Benefits of RAC

- Reliability
  - Eliminate the Instance (and server) as single point of failure
- Scalability
  - Exploit low cost commodity hardware
  - Add servers as necessary

Benefits of RAC

- Part of complete High Availability (HA) solution
- Application Failover (TAF and FCF)
- Load Balancing
- Incremental Scaling
- Lower Initial Cost / Higher Return On Investment
- Lower Maintenance Costs
- Server Consolidation

Challenges of RAC

- Complex architecture
- Setting and meeting realistic expectations
- Multiple technologies & skills required…
- Traditional DBA skills plus RAC specific
- Various personnel must understand the architecture and ***cooperate***

Planning a RAC installation

People and process issues

- Understand the architecture
- Set expectations appropriately
- Assemble the technical team
- Plan and budget for training
- Cooperation and management
- Define and measure success factors

Planning a RAC installation

Technical issues

- Choose hardware / OS
- Design networks
- Design storage
- Install & configure software
- Modify applications
- Implement Services and Resource Plans
- Implement Load Balancing & Failover
- Performance monitoring & tuning

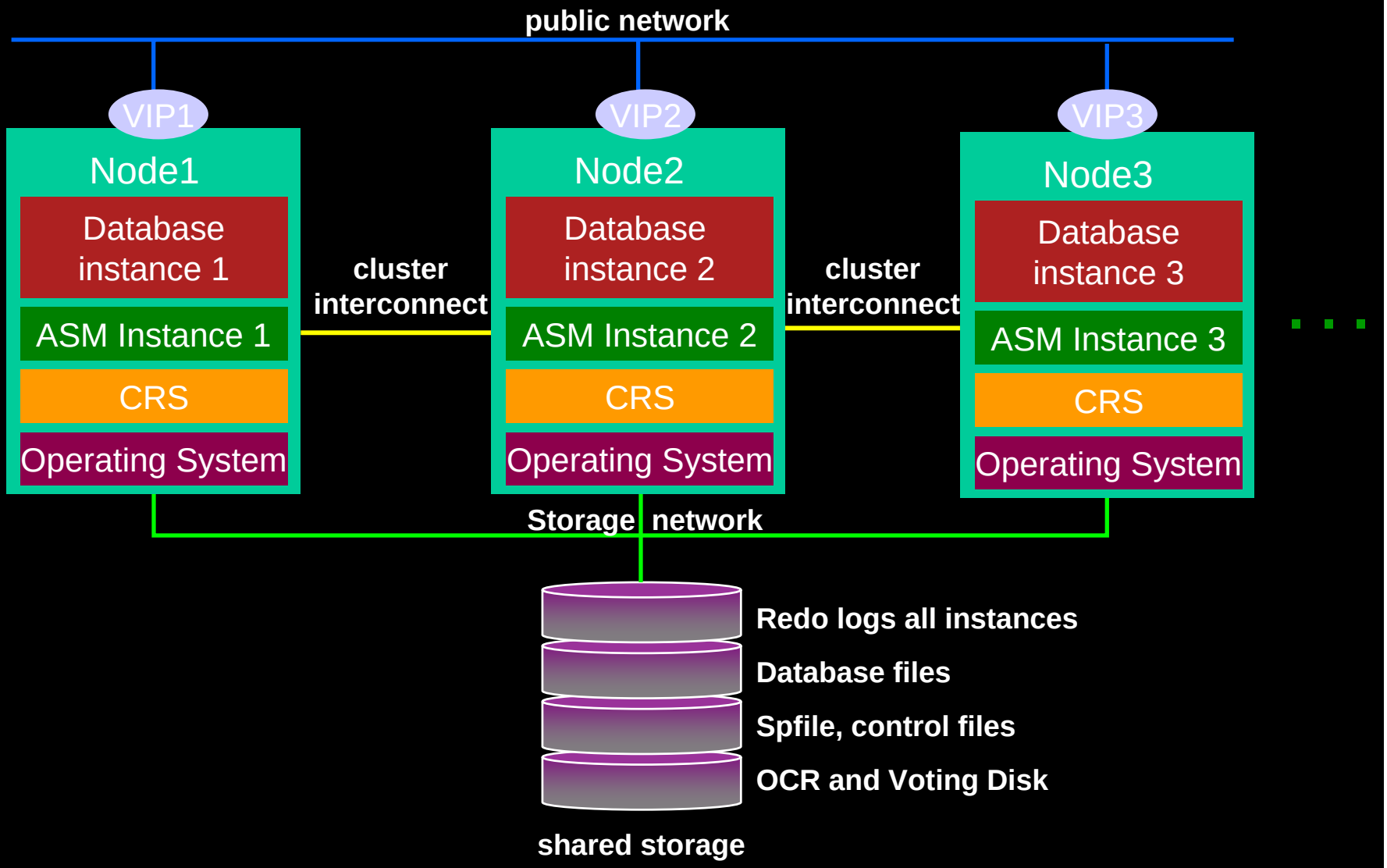
RAC Architecture

- RAC != “Clusters” the database object
- Database != Instance
- A database is a set of files containing all persistent resources
- An instance is a set of memory structures and processes
- Databases and Instances can be started and stopped independently
- In a stand-alone database the ratio is 1:1
- In a RAC database the ratio is 1:Many

What is RAC?

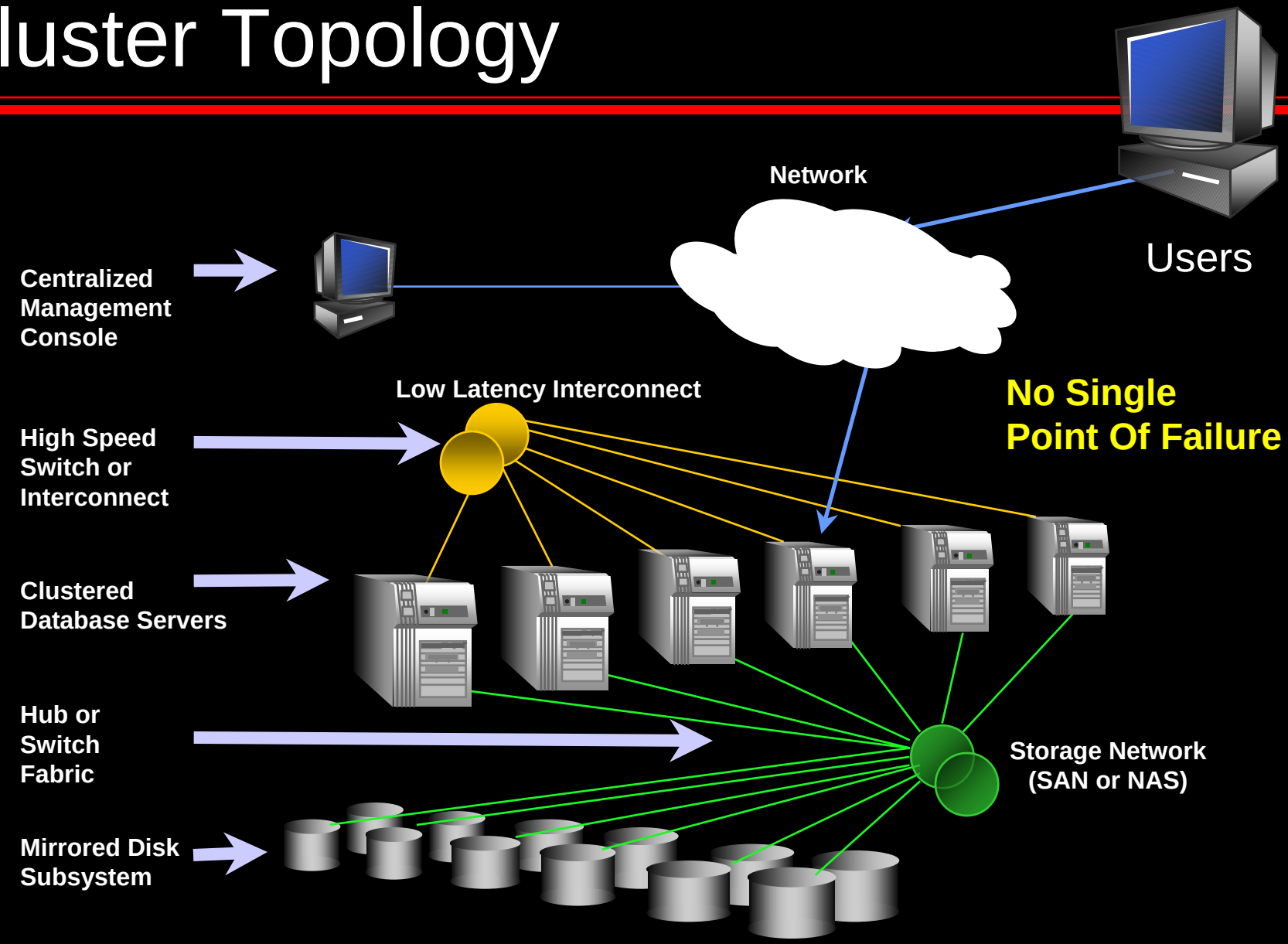
- Multiple instances running on separate servers (nodes)
- No instance is the parent or is privileged
- Single database on shared storage accessible to all nodes
- Instances exchange information over a private interconnect network

Cluster Topology



Caleb@Caleb.com

Cluster Topology



Caleb@Caleb.com

Infrastructure

Application Servers

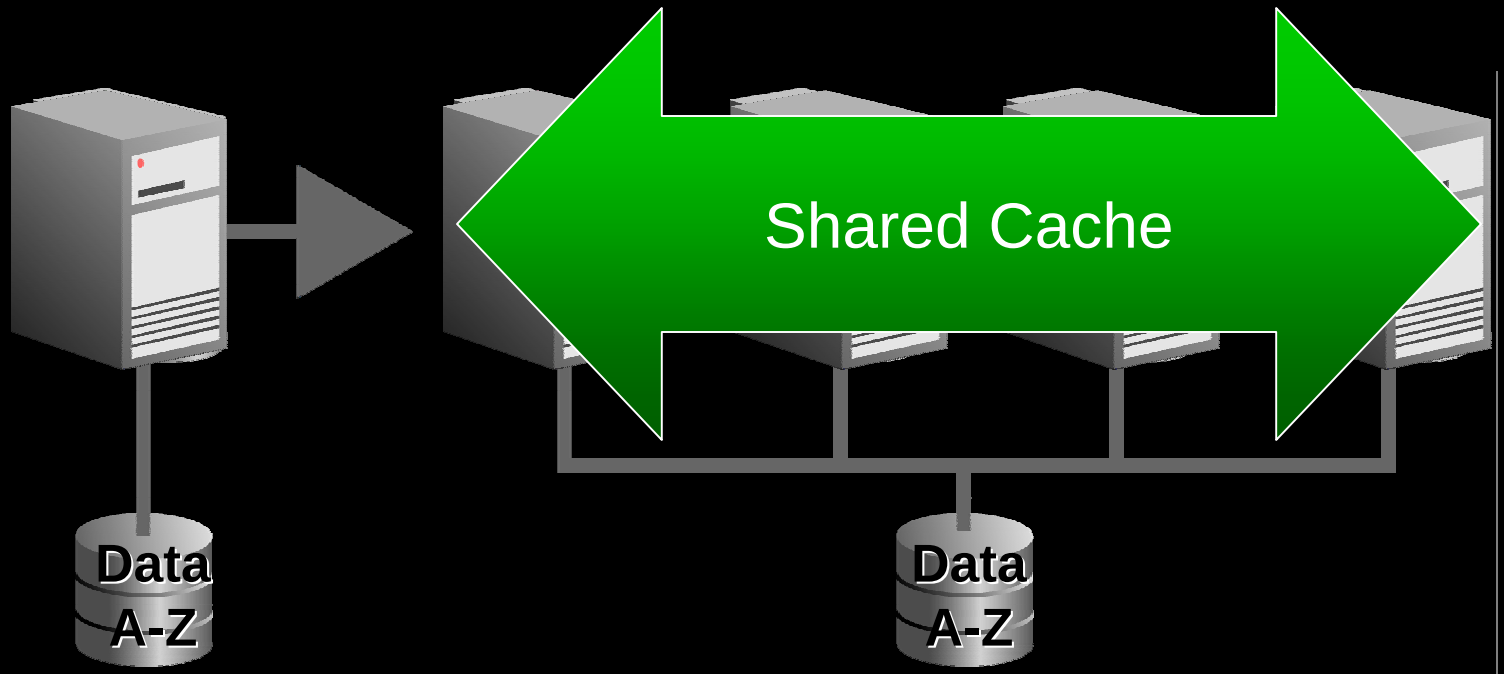
Data Servers

Storage Arrays

Caleb@Caleb.com

Shared Cache Databases

- Oracle RAC is the only Shared Cache database
- Provide BOTH availability and scalability
- Maintained like a single database
- More reliable with more servers



Caleb@Caleb.com

Set Expectations

Understand what RAC is,
and what it is NOT.

Caleb@Caleb.com

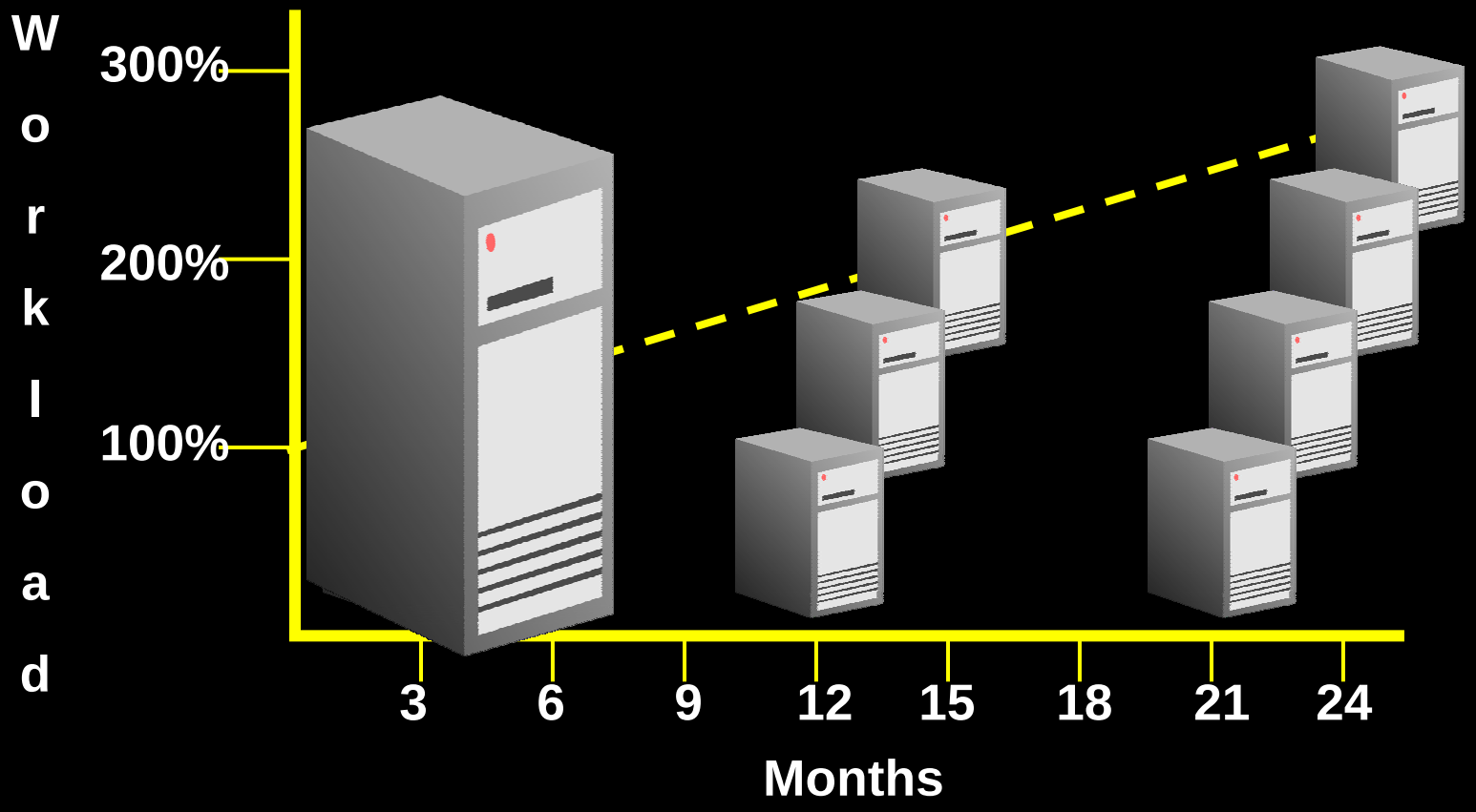
Scalability

- Required for performance and customer service
- Scaling up to another machine is expensive
- Plug-in scale-up and scale-down with the addition/removal of servers (nodes)
- Scale Up (the Old Approach)
  - Add CPU or memory to a server
    - Eg: Add 2 850 MHz procs and 2GB RAM quoted at $150K
- Scale Out (the New Approach)
  - Add more servers
    - Eg: Add dual 3.0 GHz server w/4GB RAM quoted at $24K

Value Proposition

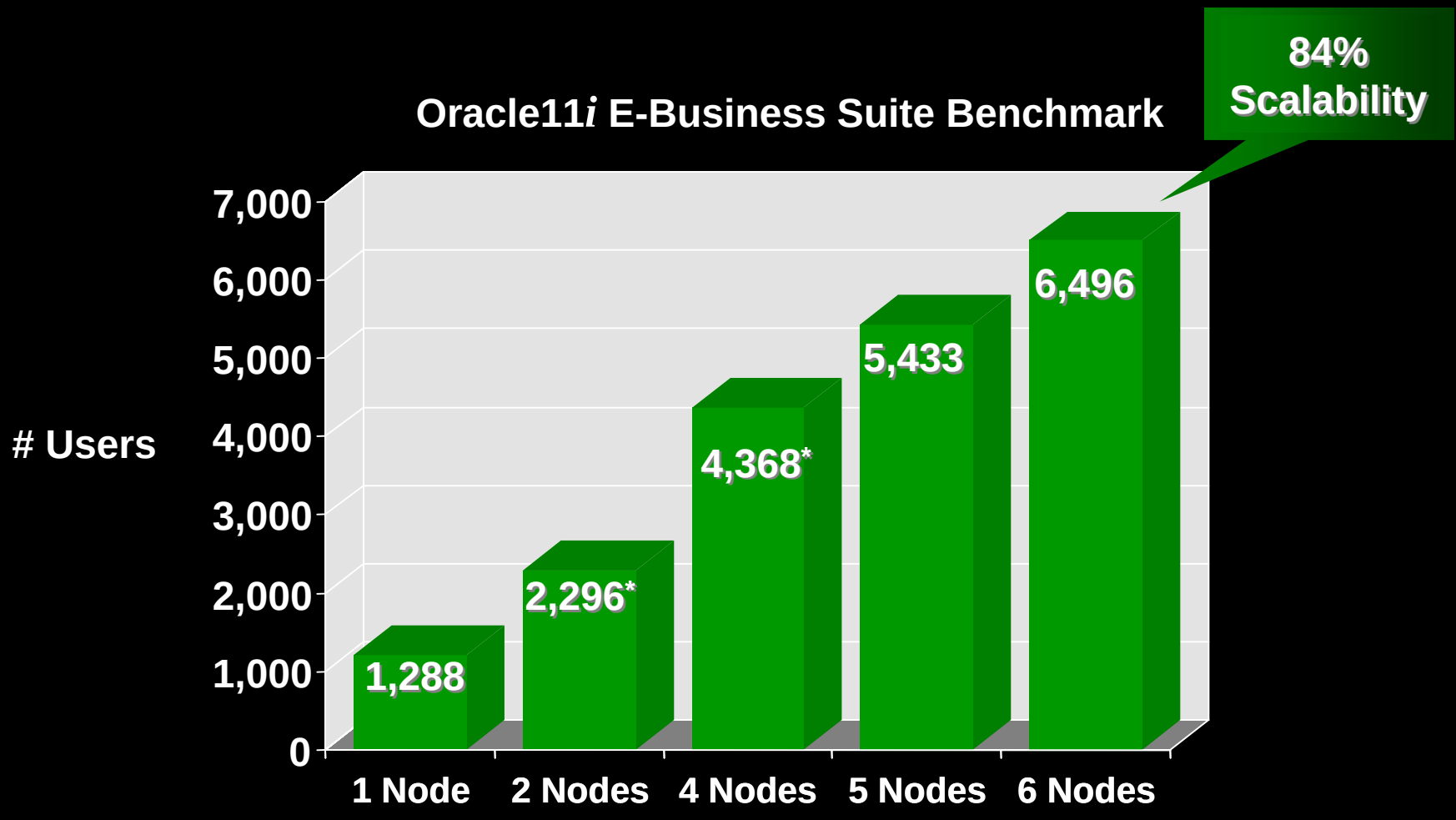
- Exploit commodity hardware
- Lower maintenance costs
- Lower down-time costs

Pay and Scale Incrementally



Caleb@Caleb.com

E-Business Suite RAC Scalability



Caleb@Caleb.com

Scalability Expectations

- If your application will scale transparently on SMP, then it is realistic to expect it to scale well on RAC, without having to make any changes to the application code.

- The “Spotlight Effect”
  - RAC is very good at putting a bright “spotlight” on poorly designed databases and applications.

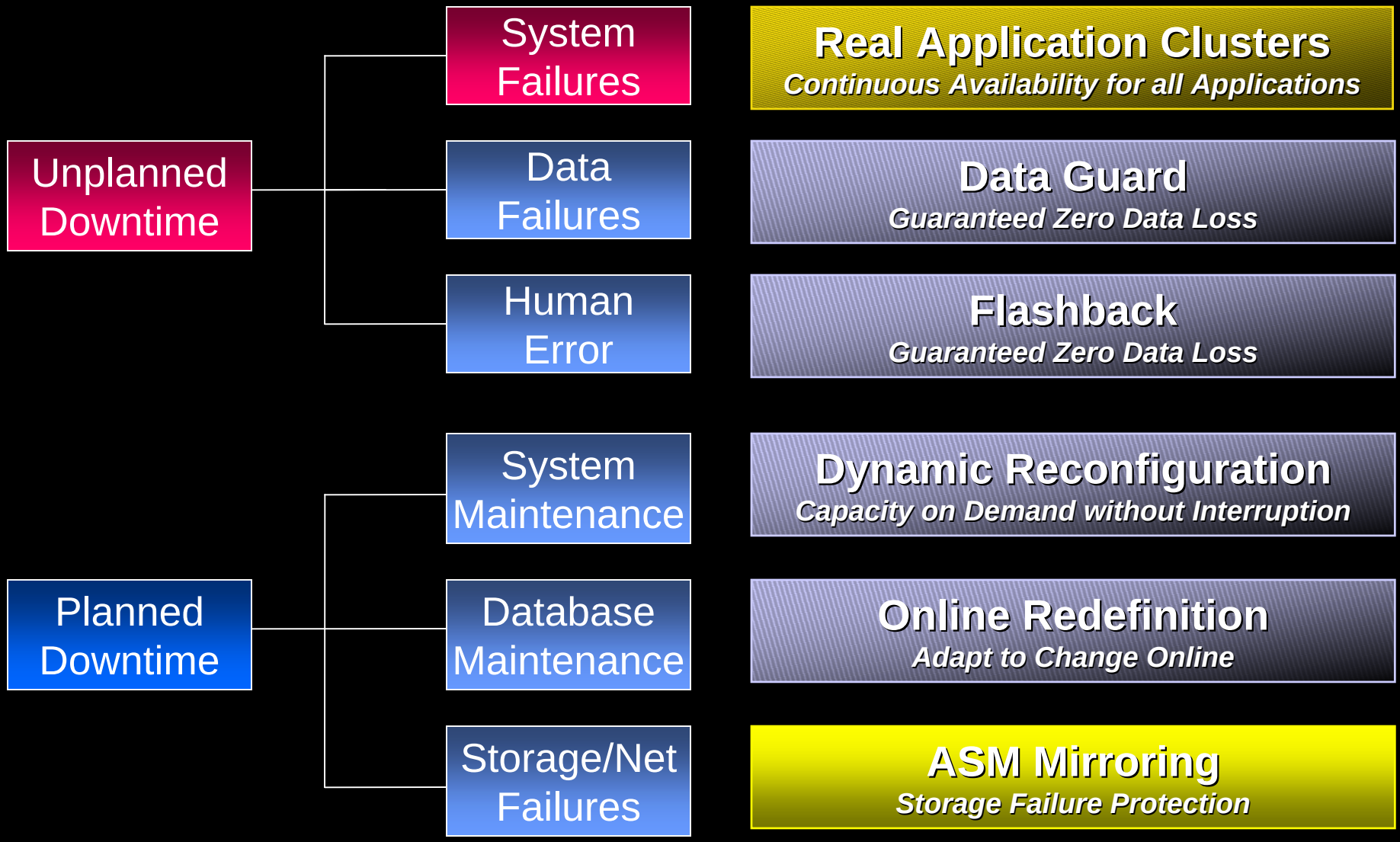
Reliability

- RAC eliminates the database instance, and the node itself, as a single point of failure, and ensures database integrity in the case of such failures
- RAC is ***one part*** a complete High Availability solution
- All the technology in the world will not make a poorly run shop reliable

Plan For Failures

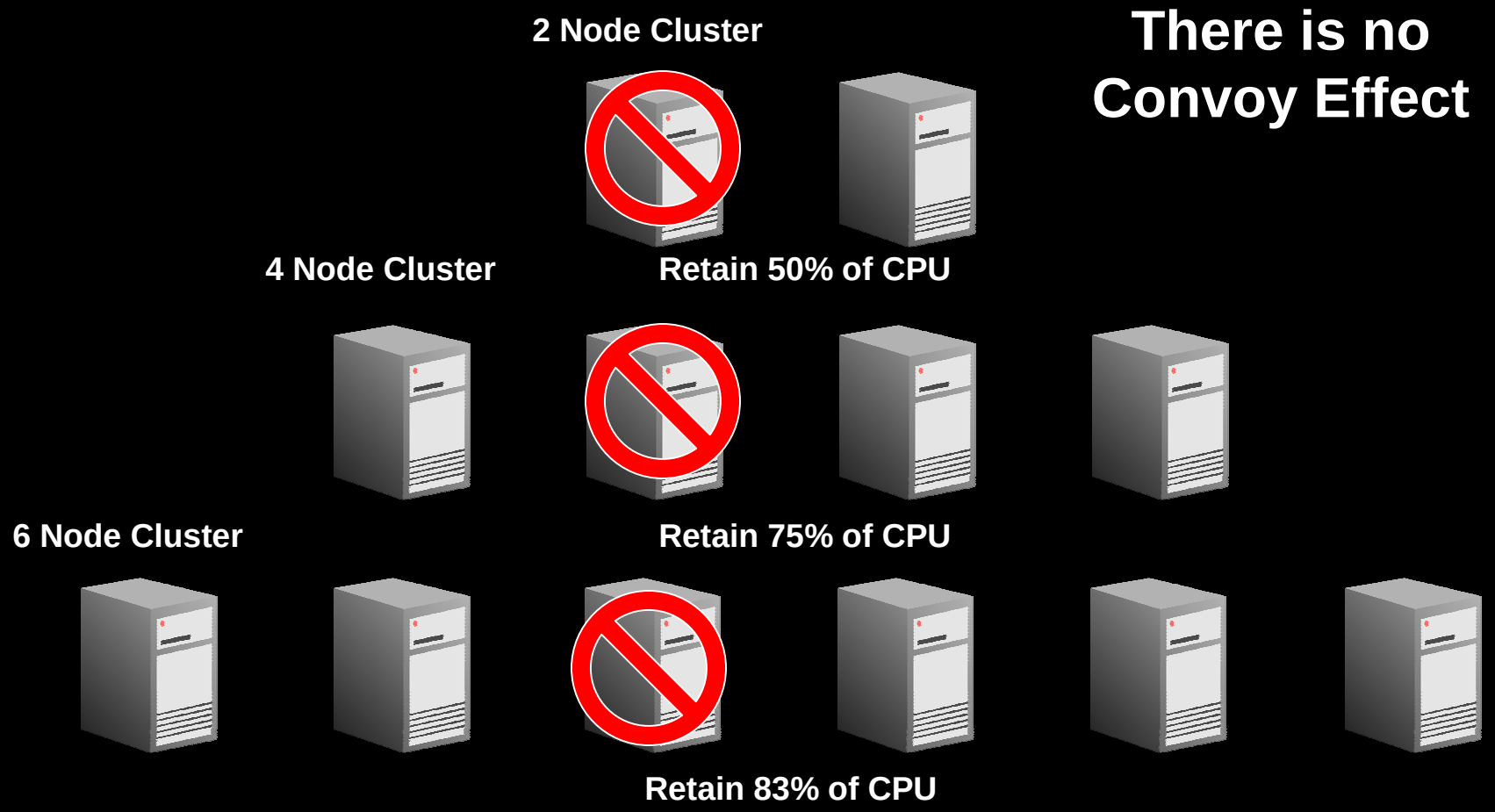
- Eliminate Single Points of Failure
  - Cluster interconnect redundancy (NIC bonding / teaming)
  - Implement multiple access paths to the storage array
- Sufficient nodes to provide adequate CPU in the event of failover
- Scalable I/O subsystems
- Plan for downtime
- Establish realistic Service Level Agreements

Fault Tolerance



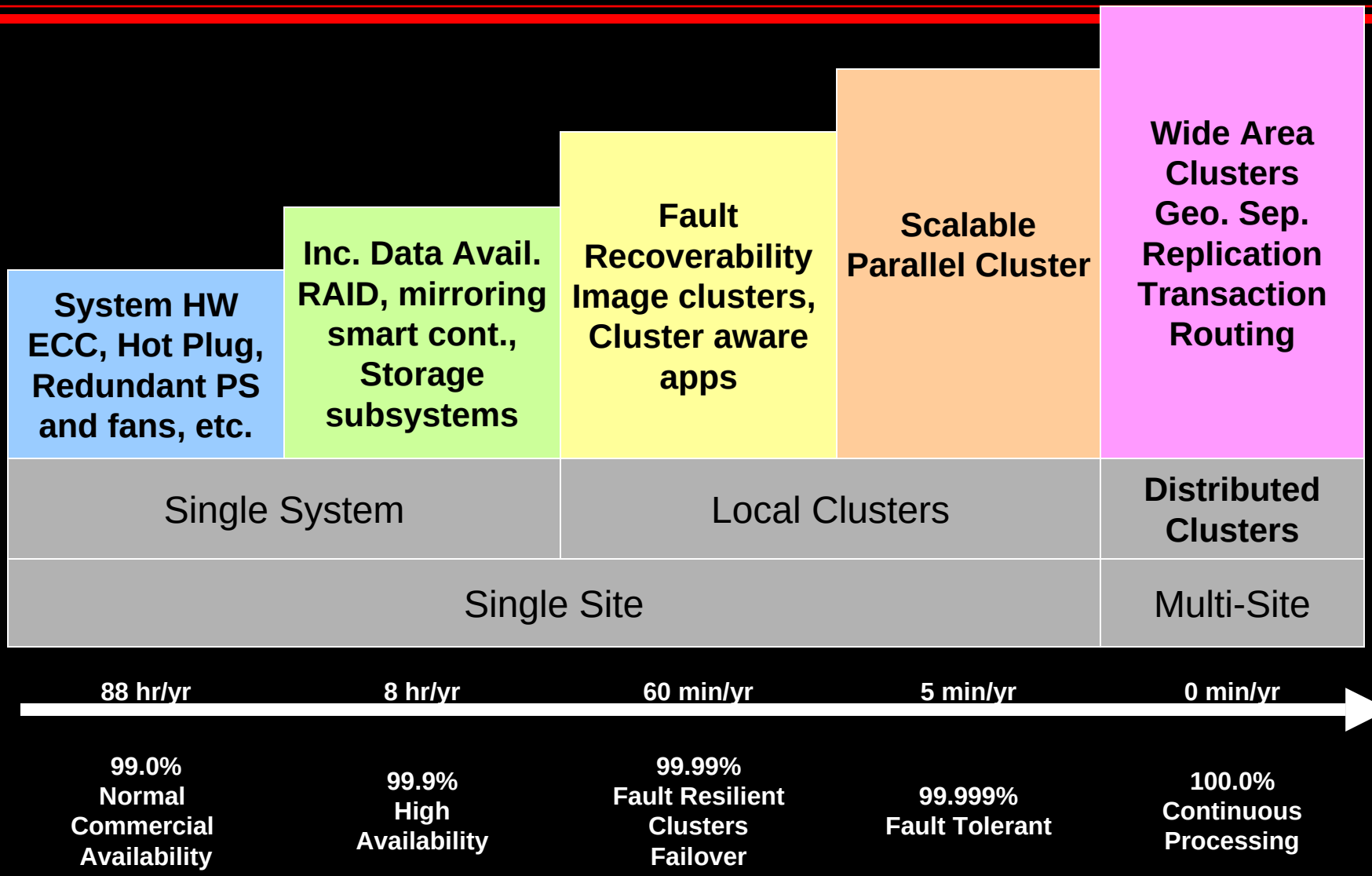
Caleb@Caleb.com

More Nodes = Higher Availability



Caleb@Caleb.com

High-Availability (HA)



Caleb@Caleb.com

A Complete High-Availability Solution

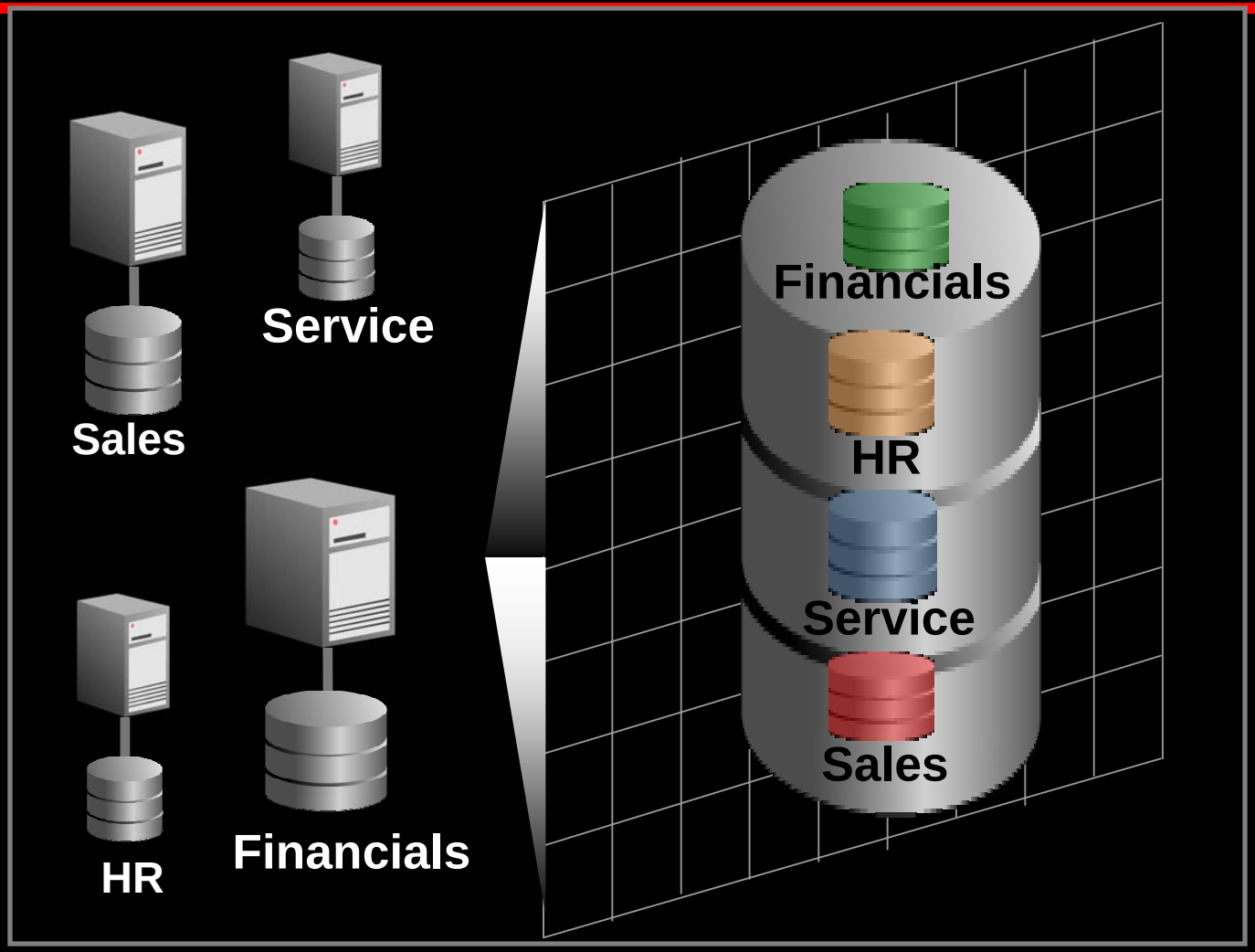
- Recovery Manager (RMAN)
- Flashback Database / Query / Table
- Streams Replication
- Grid Control
- Real Application Clusters (RAC)
- Data Guard

RAC is one part of a complete HA solution

Caleb@Caleb.com

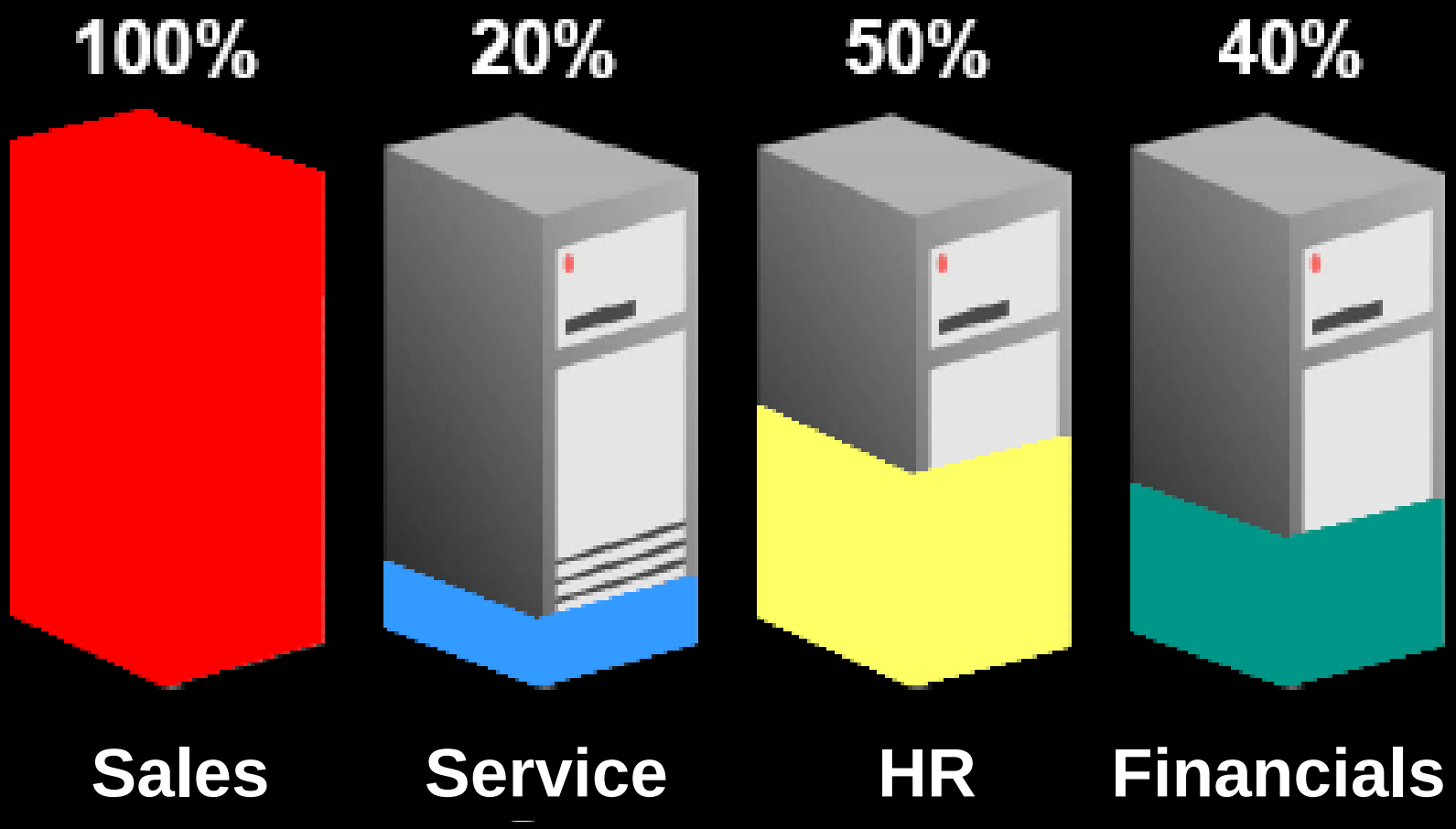
Server Consolidation

Single system image for competing workloads running across multiple systems



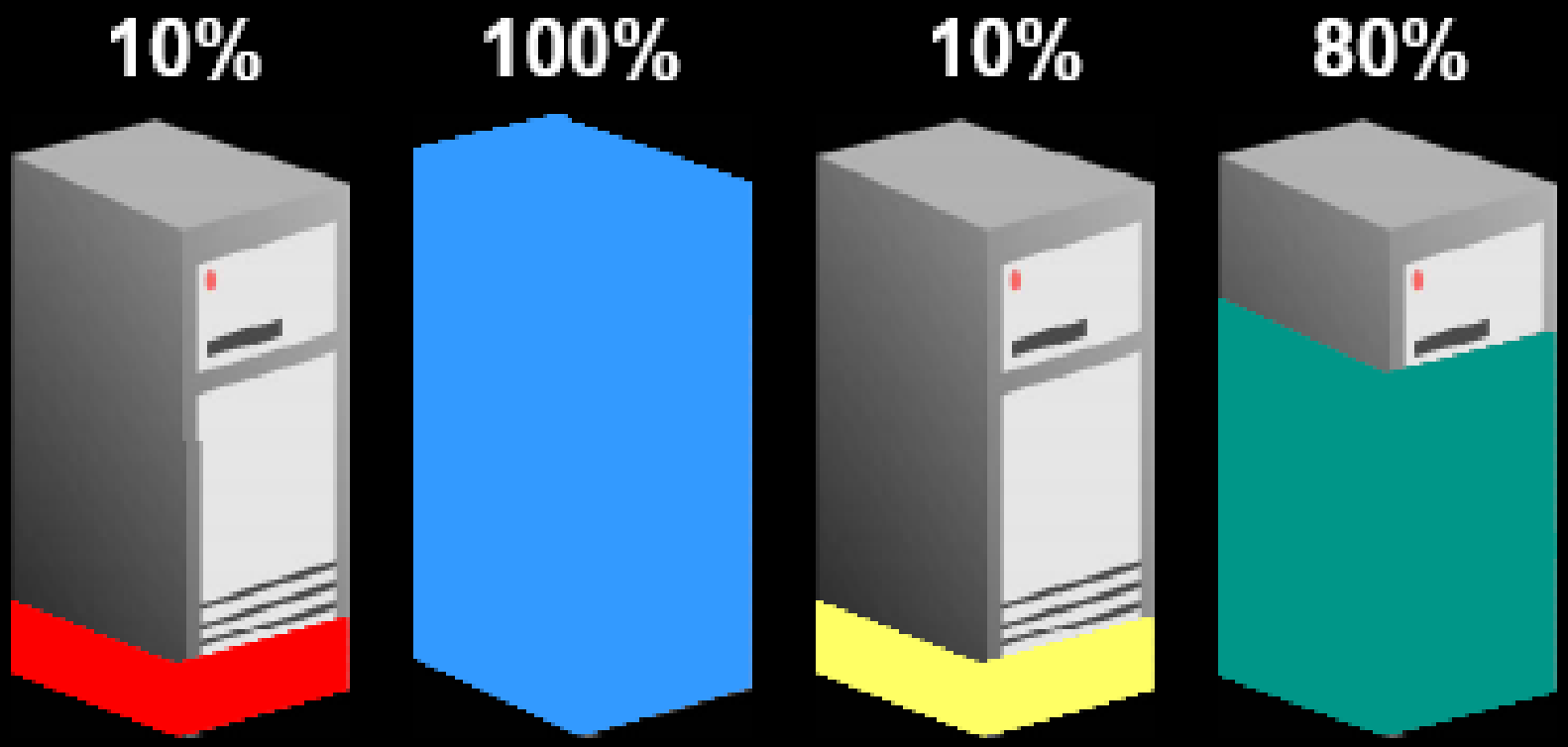
Caleb@Caleb.com

Discrete Servers



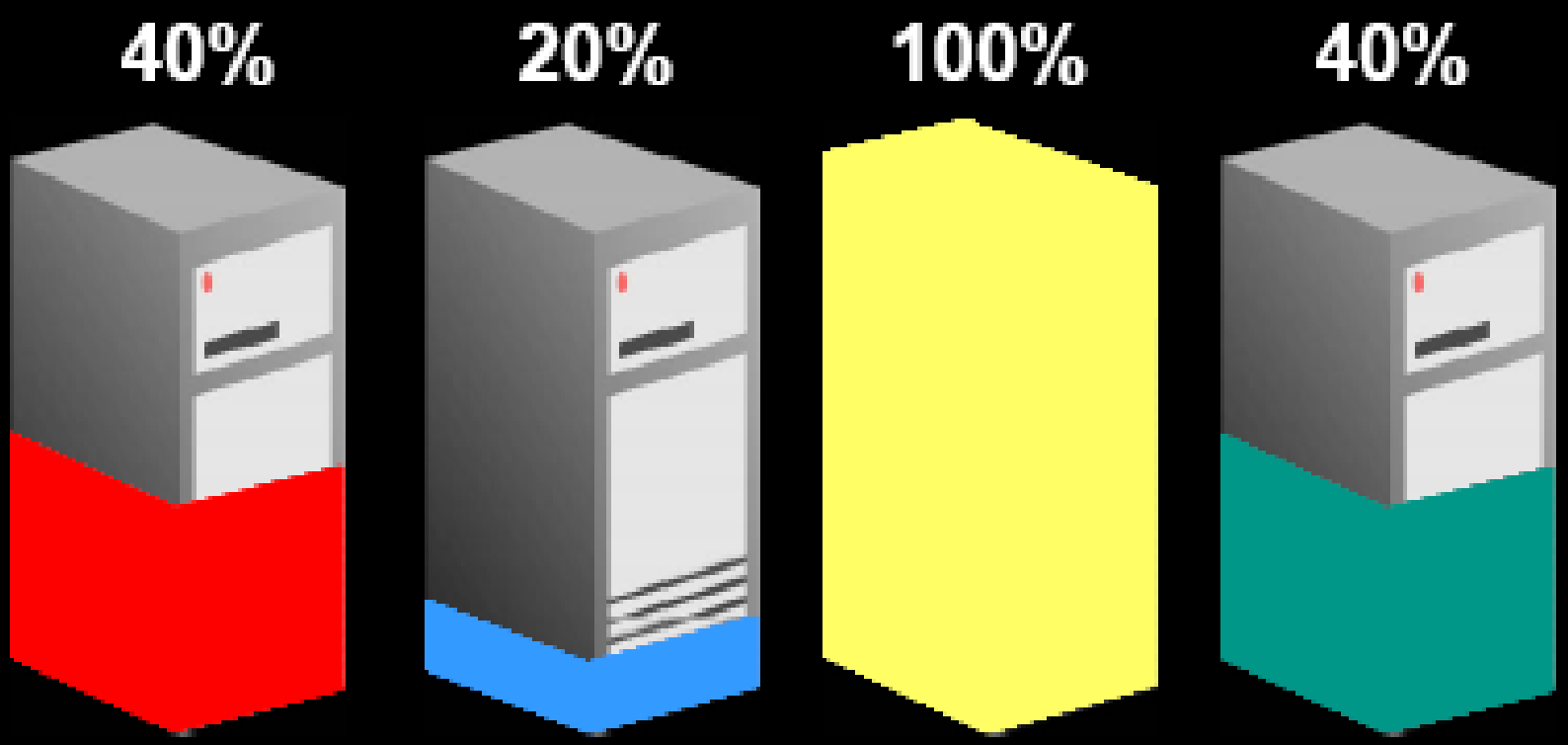
Caleb@Caleb.com

Discrete Servers



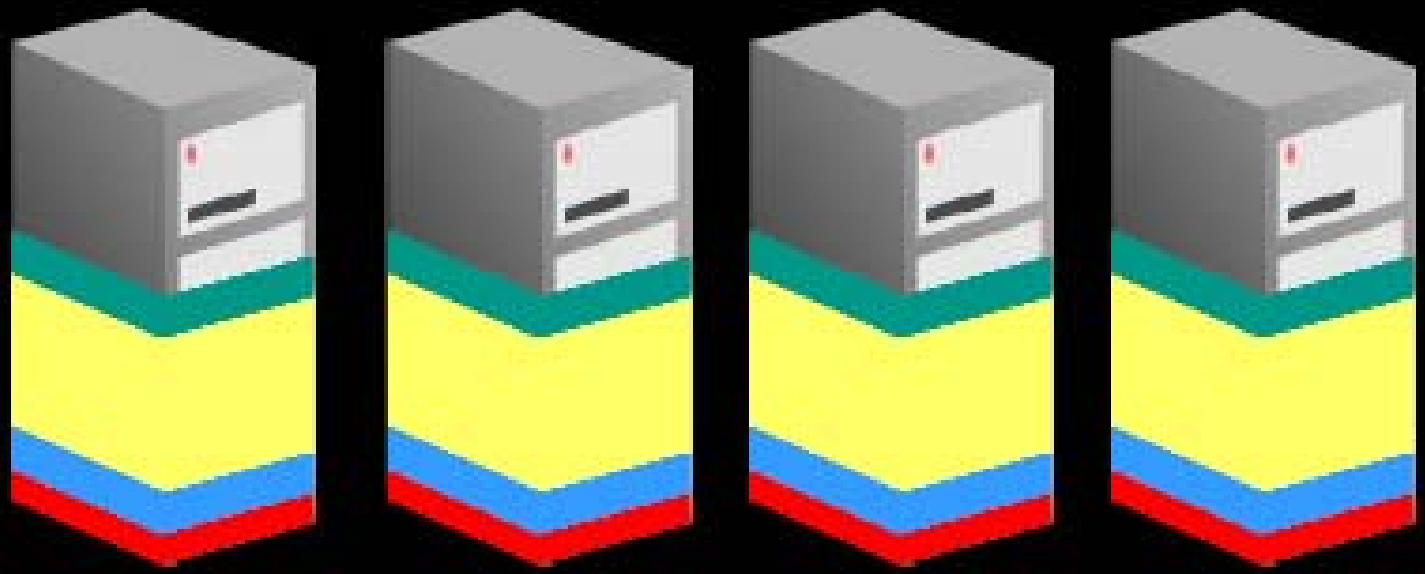
Caleb@Caleb.com

Discrete Servers



Caleb@Caleb.com

RAC Cluster



Finance

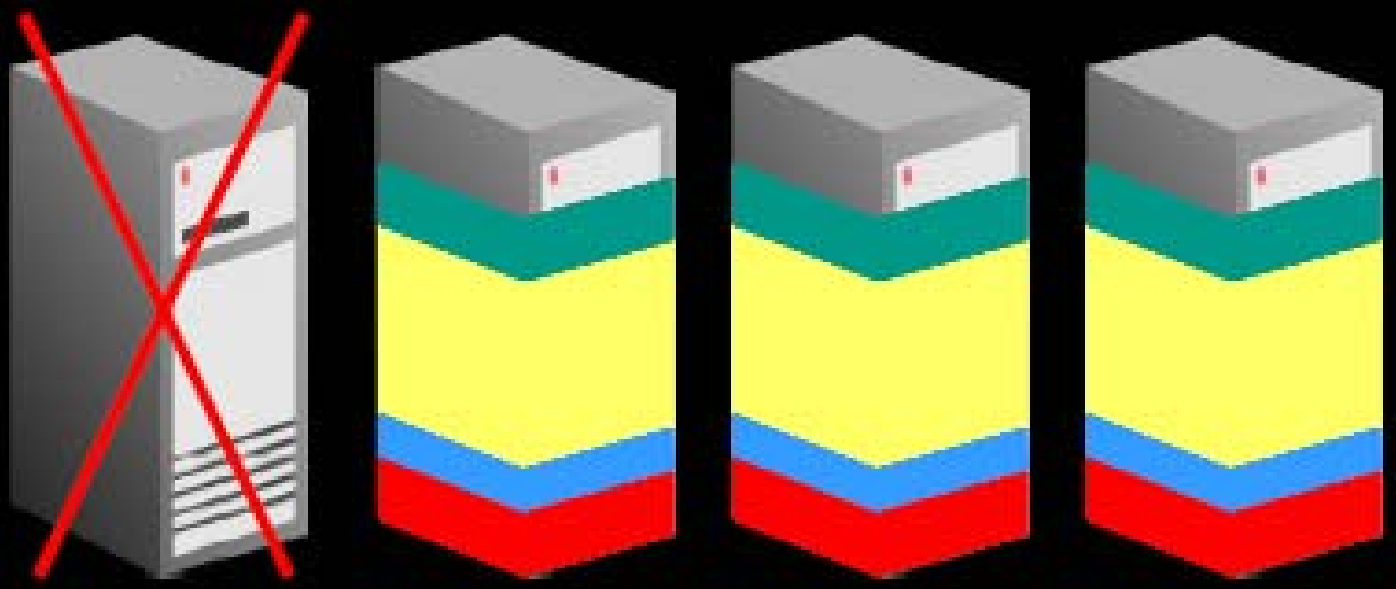
HR

Service

Sales

Caleb@Caleb.com

RAC Cluster After Failure



Caleb@Caleb.com

Assembling the Team

- Multiple skills are required
  - Traditional DBA plus RAC specific
  - System Administrator
  - Network Engineer
  - Storage Engineer
  - Application Developer
  - …

Assembling the Team

- Unusual architectures are involved
  - Private networks
  - Virtual IP addresses
  - Network bonding
  - Cluster File Systems
  - RAW stripes on SAN
  - Root access to servers
  - Multiple instance connections
  - Load balancing & failover technologies

Assembling the Team

- Managing the Team
  - Who's the boss?
  - The DBA "get's it" when things break
  - Someone (DBA) must have working knowledge of all technical aspects:
    - Hardware
    - Operating System
    - Networks
    - Storage
    - Clusterware
    - Database
    - Application Design

Choosing Hardware & OS

- Existing expertise
- Relationships with vendors
- Political reasons
- Price and value
- Technical specifications
- Be sure it is certified for RAC!

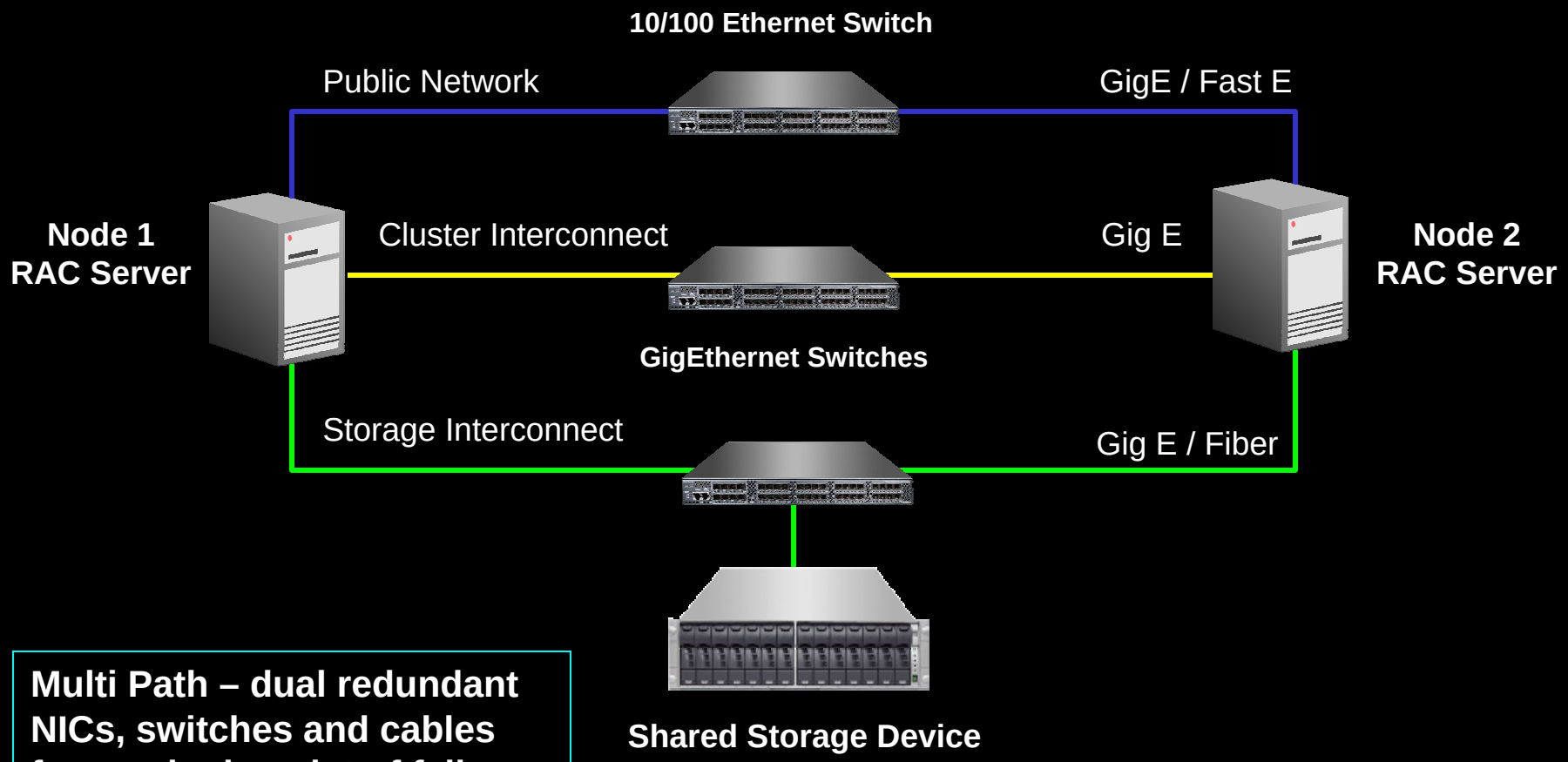
Planning the Networks

- Three separate & distinct networks:
  - "Public" interface
  - Cluster interconnect
  - Storage network
- Gig ether
- Fiber channel
- Each network should be fully redundant and bonded.

General Networking Concepts

- Public Networks
  - Used for application communication
  - Not recommended for cluster communication
- Private Networks
  - Used for cluster communication
  - Can not be seen by machines not connected directly to the private network
- Storage Networks
  - Database I/O

Typical RAC Network Topology



Multi Path – dual redundant NICs, switches and cables for no single point of failure

Caleb@Caleb.com

Planning the Storage

- Hardware storage solution
  - SAN (storage area network) eg. EMC
  - NAS (network attached storage) eg. NetApp
  - JBOD (just a bunch of disks)
- Shared File System solution
  - NFS (*certified* Network File System)
  - ASM (Oracle Automatic Storage Management)
  - OCFS (Oracle Cluster File System)
  - RAW (raw devices)
  - Proprietary (Veritas, IBM GPFS, etc)

Supported Storage Options

Storage Option	OCR and Voting Disks	Oracle Software Installation	Database	Recovery
Automatic Storage Management (ASM)	No	No	Yes	Yes
OCFS	Yes	No	Yes	Yes
OCFS2 (not cert.)	Yes	Yes	Yes	Yes
GPFS* for Linux on IBM POWER PC)	Yes	Yes	Yes	Yes
Local Storage	No	Yes	No	No
NFS File System	Yes	Yes	Yes	Yes
Shared Raw Partitions	Yes	No	Yes	No

* IBM General Parallel File System. Other CFS supported include DBE/AC (Veritas) and Tru64 CFS

Install and Configure Software

- OS install
- Oracle specific mods
- Network interface configuration
- CFS / ASM install
- Oracle Clusterware install
- Oracle Database binaries install
- Configure Oracle Net
- Create RAC database
- Configure Services, Resource Plans, Load Balancing & Failover…

OS Install

- Typically involves sys admin
- Select packages / components
- Install device drivers
- Recognize NICs / HBAs
- Configure NTP (network time protocol)
- Root access

Oracle Specific Mods

- Users & groups with consistent IDs
- Directories & permissions
- Kernel parameters
- Additional packages (eg. cvuqdisk)
- Configure hangcheck timer
- Environment variables
- Configure SSH

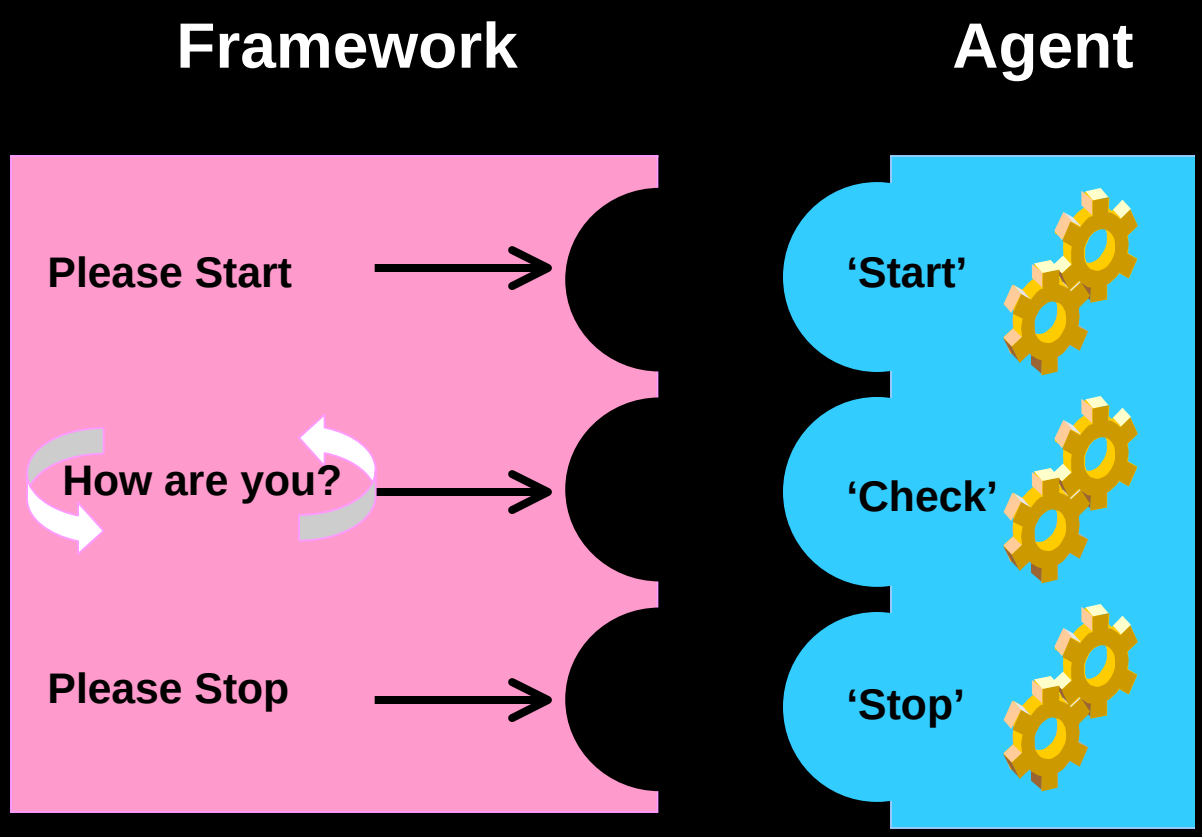
Network Interface Configuration

- Public Network
  - Physical (real) IPs
  - Virtual IPs
- Private Interconnect Network
  - Physical, private, non-routable IPs
  - Dual port, NIC bonding (different cards)
- Storage Interconnect
  - Dual gig E or fiber, bonded
  - Non-routable IPs

Oracle Clusterware Install

- Requires access to shared storage
- CFS or RAW but not ASM!
- Creates two critical shared files:
  - OCR (Oracle Cluster Registry)
  - Voting Disk
- Requires interconnect, storage and public networks in place
- Requires VIPs in place
- Will fail if any previous step missed

Clusterware Can Provide Application HA

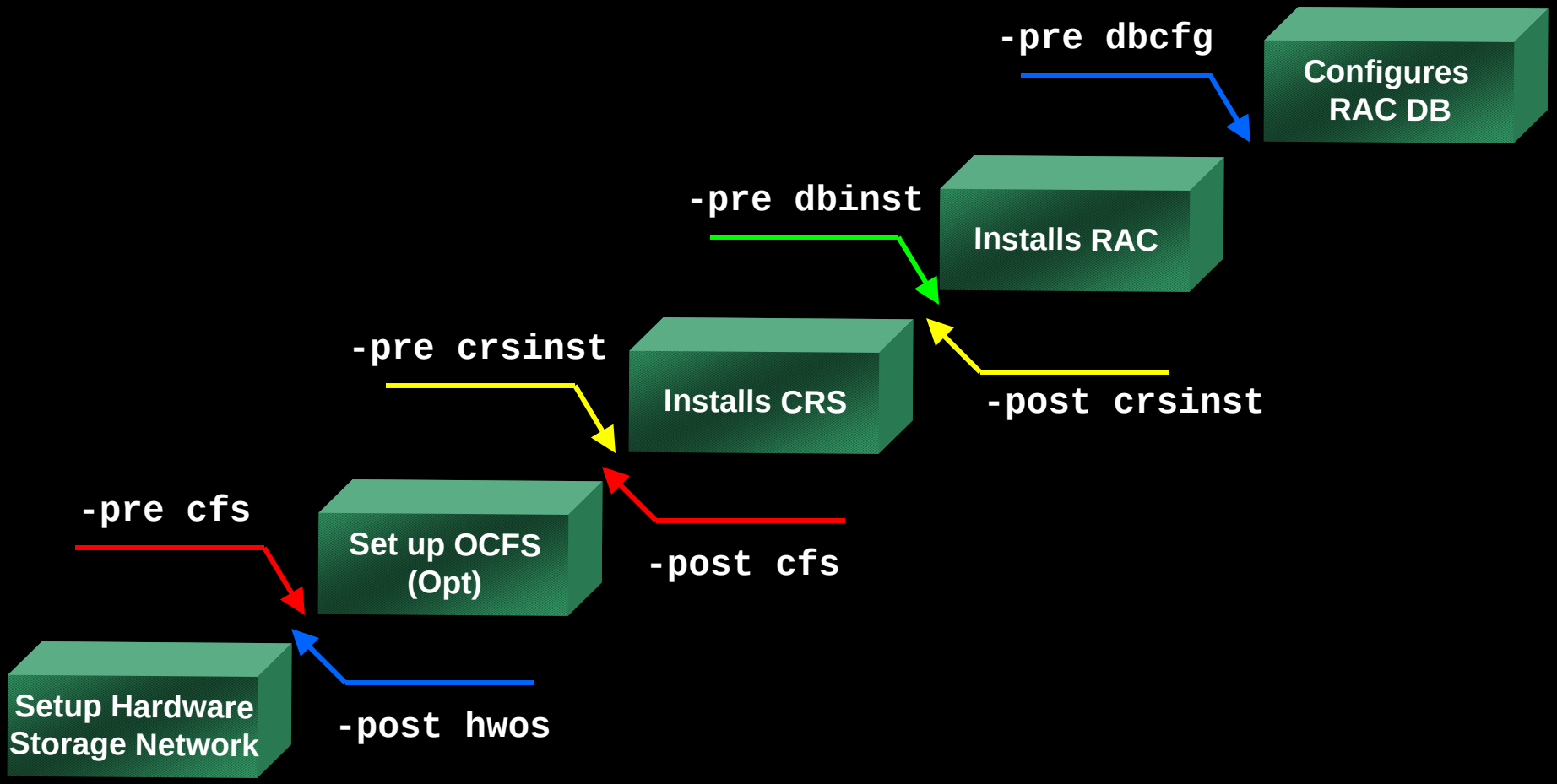


Caleb@Caleb.com

Oracle Clusterware Install

- Clusterware installs on top of CFS/ASM
- Separate ORACLE_HOME
- Installed from primary node only
- Required for a RAC database
- Provides the HA framework
- Runs as root, managed as root
- Monitors & controls the node
- Can be configured to control other apps

Cluster Verification Utility

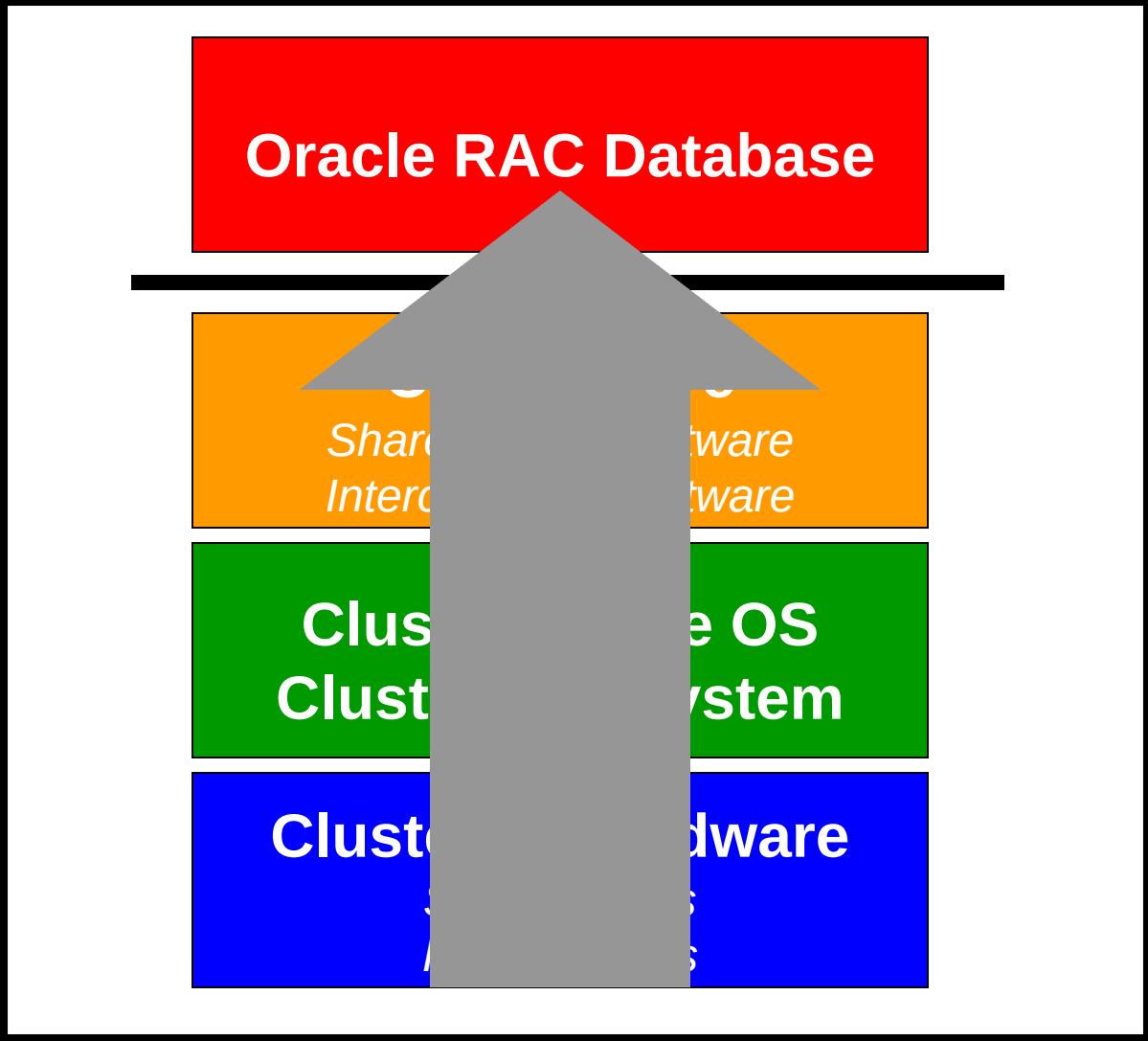


Caleb@Caleb.com

CFS / ASM Install

- Installs on top of shared storage
- Provides read / ***write*** file sharing
- NFS is plug-n-play, no installation
- Others require installation / formatting
- ASM is a separate Oracle product install in a separate ORACLE_HOME
- ASM is another whole topic….

The “Black Line”



Caleb@Caleb.com