

Successful deployment of Metrocluster for RAC



HP Integrity – The Most Trusted. Always.



©2009 Hewlett-Packard Development Company, L.P.
The information contained herein is subject to change without notice.

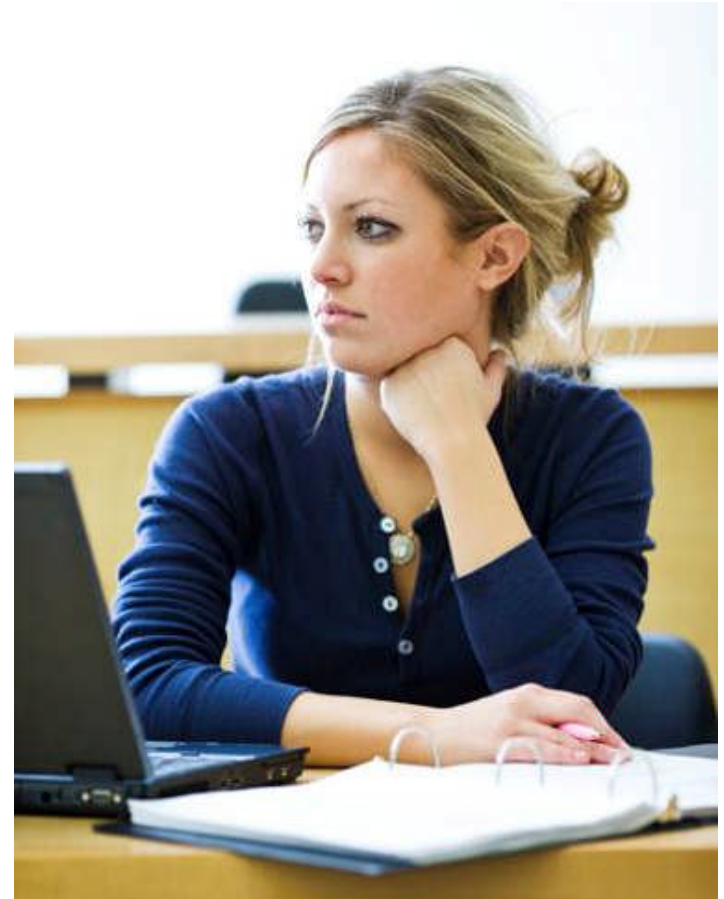
Introducing the speaker: John Foxcroft

- John Foxcroft has been with HP for 25 years, working in the Availability Clusters Solutions Lab for the past 10 years.
- His current responsibilities include training development, consulting and support planning which enables our HP field, channel partners and customers to architect, implement and support mission-critical highly available and disaster tolerance solutions on HP-UX and Linux.

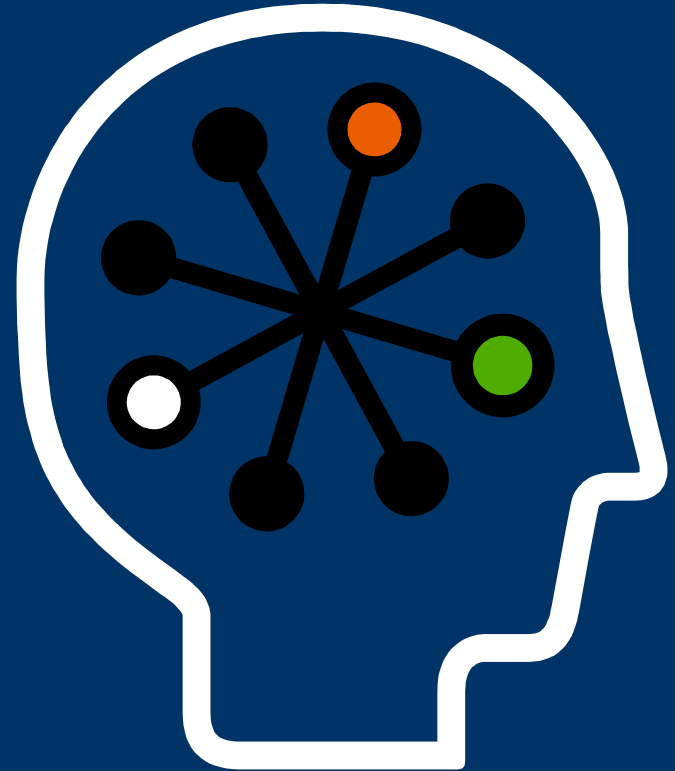


Agenda

- HP Serviceguard portfolio
- Metrocluster for RAC – a new Disaster Tolerant Solution designed for Oracle RAC
- References



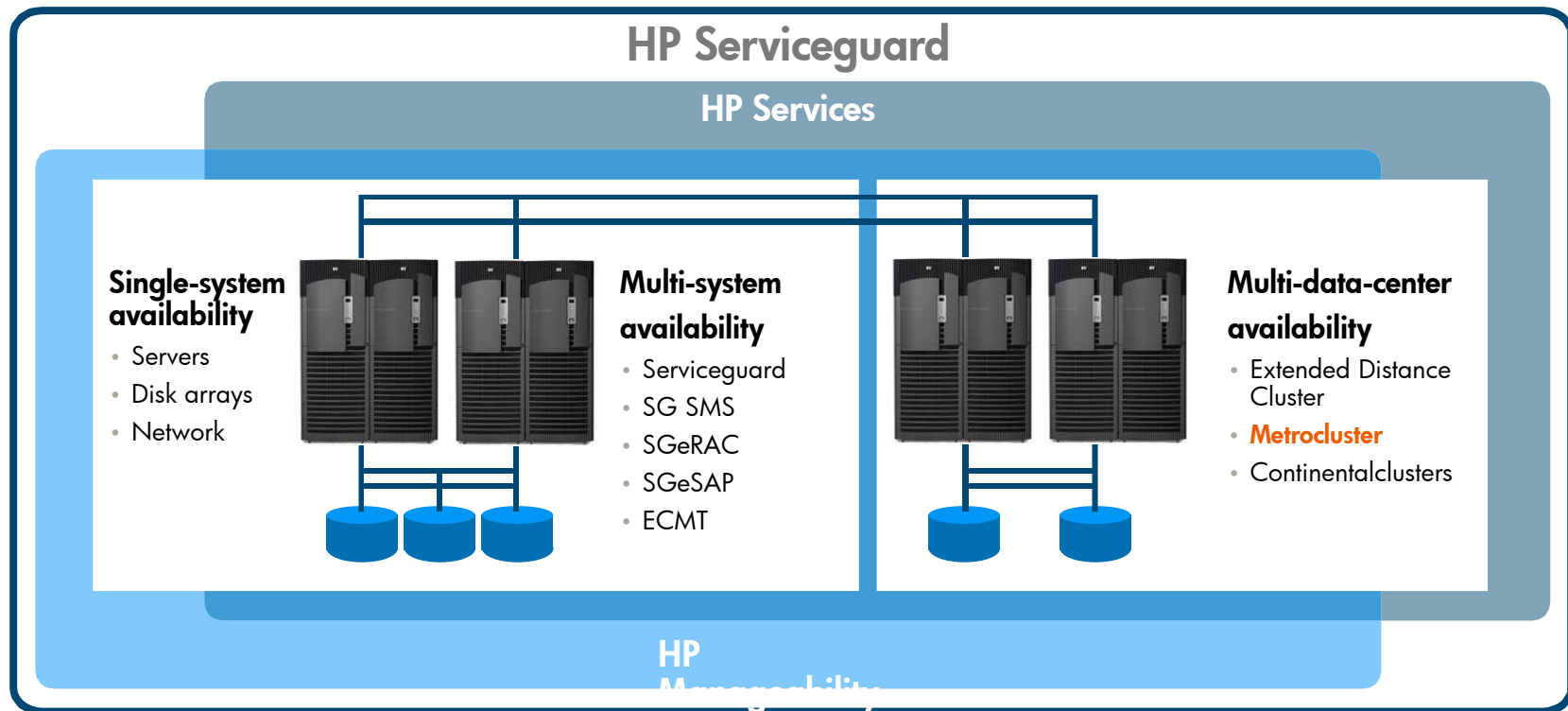
HP Serviceguard
portfolio for
HP-UX 11i v3



Keep your applications up and running

HP Serviceguard portfolio for HP-UX 11i

- Preserve your data integrity to protect your business
- Optimise planned and unplanned downtime



HP-UX Disaster Tolerant Solutions

Overview

- Extended Distance Cluster
 - A single-cluster architecture
 - Does not require any add-on software
- Metrocluster
 - A single-cluster architecture
 - Requires add-on software and special storage hardware
 - Supports Continuous Access XP, Continuous Access EVA, EMC SRDF as data replication
 - Metrocluster for RAC – Support for Oracle 10gR2 and 11gR1 RAC
 - Cross-subnet support
- Continentalclusters
 - A multi-cluster architecture, e.g., a primary and recovery cluster
 - Requires add-on software
 - Depending on data replication scheme, may require special storage hardware
 - Allow model for integrating any arbitrary data replication scheme
 - For example, for Oracle Data Guard, Veritas Volume Replicator (VVR)

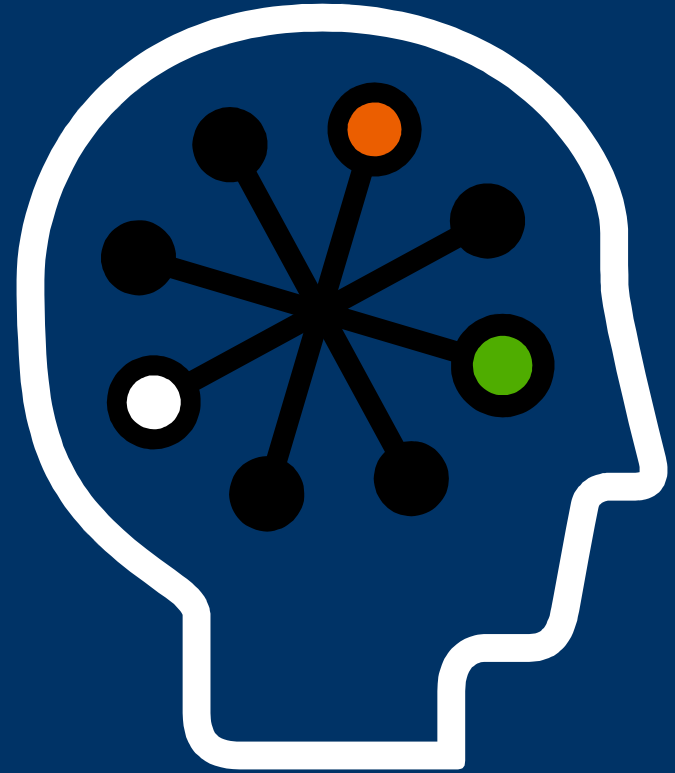


HP-UX Disaster Tolerant Solutions

Details

	Extended Distance Cluster	Metrocluster	Continentalclusters
# of Clusters	One cluster	One cluster	Multiple clusters
# of Nodes	Up to 16	Up to 16	Up to 64 nodes total
Distance	100KM – Software Mirroring	300KM – XP, EMC, EVA	Unlimited
Volume Manager	LVM, SLVM, VxVM, CVM	LVM, SLVM, VxVM, CVM	LVM, SLVM, VxVM, CVM (HP-UX 11i v2 only)
Data Replication	MirrorDisk UX VERITAS Mirroring	Continuous Access SRDF	Continuous Access, SRDF, Oracle Data Guard
Serviceguard Extension for RAC Support	Yes	Yes	Yes
Failover Direction	Bi-directional	Bi-directional	Bi-directional
Storage	XP, EVA, EMC	XP, EVA, EMC	XP, EVA, EMC
Servers OS version	HP Integrity and HP 9000 HP-UX 11i v2, 11i v3	HP Integrity and HP 9000 HP-UX 11i v2, 11i v3	HP Integrity and HP 9000 HP-UX 11i v2, 11i v3

Metrocluster for RAC introduction



Metrocluster for RAC detailed agenda

- Metrocluster for RAC introduction
- Site Aware Disaster Tolerant Architecture
- Metrocluster for RAC – key benefits
- New technologies
- Supported architectures and examples
- Release schedule
- Software and hardware requirements
- References



HP's new Disaster Tolerant Technology: Metrocluster for RAC

- Based on existing Metrocluster solution
 - A new feature of Metrocluster (not a new product)
 - Adds support for Oracle RAC (multi-instance database)
 - Supports multiple IP subnets
 - Supported by HP and Oracle
 - Currently deployed in production by HP-IT and a European government organisation
- Array-based replication for robust data replication
 - HP XP Continuous Access
 - HP EVA Continuous Access
 - EMC DMX Symmetrix Remote Data Facility
- Replicated data: CFS, CVM or SLVM



Metrocluster for RAC introduction

Database Stacking, Static Load Balancing, and Disaster Tolerance for Oracle RAC

- Fully automatic and automated failover
- Choice of raw volumes (CVM, SLVM) or Cluster File System (CFS) storage deployment
- Simplified deployment and administration
- Active/active RAC database environment
 - Multiple nodes within a data center (site) running the same RAC databases
 - Different databases running concurrently at each site
- Site controller package to control package and storage failover
 - In a multi-instance application environment
- Allows operator-initiated static load balancing
 - By moving RAC databases between sites based on performance/load measurements
- Supports many stacked RAC databases



Metrocluster for RAC introduction

Continued

- Design goals
 - Have no impact to performance of RAC's Cache Fusion network due to distance
 - Have no impact to performance of CFS Interconnect due to distance
 - Provide scale-up and scale-out environment for
 - The best performance
 - Highest level of availability and disaster tolerance
 - Rapid recovery
 - 16-node cluster (up to 8 nodes at each site)
 - Nodes need be attached only to the storage within the site
- Network provisioning
 - No need to provision cross-site network for storage access by nodes in a different site
 - Cross-site network supports
 - Multiple IP subnets
 - Serviceguard cluster membership protocols
 - Array-based data replication (FibreChannel)
 - Oracle Cache Fusion and Serviceguard Cluster File System protocols stay within a site

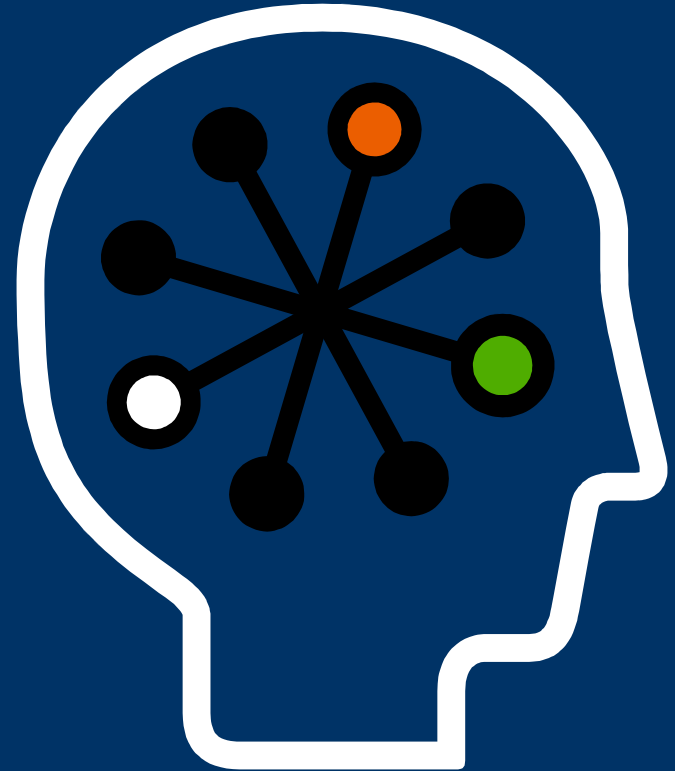


Site Aware Disaster Tolerant Architecture

Site Aware Disaster Tolerant Architecture (SADTA)

- SADTA is an architecture that enables deploying complex workloads in a Metrocluster environment.
- Complex workloads are applications that are configured using multiple inter-related MNP and failover packages that must be managed collectively. For example, Oracle Database RAC is considered a complex workload using this architecture.
- Metrocluster now supports a Site Aware Disaster Tolerant Architecture. Currently, the Oracle Database 10gR2/11gR1 RAC workload or other generic complex workloads are supported in SADTA.
- Metrocluster support of Oracle RAC is also called "Metrocluster for RAC".

Metrocluster for RAC benefits



Metrocluster for RAC: Key benefits

A site-aware disaster tolerant architecture for Oracle RAC

- Fully integrated into the Serviceguard portfolio of products
 - Serviceguard Storage Management Suite for CFS RAC
 - Added value of solution stack with SGeRAC
 - Manageability (HP SIM/SMH) and VSE integration benefits
- Disaster tolerance without impact to Oracle RAC performance
- Fully automated, bidirectional solution
- Storage and network flexibility to suit many customer environments
- Fully integrated and supported by HP and Oracle



Metrocluster for RAC: Fully-automated*

Features

- Automates the orderly startup and shutdown of all components
- Automatically monitors health of servers, storage and networking to provide rapid response to failures
- Automatically coordinates the complete failover process within the primary datacentre and to a remote recovery datacentre for major failure events

Benefits

- Minimizes unplanned downtime
- Mitigates otherwise error-prone and time-consuming manual processes

* Note: Recovery can be triggered manually if necessary, by simple configuration of auto-run and package node-switching flags

Metrocluster for RAC: Full utilization of assets

Features

- Oracle RAC databases may be highly stacked to support consolidation initiatives
- Servers in each datacentre may be fully utilised and actively run distinct Oracle RAC databases, each with a standby replica in the other datacentre for disaster tolerance
- Failover and site recovery is fully bidirectional
- Operator-initiated load-balancing allows RAC databases to be moved easily between sites

Benefits

- Effective use of server resources
- Remote site can perform useful work

Metrocluster for RAC: Storage flexibility and performance

Features

- Flexibility to use a variety of Storage arrays to match their business needs
- Supports EMC SRDF, EVA and XP Continuous Access hardware data replication
- Use the Serviceguard Cluster File System with Oracle Disk Manager for ease of manageability without significant loss of performance
- Can also deploy Oracle RAC on raw volumes (CVM or SLVM) for the best performance

Benefits

- Solution can be sized and scaled to customer's business needs
- Flexibility makes it easier to fit the solution into existing maintenance and management processes

Metrocluster for RAC: Network flexibility and performance

Features

- Supports different subnets in primary and remote sites
 - Frequently an organisational requirement for protecting networking assets and meeting auditing requirements
- Data replication distances supported up to 300km to provide geographical separation
 - No impact to Oracle RAC performance because Cache Fusion interconnect is always routed within a single datacentre only
- Redundant networking and rapid automatic network failure detection minimises downtime due to outages

Benefits

- Metrocluster for RAC can meet a flexible choice of existing network infrastructure that exist within the infrastructure
- Oracle RAC availability and performance unaffected by Disaster Tolerance architecture

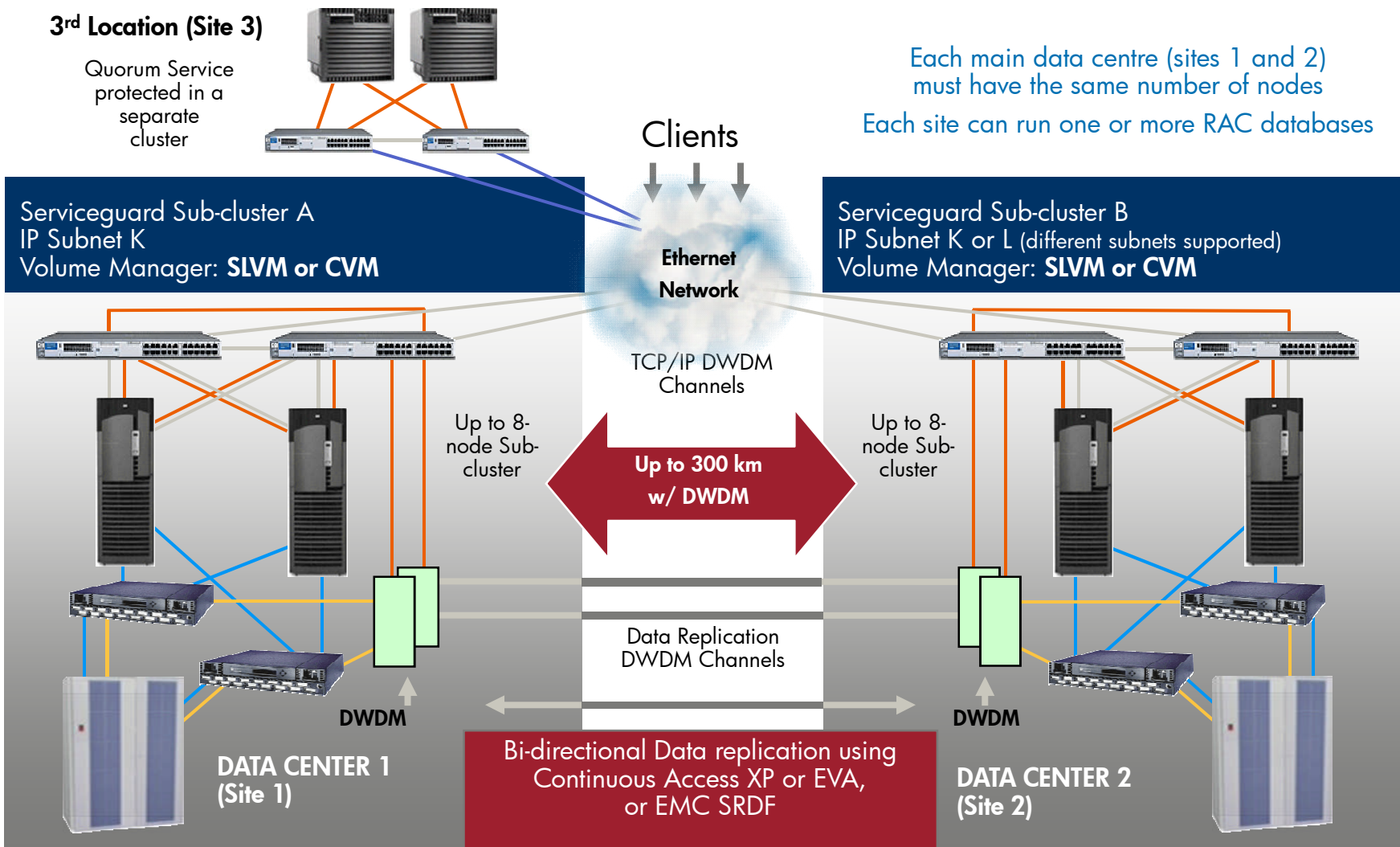
Key new technologies with Metrocluster for RAC

- Serviceguard
 - Site-awareness
 - Cluster File System sub-clustering
 - Cross-subnet support within a single cluster architecture
- SGeRAC
 - Oracle Clusterware sub-clustering
- Metrocluster
 - Shared volume activation by site
 - Site controller for automatic site failover and storage preparation
 - Smart control of storage according to primary or replica role
 - Conversion of replica from read-only to read-write upon disaster failover
 - Site Safety Latch
 - Prevents redundant workloads from running simultaneously on both sites

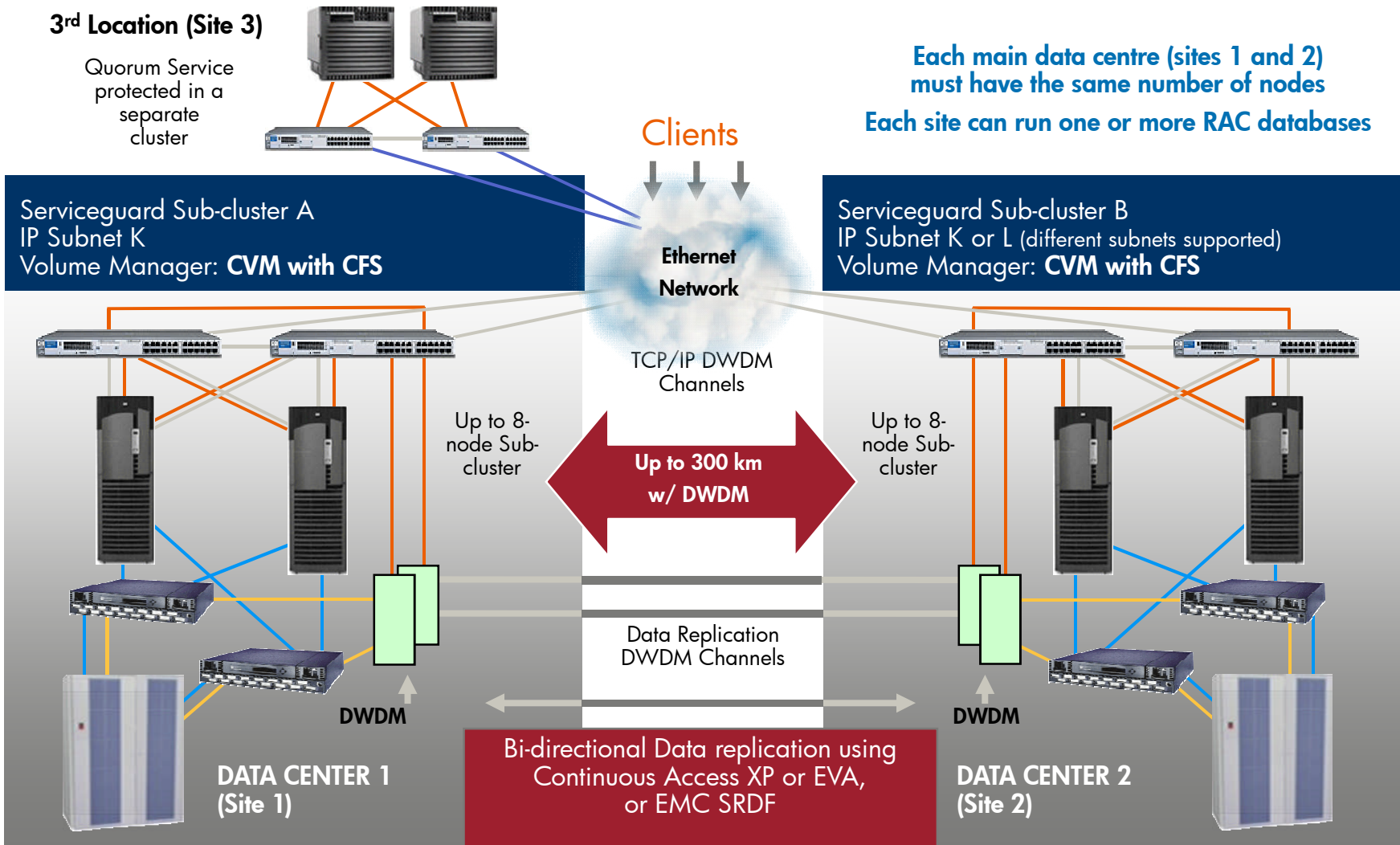
Metrocluster for RAC supported architectures and examples



Metrocluster for RAC with raw volumes



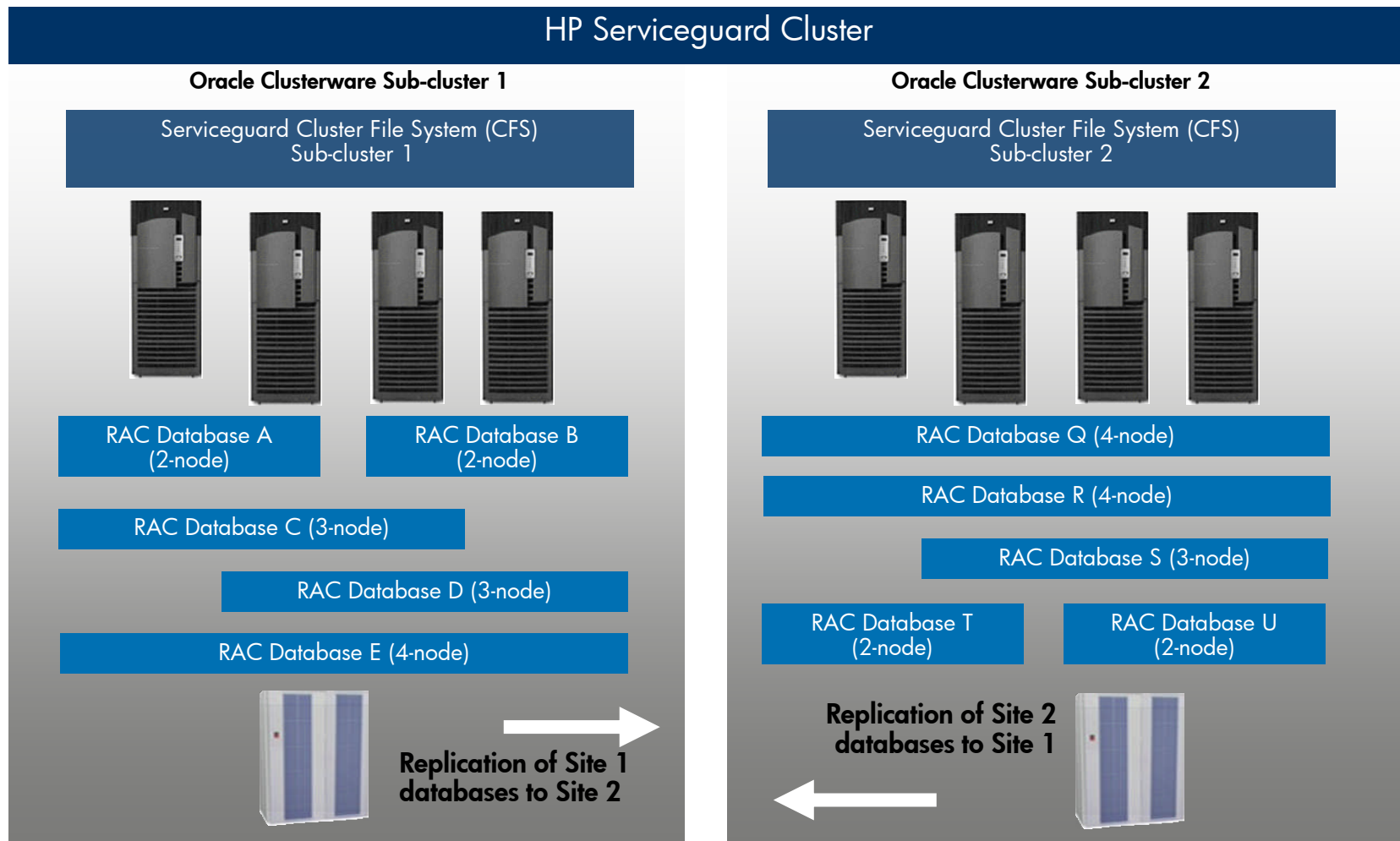
Metrocluster for RAC with Cluster File System (CFS)



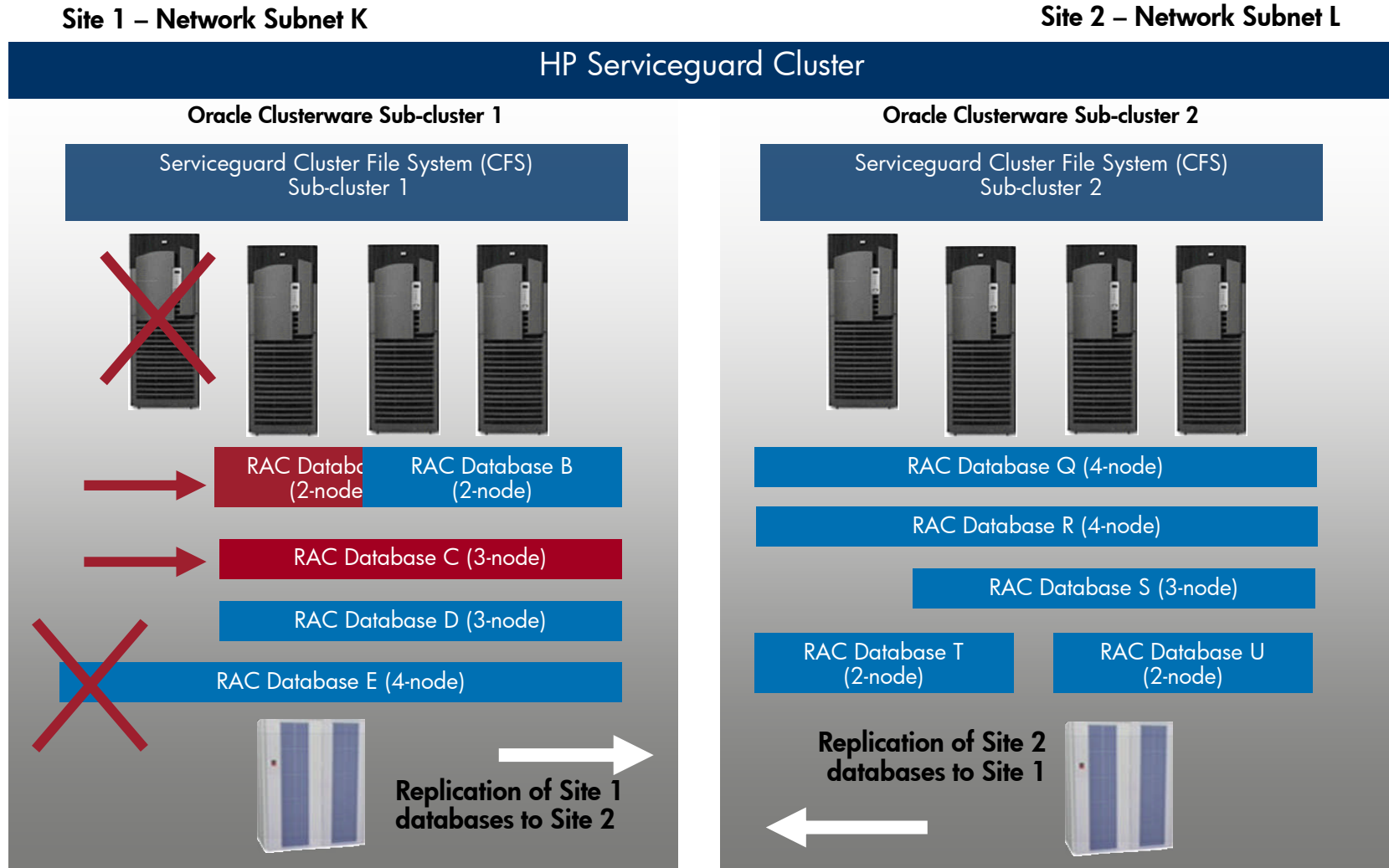
Stacked Deployment example

Site 1 – Network Subnet K

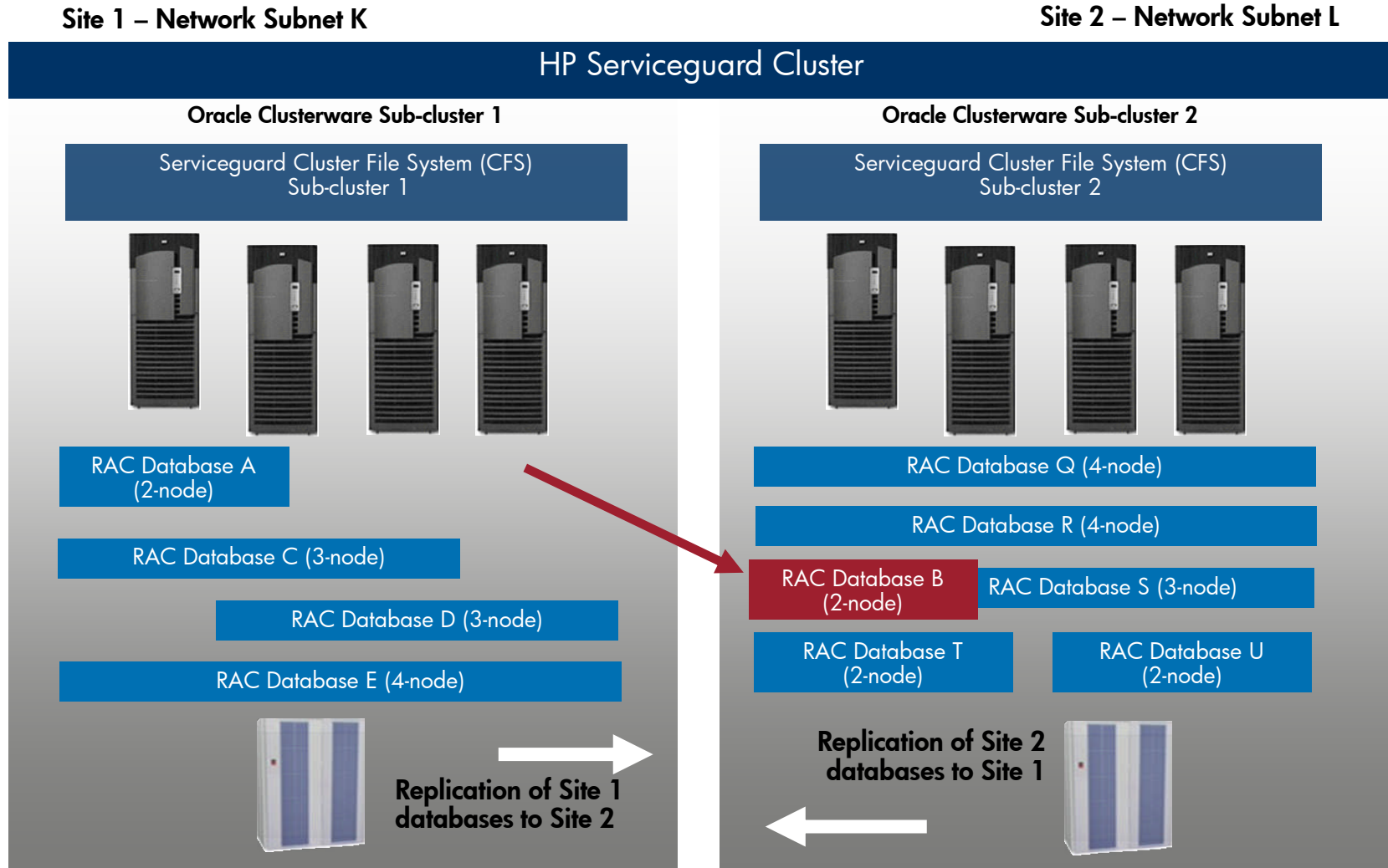
Site 2 – Network Subnet L



Stacked Deployment example: Node failure



Stacked Deployment example: Static load balancing



Stacked Deployment example: Dynamic load balancing

Site 1 – Network Subnet K

Site 2 – Network Subnet L

HP Serviceguard Cluster

Oracle Clusterware Sub-cluster 1

Oracle Clusterware Sub-cluster 2

Serviceguard Cluster File System (CFS)
Sub-cluster 1

Serviceguard Cluster File System (CFS)
Sub-cluster 2



RAC Database A
(2-node)

RAC Database B
(2-node)

RAC Database Q (4-node)

RAC Database R (4-node)

RAC Database C (3-node)

ADD 4th node

RAC Database S (3-node)

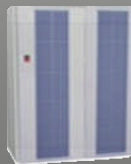
RAC Database D (3-node)

RAC Database T
(2-node)

RAC Database U
(2-node)

RAC Database E (4-node)

Oracle services can
be used to
dynamically
change nodes
where DBs run

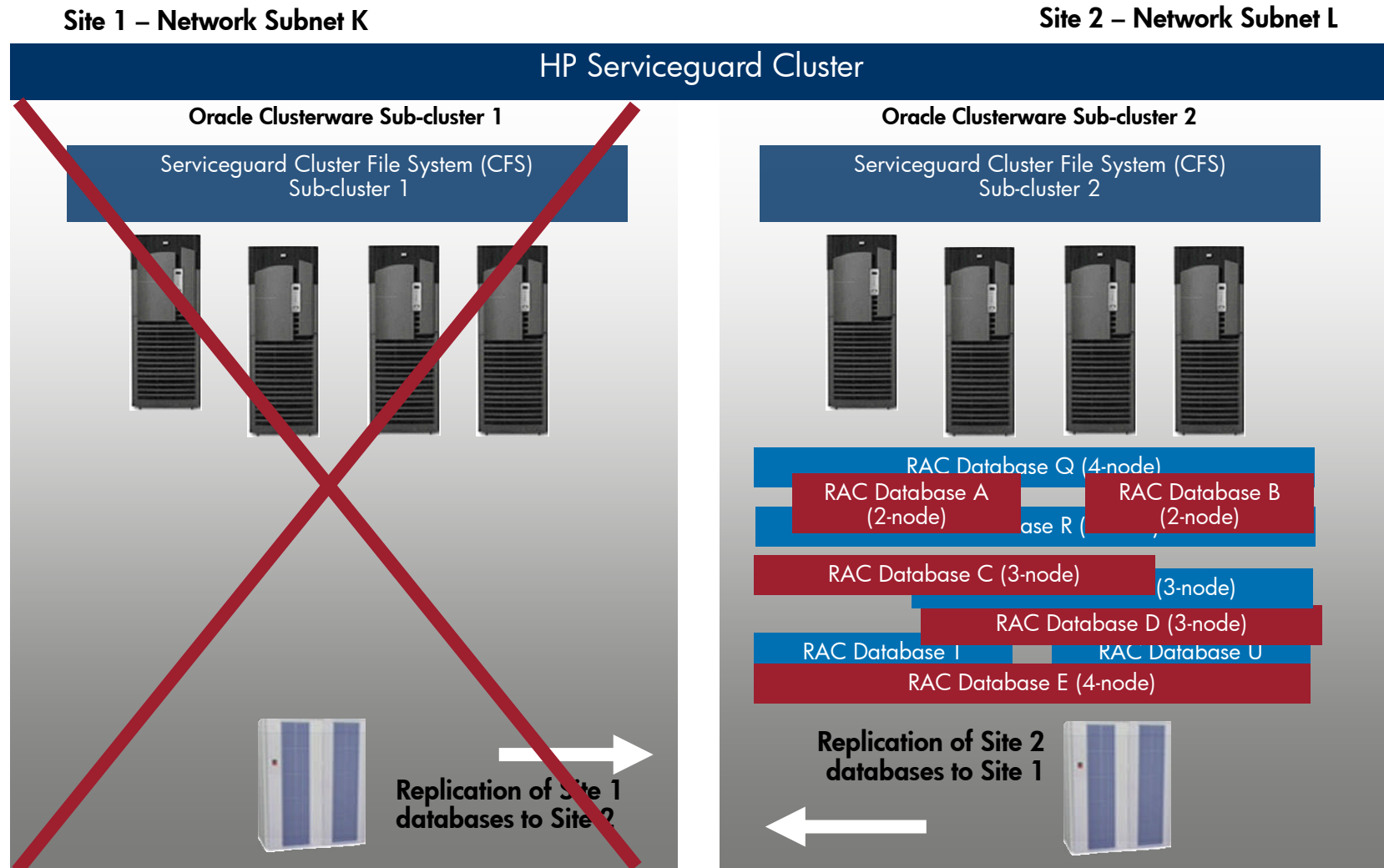


Replication of Site 1
databases to Site 2

Replication of Site 2
databases to Site 1



Stacked Deployment example: After a disaster



Metrocluster for
RAC release
schedule and
support information



Metrocluster for RAC release schedule

- General availability in September 2008 Application Release (AR0809)
 - SG/SGeRAC A.11.18 on HP-UX 11i v2 or 11i v3
 - Oracle RAC 10gR2
 - Maximum cluster size
 - 16 nodes with SLVM and Raw Volumes
 - 16 nodes with CVM and CFS 4.1 and 5.0
 - HP XP, HP EVA, and EMC DMX storage
- **March 2009 Application Release (AR0903)**
 - SG/SGeRAC A.11.18 or A.11.19 on HP-UX 11i v2 or 11i v3
 - Oracle RAC 10gR2 or 11gR1
 - Maximum cluster size
 - 16 nodes with SLVM and Raw Volumes
 - 16 nodes with CVM and CFS 4.1 and 5.0
 - HP XP, HP EVA, and EMC DMX storage
- Patches required



Metrocluster for RAC: Software requirements for Sep 2008 release

For HP-UX 11i v2 or HP-UX 11i v3

- **Metrocluster**

- Metrocluster with Continuous Access XP version A.08.00
- Metrocluster with Continuous Access EVA version A.03.00
- Metrocluster with EMC SRDF version A.07.00

- **SLVM**

- SLVM 1.0 (HP-UX 11i v2 and 11i v3)
- SLVM 2.0 (HP-UX 11i v3 only)
- Up to 16 node Metrocluster for RAC cluster, 8 nodes per site

- **Serviceguard Storage Management Suite**

- A.01.01 bundle 7, includes:
 - Serviceguard version A.11.18, SGeRAC A.11.18
 - CVM/CFS version 4.1 on HP-UX 11i v2
- A.02.00 bundle 7, includes:
 - Serviceguard A.11.18, SGeRAC A.11.18
 - CVM/CFS version 5.0 support on HP-UX 11i v2 and HP-UX 11i v3
- Up to 16 node Metrocluster for RAC cluster, 8 nodes per site



Metrocluster for RAC: Software requirements for Mar 2009 release

For HP-UX 11i v2 or HP-UX 11i v3

- **Metrocluster**

- Metrocluster with Continuous Access XP version **A.09.00**
- Metrocluster with Continuous Access EVA version **A.04.00**
- Metrocluster with EMC SRDF version **A.08.00**

- **SLVM**

- SLVM 1.0 (HP-UX 11i v2 and 11i v3)
- SLVM 2.0 (HP-UX 11i v3 only)
- Up to 16 node Metrocluster for RAC cluster, 8 nodes per site

- **Serviceguard Storage Management Suite**

- A.01.01 bundle 7, includes:
 - Serviceguard version A.11.18, SGeRAC A.11.18
 - CVM/CFS version 4.1 on HP-UX 11i v2
- A.02.00 bundle 7, includes:
 - Serviceguard A.11.18, SGeRAC A.11.18
 - CVM/CFS version 5.0 support on HP-UX 11i v2 and HP-UX 11i v3
- **A.02.01 bundle 7, includes:**
 - **Serviceguard A.11.19, SGeRAC A.11.19**
 - **CVM/CFS version 5.0 support on HP-UX 11i v2 and HP-UX 11i v3**
- Up to 16 node Metrocluster for RAC cluster, 8 nodes per site



Metrocluster for RAC: Hardware requirements for all releases

For HP-UX 11i v2 or HP-UX 11i v3

- HP Integrity Servers and HP 9000 Servers
- HP StorageWorks XP and EVA disk arrays with Continuous Access License and firmware versions supported by Metrocluster
- EMC Symmetrix disk arrays with SRDF License and firmware versions supported by Metrocluster
- Configurations supported by Metrocluster (e.g. replication distance supported)
- Cross-subnet Serviceguard cluster configuration (optional)



Metrocluster for RAC: Documentation

All available on <http://docs.hp.com/en/ha.html>

- **Disaster Tolerant Clusters**

- Designing Disaster Tolerant HA Clusters Using Metrocluster and Continentalclusters Manual
- Understanding and Designing Serviceguard Disaster Tolerant Architectures Manual
- Metrocluster with Continuous Access XP Version A.09.00 Release Notes
- Metrocluster with Continuous Access EVA Version A.04.00 Release Notes
- Metrocluster with EMC SRDF Version A.08.00 Release Notes
- Disaster Tolerant Clusters Products Compatibility and Feature Matrix (Metrocluster CA XP)
- Disaster Tolerant Clusters Products Compatibility and Feature Matrix (Metrocluster CA EVA)
- Disaster Tolerant Clusters Products Compatibility and Feature Matrix (Metrocluster/EMC SRDF)

- **Serviceguard and SGeRAC**

- Managing Serviceguard Manual
- Serviceguard A.11.19 Release Notes
- Using Serviceguard Extension for RAC Manual
- Serviceguard Extension for RAC A.11.19 Release Notes
- Serviceguard/SGeRAC/SMS/Serviceguard Manager Plug-in Compatibility and Feature Matrix
- Sample Configurations with SGeRAC and Oracle RAC 10gR2 Whitepaper
- Sample Configuration with SGeRAC and Oracle RAC 11gR1 Whitepaper
- SGeRAC and Oracle RAC on HP-UX 11i Best Practices Whitepaper
- Use of Serviceguard Extension For RAC Toolkit with Oracle RAC 10g Release 2 or later Whitepaper



References



For more information

- HP High Availability & Disaster Tolerance Solutions
 - <http://www.hp.com/go/dt>
 - <http://www.hp.com/go/serviceguard>
- HP Serviceguard Solutions
 - <http://h71028.www7.hp.com/ERC/downloads/5982-5163EN.pdf>
- HP Metrocluster
 - <http://h71028.www7.hp.com/ERC/downloads/4AA2-1532ENW.pdf>
- HP Extended Cluster for RAC
 - <http://h71028.www7.hp.com/ERC/downloads/4AA2-2306EEW.pdf>
- Technical documentation
 - <http://www.docs.hp.com/en/ha.html>



HP High Availability and Disaster Tolerance Solutions demos

- Watch Serviceguard Manageability demos at:
 - <http://www.hp.com/go/HADemos>
Click on: Serviceguard Manageability
- Watch other HA/DT demos at:
 - <http://www.hp.com/go/HADemos>
Click on: Continuity & Availability
 - Apache Web Server Farm Utilising Serviceguard CFS
 - Avoiding Outages Due to Hardware Failure
 - Handling Surges in Business Transactions
 - High-Availability Solutions
 - Lowering the Cost of High Availability
 - Serviceguard CFS for RAC integrated with VSE
 - The Virtual Server Environment Made Real (incl. Metrocluster)
 - And more...



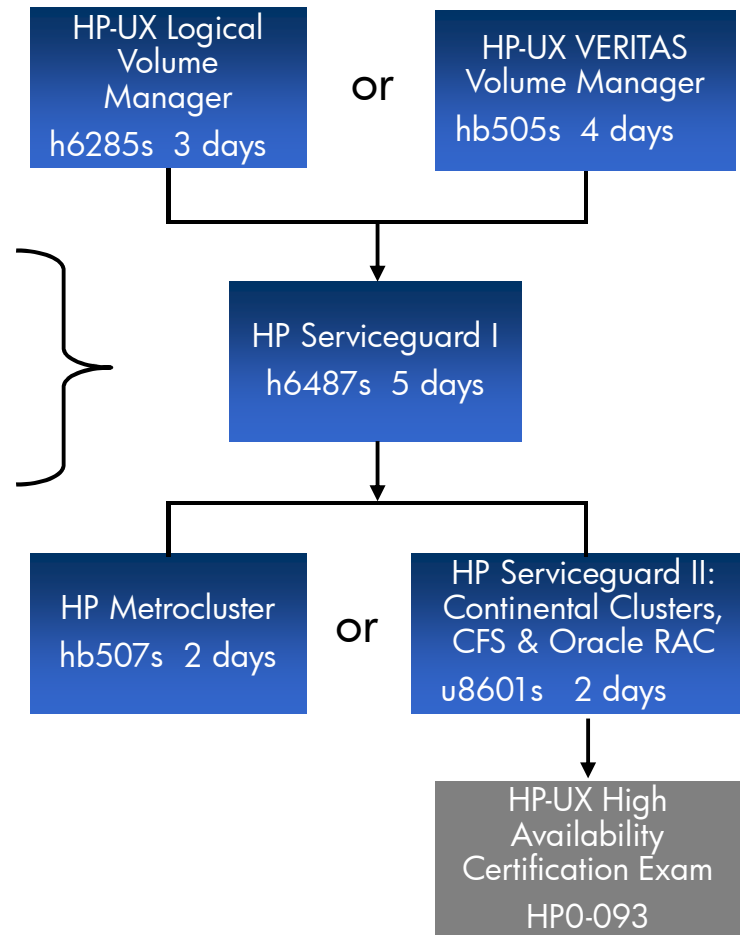
Product features and compatibility matrices

- Serviceguard, SGeRAC, Serviceguard Storage Management Suite, Quorum Server and ECMT:
 - <http://docs.hp.com/en/5971/SG-SGeRAC-SMS-EMSSupport.pdf>
 - <http://docs.hp.com/en/6202/QSSupportMatrix.pdf>
 - <http://docs.hp.com/en/4446/ECMToolkitCompatibilityMatrix.pdf>
- Disaster Tolerant Clusters products:
 - <http://docs.hp.com/en/6096/DT-matrix-cc.pdf>
 - <http://docs.hp.com/en/6120/DT-matrix-mcxp.pdf>
 - <http://docs.hp.com/en/6192/DT-matrix-mceva.pdf>
 - <http://docs.hp.com/en/6095/DT-matrix-mcsrdf.pdf>

HP High Availability training curriculum

H6487S
Sept'08
Update!

Recently updated for Serviceguard 11.18, including content on legacy and modular packages, cross-subnet support, APA and lock LUN, DSAU and labs using SG MGR SMH plug-in covering many of the latest enhancements. Students can use VxVM or LVM for labs involving storage.



- For more information, visit www.hp.com/learn/unix
- Note that POSIX Shell Programming (h4322s) is recommended as a precursor to HP Serviceguard (h6487s)
- Check out our Virtual Server Environment curriculum at www.hp.com/learn/vse

Customised training (e.g. HP Serviceguard Storage Management Suite (SMS))

Technology for better business outcomes

