

Multi-Site Failover Clusters with SQL Server 2008 and Denali

Allan Hirt

Cluster MVP

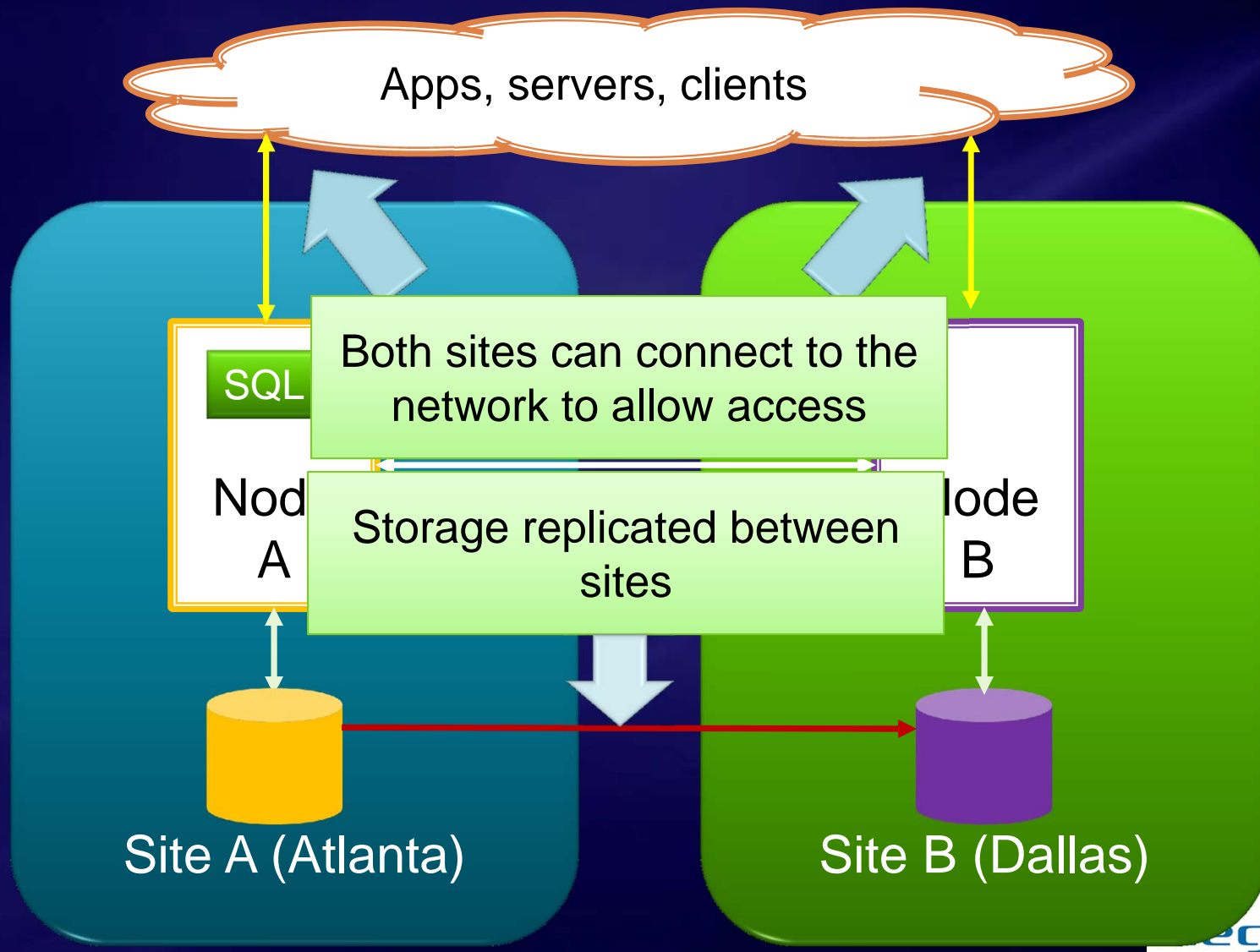
E-mail: allan@sqlha.com

Twitter: @SQLHA

Website and Blog: <http://www.sqlha.com>



The 10,000 Foot View



It's Still A Failover Cluster

- Standard cluster rules apply (e.g. must pass validation)
 - Since SQL Server requires shared storage, *must* pass the storage tests (KB943984)
- SQL Server 2005, 2008, and 2008R2 *only* on Windows Server 2008/R2
- Still one default instance per cluster (up to 25 supported per cluster)

	Windows Server 2008 RTM/SP2	Windows Server 2008 R2
SQL Server 2005	SP2+	SP3+
SQL Server 2008	RTM	SP1+
SQL Server 2008 R2	RTM	RTM

Quorum

- Need the right model
- Shared disk (No Majority/Node Majority with Disk Witness) not recommended
 - If possible, use Node Majority with File Share Witness
 - Place the file share on a server that is NOT any of the nodes and accessible by all nodes
- Make sure that you maintain quorum despite the distance
 - Put proper monitoring in place

Storage Configuration 1

- Mirror at a hardware level
 - 2x the storage
 - Vendor provided solution; Windows and SQL Server have nothing to do with it
 - Not transactional based; usually block level
- Vendor solution may or may not have some SQL Server awareness
 - Check before configuring as it may impact the order of installation
 - Follow vendor guidelines for configuration and installation
 - Test the base configuration before installing SQL Server

Storage Configuration 2

- Vendor solution may not switch sites automatically
 - Adds some manual steps you would need to do in a site-based failover
 - May prolong SQL Server coming online due to disk dependencies if storage not yet owned by other site
- Can configure tempdb on a local drive with Denali
 - Do not need to replicate that storage

Asynchronous or Synchronous?

- Key factors:

- SLAs/RTOs/RPOs
- Infrastructure capability
 - Wide enough pipe to handle a synchronous load?
 - Same configuration on both sides if possible (i.e. do not get slower/cheaper disks for the “standby” node(s); can add to latency, etc.)
- Some vendors will only support synchronous up to a certain distance
- Cost
 - Expensive to lay fibre over a distance and ensure a robust backend storage solution
 - May need to revise things like RTOs and RPOs accordingly

Networking – SQL Server 2008/R2

- Does not supports the multi-site enhancements of Windows Server 2008/R2
- Requires a virtual local area network (VLAN)
 - Underlying servers/devices on different networks
 - Allows communication as if they were on the same network
- Example configuration (2 Node)

Network Purpose	IP Address
Public Network (Node A)	181.182. 183 .100
Public Network (Node B)	181.182. 184 .100
Cluster-only Network (Node A)	10.10. 10 .100
Cluster-only Network (Node B)	10.10. 20 .100
Windows Failover Cluster	181.182. 183 .200 181.182. 184 .200
SQL Server instance (<i>on VLAN</i>)	222.223.224.200

Networking – Denali

- Fully supports the multi-site enhancement of Windows Server 2008/R2
- Does not require a VLAN
- Example configuration (2 Node)

Network Purpose	IP Address
Public Network (Node A)	181.182. 183 .100
Public Network (Node B)	181.182. 184 .100
Cluster-only Network (Node A)	10.10. 10 .100
Cluster-only Network (Node B)	10.10. 20 .100
Windows Failover Cluster	181.182. 183 .200
	181.182. 184 .200
SQL Server instance	181.182. 183 .201
	181.182. 184 .201

General Networking Tips

- All nodes must be part of the same or trusted domains
 - Especially since SQL Server requires a domain account for its service accounts
- Routing must work properly
 - Only way the different subnets can communicate with each other
- Always test (ping) *all* names and IP addresses (public/private/etc) from:
 - Inside the cluster
 - Verifies node-to-node communication
 - Outside the cluster
 - From all subnets (workstation, other server, etc.)
 - Should be able to see IP addresses in that subnet and the SQL Server VLAN (SQL 2K8 R2 and earlier)

Client Connectivity with No VLAN

- Being able to resolve names matters after a failover
 - How quickly a client/user/application can reconnect
- May want to alter a property on the network name
 - HostRecordTTL – value is in seconds
 - To see:
Get-ClusterResource “Resource Name” | Get-ClusterParameter HostRecordTTL
 - To change:
Get-ClusterResource “Resource Name” | Set-ClusterParameter HostRecordTTL *Value*

What About the Cluster Group?

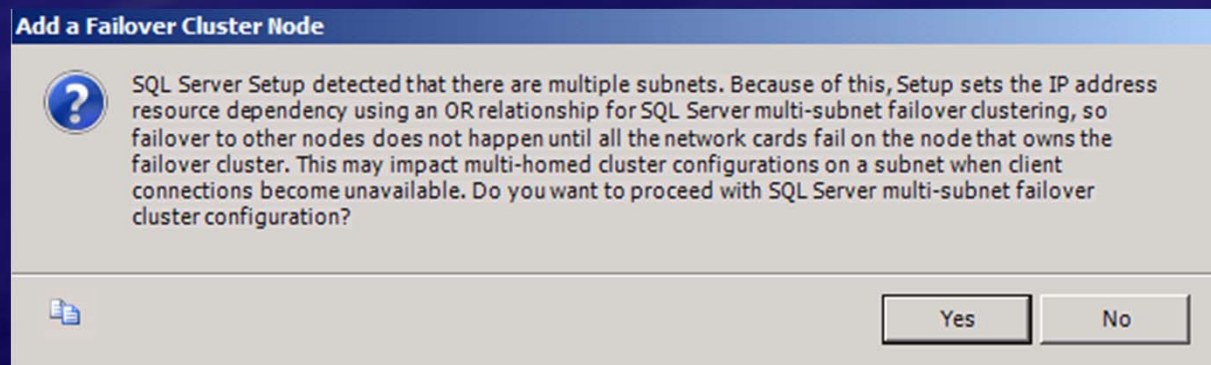
- If you didn't know, it's still there
 - Not displayed in Failover Cluster Manager
 - Needs to be up regardless of where SQL is running but ...
- Example of why it matters
 - Site A (main), Site B (failover), file share witness in Site C
 - Everything up: cluster group owned by B, SQL in A
 - Network problem occurs between A and B
 - Since B and C can still connect and B "owns" the main cluster group, B tries to gain control of SQL Server in arbitration process; SQL outage occurs in failover
 - If A owned the cluster group, would have had enough votes and no interruption to SQL
- Two things:
 - Recommendation: have it where you run SQL (best not to "load balance" across sites); you must script or do it yourself
 - Always monitor your network connectivity between sites

Installing SQL Server 2008/R2

- The “problem” – the cluster validation checks in SQL Server Setup
 - Will most likely fail on storage during Add Node
- The solution
 - Skip the cluster validation tests using the command line or ini parameter SkipRules
SkipRules=Cluster_VerifyforErrors
 - Documented in KB953748
- **WARNING: DO NOT USE THIS SWITCH FOR A NORMAL CLUSTERED INSTALLATION TO GET AROUND REAL ISSUES**

Installing a Denali Instance

- Multi-site done via Setup
- Two components: one in the first node, other in add node
 - First node: only configure the IP address for the subnet of that node's public network
 - Add node: only configure the IP address for the different subnet of that node (extra step new to Denali)



AlwaysOn Availability Groups and Multi-Site

- All nodes regardless of location part of the *same* Windows failover cluster
- NO storage mirroring requirement
 - AGs do not use/require shared storage
 - Exception: if an instance participating is clustered
- Watch quorum model and make sure to maintain it
- AG virtual network name may have multiple OR dependencies depending on the number of subnets/nodes

Summary

- The multi-site/geographically dispersed cluster experience in Denali is greatly improved
- Can take advantage of how Windows works with no VLAN requirement
- Watch where the cluster group and witness are located
- Similar planning techniques for a standard failover cluster as well as availability groups

Conditions and Terms of Use

This training package content is proprietary and confidential, and is intended only for users described in the training materials. Content and information designated for limited distribution is provided to you under a Non-Disclosure Agreement and cannot be distributed. Copying or disclosing all or any portion of the content and/or information included in such packages is strictly prohibited.

The contents of this package are for informational and training purposes only and are provided "as is" without warranty of any kind, whether express or implied, including but not limited to the implied warranties of merchantability, fitness for a particular purpose, and non-infringement.

Training package content, including URLs and other Internet Web site references, is subject to change without notice. Because Megahirtz LLC must respond to changing market conditions, the content should not be interpreted to be a commitment on the part of Megahirtz LLC, and Megahirtz LLC cannot guarantee the accuracy of any information presented after the date of publication. Unless otherwise noted, the companies, organizations, products, domain names, e-mail addresses, logos, people, places, and events depicted herein are fictitious, and no association with any real company, organization, product, domain name, e-mail address, logo, person, place, or event is intended or should be inferred.

Copyright and Trademarks

© 2011 Megahirtz LLC All rights reserved.

Megahirtz LLC may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in written license agreement from Megahirtz LLC, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Megahirtz LLC.

All products mentioned are either registered trademarks or trademarks for their respective manufacturers and owners in the United States and/or other countries.

