Windows Failover Clustering Basics for the DBA

Allan Hirt
Cluster MVP
e-mail: allan@sqlha.com
Website: http://www.sqlha.com
Blog: http://www.sqlha.com/blog
Twitter: SQLHA
Two kinds:

- **Failover Clustering**
  - SQL Server’s clustering is built on top of this
  - Purely an availability solution; scale up (scale out via application specific methods)

- **Network Load Balancing (NLB)**
  - Limited scenarios for SQL Server
  - Used for scalability as well as availability (example: web servers)
Terminology Changes Over the Years

<table>
<thead>
<tr>
<th>Year</th>
<th>Terminology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta</td>
<td>Wolfpack</td>
</tr>
<tr>
<td>1997</td>
<td>Microsoft Cluster Service (MSCS)</td>
</tr>
<tr>
<td>2000</td>
<td>Server Clustering</td>
</tr>
<tr>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
</tr>
<tr>
<td>R2</td>
<td>Failover Clustering (WSFC)</td>
</tr>
</tbody>
</table>
W2K8, Clustering, and SQL Server

- Use Windows Server 2008 R2 at this point
  - 64-bit only
  - Will need Windows Server 2008 RTM-based install if still need a 32-bit
  - 32-bit clustered SQL instances not supported on a 64-bit OS
- Clustering still an Enterprise or Datacenter feature of Windows
- 2005/2008/2008 R2 supported in a side-by-side configuration
- Core variant not supported for SQL Server 2008 R2 or earlier
  (slated for support with Denali; watch the CTPs)
- No rolling upgrade from W2K3 → W2K8 (RTM or R2)

<table>
<thead>
<tr>
<th></th>
<th>Windows Server 2008 RTM/SP2</th>
<th>Windows Server 2008 R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQL Server 2005</td>
<td>SP2+</td>
<td>SP3+</td>
</tr>
<tr>
<td>SQL Server 2008</td>
<td>RTM</td>
<td>SP1+</td>
</tr>
<tr>
<td>SQL Server 2008 R2</td>
<td>RTM</td>
<td>RTM</td>
</tr>
</tbody>
</table>
The Clustering Stack (High Level)

- Hardware (network, storage, etc.)
- Operating System (Windows)
- Failover Clustering
- Clustered Applications

This is where most of the problems will be.
Failover Cluster Components 1

Node = server

Networks
- Public – used by external connections (and the cluster)
- Private/Heartbeat – cluster communications only

Storage
- “Shared nothing”
- Cluster Shared Volume (Hyper-V only)
Quorum

- Determines at what point the cluster will stop running based on the number of available voters
- Helps avoid “split brain”

Quorum types in W2K8

- No majority (same as the old “disk only”)
  - All nodes can fail except 1 if the witness disk is online
  - Not recommended; witness disk is single point of failure
Failover Cluster Components 3

Quorum types cont’d

- **Node Majority** (same as W2K3 Majority Node Set)
  - Failure calculation: $\frac{1}{2}$ nodes (rounding up) – 1

- **Node and Disk Majority**
  - Combination of No Majority + Node Majority; removes single point of failure
  - Failure calculations:
    - Witness disk available: $\frac{1}{2}$ nodes (rounding up)
    - Witness disk unavailable: $\frac{1}{2}$ nodes (rounding up) – 1

- **Node and File Share Majority**
  - Same as Node and Disk Majority, but with file share-based witness instead of a shared disk
Resource
- Physical or logical component providing some use in the cluster
- Owned by one node at a time

Resource Groups = Services/Applications in W2K8
- Collection of resources
- Resources cannot be shared across groups

Dependencies
- A resource can be a dependency of another
- If a resource is a dependency, the child resource cannot come online until the parent is online
Two main health checks

- LooksAlive – cluster only (“Are you there?”)
- IsAlive – Application specific
  - For SQL Server, it is a SELECT @@version query
  - Not configurable

Cluster Service service on each node

- Sub-components: event processor, database manager, node manager, global update manager, communication manager, resource manager
Clustered Applications

Old terminology: virtual server
Comprised of:
- 1+ cluster disks (NOTHING LOCAL)
- Network Name
- 1 or more IP addresses
- Cluster resources which reside in a resource group
  - Resources can have dependencies
What’s In A Name?

You need lots of ‘em – all unique

- One for each cluster node (unique in the domain)
- One for the Windows failover cluster (unique in the domain)
- One for each clustered application

This is what applications/users connect to
How Does It All Work?

- Nodes monitor health of other nodes.
- If that node fails, health monitoring will cause a failover of the resource.
- Another node starts the application and reads the last saved information from the storage.
- Clients experience a slight interruption in service.
What Happens in SQL Server During a Failover?

- Resources start in dependency order
- Stop and start, so SQL Server goes through normal recovery
  - SQL Server considers itself up when the system databases are up
  - Reality: not up until the user databases are online
  - Crucial not to have a lot hanging around in the t-logs
- No name or IP address changes
Due to the recovery process in a failover, SQL Server has no data loss and is transactionally consistent to the point of failure.

Just to reiterate: connect to the clustered SQL Server instance name or IP.
- Not the node name/IP or the Windows failover cluster name/IP.

Not an automatic reconnect.
- Retry logic
- Code application to be cluster-aware (Platforms SDK)
Networking 1

- Need a unique IP address for:
  - Each node
  - Windows failover cluster
  - Each SQL Server instance
- NIC teaming fully supported under W2K8 and up (not under W2K3) – KB254101
  - Test before going into production
- SQL Server 2008 and earlier requires a VLAN for a geographically dispersed cluster
  - Does not support OR
  - Denali fixes this
Networking 2

- Minimum of 2 networks + corresponding NICs
  - Externally facing/public
  - Internal*
  - Storage (iSCSI) requires its own dedicated network
  - Hyper-V should get its own (if using it)

- Rename the networks with friendly names

- Need redundancy at the physical layer, too
  - Example: Don’t plug all NICs into the same switches
Windows

- Domain account needed for cluster creation and administration – THAT’S IT
  - Requires Create Computer Objects right on Computers OU
  - If cannot give CCO, create the Cluster Name Object (CNO) and Virtual Computer Object (VCO) manually
- Domain account NOT used to run the cluster service – runs under a special context

SQL Server

- Still requires service accounts
- Use a Service SID; do not have to use domain groups
- No local Administrator requirements unless using something like xp_cmdshell (off by default)

Never use a domain admin account or assign domain admin privileges – THEY ARE NOT NEEDED
Security 2

- Still prefers dynamic DNS
- Firewall
  - By default, it’s on – and there can be up to 3
  - Use common sense and follow your corporate security guidelines
  - If on, can cause things like:
    - Unwanted RPC errors
    - Inability for Management Studio to see SQL Server instances hosted on a particular cluster node
  - If left on, make exceptions (both inbound and outbound rules) for things like the SQL Server port number
- New cluster security for users
- IPv6 support in W2K8/R2
Anti-virus and SQL Server Failover Clustering

- Not recommended if not needed
- If required, set exclusions:
  - .mdf (data), .ldf (log), .ndf (additional data)
  - .bak (default backup extension), .trn (default t-log backup extension)
  - All directories with Analysis Services data, log, temporary files, backups
  - Entire quorum/witness disk
  - \MSDTC directory for MSDTC disks (if used)
  - \Cluster subdirectory under %windir%

SQL Server & anti-virus KB 309422
Cluster & anti-virus KB 250355
Drive types supported: only SAS, iSCSI, Fibre; no old parallel SCSI

- Storage must be SCSI-3 SPC-3 command compliant
- Storage must support persistent reservations

Still need at least one drive letter per clustered instance of SQL Server 2008
- Can use mount points for everything else

Cluster Shared Volumes (CSV) *not* supported for SQL Server use; Hyper-V only

Can select multiple drives during SQL Server install
- Rename drives to make them easier to discern
Cluster Validation 1

- W2K8 no longer relies on the HCL/Server Catalog
- Hardware must be logoed for W2K8/R2
- Must re-run with hardware or major config change
- Run via Failover Cluster Manager or PowerShell (PS is W2K8 R2 only)
- TIP: great way to document your servers and check for changes over time
Cluster Validation 2

- Different types of tests run (storage, networking, etc.)
  - Report location: `%windir%\Cluster\Reports`
  - Pass, Warning, Fail
  - Warning not fatal, but must investigate
  - Fix any problems and re-run
  - TIP: Re-run failed tests first, then re-run full suite

- Additional validation tests if run after configured in Windows Server 2008 R2

- SQL Server relies on a successful result
  - Prevent a false positive
OS supports multiple DTCs – no longer need to share one per Windows failover cluster.

Installation options:

- Old way: DTC in its own group (with its own IP, disk)
- One DTC per SQL Server instance
  - Old way + bind to instance
  - Put in same group with SQL Server
- Do not cluster DTC at all; clustered SQL Server will “negotiate down” to use a local DTC if nothing clustered.
MSDTC, SQL Server 2008/R2, and Windows Server 2008

- OS supports multiple DTCs – no longer need to share one per Windows failover cluster
- Installation options
  - Old way: DTC in its own group (with its own IP, disk)
  - One DTC per SQL Server instance
    - Old way + bind to instance

```
PS C:\users\cluadmin> msdtc -tmMappingview *
PS C:\users\cluadmin>
```
OS supports multiple DTCs – no longer need to share one per Windows failover cluster

Installation options

Old way: DTC in its own group (with its own IP, disk)

One DTC per SQL Server instance

Old way + bind to instance
OS supports multiple DTCs – no longer need to share one per Windows failover cluster.

Installation options:

Old way:
- DTC in its own group (with its own IP, disk)
- One DTC per SQL Server instance
- Do not cluster DTC at all; clustered SQL Server will "negotiate down" to use a local DTC if nothing clustered.

New way:
- Use TMMapping view
- Key: HKLM\Cluster\MSDTC\TMMapping\Service\EquinoxDTC
- Num of Values = 3
  - ValueName: ClusterResourceId
    - Value Type: 1
    - Value: SQL Server
  - ValueName: ApplicationType
    - Value Type: 4
    - Value: 1
  - ValueName: Name
    - Value Type: 1
    - Value: MSSQLSERVER
- Operation succeeded
OS supports multiple DTCs – no longer need to share one per Windows failover cluster

Installation options

Old way: DTC in its own group (with its own IP, disk)
One DTC per SQL Server instance
  Old way + bind to instance
  Put in same group with SQL Server
OS supports multiple DTCs - no longer need to share one per Windows failover cluster.

Installation options:

Old way: DTC in its own group (with its own IP, disk)

One DTC per SQL Server instance

Put in same group with SQL Server

Do not cluster DTC at all; clustered SQL Server will “negotiate down” to use a local DTC if nothing clustered.
VMs as a “Guest Cluster”

- Supported as of May, 2009 (KB 956983)
- W2K8 OS for guests only
- SQL 2K5 or SQL 2K8
- Will need some sort of shared storage solution (most likely iSCSI)

Recommendations:
- Map virtual NICs to different networks/NICs at the hypervisor level
- Place the VMs on separate hypervisors to prevent a single point of failure

Bottom line: same rules as a physical cluster apply
- Yes, you can technically have one note physical, one node virtual if you pass Validation
Administration

Tools

- Failover Cluster Management (RTM)/Manager (R2)
  - No more Cluster Administrator
  - MMC snap-in
- cluster.exe
  - Going away after W2K8 R2
- PowerShell cmdlets (W2K8 R2 only)

Debug logging

- Based on event tracing – old cluster.log gone
- Must turn on
- Logs stored in %windir%\System32\winevt\logs
- Up to 3 files (incremented on reboot)
- Get-ClusterLog to dump to file similar to old cluster.log
- Tracerpt.exe to dump trace session
Resources 1

- Pro SQL Server 2008 Failover Clustering (Apress, 2009)
- My blog – updated regularly with new stuff!
- SQL Server 2008 failover clustering whitepaper
- Windows Clustering Team Blog
- Microsoft Enterprise Platforms Support: Windows Server Core Team blog
Resources 2

KB Articles:

- **956893** “Support policy for Microsoft SQL Server products that are running in a hardware virtualization environment”
- **897615** Support policy for Microsoft software running in non-Microsoft hardware virtualization software
- **943984** The Microsoft Support Policy for Windows Server 2008 Failover Clusters
- **327518** The Microsoft SQL Server support policy for Microsoft Clustering
- **958734** SQL Server 2008 failover clustering rolling patch and service pack process