SQLskills Immersion Event

IECAG: Clustering and Availability Groups

Module 3: AlwaysOn Availability Groups

Jonathan Kehayias
Jonathan@SQLskills.com

Overview

- Availability group overview
- Readable secondaries
- Application connection support
- Partial database containment
- Migrating from database mirroring
- Quorum model and node votes
- Asymmetric storage considerations
- Security
- Performance

© SQLskills, All rights reserved.
https://www.SQLskills.com

HA Features with Availability Groups

High-Availability Feature	Availability Groups
Standby type	Hot, Hot + Warm, or just Warm
Failure detection	Automatic
Automatic failover	Yes
Protects against failed server process	Yes (but only the replicated databases are protected)
Protects against failed disk	Yes (but only the replicated databases are protected)
Transactionally consistent	Yes
Transactionally current	Yes (if in synchronous mode)
Performance impact	It depends (varies from no performance impact to possibly severe impact)
Time to failover	Depends on restart recovery
Locations	Dispersed (performance could be a limiting factor)
Hardware Requirements	Disk space, disk speed, network capacity
Multiple standby/secondary nodes	8
Standby available for reporting, etc.	Yes
Cost	Medium
Complexity	Medium-High

SQLskills immerse yourself in sql server

3

© SQLskills, All rights reserved.
https://www.SQLskills.com

Availability Group Types

- **Basic Availability Groups (2016 and higher Standard Edition)**
  - Supports a single database only
  - Only one secondary replica supported
  - No readable access to secondary database replica
- **AlwaysOn Availability Groups (2012 and higher)**
  - High availability, disaster recovery, read-scale load balancing
  - Requires external cluster manager
- **Read-Scale Availability Groups (2017 and higher)**
  - Secondary replicas for read-only workloads without high availability
  - Replication of data between nodes only
  - No cluster manager
- **Distributed Availability Groups (2016 and higher)**
  - An Availability Group of Availability Groups (more later)

SQLskills immerse yourself in sql server

4

© SQLskills, All rights reserved.
https://www.SQLskills.com

Basic Availability Groups

- **SQL Server 2016+ Standard Edition only**
- **Replaces deprecated Database Mirroring with similar functionality**
- **Supports failover of a single database maintaining a single replica**
  - May be either synchronous or asynchronous commit
- **No read access to secondary replica**
- **No backups on secondary replica**
- **Cannot add or remove a replica to an existing Basic Availability Group**
  - Must specify both replicas during creation – changes required drop and recreate
- **Cannot be upgraded to advanced availability groups**
  - Must drop the Basic Availability Group before changing to Enterprise Edition to join an existing Availability Group

SQLskills immerse yourself in sql server

5

© SQLskills, All rights reserved.
https://www.SQLskills.com

AlwaysOn Availability Groups Overview (1)

- **AlwaysOn was the over-arching marketing term for 2012+ HA/DR**
  - AlwaysOn Availability Groups <> AlwaysOn Failover Cluster Instances
- **Availability group contains one or more user databases that fail over together**
- **Availability database is a database that belongs to an availability group**
  - Primary database is the read-write copy (limit 1)
  - Secondary database is the read-only or non-readable copy
    - Up to four on SQL Server 2012 and eight on SQL Server 2014+
    - Distributed AG's allow further scale out of readable secondaries
- **Availability replica is the SQL Server instance (standalone or SQL FCI) that hosts the availability databases**
  - Primary replica hosts the read-write databases
  - Secondary replica(s) host the secondary copies, non-writeable copies

SQLskills immerse yourself in sql server

6

© SQLskills, All rights reserved.
https://www.SQLskills.com

AlwaysOn Availability Groups Overview (2)

- **Availability group listeners support automatic client redirection to the primary replica or redirection to available readable secondaries**
  - Removes the need to designate a failover partner in the client connection string
- **Secondary replicas can also be configured to permit read-only workloads, with low latency updates from the primary**
  - Offload full (copy-only) database and transaction log backup operations
- **Leverages external cluster manager for health detection, failover coordination, and to provide the availability group listener functionality for connectivity across any node**
  - WSFC on Windows
  - Pacemaker on Linux
- **You can use standalone instances, SQL FCIs or a mix**
- **Using SQL FCIs means *no automatic failover* of the AG**

SQLskills immerse yourself in sql server

7

© SQLskills, All rights reserved.
https://www.SQLskills.com

Read-Scale Availability Groups

- **SQL Server 2017+ Enterprise Edition**
- **Secondary replicas for read-only workloads without high availability or disaster recovery**
- **Replication of data between nodes only**
- **No cluster manager required on Windows or Linux**
  - Availability Group CLUSTER_TYPE = NONE
- **Supports read only routing through a listener**
  - Not created in cluster – primary replica only and must specify local IP
  - Not required in DNS – simply used to support the rerouting of connections

SQLskills immerse yourself in sql server

8

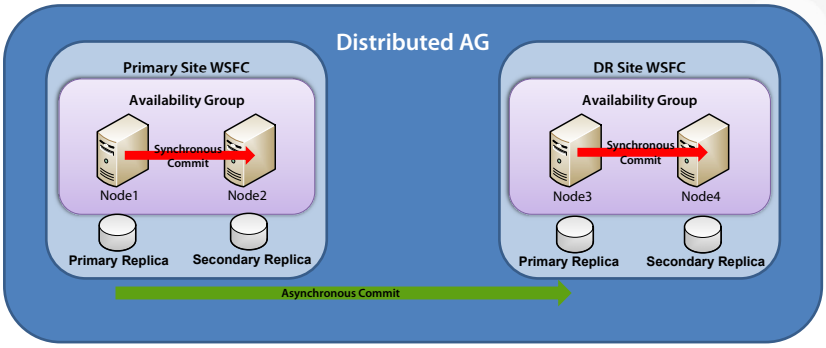
© SQLskills, All rights reserved.
https://www.SQLskills.com

Distributed Availability Groups (1)

- New feature in SQL Server 2016+ that can be considered an "availability group of availability groups"
- Associates two availability groups on different WSFCs together for disaster recovery between geo-dispersed sites
  - Allows data to be continuously replicated without network problems affecting the primary site in any way
- Only the primary replica for the primary availability group is writeable
  - Secondary availability group becomes read-only once joined to the Distributed Availability Group
- Only supports manual failover to the secondary availability group

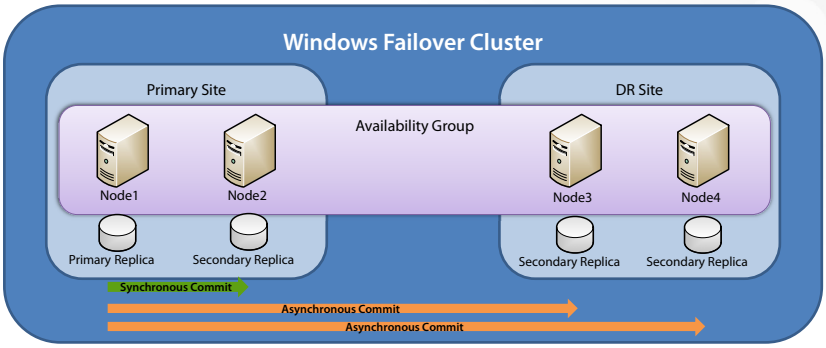
© SQLskills, All rights reserved.
https://www.SQLskills.com

Distributed Availability Group



© SQLskills, All rights reserved.
https://www.SQLskills.com

Multi-Site Availability Group



© SQLskills, All rights reserved.
https://www.SQLskills.com

Distributed Availability Group Benefits

- **Isolation of each WSFC**
  - Each WSFC maintains its own quorum model and voter configuration
  - Secondary WSFC health has no impact on Primary WSFC status
- **Data is only sent once over the network to the Secondary WSFC, regardless of number of replicas, then replicated within the WSFC to each replica**
  - A single WSFC Availability Group sends the data to each replica separately
- **Supports rolling upgrades of the OS between datacenters**

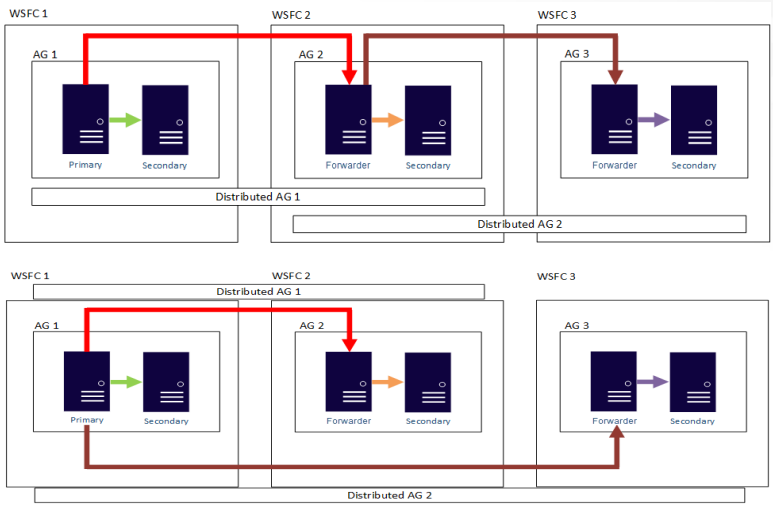
© SQLskills, All rights reserved.
https://www.SQLskills.com

Scale-out Readable Replica Distributed AG

- **Allows multiple sites to have near-real-time readonly access to data**
- **Allows the primary replicas to participate in two different distributed availability groups**
  - Primary replica of the primary availability group can create two distributed availability groups
  - Each primary in a secondary availability group can act as a forwarder and create a second distributed availability group for read-scale-out
- **Read-only routing only functions on the primary availability group in any distributed availability group configuration**
  - All secondary availability group listeners will route connections to the primary replica in the secondary availability group only

© SQLskills, All rights reserved.
https://www.SQLskills.com

Scale-out Readable Replica Distributed AG (2)



© SQLskills, All rights reserved.
https://www.SQLskills.com

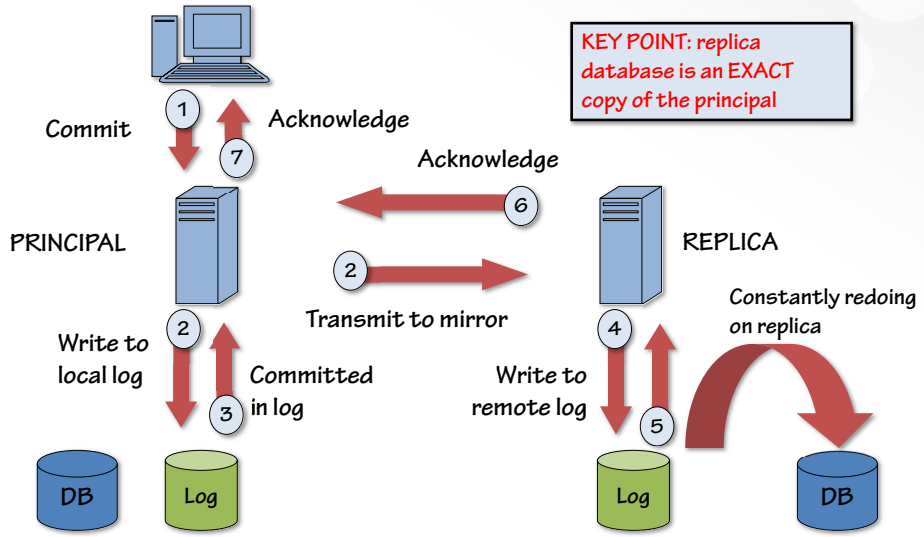
Valid Usage Scenarios

- **User-database high availability and disaster recovery**
- **Offloading read-only activity**
  - But remember – there is still only *one* read-write replica
- **Offloading backups**
  - Transaction log backups
  - Full backups (copy only)

© SQLskills, All rights reserved.
https://www.SQLskills.com

Principles of Availability Groups

Synchronous, High-Protection Configuration



© SQLskills, All rights reserved.
https://www.SQLskills.com

Synchronous vs. Asynchronous

- **No data loss?**
  - Up to three synchronous replicas
    - Primary replica
    - Two secondaries
  - Beyond that, asynchronous
- **Some data loss allowed? No tolerance for synchronous overhead?**
  - Up to four asynchronous replicas in 2012, up to eight in 2014+, unlimited in 2016+ with distributed option
- **Failover mode**
  - Automatic or manual
  - If automatic, up to two replicas (including current primary)

SQLskills immerse yourself in sql server

17

© SQLskills, All rights reserved.
https://www.SQLskills.com

Standalone vs. Failover Cluster Instance

- **Availability Groups support multiple options:**
  - All standalone SQL Server instances on each node (non-shared storage)
  - All Failover Cluster Instances (shared storage)
    - High Availability within FCI not the Availability Group
  - Mix of standalone and Failover Cluster Instances
- **But just because you can…**
  - Consider the complexity of availability groups + Failover Cluster Instances
  - Can you achieve your goals using standalone instances?
- **Failover Cluster Instance with asynchronous Availability Groups provides High Availability without the transaction impact of synchronous commit**

SQLskills immerse yourself in sql server

18

© SQLskills, All rights reserved.
https://www.SQLskills.com

REQUIRED_SYNCHRONIZED_SECONDARIES_TO_COMMIT

- **Introduced in SQL Server 2017**
- **Relates to replicas in synchronous commit mode**
- **Sets a minimum number of synchronous secondary replicas required to commit before the primary commits a transaction**
  - Default is 0 which gives the same behavior as SQL Server 2016
  - The maximum value is the number of replicas minus 1
- **Guarantees that the primary replica will not proceed until the minimum number of replicas have committed each transaction**
  - If the minimum number of replicas is not available then commits on the primary will fail
- **Use with caution as this will block writes if correct number of secondaries are unavailable**

SQLskills immerse yourself in sql server

19

© SQLskills, All rights reserved.
https://www.SQLskills.com

Readable Secondaries

- **Allow offloading of read-only workloads to secondary replicas in the environment**
- **Uses snapshot isolation to avoid blocking replay of transactions on the secondary**
- **If additional statistics are required, they are created as temporary objects in tempdb**
  - If additional indexes are required, they must be created in the primary database
- **Allows offloading of log backups to secondary replicas**
  - All log backups form a single chain regardless of what server the backup occurred from
- **FULL backups allowed with COPY_ONLY**
- **Readable secondaries unavailable in SQL Server 2012 during failovers or loss of quorum for the WSFC (fixed in SQL Server 2014+)**

SQLskills immerse yourself in sql server

20

© SQLskills, All rights reserved.
https://www.SQLskills.com

Readable Secondary Options

- **ALLOW_CONNECTIONS = "ALL"**
  - Supports TDS clients, older clients
  - Client is responsible for ensuring replica is readable secondary
- **ALLOW_CONNECTIONS = "READ_ONLY"**
  - Only connections with "ApplicationIntent" set to "ReadOnly" are accepted
  - "Intent" means you should use read-only operations, but read/write connections are necessarily prevented
    - Will fail on the first DML or DDL operation

SQLskills immerse yourself in sql server

21

© SQLskills, All rights reserved.
https://www.SQLskills.com

Read-Only Impact on Recovery Time Objective

- **Page access on secondary replica can increase I/O contention with REDO thread activity, assuming collocation**
- **REDO thread on secondary replica can be blocked by DDL operations**
  - Sch-S lock held by read-only query on a table
  - DDL transaction on the primary replica rebuilds an index for that same table
  - Sch-M lock on secondary replica
- **Monitor via lock_redo_blocked Extended Event**

SQLskills immerse yourself in sql server

22

© SQLskills, All rights reserved.
https://www.SQLskills.com

Read-Only Impact on Data Latency

- Reporting workloads unlikely to affect the acknowledgement of synchronous commit mode
- Latency between primary and secondaries can be a few seconds to minutes depending on workload and network latency
- Data latency can still exist for synchronous mode, as the REDO thread can fall behind
- No data latency for reporting allowed?
  - Run on the primary replica

© SQLskills, All rights reserved.
https://www.SQLskills.com

Read-Only Snapshot Isolation (1)

RCSI / SI	Read-Only	14-byte primary	14-byte secondary	Row version primary tempdb	Row version secondary tempdb
No	No	No	No	No	No
No	Yes	Yes	Yes	No	Yes
Yes	No	Yes	Yes	Yes	No
Yes	Yes	Yes	Yes	Yes	Yes

© SQLskills, All rights reserved.
https://www.SQLskills.com

Read-Only Snapshot Isolation (2)

- Existing rows are not modified with 14-byte overhead until existing row is updated or deleted
- 14-byte overhead included when a new row is added
- Watch out for increased page splits and fragmentation on the primary
  - Mitigate with index fill factors

SQLskills immerse yourself in sql server

25

© SQLskills, All rights reserved.
https://www.SQLskills.com

Cross-database Transactions
SQL Server 2012/2014

- As with database mirroring, cross-database transactions and distributed transactions for availability group databases are 'not supported'
- For cross-database transactions, each database commits independently, even for databases in the same availability group, as a failover can happen after one database commits a transaction but the other database does not
- For distributed transactions, post-failover the new primary replica can't connect to the distributed transaction coordinator of the prior primary replica (unable to get transaction status)

SQLskills immerse yourself in sql server

26

© SQLskills, All rights reserved.
https://www.SQLskills.com

Cross-database Transactions
SQL Server 2016+

- **Cross-database transactions and distributed transactions are not supported prior to Service Pack 2:**
  - When more than one database involved is in the same Availability Group
  - Within the same instance even if the databases are not part of the same Availability Group
  - When the Availability Group was not enabled for distributed transactions
- **Distributed transactions between databases hosted by two different instances of SQL Server are supported prior to Service Pack 2 if:**
  - Running on Windows Server 2016 or on Windows Server 2012 R2 with hotfix KB3090973 applied
  - Availability Group is created WITH DTC_SUPPORT = PER_DB
    - There is no current method to ALTER an existing Availability Group for this option
  - All instances in the Availability Group must be SQL Server 2016 or later
- **Full distributed transaction support on Service Pack 2 and later**

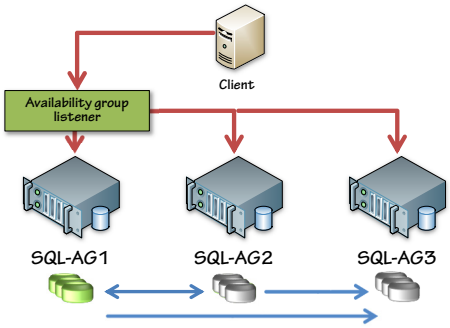
SQLskills immerse yourself in sql server

27

© SQLskills, All rights reserved.
https://www.SQLskills.com

Application Connection Failover

- **Clients connect using the availability group listener virtual network name**
- **Can specify the Application Intent=ReadOnly parameter**
- **During a failover the ReadOnly connections can be automatically rerouted by the Listener to a new node if the current node becomes the primary**



SQLskills immerse yourself in sql server

28

© SQLskills, All rights reserved.
https://www.SQLskills.com

Read-Only Routing

- **Prerequisites**
  - Must connect using an availability group listener
  - One or more replicas need to be read-only secondaries
- **Read-only routing setup**
  - Connect to the primary replica
  - Configure read-only for the secondary role
  - Configure a READ_ONLY_ROUTING_URL for each replica
  - Designate a READ_ONLY_ROUTING_LIST for each replica
- **Leveraging read-only routing**
  - Connect via TCP with the ReadOnly application intent, referencing the listener and a database within the availability group
- **Round-robin "load balancing" in SQL Server 2016**

SQLskills immerse yourself in sql server

29

© SQLskills, All rights reserved.
https://www.SQLskills.com

Read-Only Routing Credentials

- **Don't forget about the required logins on the secondary replicas**
- **Options:**
  - Partially contained databases
  - CREATE LOGIN with SID option
  - KB 918992 sp_help_revlogin to transfer logins (applies to 2012+)

SQLskills immerse yourself in sql server

30

© SQLskills, All rights reserved.
https://www.SQLskills.com

Partial Database Containment

- **Contained Database Authentication stores credentials inside the database removing the need to manually manage logins across environments or availability group nodes**
  - Requires enabling the 'contained database authentication' sp_configure option for the instance
  - Individual databases can be converted to a contained database using ALTER DATABASE <database_name> SET CONTAINMENT = PARTIAL
    - Requires a lock
- **Existing SQL Server Logins can be converted to a contained database user with the *sp_migrate_user_to_contained* system stored procedure**
- **New contained database users can be created using the new DDL feature CREATE USER <username> WITH PASSWORD = 'password'**

SQLskills immerse yourself in sql server

31

© SQLskills, All rights reserved.
https://www.SQLskills.com

Partial Database Containment

- **Temporary tables**
  - Uses the collation of the contained database instead of the tempdb collation
  - Cannot contain named constraints, or refer to user-defined types, XML schema collections, or user-defined functions
- **Instance-level features like replication, change data capture, or change tracking can not be used**
- **Feature usage that crosses the application boundary can be tracked with sys.dm_db_uncontained_entities or the database_uncontained usage Extended Event**
- **Any dynamic SQL used will be flagged as uncontained usage since its actual containment cannot be determined until run time**

SQLskills immerse yourself in sql server

32

© SQLskills, All rights reserved.
https://www.SQLskills.com

Non-contained Objects

- **Similar to database mirroring, you'll need to consider non-contained object synchronization:**
  - Logins (transfer via sp_help_revlogin)
  - SQL Server Agent jobs
    - Custom execution logic to execute as appropriate
  - Linked servers
  - SQLCLR Assembly Permissions
    - UNSAFE/EXTERNAL_ACCESS permissions and key are maintained in master not the user database – create on all nodes
    - TRUSTWORTHY bit set by master database not user database and doesn't replicate
    - DBO for the database must exist with same SID on both servers

SQLskills immerse yourself in sql server

33

© SQLskills, All rights reserved.
https://www.SQLskills.com

Backup Considerations

- **Do you want to offload backups to your secondary replicas?**
- **Even if you do not, how will you handle a failover in your maintenance jobs?**
- **Your maintenance must be "aware"**
  - Ola Hallengren's scripts handle this (http://bit.ly/3Jf3Go)
  - Maintenance plans do too
  - sys.fn_hadr_backup_is_preferred_replica
- **Availability group configurations**
  - Automated backup preference
  - Backup priority

SQLskills immerse yourself in sql server

34

© SQLskills, All rights reserved.
https://www.SQLskills.com

Multi-subnet Connection Support

- **SQL Server 2012+ supported on Windows Server Failover Clusters with nodes that cross-subnets**
- **Microsoft recommends using the client MultiSubnetFailover attribute for both single and multi-subnet topologies that reference the availability group listener name**
- **Opens up TCP sockets for availability group listener IP addresses in parallel**
- **Older client libraries?**
  - Recommendation is to increase the client login timeout
  - Consider adjusting the HostRecordTTL value to a lower value

SQLskills immerse yourself in sql server

35

© SQLskills, All rights reserved.
https://www.SQLskills.com

Automatic Seeding (Initialization)

- **New feature in SQL Server 2016 that allows automatic creation of secondary replicas for every database in an availability group**
  - Eliminates the need to perform backup of primary databases to restore on secondary replicas
- **Requires the data and log paths be identical on all availability replicas**
- **Must create the availability group with SEEDING_MODE=AUTOMATIC**
  - Join secondary replicas to the availability group
  - GRANT CREATE ANY DATABASE to the availability group
  - ALTER AVAILABILITY GROUP ADD REPLICA does not support SEEDING_MODE=AUTOMATIC even though it does not fail
    - Requires modifying the replica after it is added to set the SEEDING_MODE

SQLskills immerse yourself in sql server

36

© SQLskills, All rights reserved.
https://www.SQLskills.com

Log Stream Compression Defaults

- **Synchronous-commit – Not compressed**
  - May be enabled with Trace Flag 9592 – disabled by default since compression adds latency
- **Asynchronous-commit – Compressed**
  - May be disabled with Trace Flag 1462
- **During Automatic Seeding – Not compressed**
  - May be enabled with Trace Flag 9567 – compression can reduce the seeding transfer time but may increase CPU usage on the primary replica

SQLskills immerse yourself in sql server

37

© SQLskills, All rights reserved.
https://www.SQLskills.com

Integration with Azure VMs

- **Azure configuration options included in the Availability Group Wizard in SQL Server 2014+ Management Studio**
  - Allows creation of a asynchronous secondary replica on the Windows Azure Infrastructure as a Service (IaaS) services
  - Supports manual failovers to allow services to run in Azure VM for disaster recovery
- **Supports synchronous secondary replicas within Azure for high availability and automatic failover in SQL Server 2014+**
- **Note: Availability Groups can be deployed on AWS EC2 VMs manually for high availability as well**

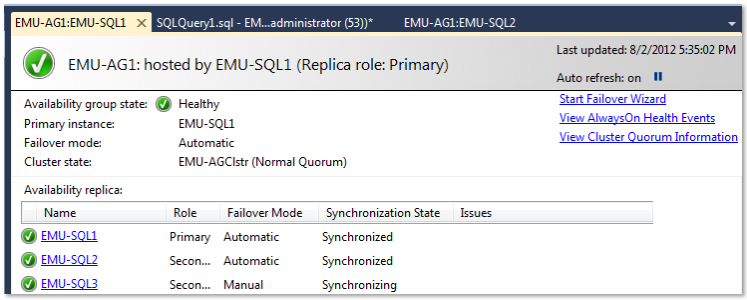
SQLskills immerse yourself in sql server

38

© SQLskills, All rights reserved.
https://www.SQLskills.com

Monitoring

- **Requires a combination of data sources and tools:**
  - AG Dashboard in SQL Server Management Studio
  - Object Explorer state information
  - Performance counters
  - Catalog and Dynamic Management Views
  - Extended Events session that tracks recent availability group critical activity



© SQLskills, All rights reserved.
https://www.SQLskills.com

SQL Server: Database Replica Counters (1)

- **Log Bytes Received/sec (Secondary)**
  - Amount of log records received by the secondary replica for the database in the last second.
- **Log remaining for undo (Secondary)**
  - The amount of log in kilobytes remaining to complete the undo phase
- **Log Send Queue (Secondary)**
  - Amount of log records in the log files of the primary database, in kilobytes, that has not yet been sent to the secondary replica
  - Value is sent to the secondary replica from the primary replica
  - Does not include FILESTREAM files that are sent to a secondary
- **File Bytes Received/sec**
  - Amount of FILESTREAM data received by the secondary replica for the secondary database in the last second

© SQLskills, All rights reserved.
https://www.SQLskills.com

SQL Server: Database Replica Counters (2)

- **Recovery Queue (Secondary)**
  - Amount of log records in the log files of the secondary replica that has not yet been redone.
- **Redo Bytes Remaining (Secondary)**
  - The amount of log in kilobytes remaining to be redone to finish the reverting phase
- **Redone Bytes/sec (Secondary)**
  - Amount of log records redone on the secondary database in the last second
- **Total Log requiring undo (Secondary)**
  - Total kilobytes of log that must be undone
- **Transaction Delay (Primary)**
  - Delay in waiting for unterminated commit acknowledgement, in milliseconds

SQLskills immerse yourself in sql server

41

© SQLskills, All rights reserved.
https://www.SQLskills.com

SQL Server: Availability Replica

- **This counter object applies to primary and secondary replicas**
- **Bytes Received from Replica/sec**
  - Number of bytes received from the availability replica per second
  - Will generate traffic even for quiesced databases
- **Bytes Sent to Replica/sec**
  - Number of bytes sent to the remote availability replica per second
- **Flow Control Time (ms/sec)**
  - Time in milliseconds that log stream messages waited for send flow control, in the last second
- **Flow Control/sec**
  - Number of times flow-control initiated in the last second
  - Flow Control Time (ms/sec) divided by Flow Control/sec is the average time per wait

SQLskills immerse yourself in sql server

42

© SQLskills, All rights reserved.
https://www.SQLskills.com

DMVs (1)

- **sys.dm_hadr_auto_page_repair**
  - Returns a row for every automatic page-repair attempt on any availability database on an availability replica that is hosted for any availability group by the server instance
- **sys.dm_hadr_availability_group_states**
  - Returns a row for each availability group that possesses an availability replica on the local instance of SQL Server and shows health descriptions
- **sys.dm_hadr_availability_replica_states**
  - Returns state information for each local availability replica and a row for each remote availability replica in the same availability group as a local replica

© SQLskills, All rights reserved.
https://www.SQLskills.com

DMVs (2)

- **sys.dm_hadr_cluster**
  - Shows WSFC cluster name, quorum type and quorum state
- **sys.dm_hadr_cluster_members**
  - Shows WSFC nodes, their current state and quorum vote
- **sys.dm_hadr_cluster_networks**
  - Networks associated with the WSFC nodes (IP address, subnet information)
- **sys.dm_hadr_database_replica_states**
  - Database replica state information, last hardened LSN and time, last redone LSN and time, log send queue size, log send rate, redo queue size, redo rate, last commit LSN, last commit time

© SQLskills, All rights reserved.
https://www.SQLskills.com

Migrating from Database Mirroring

- **Non-SQL Server staff must be involved as there are WSFC, networking, and storage considerations**
- **Relying on transparent client redirect?**
  - Once you move past a second replica, you'll need to move to an AG Listener or other approach
- **Existing database mirroring sessions?**
  - SQL Server instances need to exist on nodes of the same WSFC
    - New considerations around WSFC Quorum Mode and Node Votes
  - Enable AlwaysOn HA for each SQL Server service (requires restart)
  - Remove database mirroring(SET PARTNER OFF) on principal
  - Temporarily stop log backup jobs on principal database
  - Create AG on principal instance and the existing "mirror" recovering database(s) can be "joined only" to w/o having to restore again
  - Create AG Listener

SQLskills immerse yourself in sql server 45

© SQLskills, All rights reserved.
https://www.SQLskills.com

Cross-cluster Migration of AGs

- **Windows Server Failover Clustering (WSFC) does *not* support operating system rolling upgrades**
- **SQL Server 2012 SP1 introduced support for cross-cluster migration of availability groups for deployments to a new WSFC cluster**
- **A cross-cluster migration moves one availability group or a batch of availability groups to the new, destination WSFC cluster with minimal downtime**
- **Path**
  - Source: Windows Server 2008 or 2008 R2 + SQL Server 2012 RTM or SP1
  - Destination: Windows Server 2012 + SQL Server 2012 SP1
- **See "Cross-cluster Migration of AlwaysOn Availability Groups for Operating System Upgrades" http://bit.ly/10cFqPU**
- **Not supported after SQL Server 2014 – use Distributed AG's for migration**

SQLskills immerse yourself in sql server 46

© SQLskills, All rights reserved.
https://www.SQLskills.com

Quorum Model

- **Quorum model is managed via the WSFC**
  - Number of 'elements' needed to keep the WSFC running, and also avoid 'split brain' by ensuring majority of elements are visible
- **Must consider non-shared vs. shared-storage models**
  - Node and Disk Majority (shared disk)
  - No Majority: Disk Only (shared disk)
  - Node Majority
  - Node and File Share Majority
- **Odd vs. even number of nodes drives quorum decisions**
  - Node Majority
  - Node and File Share Majority

SQLskills immerse yourself in sql server

47

© SQLskills, All rights reserved.
https://www.SQLskills.com

Node Votes

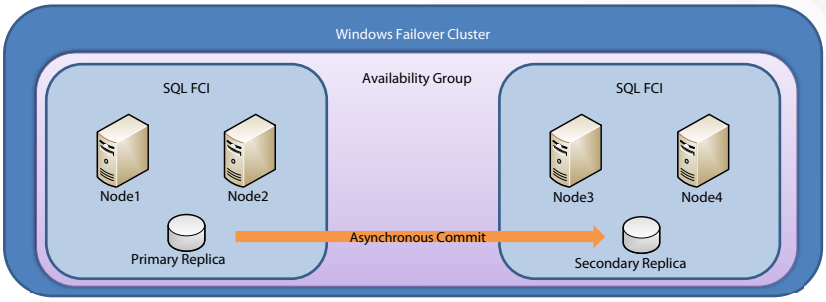
- **By default every node has a 'vote' in quorum, but this may not be appropriate for your topology**
  - Example: you may only want nodes in the primary data centers to have votes, and then no votes in the disaster recovery data center
- **Windows Server hotfix KB 2494036 allows assignment of votes via NodeWeight property**
  - Can configure PowerShell cmdlet or Cluster.exe
- **Windows Server 2012+ offers:**
  - NodeWeight handling by default
  - Dynamic Quorum

SQLskills immerse yourself in sql server

48

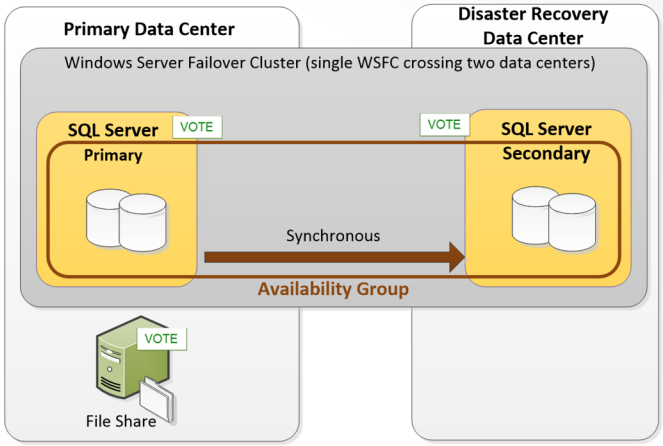
© SQLskills, All rights reserved.
https://www.SQLskills.com

Node Votes Example



© SQLskills, All rights reserved.
https://www.SQLskills.com

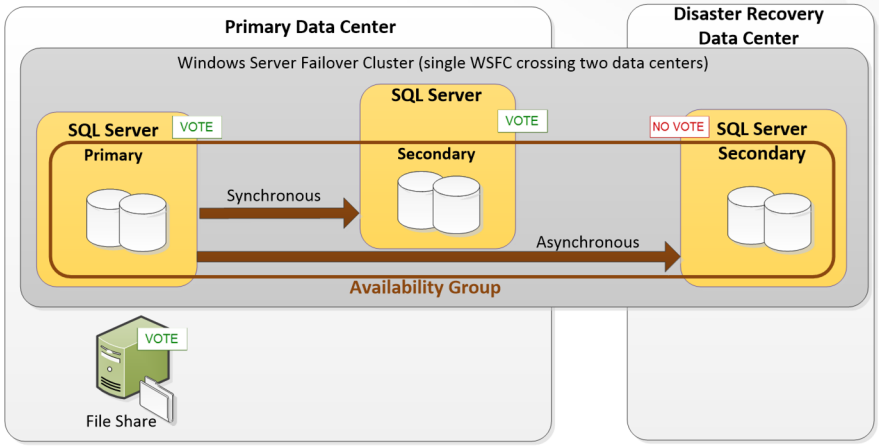
Node Votes Example



Source: "Building a High Availability and Disaster Recovery Solution using AlwaysOn Availability Groups (Joe Sack, Sanjay Mishra), http://bit.ly/L9RbLg

© SQLskills, All rights reserved.
https://www.SQLskills.com

Node Votes Example



Source: "Building a High Availability and Disaster Recovery Solution using AlwaysOn Availability Groups (Joe Sack, Sanjay Mishra), http://bit.ly/L9RbLg

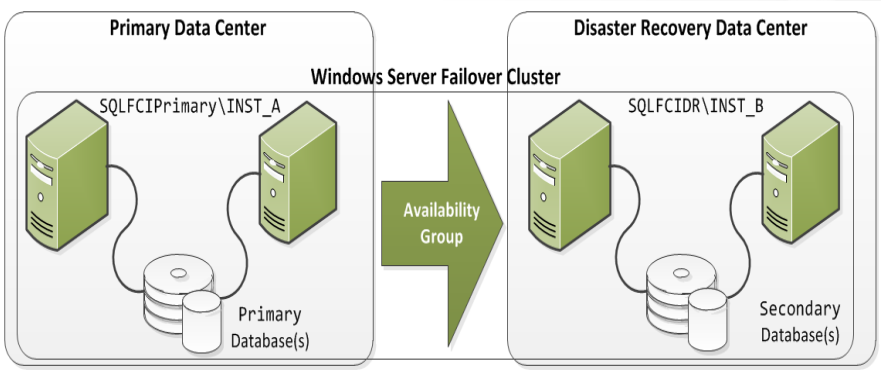
© SQLskills, All rights reserved.
https://www.SQLskills.com

Asymmetric Storage

- Consider a scenario where you have an AG that has replicas on two SQL FCIs across a multi-site WSFC…
- Shared storage on the primary data center may not be visible to storage on the second data center
- Functionality to share storage between just a subset of nodes is called asymmetric storage
  - Symmetric storage used to be required
- Asymmetric storage supported in Windows Server 2008 R2 SP1 and Windows Server 2008 via hot-fix (KB 976097)

© SQLskills, All rights reserved.
https://www.SQLskills.com

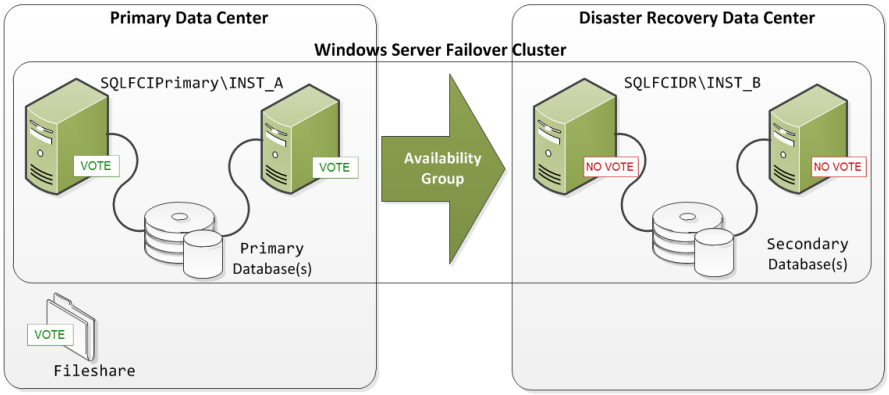
SQL FCI for HA and AG for DR



Source: "AlwaysOn Architecture Guide: Building a High Availability and Disaster Recovery Solution by Using Failover Cluster Instances and Availability Groups (Joe Sack, Sanjay Mishra), http://bit.ly/Rjql54

© SQLskills, All rights reserved.
https://www.SQLskills.com

SQL FCI + AG Vote Assignments

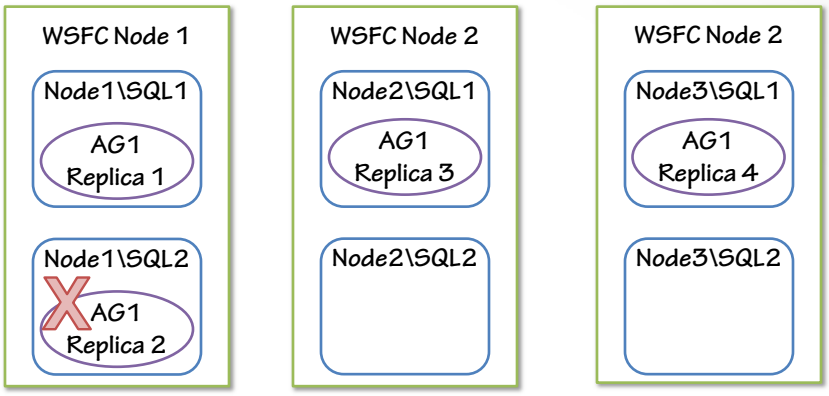


Source: "AlwaysOn Architecture Guide: Building a High Availability and Disaster Recovery Solution by Using Failover Cluster Instances and Availability Groups (Joe Sack, Sanjay Mishra), http://bit.ly/Rjql54

© SQLskills, All rights reserved.
https://www.SQLskills.com

One WSFC Node per Availability Replica

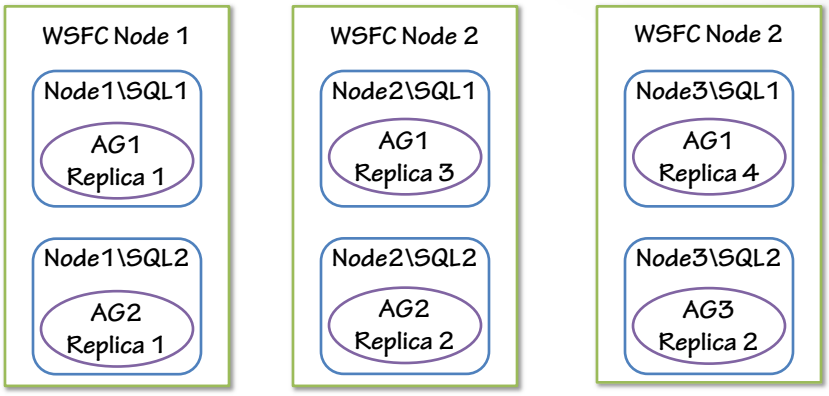
- A single WSFC node can only host one availability replica per availability group



© SQLskills, All rights reserved.
https://www.SQLskills.com

Instances and Replicas

- On one WSFC node, greater than one or more instances can host different availability group replicas



© SQLskills, All rights reserved.
https://www.SQLskills.com

Using Failover Cluster Manager (1)

- **Failover Cluster Manager (FCM) will let you do bad things to your availability group resources**
- ***Don't* use FCM to change the following availability group and availability group listener resource settings:**
  - Preferred owners
  - Possible owners
- ***Don't* use FCM to fail over the availability group or availability group listener**
- ***Don't* use FCM to add or remove resources contained within the availability group and availability group listener resource groups**

© SQLskills, All rights reserved.
https://www.SQLskills.com

Using Failover Cluster Manager (2)

- **For Failover Cluster Instances, you may need to configure possible node owners for asymmetric storage purposes, removing those nodes that don't have access to the shared storage**
  - FCI network name, IP address, SQL Server service, Agent service, other FCI resources
  - Not configuring this will cause the availability group creation to fail as the nodes can be possible owners to both replicas ("crossing streams")

© SQLskills, All rights reserved.
https://www.SQLskills.com

Security Considerations

- **Dependent on Kerberos?**
  - Use the same service account across replicas
  - Domain admin will need to manually register a Service Principal Name (SPN) for the SQL Server service account for the virtual network name (VNN) of the availability group listener
  - Changing the account requires a manual SPN re-register
- **Deploying availability groups without Active Directory permissions?**
  - Pre-stage the Cluster Name Object

SQLskills immerse yourself in sql server

59

© SQLskills, All rights reserved.
https://www.SQLskills.com

Performance Considerations

- **You can have hardware disparity across nodes – but should you?**
  - I/O path and disk subsystem
  - Cores
  - Memory
- **Consider dedicated NICs for AG traffic**
  - See "Configuring a Dedicated Network for Availability Group Communication", http://bit.ly/STMCMQ
  - Even on latent availability groups, you'll see Bytes Received from Replica/sec performance counter generate persistent traffic *per database*
- **Failover client connectivity**
  - Newer providers .NET 4.0, SNAC can cache IPs and don't have to wait for DNS
    - MultiSubnetFailover = True
    - Older clients, consider modifying HostRecordTTL value (reducing to 5 minutes, for example) and increasing connection timeout

SQLskills immerse yourself in sql server

60

© SQLskills, All rights reserved.
https://www.SQLskills.com

Worker Pool

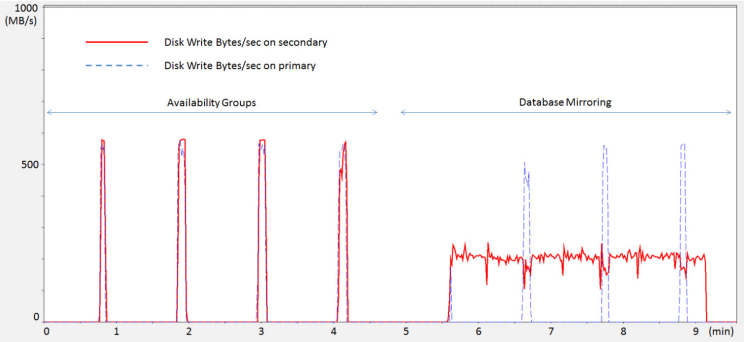
- **Maximum threads used by availability groups is configured max worker threads minus 40**
  - Availability replicas share a single thread pool
  - Idle threads get released to the thread pool
- **Thread distribution**
  - 1 log capture thread per primary database on primary replica
  - 1 log send thread for each secondary database
  - 1 redo thread per secondary replica
  - Secondary replica backup holds a thread on the primary replica for duration of the operation

© SQLskills, All rights reserved.
https://www.SQLskills.com

I/O Redo Improvements

- **Database mirroring continually flushes dirty pages**
  - Facilitates faster 'restoring' to 'online' state
- **Availability groups flush on checkpoint or buffer pool pressure**
  - Flush is not required on a role change given "online" state (readable or not)

Source: SQLCat.com "Comparing I/O characteristics in AlwaysOn Availability Groups and Database Mirroring" http://bit.ly/KwuoZH

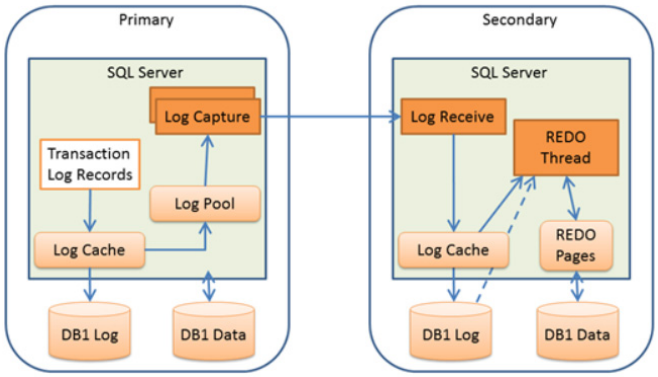


© SQLskills, All rights reserved.
https://www.SQLskills.com

Log Pool Cache Mechanism (1)

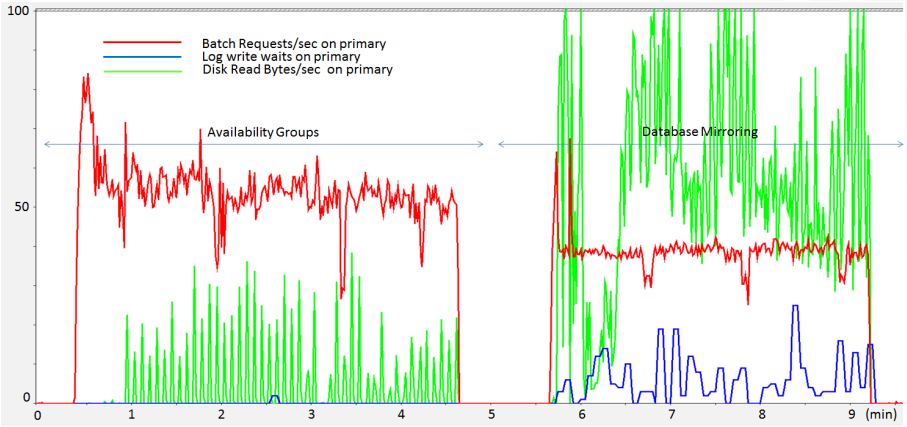
- New log pool cache mechanism for log record consumers
- If an unsent log is not in the cache, threads have to read from the file
- The log pool is a dynamic cache that can grow if no memory pressure

Source: SQLCat.com "Comparing I/O characteristics in AlwaysOn Availability Groups and Database Mirroring" http://bit.ly/KwuoZH



© SQLskills, All rights reserved.
https://www.SQLskills.com

Log Pool Cache Mechanism (2)



Source: SQLCat.com "Comparing I/O characteristics in AlwaysOn Availability Groups and Database Mirroring" http://bit.ly/KwuoZH

© SQLskills, All rights reserved.
https://www.SQLskills.com

Configuring Replication with AG's

- Does not support Peer-to-Peer Replication publishers or the distribution database
- Configure distribution on remote distributor
  - Add all AG replicas as publishers to distributor
- Configure publisher on original primary replica
- Configure remote distribution on all possible publishers
- Configure replica hosts as replication publishers
  - Add linked server to the distribution database server
- Redirect publication to AG Listener name with sys.sp_redirect_publisher in the distribution database
- Validate configuration in the distribution database using sys.sp_validate_replica_hosts_as_publishers
- Create/configure subscriptions

SQLskills immerse yourself in sql server

65

© SQLskills, All rights reserved.
https://www.SQLskills.com

Replicated Database Consistency with AG's

- The log reader will not process log records that have not been hardened at all availability group secondary replicas by default to ensure that on failover, all rows replicated to a subscriber also are present at the new primary.
- When the publisher is configured with two replicas (one primary and one secondary) and a failure happens, if the original primary replica remains down, the logreader does not move forward until all secondary databases are brought back online or until the failed replica is removed from the availability group.
  - To allow the logreader to proceed, remove the original primary replica using ALTER AVAILABITY GROUP <group_name> REMOVE REPLICA, then add a new secondary replica to return high availability or disaster recovery
- Trace flag 1448 - enable the log reader to move forward when asynchronous replicas have not acknowledged receipt of log record

SQLskills immerse yourself in sql server

66

© SQLskills, All rights reserved.
https://www.SQLskills.com

Collaboration is a Requirement

- **Not just for initial configuration, but for high availability and disaster recovery scenarios**
  - DBA, Windows Server, network, storage
  - More involved than database mirroring
- **Disaster-recovery scenario definitely requires a run-book and a detailed assignment of duties, for example, on bringing up service at the disaster recovery data center:**
  - Confirm state of primary data center (DBA, Windows, network)
  - Force quorum on DR site node (Windows)
  - Force AG group to DR site (DBA)
  - Add votes to DR nodes (Windows)

SQLskills immerse yourself in sql server

67

© SQLskills, All rights reserved.
https://www.SQLskills.com

SQL Server 2019 Enhancements

- **Supports up to five synchronous commit replicas**
  - One primary replica, plus four synchronous secondary replicas
- **Secondary connection redirection for configurations without a listener**
  - Cluster technology doesn't offer listener capability
  - Read Scale Availability Groups with CLUSTER_TYPE = NONE
  - Cloud multi-subnet configurations with floating IP's
    - Complex, prone to errors and difficult to troubleshoot
  - The replica PRIMARY_ROLE must include a READ_WRITE_ROUTING_URL specification
  - The connection string must define ApplicationIntent=ReadWrite (default)
- **Support for Availability Groups on containers in a Kubernetes cluster**

SQLskills immerse yourself in sql server

68

© SQLskills, All rights reserved.
https://www.SQLskills.com

Key Takeaways

- Successful deployment of an Availability Group solution requires good team dynamics, especially for multi-site/subnet configurations to troubleshoot problems when they occur
- Windows Server 2012+ should be used to leverage Dynamic Quorum and last man standing for Availability Group nodes
- SQL Server 2014+ enhancements for readable secondaries allow read-only workloads to continue during a loss of quorum or failover of the Availability Group
- Understand the impacts of read-only workloads and row versioning on the primary replica's index fragmentation

SQLskills immerse yourself in sql server

69

© SQLskills, All rights reserved.
https://www.SQLskills.com

Availability Group Resources

- Whitepapers
  - AlwaysOn Architecture Guide: Building a High Availability and Disaster Recovery Solution by Using AlwaysOn Availability Groups
    - http://msdn.microsoft.com/en-us/library/jj191711
  - AlwaysOn Architecture Guide: Building a High Availability and Disaster Recovery Solution by Using Failover Cluster Instances and Availability Groups
    - http://msdn.microsoft.com/en-us/library/jj215886
  - AlwaysOn Solution Guide: Offloading Read-Only Workloads to Secondary Replicas
    - http://msdn.microsoft.com/en-us/library/jj542414

SQLskills immerse yourself in sql server

70

© SQLskills, All rights reserved.
https://www.SQLskills.com

Review

- Availability group overview
- Readable secondaries
- Application connection support
- Partial database containment
- Migrating from database mirroring
- Quorum model and node votes
- Asymmetric storage considerations
- Security
- Performance

SQLskills immerse yourself in sql server

71

© SQLskills, All rights reserved.
https://www.SQLskills.com

Questions?

SQLskills immerse yourself in sql server