New Features in 2016's SQL Server AlwaysOn Availability Groups

SQL Server AlwaysOn Availability Groups were first introduced in 2012 and continue to receive enhancements as each new version of the database is released. For SQL Server 2016, Microsoft focused on addressing customer pain points with the high-availability and disaster recovery feature, rather than adding new functionality.

Microsoft has added Distributed Transaction Coordinator support, optimized log transport and other features to AlwaysOn Availability Groups in SQL Server 2016.

Failover on Database Health

Even though the focus of Availability Group (AG) failover has been to provide redundancy at the database level, or group of related databases, the actual failover has been triggered by the health of the SQL Server instance as a whole. For example, a database in an availability group could be offline, perhaps due to a disk failure, and the rest of the instance is up and operational. With SQL Server 2014 and earlier implementations, failover would not be triggered because the overall health of the instance is still good. In SQL Server 2016, all databases within the affected AG would failover.

This behaviour is optional. By default, it won't be enabled; the database administrator will need to enable the policy in order to have failover whenever a database becomes unhealthy.

Unified HA Solution Greater scalability: Load balancing readable secondaries Increased number of auto-failover targets Log transport performance Improved manageability: DTC support Database-level health monitoring Group managed service account

Distributed Transaction Coordinator Support

One of the most requested enhancements for SQL Server AlwaysOn Availability Groups was to support the **Distributed Transaction Coordinator (DTC)** to manage transactions across multiple databases and instances. The DTC is not supported in availability groups in SQL Server 2014 and earlier.

Since the DTC is part of the Windows operating system, Microsoft made modifications to Windows Server 2016 and the API so that SQL Server 2016 can fully support it. Therefore, your SQL Server 2016 instance must reside on a Windows Server 2016 OS in order to have the DTC supported. This will also be supported on Windows Server 2012 Release 2 after rollup patch KB3090973 is installed.

Synchronous Replicas

There are two types of secondary replicas within an AG -- asynchronous and synchronous. With **Synchronous Replication**, the transaction must first be recorded in the replica's log before it can send an acknowledgement to the primary replica. In order to have a secondary replica be an automatic failover target, it must be a synchronous replica. In SQL Server 2014, there can be three synchronous replicas, but only two of those can be designated as automatic failover targets. This means those two synchronous replicas would be the primary failover target for each other. In SQL Server 2016, all three synchronous replicas can now be designated as failover targets. As a result, if an issue occurred during or even after a failover, the third synchronous replica could also participate as a failover target.

Optimized Log Transport

Some customers have deployed AGs in a high-volume, highly transactional environment with solid-state disks. In these environments, it has been a challenge at times for the synchronous communication of the log data between replicas to keep up with the high-speed hardware. This can be problematic in this type of environment as time moves forward, and the latency from the synchronous commit process can start to affect the performance of the primary.

Microsoft has worked to streamline the pipeline between the synchronous replicas to gain better log-data throughput when utilizing SQL Server AlwaysOn Availability Groups. But if the throughput is faster, there will be that much more "redo" activity to be processed when replicating log records for database recovery purposes. So, Microsoft is also working on enabling the redo operation to become parallel in order to keep up. Instead of a percent-improvement goal over the throughput in SQL Server 2014 to SQL Server 2016, the goal is to be as close as possible to a non-AG standalone instance in terms of throughput.

Load Balancing Across Readable Secondary Replicas

In SQL Server 2014, using the availability group listener mechanism to offload reads to the secondary replicas is supported via the read-only routing list. But the first replica in the list gets the most activity because it's the one that is always tried first.

In SQL Server 2016, the list of readable secondary replicas offers up connection information on a round robin basis. Also, each replica has its own read-only routing list so that read-only balancing via the availability group listener could route traffic to secondary replicas that may be more "local" to the current primary replica.

Group Managed Service Accounts

A group Managed Service Account (gMSA) is a type of security account released in Windows Server 2012 that SQL Server 2016 can fully leverage -- especially with AGs. The group account provides similar security capabilities as a local managed service account on an individual server, but it has a domain scope. When an SQL Server instance is installed with the default settings, a local service account is used to manage passwords for applications and processing services. If the instance is part of an AG, setting up and maintaining permissions to access common resources such as file shares often becomes complex as security provisions need to be established for the service account from each instance within the AG. A gMSA addresses this issue. In addition, some organizations use a regular domain user account to manage service passwords -- a practice that could be flagged in a security audit, as someone lacking authorization could log in with those credentials. Using a gMSA prevents that and also automatically manages password synchronization across all nodes.

Overall, there isn't a really flashy new element to point to among the additions to AGs in SQL Server 2016. But the planned enhancements provide real value in terms of availability, compatibility, scalability and manageability. These are the types of enhancements that address real-world implementation issues discovered in production applications.

References

- https://msdn.microsoft.com/en-us/library/hh213151.aspx
- https://msdn.microsoft.com/enus/library/windows/desktop/ms684146(v=vs.85).aspx