SIOS

SIOS WHITEPAPER

Understanding Disaster Recovery Options for SQL Server

us.sios.com

One of the challenges IT and database administrators confront when implementing disaster recovery provisions is choosing from among the myriad options available. Existing high availability configurations designed to minimize downtime for critical applications may not be adequate for recovering fully from a widespread disaster. And existing disaster recovery provisions may not be as comprehensive or cost-effective as they could be.

Assuring business continuity requires careful planning. The plan must address all aspects of the business, and one of the most important involves Information Technology. Businesses today run on data. Without a solid business continuity strategy and a disaster recovery plan for IT, the organization risks losing access to data or, worse yet, actually losing valuable data.

This white paper provides some practical guidance to help system and database administrators tasked with creating business continuity and disaster recovery (DR) plans. The first section outlines seven steps for creating a business continuity plan. The second section offers some helpful context for creating a comprehensive and cost-effective DR plan. The third section highlights the most popular options for providing DR protection for SQL Server databases. The fourth and final section describes adding DR to an existing high-availability failover cluster.

Business Continuity Planning

For those administrators who hate to plan, General George Patton offers this advice: “A good plan today is better than a perfect plan tomorrow.” No business continuity (BC) or disaster recovery plan can possibly address every possible event or set of circumstances, which is why both should continually evolve as lessons learned inform various improvements.

Providing the guidance needed to create a solid BC plan would fill a book. But because the business continuity plan forms the foundation for the disaster recovery plan, at least some discussion is warranted here. What follows is a summary of seven steps that have proven to be useful when creating and enhancing BC plans.

Step #1: Prepare to Plan – This step mostly involves gathering pertinent information about key personnel, customers, suppliers, facilities, utilities, security provisions, records, operating procedures and processes, service and licensing agreements, applicable privacy regulations, etc. If the business depends on it for anything critical to operations, it should be included.

Step #2: Establish Plan Objectives – The BC plan must support the organization’s core mission, and that requires establishing a set of objectives based on an assessment of possible disruptions. Of particular interest to IT are the recovery time and recovery point objectives (covered below), as well as the budget available before, during and after a disruption.

Step #3: Identify and Prioritize Potential Threats and Impacts – While it is not possible to foresee every way business might someday be disrupted, there are likely threats based on the organization’s locations and circumstances. Every facility could lose power, but only some might experience a tornado, hurricane or earthquake. Use probabilities to determine

priorities and estimate the potential duration of every threat.

Step #4: Develop Mitigation and Business Continuity Strategies – This is the core of the BC plan, and should include ways to minimize business impacts before, during and after recovering from a disruption. For IT, the mission-criticality of each application will be used to determine its priority in the DR plan. For all departments, the ability to maintain communications will be key, especially in the event some aspect of the plan fails and a contingency is urgently needed.

Step #5: Identify Teams and Tasks – This step could be included in Step #4, but is kept separate here to emphasize its importance. After all, it is people who will implement the BC plan and people who will take action to compensate for any of the plan's deficiencies, such as critical tasks not included in a checklist. This step should also establish a line of succession with alternate members or teams identified should the primary ones be unavailable.

Step #6: Test the Plan – The best way to uncover holes in the plan and prepare teams for implementing it is to test it—thoroughly and regularly—by simulating business disruptions caused by the threats identified. Scheduled power outages or major upgrades can serve as ideal opportunities to conduct these tests, but some should also occur unannounced.

Step #7: Maintain/Enhance the Plan – This step is ongoing and serves as the feedback loop for adjusting, updating, enhancing and otherwise maintaining the plan based on lessons learned during the tests and actual disruptions. Anything new, such as a new facility, application or service, should also go through the planning process separately or as part of this ongoing step.

Disaster Recovery Planning

The disaster recovery plan for IT builds on the BC plan with specific strategies to protect all data and ensure critical applications can continue to run with minimal or no disruption. There are two aspects of the BC plan that are fundamental to DR planning: the threat assessment in Step #3 and the business impact analysis in Step #4. The former identifies those disasters the organization is most likely to experience, while the latter determines which applications are mission-critical.

The plan should recognize the difference between "failures" and "disasters" because that difference determines the different provisions needed for high availability (HA) and disaster recovery. Failures are short in duration and small in scale, affecting a server, rack, or the power or cooling in a datacenter. Disasters have enduring impacts and are more widespread, affecting entire datacenters in ways that preclude rapid localized recovery. For example, a tornado, hurricane or earthquake might knock out power and networks, and close roads, making the datacenter—and corporate offices—inaccessible for days.

Perhaps the biggest difference involves replication to redundant resources (systems, software and data), which can be local—on a Local Area Network—to recover from a failure. By contrast,

the redundancy required to recover from a widespread disaster must span a Wide Area Network. For database applications that require high transactional throughput performance, the ability to replicate the active instance's data synchronously across the LAN enables the standby instance to be "hot" (in synch with the active instance), ready to take over immediately and automatically in the event of a failure. Such rapid response should be the goal of all HA provisions.

Because latency inherent in the WAN would adversely impact on the throughput performance in the active instance when using synchronous replication, data is usually replicated asynchronously in DR configurations. This means that updates to the standby instance always lag behind updates being made to the active instance, which makes the standby instance "warm" and could result in some data loss with an automatic failover. A manual recovery process, while taking longer to complete the failover, can assure there is no data loss.

Another difference is the impossibility of having a Storage Area Network (SAN) or other form of shared storage that spans the WAN. Some failover clustering solutions, most notably Windows Server Failover Clustering (WSFC) and SQL Server's Failover Cluster Instances (FCIs) require shared storage, which is also not available in the public cloud. This means that WSFC and FCIs require, at a minimum, a separate data replication solution to be used for both HA and DR purposes in the cloud.

These differences lead to differences in the Recovery Time Objectives and Recovery Point Objectives established for HA and DR purposes. RTO is the maximum tolerable duration of an outage. Mission-critical applications have low RTOs, normally on the order of a few seconds for HA, and high-volume online transaction processing applications generally have the lowest. For DR, RTOs of many minutes or even hours are fairly common owing to the extraordinary cost of implementing provisions capable of fully recovering from a widespread disaster in mere minutes.

RPO is the maximum period during which data loss can be tolerated. If no data loss is tolerable, then the RPO is zero. Because most data has great value (Otherwise there would be no need to capture and store it.) low RPOs are common for both HA and DR purposes. For HA, synchronous data replication makes it relatively easy to satisfy a low or zero RPO.

The situation for DR is substantially different, however, with a low RPO creating the need for a potential tradeoff with RTO. Here's why: For applications with an RPO of zero, manual processes are required to ensure that all data (e.g. from a transaction log) has been fully replicated and verified on the standby instance before the recovery—in the form of a failover—can occur. This additional, potentially considerable effort has the effect of increasing the recovery times.

The DR Options

With a recognition that DR is different from HA, and that longer RTOs of many minutes or even hours are to be expected when recovering from a disaster, system and database administrators have considerable latitude when choosing different DR provisions for different applications.

In the cloud, all three major cloud service providers (CSPs) have DR offerings suitable for most applications. Google has what could be called DIY (Do-It-Yourself) DR guided by templates, cookbooks and other tools. DIY is a viable option because, compared to HA provisions, DR is relatively easy to implement with data backups or snapshots and "warm" standby instances, all of which are available in every cloud. Microsoft and Amazon have managed DR-as-a-Service (DRaaS) offerings: Azure Site Recovery (ASR) and CloudEndure Disaster Recovery, respectively. For all three CSPs it is important to note that at least some manual intervention is required to affect a full recovery.

The DIY DR option leverages procedures that should already be in place for most applications. For example, all organizations routinely backup data and/or take snapshots for recovery and/or archiving purposes. For database applications, it is common to create transaction logs that can be applied, much like incremental backups are, to a "warm" standby version or the most recent full backup of the database. A best practice is to store these duplicates of the data at a remote location, where there are also standby resources (hardware, and system and application software) capable of running the application. It takes more time to recover from outages with DIY DR, but the relatively low cost can make this a viable option for many applications.

While DR is different from HA, it is possible (and generally preferable) to add DR to an existing HA configuration, which is covered in the next section. There are two popular options for combining HA and DR provisions for SQL Server: SQL Server's own Always On Availability Groups feature and third-party failover clustering software.

Always On Availability Groups replaced database mirroring in SQL Server 2012 Enterprise Edition, and this feature is also included in SQL Server 2017 for Linux. This is SQL Server's most robust HA/DR offering, capable of delivering rapid, automatic failovers with no data loss for HA, and/or protecting against widespread disasters by leveraging asynchronous replication with minimal or no data loss. But it requires licensing the more expensive Enterprise Edition, making it cost-prohibitive for many applications, and it lacks protection for the entire SQL instance. For Linux, which lacks a feature equivalent to Windows Server Failover Clustering, there is a need for additional commercial and/or open source software to provide HA and DR protections.

A notable disadvantage with application-specific options like Always On Availability Groups is the need for administrators to use other HA and/or DR solutions for all non-SQL Server applications. Having multiple HA/DR solutions inevitably increases complexity and costs (for licensing, training, implementation and ongoing operations), which is why many organizations prefer using application-agnostic third-party solutions.

Third-party failover clustering solutions are the second HA/DR combo option. These are purpose-built to support virtually all applications running on Windows Server and Linux in public, private and hybrid clouds. They are implemented entirely in software and usually include real-time data replication, continuous monitoring for detecting failures at the system and application levels, and configurable policies for failover and failback.

SANless failover clustering solutions that integrate with Windows Server Failover Clustering

enable the use of SQL Server Failover Cluster Instances (FCIs) spanning datacenters and cloud regions. SQL Server FCIs are supported by both the Standard and Enterprise Editions of SQL Server for Windows. Because SQL Server 2017 for Linux lacks the equivalent of WSFC, the failover clustering solution must handle all data replication and other functionality directly.

SIOS Technology offers two separate SANless failover clustering solutions—one for Windows Server and one for Linux—that are both designed to provide complete and cost-effective HA and DR protections.

SIOS DataKeeper for Windows Server is available in both a Standard Edition and a more robust Cluster Edition. The Standard Edition provides real-time data replication for DR protection in a Windows Server environment. The Cluster Edition provides seamless integration with WSFC, making it possible to create SANless clusters in the cloud. The ability to deploy robust HA configurations with FCIs in SQL Server's Standard Edition eliminates the need to upgrade to the Enterprise Edition just for Always On Availability Groups. SIOS DataKeeper supports all versions of SQL Server back to SQL Server 2008.

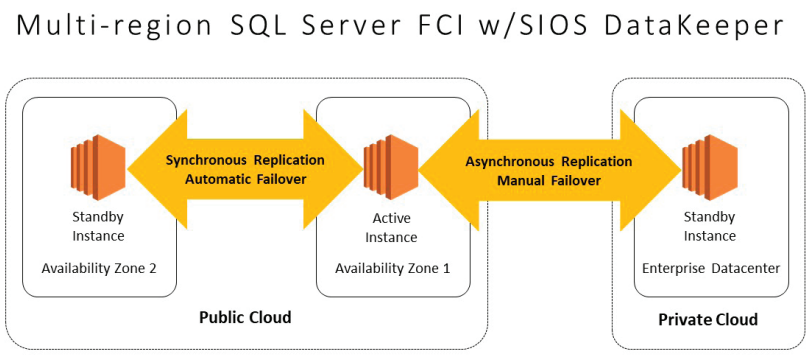
SIOS Protection Suite for Linux provides the equivalent of the DataKeeper Cluster Edition in a complete DR/HA solution that combines real-time data replication with application-level failover clustering comparable to that provided by WSFC. The suite eliminates the need for organizations to struggle with do-it-yourself open source software projects. SIOS Protection Suite supports the only version of SQL Server currently available for Linux, SQL Server 2017.

Like most other third-party failover clustering software, both SIOS solutions are application-agnostic, which eliminates the need to have different HA/DR provisions for different applications. Being SANless overcomes impediments caused by the lack of shared storage in the cloud, while making it possible to leverage the cloud's many resiliency-related capabilities, including availability zones and regions. It is for this reason that SIOS SANless failover clusters are able to operate seamlessly in private, public and hybrid cloud environments.

Adding DR to a SANless Failover Cluster

Any application that requires high availability to prevent downtime caused by failures also needs the ability to recover from a widespread disaster. One way to do that is to add DR protection to an existing HA failover cluster, and such configurations afford some compelling advantages. Having a combined solution is easier to manage and can be more cost-effective than having separate solutions for HA and DR. Integrated management makes it easier to test various failure and disaster scenarios. And it more easily facilitates minimizing downtime during routine hardware and software upgrades.

The diagram shows a popular configuration for a SANless failover cluster that provides both HA and DR protections in a hybrid cloud. The cluster spreads three SQL Server instances across two availability zones in a public cloud and a distant datacenter in a private cloud. For the two-node HA cluster in the public cloud, data replication is synchronous, and failovers can be configured to occur automatically. The third instance in the private cloud uses asynchronous data replication and a manual recover process to protect against widespread disasters.



This hybrid cloud configuration of a SIOS SANless failover cluster consists of a two-node HA cluster spanning two availability zones in a public cloud, with DR protection provided by a third instance deployed in an on-premises private cloud.

Note that the two-node HA configuration could be in the private cloud with the DR instance in the public cloud. Note also how this configuration overcomes yet another limitation—this one in the Standard Edition of SQL Server—of being able to have a maximum of only two FCI nodes in a failover cluster.

It is true that using a third-party failover clustering solution increases costs. But weighing that relatively modest increase against the cost of downtime, plus avoidance of needing to license the more expensive Enterprise Edition just for HA/DR, plus the savings afforded by the cloud make a compelling case for using a SANless failover clustering solution to implement or improve HA and/or DR protections for your SQL Server databases.

To help you get started, SIOS offers free trial versions of both SIOS DataKeeper for Windows Server and the SIOS Protection Suite for Linux, and these are available on the Web at us.sios.com. SIOS also offers comprehensive documentation, an assortment of templates that automate all or part of application-specific and/or cloud-specific configurations, responsive support, and variety of other useful resources to help ensure successful deployments. To learn more about how your organization can benefit from the carrier-class HA and DR protection afforded by SANless failover clustering from SIOS Technology, please contact SIOS by phone at (650)645-7000 or by email at info@us.sios.com.

However you choose to protect your SLQ Server databases, keep in mind that the only thing harder than doing something—anything—to better prepare for recovering from a disaster is trying to explain why you didn't.

.

SIOS Technology Corp.
155 Bovet Road, Suite 476
San Mateo, CA 94402
Tel: 650-645-7000

info@us.sios.com
https://us.sios.com

© 2023 SIOS Technology Corp. All rights reserved. SIOS, SIOS Technology, SIOS DataKeeper, SIOS LifeKeeper, SIOS Protection Suite and associated logos are registered trademarks or trademarks of SIOS Technology Corp. and/or its affiliates in the United States and/or other countries. All other trademarks are the property of their respective owners. WP-1011-B