TR-4891: SAP HANA disaster recovery with Azure NetApp Files

NetApp Solutions

Dorian Henderson, Ivana Devine
July 21, 2021

Table of Contents

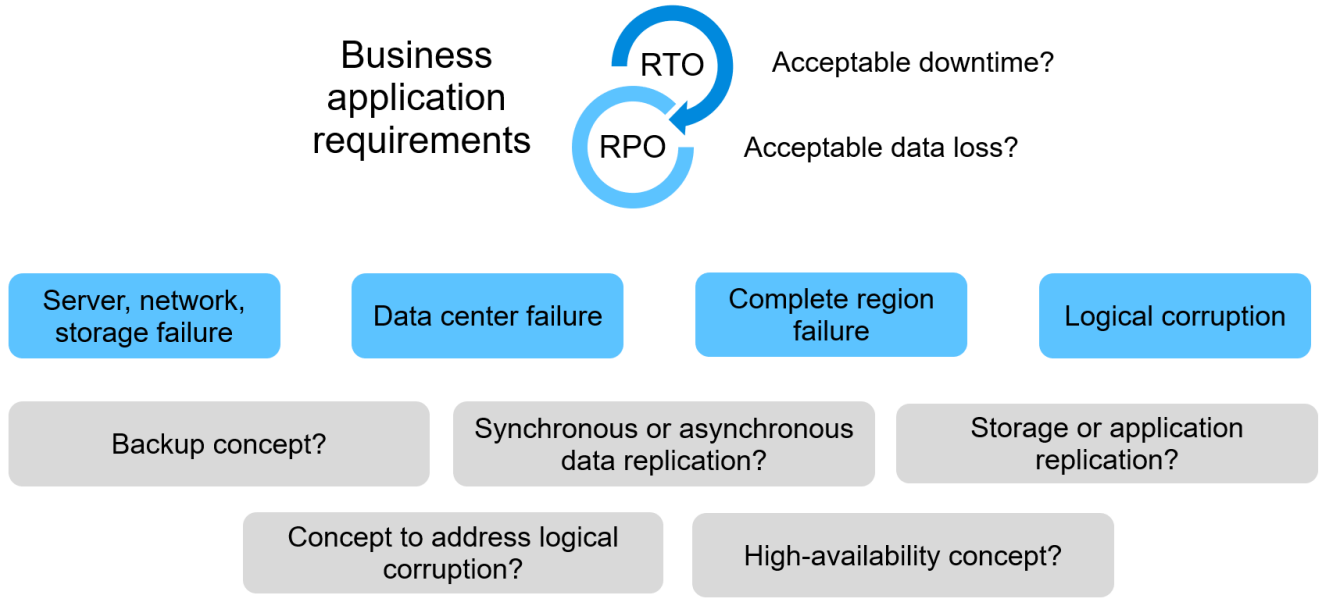
TR-4891: SAP HANA disaster recovery with Azure NetApp Files

Nils Bauer, NetApp
Ralf Klahr, Microsoft

Studies have shown that business application downtime has a significant negative impact on the business of enterprises. In addition to the financial impact, downtime can also damage the company's reputation, staff morale, and customer loyalty. Surprisingly, not all companies have a comprehensive disaster recovery policy.

Running SAP HANA on Azure NetApp Files (ANF) gives customers access to additional features that extend and improve the built-in data protection and disaster recovery capabilities of SAP HANA. This overview section explains these options to help customers select options that support their business needs.

To develop a comprehensive disaster recovery policy, customers must understand the business application requirements and technical capabilities they need for data protection and disaster recovery. The following figure provides an overview of data protection.



Business application requirements

There are two key indicators for business applications:

- The recovery point objective (RPO), or the maximum tolerable data loss
- The recovery time objective (RTO), or the maximum tolerable business application downtime

These requirements are defined by the kind of application used and the nature of your business data. The RPO and the RTO might differ if you are protecting against failures at a single Azure region. They might also differ if you are preparing for catastrophic disasters such as the loss of a complete Azure region. It is important to evaluate the business requirements that define the RPO and RTO, because these requirements have a significant impact on the technical options that are available.

High availability

The infrastructure for SAP HANA, such as virtual machines, network, and storage, must have redundant components to make sure that there is no single point of failure. MS Azure provides redundancy for the different infrastructure components.

To provide high availability on the compute and application side, standby SAP HANA hosts can be configured for built-in high availability with an SAP HANA multiple-host system. If a server or an SAP HANA service fails, the SAP HANA service fails over to the standby host, which causes application downtime.

If application downtime is not acceptable in the case of server or application failure, you can also use SAP HANA system replication as a high-availability solution that enables failover in a very short time frame. SAP customers use HANA system replication not only to address high availability for unplanned failures, but also to minimize downtime for planned operations, such as HANA software upgrades.

Logical corruption

Logical corruption can be caused by software errors, human errors, or sabotage. Unfortunately, logical corruption often cannot be addressed with standard high-availability and disaster recovery solutions. As a result, depending on the layer, application, file system, or storage where the logical corruption occurred, RTO and RPO requirements can sometimes not be fulfilled.

The worst case is a logical corruption in an SAP application. SAP applications often operate in a landscape in which different applications communicate with each other and exchange data. Therefore, restoring and recovering an SAP system in which a logical corruption has occurred is not the recommended approach. Restoring the system to a point in time before the corruption occurred results in data loss, so the RPO becomes larger than zero. Also, the SAP landscape would no longer be in sync and would require additional postprocessing.

Instead of restoring the SAP system, the better approach is to try to fix the logical error within the system, by analyzing the problem in a separate repair system. Root cause analysis requires the involvement of the business process and application owner. For this scenario, you create a repair system (a clone of the production system) based on data stored before the logical corruption occurred. Within the repair system, the required data can be exported and imported to the production system. With this approach, the productive system does not need to be stopped, and, in the best-case scenario, no data or only a small fraction of data is lost.

The required steps to setup a repair system are identical to a disaster recovery testing scenario described in this document. The described disaster recovery solution can therefore easily be extended to address logical corruption as well.

Backups

Backups are created to enable restore and recovery from different point-in-time datasets. Typically, these backups are kept for a couple of days to a few weeks.

Depending on the kind of corruption, restore and recovery can be performed with or without data loss. If the RPO must be zero, even when the primary and backup storage is lost, backup must be combined with synchronous data replication.

The RTO for restore and recovery is defined by the required restore time, the recovery time (including database start), and the loading of data into memory. For large databases and traditional backup approaches, the RTO can easily be several hours, which might not be acceptable. To achieve very low RTO values, a

backup must be combined with a hot-standby solution, which includes preloading data into memory.

In contrast, a backup solution must address logical corruption, because data replication solutions cannot cover all kinds of logical corruption.

Synchronous or asynchronous data replication

The RPO primarily determines which data replication method you should use. If the RPO must be zero, even when the primary and backup storage is lost, the data must be replicated synchronously. However, there are technical limitations for synchronous replication, such as the distance between two Azure regions. In most cases, synchronous replication is not appropriate for distances greater than 100km due to latency, and therefore this is not an option for data replication between Azure regions.

If a larger RPO is acceptable, asynchronous replication can be used over large distances. The RPO in this case is defined by the replication frequency.

HANA system replication with or without data preload

The startup time for an SAP HANA database is much longer than that of traditional databases because a large amount of data must be loaded into memory before the database can provide the expected performance. Therefore, a significant part of the RTO is the time needed to start the database. With any storage-based replication as well as with HANA System Replication without data preload, the SAP HANA database must be started in case of failover to the disaster recovery site.

SAP HANA system replication offers an operation mode in which the data is preloaded and continuously updated at the secondary host. This mode enables very low RTO values, but it also requires a dedicated server that is only used to receive the replication data from the source system.

Next: Disaster recovery solution comparison.