Whitepaper

Silk Highly Available Architectures for SQL Server



Contents

About Silk 2
Defining RPO and RTO Targets for Your Silk Application 3
Tier 1: Premium HA and DR 4
  Tier 1 Use Case 4
Tier 2: Critical HA and DR 5
  Tier 2 Use Case 6
Tier 3: Failover and Recover 6
  Tier 3 Use Case 7
Tier 4: Non-Critical DR 7
  Tier 4 Use Case 8
Conclusion 8

About Silk

Silk is more than just a storage solution or a cloud migration tool. We're an intelligent data platform that optimizes cloud resources to significantly improve performance, resiliency, data lifecycle management, and cost efficiency in the cloud. We operate between databases and cloud infrastructure to let our customers manage their data in ways cloud providers can't. Silk gives you the means to shift your mission-critical data to the cloud and operate on par with, or beyond, even the fastest on-prem environments.

Silk lets you use one platform to run multiple databases concurrently, giving you more ways to improve management of your data. Our platform features two-tier virtualization with a fully distributed, cloud-native, multi-cloud framework that decouples your compute and storage and promotes self-optimizing for scalability. Our dynamic database block size alignment uses an adaptive, log-structured datastore to improve performance, while our elastic architecture optimizes latency and throughput. We use a shared multi-tenant datastore so transaction, analytic, and mixed workloads can run together on pooled resources for greater efficiency. Silk is a database-, application-, and cloud-agnostic platform that helps companies harness the full power of the cloud.

The Silk Intelligent Data Platform can give you the tools you need to make your cloud environment run up to 10x faster. Contact our Sales team for a demo to see how Silk can help you make the most of your cloud resources.

Defining RPO and RTO Targets for Your Silk Application

Maintaining high availability and planning for disaster recovery are two critical business continuity components in your cloud data journey. Building a robust resiliency strategy lets you minimize workflow and operational interruptions and achieve faster recovery times when your applications experience downtime or adverse events. When defining your applications' recovery plans and recovery targets, you can choose how to leverage your SQL Server and Silk resources depending on your needs, offering more flexibility and control over your SLAs and cloud costs.

Silk is designed for resiliency and availability, serving as a reliable foundation to build a broader, layered resiliency plan for your applications. Features like replication are configurable based on the high availability and disaster recovery SLAs required for your SQL Server workloads. Our active-active system architecture is purpose-built with no single points of failure, ensuring high fault tolerance and fast disaster recovery. Our range of data management options lets you choose the performance, resiliency, and data access levels you need while maintaining your system's cost requirements.

Silk works with SQL Server's Always On availability groups to support one or more SQL Server instances in an availability zone and synchronous and asynchronous replication between SQL Server instances. The Always On feature is SQL Server's native high availability and disaster recovery solution, allowing you to define and manage primary and secondary databases within a failover environment. For more details on the Always On feature, check out Microsoft's What is an Always On availability group article.

Silk complements Always On to enhance your DR strategy through our native asynchronous, snapshot-based replication. Silk replication can copy data from one SDP to another across availability zones and to different destinations while maintaining the highest throughput levels, ensuring your data is rapidly available during recovery. For more details on Silk replication, check out the **Support Disaster Recovery through Replication** section of our Silk Cloud Data Platform Architecture whitepaper.

Silk's high availability architectures for SQL Server help you establish recovery targets and build a recovery plan based on how much data loss (RPO) and downtime (RTO) your environment can sustain in an outage. While Silk's data management options are customizable depending on your workload requirements, we offer four base architectures:

- **Tier 1 (Premium HA and DR):** Best for core operational applications, Tier 1 offers the highest availability and lowest recovery times using SQL Server Always On for both high availability and disaster recovery. Tier 1 allows applications to run in two regions with cross-failover capability. With database distribution between all the primary nodes, HA switchover scenarios only affect 20-30% of the production database, while DR scenarios affect no more than 50% of the database workloads.

- **Tier 2 (Critical HA and DR):** Best for most mission-critical applications, Tier 2 offers the most cost-effective solution for Always On for high availability and disaster recovery, supporting one-way replication for production applications running in one region.

- **Tier 3 (Failover and Recover):** Best for business-critical failover and recovery operations, Tier 3 only uses SQL Server for high availability, while using Silk replication for disaster recovery.

- **Tier 4 (Non-Critical DR):** Best for non-critical applications, Tier 4 is the most cost-effective solution if SQL Server Always On isn't required, using only Silk replication for disaster recovery.

Architecture	SQL Server Always On	Silk Replication	High Availability		Disaster Recovery		Fault Tolerance
			RPO	RTO	RPO	RTO	
Tier 1: Premium HA and DR	✓	✗	0	~5 seconds–1 hour	<5 minutes	Customer defined	3 failures
Tier 2: Critical HA and DR	✓	✗	0	~5 seconds–1 hour	<5 minutes	Customer defined	2 failures
Tier 3: Failover and Recover	✓ (Only for HA)	✓	0	~5 seconds–1 hour	~5 minutes–1 hour	1–4 hours	2 failures
Tier 4: Non-Critical DR	✗	✓	--	--	~5 minutes–1 hour	1–8 hours	2 failures

Tier 1: Premium HA and DR

The Tier 1 architecture provides the highest availability and the shortest recovery time for complex workloads that run applications in more than one region and require cross-region replication for HA and DR. In this architecture, SQL Server Always On replicates synchronously from your primary database instance to secondary instances across availability zones within the same region. Data is also cross replicated asynchronously from an availability zone in one region to one or more availability zones in another region or to multiple regions. The Tier 1 architecture can sustain a regional failure in the primary region (affecting multiple or all zones in that region) as well as a zonal failure in the secondary region. This solution creates at least three copies of the production database, ensuring the highest protection of your data.

High Availability: SQL Server Always On ensures zero RPO, which means your environment sustains no data loss during an outage because your data is synced across availability zones. RTO is as little as seconds with automatic recovery procedures but could take up to an hour with manual recovery, depending on your organization's recovery requirements.

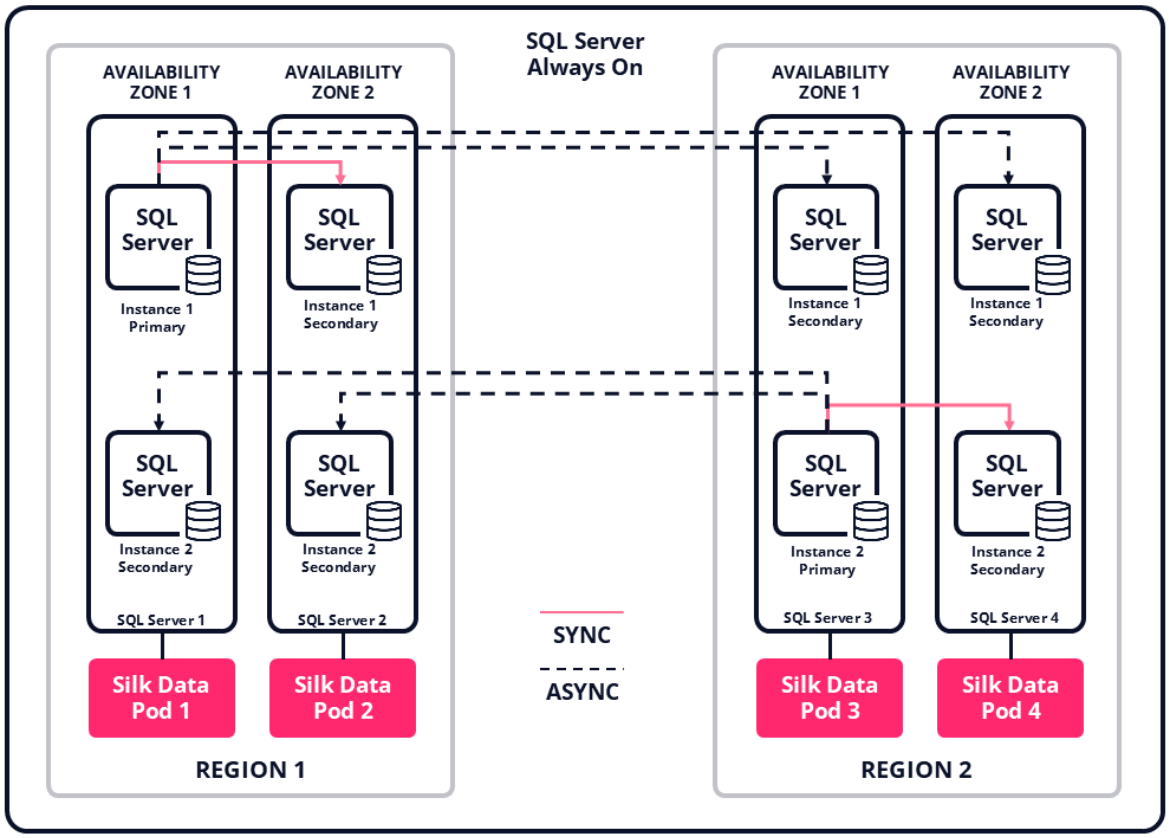
Disaster Recovery: This architecture uses SQL Server Always On for DR instead of Silk replication, so your DR targets (both RTO and RPO) are based on how quickly database replication occurs. Generally, RPO is <5 minutes. RTO is a manual process in this architecture, which means it takes longer unless automated with a third-party tool.

Tier 1 Use Case

In the example below, there are two availability zones in Region 1 and two availability zones in Region 2. Each availability zone has a SQL Server database with two instances, and each database is connected to a Silk Data Pod (SDP). There are two primary instances (**Instance 1 Primary** and **Instance 2 Primary**), each with three secondary instances for redundancy.

Data is replicated synchronously from the **Instance 1 Primary** instance to the secondary instance in Region 1, and data is replicated asynchronously from the **Instance 1 Primary** instance in Region 1 to the two secondary instances in Region 2.

Likewise, data is replicated synchronously from the **Instance 2 Primary** instance to the secondary instance in Region 2, and data is replicated asynchronously from the **Instance 2 Primary** instance in Region 2 to the two secondary instances in Region 1.



Tier 2: Critical HA and DR

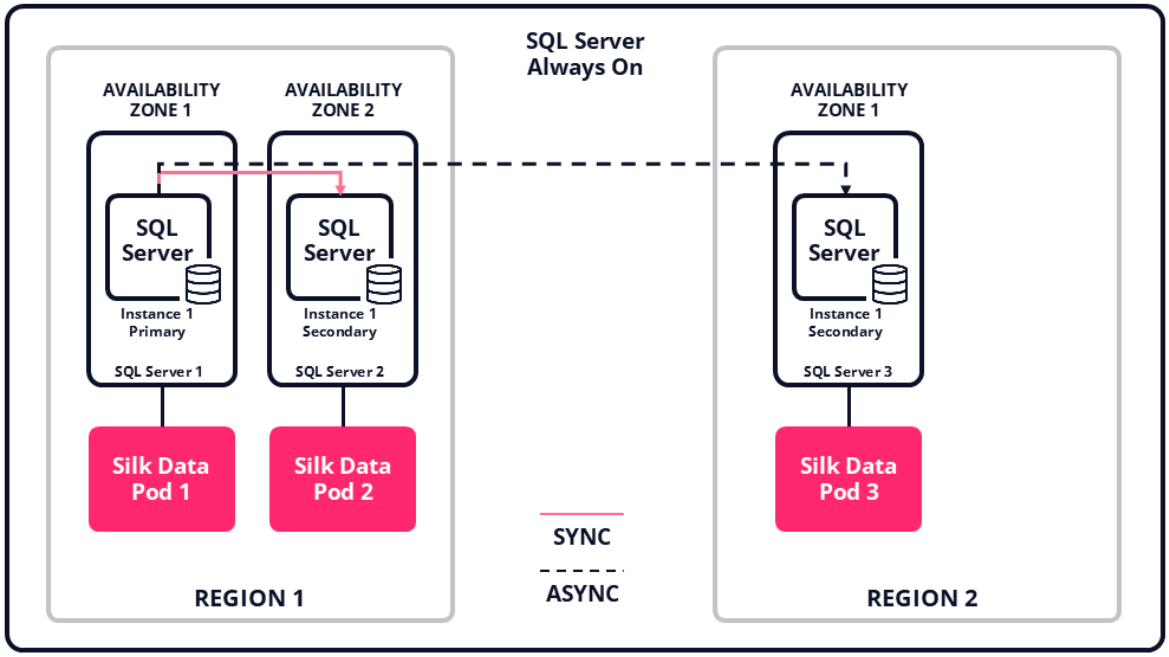
The Tier 2 architecture is the most cost-effective solution when using SQL Server Always On for high availability and disaster recovery. This architecture is designed to give mission-critical applications HA and DR solutions that have excellent SLAs for RPO and RTO requirements. Tier 2 supports one-way synchronous replication in one region and one-way asynchronous replication to a single SQL server. It offers the same high availability as the Tier 1 architecture, with the entire application stack for the primary database running in one region while the second region is used solely for disaster recovery. The Tier 2 architecture can sustain a regional failure in the primary region (affecting multiple or all zones in that region), but unlike Tier 1, it cannot sustain additional failures in the secondary region.

High Availability: Like Tier 1, SQL Server Always On ensures zero RPO with no data loss during an outage. RTO is as little as seconds with automatic recovery procedures but could take up to an hour with manual recovery, depending on your organization's recovery requirements.

Disaster Recovery: Tier 2 disaster recovery is the same as Tier 1, relying on SQL Server Always On instead of Silk replication. Recovery targets (both RTO and RPO) are based on how quickly database replication occurs. Generally, RPO is <5 minutes. RTO is a manual process in this architecture, which means it takes longer unless automated with a third-party tool.

Tier 2 Use Case

In the example below, there are two availability zones in Region 1 and one available zone in Region 2. Each availability zone has a SQL Server database with one instance, and each database is connected to an SDP. There is one primary instance and there are two secondary instances (one in each region). Data is replicated synchronously from the primary instance to the secondary instance in Region 1, and data is replicated asynchronously from the primary instance to the secondary instance in Region 2.



Tier 3: Failover and Recover

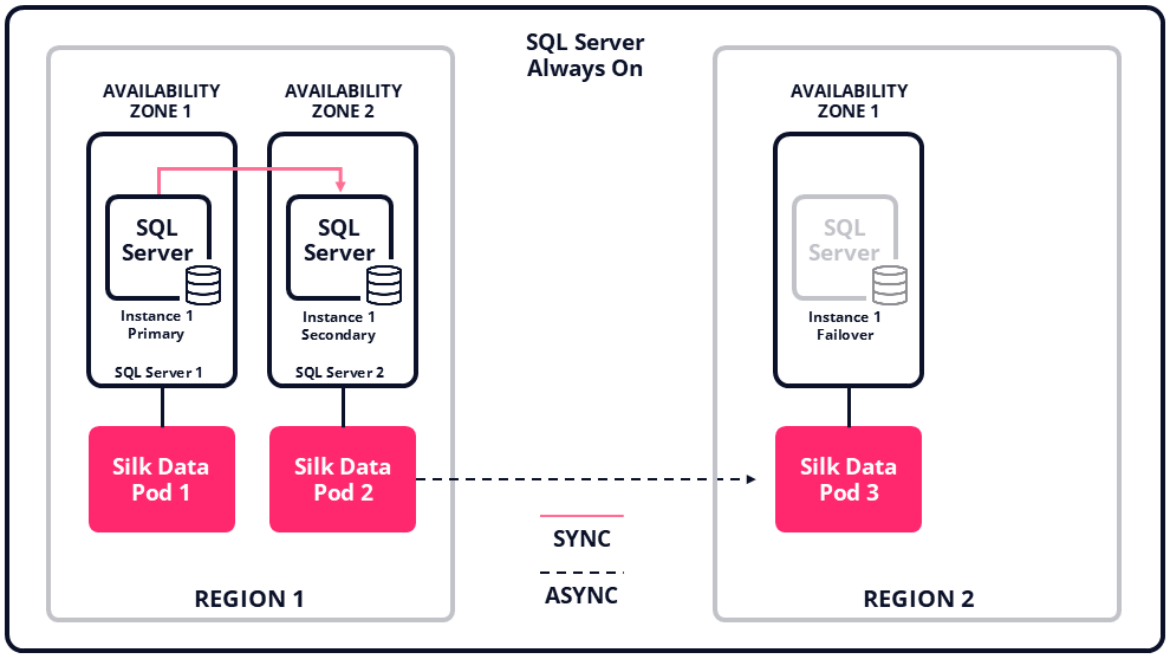
The Tier 3 architecture uses SQL Server for high availability only and uses Silk replication for disaster recovery. This provides a cost-effective solution for business-critical workloads that require minimal data loss but can sustain operational outages for a few hours. Data is replicated synchronously from one availability zone to another within the same region. Data is also replicated asynchronously from one SDP to another across regions. This architecture is best for failover and recovery, using Silk-defined SLAs and data protection to provide RTO within hours instead of your cloud provider's native backup and restore functionality, which can take weeks. If SLAs allow the use of Tier 3, it can be a very cost-effective solution for compute (VM) and DB licensing costs.

High Availability: Like Tiers 1 and 2, SQL Server Always On ensures zero RPO and no data loss during an outage. RTO is as little as seconds with automatic recovery procedures but could take up to an hour with manual recovery, depending on your organization's recovery requirements.

Disaster Recovery: In this architecture, Silk replication is enabled, and recovery SLAs are tunable based on your Silk snapshot policies. Depending on the frequency of snapshot creation and replication, RPO ranges from 5 minutes–1 hour, while RTO is typically 1–4 hours.

Tier 3 Use Case

In the example below, there are two availability zones in Region 1 and one availability zone in Region 2. Each availability zone has a SQL Server database with one instance, and each database is connected to an SDP. There is a primary instance, a secondary database, and a failover instance for disaster recovery. Data is replicated synchronously, using SQL Server Always On, from the primary instance to the secondary instance in Region 1. Data is also replicated asynchronously from an SDP in Region 1 to the SDP in Region 2 using Silk replication. In this example, Region 2 is experiencing an outage, and your failover instance is powered off. However, your data is still protected and recoverable through your SDPs, thanks to Silk replication.



Tier 4: Non-Critical DR

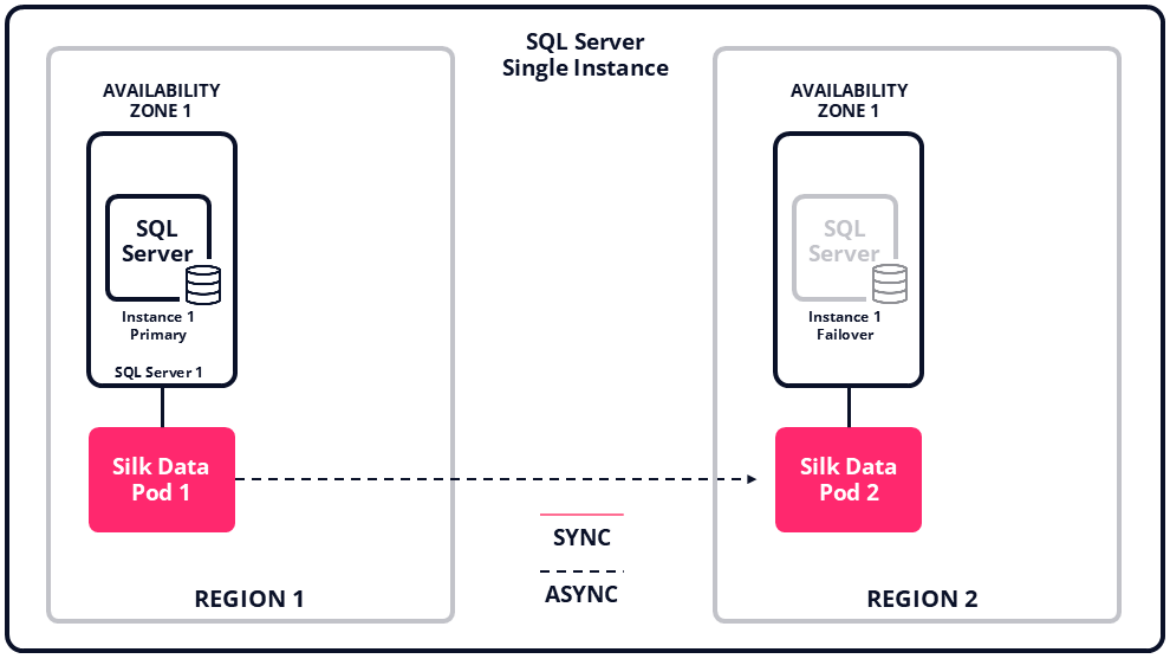
Unlike the other tiers, the Tier 4 architecture does not offer high availability with SQL Server Always On. Disaster recovery scenarios are fully managed using snapshot-based asynchronous replication within Silk. This provides the most cost-effective solution for applications that can sustain some data loss for DR and downtime to initiate a recovery. Tier 4 only supports a single SQL Server instance plus a failover instance for disaster recovery. However, you can use asynchronous replication at the database level to use SQL Server in Standby mode or leave the server powered off until it is needed for a DR event. SQL Server licensing generally allows for scenarios to fail over the DB licenses to the new host.

This architecture also uses Silk replication to replicate data asynchronously from one SDP to another across regions. Tier 4 is best for failover and recovery with no HA requirement, using Silk-defined SLAs and data protection to provide RTO within hours instead of your cloud provider's native backup and restore RTO functionality, which can take weeks.

Disaster Recovery: Like Tier 3, Silk replication is enabled in this architecture, and recovery SLAs are tunable based on your Silk snapshot policies. Depending on the frequency of snapshot creation and replication, RPO ranges from 5 minutes–1 hour, while RTO is typically 1–4 hours.

Tier 4 Use Case

In the example below, there is one availability zone in Region 1 and one availability zone in Region 2. The availability zone in Region 1 has a SQL Server database with a primary instance, and the availability zone in Region 2 has a failover instance. Each instance is connected to an SDP. Data is replicated asynchronously from the SDP in Region 1 to the SDP in Region 2 using Silk replication. Like the Tier 3 example, this example shows an outage in your failover instance, but your data is still protected and recoverable through your SDPs since they're using Silk replication.



Conclusion

Silk works with SQL Server Always On to support your organization's high availability and disaster recovery SLAs. We offer four tiers of support to accommodate your RPO and RTO requirements in the most cost-effective way possible:

- **Tier 1: Premium HA and DR** provides the most coverage for HA and DR plans, supporting cross-failover across two regions and recovery times of just minutes.
- **Tier 2: Critical HA and DR** is the most cost-effective use of SQL Server Always On for HA and DR, supporting one-way replication in the primary region and up to two failures while still achieving fast recovery times.
- **Tier 3: Failover and Recover** uses SQL Server Always On for HA while using Silk replication for DR, offering a practical solution for compute and database licensing needs.
- **Tier 4: Non-Critical DR** does **not** use SQL Server Always On, offering a highly cost-efficient solution for non-critical applications by providing only Silk replication for DR.

These HA/DR solutions provide a scalable layer of protection in your resilience strategy, allowing you to add more layers as you build a robust, fault-tolerant system.