



White Paper

Deploying a SQL Database on Azure with the Silk Data Virtualization Platform

How To Guide
Version 1
March 2023

Table of Contents

- 1. Overview 3
- 2. Purpose of this How To Guide 3
- 3. Abbreviations 3
- 4. High Level Architecture Overview of Silk 4
 - 4.1 VisionOS 4
 - 4.2 Flex 4
 - 4.3 Clarity 4
- 5. SQL Databases and Silk 4
- 6. Steps Required to deploy Silk onto Azure. 5
 - 6.1 Detailed Steps required to deploy Silk on SQL 5
- 7. Choose a VM Shape and Image 5
 - 7.1 Choosing the VM image 5
 - 7.2 Choosing a VM shape 5
 - 7.3 An example of a VM shape 5
- 8. Apply Silk & SQL best practices to VM 6
 - 8.1 VM Configuration for Silk 6
 - 8.2 Azure SQL Licensing Considerations 6
- 9. Create, Connect, and ConfigureConnect Silk Volumes 6
 - 9.1 SQL Volume Filesystems Silk Volume group recommendations. 6
 - 9.2 Silk Volume Geometry 7
 - 9.3 Silk Data Reduction 7
 - 9.4 Disk Partition Type. 7
 - 9.5 NTFS Format Options 7
 - 9.5.1 Example: Power Shell (PS) Command Prompt. 7
 - 9.6 Compression and Data Reduction and Encryption 8
 - 9.7 Volume expansion 8
 - 9.8 Silk Networking Configuration 8
- 10. Options for replicating and backing up your SQL environment on Silk 8
- 11. Additional Resources 8

1. Overview

As more organizations move their data to the cloud, Silk is invested in ensuring that its Platform makes it easy to migrate and maintain data on the public cloud.

For organizations looking to leverage the public cloud, Silk offers the flexibility and high performance that users have come to expect, through its cloud agnostic platform.

With Silk, customers can dynamically compose data resources based on application requirements: allocating just the right amount of both performance and capacity while minimizing overprovisioning and waste. Furthermore, Silk offers a disaggregated, dynamic cloud architecture. Customers can scale capacity or performance resources independent of one another, thus delivering further flexibility to the business.

Silk's Cloud Data Virtualization Platform is a combination of tested and packaged software and services. The software stack is built on the software components of VisionOSTM, Clarity, and Flex which provide a rich set of data services, machine learning, analytics, and policy- based automation and orchestration.

2. Purpose of this How To Guide

This guide describes how to integrate the Microsoft Azure Database with the Silk Cloud Data Virtualization Platform. This document is for Database Administrators (DBAs), architects and database managers deploying SQL on Azure IaaS solution with the Silk Cloud Platform. This document provides the steps required to build an SQL on Silk IaaS deployment on Azure for SQL.

3. Abbreviations

Abbreviation	Definition	Notes
AIOps	Artificial Intelligence for IT Operations	
API	Application Programming Interface	
CPU	Central Processing Unit	
DB	Database	
DBA	Database Administrator	
DNS	Domain Name System	
DR	Disaster Recovery	
GPT	GUID partition table	
GUID	Global Unique Identifier	
I/O	Input / Output	
IaaS	Infrastructure as a Service	
iSCSI	Internet Small Computer System Interface	
K-RAID	RAID proprietary to Silk	See RAID
KB	Kilobytes	
MBPS	Megabytes per second	
MBR	Master Boot Record	
NIC	Network Interface Card	
NTFS	New Technology File System	
OS	Operating System	
PS	Power Shell command prompt	
RAID	Redundant Array of Independent Disks	See K-RAID
REST API	Representational State Transfer	REST API is the API that conforms to the REST architectural style
SDP	Silk Data Pods	
SQL	Structured Query Language	
vCPU	virtual Central Processing Unit	
VG	Volume Group	
VM	Virtual Machine	

4. High Level Architecture Overview of Silk

A high level overview of the Silk architecture is described below.

4.1 VisionOS

Vision OS is a data hypervisor for resource virtualization and helps customers accelerate their cloud adoption journey. Running on any cloud, VisionOS delivers enterprise-class data services with a highly flexible symmetric active-active scale-out shared data framework. VisionOS turns any underlying cloud infrastructure into the world's most capable high performance scale-out data virtualization and mobility platform.

VisionOS architecture is comprised of the following software elements:

1. c.node (compute/controller node) - Virtualization for data management and optimization
2. d.node (data node) - Virtualization of a data endpoint
3. m.node (media node) - Virtualization of data nodes management protected under K-RAID

4.2 Flex

Flex is a data orchestration platform for managing resources in the public cloud. Flex delivers an on-demand ability to compose, optimize, manage, and decommission resources as needed to support application SLAs from within the Silk Cloud Platform. Flex orchestrates on-premises and cloud resources as new workloads emerge, move, and evolve over time.

4.3 Clarity

Clarity is a cloud-based Artificial Intelligence for IT Operations (AIOps) engine that delivers predictive analytics through a comprehensive set of management and monitoring functionality which analyze millions of data points from our install base. This includes the unique capability to leverage application-level intelligence, machine learning, and big data predictive analytics.

Clarity is designed to maximize business agility while minimizing risk. Clarity's advanced analytics provide recommendations on self-healing policies and preemptive resource optimization while enabling a whole new automated approach to managing business and application SLAs.

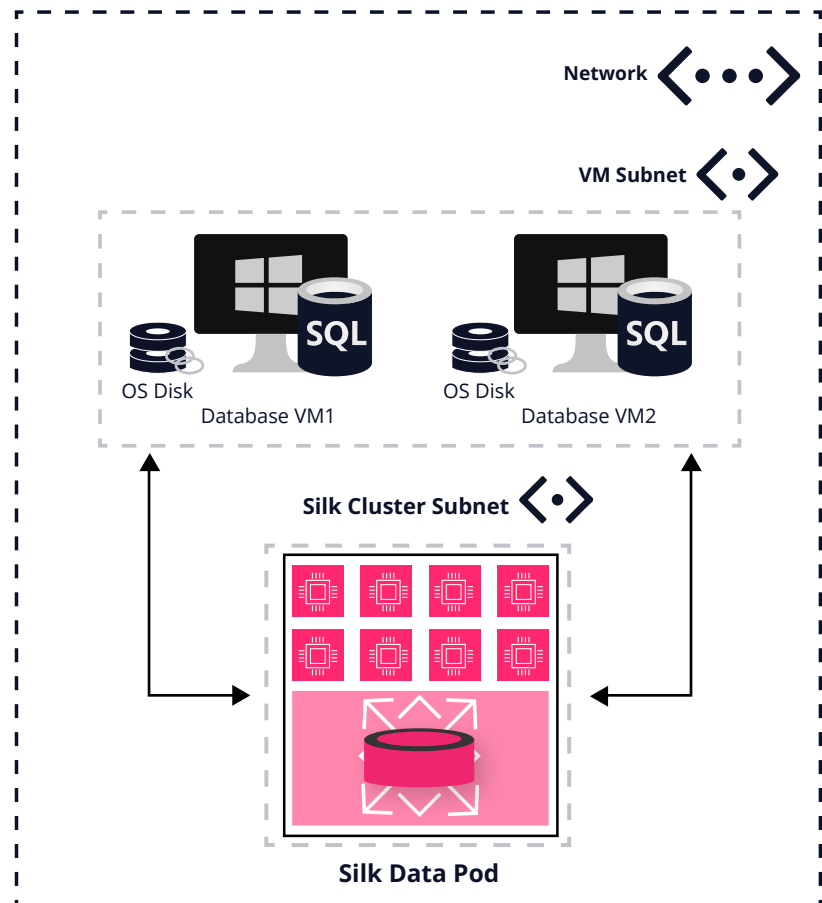
5. SQL Databases and Silk

SQL databases are among the highest performing, largest capacity databases both on premises and in the cloud. SQL databases requires large amounts of I/O bandwidth, low latency, and high I/O rates for both read and write operations.

These demands are where Silk comes in, by providing a robust high-performance invisible layer of storage and database services (Figure 5.1).

Users do not need to modify the SQL stack for successful deployment. The Silk Platform recognizes SQL databases as one or more standard cloud volumes. SQL admins can perform all tuning, administration and life cycle management activities as normal.

Figure 5.1: High Level Architecture Drawing of the Microsoft Azure Database running on the Silk Platform.



6. Steps Required to deploy Silk onto Azure

An overview of the steps required to deploy Silk onto your Microsoft Azure database is as follows:

1. Choose a Virtual Machine (VM) shape and Image (Section 7)
2. Apply Silk & SQL Best practices to VM (Section 8)
3. Create, Connect, and Configure Silk Volumes (Sections 9 to 16)

6.1 Detailed Steps required to deploy Silk on SQL

On the Silk Data Pod (SDP) – See Silk Cloud Platform User Guide for details [here](#):

1. Create the SQL host.
2. Create the Volume Group (VG) for the SQL host.
3. Create volumes (for data, log etc.).
4. Keep all volumes in the same VG to leverage snapshot and replication features.
5. Map volumes to host.

7. Choose a VM Shape and Image

7.1 Choosing the VM image

The Microsoft Azure Marketplace offers images containing SQL database software. Alternatively, you can choose a standard image and install SQL separately. One advantage of running SQL on Azure deployed on Silk is that you can choose the flavor of Windows or Linux and version of SQL to match your internal business requirements.

You can find further details about SQL specific VM images on Azure [here](#).

7.2 Choosing a VM shape

Silk works differently from standard cloud storage options. Silk's architecture bypasses the I/O restrictions that apply to standard cloud storage options, enabling users to achieve higher levels of I/O performance. You should choose a VM shape depending on the required CPU and memory characteristics. The peak write I/O bandwidth just needs to be lower than the network bandwidth limit for the VM shape.

7.3 An example of a VM shape

A database requires 3,000 MBPS, but only needs 32 vCPU cores to successfully run its workload. With standard Azure storage I/O capabilities, you would have to use at least a 96 vCPU VM shape to reach that level of I/O capacity. This results in overprovisioning of both infrastructure and licensing, which dramatically increases SQL licensing costs. By using the Silk Platform to provide data I/O, you can easily provision enough throughput using only a 32 vCPU VM (Table 7.1).

Table 7.1: Comparison of provisioning using cloud native storage limits versus Silk

VM-Shape	Memory	Cores	Cloud Native storage limits (uncached)	Network BW limit	Effective max Silk Limit per vm
Standard_E32s_v5	256GB	32	<ul style="list-style-type: none">• 51,200 IOPS• 865 Mbytes/s combined	<ul style="list-style-type: none">• 2,000 Mbytes/s• 2,000 Mbytes/s egress	<ul style="list-style-type: none">• 2,000 Mbytes/s write• 5,500 Mbytes/s read• IOPS 400k per c.node
Standard_E64s_v5	512GB	64	<ul style="list-style-type: none">• 80,000 IOPS• 1,735 Mbytes/s combined	<ul style="list-style-type: none">• 3,750 Mbytes/s• 3,750 Mbytes/s egress	<ul style="list-style-type: none">• 2,000 Mbytes/s write• 5,500 Mbytes/s read• IOPS 400k per c.node
Standard_E104s_v5	672GB	104	<ul style="list-style-type: none">• 120,000 IOPS• 4,000 Mbytes/s combined	<ul style="list-style-type: none">• 12,500 Mbytes/s• 12,500 Mbytes/s egress	<ul style="list-style-type: none">• 12,500 Mbytes/s write• 12,500 Mbytes/s read• IOPS 400k per c.node

You can find full details of some suitable available VM shapes [here](#).

8. Apply Silk & SQL best practices to VM

8.1 VM Configuration for Silk

To connect an IaaS VM to a Silk Pod, you must follow the steps detailed in the Silk Configuration Guide, available [here](#).

8.2 Azure SQL Licensing Considerations

Silk recommends using the SQL server IaaS Agent extension to integrate your SQL on VMs to the Azure portal. Full details on this functionality and how to ensure its operation can be found [here](#).

9. Create, Connect, and Configure Connect Silk Volumes

9.1 SQL Volume Filesystems Silk Volume group recommendations

In general Silk SDPs do not add any further configuration requirements surrounding volumes used for SQL. Silk Systems can deliver full performance of an SDP from a single Volume. Single disk and single volume configurations will provide maximum performance on Silk SDP. Users no longer need to overprovision disk sizing to meet specific performance goals. A 200GB and 20TB volume on Silk SDP provides the same performance.

Users who previously used any volume aggregation to achieve performance goals (such as Storage Spaces or volume raid configurations in disk manager), may no longer need to implement these to maximize performance.

However, there can be benefits to both the OS performance and the management/reporting of performance to separate the database files of a large database on to separate volumes. Refer to **Figure 9.1.1** and **Figure 9.2.1** for an example of a typical large individual DB typical volume layout.

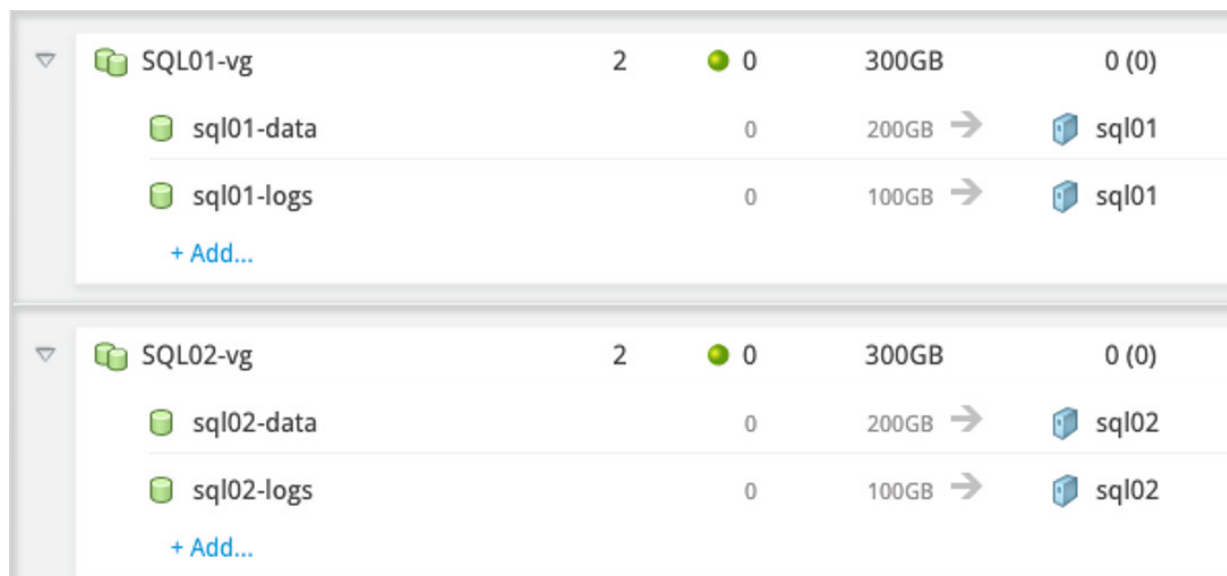


Figure 9.1.1: Sample Silk Volume.

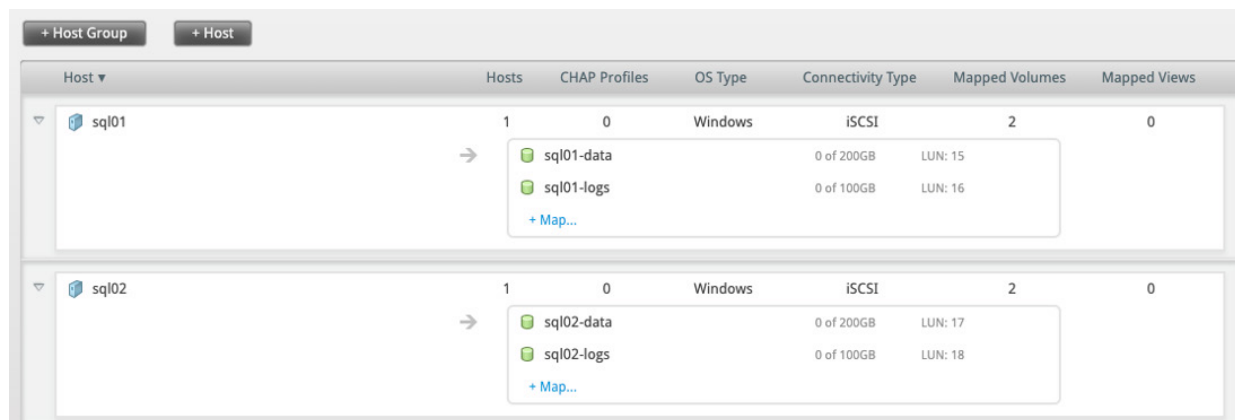


Figure 9.1.2: Sample Volume Group configuration.

9.2 Silk Volume Geometry

The Silk Cloud Data Virtualization Platform provides cloud resources as virtual volumes. Silk uses a Write Aligned, variable block size architecture. As a result, there is no advantage to Silk volumes of block size alignment or a larger logical block size. Silk always presents volumes using 512-bytes sectors as this ensures maximum compatibility and minimizes wastage. This is not a configurable option.

9.3 Silk Data Reduction

The Silk Platform will attempt to reduce the amount of physical space required to host databases to a minimum. Silk features technologies such as deduplication, compression, and thin provisioning, redirect on write, which help to minimize the cost of your cloud infrastructure. Both Silk and Microsoft support these technologies for SQL databases.

Data deduplication is an optional feature from Silk that can be enabled per Volume Group (VG). Disabling this feature can help to reduce the CPU load on Silk c.nodes. For Databases containing pre-encrypted or compressed data it may make sense to disable the deduplication Silk feature. Users can non-disruptively change this setting by moving volumes between VGs after deployment, after confirming the quantity of unique data that the Silk Data Pod is storing.

9.4 Disk Partition Type

By default, Windows should use a GUID partition table (GPT) partition type on all Silk volumes. Silk recommends that you use this default setting. Some organizations used the old Master Boot Record (MBR) type for 32-bit Oses and removable drives such as memory cards.

9.5 NTFS Format Options

Silk makes no changes to the recommended format options for Microsoft SQL. The unit of storage in SQL Server is a page, which is 8 KB. Eight contiguous pages make up an extent (which is 64 KB in size). SQL Server uses extents to store data. Therefore, on a SQL Server machine, the New Technology File System (NTFS) allocation unit size for hosting all SQL database files should be set to 64 KB.

9.5.1 Example: Power Shell (PS) Command Prompt

```
PS C:\> Format-Volume -DriveLetter D -FileSystem NTFS -AllocationUnitSize 64k
```

The following pop-up window is displayed (**Figure 9.5.1**).

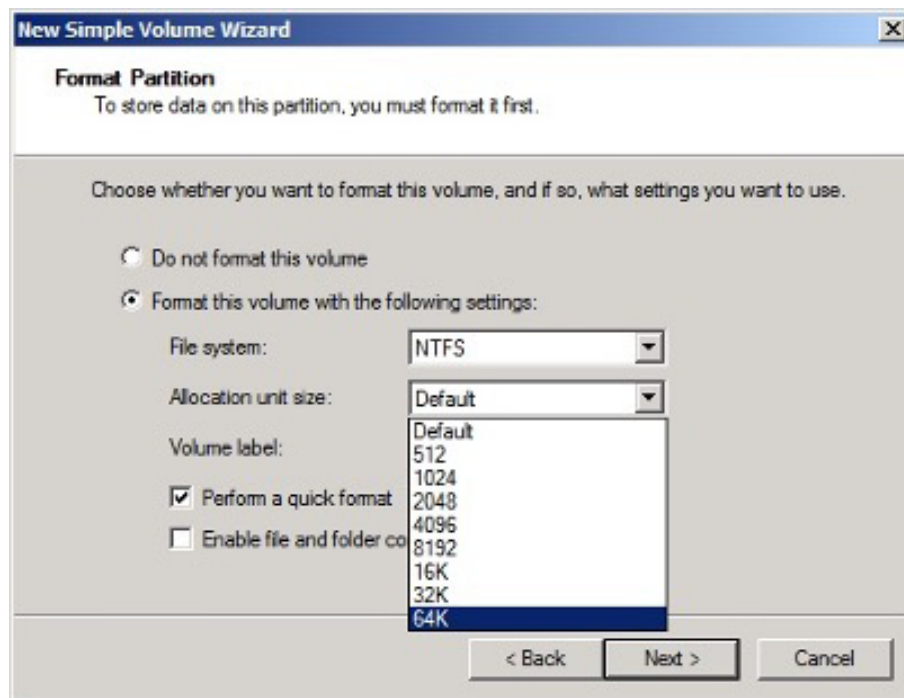


Figure 9.5.1: New Simple Volume Wizard pop up window appears after entering the command prompt.

9.6 Compression and Data Reduction and Encryption

Cloud service providers encrypt the data prior to placing the data onto physical media. In most cases, users do not find it productive to compress or encrypt data on a database host prior to sending to the Silk Data Pod.

9.7 Volume expansion

Users can add additional space to an SQL DB simply by expanding a volume from one of the Silk UIs then also expanding the NTFS Filesystem within Windows.

9.8 Silk Networking Configuration

The Silk platform supports a wide range of networking configurations. However, for maximum performance, reporting, reliability, and manageability, Silk recommends the following configuration settings.

Utilize a dedicated additional NIC for data traffic to Silk on the database host(s) as follows:

1. Attach this Network Interface Card (NIC) to a subnet just for iSCSI data traffic.
2. Enable traffic to route from this subnet to the Silk SDP data 1 subnet and lock access to only iSCSI traffic.
3. Open the following port: 3260, between the Silk data subnet and the host data subnet.

Where management or snapshot integration operations need to originate from the DB host, REST API traffic should be enabled to flow on port 443 and from the host that needs to perform REST API queries or commands.

10. Options for replicating and backing up your SQL environment on Silk

Silk's patented snapshot technology allows you to create zero-footprint clones without the need to provision another virtual machine. You can use snapshots to accelerate backup methodologies and integrate directly with major backup vendors like CommVault and Netbackup.

Silk recommends the SQL always-on database replication technology where zero data loss replication, or automatic failover between zones or regions are required.

You can also asynchronously replicate Silk's snapshots to another Region or zone, providing Disaster Recovery (DR) options as alternatives to SQL Always-On. Silk-based replication can help reduce SQL licensing costs and limit the need to keep the DR SQL instance online.

11. Additional Resources

For more information on how Silk can optimize your SQL database experience in the cloud, visit silk.us.

About Silk

The Silk Data Virtualization Platform gives demanding workloads 10x faster performance on the cloud compared to native cloud alone. It is a virtualization layer that sits between the underlying cloud infrastructure and customers' workloads. Without refactoring, workloads such as Oracle, Microsoft SQL Server, and industry-specific applications can move onto public cloud while massively improving user experience. Industry leaders in e-commerce, SaaS, FinTech, and healthcare trust Silk with their business-critical workloads to get the ultra-fast speeds their customer demand. Silk is headquartered outside of Boston, MA.

To learn more, visit silk.us.