

Hybrid and Hyperscale Cloud with SQL Server 2016

Technical White Paper

Published: April 2016

Applies to: Microsoft SQL Server 2016 and Microsoft Azure

Summary: With cloud computing comes a new paradigm shift as organizations continue to realize the potential cost benefits of running their database applications in the cloud with greater scale and flexibility. Microsoft SQL Server 2016 is built for the cloud first, but not cloud only—organizations can easily deploy SQL Server in a private cloud, hybrid cloud, or public cloud, and can use familiar tools for development and management. Scalability features such as Stretch Database, backup to Microsoft Azure, and availability groups span across on-premises, private, and public cloud infrastructure. The ability to seamlessly cross these logical boundaries and integrate operations with data and services from virtually anywhere is one of the most exciting and impactful features of SQL Server, SQL Server 2016, and Microsoft Azure.

Copyright

The information contained in this document represents the current view of Microsoft Corporation on the issues discussed as of the date of publication. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information presented after the date of publication.

This white paper is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, AS TO THE INFORMATION IN THIS DOCUMENT.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in, or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Microsoft Corporation.

Microsoft may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Microsoft, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

© 2016 Microsoft Corporation. All rights reserved.

Microsoft, Active Directory, Microsoft Azure, Excel, SharePoint, SQL Server, Windows, and Windows Server are trademarks of the Microsoft group of companies.

All other trademarks are property of their respective owners.

Contents

- Microsoft SQL Server evolution 4**
- Hyperscale with SQL Server.....6**
 - Hybrid solutions.....6**
 - Traditional non-virtualized deployment 6
 - Private cloud deployment.....7
 - Public cloud deployment 7
 - Hybrid deployment.....7
 - Consistent platform8**
 - Common development, management, and identity tools 9
 - SQL Server and Microsoft Azure 10
 - Hyperscale features in SQL Server.....12**
 - Stretch Database 12
 - Enhanced backup to Azure 15
 - High availability and disaster recovery 16
- Conclusion..... 17**
- More information..... 18**

Microsoft SQL Server evolution

For many years, Microsoft SQL Server has moved toward easier integration with outside data sources while simplifying management and administration. With the recent development of Microsoft cloud services and the shift toward larger and more varied data sources, the cloud's capabilities have grown dramatically. Since SQL Server 2012, cloud strategy has been a central focus of development. SQL Server 2014 added significant new cloud features, and these features have been enhanced with SQL Server 2016. *Hybrid cloud* is a term that resonates with SQL Server developers, administrators, and business leaders because of the benefits that hybrid clouds deliver to data management. Analysts have also recognized the dedication of SQL Server to cloud development and its ability to execute on that vision, as illustrated by the Gartner Magic Quadrant for infrastructure as a service (Figure 1).



Figure 1: Gartner Magic Quadrant for infrastructure as a service¹

¹ Gartner, <https://aws.amazon.com/resources/gartner-2015-mq-learn-more/>, May 2015.

In addition, SQL Server has consistently added innovative features over the last 15 years (Figure 2).

The evolution of SQL Server continues...

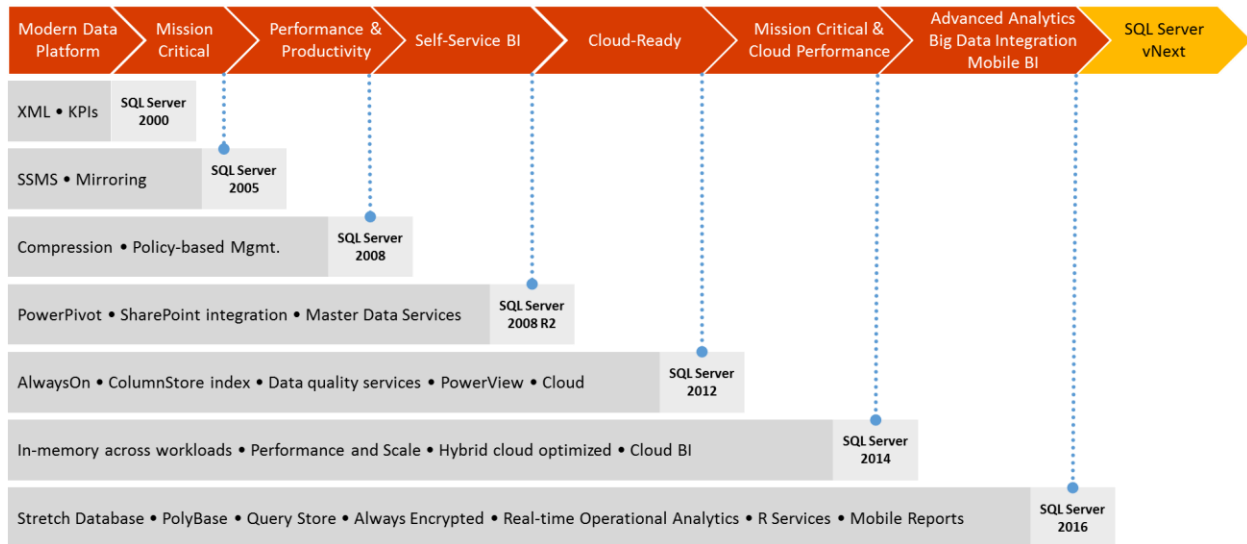


Figure 2: Major SQL Server functionalities added across releases

SQL Server 2016 adds the following to this existing functionality:

- Stretch Database for transparently stretching warm and cold online transaction processing (OLTP) data to Microsoft Azure on demand without application changes
- Faster hybrid backups, high availability, and disaster recovery scenarios to back up and restore your on-premises databases to Microsoft Azure
- Enhanced availability with cloud replicas
- Enhanced consistency and experience in working with SQL Server on-premises, SQL Server on Azure Virtual Machines, and Azure SQL Database

Hyperscale with SQL Server

Hyperscale is a term coined to describe the ability of a system to scale in response to demand. In the realm of data management, this entails scalable and distributed data in addition to highly scalable computing. This scalability is the core for enabling cloud-based data, big data, or distributed storage.

SQL Server 2016 is being developed as “cloud first, but not cloud only,” meaning on-premises and cloud products are designed to work together in a hybrid environment. New tools in SQL Server and Microsoft Azure make it even easier to scale to the cloud; to build patching, backup, and disaster recovery solutions; and to access resources wherever they are—on-premises, in a private cloud, or in a public cloud. These tools provide an easy on-ramp to the cloud for on-premises SQL Server databases, enabling customers to use their existing skills to take advantage of Microsoft global datacenters. The architecture is designed to facilitate hybrid solutions while also maintaining simplicity in common tasks and providing a set of consistent tools that work in a similar manner whether operating on-premises or in the cloud.

Hybrid solutions

In almost any application scenario, from the largest public website to the smallest departmental application, the vast majority of applications rely on some sort of database management system. The relational database is so ubiquitous that many developers and IT professionals have become desensitized to it, considering it as simply part of the stack for a modern application. As organizations look to take advantage of cloud computing, the availability of cloud-enabled database systems will be critical to their success.

This white paper explains the Microsoft vision of relational database management systems in the context of cloud computing. It is a hybrid IT vision that gains leverage from the industry-standard SQL Server technology set and makes it available across the spectrum of deployment approaches that organizations adopt today (Figure 3).

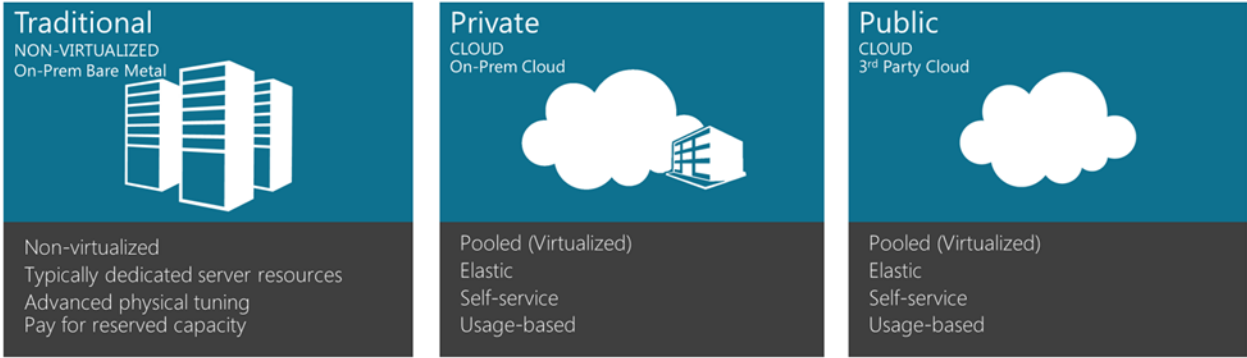


Figure 3: A combination of on-premises and cloud solutions to help modern IT departments meet business needs

Traditional non-virtualized deployment

Despite massive improvements in virtualization technology in the past 10 years, there is still a significant performance penalty for virtualizing certain workloads. Large, complex, and mission-critical OLTP systems

remain the province of dedicated servers that have the operating system and database platform installed directly. Traditionally deployments are often chosen for complex, mission-critical scenarios where every bit of extra performance is needed.

Private cloud deployment

While a private cloud *may* have all the characteristics of a public cloud, that does not necessarily mean it *must* have all of those characteristics. For example, many private clouds do not implement a full chargeback accounting mechanism. Nevertheless, as organizations mature their private cloud strategy, the service and service levels offered by private clouds begin to align more closely with those offered by public cloud providers. With a private cloud deployment, organizations can enable flexibility with a virtualized environment while maintaining ownership of hardware and data. Private cloud deployments consolidate compute power and storage, yielding better utilization and cost efficiency.

Public cloud deployment

Public cloud deployments offer several characteristics that are attractive to organizations with constantly growing infrastructure needs. The scalability and elasticity of a public cloud infrastructure allow for dynamic growth (and reduction) to meet demands quickly—in a matter of minutes. With pay-for-use pricing models offered by infrastructure as a service (IaaS) and platform as a service (PaaS), the total cost of ownership is often significantly lower than an on-premises deployment. Consumers of pay-for-use services are charged only for usage, and unused compute and storage resources do not contribute to operating costs.

Hybrid deployment

The concept of hybrid cloud recognizes that organizations typically have a portfolio of different applications deployed across their business and a breadth of environments with unique requirements. Some applications require detailed and complex hardware configurations that defy deployment into the type of commoditized, “one-size-fits-all” environment offered by cloud computing. Equally, workloads in many businesses are extremely compelling for massive-scale public clouds—it can be economically infeasible to allocate sufficient levels of hardware for applications that experience massive peaks and troughs in demand. The Microsoft goal for hybrid cloud is to offer organizations breadth of choice in how and where they run their applications, while at the same ensuring they can use a common set of server products, tools, and expertise across a portfolio of solutions (Figure 4).








	Traditional NON-VIRTUALIZED	Private CLOUD	Public CLOUD
Database Scenarios	Large, Mission-Critical OLTP Data Warehousing Big Data	OLTP BI Workloads Datamarts	OLTP BI Workloads Big Data
Rationale for Usage	 Compliance  Complex, Mission-Critical Applications	 Compliance  Economics Consolidation Cost Allocation	 Rapid Development  Economies of Scale Pay-for-Use  Scalability & Elasticity

Figure 4: The unique benefits and challenges of database deployment options, as organizations increasingly move to the cloud

Consistent platform

The foundations of common Microsoft development, management, and identity tools are Windows Server and Microsoft Azure. The full breadth of Microsoft technology solutions complement these, including Microsoft SQL Server, Microsoft System Center, Microsoft Active Directory, and Microsoft Visual Studio. Together, these technologies provide one consistent platform for infrastructure, applications, and data that can span your datacenter, service provider datacenters, and the Microsoft public cloud (Figure 5).

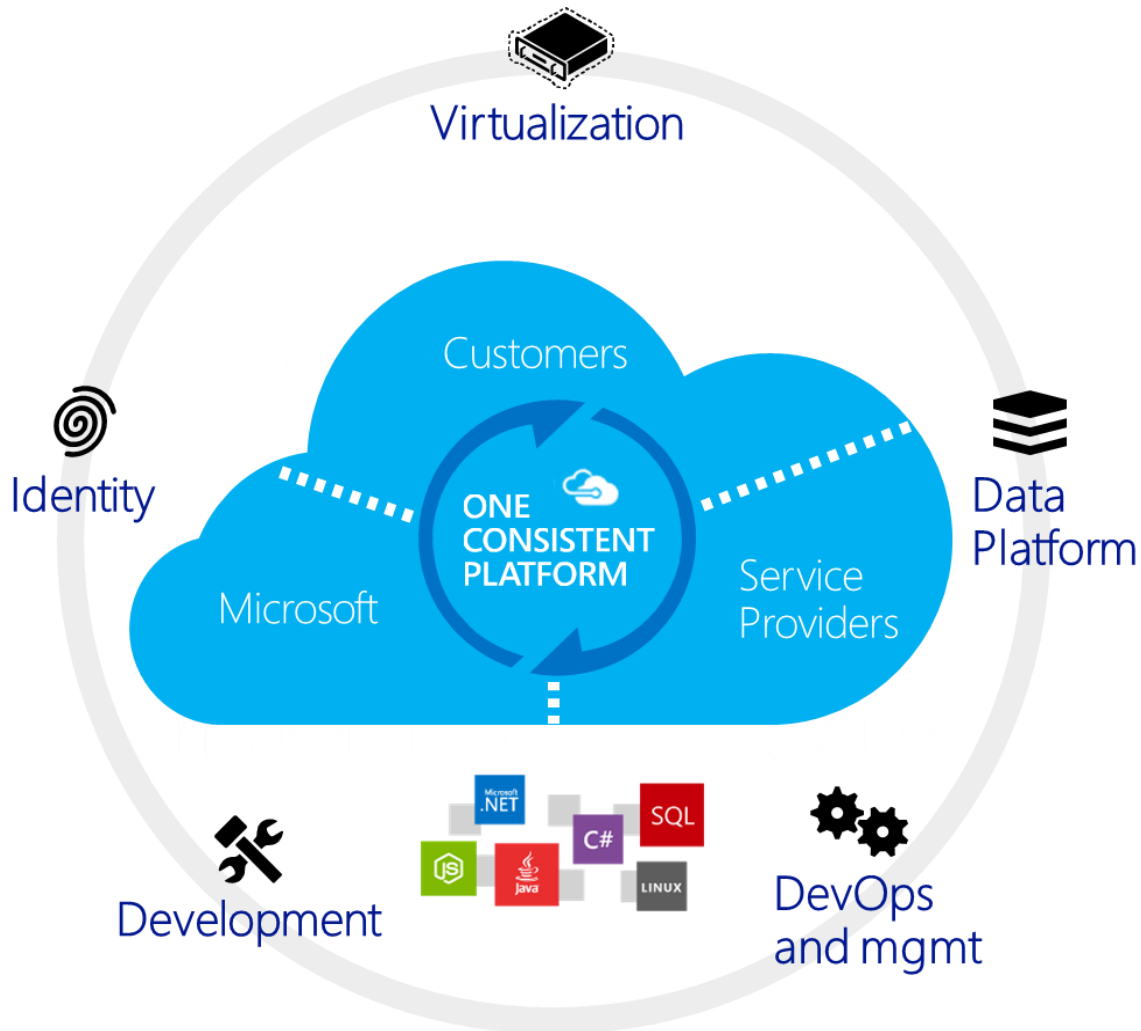


Figure 5: Consistent and integrated platform

Common development, management, and identity tools

This consistent platform removes complexity so enterprises can focus on identifying business needs and determining how applications can address them. This vision of consistency can help enterprises be flexible in thinking about the best ways to deploy applications across the three clouds. The main components of this platform are integrated virtualization, a complete data platform, unified management and DevOps, a flexible development paradigm, and a common identity.

Integrated virtualization

The integrated and portable virtualization built into Windows Server allows your team to virtualize servers as well as the network, storage, and applications across clouds. With integrated virtualization, enterprises can generate compute capability for software-defined networking and virtualization of storage, regardless of where the virtual machines reside.

Complete data platform

A complete data platform powered by SQL Server and Azure SQL Database allows you to manage petabytes of data, power mission-critical applications, and give business users BI solutions with a range of tools—from Microsoft Excel to Apache Hadoop.

Unified management

With the power of unified management, enterprises can employ tools like Microsoft System Center to administer all of the automation scripts used to manage and pool resources in a datacenter. For example, you can provision and move virtual machines, objects, or data from the datacenter through a service provider or to Azure for bottomless storage, redundancy, and disaster recovery.

Flexible development

Flexible development allows your organization's developers to use their choice of tools, languages (Microsoft or open source), and open standards to quickly build applications, connect them with other apps and data, and then deploy on-premises, in the cloud, or in a hybrid model. Microsoft Visual Studio and Microsoft Team Foundation Server enable application lifecycle management—from the idea through the deployment of an app.

Common identity

The Microsoft vision for a consistent hybrid cloud platform gives enterprises the ability to manage and connect all of their applications seamlessly—regardless of where those applications are hosted. Active Directory and Microsoft Azure Active Directory provide a powerful base for single identity across clouds to securely extend applications to people and their devices. For example, an enterprise can leverage its existing Active Directory group policies with Active Directory Federation Services and Azure Active Directory so group policies will extend to the cloud automatically.

SQL Server and Microsoft Azure

The Microsoft data platform leverages SQL Server technology and makes it available across physical on-premises machines, private clouds, private clouds hosted by a third party, and the public cloud. This enables you to meet unique and diverse business needs through a combination of on-premises and cloud-hosted deployments while using the same set of server products, development tools, and expertise across these environments.

Each offering can be characterized by the level of administration you have over the infrastructure and by the degree of cost efficiency achieved by database-level consolidation and automation (Figure 6).

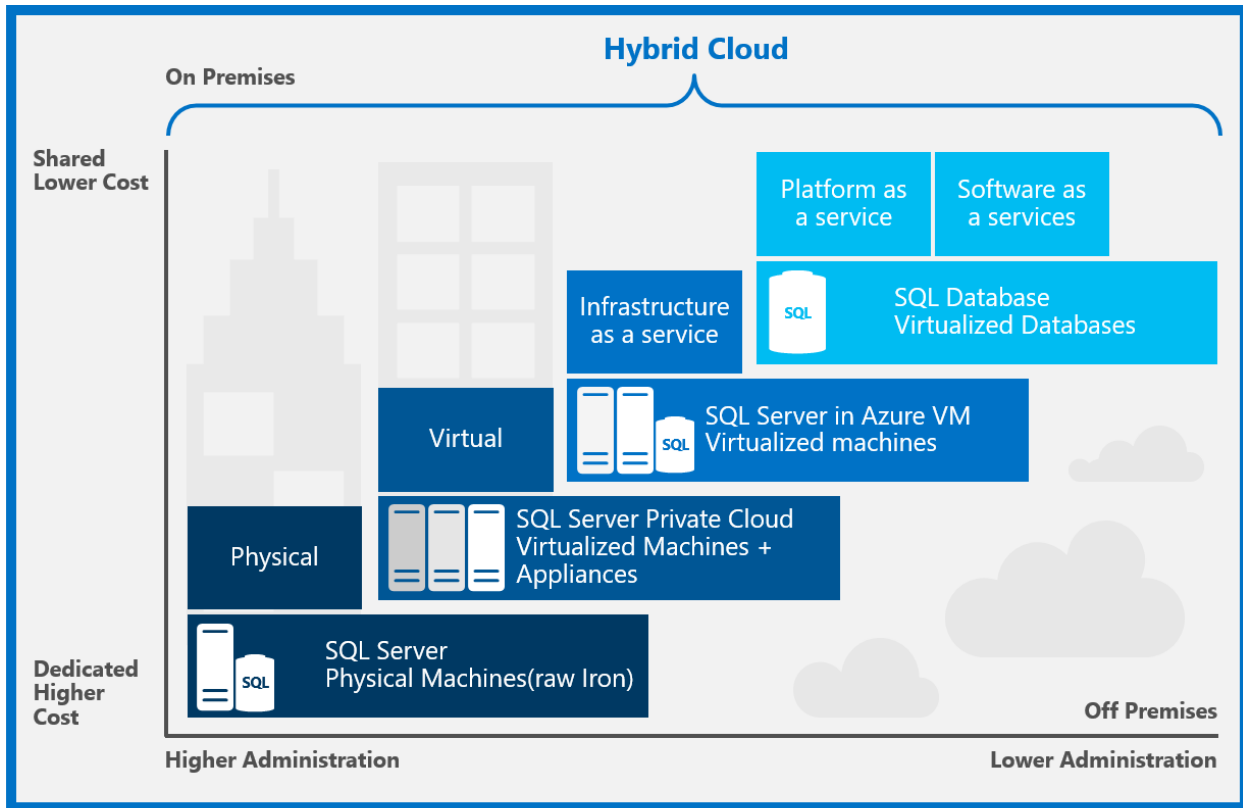


Figure 6: SQL Server technology, from on-premises to the public cloud

When designing an application, you have four basic options for hosting the SQL Server portion:

- SQL Server in non-virtualized physical machines
- SQL Server in on-premises virtualized machines (private cloud)
- SQL Server in Azure Virtual Machines (public cloud)
- Azure SQL Database (public cloud)

SQL Server 2016 creates new functionality and enhances existing services to facilitate moving to a hybrid cloud. Scalability, availability, security, identity, backup and restore, and replication all have enhancements for working with SQL Server in Azure Virtual Machines, Azure SQL Database, or other Azure services.

In addition, a plethora of new data sources available (many residing natively in the cloud) means that integration, processing, and analytics are also impacted by enhancements in SQL Server 2016. Analysis Services, Reporting Services, and Integration Services all have additional features that make working with cloud data and hybrid solutions easier. SQL Server 2016 enhances the consistency of your development and administration efforts by making it much easier to work within a hybrid cloud environment.

Hyperscale features in SQL Server

Stretch Database

Data is continuously growing, and users generally want to retain all data for many possible reasons, including the following:

- Regulatory compliance (for example, taxes)
- Auditing (for example, fraud investigation)
- Planning (for example, comparing past results)
- Nature of business (for example, retailer transaction details history)
- Inability to determine what can be safely deleted (for example, what might a government agency or major institutional investor ask for?)

Traditional archiving solutions typically require third-party software and a completely different data store and application to access. Some archiving solutions depend on backups or offline storage. These may be acceptable for some environments, but many enterprises want their archive stored where the data was born. They also want it accessible using the same application, as needed, without having to wait for the data to be restored or brought online.

For these users, the only option is to keep all data in the production database. This results in other issues, some of which become increasingly difficult to solve. Those issues include:

- Maintenance operations become more difficult—consider the cost of re-indexing a 1 billion row, 1 TB table
- Restoring a 5 TB database could take a day or more
- Tiered storage solutions (such as SSD + SAS + SATA) can save significant costs compared to physical disks, but often little on overall total cost of ownership (TCO) because they usually reside on the same SAN infrastructure (rack, trays, interconnects, power, management software) and require the same licenses, maintenance contracts, and administration effort

With SQL Server 2016 Stretch Database, you can stretch large operational tables from on-premises to Azure with the ability to query with near-infinite capacity and lower TCO storage. Applications continue to work without code changes, existing database administrator (DBA) skills and processes remain relevant, and developers can continue using current tools and APIs. With Stretch Database, you do not have the complexity or costs associated with managing external archiving and hardware. Stretch Database is a unique feature that solves these common problems.

Transactional databases with large amounts of historical data, typically stored in a small number of tables, are the ideal candidates for stretch databases. These tables may contain more than a billion rows. Stretch Database Advisor, a feature of SQL Server 2016 Upgrade Advisor, can be used to identify databases and tables for Stretch Database. Stretch Database Advisor helps you adopt Stretch Database by analyzing existing database tables based on adjustable table size thresholds to identify candidates for Stretch Database.

Stretch Database can help you selectively begin migrating data to the cloud as needed (Figure). It can also enhance existing applications and facilitate DBA tasks, especially on larger databases.

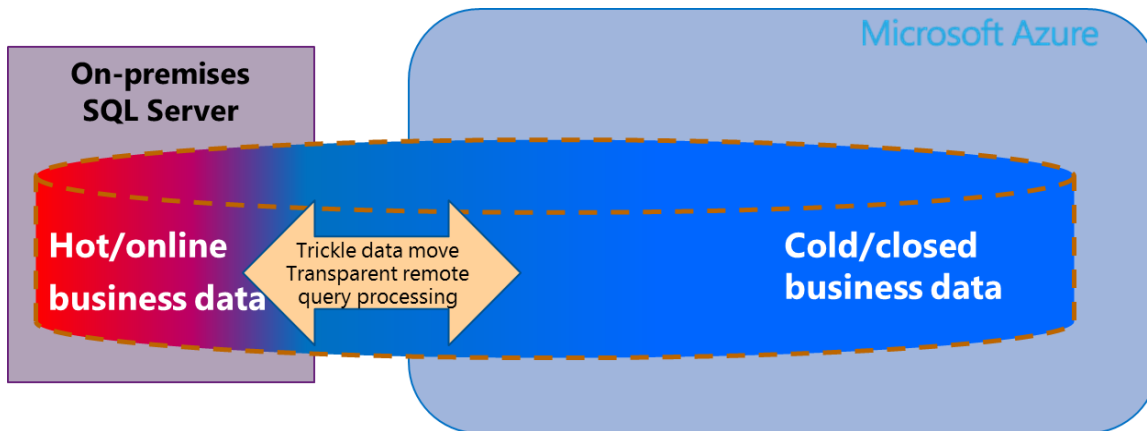


Figure 7: Migration of selected data to the cloud with Stretch Database from on-premises SQL Server

You can now add security policies and Always Encrypted to a table that has been configured for Stretch Database. You can also configure a table that uses Row-Level Security for Stretch Database.

How it works

When you enable Stretch Database on a database, it creates a secure linked server definition in the on-premises SQL Server. This linked server definition has the remote endpoint as the target (Figure). When you enable Stretch Database on a table in the database, it provisions remote resources and begins to migrate eligible data, if migration is enabled.

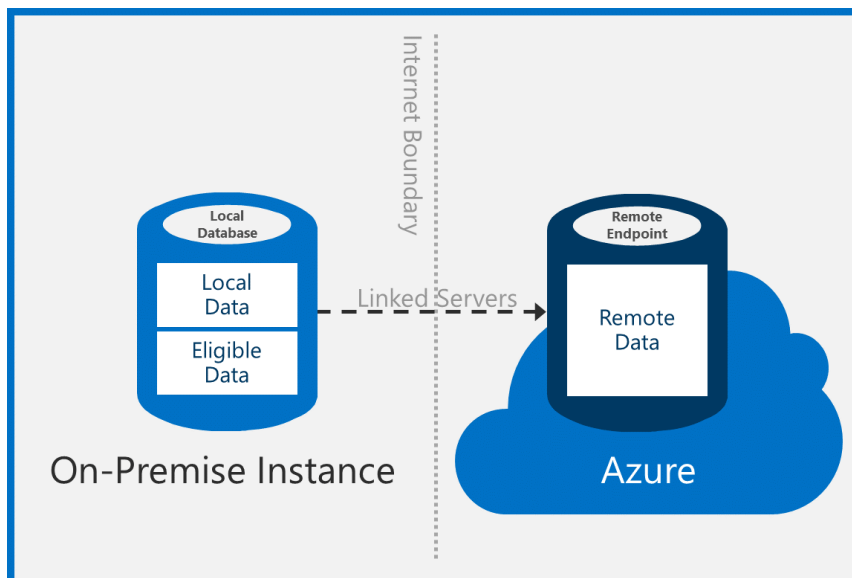


Figure 8: The Stretch Database architecture

Queries against tables with Stretch Database enabled automatically run against both the local database and the remote endpoint. Stretch Database leverages processing power in Azure to run queries against

remote data by rewriting the query. You can see this rewriting as a "remote query" operator in the new query plan.

Cloud on your terms

You can choose to stretch specific SQL Server database tables to Azure. You have the option to move the entire table to Azure, or you can set it up to store only cold data there. Moving the entire table would be a good choice if the table stored only archive data. Moving cold data would be a good choice if you had a single table that contained both hot and cold data. Once this is set up, Stretch Database uses trickle data migration to move the data to Azure. It can also be moved back to the on-premises SQL Server database, so you can achieve transparent, bi-directional data movement without any changes to your stored procedures, functions, other application code, or user access control. You are in control of what is stored in the cloud and what is stored on-premises.

Additionally, the data is kept secure at all times during this process. Transparent Data Encryption (TDE) is applied at rest, in transit, and in use by selective encryption of individual columns in a table with keys that are never given to the database system or cloud service provider.

Enhance existing applications

By choosing which data resides where, Stretch Database can help you enhance existing applications. For example, if you have company policies, service level agreements (SLAs), or TCO goals that may impact how data is archived, Stretch Database can help meet those goals or comply with those policies. A limited amount of computation can occur on Azure, and Stretch Database automatically determines that the compute should happen at the location where it makes the most sense. Further, Stretch Database will transparently scale out by creating additional shards on Azure as needed.

Empower DBAs

Stretch Database has several features that facilitate database administration. There is symmetry in manageability functions and tools (security, access control, monitoring, maintenance, policies, and so on) between on-premises and stretched tables. You can still back up and restore as usual, but because the archived rows are on Azure, the backup takes less time—which is especially important with very large databases. The same principle applies to index maintenance, which is also less costly in a Stretch Database scenario. Most importantly, you can get almost infinite database storage without increasing the complexity or TCO of an external archiving solution.

With Stretch Database, you get a set of elegant solutions to problems common to larger databases. That being said, some moderate performance reductions may occur when accessing the cold data compared to keeping all data on-premises, and any updates or deletes are handled as an administrative function. However, based on the nature of cold data, those tradeoffs will generally be worth the lower TCO, reduced complexity, and other benefits that Stretch Database provides.

Enhanced backup to Azure

Your backup strategy can be greatly enhanced with a hybrid cloud approach. SQL Server has several options for backing up to Azure, including managed backup, backup to Azure block blobs, and Azure Storage snapshot backup. SQL Server 2016 has made enhancements to each of these backup options.

Managed backup

Managed backup provides the capability to easily manage and automate SQL Server backups to Microsoft Azure Blob Storage. You can manage the whole instance or particular databases with an easy interface that can be accessed directly in the SQL Server Management Studio Object Explorer in the management node. It provides off-site, geo-redundant backups with the ability to control the retention period, and supports point-in-time restore for the retention time period specified.

SQL Server 2016 provides several enhancements to this tool. For example, you now have more granular control of the backup schedule for both full and log backups. You can specify a custom backup schedule using the system-stored procedure. Also, backups are made locally and then uploaded to the cloud, allowing for faster recovery and resiliency to transient network issues. You can also now run managed backups on system databases. Additionally, SQL Server 2016 supports simple recovery mode.

Managed backup can also be configured at the database level or at the SQL Server instance level. When configuring at the instance level, any new databases are also backed up automatically. Settings at the database level can be used to override instance-level defaults. You can also encrypt the backups for additional security, and you can set up an automated and custom schedule to control when the backups are taken.

Backup to Azure block blobs

Backup to Azure block blobs allows you to manage backups to Azure Blob Storage with fine-grained control over the process. This service, which was called *backup to URL* in previous versions, has been enhanced with SQL Server 2016 and now supports using block blobs instead of page blobs. The process has been incorporated into the database engine, and the type of blob storage used has been altered from three-copy page blobs to significantly less expensive parity-based block blobs. Block blobs have a size limitation of 200 GB per blob, whereas page blobs have a size limitation per blob of 1 TB. With block blobs, you can stripe your backup set across multiple blobs, resulting in a maximum backup size of 12.8 TB.

Two steps are required to stripe your backup set to block blobs. First, you must obtain a shared access signature (SAS) token to a container in your Azure Storage account. This SAS token is used to create a SQL credential and then again when the Azure container is specified as the backup device. One way to obtain this SAS token is to use Azure PowerShell commands, and then use the same Transact-SQL or PowerShell backup commands without specifying the `WITH CREDENTIAL` clause that was formerly required when using *backup to URL*.

Backup to Azure with file snapshot

SQL Server 2014 introduced data files in Microsoft Azure, which enables native support for SQL Server database files stored as Azure blobs. This allows you to create a database in SQL Server running on-premises or in a virtual machine on Azure with a dedicated storage location for your data in Azure Blob Storage. This also provides an alternative storage location for your database backup files by allowing you to restore from or to Azure Storage.

SQL Server 2016 builds on this capability with backup with file snapshot, which provides the fastest and cheapest method for creating backups and running restores. It uses Azure snapshots to provide nearly instantaneous backups and quicker restores for database files stored using Azure Blob Storage. This capability enables you to simplify your backup and restore policies.

Unlike Azure backup with block blobs, data is not actually moved. Instead, when SQL Server database files are directly stored in Azure Storage, the snapshot of those files is created. You only need to run a full backup once to establish the backup chain. Snapshot backups minimize the use of SQL Server resources to create the backup. This is especially useful for moderate to very large databases, where the impact of backups can be significant.

Because each file snapshot backup set contains a snapshot of each database file, a restore process requires at most two adjacent backup sets. This is true regardless of whether the backup set is from a full database backup or log backup. This is very different than the restore process when using traditional streaming backup files to perform the restore process. With traditional streaming backup, the restore process requires the use of an entire chain of backup sets: the full backup, a differential backup, and one or more transaction log backups. The recovery portion of the restore process remains the same, regardless of whether the restore is using a file snapshot backup or streaming backup set.

High availability and disaster recovery

Availability groups is a high availability and disaster recovery solution that provides an enterprise-level alternative to database mirroring. An availability group is an integrated set of options that includes automatic and manual failover of a group of databases, support for as many as eight secondary replicas (secondaries), faster failover for applications, and automatic page repair. An availability group fails over at the availability replica level. That is, an availability group fails over to one of its secondary replicas (the current failover target). An availability group needs a group listener to provide connectivity to the database.

Add Azure Replica Wizard

The Add Azure Replica Wizard adds a replica of your databases to Azure. Use the Add Azure Replica Wizard to help create a new Windows Azure Virtual Machine in hybrid IT and configure it as a secondary replica for new or existing AlwaysOn Availability Groups. In SQL Server 2016, the Availability Group Listener is created and configured within the wizard. This allows clients to connect seamlessly to the Azure replica after failover, as soon as the wizard completes its setup and without additional complex steps.

Distributed Availability Groups

With Windows Server 2016, members of a Windows Server Cluster no longer need to be joined to the same domain. This enables Distributed Availability Groups, which are “loosely coupled” groups. Secondary replicas of an availability group can exist in different geographical regions than the primary. This enables serving read-only workloads for remote regions, as well as disaster recovery support.

Conclusion

SQL Server has a long history of moving toward easier integration with outside data sources while simplifying management and administration. SQL Server 2016, specifically, has been designed to facilitate working in a hybrid cloud environment in which data and services can reside in many locations. The Stretch Database feature in SQL Server 2016 illustrates the scalability that this kind of interactivity can provide. Administrative functions—such as backups, migrations, and availability—are all enhanced and made simpler with features in SQL Server 2016. SQL Server and Microsoft Azure work better together because the Microsoft hybrid cloud concept drives consistency in tools and processes, making it easier to realize the benefits of hybrid and hyperscale cloud.

More information

The following websites offer more information about topics discussed in this white paper:

- <http://www.microsoft.com/sqlserver/>: SQL Server website
- <http://technet.microsoft.com/en-us/sqlserver/>: SQL Server TechCenter
- <http://msdn.microsoft.com/en-us/sqlserver/>: SQL Server DevCenter
- <http://azure.microsoft.com/en-us/>: Microsoft Azure

Feedback

Did this paper help you? Please give us your feedback by telling us on a scale of 1 (poor) to 5 (excellent) how you rate this paper and why you have given it this rating. More specifically:

- Are you rating it highly because of relevant examples, helpful screen shots, clear writing, or another reason?
- Are you rating it poorly because of examples that don't apply to your concerns, fuzzy screen shots, or unclear writing?

This feedback will help us improve the quality of white papers we release in the future.

Please send your feedback to: sqlsrvwfeedback@microsoft.com