

Microsoft SQL Server 2016: An Initial Assessment

Renewed focus on security and extensibility

This is a licensed reprint from Ovum Research. Additional reports referenced in the Appendix available on the Ovum Knowledge Center

Publication Date: 10 Jun 2016 | Product code: IT0014-003125

Tony Baer



Summary

Catalyst

Microsoft recently announced that SQL Server 2016 would hit final release on June 1. Highlights of the release are in several areas: advanced analytics leveraging PolyBase and in-database R support; expanded in-memory options; a new phased migration option to the cloud; and several security features that make data protection more granular. And Microsoft has extended SQL Server's Reporting Services utility to support the mobile clients that enterprises are expecting as part of the BI stack.

Ovum view

Given that there have been four release candidate offerings, there has been little mystery as to the ultimate makeup of the forthcoming product. With the important exceptions of the Stretch Database capability, which allows portions of a database to migrate to the Azure cloud, most of the enhancements in SQL Server 2016 are supported in various forms in rival database platforms. The significance instead is Microsoft's core competency to popularize and commoditize technologies that debuted elsewhere through simplified visual tooling, aggressive price points, and compatibility with tooling that already reaches a large professional skills base. That continues in the current release, which brings important capabilities to Windows: expanded in-memory computing, pushdown in-database analytics, more granular and dynamic data security, hybrid cloud support, and housekeeping features aimed at productivity and compliance. The teaser announcement of a planned Linux release is still the stuff of product roadmap.

Key messages

- SQL Server 2016 represents a solid update that keeps current with industry trends with in-memory computing, in-database analytics, native mobile BI reporting clients, and more granular security.
- The Stretch Database capability represents a breakthrough for phasing migration of individual databases to the cloud, but as first-generation technology, is feature-incomplete.
- New security options such as Always Encrypted and row-based access are additive, not replacements for existing data protection and access control practices.

Recommendations

Recommendations for enterprises

Microsoft SQL Server 2016 provides a solid update for OLTP and mixed OLTP/analytic workloads that, for the most part, keep current with scaling, in-memory processing, and security features in the database market. It also provides a logical extension of reporting services to support native clients on iOS, Android, and Windows Phone. The exception is the stretch capability, which breaks new ground in providing the ability to migrate parts of a database to the cloud, eliminating the need for all-or-nothing decisions. That capability is potentially very useful for piecemeal cloud migrations or hybrid

scenarios, but for now the uses will be limited to archival (nonchanging) data, but keep your ear to the ground on this one because we expect Microsoft will target early enhancements.

Other highlights include lifting a number of limitations from the Hekaton in-memory technology (e.g., in capacity and ability to accept JSON data), opening it up to a wider range of use cases involving hot spots with frequently updated tables or ETL staging servers. In-database R analytics brings SQL Server to parity with rivals, meaning you can now unleash your data scientists to run programmatic analytics with the benefit of full multithreading. With the announcement of SQL Server 2016, Microsoft is boasting highly competitive TPC-E (transaction) and TPC-H (decision support) benchmarks. While we take all vendor-supplied benchmarks with a grain of salt, our takeaway is that SQL Server 2016 performance is certainly competitive.

Scope

Microsoft's database business has diversified considerably over the years. For SQL relational databases, SQL Server is no longer the only game in town. Microsoft also offers a platform optimized for the Azure cloud as platform-as-a-service: Azure SQL Database. And it offers a massively parallel processing (MPP) analytic appliance that is branded Analytics Platform System (APS). The focus of this release, and this report, is on the newest refresh of the flagship platform, SQL Server 2016.

Extensibility is a core theme

"Stretching" to the cloud

SQL Server 2016 adds a clever option to "stretch" the database so some of it can be stored in the cloud on a new dedicated SQL Server Stretch Database Azure service. Note that while early prerelease versions supported Azure SQL Database, we expect that customers will prefer the full SQL Server route to ensure feature compatibility. The new feature, Stretch Database Advisor, allows selected database tables (or filtered data from a specific table) to move to the cloud based on preset rules for data archiving. In its initial release, the stretch tables would be archived data that would be immutable. By comparison, Oracle Database 12c implements such a feature at the pluggable database level. Aimed at hybrid cloud deployments for transaction systems that keep large amounts of historical data online, the Stretch Database feature may be useful for scenarios such as

- phasing migration of a database from on premise to the cloud (aka "gradual cloud adoption"); and/or
- selectively archiving older data to the cloud while keeping current data on premise, without changing the application.

We like the idea of Stretch Database for cloud migration of a database for mixed transaction and analytic workloads involving current and almost-current data. However, for analytics that will rely heavily on long-term visibility (e.g., going back multiple years for an "active archiving" use case), Hadoop (HDInsight in the Azure cloud) would be the more economical option.

The stretch capability is first generation; it does not yet support

- Tables that are memory-optimized; use change data tracking.
- Specialized data types such as timestamp, XML, spatial, or geography.

- Update or Delete functions (at this point, stretched data in the cloud must be treated as an immutable archive). Getting there will require some more complex engineering for full ACID support.
- Automated bursting of data to the cloud. At this point, compute (and storage) capacity can be ramped up rapidly when there is a spike in activity, but that must be handled manually.

Here's another feature for the wish list: For closely coupled applications (applications with direct reliance on specific SQL Server tables), it would be beneficial to have a feature that could automatically migrate applications in containers along with the stretched tables, but that would require that the limitation to historical data be lifted.

Our verdict? SQL Server 2016's Stretch Database feature is a useful first step for making cloud migration seamless that we expect in future releases will become more feature-complete.

Extending to mobile analytics

SQL Server Reporting Services (SSRS) has always been one of the capabilities that differentiated SQL Server from its principal rivals; by contrast, Oracle and IBM DB2 reporting tools are à la carte (included as part of separate BI suites), while Teradata does not provide native reporting tools. SSRS allows SQL Server customers to avoid the need for buying à la carte BI or visualization tools. SSRS has been Microsoft's on-premise solution for building paginated reports, while Power BI has been the pure cloud-based solution for building interactive visualizations. Microsoft also offers SQL Server Analysis Services (SSAS) as an optional add-on for providing multidimensional OLAP cubes for faster report generation. Outside the database, there is Excel, which can be used for generating analytic reports through its charting capabilities (for that reason, Excel has often been referred to as the world's de facto most popular BI tool).

To this portfolio, SQL Server 2016 adds mobile reports that generate native reports in Android, iOS, and Windows Phone formats. SSRS mobile reports are dynamically adjusted to the footprint – from tablet to smartphone form factor. This capability has come through Microsoft's 2015 acquisition of partner Datazen Software. The company developed a server that converted SSRS and Microsoft Power BI visualizations to mobile form factors. Since the acquisition, Microsoft customers with SQL Server 2008 and newer were entitled to free downloads of the server. SQL Server 2016 integrated that capability, eliminating the need for a separate mobile report server. Obviously, Microsoft is not the first to add mobile BI reporting – those capabilities have become widely available with many third-party BI, reporting, and visualization tools. But adding it to the core database allows Microsoft customers to get basic reporting capabilities – now on mobile – without additional third-party tool cost or architectural complexity.

Also introduced is an upgraded HTML5-based Web Report Manager portal that adds support for the new mobile report-generation capability (mobile and desktop reports can be accessed from the same portal).

A new focus on advanced analytics

In-database R analytics

Although Microsoft does offer a separate SKU for analytics, one of the themes of the 2016 release is extending the envelope for analytics for mixed workload environments. The new release bundles in

several capabilities that bring advanced analytics into the standard editions that most enterprises will buy. Getting R to the mainstream is what this is all about; the capability is timely given the growing popularity of the R programming language.

We reviewed a few months back the integration of R language for in-database analytics. The Enterprise Edition will support full, multithreaded parallelism. Admittedly, Microsoft is not the first to deliver this; IBM, Oracle, and Teradata offer multithreaded R capabilities with their Hadoop offerings. In-database R analytics provides an alternative to SQL for data scientists and data engineers who prefer programmatic paths that could take advantage of advanced analytic approaches such as machine learning. From an administrative perspective, pushing the analytics into the database reduces or eliminates data movement and places it under the security umbrella of the data platform.

PolyBase

Likewise, PolyBase, a capability that allows you to federate and/or push down query processing to Azure targets and provide virtualized views of remote data as extended tables, formerly required a separate SKU (Microsoft's parallel Analytics Platform System appliance, or APS); it is now part of the core platform. That does not necessarily eliminate the need for APS, but it means that you can now federate or push down queries to multiple targets without requiring a specialized MPP data warehouse appliance.

As noted, most of Microsoft's rivals already offer this capability in their latest versions (e.g., Teradata QueryGrid, Oracle Big Data SQL). Specialized MPP data warehouses like Vertica and Pivotal (Greenplum database) also can push down query processing to Hadoop. And increasingly, pushdown is also becoming part of the vocabulary for analytics tools (e.g., Alteryx, Zoomdata). Bundling PolyBase with core SQL Server therefore is the latest evidence that multi-platform or federated query is becoming commoditized in much the same way that SQL Server Analysis Services did for OLAP a decade or more ago. Not all federated/pushdown query capabilities, however, are created equal, or are licensed equally. PolyBase provides good value in that it is bundled with the Enterprise Edition and connects SQL Server 2016 with Hadoop (on premise and/or in the Azure cloud through HDInsight via Azure BLOB storage) for no extra cost. By contrast, Teradata QueryGrid supports a greater variety of targets; beyond Hadoop, it also supports Oracle and SQL Server, but you must pay extra for the connectors.

Other key analytics-related enhancements

They include:

- Unified Visual Studio support of analytics and database development tools. There are no longer two distinct sets of tooling for developing SQL Server database and BI server tools. They have been converged, a positive development that should clear the way for development of hybrid applications that embed analytics with transaction processing.
- Updatable, columnar indexes that integrate with on-disk row stores, which will make operational analytics more seamless.

What about APS?

As the pure analytic counterpart to SQL Server, Microsoft offers the APS MPP appliance. In the SQL Server release cycle, Microsoft is not updating APS. Instead, the emphasis on this go-round is that SQL Server 2016 is borrowing some features from APS, making it a more versatile platform for mixed

OLTP and analytics workloads. Nonetheless, for core analytics use cases (where platforms such as Teradata, IBM Pure Data/Netezza, Vertica, Greenplum/Pivotal, or Oracle Exadata might be considered), APS (or Azure Data Warehouse in the cloud) remains the Microsoft product of choice.

Riding the in-memory wave

The economics of silicon-based storage (including dynamic random-access memory (DRAM) and flash) have in the past few years expanded the array of options in the database arena, both on premise and in the cloud. The impact of in-memory is not simply speed, but the ability to orchestrate embedded analytics into transaction processing or enable smart analytics that leverages machine-learning capabilities. Microsoft is following a long line of players, including:

- established providers like Oracle, IBM, and Teradata, which offer columnar in-memory options to their row-based, transaction stores
- SAP, which notably introduced the HANA pure in-memory database as the foundation for its next-generation enterprise applications and as the turbocharger for Business Information Warehouse
- upstarts such as MemSQL, Redis, and Aerospike, which offer platforms based wholly or jointly on memory and/or flash storage optimized specifically for distributed Internet transaction-processing applications
- Hadoop, where emerging compute engines such as Spark are placing a premium on memory-intensive processing nodes.

Hekaton, the code name for Microsoft's in-memory configuration of SQL Server, was introduced in the 2014 edition. For SQL Server 2016, the capabilities have significantly expanded. For instance, while the original version supported memory-optimized tables and natively compiled stored procedures, the new release allows you to alter them. Other enhancements include support for large data objects – which could include variably structured data entities such as small JSON documents – and the option for making in-memory data nondurable (making it useful as a caching option for volatile data that would otherwise not be persisted). As would be expected, maximum capacity has burst through the 128GB/instance limit of the previous release to whatever is supported by the operating system.

These enhancements make SQL Server useful for mixed use cases such as embedding analytics with some transaction-processing operations. However, the new features do not necessarily make SQL Server a replacement for more specialized platforms, such as the niche platforms from start-up providers (e.g., MemSQL, Aerospike) that are engineered strictly for high-throughput, distributed transaction use cases that may tolerate eventual consistency (e.g., modified ACID), and in some cases, are offered with more aggressive pricing.

Extended JSON support: Not a MongoDB killer

SQL Server 2016's extended JSON support is the latest evidence of a trend toward overlap and convergence of formerly specialized database platforms – SQL, NoSQL, and Hadoop platforms are blending together. But even if a database platform can support multiple data formats and structures, it does not make their strengths and weaknesses equal.

So, as we stated with in-memory processing, the "native" JSON support introduced in SQL Server 2016 is not intended to replace a JSON-based NoSQL database such as MongoDB, Couchbase, or

Microsoft's own cloud-based Azure DocumentDB. JSON in SQL Server is, in essence, a "flavor" that extends the database to address applications that mix transaction and JSON document data, and allows JSON data to be managed under a common data management umbrella. SQL Server treats JSON not as a new data type (which would disrupt the schema for SQL Server users) but as a form of the variable character type it already supported. The data can be queried via SQL by flattening complex JSON documents to single rows, and can be reconstituted as full JSON output. There is flexibility in how the data can be stored – it can go into Hekaton, row, temporal, and column store tables. But there is one important limitation: For now, in-memory storage is restricted to smaller JSON objects.

This is akin to how IBM, Oracle, and Teradata treat JSON. But if your application is strictly confined to JSON data, you are best off using one of the specialized platforms that provides indexing optimized for JSON's complex structures and also allows the flexibility of operating without preset schema.

Stepping up security

Hear no evil, see no evil: Always-on encryption shifts protection to the client

Microsoft has been steadily upping its game over the years with protecting data in SQL Server. SQL Server 2012 introduced encryption as an option; you could encrypt the full database, via the Transparent Data Encryption option; specific columns, files (a capability introduced at the OS level in Windows 2000), or volumes (via BitLocker); or have the application handle it. In turn, SQL Server 2014 added support for Common Criteria security certifications and addition of the capability to encrypt backups. Those features used the database engine to manage encryption.

By contrast, SQL Server 2016 adds a new option where encryption is actually managed *outside* the database. The feature, Always Encrypted, moves control over encryption and decryption to the client tier and will be available with the high-end Enterprise Edition. It doesn't replace existing Transparent Encryption options that are enforced by the database, but it adds a new client-side alternative for data in motion. By comparison, Oracle provides external encryption at the network level, while IBM DB2 does not offer this option (it supports traditional approaches managed by the database itself).

The good news is that Always Encrypted mode prevents decryption exploits that compromise the entire database and keeps the data encrypted both at rest and while it is being accessed or processed. It enforces separation of duties between those who own the data (the client) and those who manage it (the database administrators), because only the clients have access to the keys. The not-so-good news is that the attack vector is shifted to the client, meaning the onus of protection moves outside the database to the edge, where it would have to be protected via either an external tool such as Azure Key Vault or the application.

Data access getting more granular

Row-based access

Paid SQL Server 2016 editions (Enterprise and Standard) have added more granularity in protecting access to data. They have implemented two related features that regulate the data that can be

accessed and by whom. The features include row-level security and dynamic masking that are both implemented by the database engine, as opposed to an external tool or application. Here, Microsoft is playing catch-up to features that are already present in Oracle and IBM DB2. In SQL Server, row-level security is implemented through SQL statements.

Dynamic masking

While row-level security guards access to specific records, dynamic masking handles how specific columns of unencrypted data are displayed. This is an example of Microsoft's "cloud-first" strategy for introducing new features to databases (it is hardly alone; Oracle is also taking this tack). Dynamic masking debuted in Azure SQL Database before coming to the flagship product in the Enterprise and Standard editions of the SQL Server 2016 release.

With dynamic masking, the data may be displayed as is or obfuscated based on the user's (or user role's) privileges. Note that even if a user is entitled to see only masked data, authorization for making updates is managed separately.

Comparisons with other tools are not exact. For instance, IBM offers dynamic masking and Oracle does not, but both offer broader capabilities not (yet) matched by Microsoft, such as discovering sensitive data and data dependencies and applying conditional masking. And there is an active third-party market for data protection tools (e.g., Protegrity) that offers a broad range of policy-based data protection capabilities such as encryption and tokenization.

For Microsoft, row-based dynamic data masking in SQL Server 2016 is the first step toward broader, policy-based protection.

Rounding out the platform

Scaling

As would be expected, the 2016 version raises the upper limits on scale. The Enterprise Edition can now support as many cores and as much memory as is supported by the underlying OS, a specification that will prompt customers to upgrade to Windows 10 (support for Linux, as noted below, is still in the future). That compares to upper limits of 640 cores and 4TB of memory for SQL Server 2014. The theoretical limit for overall storage size for SQL Server 2016 is 524PB; while Microsoft has tested the new product in that configuration, as a practical (and economic) matter, we would expect customers to be running at hundreds of petabytes at scale to use Hadoop. But recognizing that typical enterprise installs will continue inflating in size (both in storage and processing capacity), Microsoft is shifting pricing of SQL Server from processors to cores.

Keeping the trail of breadcrumbs

Lineage has always been an important capability for understanding what happens to data over time; it can be a useful productivity aid to track the efficacy of data manipulation, and for highly regulated scenarios, may be required for proving compliance for requirements such as segregation of duties or proving that read/write access controls have been properly enforced. SQL Server 2016 adds several lineage features that should prove useful for DBA productivity.

They include a new Query Store feature that provides more comprehensive access to query execution plans. Until now, you could only view plans that were active in cache; with 2016, you can see the history of all plans that have been committed to the plan cache and the execution statistics that accompany them.

Another lineage-related new feature is temporal tables that have snapshots of old rows in a base table; each time the rows are updated, old versions are preserved in a temporal store. For DBAs, this eliminates the need to manually construct their own row-version solutions.

Linux will wait

There was considerable excitement after Microsoft announced plans to port SQL Server to Linux back in March. It is a logical move given Microsoft's clear evolution toward becoming a multi-platform company, as evidenced by the unmistakably growing presence of Linux in the Azure cloud (e.g., availability of Red Hat Enterprise Linux VMs); embrace of iOS and Android (alongside Windows Mobile) for mobile versions of Microsoft Office; and continuation of R Server (the rebranded product from the Revolution Analytics acquisition) support for Linux.

But Linux is not part of the initial SQL Server 2016 rollout. We expect that there will be public previews of SQL Server on Linux initially in Azure, and later on premise, over the next six to 12 months. And we also expect that initial versions of SQL Server on Linux will carry a subset of functionality that is available on Windows, with a subsequent ramp-up to parity.

Appendix

Methodology

This report was compiled through ongoing discussions with Microsoft customers, briefings with Microsoft, and research of data sheets and documentation for SQL Server and competing products.

Further reading

"Microsoft extends the path with enterprise R," IT0014-003105 (February 2016)

Microsoft and Salesforce: Past, Present, and Future, IT0021-000165 (April 2016)

SWOT Assessment: Microsoft Dynamics CRM 2016, IT0020-000188 (April 2016)

"Microsoft Azure Data Lake takes big step in taming big data," IT0014-003078 (November 2015)

Author

Tony Baer, Principal Analyst, Information Management

tony.baer@ovum.com

Ovum Consulting

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Ovum's consulting team may be able to help you. For more information about Ovum's consulting capabilities, please contact us directly at consulting@ovum.com.

Copyright notice and disclaimer

The contents of this product are protected by international copyright laws, database rights and other intellectual property rights. The owner of these rights is Informa Telecoms and Media Limited, our affiliates or other third party licensors. All product and company names and logos contained within or appearing on this product are the trademarks, service marks or trading names of their respective owners, including Informa Telecoms and Media Limited. This product may not be copied, reproduced, distributed or transmitted in any form or by any means without the prior permission of Informa Telecoms and Media Limited.

Whilst reasonable efforts have been made to ensure that the information and content of this product was correct as at the date of first publication, neither Informa Telecoms and Media Limited nor any person engaged or employed by Informa Telecoms and Media Limited accepts any liability for any errors, omissions or other inaccuracies. Readers should independently verify any facts and figures as no liability can be accepted in this regard – readers assume full responsibility and risk accordingly for their use of such information and content.

Any views and/or opinions expressed in this product by individual authors or contributors are their personal views and/or opinions and do not necessarily reflect the views and/or opinions of Informa Telecoms and Media Limited.

CONTACT US

www.ovum.com

analystsupport@ovum.com

INTERNATIONAL OFFICES

Beijing

Dubai

Hong Kong

Hyderabad

Johannesburg

London

Melbourne

New York

San Francisco

Sao Paulo

Tokyo

