

The background of the top section is a low-angle photograph of several modern skyscrapers with glass facades, reaching towards a cloudy sky. The image is partially overlaid by orange and grey geometric shapes.

CosmosDB in the new emerging world - Use cases and applications

INTRODUCTION

In today's globalized world, businesses have customers and suppliers that are spread across the globe, accessing digital platforms, conducting transactions and updating information in real-time.

A shift in user expectations and behavior, aggressive competition and innovations in the market show that consumers are no longer satisfied with slow-performing legacy platforms.

End-users have come to expect instant gratification from snappy B2C and B2B platforms that respond to user interactions with close to zero latency. Latency creates poor brand experiences, drives user dropouts and loss of customers.

Businesses need to always be available, have the ability to scale, provide instant access and store information across global locations with close to zero latency. This universal speed expectation has become table stakes.

Welcome Azure Cosmos DB to the game! Azure Cosmos DB is Microsoft's globally distributed, multi-model database service. Consider the business use cases it enables, and its value proposition in the customer environment.

A global Fortune 100 firm runs an E-commerce B2C portal. The portal has huge volumes of worldwide users accessing the digital platform daily. Alternatively, consider a large retail firm that has its supply chain platforms being accessed at all times by global suppliers and vendors, and updated in real-time through IOT systems.

Think about many data flows coming in real-time to your inventory tracking system from your distribution centers, trucking systems, and point of sale (POS) systems across different locations spanning multiple geographies. Combine this with the need to process, update and generate useful business information in real-time. The volume, technical complexity and processing speed requirements could be overwhelming.

WHAT IS COMMON AMONG THESE DIFFERENT CONTEXTS?

- a They are mission-critical applications where the digital platforms are at the core of the business.
- b They are globally distributed applications.
- c Users expect the systems to be always available on-demand. If the system is unavailable even for a very minor amount of time, it can have a drastic impact on the business, and pileup of potential losses.
- d Users expect the system to be highly responsive. Imagine the negative perception of a global E-commerce portal that takes 20 seconds to load or a supplier who must wait 30 seconds to get real-time updates of the inventory supply connected through IOT platforms.
- e The systems are update intensive. Continuous updates to inventory are happening in real-time, which then has to get reflected across globally distributed database systems.

Combine this with the need to process, update and generate useful business information in real-time. The volume, technical complexity and processing speed requirements could be overwhelming.

In SHORT, GLOBAL FIRMS with DISTRIBUTED APPLICATIONS



need to be
always online



need to be
highly responsive



need very
low latency



need for elastic
on-demand
scalability

Azure Cosmos DB is a globally distributed database service that's designed to provide low latency, elastic scalability of throughput, well-defined semantics for data consistency, and high availability. In short, if your application needs guaranteed fast response time anywhere in the world, if it's required to be always online, and needs unlimited and elastic scalability of throughput and storage, you should consider building your application on Azure Cosmos DB. With Azure Cosmos DB, you can add or remove the regions associated with your account at any time. Your application doesn't need to be paused or redeployed to add or remove a region. It continues to be highly available all the time because of specific capabilities that the service natively provides.

UNIQUE VALUE PROPOSITION for BUSINESS

In brief, as Microsoft Cosmos DB architect and product manager Rimma Nehme puts it very succinctly - "If you have write-heavy workloads, spanning multiple geos, and you need this near real-time ingestion of your data, this becomes extremely attractive for IoT, web, mobile gaming scenarios, Cosmos DB is for you."



Automatic
multi-region
geo-replication



High
availability



Applications
don't need
redemption
during regional
failover



Guaranteed Low
latency (including
write) with
< single-digit
milli-secs

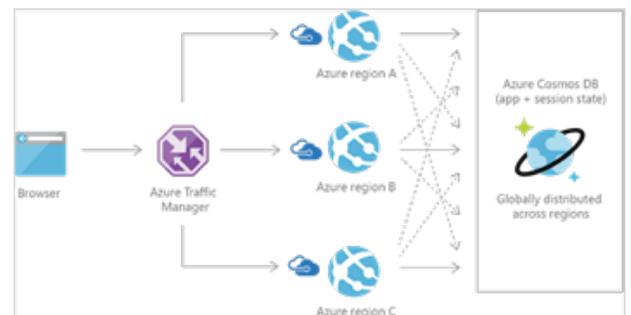


Comprehensive
service level
agreements
(SLA)

ARCHITECTURAL VIEW to COSMOS DB

A simplified reference architecture for a global-scale, real-time, high demand variability application considerations described above is shown here.

Cosmos DB allows these capabilities in a seamless manner. Conceptual architecture is referenced below.



Courtesy: <https://github.com/MicrosoftDocs/azure-docs/blob/master/articles/cosmos-db/use-cases.md>

HOW is COSMOS DB DIFFERENT/BETTER THAN the ALTERNATIVES?

The “fit” of each data store solution is based on the business use case and the overall technology space of “Cloud Databases”

The business usecases described so far give you a good high-level understanding of the business uses of Cosmos DB. One alternative that maps very closely to Cosmos for very similar needs is Amazon Dynamo DB. Both platforms support performance at scale, and high-availability and globally distributed applications. From a customer standpoint, the key differentiator in addition to the different technology stack is the ability for Cosmos DB to have more diverse multi-model capabilities (allowing it to meet various needs in addition to the NOSQL use case).

USE CASES for LIFT & SHIFT to COSMOS DB

In customer scenarios, where NOSQL databases are in-use, the customers are on MongoDB running it on Azure VMs or they are running on-premise models. This requires the customer IT teams to manage all of the provisioning, scaling and managing the data needs of the application. These scenarios are ripe scenarios for seamless lift and shift migration to Cosmos DB, since the existing set of back-end APIs for NOSQL Databases can be used as-is, and rewired to Cosmos DB.

HOW DOES COSMOS DB ENABLE THESE BENEFITS?



Global distribution

In the traditional world, creating any large-scale globally distributed application would have meant a significant hassle in terms of planning for complex and multi-data center environment configurations.

Since Cosmos DB has been designed purposefully as a globally distributed database, it can be configured to automatically replicate all of the data into different regions.

A key capability in this context is the concept of multi-homing: Cosmos DB is automatically aware of the nearest region to that request, and sends the request automatically to that region. So, when additional regions are configured, there is no re-deployment needed, and the application continues to be available at all times.



Elastic scale out of storage and throughput

Cosmos DB allows you to elastically scale and provision the throughput across one or multiple global regions. Cosmos DB allows almost unlimited scale for throughput as well as storage capacity, based on demand that scales up or down. And you only pay for the storage you need and the throughput that you provision.



Guaranteed single-digit-millisecond latency

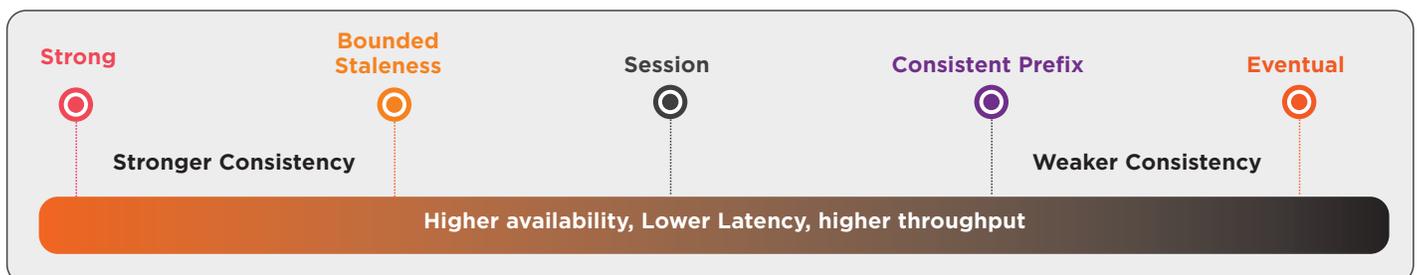
Cosmos DB supports 99.99% availability SLA for all single region database accounts, and all 99.999% read availability on all multi-region database accounts.



Choice of consistency

From a business application standpoint, consistency is a key consideration for globally distributed applications. Azure Cosmos DB has the capability to adjust the consistency level based on the customer's needs.

This is a key benefit since Cosmos allows triaging the customer's needs to the combination of latency, throughput, and availability specific to customer requirements. The different levels of consistency offered are Strong, Bounded Staleness, Session, Consistent Prefix and Eventual.



Enterprise-level SLAs

Cosmos DB provides comprehensive SLAs (four 9s) for availability, throughput, consistency and latency. This is key for businesses that run the mission-critical applications and cannot afford operational risk.

UNIQUE VALUE PROPOSITION for TECHNOLOGY

Based on the developer's needs, Cosmos DB allows the data to be treated in different ways. This built-in multi-model capability means that regardless of the data-model chosen, all of the benefits like scale, latency & availability and global distribution are available to the business.

Cosmos DB natively supports document, key-value, graph, and column-family data models. With Azure Cosmos DB, you can access your data using APIs of your choice: such as SQL (document), MongoDB (document), Azure Table Storage (key-value), Gremlin (graph), and Cassandra (column-family) are all natively supported.

Another key capability in Cosmos DB from a technology standpoint is the capability of multi-master replication. Cosmos DB is one of the first operational databases out there in the marketplace that runs on such a scale and enables globally scalable multi-master available to the customers.

This capability allows the writes from the application to the data-base to happen to the closest region, significantly reducing the write latency. It also provides for a significantly more scalable architecture, since the developer does not need to have a single region identified as a master, and then have replication of data updates pushed to all other regions. All of this is automatically taken care of by Cosmos DB.

CONCLUSION

Cosmos DB can stake a claim as the database of today that anticipates the future as a common back-end for a range of mission-critical globally distributed data storage applications.

REFERENCES

<https://docs.microsoft.com/en-us/azure/cosmos-db/distribute-data-globally>

<https://github.com/MicrosoftDocs/azure-docs/blob/master/articles/cosmos-db/use-cases.md>

<https://techcrunch.com/2018/09/24/microsoft-updates-its-planet-scale-cosmos-db-database-service/>

About Orion

Headquartered in Edison, New Jersey, with 3500 associates across 12 delivery centers, Orion is uniquely positioned as a high touch, high quality and nimble solutions provider with the proven expertise and scale to solve complex problems of Fortune 1,000 customers. Orion provides digital transformation, artificial intelligence, machine learning, advanced data analytics and automation services to large enterprise customers across the financial and professional services, communications and media, professional sports and entertainment, education and healthcare industries.

Contact Us

Orion Business Innovation

333 Thornall Street, 7th Floor, Edison, NJ 08837

Tel: +1-877-456-9922 | Email: info@orioninc.com

📍 USA | UK | Germany | Russia | Switzerland | CEE | India | Australia

