Azure Cosmos DB was built from the ground up with global distribution and horizontal scale at its core – it offers turn-key global distribution across any number of Azure regions by transparently scaling and replicating your data wherever your users are.

Global distribution
Easily build globally-distributed applications without the hassle of complex, multiple-datacenter configurations. Designed as a globally distributed database system, Cosmos DB automatically replicates all of your data to any number of regions of your choice, for fast, responsive access. Cosmos DB supports transparent multi-homing and guarantees 99.99% high availability.

Multi-model + multi-API
Only Cosmos DB allows you to use key-value, graph, and document data in one service, at global scale and without worrying about schema or index management. Cosmos DB automatically indexes all data, and allows you to use your favorite NoSQL API including SQL, JavaScript, Gremlin, MongoDB, and Azure Table storage to query your data.

Choice of consistency
Cosmos DB offers five well-defined consistency levels—strong, bounded staleness, session, consistent-prefix and eventual—for an intuitive programming model with low latency and high availability for applications spanning the world.

Elastic scaleout of storage and throughput
With Cosmos DB, you only pay for the throughput and storage you need. Cosmos DB allows you independently and elastically scale storage and throughput across one or multiple global regions.

Guaranteed single-digit millisecond latency
Serve read and write requests from the nearest region while simultaneously distributing data across the globe. With its latch-free and write-optimized database engine, Cosmos DB guarantees less than 10-ms latencies on reads and less than 15-ms latencies on (indexed) writes at the 99th percentile.

Comprehensive SLAs
Rest assured your apps are running on world-class infrastructure, with "battle-tested" service, in the most trusted cloud. Cosmos DB is the first and only service to offer industry-leading comprehensive 99.99% SLAs for latency at the 99th percentile, guaranteed throughput, consistency and high availability.
Apps with global reach

Elastically scale out storage and throughput by transparently adding more partitions under each Cosmos DB collection to meet the bursty usage patterns of internet scale web and mobile applications. Replicate data across any additional regions of your choice to deliver low-latency access for a global user base.

Product catalog

Product catalog attributes have high variability and change often over time. Cosmos DB’s automatic indexing over flexible schemas is a perfect fit for storing product catalogs, IoT device registries, and other catalog systems.

Real-time personalization

Cosmos DB offers tunable consistency levels and single-millisecond latencies for reads and writes, enabling applications to retrieve user profiles and personalized settings efficiently to render UI experiences quickly.

IoT and telematics

Easily store high-volume Internet of Things (IoT) device data, and leverage real-time change feeds to respond quickly to anomalies.

Gaming

Modern games often rely on cloud databases to deliver social and personalized content like in-game stats and high-score leaderboards. Games databases often require single-millisecond latencies for reads and writes to deliver lag-free experiences, and handle massive spikes in request rates during new game and feature launches.