Azure Cosmos DB

Globally distributed multi-model database

Agenda

- What
  - Walkthrough
    - CLI/PS, Portal
- IT Pro issues
  - Availability
  - Capacity management
  - Security
  - Performance management
    - Alerts/Metrics

InfoWorld FROM IDG INSIDER

Home > Databases

INFOWORLD TECH WATCH

By Serdar Yegulalp, Senior Writer, InfoWorld | MAY 11, 2017

About | Informed news analysis every weekday

Why Microsoft's Cosmos DB represents the future of cloud databases

Here are four reasons why there's more to Microsoft's new database as a service

News Conference Index TQ Deals Answers Tech5 Expand to Asia

CATEGORIES

Azure Cosmos DB is Microsoft's new database for globally-distributed applications

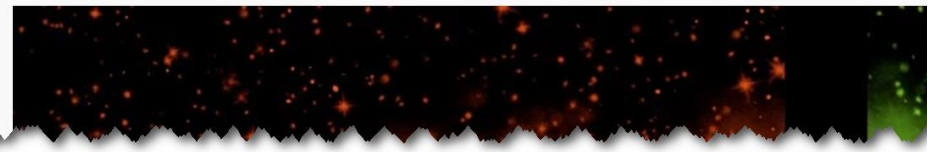
by MATTHEW HUGHES — 12 days ago in DESIGN & DEV

TC News Startups Mobile Gadgets Enterprise Social Europe Trending Facebook Tesla Snap

Cloud developers databases cloud

With Cosmos DB, Microsoft wants to build one database to rule them all

Posted May 10, 2017 by Frederic Lardinois (@fredericl)



PCWorld FROM IDG

NEWS REVIEWS HOW-TO VIDEO BUSINESS LAPTOPS TABLETS PHONES HARDWARE SECURITY SOFTWARE GADGETS

Business Software Mobile Network Storage **Cloud & Services**

Home / Cloud & Services

NEWS

Microsoft takes its databases worldwide with Cosmos DB

The update to DocumentDB provides global scale for app data

At Build conference last month, Microsoft announced Azure Cosmos DB, a brand new cloud database service built from the ground up to power planet-scale cloud services and data-intensive applications. Microsoft claims that Azure Cosmos DB is the industry's first globally distributed, multi-model database service to deliver horizontal scale with guaranteed uptime, throughput, consistency and single-digit millisecond latency at the 99th percentile. Microsoft is also adding several new features to its cloud database services. Back in April, Microsoft announced Azure SQL Database Threat Detection, a new security intelligence feature built into the Azure SQL Database service. I think Microsoft will further move towards Amazon in the Forrester Wave for DBaaS in the coming months.

Azure Cosmos DB

A globally distributed, massively scalable, multi-model database service

Table API

MongoDB API

Key-value

Column-family

Document

Graph

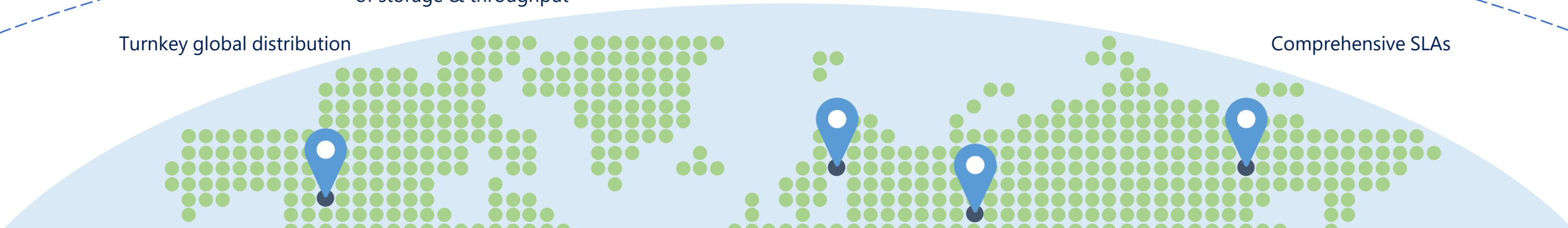
Guaranteed low latency at the 99th percentile

Elastic scale out
of storage & throughput

Five well-defined consistency models

Turnkey global distribution

Comprehensive SLAs



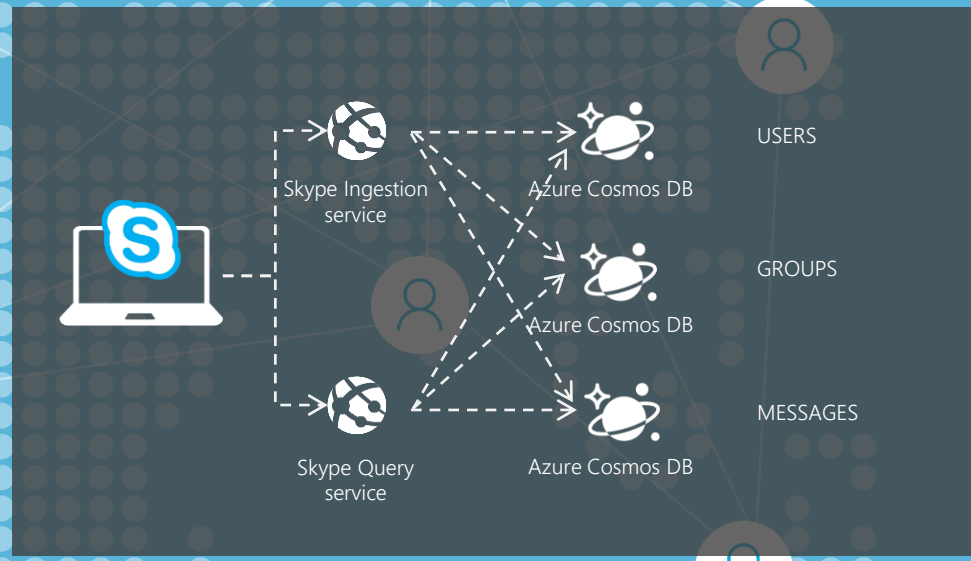
Skype powers 1M searches per second over conversation data

Business need

- Provide search capabilities over TBs-PBs of Skype and Teams conversations
- Fast ingestion with multiple writes, overlay group memberships
- Secure & compliant data storage with high privacy requirements

Key benefits

- Cosmos DB supports fast ingestion of message data from 1:1 communication, group chats
- Cosmos DB enables real-time query over message and group conversations, with custom filters on when user enters/leaves thread



skype™

6TB
User data

1TB
Group data

44TB
Message data

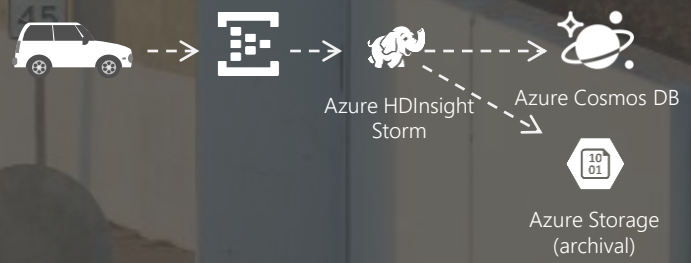
Toyota drives connected car push forward with Azure Cosmos DB

Business need

- Need to ingest massive volumes of diagnostic data from vehicles and take real-time actions as part of connected car platform
- Management and operations of database infrastructure to handle exponential growth of data

Key benefits

- Cosmos DB can scale elastically without operational overhead of MongoDB
- Perform fast queries over events to deliver recommended services, safety notices to vehicles
- Perform staged migration via MongoDB APIs



8TB
Vehicle Telemetry

250K
Lexus Cars

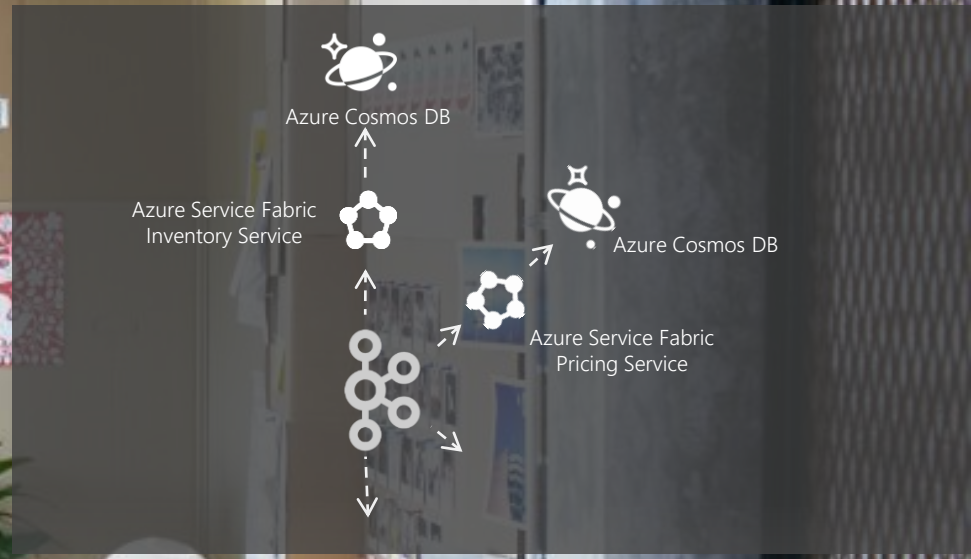
E-commerce challenger eyes the top spot, runs on Azure Cosmos DB

Business need

- Process Ms of retail transactions per second in milliseconds in inventory pipeline during peak ("Black Friday")
- Fast development cycles and loosely coupled micro-services to keep up with a competitive marketplace

Key benefits

- Cosmos DB provides elastic scalability from 1-10M requests per second
- Improved reliability, and faster order processing times than previous OSS solution
- Reduced development time and operational overhead



jet

10M
Peak RPS

12TB
Provisioned

64
Databases

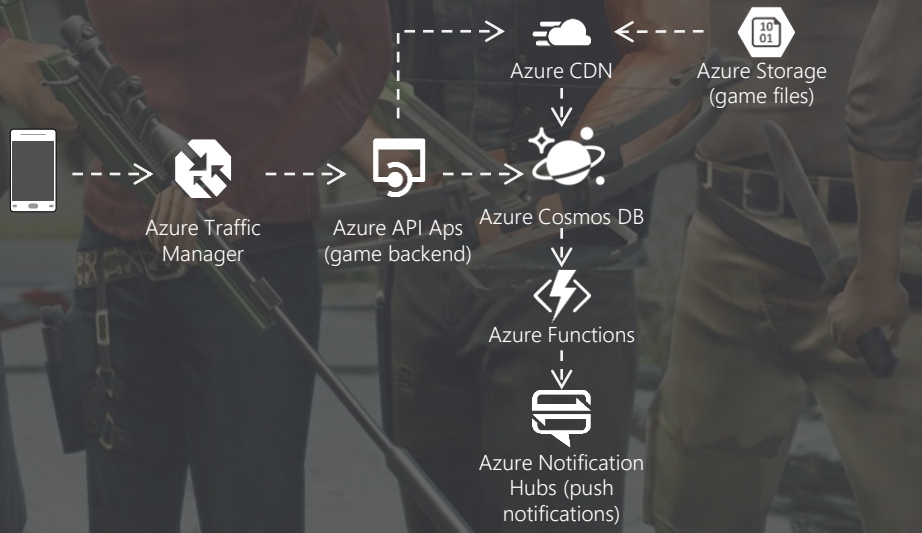
Performance at massive scale allows millions to play mobile game

Business need

- Handle millions of players on Day 1 due to popularity of the TV series
- Match-making of players for competitive and lag-free experience
- Provide new content weekly, and iterate on social functionality
-

Key benefits

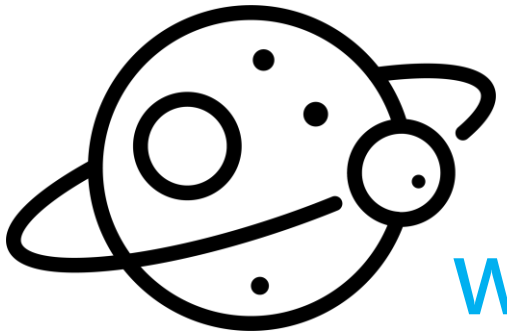
- Cosmos DB provides elastic scalability for millions of users and flexible schema to support social features and gameplay
- Global distribution allows for low latency for players spread worldwide
- Automatic indexing used to build real-time leaderboards



#1
iOS App Store

1B
Daily Queries

1M
Peak Active



What sets Azure Cosmos DB apart

Turnkey Global Distribution

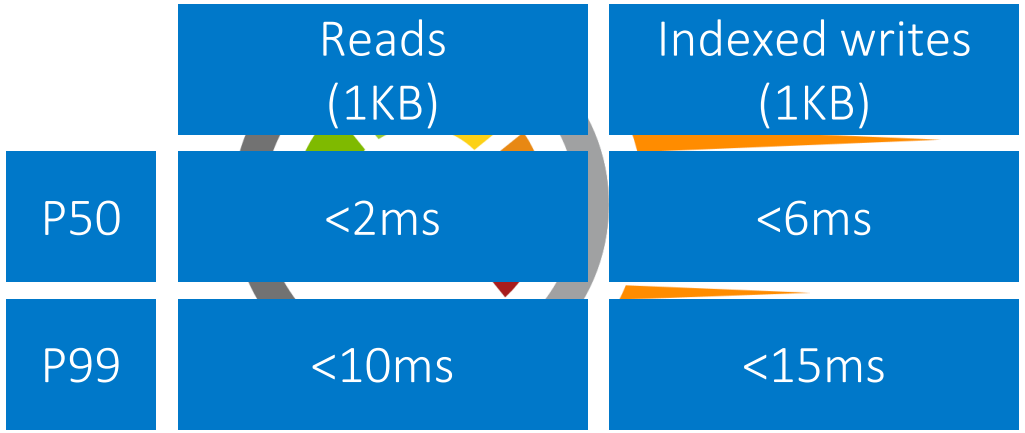
Worldwide presence as a Foundational Azure service

Automatic multi-region replication

Multi-homing APIs

Manual and automatic failovers

Designed for High Availability



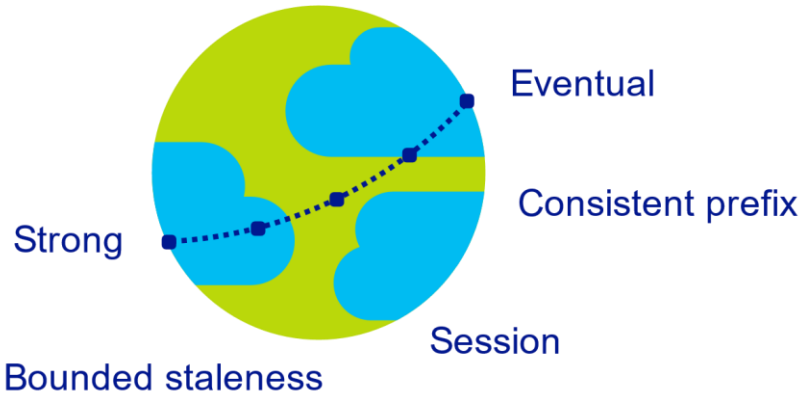
Guaranteed low latency at P99 (99th percentile)

Requests are served from local region

Single-digit millisecond latency worldwide

Write optimized, latch-free database engine designed for SSD

Synchronous automatic indexing at sustained ingestion rates



Multiple, well-defined consistency choices

Global distribution forces us to navigate the CAP theorem

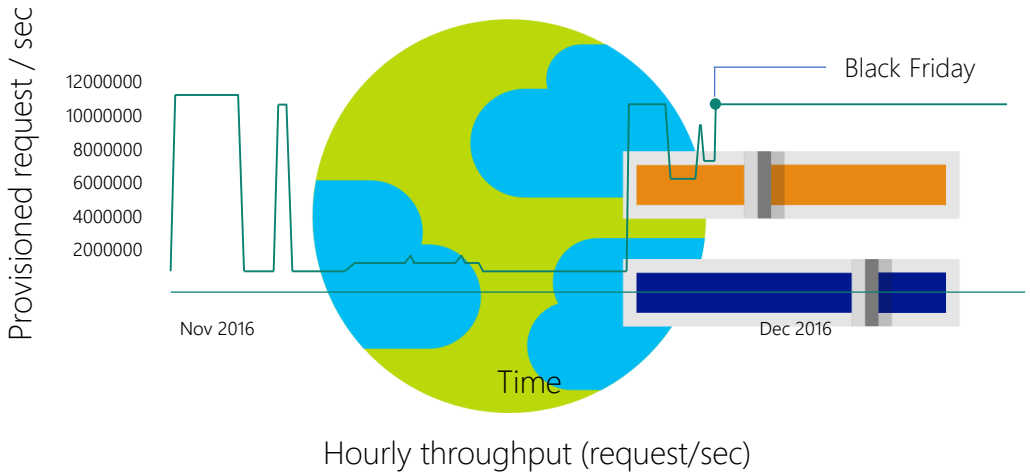
Writing correct distributed applications is hard

Five well-defined consistency levels

Intuitive and practical with clear PACELC tradeoffs

Programmatically change at anytime

Can be overridden on a per-request basis



Hourly throughput (request/sec)

Elastically scalable storage and throughput

Single machine is never a bottle neck

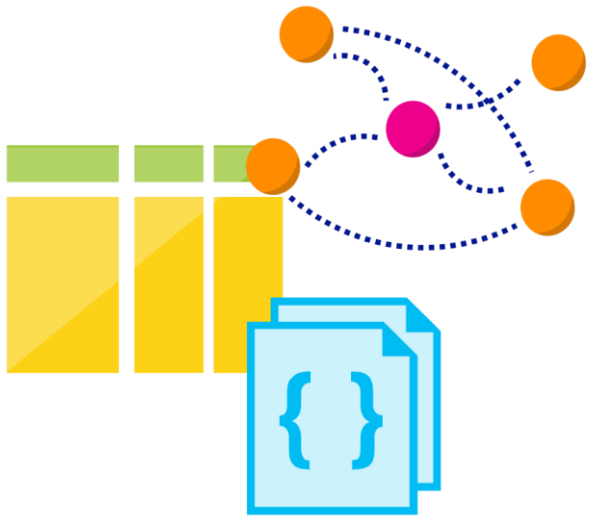
Transparent server-side partition management

Elastically scale storage (GB to PB) and throughput (100 to 100M req/sec) across many machines and multiple regions

Automatic expiration via policy based TTL

Pay by the hour, change throughput at any time for only what you need

Support for both request per second and requests per minute to handle spikes cost-effectively



Multi-model, multi-API

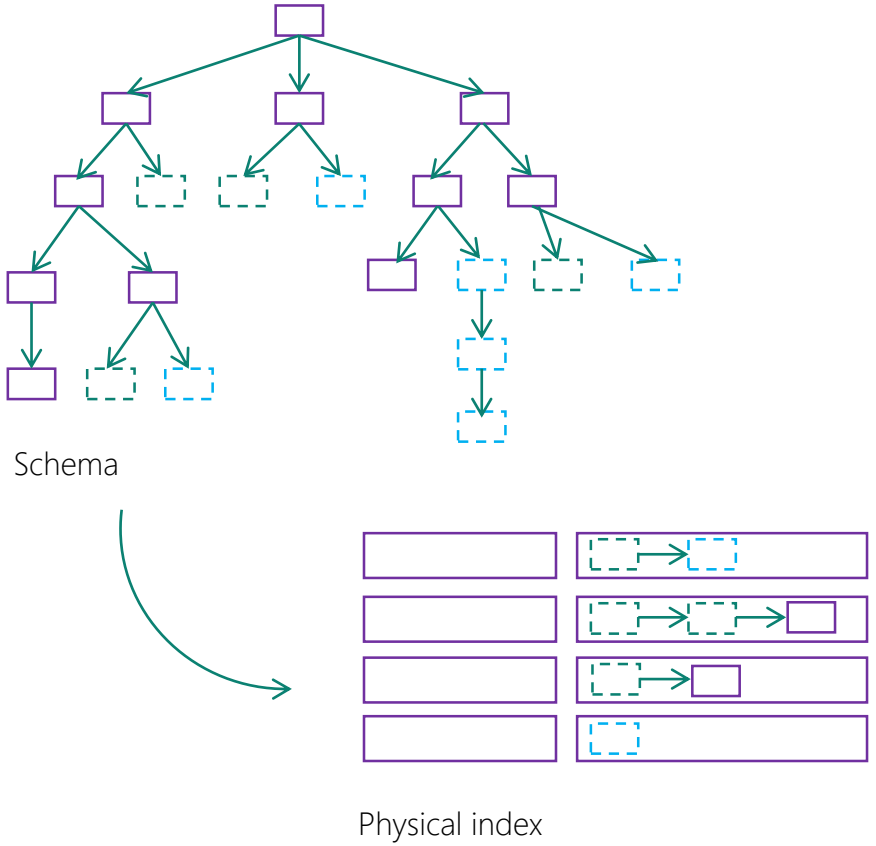
Database engine operates on Atom-Record-Sequence type system

All data models can be efficiently translated to ARS

Multi-model: Key-value, Document, and Graph

Multi-API: SQL (DocumentDB), MongoDB, Table, and Gremlin

More data-models and APIs to be added



Schema-agnostic, automatic indexing

At global scale, schema/index management is painful

Automatic and synchronous indexing

Hash, range, and geospatial

Works across every data model

Highly write-optimized database engine

Industry-leading, enterprise-grade SLAs

99.99% availability – even with a single region

Made possible with highly-redundant storage architecture

Guaranteed durability – writes are majority quorum committed

First and only service to offer SLAs on:

- Low-latency
- Consistency
- Throughput

Security & Compliance

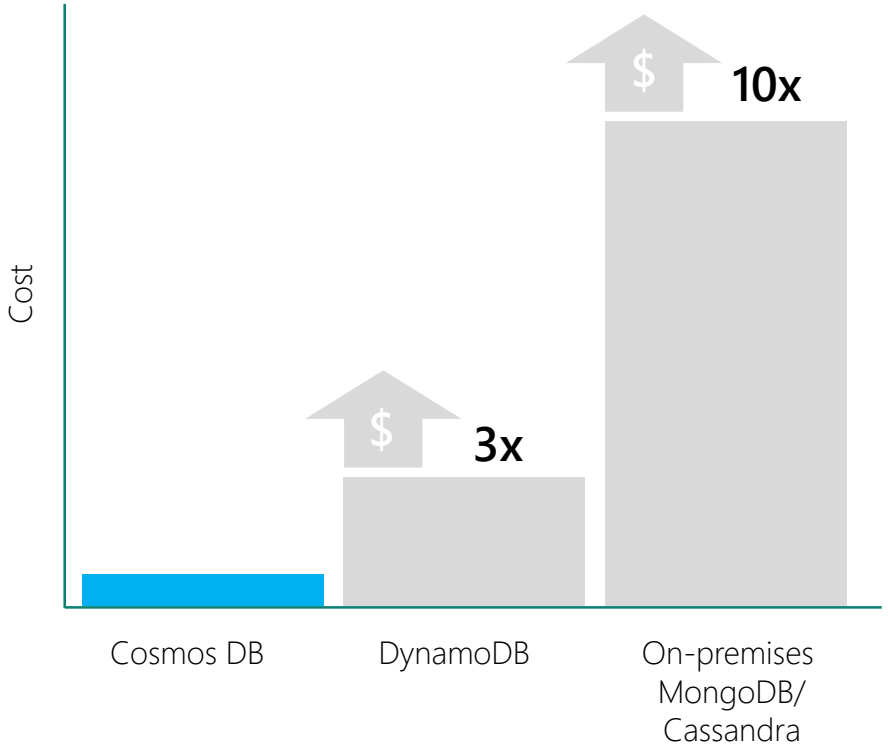
Always encrypted at rest and in motion

Fine grained “row level” authorization

Network security with IP firewall rules

Comprehensive Azure compliance certification:

- ISO 27001
- ISO 27018
- EUMC
- HIPAA
- PCI
- SOC1 and SOC2



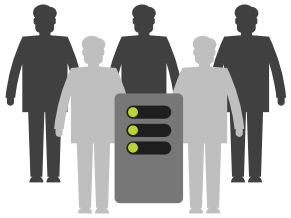
Lowest Total Cost of Ownership (TCO)

Deeply exploit cloud core properties and economies of scale

Significantly cheaper than DynamoDB, Cassandra, Cloud Spanner and MongoDB

Designed from the ground up as a multi-tenant service with end-to-end resource governance to provide performance isolation.

Fully managed as a service - no dev/ops expenses needed



Availability

- No need to setup replicas (local) other DC
- Failover - Assign priority to locations
- Automatic failover
- Manual Failover
- Distribution (low latency)
  - Add PreferredLocation to read the data locally
  - Choose right consistency
- SLA
- How to get Always-On

Capacity Management

- RUs
- # of operations ps
  - READ , WRITE everything
- For smoothing out spikes
  - # of operations pm
- Things to watch out for
  - Throttled operation –
    - Results in extra latency as SDK automatically retries
  - Identify throttled operations by logging them in appinsight

Performance

- Index
  - Auto index
  - Change index
- RU
- Latency
- Dig into the query
  - Query stats

Azure Cosmos DB

A globally distributed, massively scalable, multi-model database service

Table API

MongoDB API

Key-value

Column-family

Document

Graph

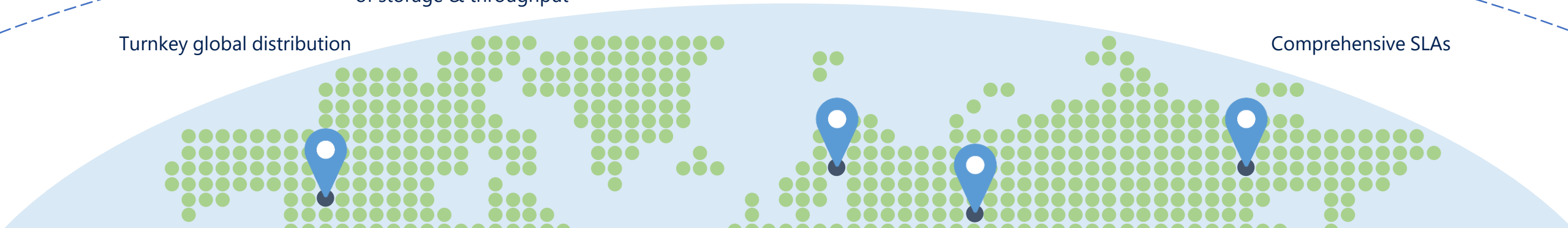
Guaranteed low latency at the 99th percentile

Elastic scale out
of storage & throughput

Five well-defined consistency models

Turnkey global distribution

Comprehensive SLAs



Next steps

When you need low-latency, global distribution, availability, elasticity

- If you have MongoDB and facing the pain of managing the cluster, perf, availability and distribution and need SLAs
  - Easy migration – just migrate the data and connection string change
- If you have Azure Table or Redis or Memcached usage
  - Easy migration for premium Table customers with more abilities(index/elasticity/distribution)
  - Redis/Memcached – data can be migrated and you can get lot more throughput/distribution
- If you have Cassandra – We can help you migrate to save you from managing clusters , doing jvm/os monitoring and yaml fiddling
- If you have DynamoDB – We can help you migrate with our migration tool and help you save money
- If you are trying to migrate from Relational DB for scale, throughput – offcourse we can help you do the right thing.

Pointers

- Getting Started
  - cosmosdb.com
  - TCO- https://azure.microsoft.com/en-us/updates/documentdb-total-cost-of-non-ownership-paper-available/
  - aka.ms/cosmosdb
  - aka.ms/cosmosdb-Tables
  - aka.ms/cosmosdb-Graph
  - aka.ms/cosmosdb-MongoDB
  - aka.ms/cosmosdb-DocumentDB
  - cosmosdb.com/capacityplanner
- Download
  - aka.ms/CosmosDB-emulator
- Re-visit Build session recordings on Channel 9.

Thank you!