Azure Cosmos DB

Rafat Sarosh
Principal Program Manager – Azure Cosmos DB

Microsoft

MODERN APPS FACE NEW CHALLENGES

Managing and syncing data distributed around the globe

Delivering highly-responsive, real-time personalization
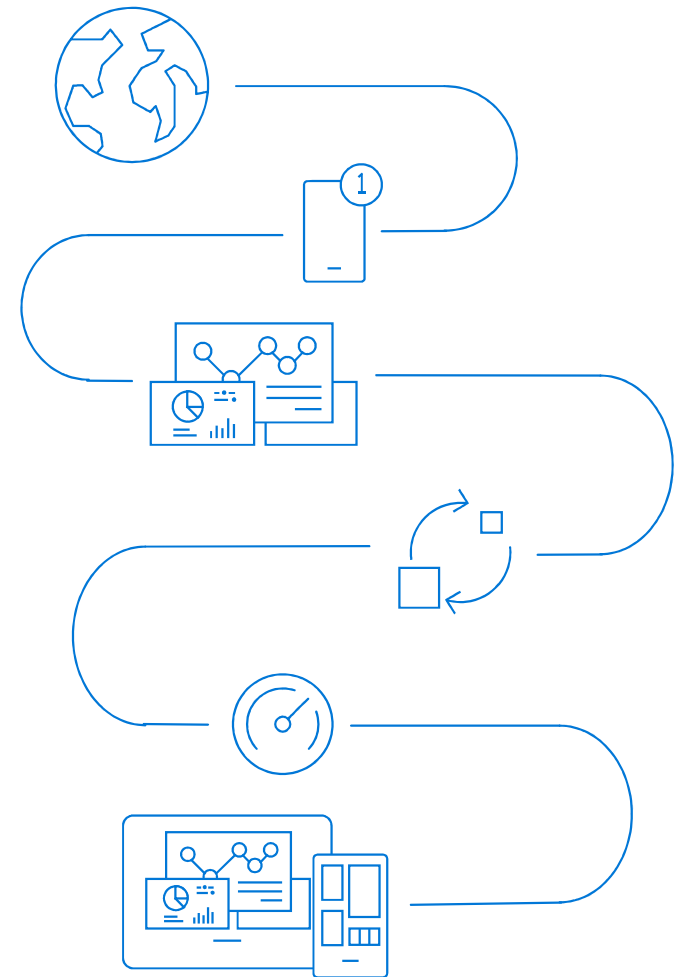
Processing and analyzing large, complex data

Scaling both throughput and storage based on global demand

Offering low-latency to global users

Modernizing existing apps and data –

Personalization, Recommendation, Social, Gammification.

AZURE COSMOS DB

A FULLY-MANAGED GLOBALLY DISTRIBUTED DATABASE SERVICE BUILT TO GUARANTEE
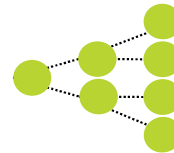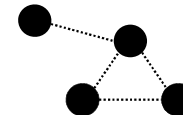EXTREMELY LOW LATENCY AND MASSIVE SCALE FOR MODERN APPS

SQL

cassandra

mongoDB

Gremlin
$G = (V, E)$

Azure Cosmos DB

… more coming soon

| Key-Value | Column-family | Documents | Graph |

A globally-distributed, multi-model database service

| Global distribution | Elastic scale out | Guaranteed low latency | Five consistency models | Comprehensive SLAs |

TURNKEY GLOBAL DISTRIBUTION

PUT YOUR DATA WHERE YOUR USERS ARE IN MINUTES

Automatically replicate all your data around the world, and across more regions than Amazon and Google combined.

- Available in all Azure regions

- Manual and automatic failover

- Automatic & synchronous multi-region replication
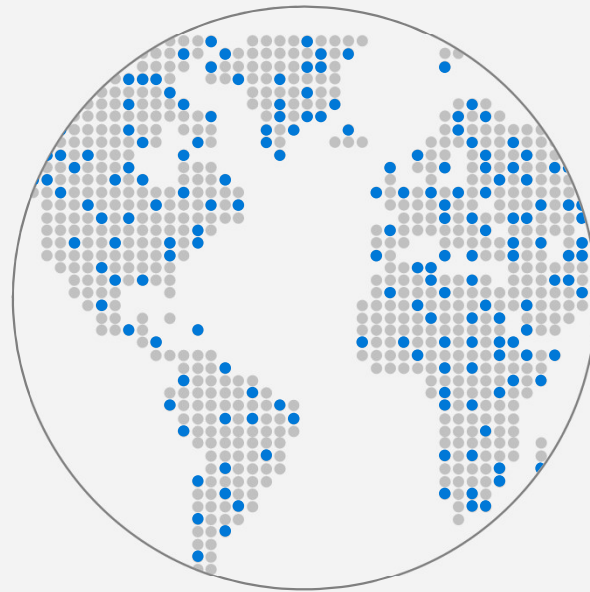
ELASTIC SCALE OUT OF STORAGE AND THROUGHPUT

SCALES AS YOUR APPS' NEEDS CHANGE

Independently and elastically scale storage and throughput across regions – even during unpredictable traffic bursts – with a database that adapts to your app's needs.

- Elastically scale throughput from 10 to 100s of millions of requests/sec across multiple regions

- Support for requests/sec for different workloads

- Pay only for the throughput and storage you need
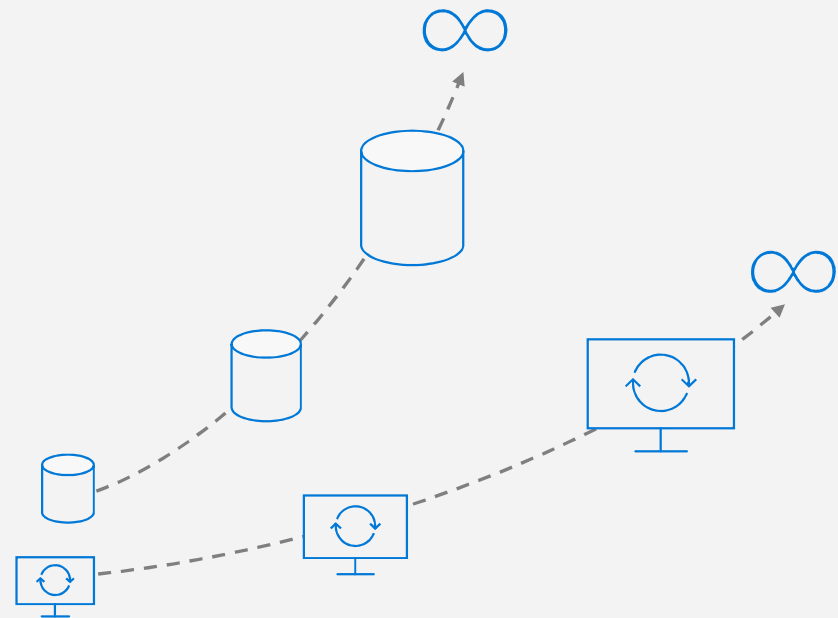
GUARANTEED LOW LATENCY

**PROVIDE USERS AROUND THE WORLD WITH
FAST ACCESS TO DATA**

Serve <10 ms read and <15 ms write requests at the
99th percentile from the region nearest to users,
while delivering data globally.
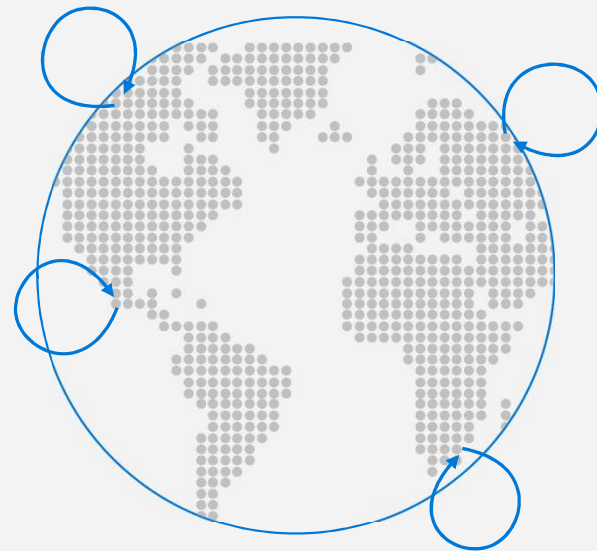
FIVE WELL-DEFINED CONSISTENCY MODELS

CHOOSE THE BEST CONSISTENCY MODEL FOR YOUR APP

Offers five consistency models

Provides control over performance-consistency tradeoffs, backed by comprehensive SLAs.

An intuitive programming model offering low latency and high availability for your planet-scale app.

Strong   Bounded-stateless   Session   Consistent prefix   Eventual

MULTIPLE DATA MODELS AND APIs

USE THE MODEL THAT FITS YOUR REQUIREMENTS, AND THE APIS, TOOLS, AND FRAMEWORKS YOU PREFER

Cosmos DB offers a multitude of APIs to access and query data including, SQL, various popular OSS APIs, and native support for NoSQL workloads.

Use key-value, tabular, graph, and document data

Data is automatically indexed, with no schema or secondary indexes required

Blazing fast queries with no lag

Table API

MongoDB

Gremlin

SQL

cassandra

JS

Key-value

Column-family

Document

Graph
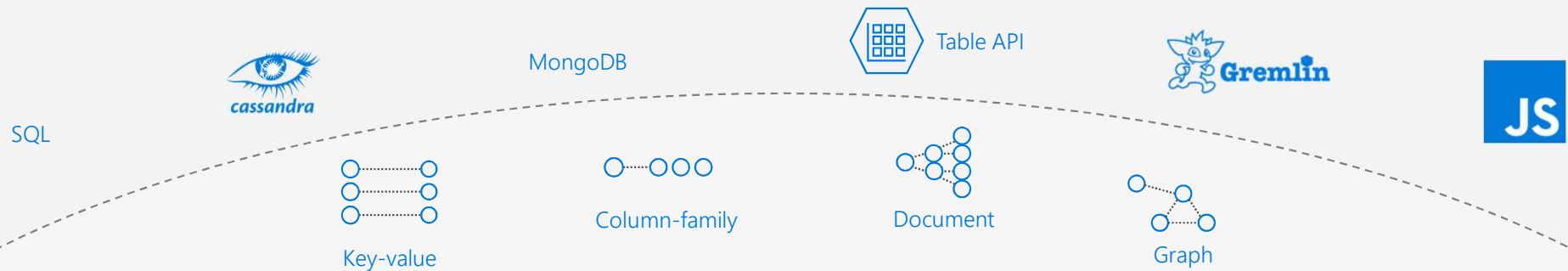
HANDLE ANY DATA WITH NO SCHEMA OR INDEXING REQUIRED

Azure Cosmos DB's schema-less service automatically indexes all your data, regardless of the data model, to delivery blazing fast queries.
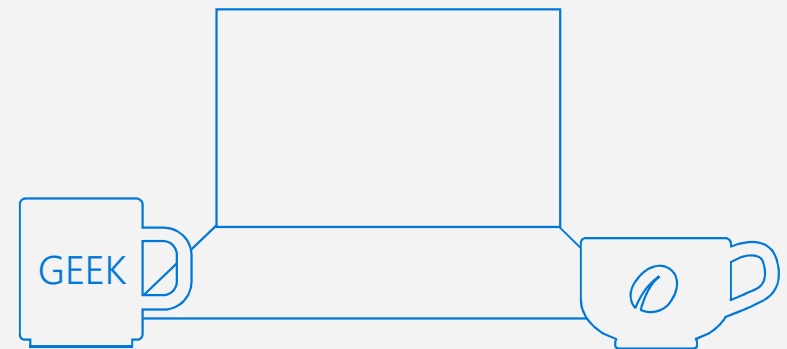
- Automatic index management

- Synchronous auto-indexing

- No schemas or secondary indices needed

- Works across every data model

Item	Color	Microwave safe	Liquid capacity	CPU	Memory	Storage
Geek mug	Graphite	Yes	16oz	???	???	???
Coffee Bean mug	Tan	No	12oz	???	???	???
Surface book	Gray	???	???	3.4 GHz Intel Skylake Core i7-6600U	16GB	1 TB SSD

COMPREHENSIVE SLAs

RUN YOUR APP ON WORLD-CLASS INFRASTRUCTURE

Azure Cosmos DB is the only service with financially-backed SLAs for millisecond latency at the 99th percentile, 99.999% HA and guaranteed throughput and consistency

Latency

<10 ms
99th percentile

HA

99.999%

Throughput

Guaranteed

Consistency

Guaranteed
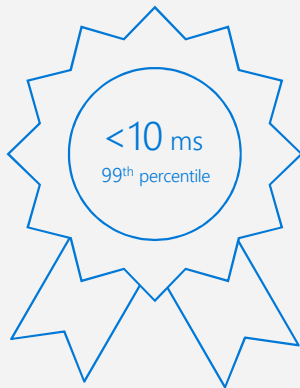
TRUST YOUR DATA TO INDUSTRY-LEADING SECURITY & COMPLIANCE

Azure is the world's most trusted cloud, with more certifications than any other cloud provider.

- Enterprise grade security
- Encryption at Rest
- Encryption is enabled automatically by default
- Comprehensive Azure compliance certification

TOP 10 REASONS WHY CUSTOMERS USE AZURE COSMOS DB

The 1st and only database with global distribution turnkey capability

Deliver massive storage/throughput scalability database

Provides guaranteed single digit millisecond latency at 99th percentile worldwide

Natively supports different types of data at massive scale

Boasts 5 well-defined consistency models to pick the right consistency/latency/throughput tradeoff

Enables mission critical intelligent applications

Gives high flexibility to optimize for speed and cost

Tackles big data workloads with high availability and reliability

Provides multi-tenancy and enterprise-grade security

Naturally analytics-ready and perfect for event-driven architectures

CUSTOMERS POWERING THEIR APPS WITH AZURE COSMOS DB

- 7000+ Azure Cosmos DB customers

- NoSQL is growing at 35.1% YoY

- 20% of customers have migrated Apps from Mongo DB to Cosmos DB

COSMOS DB WINS 2018 TECHNOLOGY OF THE YEAR AWARDS

"Do you need a distributed NoSQL database with a choice of APIs and consistency models? That would be Microsoft's Azure Cosmos DB."—InfoWorld, [Technology of the Year 2018: The best hardware, software, and cloud services](#)

MARKET LEADERSHIP AND GAINING POPULARITY

DB-Engines Ranking of Microsoft Azure Cosmos DB

February 2015
Microsoft Azure Cosmos DB: **0.585**

Overall
31 +57

© February 2018, DB-Engines.com

Multi-model 1 +5   **Graph** 2 +2   **Key Value** 3 +7   **Wide Column** 3 +1   **Document** 5 +4
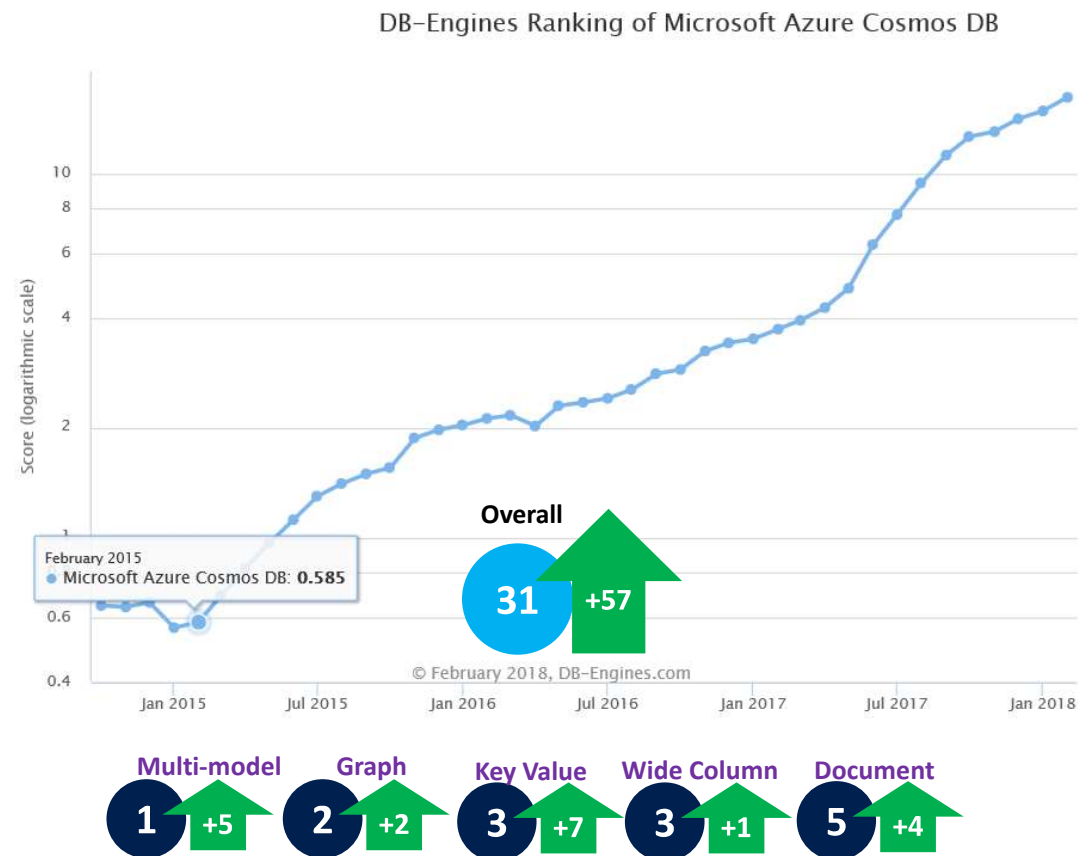
DB-ENGINES

WHAT'S ANNOUNCED IN BUILD

V-NET

BULK EXECUTOR

RX JAVA LIBRARY

SHARED THROUGHPUT
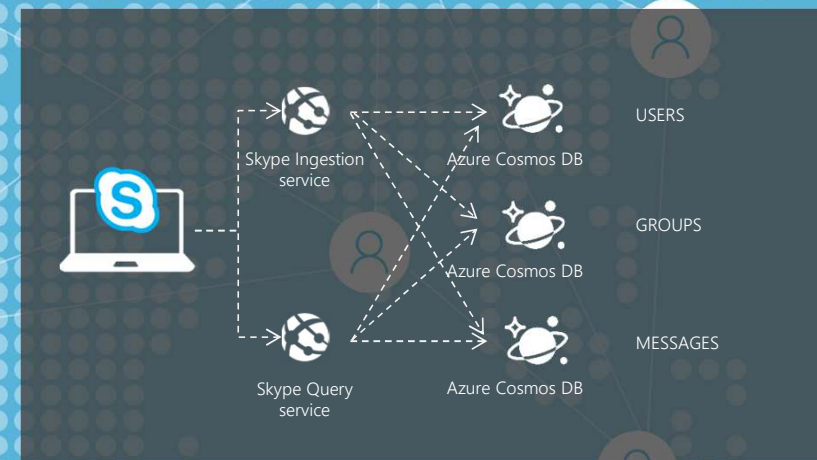
MULTI-MASTER

CUSTOMER STORIES

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 Ingestion service

Skype Query service

Azure Cosmos DB — USERS

Azure Cosmos DB — GROUPS

Azure Cosmos DB — MESSAGES

skype™

6TB
User data

44TB
Message data

TOYOTA STEERS IOT TELEMETRICS TOWARD THE FUTURE

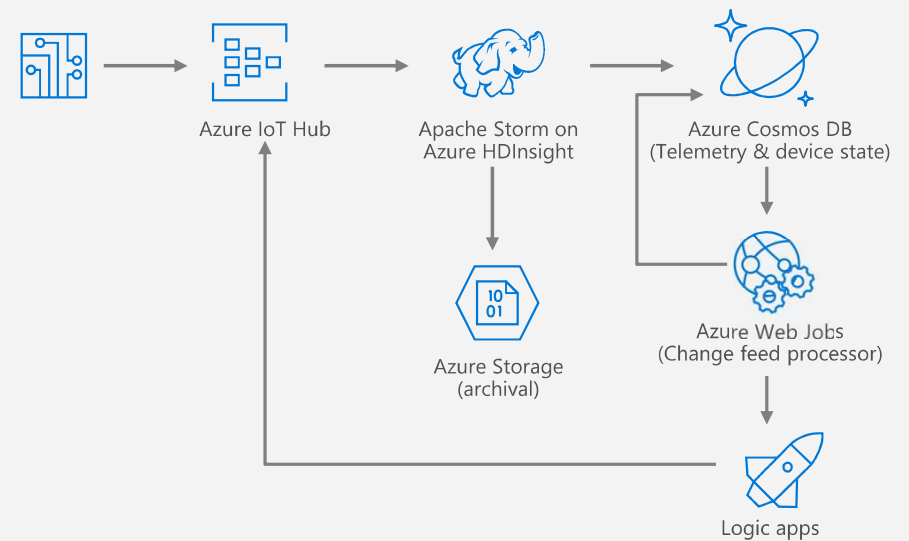Diverse and unpredictable IoT sensor workloads require a responsive data platform

- Real-time vehicle diagnostics
- Instant elastic scaling
- No loss in ingestion or query performance

8TB
Vehicle Telemetry

250K
Lexus Cars

Azure IoT Hub

Apache Storm on Azure HDInsight

Azure Cosmos DB (Telemetry & device state)

Azure Storage (archival)

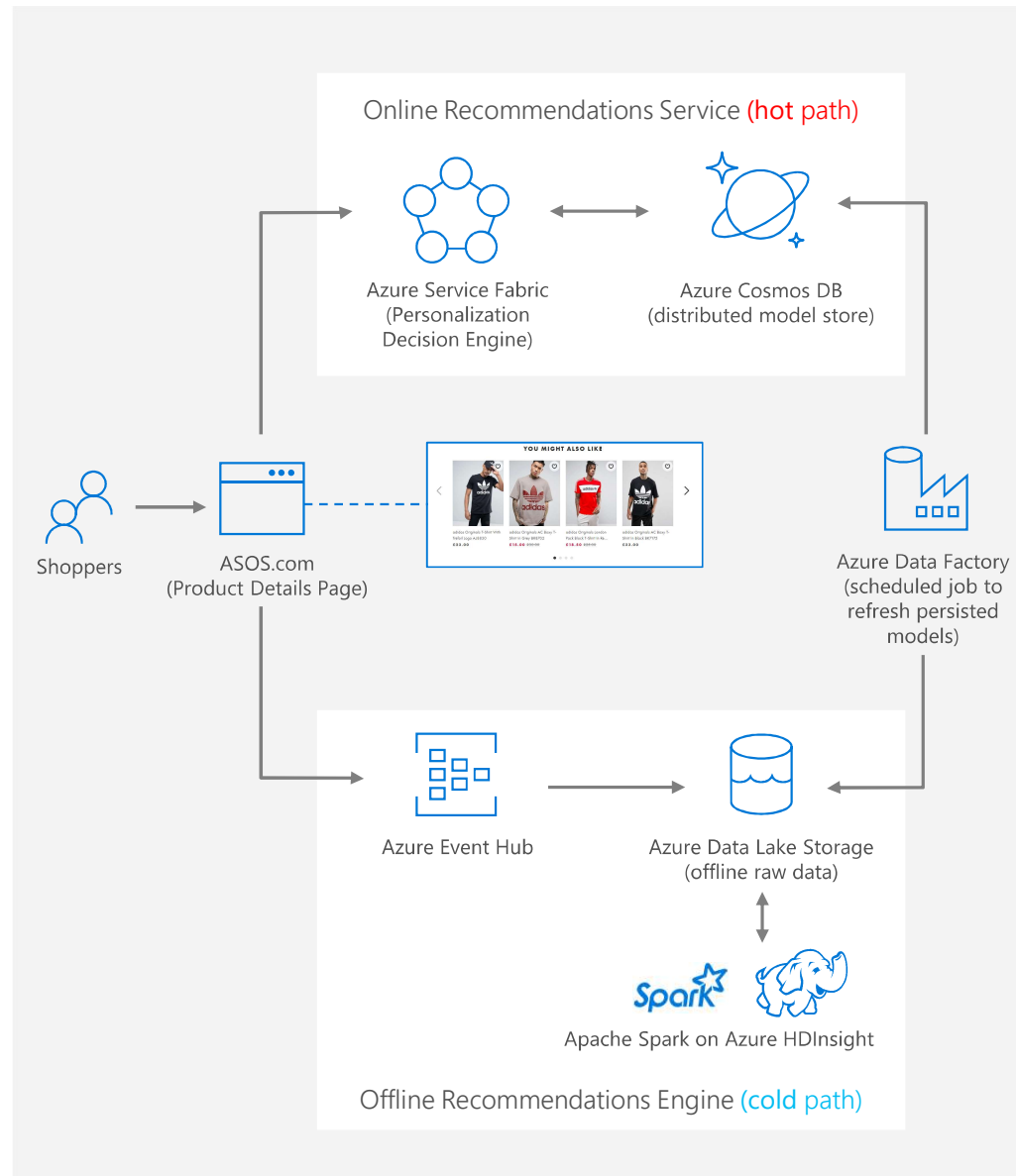Azure Web Jobs (Change feed processor)

Logic apps

ASOS DELIVERS PERSONALIZATION TO 15M CUSTOMERS

Personalized shopping experiences and real-time order updates win with millennial shoppers

- Machine learning models generate real-time recommendations across 85,000 products

- Product analysis in milliseconds

- Low-latency ensures fast page loads

Online fashion retailer **ASOS achieves +25% annual growth** with more responsive apps that deliver personalized customer experiences

asos

Online Recommendations Service (hot path)

Azure Service Fabric (Personalization Decision Engine)

Azure Cosmos DB (distributed model store)

Shoppers

ASOS.com (Product Details Page)

YOU MIGHT ALSO LIKE

Azure Data Factory (scheduled job to refresh persisted models)

Azure Event Hub

Azure Data Lake Storage (offline raw data)

Spark

Apache Spark on Azure HDInsight

Offline Recommendations Engine (cold path)

DATA DISTRIBUTED AND AVAILABLE GLOBALLY

Put your data where your users are to give real-time access and uninterrupted service to customers anywhere in the world.

- Turnkey global data replication across all Azure regions
- Guaranteed low-latency experience for global users
- Resiliency for high availability and disaster recovery

Browser

Azure Traffic Manager

Azure region A
<10 ms

Azure region B
<10 ms

Azure region C
<10 ms

Azure Cosmos DB
(app + session state)

Globally distributed
across regions

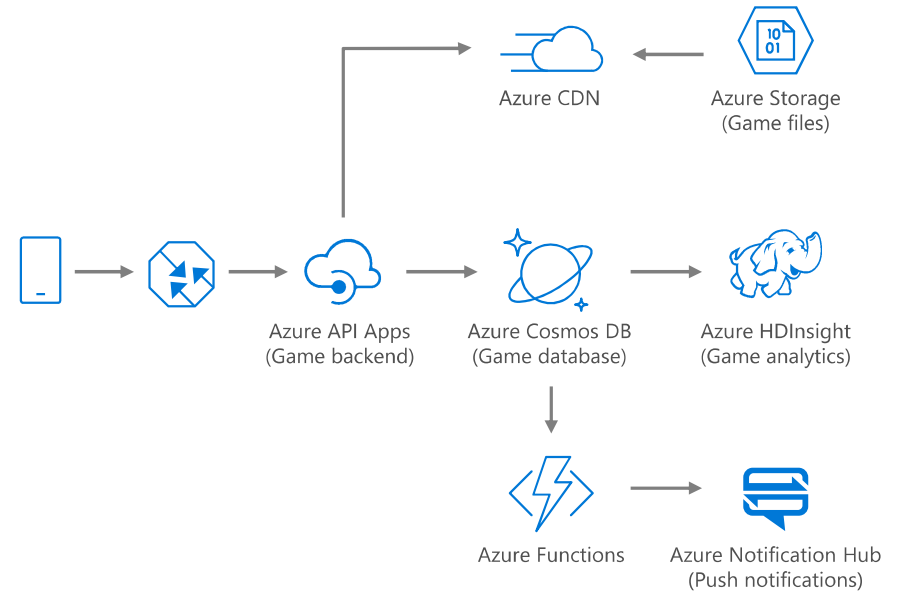NEXT GAMES
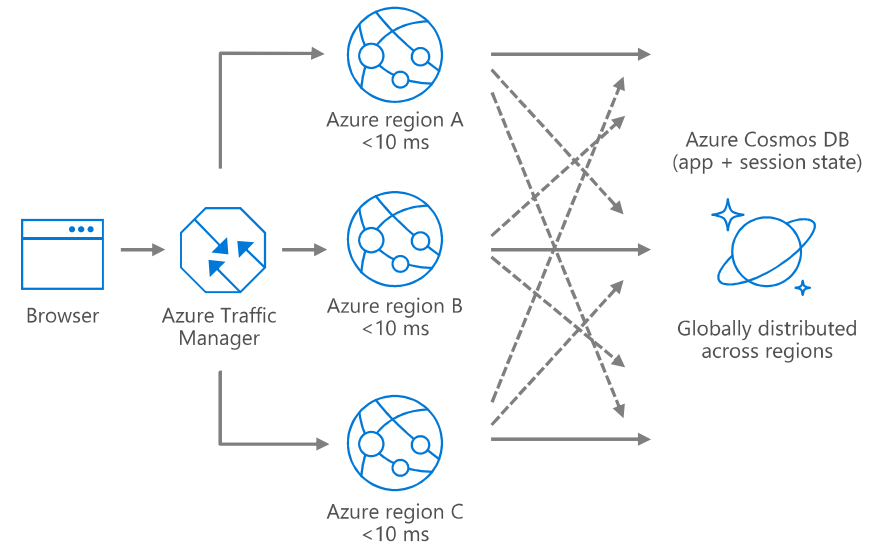
#1
iOS App Store

1B
Daily Queries

1M
Peak Active

Domino's

IDEAL FOR GAMING, IOT AND ECOMMERCE

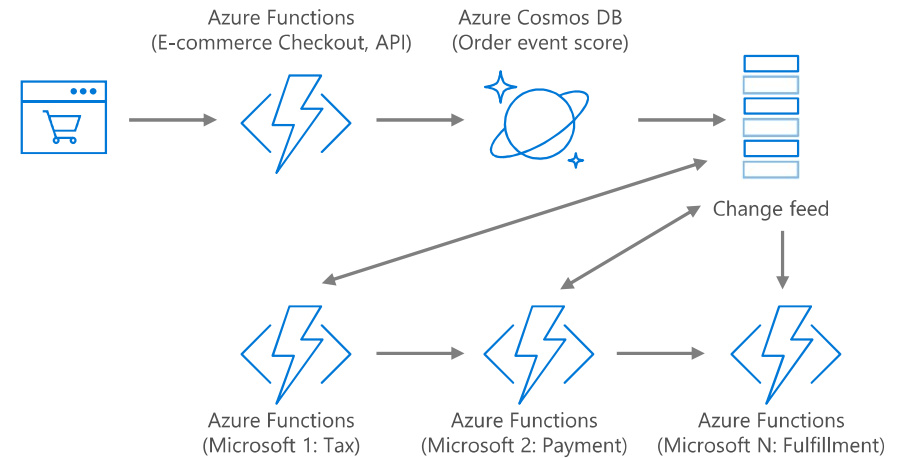Maintain service quality during high-traffic periods requiring massive scale and performance.

- Instant, elastic scaling handles traffic bursts

- Uninterrupted global user experience

- Low-latency data access and processing for large and changing user bases

- High availability across multiple data centers

Azure CDN

Azure Storage
(Game files)

Azure API Apps
(Game backend)

Azure Cosmos DB
(Game database)

Azure HDInsight
(Game analytics)

Azure Functions

Azure Notification Hub
(Push notifications)

NEXT GAMES

Domino's

jet

XBOX LIVE

SIMPLIFIED DEVELOPMENT WITH SERVERLESS ARCHITECTURE

Experience decreased time-to-market, enhanced scalability, and freedom from framework management with event-driven micro-services.
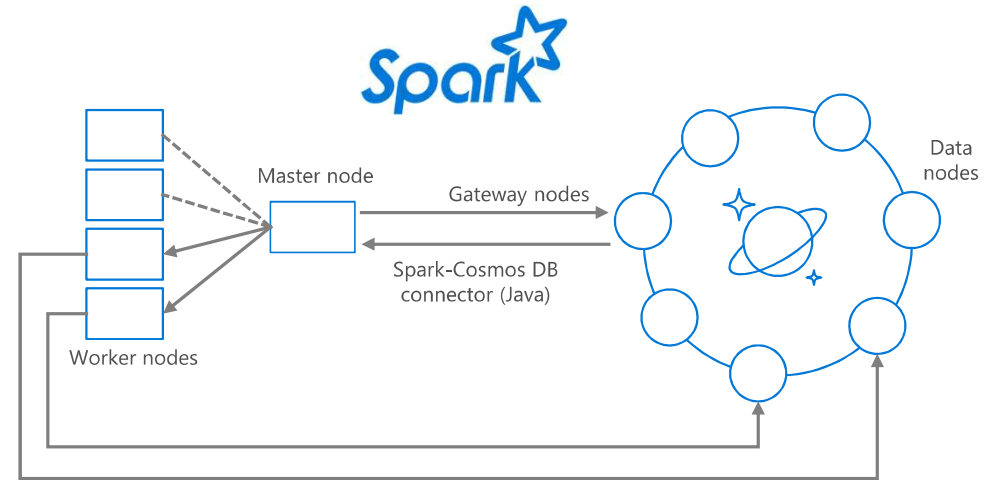
- Seamless handling of any data output or volume

- Data made available immediately, and indexed automatically

- High writes per second, with stable ingestion and query performance

- Real-time, resilient change feeds logged forever and always accessible

- Native integration with Azure Functions

Azure Functions (E-commerce Checkout, API) → Azure Cosmos DB (Order event score) → Change feed → Azure Functions (Microsoft 1: Tax) → Azure Functions (Microsoft 2: Payment) → Azure Functions (Microsoft N: Fulfillment)

asos jet

RUN SPARK OVER OPERATIONAL DATA

Accelerate analysis of fast-changing, high-volume, global data.

- Real-time big data processing across any data model

- Machine learning at scale over globally-distributed data

- Speeds analytical queries with automatic indexing and push-down predicate filtering

- Native integration with Spark Connector

QUESTIONS?