



Data Technologies Overview (with Azure!)





Hello!

Charmaine
Lee

Twitter: @charmaine_klee

Anthony
Chu

Twitter: @nthonychu

Gora
Kong

Twitter: @gorakong





Types of Data Stores

Relational

SQL Database

SQL Data

Warehouse

MySQL

PostgreSQL

Non-Relational

Table Storage

Redis

Cosmos DB

Search

Big Data

Blob Storage

Data Lake Storage

HDInsight

No-DB?

Service Fabric



Why different data stores?

- Shift towards microservices
- Domain Driven Design
- “Best tool for the job”





Relational Databases

SQL Database



SQL Server
on premises

You



SQL Server

Operating System

Hardware



SQL Database



SQL Server
on VM

You

Azure

SQL Server

Operating System

Hardware



SQL Database



SQL Database

Azure



SQL Server

Operating System

Hardware





SQL Database

- Starts at \$6/month
- Scale on the fly
- 99.99% uptime SLA





SQL Database

Point-in-time restore

Restore

achu-test

* Database name

achu-test_2017-08-24T02-40Z

Point-in-time

Long-term

Oldest restore point

2017-08-23 22:57 UTC

Restore point (UTC)

2017-08-24

2:40:00 AM

SQL Database



Geo-replication

- Readable replicas
- Manual failover*





SQL Database

- Encryption at-rest and in-flight
- Row-level security
- Auditing



SQL Database



Dynamic data
masking

SQL users excluded from masking (administrators are always excluded) ⓘ

SQL users excluded from masking (administrators are always excluded) ✓

Recommended fields to mask

SCHEMA	TABLE	COLUMN	
SalesLT	Address	AddressID	ADD MASK
SalesLT	Address	AddressLine1	ADD MASK
SalesLT	Address	AddressLine2	ADD MASK
SalesLT	Customer	LastName	ADD MASK
SalesLT	Customer	EmailAddress	ADD MASK
SalesLT	Customer	Phone	ADD MASK
SalesLT	Customer	PasswordHash	ADD MASK
SalesLT	Customer	PasswordSalt	ADD MASK



SQL Database

Automatic
threat detection

The interface displays security alerts for a SQL database. At the top, there are two columns for severity levels: 'HIGH' with a count of 2 and 'MEDIUM' with a count of 3. The date range is set from '1 Feb' to '1 Mar'. Below this is a table of alerts. The first two rows, both labeled 'Potential SQL Injection' with a count of 3, are highlighted with a red border. The other three rows show different types of alerts with counts of 1.

	DESCRIPTION	COUNT
HIGH 2	MEDIUM 3	
NEW	Potential SQL Injection	3
NEW	Potential SQL Injection	3
NEW	A possible vulnerability to SQL Injection	1
NEW	Logon by an unfamiliar principal	1
NEW	Someone logged on to your SQL server from an an unusual loc...	1



SQL Database

- Familiar tools
 - SQL Server Mgmt Studio
 - Visual Studio & VSCode
 - In-portal query



SQL Database




Query Performance Insight






SQL Database







Database Advisor

 Recommendations
IndexAdvisor_test_2

 Automate  View  Feedback
discard...

Recommendations

ACTION	RECOMMENDATION DESCRIPTION	IMPACT
 PARAMETERIZE QUERIES (PREVIEW)	Scope: Entire database Reason: Non-parameterized queries are causing performance issues	HIGH IMPACT
 FIX SCHEMA ISSUES (PREVIEW)	Error code: 201 Error message: Procedure or function 'sp_indexes' expects parameter '@table_s	HIGH IMPACT
 CREATE INDEX	Table: [test_table_0.790915] Indexed columns: [index_1],[index_2],[index_3]	MEDIUM IMPACT
 DROP INDEX (PREVIEW)	Index name: IR_[test_schema].[test_table_0.464078]_81D4FA6BE11E402E5530 Reason: Duplicate index	LOW IMPACT



SQL Database

SQLDB Managed Instance

- Addresses gap between SQLDB and SQL Server
- Migration service
- Private preview





SQL Database

Stretch Database

- SQL Server 2016
- Transparently migrate cold data to the cloud
- Query “stretched” data





SQL Data Warehouse

A SQL-based, fully-managed, petabyte-scale cloud data warehouse.

Key Features:

- ◇ Scalability
- ◇ Security & Reliability
- ◇ Integrated





SQL Data Warehouse ARCHITECTURE

- Massively Parallel Processing (MPP) engine
- PolyBase
Query Azure Blob Storage & Data Lake Store
Import/Export data directly





SQL Data Warehouse

SCALABILITY

- Decoupled storage & compute
Scale up/down, pay for only what you use
- Scale based on Data Warehouse Units (DWUs)
Used computational resources
higher DWUs → higher performance → higher costs
- Pause on demand





SQL Data Warehouse

SECURITY

- Automatic threat detection
- Authentication

RELIABILITY

- 99.9% uptime
- Disaster recovery

Data stored using geo-redundant blobs





SQL Data Warehouse INTEGRATED

- Use in conjunction with SQL Server tools
SSDT, SSIS, SSAS, SSRS
- Azure Data Factory, Machine Learning,
Stream Analytics
- Power BI
- Cortana Analytics Suite
For building analytic solutions





MySQL / PostgreSQL

- High availability
- High security
- Elastically scales





MySQL

- Relational database service
- Create DB in minutes
- Predictable & easy scaling





MySQL

- Easy DB admin & management
 - DB patching
 - Backups
 - Monitoring
 - Security





PostgreSQL

- Object-RDBMS
- Versions 9.5 & 9.6
- Can export current data to Azure






Relational Databases

- SQL DB
- SQL Data Warehouse
- MySQL
- PostgreSQL



A decorative graphic on the left side of the slide consists of a cluster of hexagons in various shades of blue and cyan. Some hexagons contain white icons: a lightbulb, a thumbs-up, a network of nodes, a smartphone, a magnifying glass, a gear, and a speech bubble. A large, solid cyan hexagon is positioned to the left of the main text.

Non-relational Databases (NoSQL)



Table Storage

- Cost-effective key-value store
- Semi-structured data
- Flexible schema
- Single index
- Fast
- 7 cents per GB
- 0.36 cents per 100k transactions



Table Storage

Vehicles

Query

Import

Export

Add

Edit

Select all

Column Options

Delete

Refresh

PartitionKey	RowKey	Timestamp	Color	Type	Wheels
car	1	2017-08-25T06:34:04.858Z	Red		4
boat	2	2017-08-25T06:33:33.375Z	White	Sailboat	
car	3	2017-08-25T06:35:37.378Z	Blue		4
boat	4	2017-08-25T06:36:04.007Z	Grey	Speed Boat	



Redis Cache

- Fully-managed Redis as a service
- Highly secure: Uses Redis authentication and supports SSL connections
- Scales up: Cache is available in sizes up to 530 GB (Premium)





Redis Cache

- 3 tiers:

Basic: Single node, multiple sizes

High-availability SLA

Standard: Two-node primary/secondary

Premium: All of the Standard features + more

- Data persistence (Premium only)
→ data persists even with hardware failure



Why Redis Cache?



- ## High Performance

Low latency, high throughput Redis engine

- ## Great features

Support for high performance data types, transactions, Lua scripting, keys with limited TTL, etc.

- ## Easy to use and manage

Monitor health & performance of cache through Azure portal





Monitoring a Redis Cache

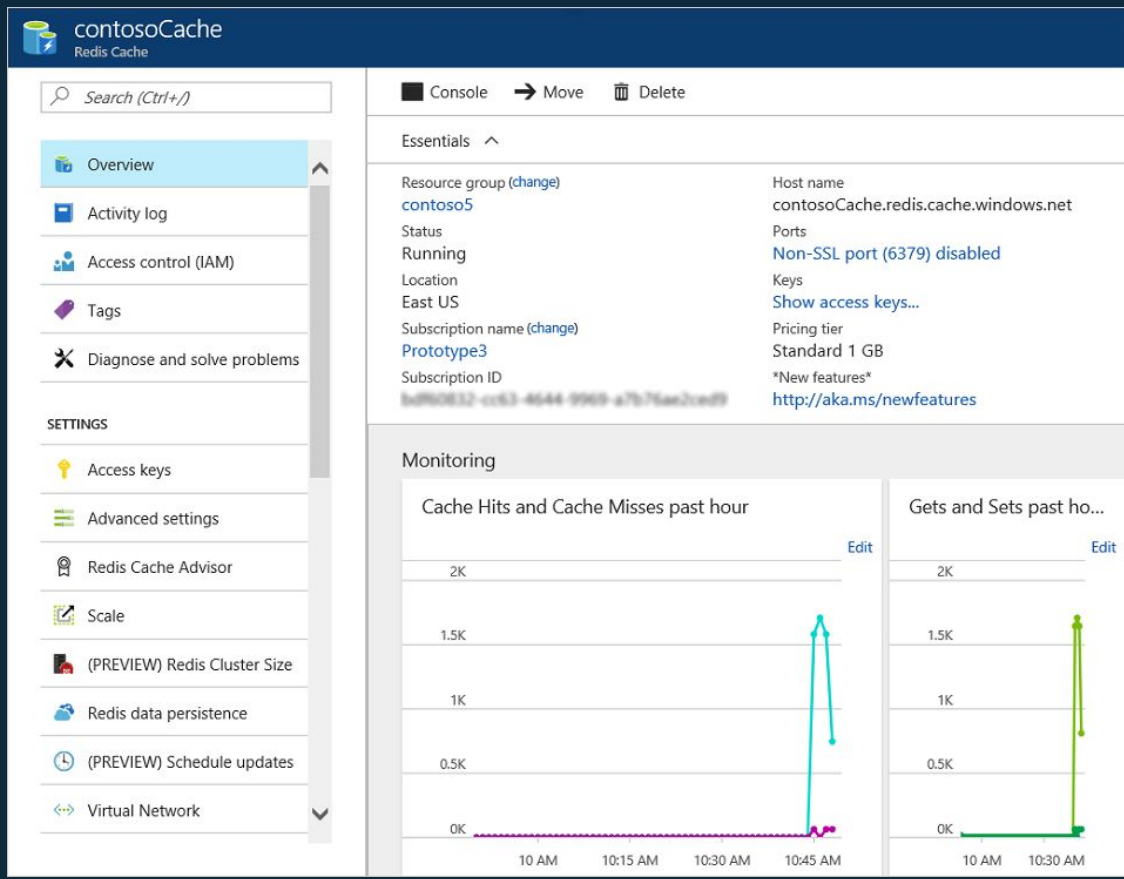
Performance metrics that can be monitored:

- ◇ Uptime and Availability
- ◇ Cache Hits & Misses
- ◇ Get/Set Commands
- ◇ Total Operations
- ◇ Evicted & Expired Keys
- ◇ Memory, Bandwidth & CPU utilization

Full support for customization of monitored parameters, & alerts



Monitoring a Redis Cache





Cosmos DB

- One step from DocumentDB, one giant leap for databases
- Globally distributed, multi-model database service
- Low latency, high availability & consistency



Turnkey Global Distribution

- Distribute data with just one click!

1



Azure Regions

A decorative pattern of hexagons in various shades of blue and teal. Some hexagons contain icons: a lightbulb, a thumbs up, a smartphone, a magnifying glass, and a gear. A large teal hexagon in the center-left contains the number '2'.

2

Elastic Scaling & High Availability

- Scale throughput and storage at anytime and anywhere in the world
- App is always “on”

Consistency

- Can choose between 5 well-defined consistency models

3





Consistency

Consistency Levels and guarantees

Consistency Level	Guarantees
Strong	Linearizability
Bounded Staleness	Consistent Prefix. Reads lag behind writes by k prefixes or t interval
Session	Consistent Prefix. Monotonic reads, monotonic writes, read-your-writes, write-follows-reads
Consistent Prefix	Updates returned are some prefix of all the updates, with no gaps
Eventual	Out of order reads

A decorative pattern of hexagons in various shades of blue and cyan. Some hexagons contain icons: a lightbulb, a thumbs-up, a network node, a smartphone, a magnifying glass, a gear, and a speech bubble.

4

No database schema/index
management



5

Data models

- Graph
- Key-value
- Table

APIs

- DocumentDB API
- MongoDB API
- Table API
- Graph (Gremlin) API



Azure Search

- ◇ **Cloud search service
for web and mobile
app development**





Azure Search Features

- Full text search
 - Searching in a full-text db instead of metadata
- Data integration
 - Azure Search Indexes accept data from any source
- 56 languages
 - Handles language-specific linguistics





Azure Search Features

- Monitoring and reporting
 - Search traffic analytics
- Geospatial support
 - Proximity-based results
- UX features
 - Search suggestions, filters, sorting, etc.





Non-Relational Databases

- Table Storage
- Redis
- Cosmos DB
- Search



A cluster of hexagons in various shades of blue and cyan, some solid and some outlined, arranged in a geometric pattern in the top-left corner.

Big Data

- Store everything
- Unstructured data
- Batch query and process



A cluster of hexagons in various shades of blue and cyan, some solid and some outlined, arranged in a geometric pattern in the top-left corner.

Blob Storage

- Cost-effective object storage
- Different access tiers
- Massively-scalable
- Geo-redundant





Data Lake Storage

- “No limits”
 - Petabyte-size files
 - Trillions of object
 - Scalable throughput
- Always encrypted, RBAC
- HDFS-compatible





Data Lake Analytics

- On-demand analytics jobs
- Query from
 - Data Lake Store
 - Blob Storage
 - SQLDB and SQLDW
- U-SQL
 - C#
 - Python / R





HDInsight

- Managed Hadoop
- Query Data Lake Store
- Spark, Storm, Kafka, HBase





Service Fabric

- Store data with the application
- Stateful Services
- Stateful Actors





Summary

Relational

SQL Database

SQL Data

Warehouse

MySQL

PostgreSQL

Non-Relational

Table Storage

Redis

Cosmos DB

Search

Big Data

Blob Storage

Data Lake Storage

HDInsight

No-DB?

Service Fabric

SQLSaturday Sponsors!

Titanium
& Global Partner



Gold



Silver



Bronze



Without the generosity of these sponsors, this event would not be possible! Please, stop by the vendor booths and thank them.





Thanks!

**Charmaine
Lee**

Twitter: @charmaine_klee

**Anthony
Chu**

Twitter: @nthonychu

**Gora
Kong**

Twitter: @gorakong

