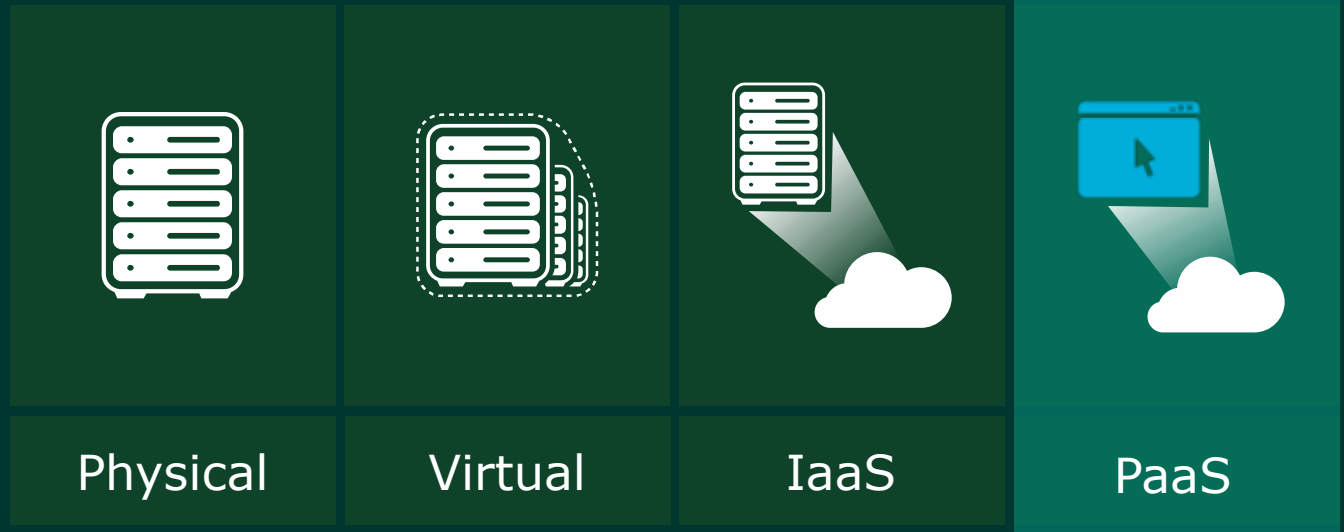
Azure Data Overview

SQL Database

Microsoft Relational Storage Options

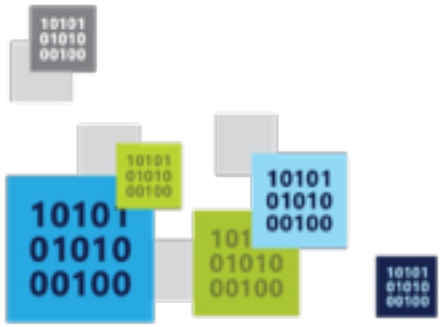
A Continuous offering



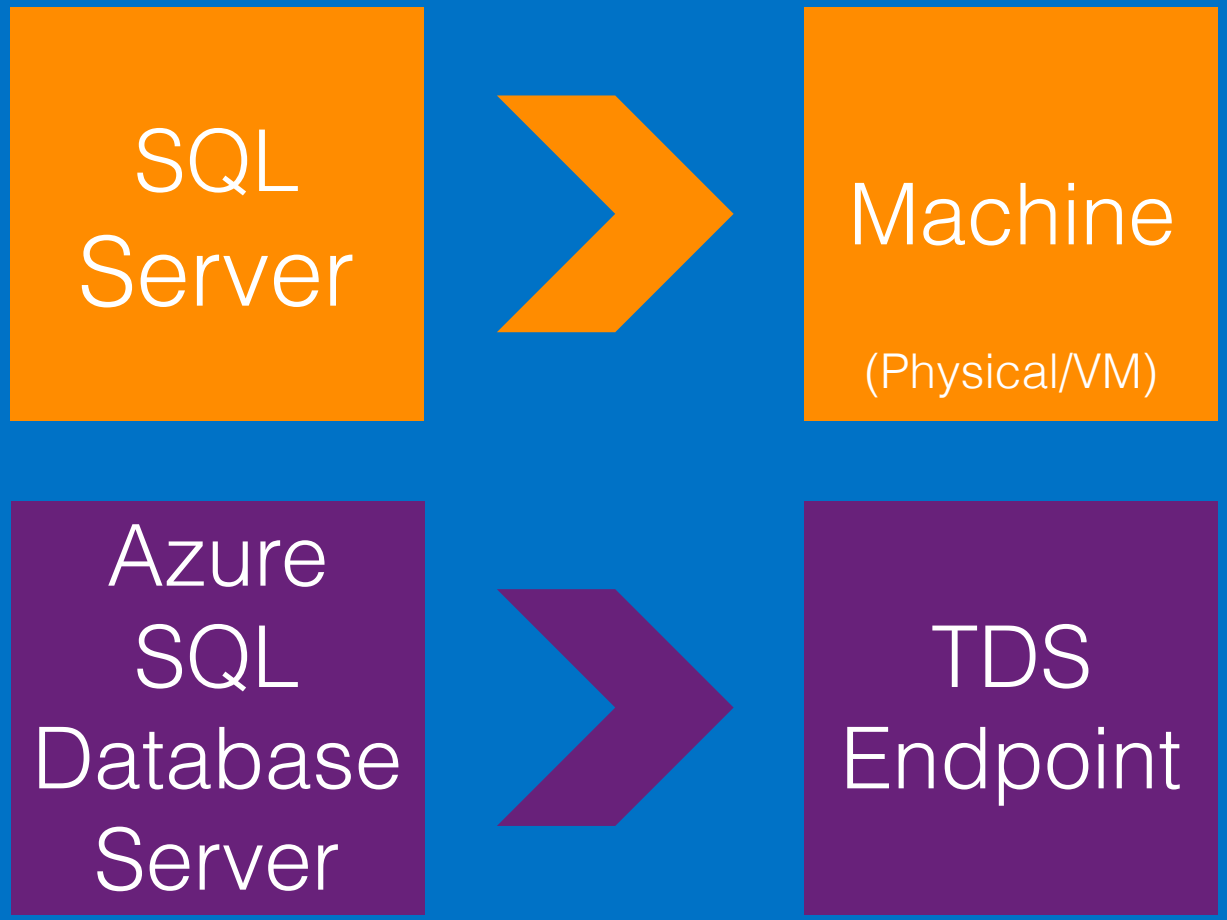
From private to public Cloud

Architecture

Starting with the basics

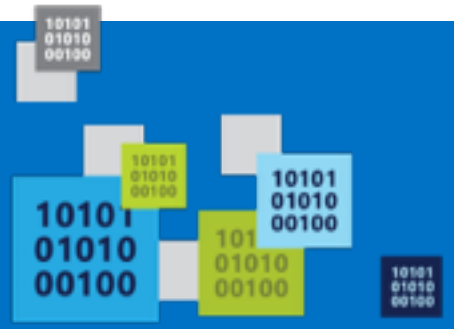


A Server is not a machine



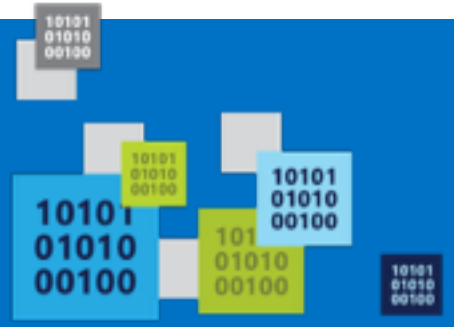
SQL Database The Basics

- SQL Server database technology “as a Service”
- Fully Managed
- Enterprise-ready with automatic support for HA, DR, Backups, replication and more



SQL Database
The Basics

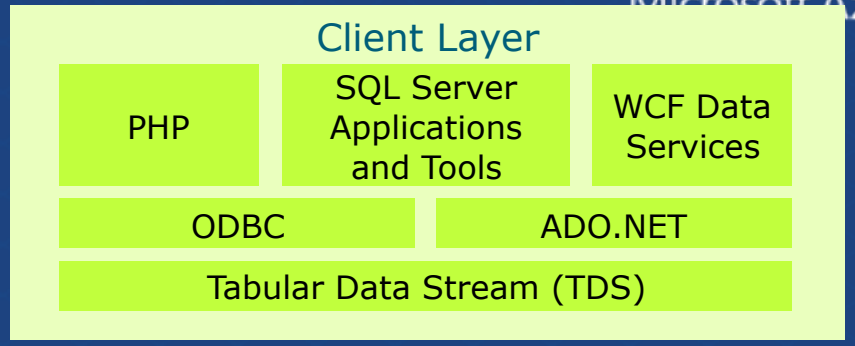
- Scale out with ElasticScale
- Built-in regional database replicas for additional protection
- Uptime SLA of 99.99%



How It Works – Architecture of the Service

Microsoft Azure

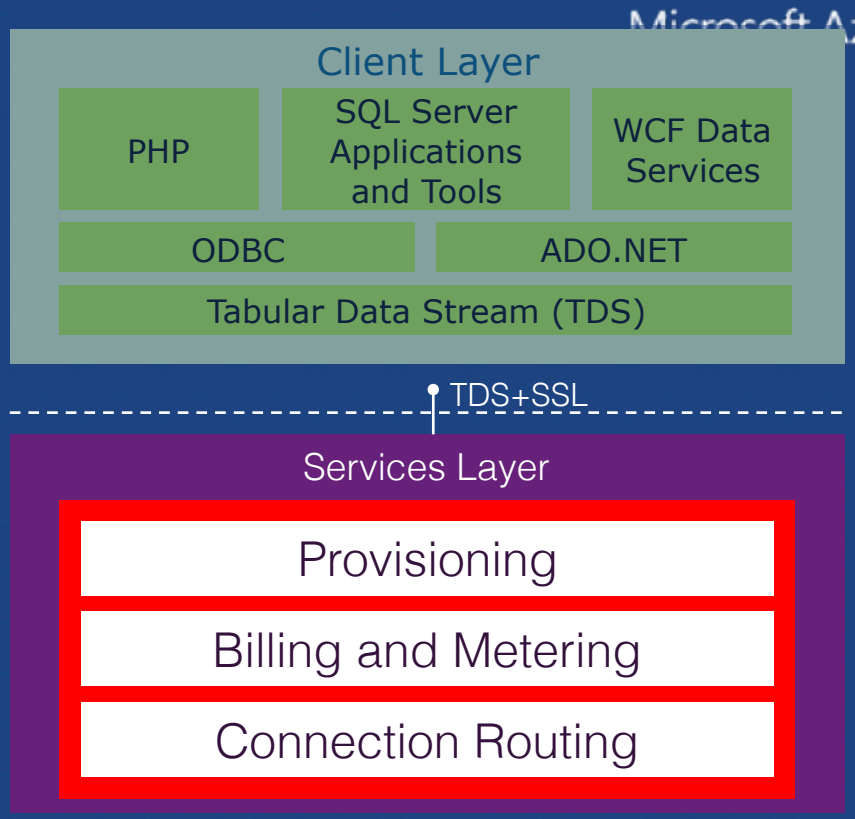
Applications communicate directly with SQL Database using TDS.



Microsoft

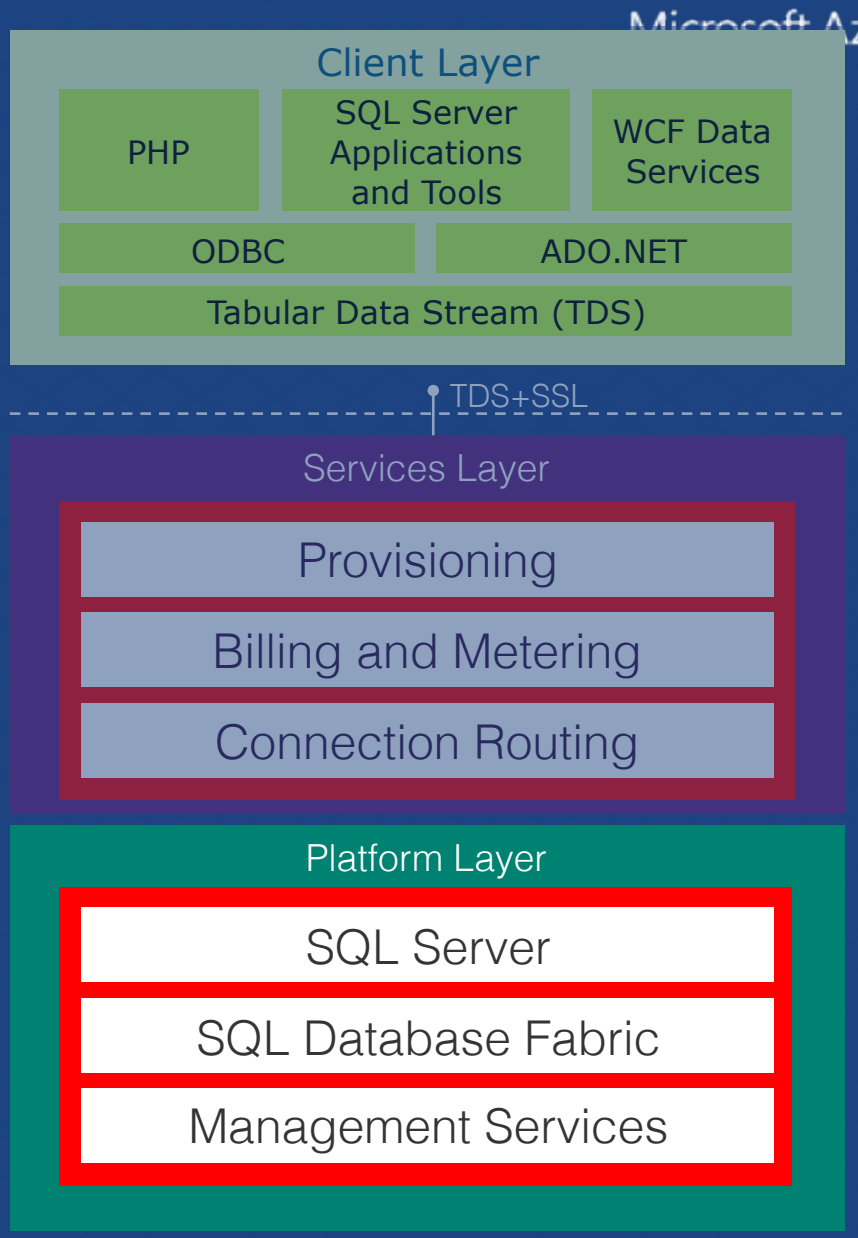
How It Works – Architecture of the Service

Gateway between Client layer and Platform layer.



Microsoft

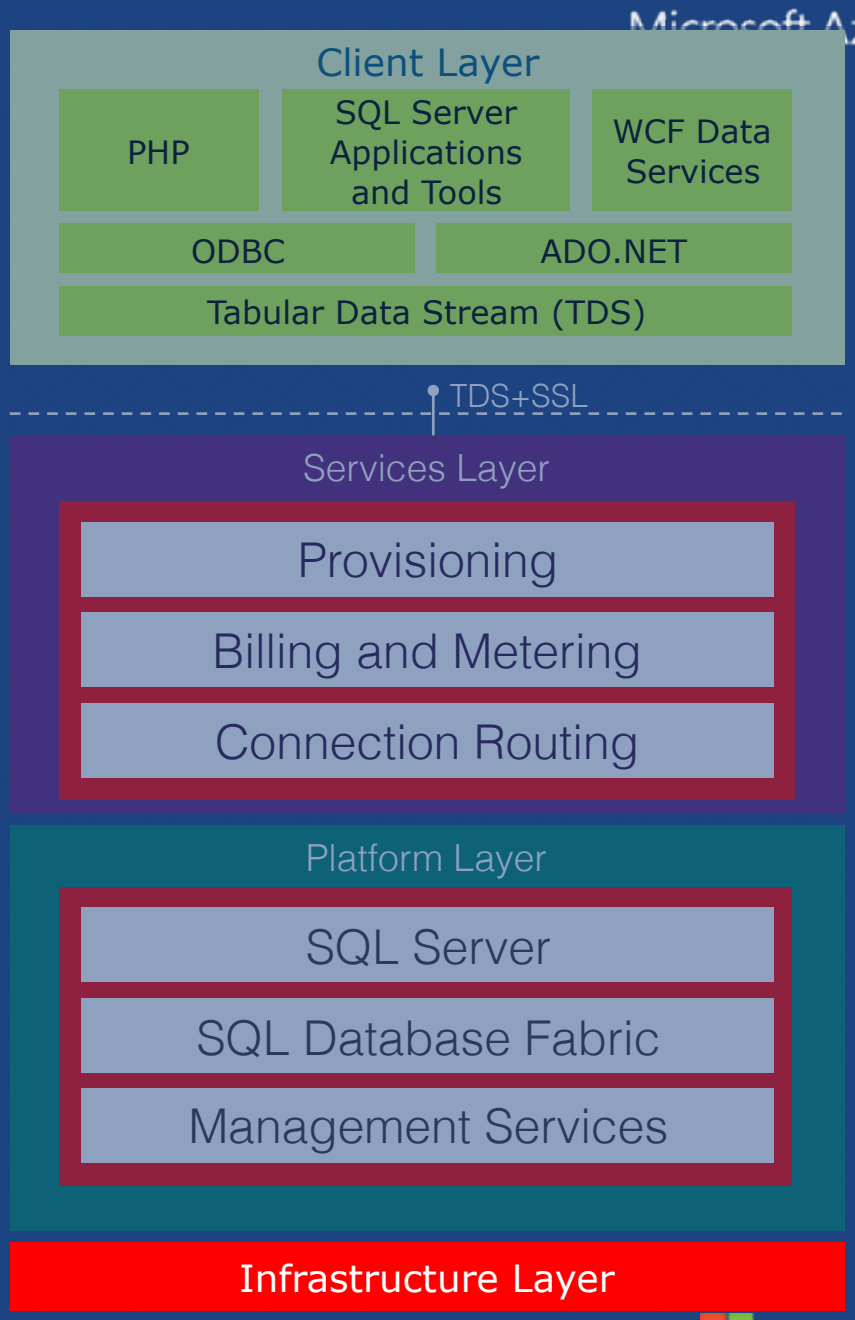
How It Works – Architecture of the Service



Includes physical servicers and services that support the Services layer.

Microsoft

How It Works – Architecture of the Service



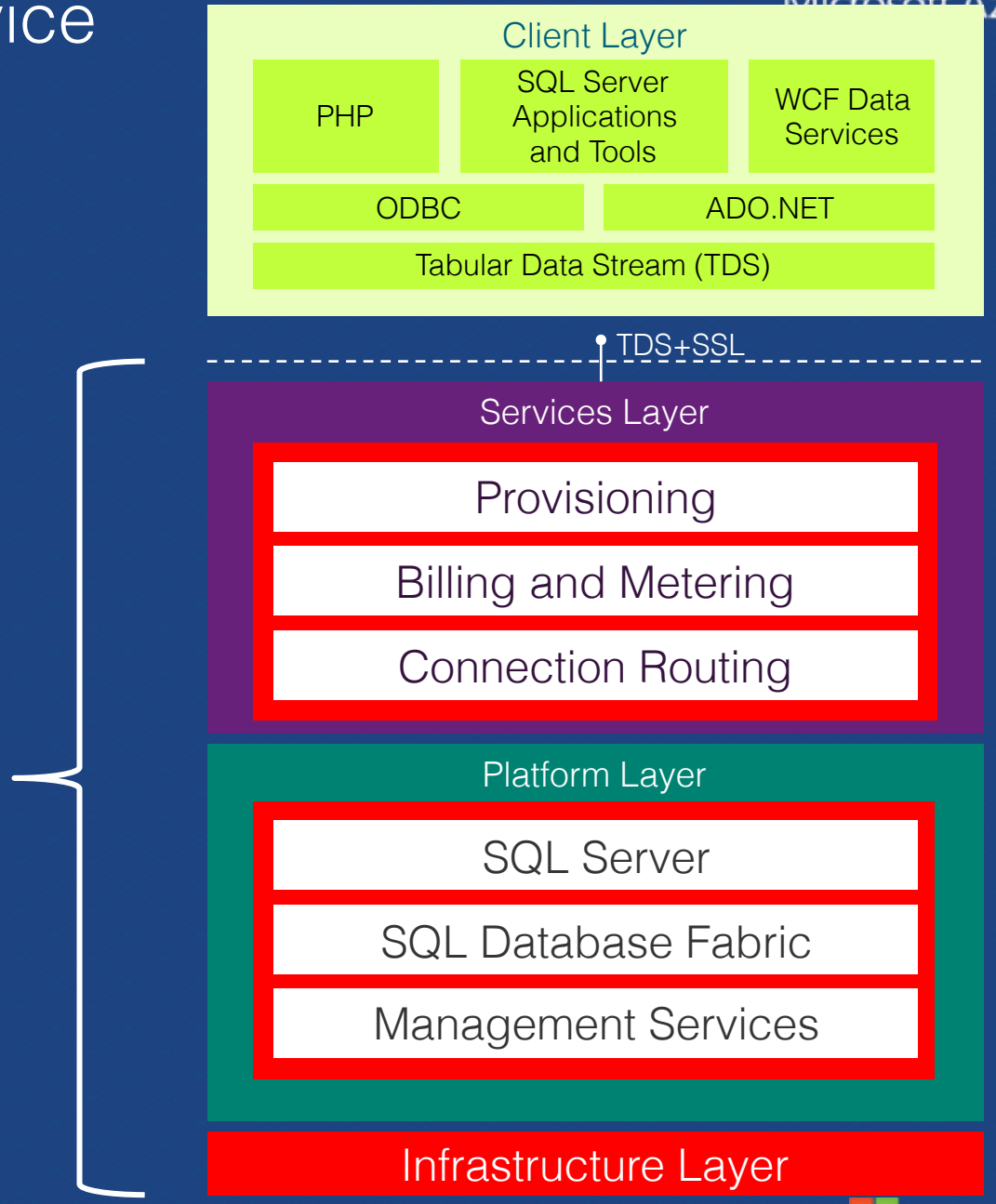
Administration of the physical HW and OS.

Microsoft

How It Works – Architecture of the Service

Microsoft Azure

Microsoft Azure SQL Database PaaS



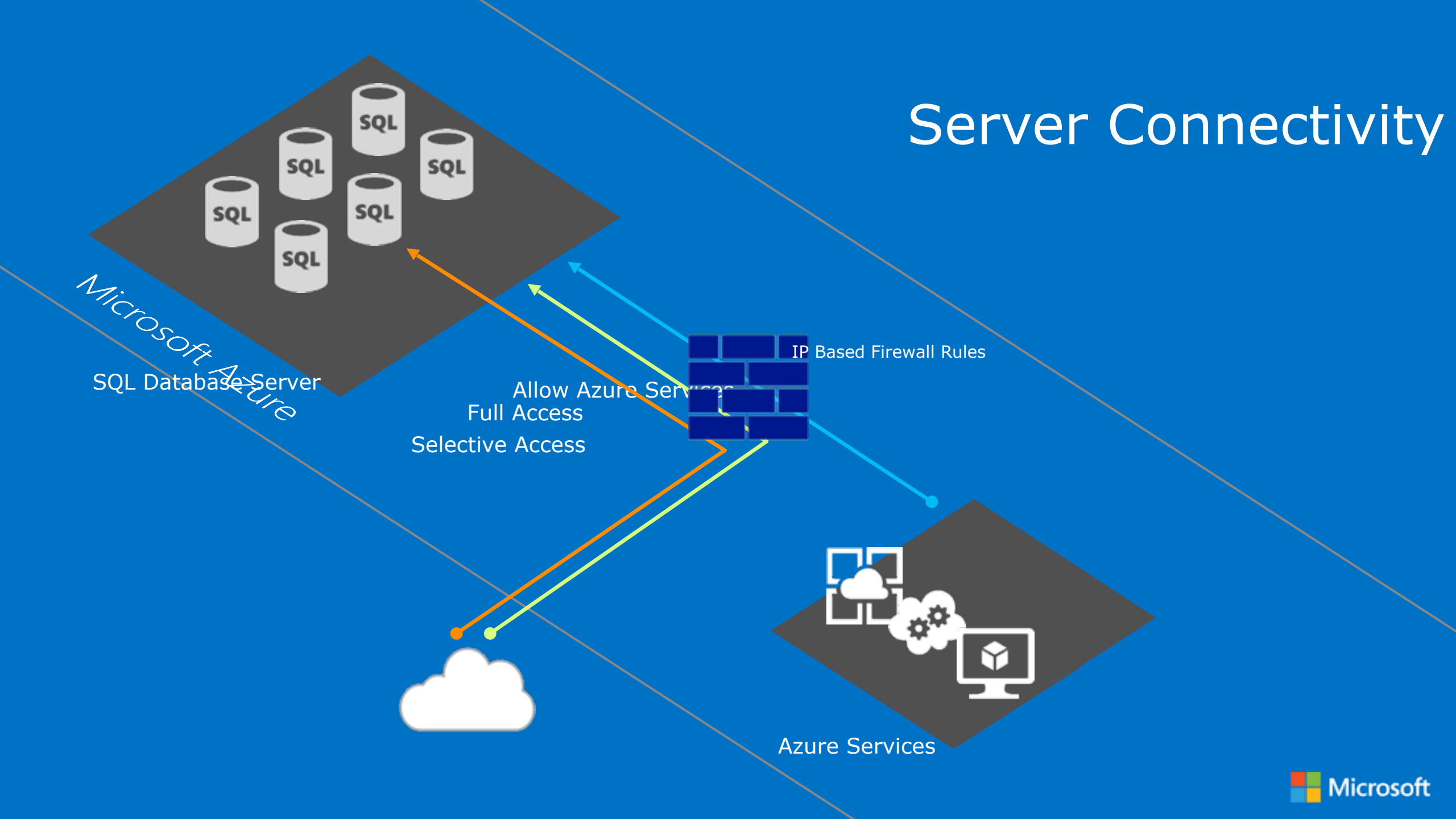
Microsoft

SQL Database Server

- The Service head contains databases
- Connect via automatically generated FQDN: {name}.database.windows.net
- Initially contains only a master database

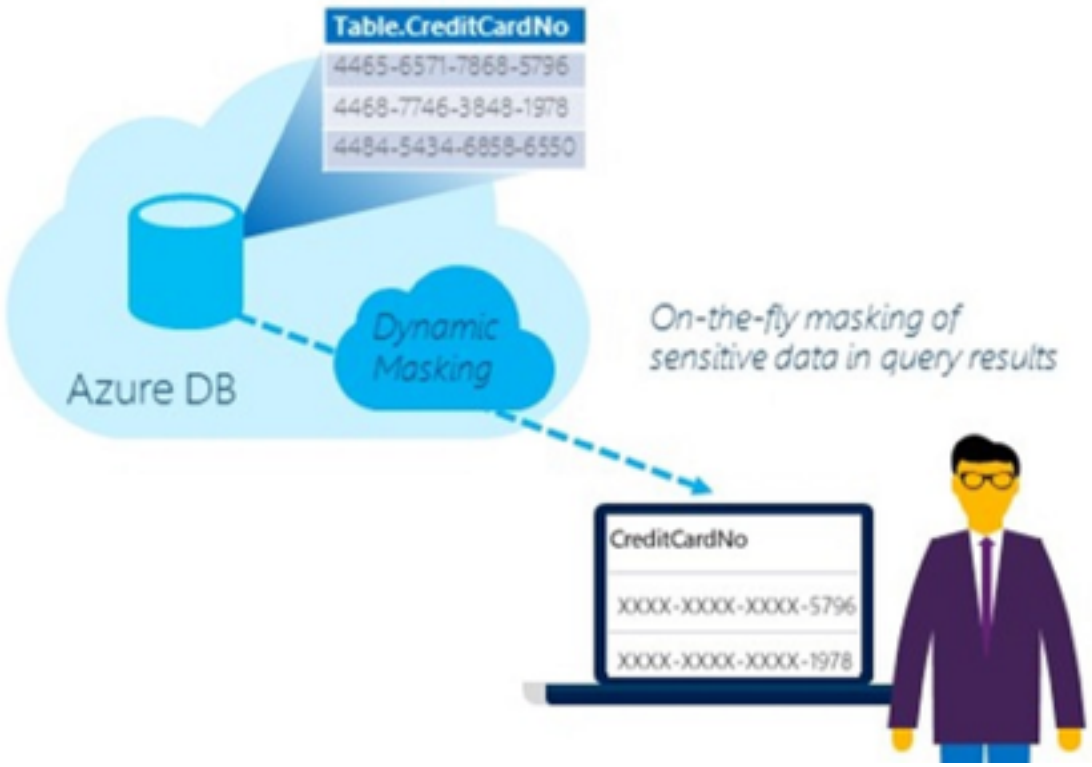
Microsoft

Server Connectivity



Microsoft

Dynamic Data Masking



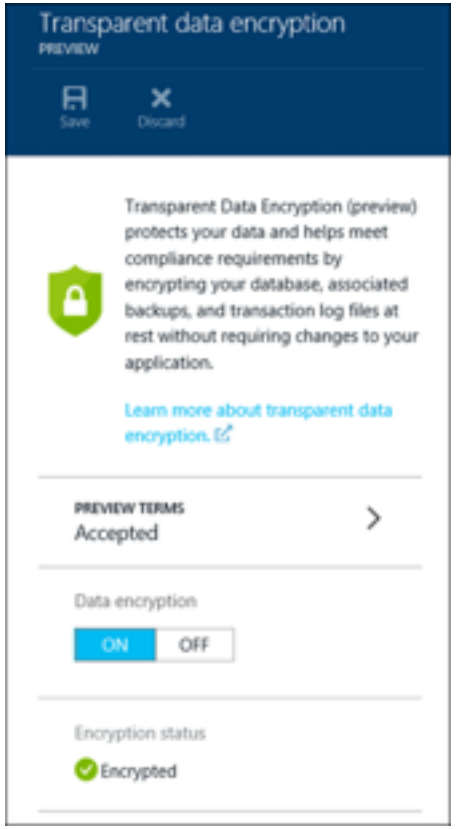
- Limits sensitive data exposure
- Prevents unauthorized access to data
- Policy-based security – no changes to data or application
- Meet regulatory compliance
- Dev/Test production data without compromising data

Protect Sensitive Data

Limit Exposure of Sensitive Data

Transparent Data Encryption (TDE)

Microsoft Azure



- Encrypted database, backups, and transaction log at rest
- 2-click provisioning
- Reduced attack surface area
- No code changes to existing applications
- Database encryption key - AES-256
- Meet regulatory compliance
- Accelerated hardware encryption

Encrypt and Protect Database

Microsoft

Row-Level Security (RLS)

- Fine-grained access over rows
- Access restrictions logic contained in database
- Simplified design and coding of security
- Meet regulatory compliance
- Reduced surface area of your security system

Fine-grained Access Over Rows

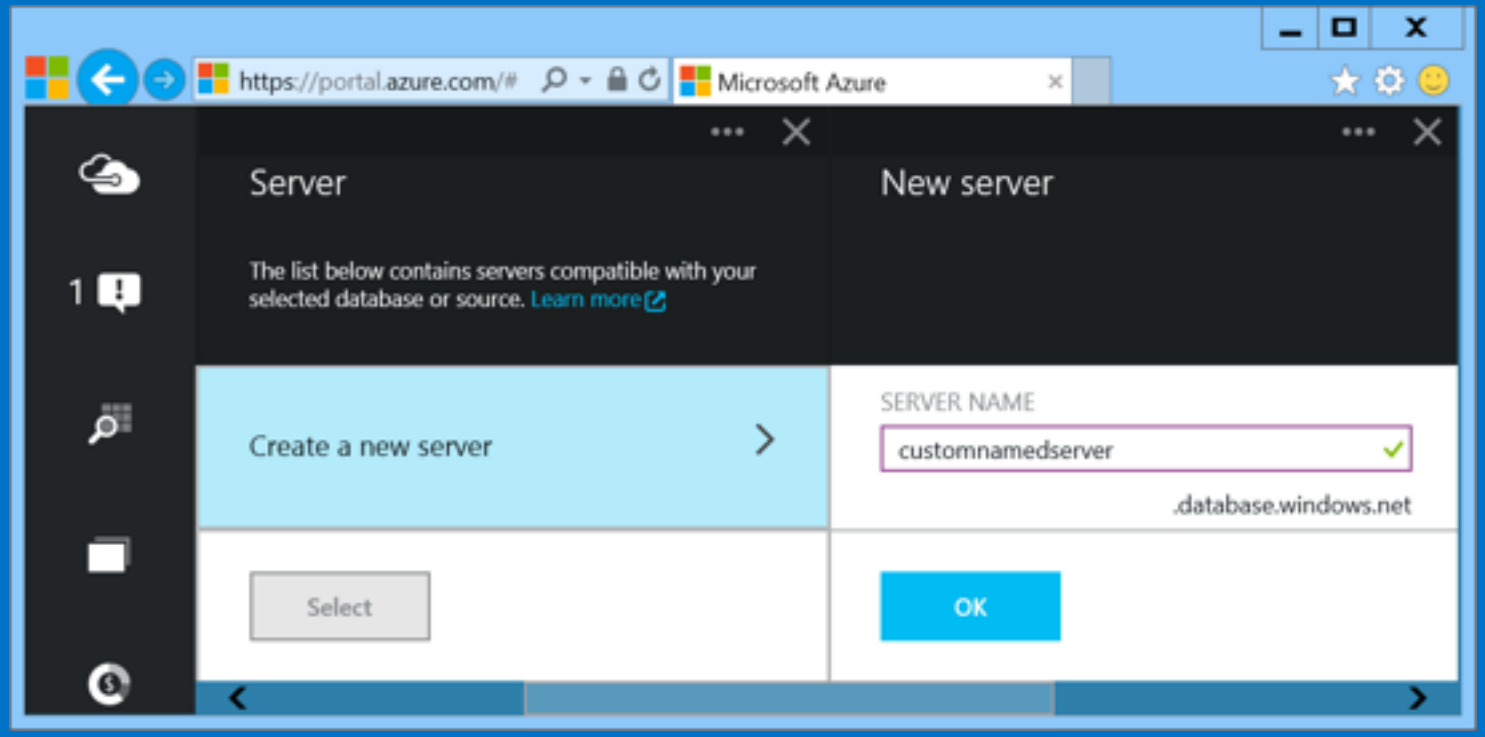
Other notable features

- Contained Database Users
- Parallel Queries
- Common Language Runtime (CLR) assemblies

Fine-grained Access Over Rows

Provision Servers Interactively

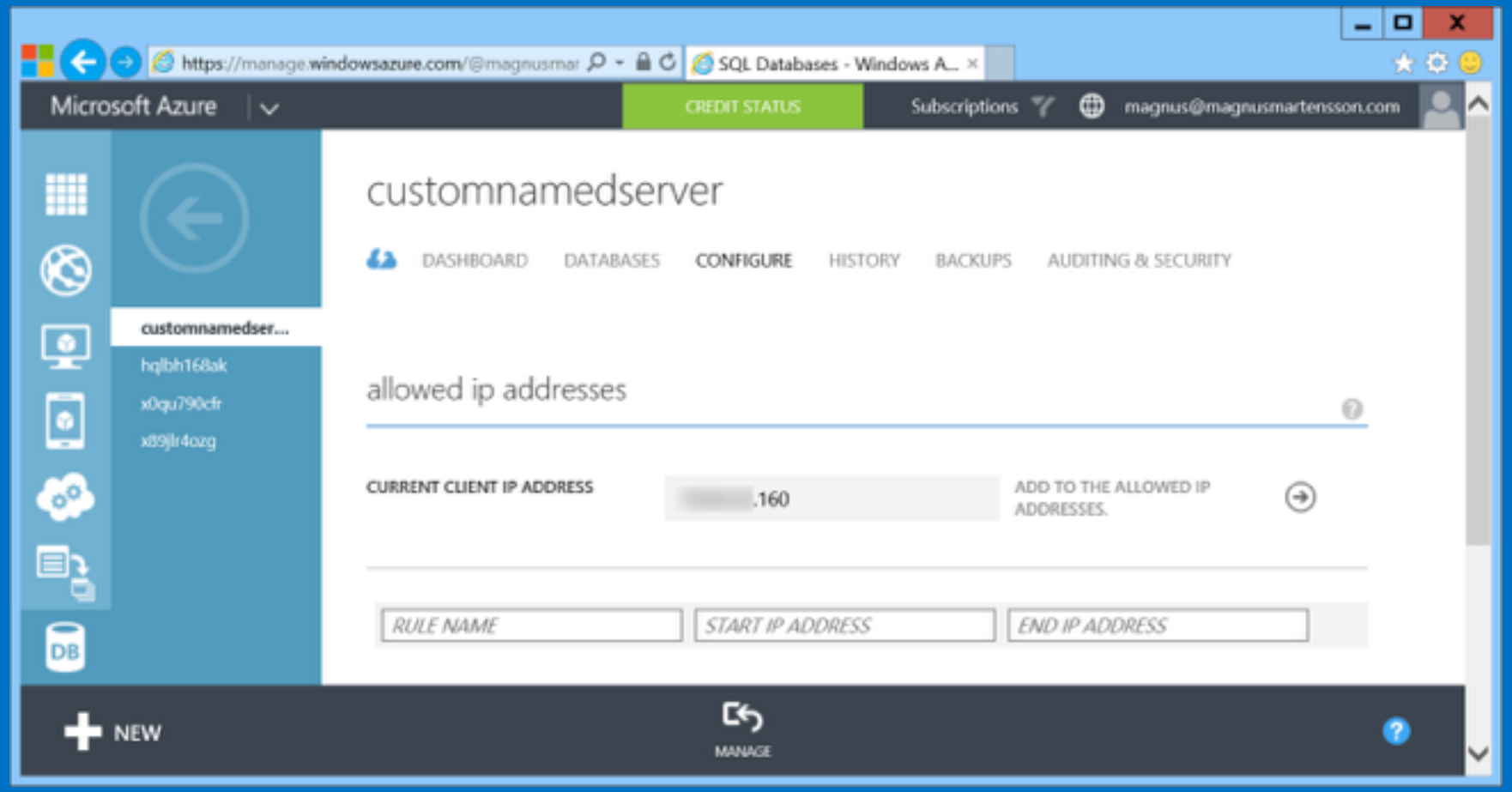
In the Preview Management Portal create a SQL Database server



Provision Servers Interactively

Microsoft Azure

In the Management Portal add firewall rules

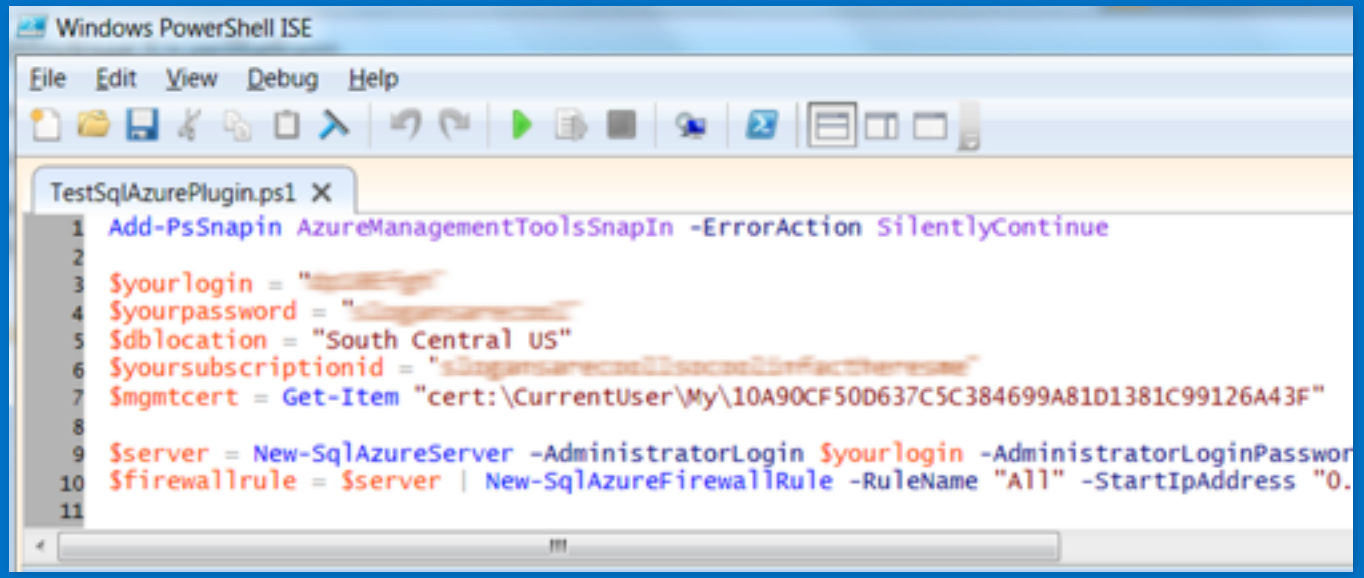


Microsoft

Automate Server Provisioning

Microsoft Azure

Microsoft Azure Platform PowerShell cmdlets



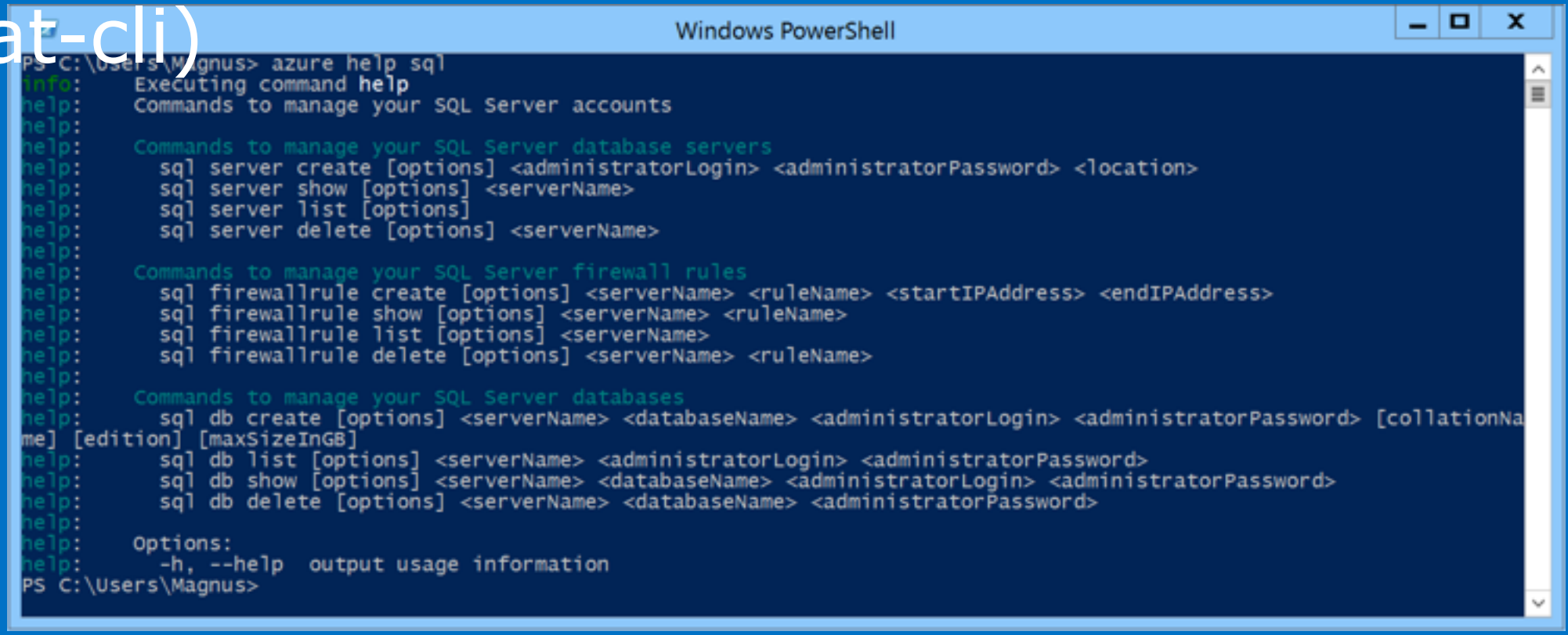
http://bit.ly/azurepowershell

Microsoft

Automate Server Provisioning

Microsoft Azure

Azure Cross-Platform Command-Line Interface (xplat-cli)



http://bit.ly/azurexplatcli

Microsoft

Selecting the right SQL Database edition

Microsoft Azure

Service Tier	Performance Level	Common App Pattern	Performance			Business Continuity	
			Max DB Size	Transaction Perf. Objective	DTU	PITR	DR / GEO-Rep
Basic	Basic	Small DB, SQL opp	2 GB	Reliability / Hr.	5	7 Days	DB Copy + Manual Export
Standard	S0 S1 S2	Wrkgp/cloud app, multiple concurrent operations	250 GB	Reliability / Min.	10 20 50	14 Days	DB Copy + Manual Export
Premium	P1 P2 P3	Mission Critical, High volume, Many concurrent Users	500 GB	Reliability / sec.	100 200 800	35 Days	Active Geo-replication

This information is subject to change over time.

Microsoft

Point-in-time restore - “oops recovery”

- Auto backups, transactional logs every 5 min
- Backups in Azure Storage and geo-replicated
- Creates a side-by-side copy, non-disruptive
- Backups retention policy: 7, 14 or 35 days
- Automated export of logical backups for long-term backup protection

Geo-restore – Emergency data recovery when you need it most

- Available in all tiers: Basic, Standard and Premium
- Built on geo-redundant Azure Storage
- Recover to any Azure region

Standard geo-replication

- Opt-in for Standard & Premium databases
- Creates a stand-by secondary
- Replicate to pre-paired Azure region
- Automatic data replication, asynchronous
- Opt-in via REST API, PowerShell or Azure Portal
- Microsoft-managed, RTO<24h, RPO<1 hr

Active geo-replication

- Self-service activation in Premium
- Create up to 4 readable secondaries
- Replicate to any Azure region
- Automatic data replication, asynchronous
- REST API, PowerShell or Azure Portal
- RTO<1h, RPO<5m, you choose when to failover

Auditing

- Configurable to track & log database activity
- Dashboard views in the portal for at-a-glance insights
- Pre-defined Power View reports for deep visual analysis on Audit log data
- Audit logs reside in your Azure Storage account
- Available in Basic, Standard, and Premium

Microsoft

Full-Text Search

- Fast and flexible indexing of textual data
- Data types: char, varchar, nchar, nvarchar, text, ntext, image, xml, varbinary(max), or FILESTREAM
- Handles high query volume
- Common use cases:
  - Searching websites, product catalogs, news items and more
  - Document management systems
  - Any applications that need to provide search capabilities over data stored in a SQL Database

XML Indexes

- XML Indexes - improves XQuery-based query performance
- Primary XML Index - speed up access to elements and attributes
  - CREATE PRIMARY XML INDEX XML_Order_Items
  - ON Sales.Order (Items);
- Secondary XML Index – help resolve specific XQuery expressions rapidly

Dynamic Management Views (DMV)

- Monitor common database, execution and transaction related events in near-real time
- Diagnose blocked or long-running queries, resource bottlenecks, poor query plans, and more
- Help improve capacity planning
- Use familiar T-SQL language

Use Familiar Technologies

Transact-SQL

(obviously)

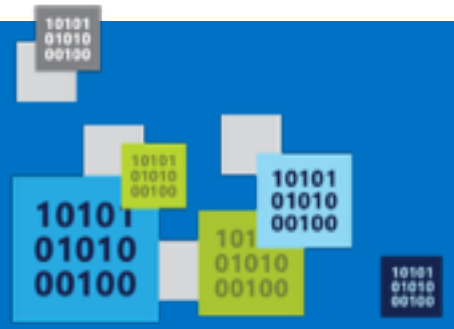
Use Familiar Technologies - Languages

.NET Framework (C#, Visual Basic, F#): ADO.NET

C / C++: ODBC

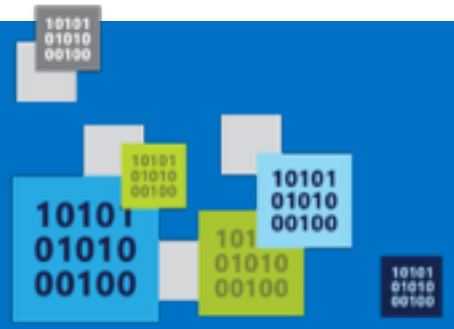
Java: Microsoft JDBC provider

PHP: Microsoft PHP provider



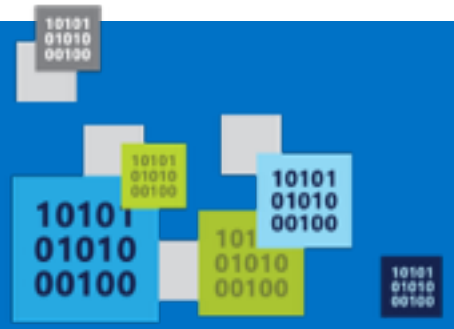
Use Familiar Technologies - Frameworks

- OData
- Entity Framework
- WCF Data Services
- NHibernate (etc.)



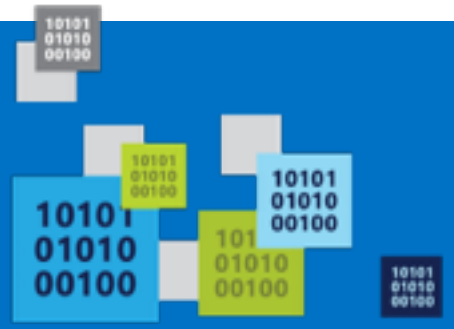
Use Familiar Technologies - Tools

- SQL Server Management Studio (>=2008 R2)
- SQL Server command-line utilities (SQLCMD, BCP)
- Visual Studio IDE for database development



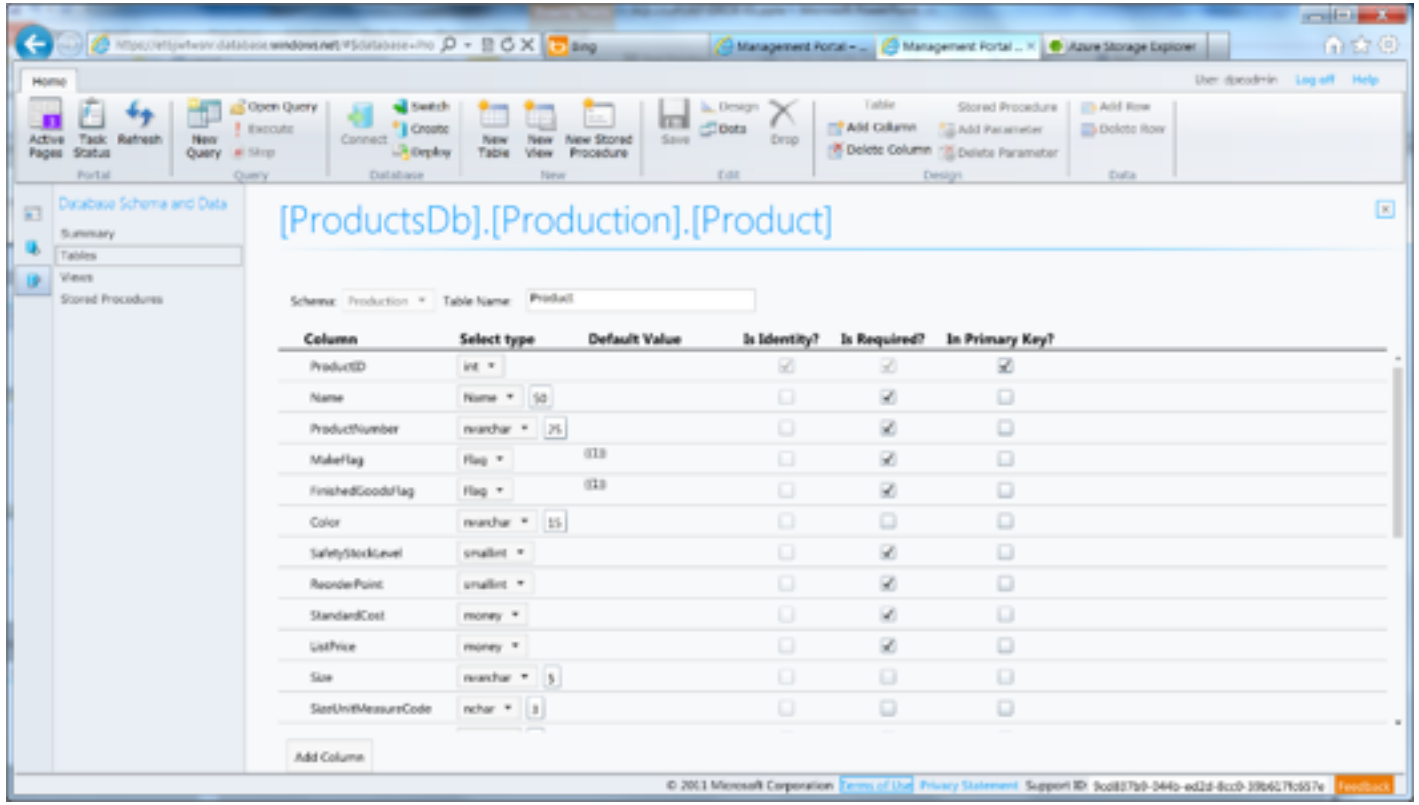
Unsupported SQL Server Features

- Use command, distributed transactions, distributed views
- Service Broker
- SQL Agent
- SQL Profiler
- Native Encryption

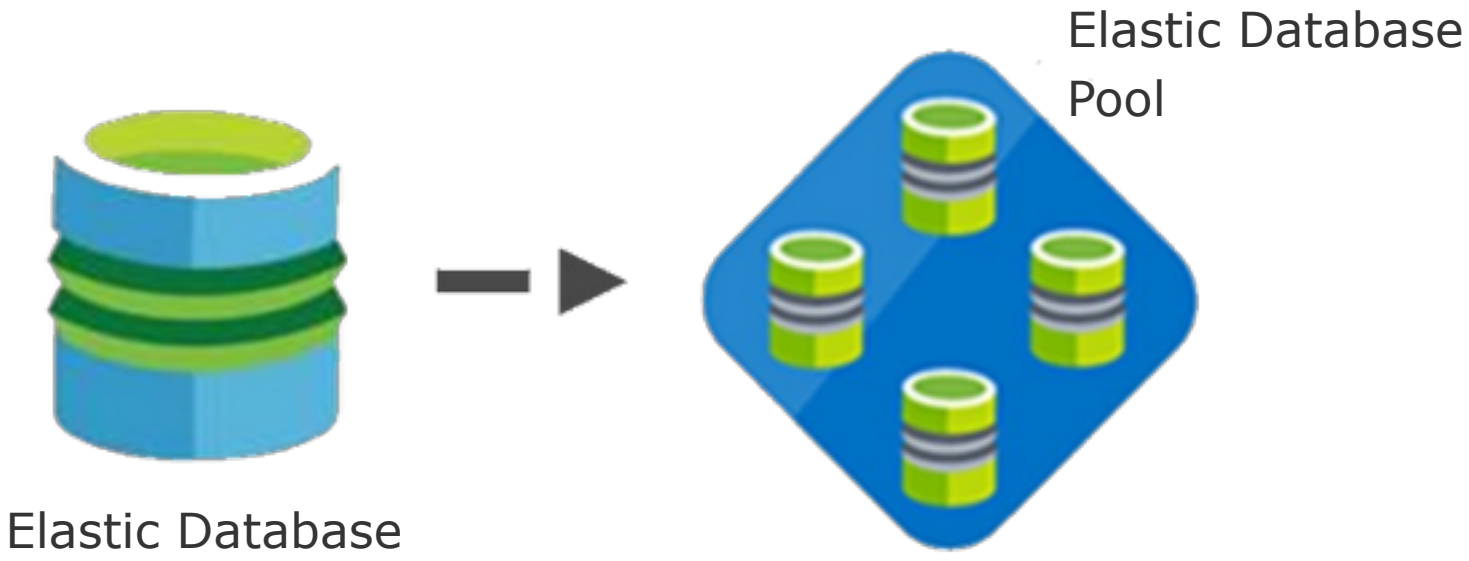


Web designers for tables, views, stored procs

Interactive query editing and execution



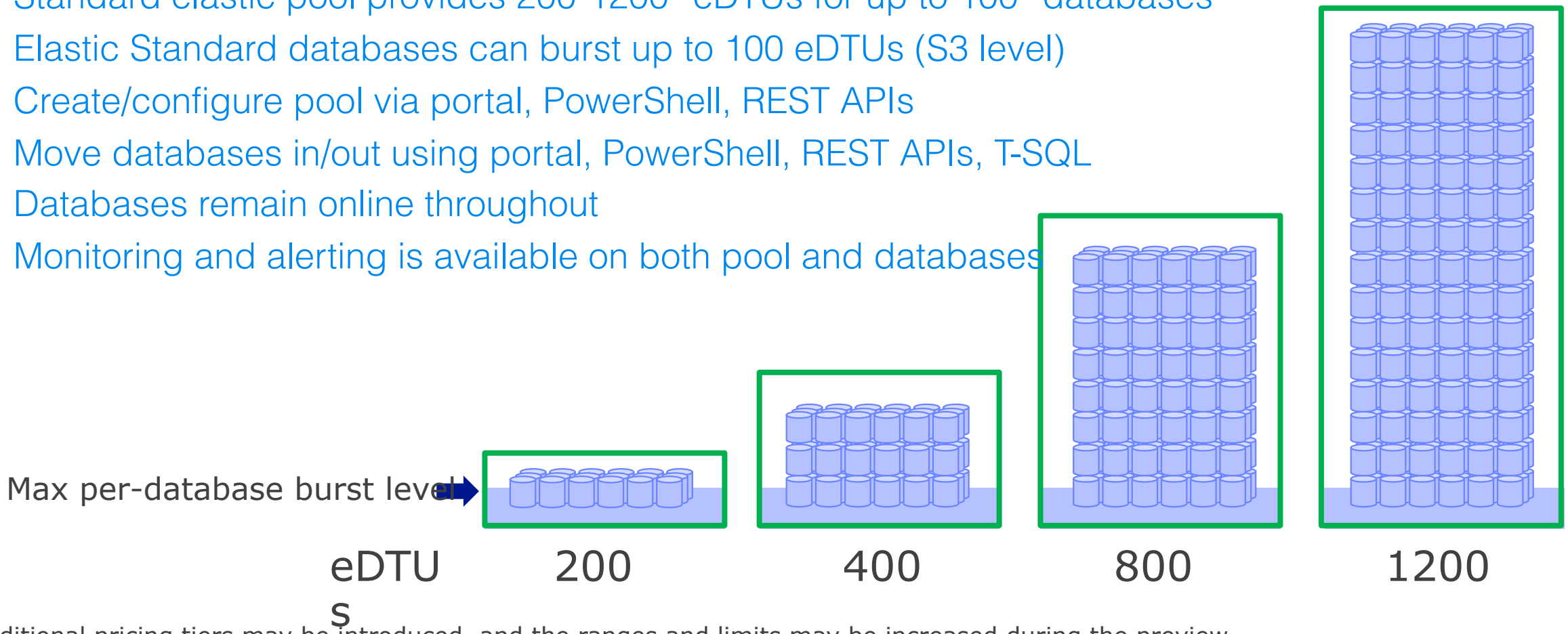
Elastic Database



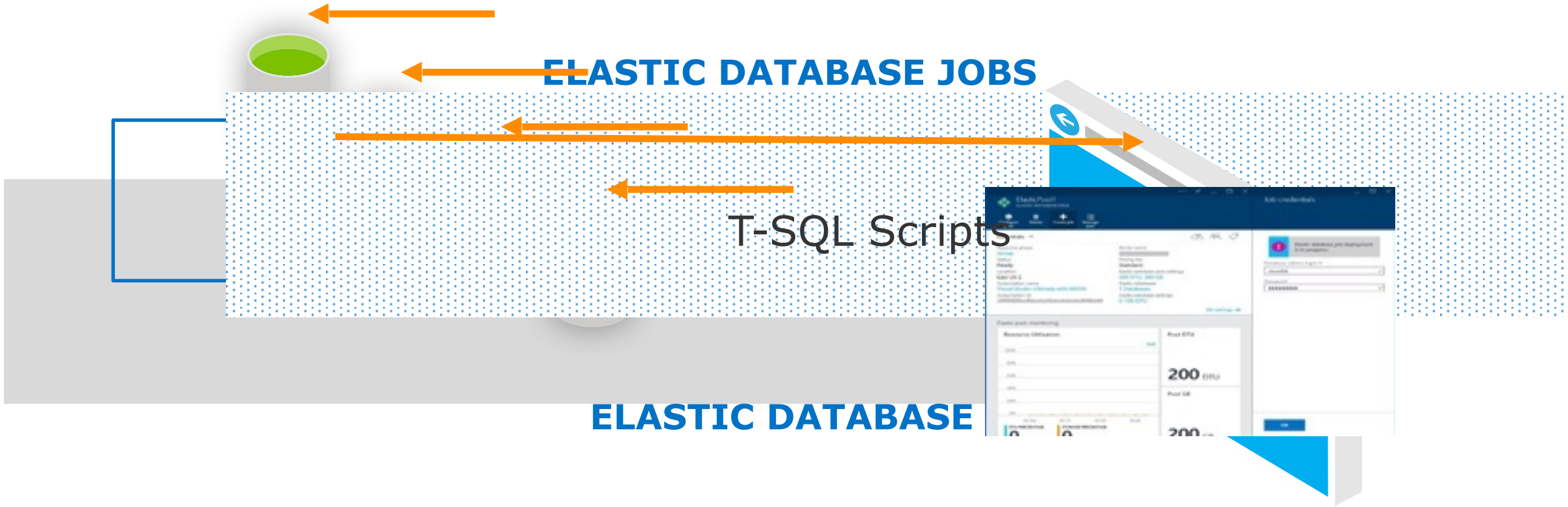
Predictable model for deploying large numbers of databases

Elastic Database Model PREVIEW

- Elastic databases, Elastic database pools
- Pooled resources leveraged by many databases
- Standard elastic pool provides 200-1200* eDTUs for up to 100* databases
- Elastic Standard databases can burst up to 100 eDTUs (S3 level)
- Create/configure pool via portal, PowerShell, REST APIs
- Move databases in/out using portal, PowerShell, REST APIs, T-SQL
- Databases remain online throughout
- Monitoring and alerting is available on both pool and databases



*Additional pricing tiers may be introduced, and the ranges and limits may be increased during the preview

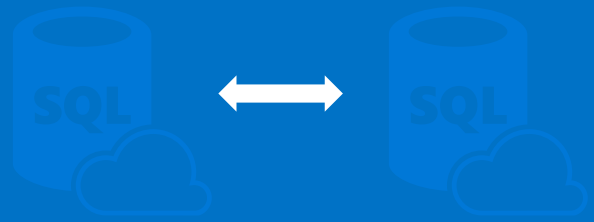


Execute administrative tasks across each database

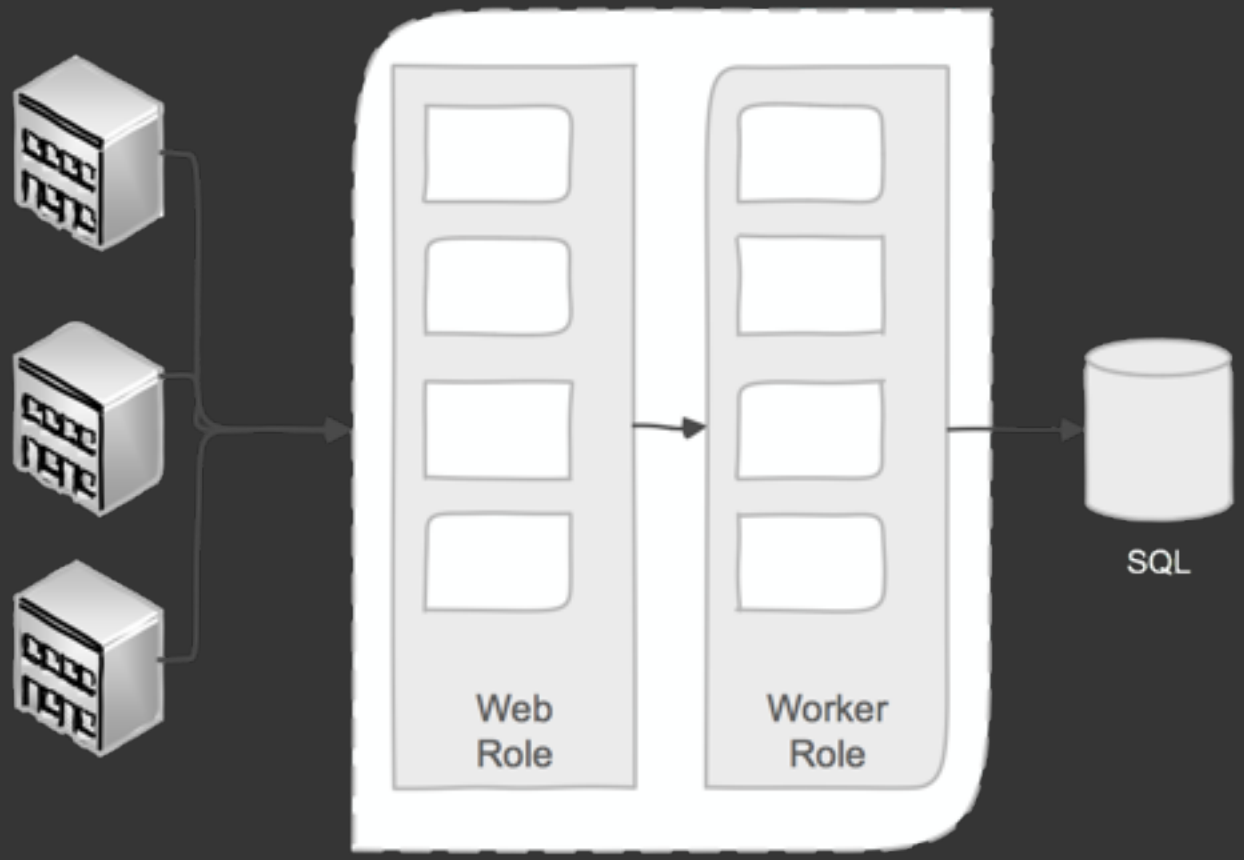
Elastic Database Jobs

- Apply changes or administrative operations to many databases
- Use familiar T-SQL scripts to define jobs
- Built-in automatic retries in case of transient failures
- Tightly integrated with elastic pools in the new Azure Portal
- Designed for batch processing

Elastic Scale



Canonical cloud app architecture



- Classic 3-tier enterprise architecture:
- Scale out the front ends to multiple instances is easy
- Scale the data-tier is more challenging

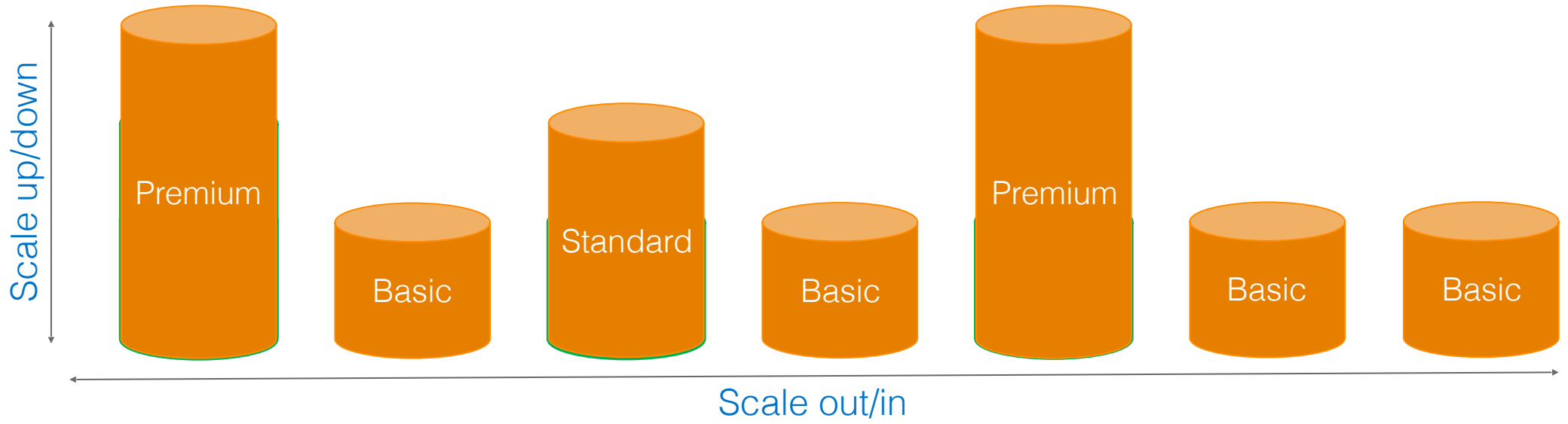
SQL Database Considerations and Best Practices

Microsoft Azure

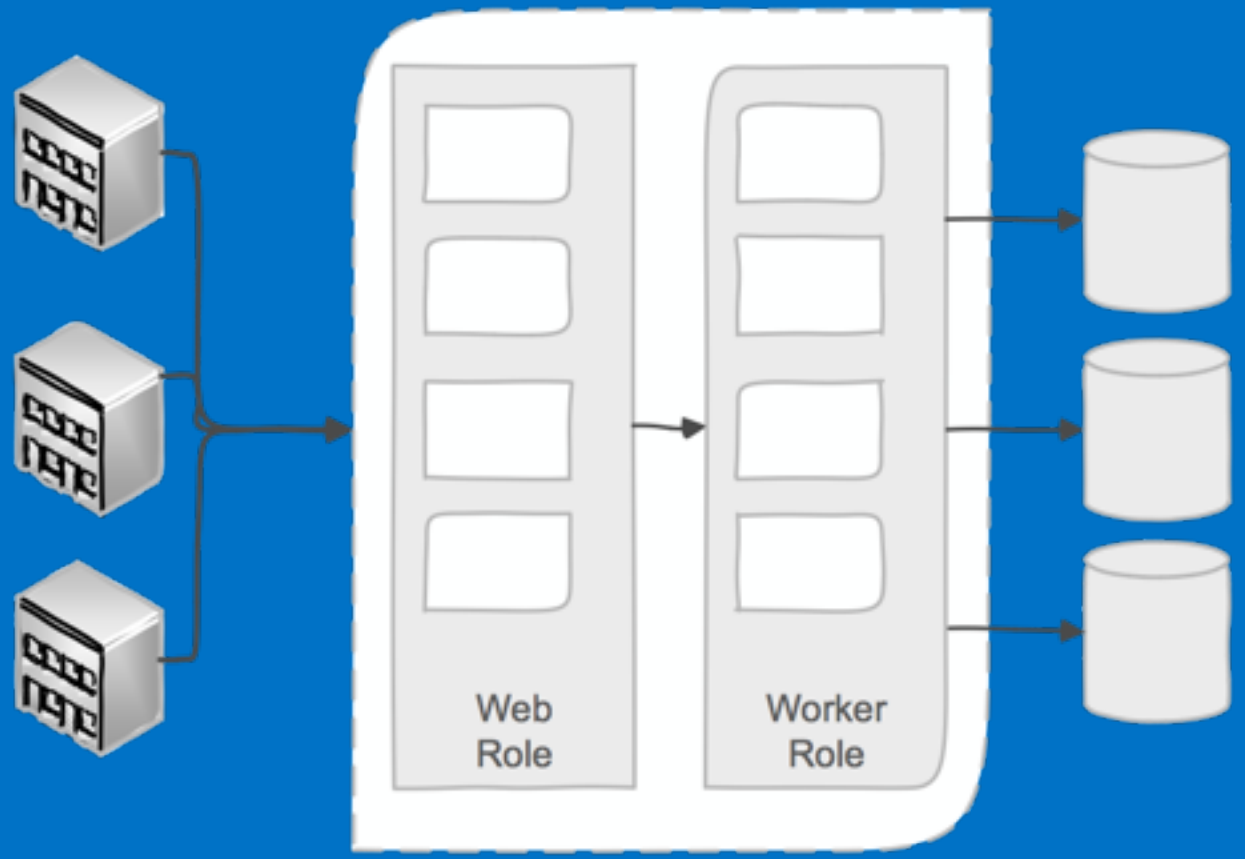
- Elastic Scale across thousands of databases via custom sharding
- Scale out via .NET Client libraries consumed by customer applications to support sharded database pattern
- Enables developer and manageability functions
- Supports split, merge, and move operations on data

Microsoft

Vertical: Scale-up or scale-down
Horizontal: Scale-out or scale-in



Elastic Scale architecture



Microsoft Azure

SQL Server VM

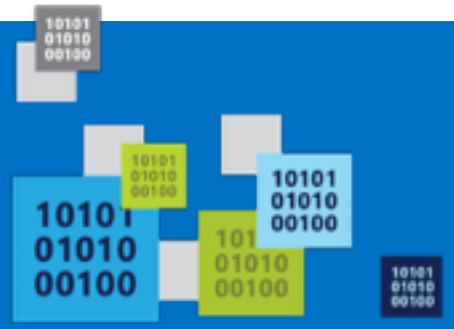
Microsoft

Run SQL on VM

- Run any SQL product on cloud VM
- Support for SQL Server, Oracle, MySql
- Ready to go VM images available in Gallery
- Persistent storage using attached disk in blob storage
- Has all the benefits and powers of VMs combined with the full features of a SQL Engine

SQL IaaS

SQL Database



Microsoft Azure

SQL Database

When you want reduced overhead and possibly need elastic scale.

Customer does not want to add additional IT resources for support and maintenance.

Avoiding CAPEX and OPEX.

SQL IaaS

Existing applications which requires full box product functionality.

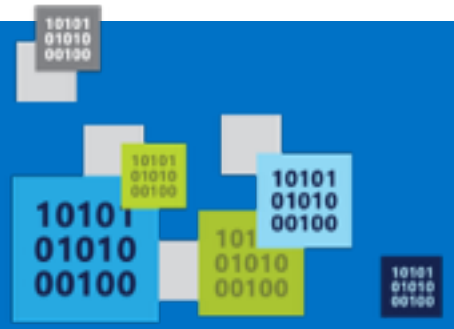
Customer has ecosystem of IT resources for support and maintenance.

Removing CAPEX.

Microsoft

Other features SQL IaaS supports that SQL Database doesn't (yet)

- Full SQL Server functionality (e.g. Reporting Services)
- Windows authentication available (requires VM to be joined to on-premises domain)
- Larger database sizes possible (16TB)

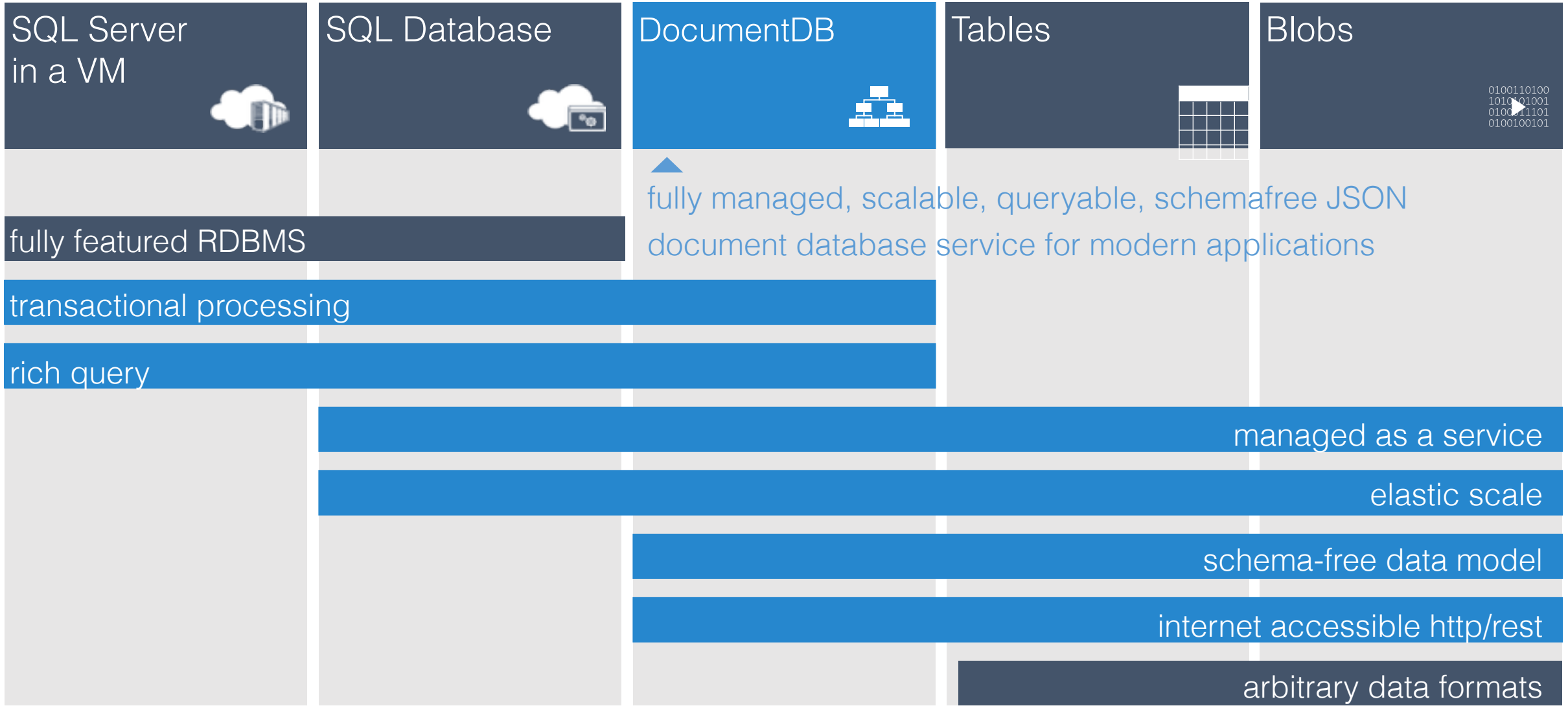


DocumentDB

Fully managed, scalable JSON document database service

Microsoft Azure

Microsoft Azure Data Services



Microsoft

Microsoft Azure

MongoDB

Existing applications which require extra capacity for scale out and can not be migrated

Customer has ecosystem of IT resources for support and maintenance

Removing CAPEX

Mongo MMS compatibility

DocumentDB

Applications that need managed elastic scale, query over schema free data, native JSON/JavaScript support

Customer does not want to add additional IT resources for support and maintenance

Avoiding CAPEX and OPEX

Built-for-the-cloud database technology

Microsoft

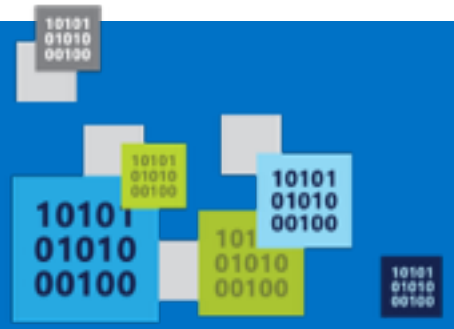
Tunable Consistency – four distinct levels

Strong
Bounded Stateless
Session
Eventual

Microsoft

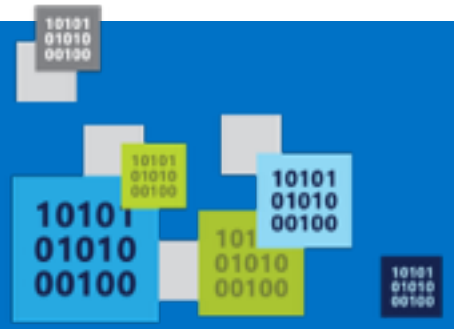
Tunable Consistency
Strong

All writes are visible to all readers. Writes synchronously committed by a majority quorum of replicas and reads are acknowledged by the majority read quorum.



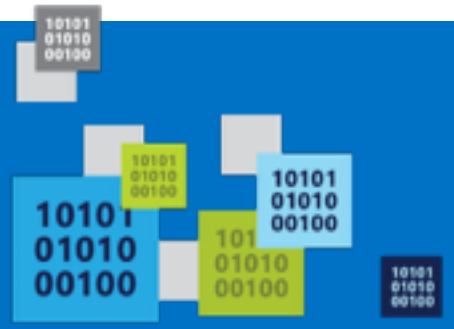
Tunable Consistency
Bounded Stateless

Guaranteed ordering of writes, reads adhere to minimum freshness. Writes are propagated asynchronously, reads are acknowledged by majority quorum lagging by at most K prefixes.



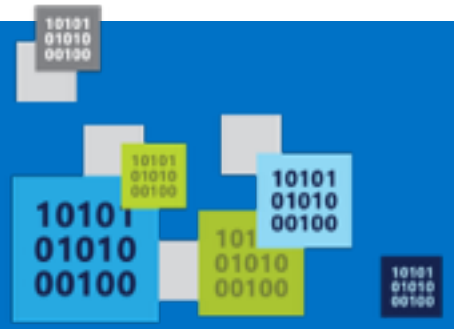
Tunable Consistency
Session

Read your own writes. Writes are propagated asynchronously while reads for a session are issued against the replica that can serve the requested version.



Tunable Consistency
Eventual

Reads eventually converge with writes. Writes are propagated asynchronously while reads can be acknowledged by any replica. Readers may view older data then previously observed.



Tunable Consistency

	Writes	Reads
Strong	sync quorum writes	quorum reads
Bounded	async replication	quorum reads
Session *	async replication	session bound replica
Eventual	async replication	any replica

* Ideal consistency and performance tradeoff for many application scenarios. High performance writes and reads with predictable consistency.

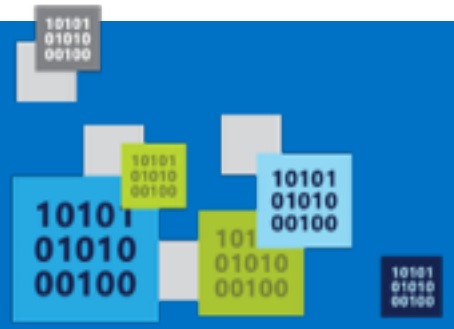
Partitioning Data

Hash
Range
Lookup

Microsoft

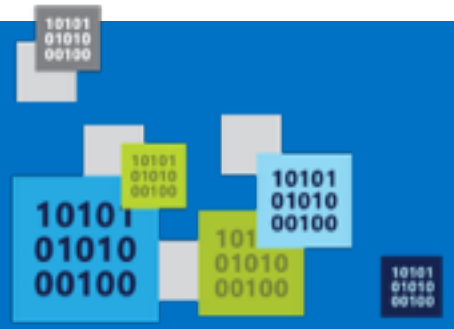
Partitioning
Hash

Partitions are assigned based on the value of a hash function, allowing you to evenly distribute requests and data across a number of partitions



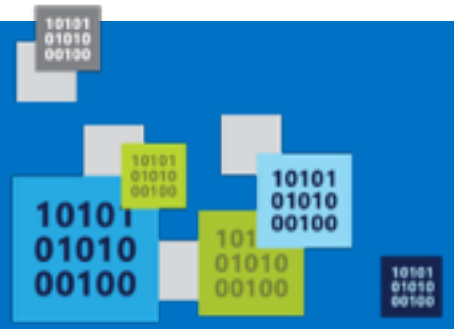
Partitioning
Range

Partitions are assigned based on whether the partition key is within a certain range.



Partitioning
Lookup

Partitions are assigned based on a lookup map that assigns discrete partition values to specific partitions a.k.a. a partition or shard map.



Search

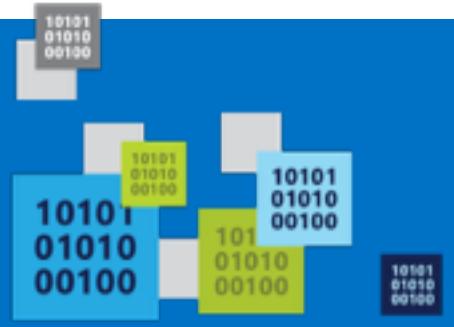
Azure Search

Embed a sophisticated search experience into web and mobile applications without having to worry about the complexities of full-text search and without having to deploy, maintain or manage any infrastructure.

Microsoft

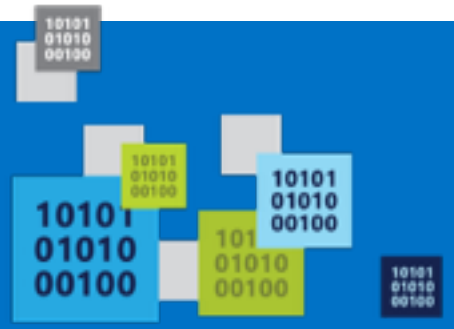
Azure Search

Perfect for enterprise cloud developers, cloud software vendors, cloud architects who need a fully-managed search solution.



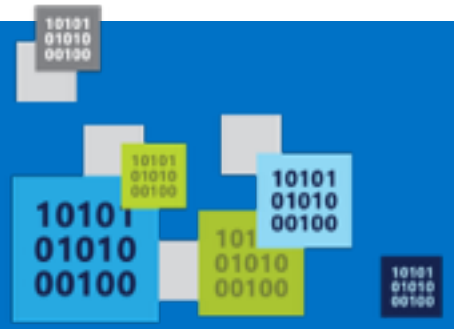
Search Functionality

- Simple HTTP/JSON API for creating indexes, pushing documents, searching
- Keyword search with user-friendly operators (+, -, *, “”, etc.)
- Hit highlighting
- Faceting (histograms over ranges, typically used in catalog browsing)



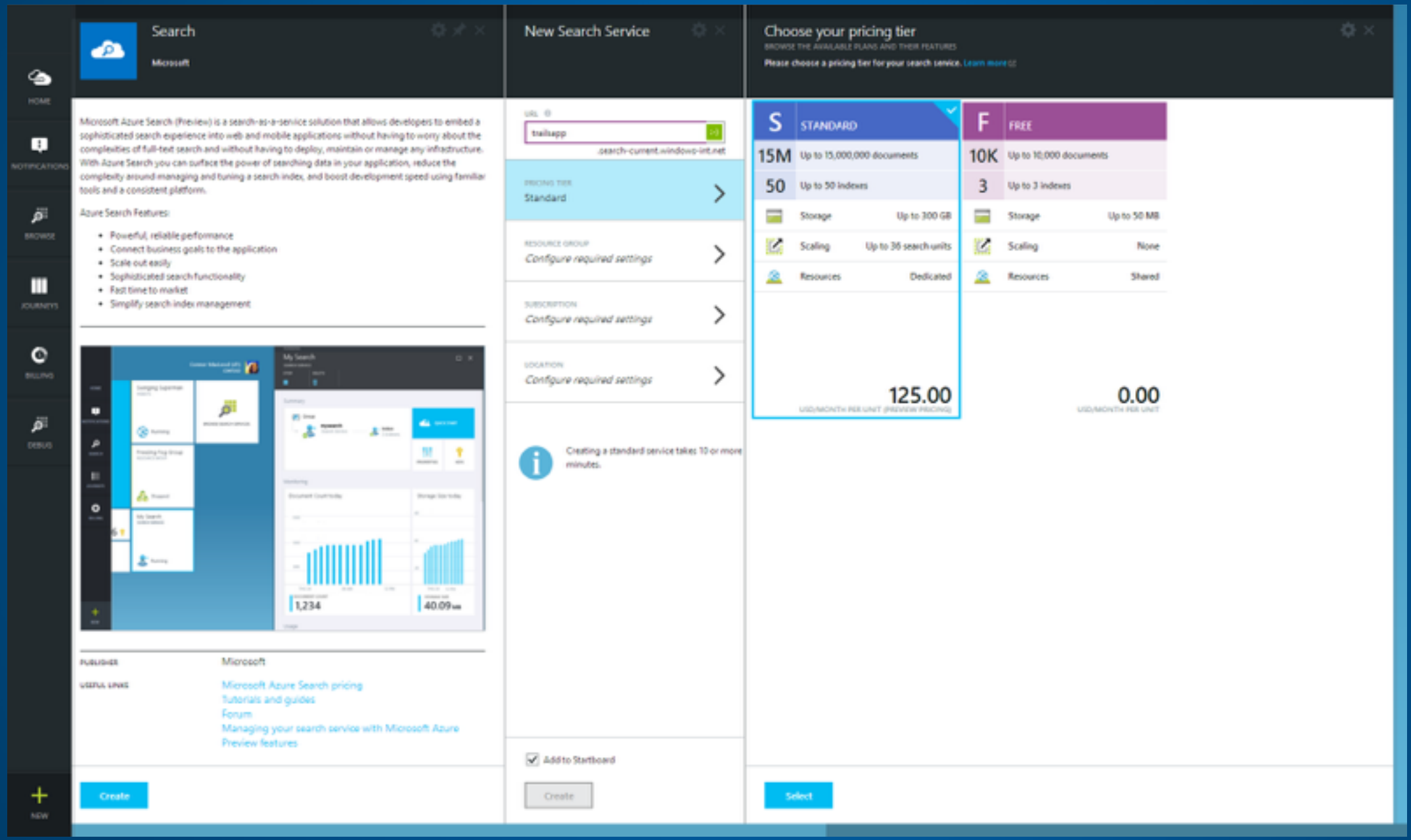
Search Functionality

- Suggestions (auto-complete)
- Rich structured queries (filter, select, sort) that combines with search
- Scoring profiles to model search result relevance
- Geo-spatial support integrated in filtering, sorting and ranking



Choose Tier

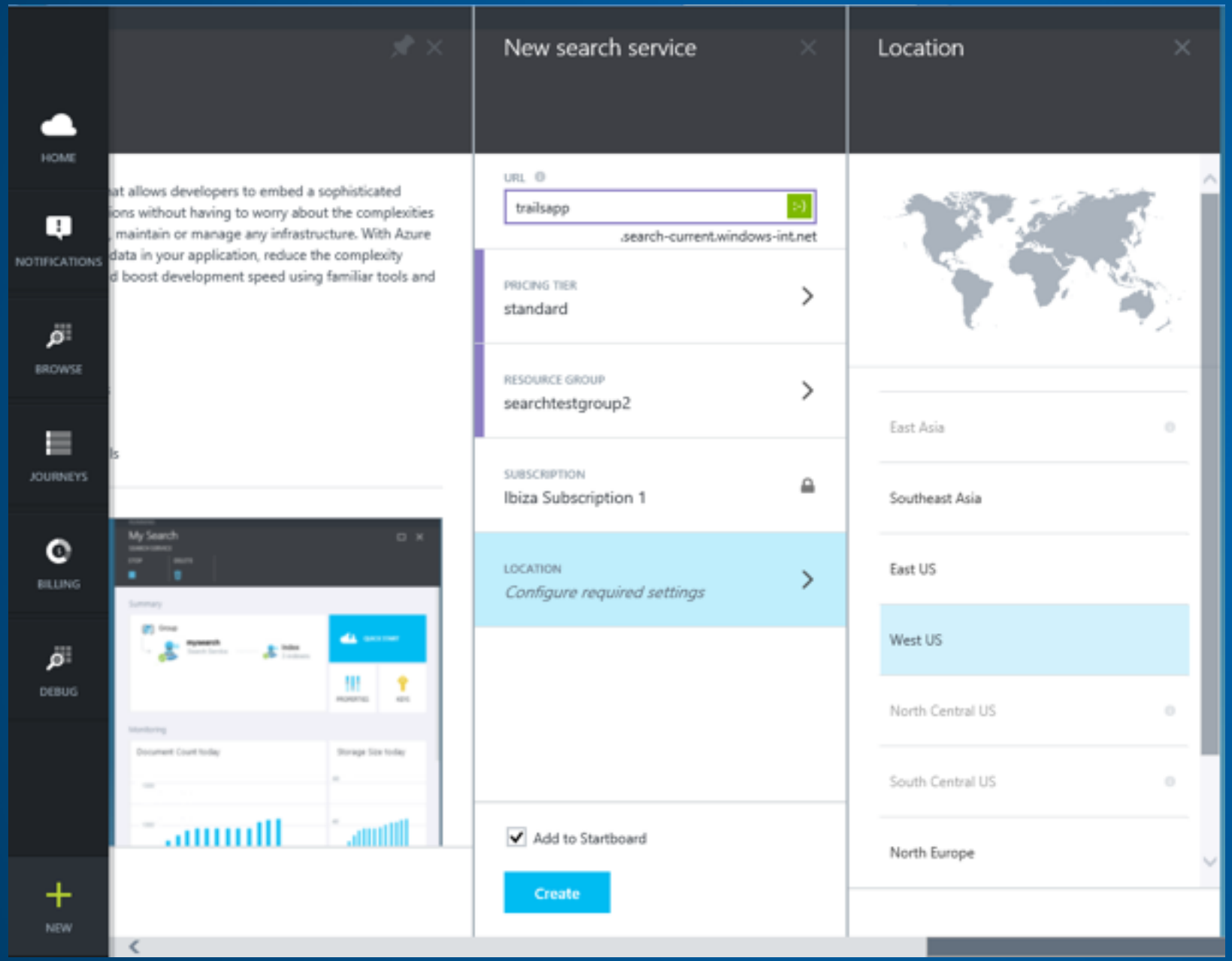
Microsoft Azure



Microsoft

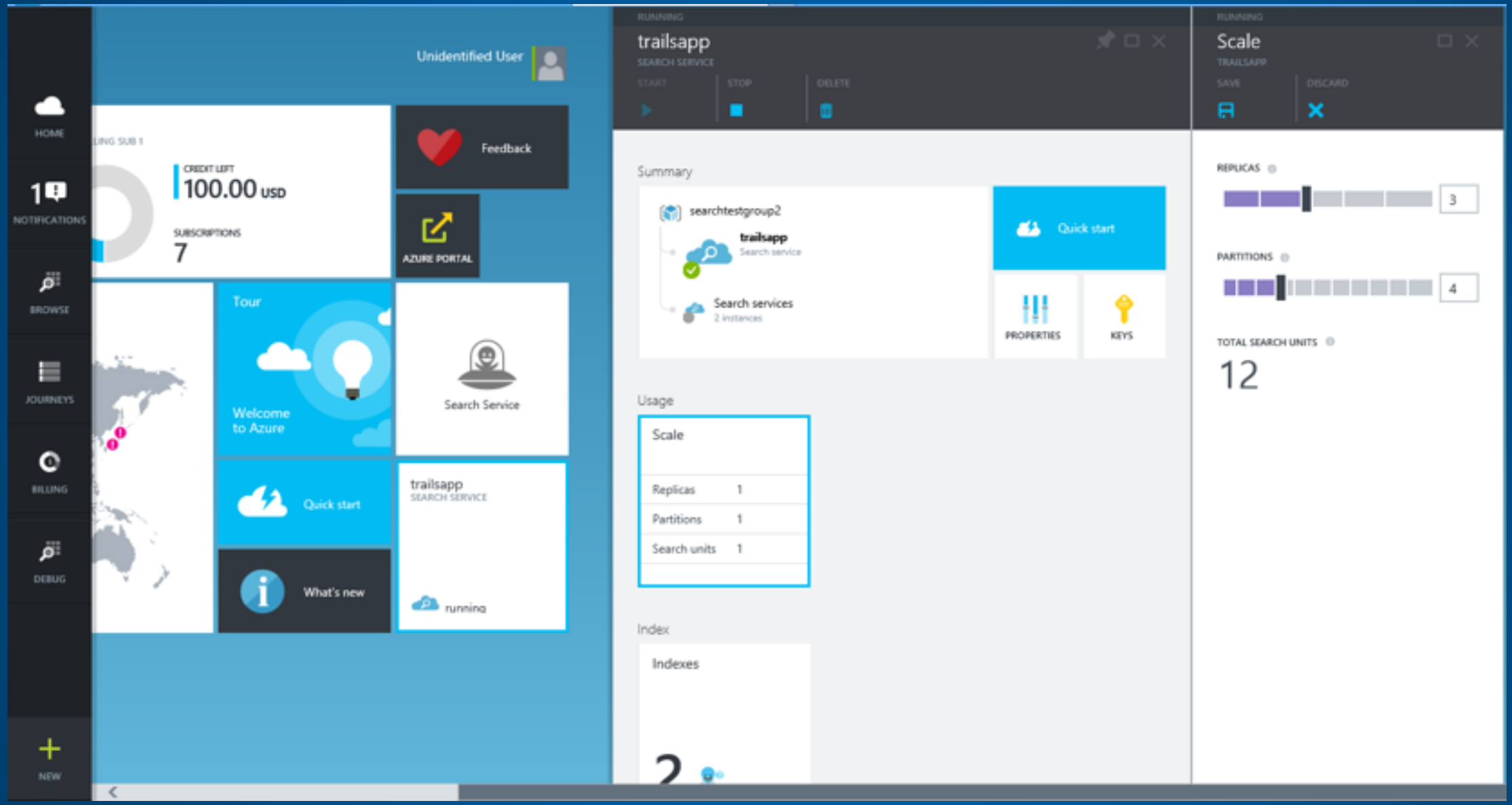
Global Availability

Microsoft Azure



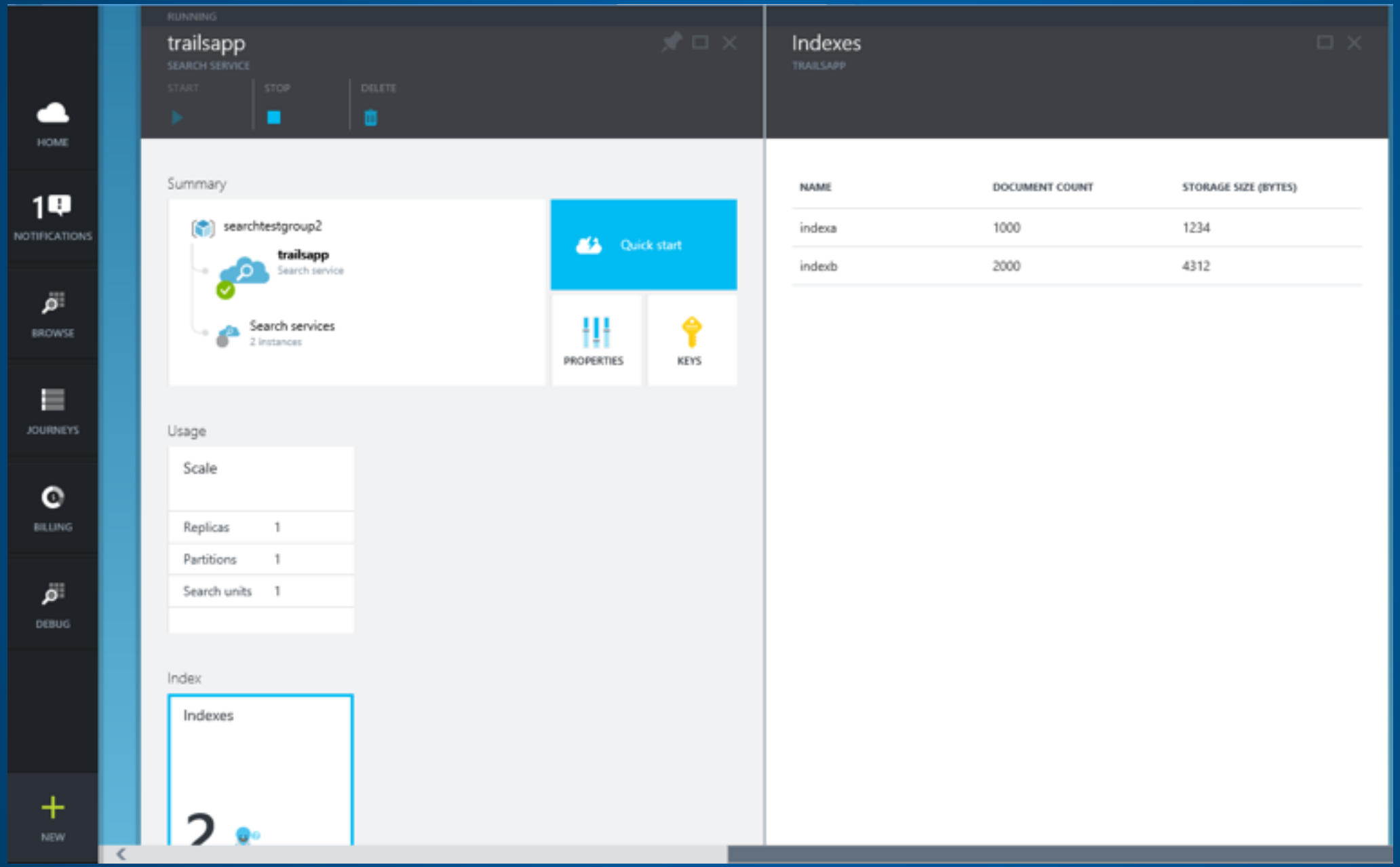
Microsoft

Scale Up


HOME
NOTIFICATIONS
BROWSE
JOURNEYS
BILLING
DEBUG
NEW
Unidentified User
CREDIT LEFT
100.00 USD
SUBSCRIPTIONS
7
Feedback
AZURE PORTAL
Tour
Welcome to Azure
Search Service
Quick start
What's new
trailsapp
SEARCH SERVICE
running
RUNNING
trailsapp
SEARCH SERVICE
START
STOP
DELETE
Summary
searchtestgroup2
trailsapp
Search service
Search services
2 instances
Quick start
PROPERTIES
KEYS
Usage
Scale
Replicas 1
Partitions 1
Search units 1
Index
Indexes
2
RUNNING
Scale
TRAILSAPP
SAVE
DISCARD
REPLICAS
3
PARTITIONS
4
TOTAL SEARCH UNITS
12

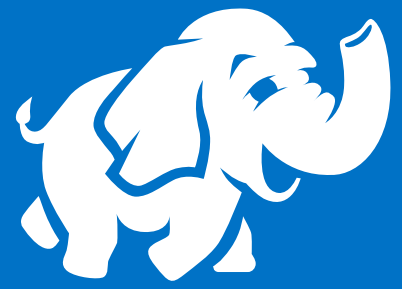
Manage Indexes

Microsoft Azure



Microsoft

HDInsight



Relational DB vs. Hadoop

Data size	Gigabytes (Terabytes)	Petabytes (Hexabytes)
Access	Interactive and Batch	Batch
Updates	Read / Write many times	Write once, Read many times
Structure	Static Schema	Dynamic Schema
Integrity	High (ACID)	Low
Scaling	Nonlinear	Linear

Reference: Tom White's Hadoop: The Definitive Guide

Programming HDInsight – Existing ecosystem

Hive

Pig

Mahout

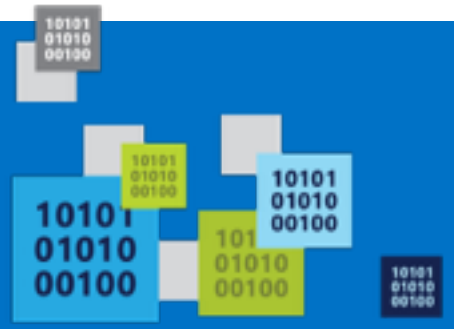
Cascading

Scalding

Scoobi

Pegasus

…



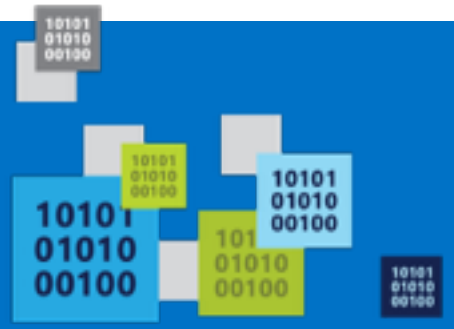
Programming HDInsight
– Microsoft .NET

C#

F#

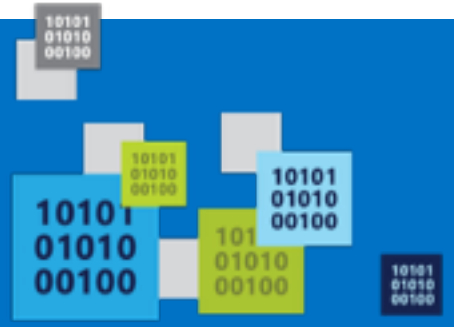
Map/Reduce

Microsoft .NET management clients

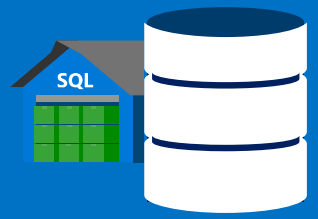


Programming HDInsight
– DevOps / IT Pros

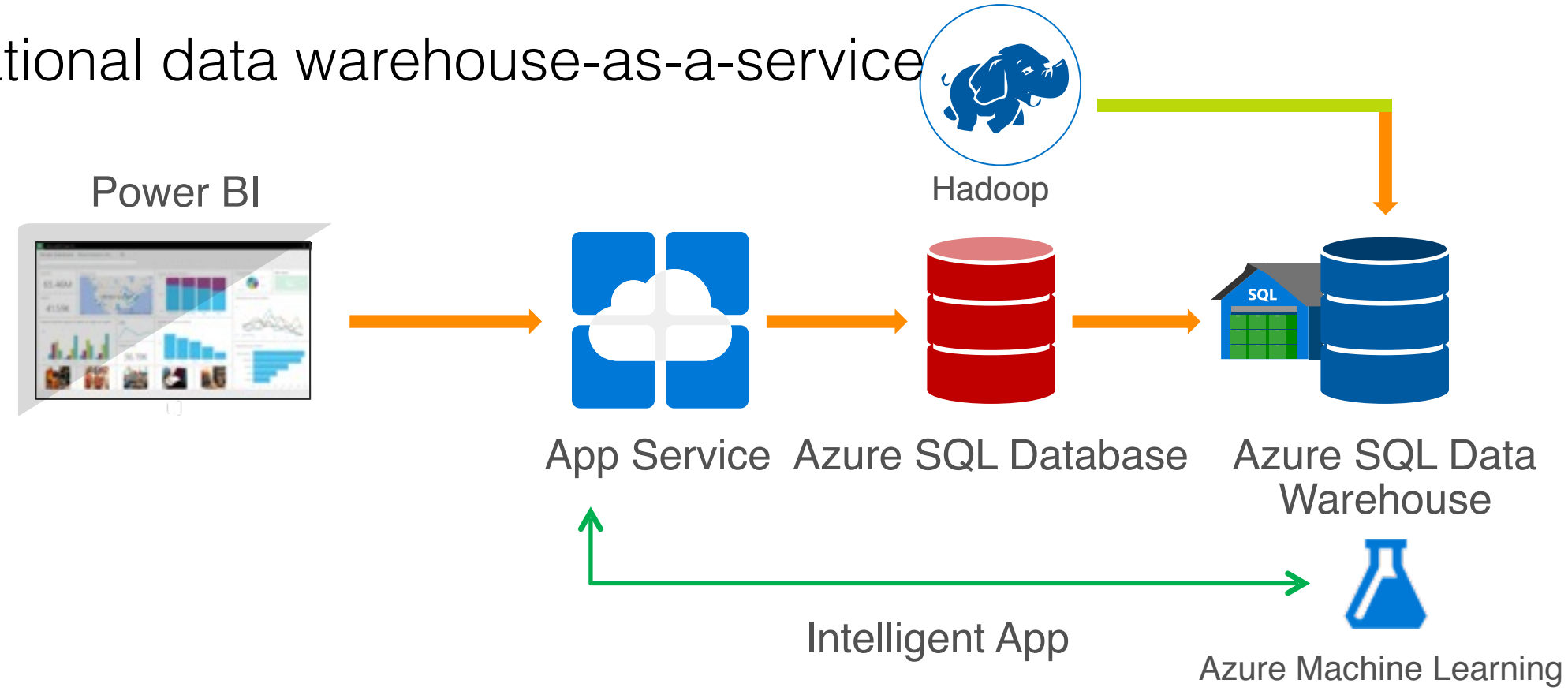
PowerShell
Cross-Platform CLI tools (xplat-cli)



Data Warehousing



A relational data warehouse-as-a-service



Elastic scale & performance

Scales to petabytes of data

Massively Parallel Processing

Instant-on compute scales in seconds

Query Relational / Non-Relational

Powered by the Cloud

Get started in minutes

Integrated with Azure ML, PowerBI & ADF

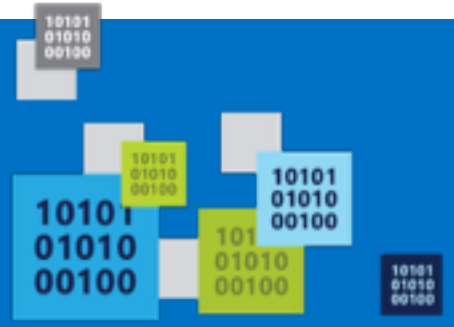
Enterprise Ready

Market Leading Price & Performance

Simple billing compute & storage

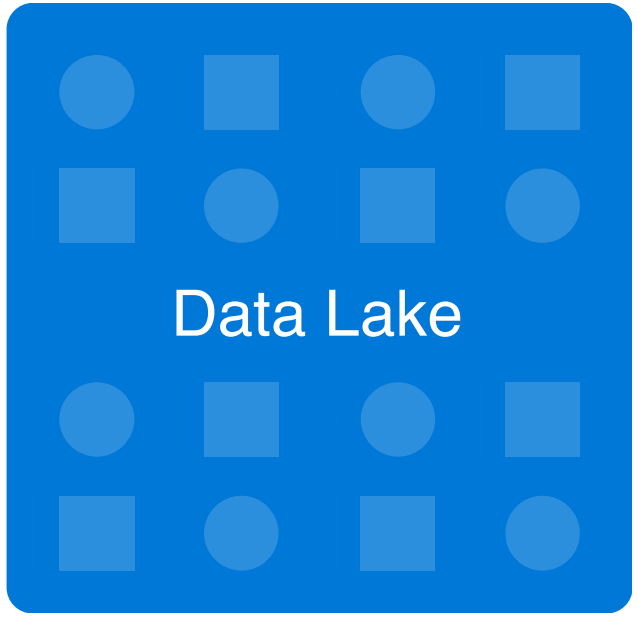
Pay for what you need, when you need it with dynamic pause

Bring DW to the Cloud without rewriting



Data Lake

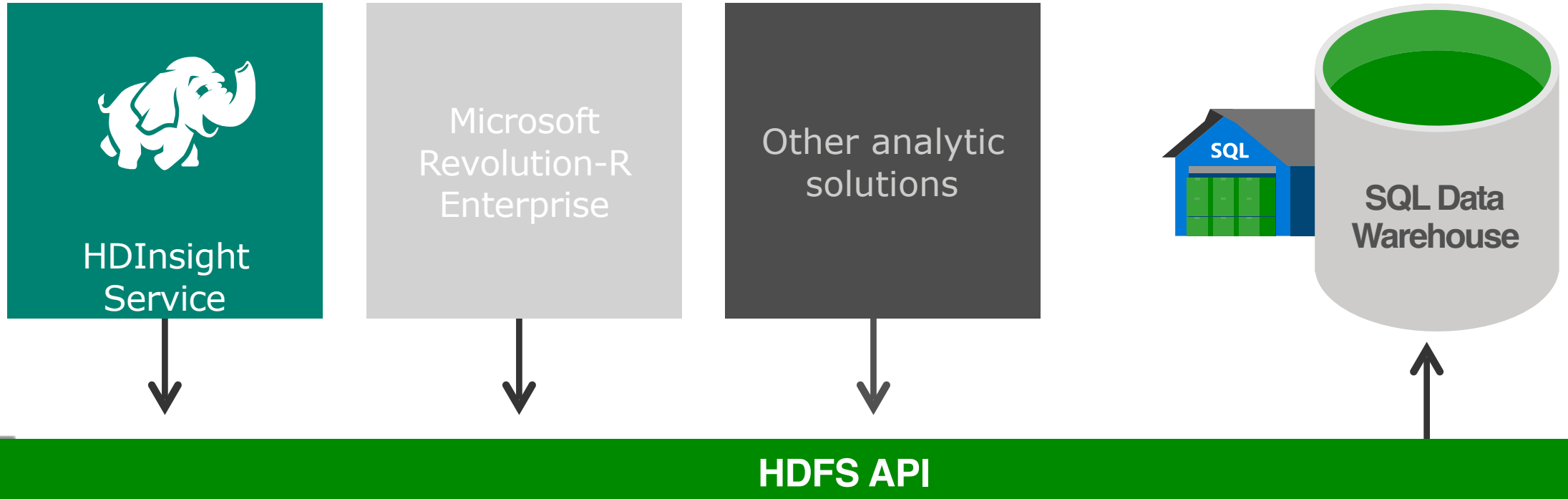
Azure Data Lake service



- Store and manage infinite data
- Keep data in its original form
- High through put, low latency analytic jobs
- Enterprise-grade security + access control

Data Lake service

Transformative way to store and process infinite data



MongoDB, MySQL, Oracle, Cassandra, Neo4j and more

Additional Database options in Azure

- Azure Table Service is a “Big Table” entity store.
- MongoDB is a document (JSON) store.
- Cassandra is a columnar store with excellent replication.
- HBase is a Big Data (Hadoop) store available in HDInsight.
- Oracle VMs are supported in Azure.
- MySQL is offered from the partner ClearDB.

Microsoft

Microsoft Azure Data Platform

Microsoft Azure

SQL Database

SQL on IaaS

DocumentDB

MongoDB, MySQL, Oracle, Cassandra, Neo4j and more

HDInsight

Search

Microsoft