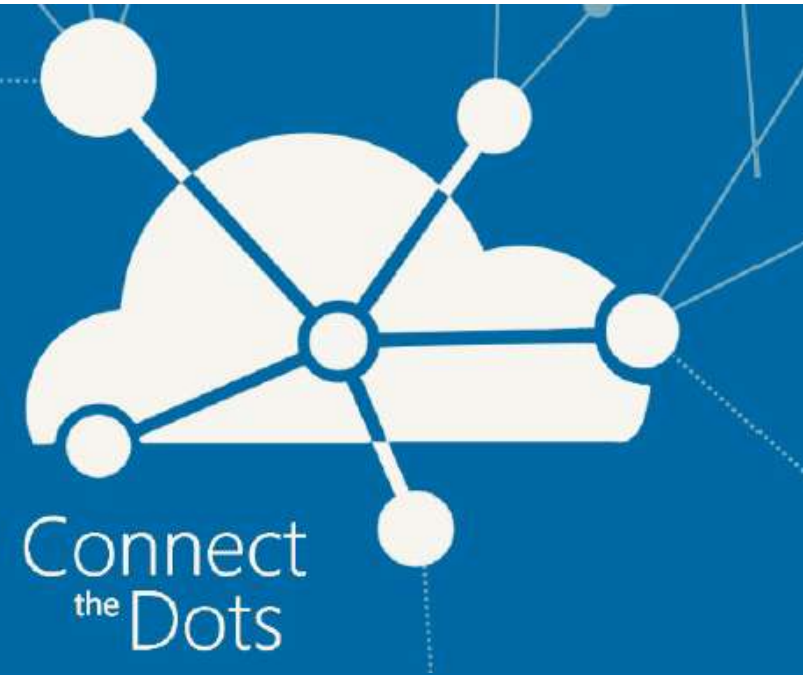
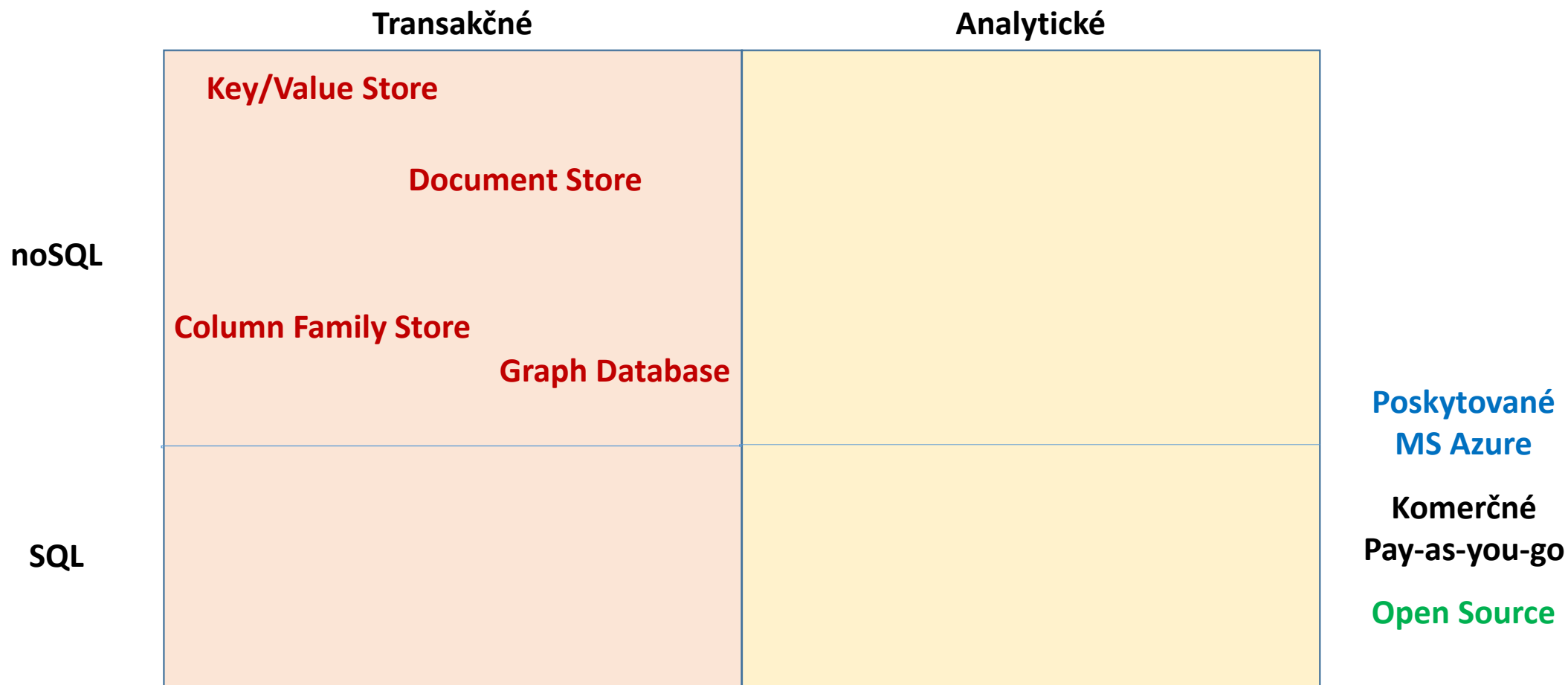


Stanislav Harvan
v-sharva@microsoft.com



Databázy v Azure


































Databázy v Azure



Databázy v Azure

	Transakčné	Analytické	
noSQL	<p>Key/Value Store (Tables, Redis, ...)</p> <p>Document Store (DocumentDB, MongoDB, RavenDB, CouchBase, ...)</p> <p>Column Family Store (HBase, Cassandra, ...)</p> <p>Graph Database (Neo4j, ...)</p>	<p>Big Data Analytics (HDInsight, Hadoop)</p>	Poskytované MS Azure
SQL	<p>Relational Database (SQL Database, SQL Server, Oracle, DB2, MySQL, MemSQL, ...)</p>	<p>Relational Analytics (SQL Data Warehouse, SQL Server, Oracle, MySQL, ...)</p>	

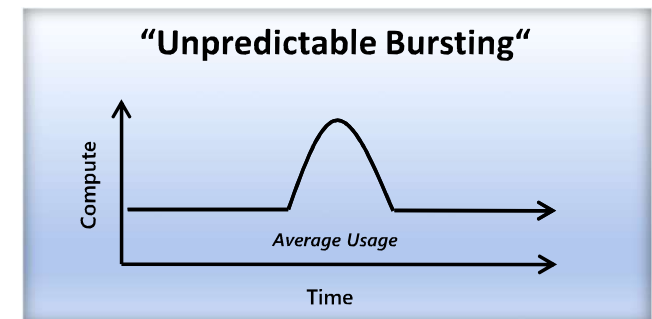
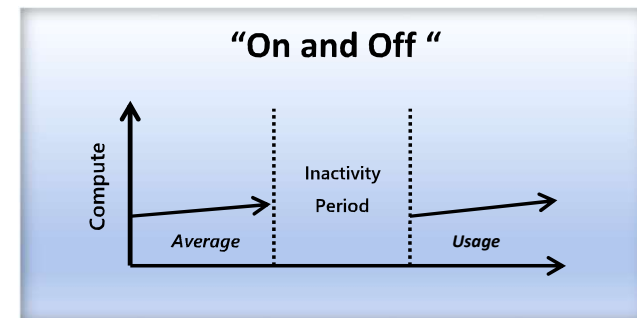
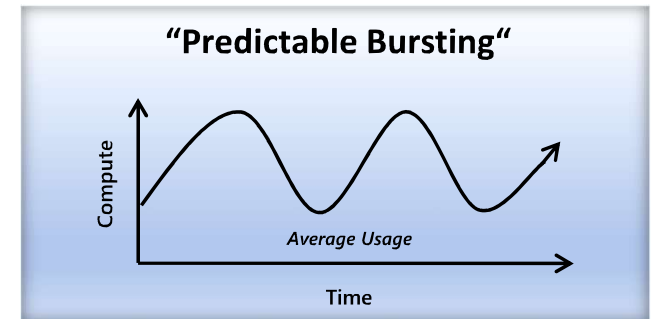
Typické přínosy cloudu

 = Managed for You	Standal one Servers	IaaS	PaaS	SaaS
Applications				
Runtimes				
Database				
Operating System				
Virtualization				
Server				
Storage				
Networking				

Optimalizácia licenčných poplatkov a poplatkov za podporu

„Pay-as-you-go“

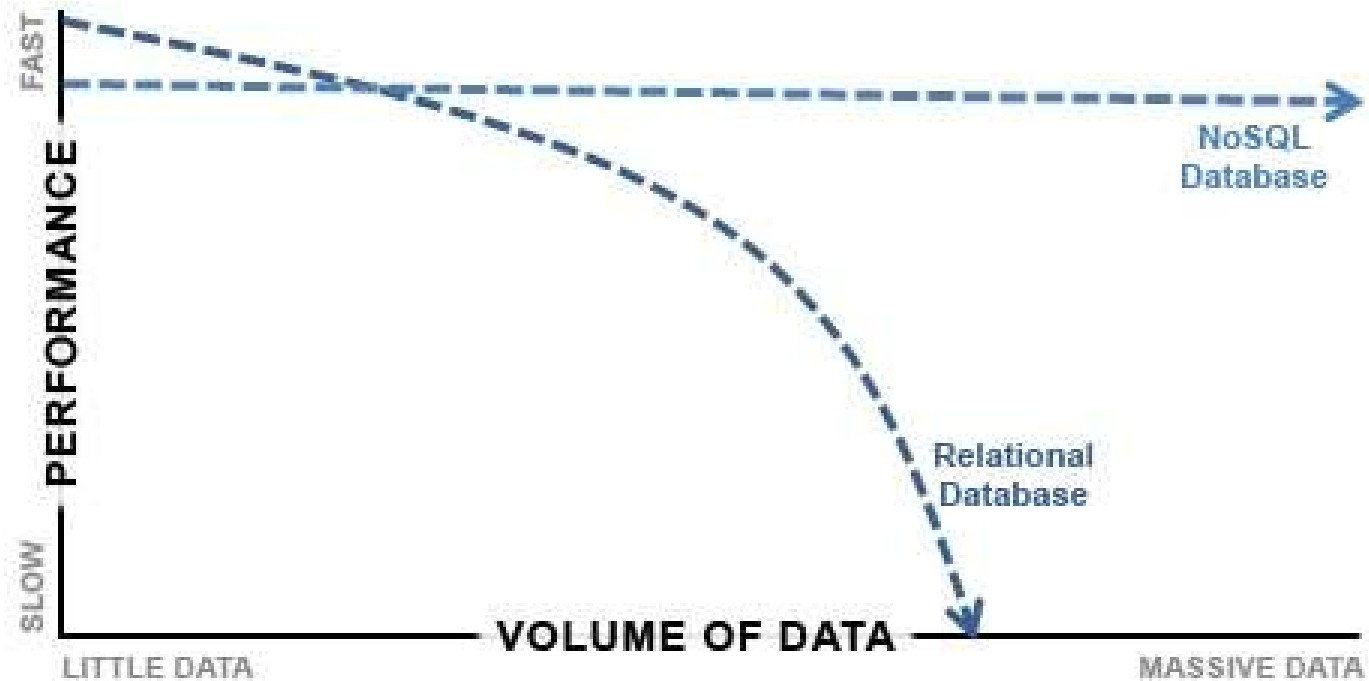
- K dispozícii predkonfigurované VM s databázovým SW
- Cena za používanie databázového SW je uvedená za časové obdobie (minúta, hodina...)
- Technickú podporu pre databázový software je k dispozícii cez Windows Azure Support plan



Databázy v Azure

	Transakčné	Analytické	
noSQL	<p>Key/Value Store (Tables, Redis, ...)</p> <p>Document Store (DocumentDB, MongoDB, RavenDB, CouchBase, ...)</p> <p>Column Family Store (HBase, Cassandra, ...)</p> <p>Graph Database (Neo4j, ...)</p>	<p>Big Data Analytics (HDInsight, Hadoop)</p>	Poskytované MS Azure
SQL	<p>Relational Database (SQL Database, SQL Server, Oracle, DB2, MySQL, MemSQL, ...)</p>	<p>Relational Analytics (SQL Data Warehouse, SQL Server, Oracle, MySQL, ...)</p>	

Škálovatelnost RDBMS a NoSQL



NoSQL

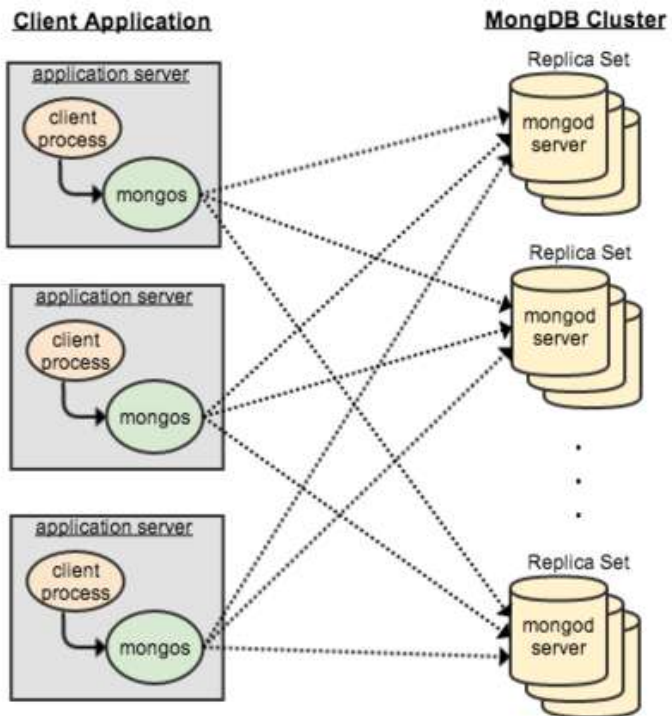
- Amazon -> [Dynamo](#)
- Google -> [Bigtable](#)

- Nerelačné databázy
- Cluster-friendly
- open-source
- nevynucujú fixnú schému
- objavili sa v poslednom desaťročí s rozšírením Webu

Prínosy NoSQL

- jednoduchší vývoj aplikácie
- neobmedzené škálovanie výkonu
- nižšia latencia
- efektívnejšia práca s vybranými typmi dát
- vyššia dostupnosť

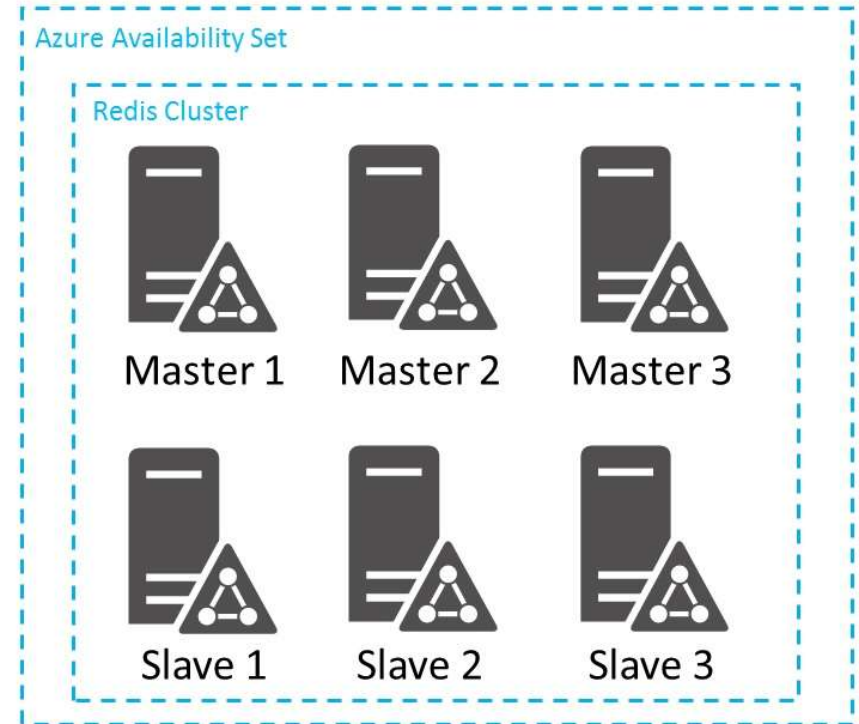
Rôzne prístupy ku clustrom



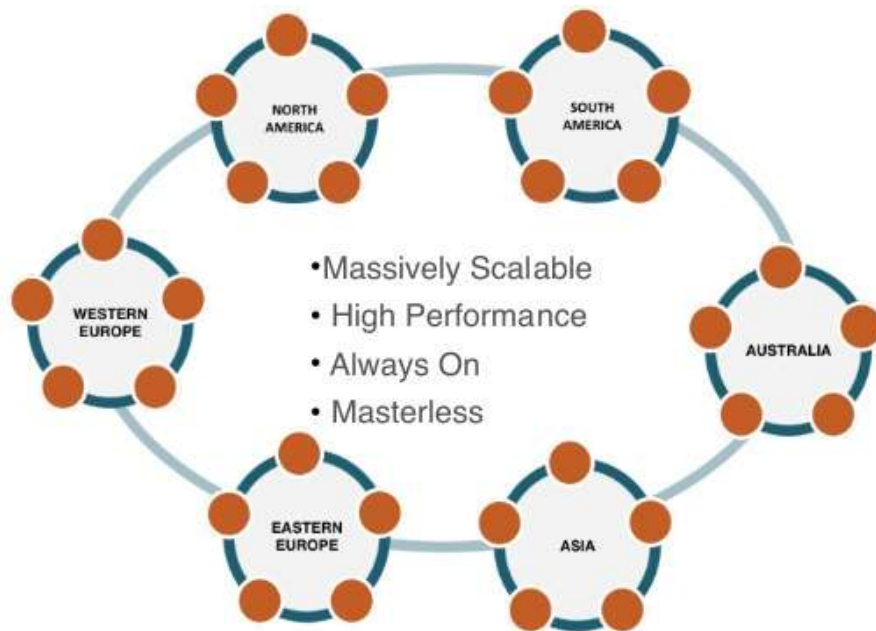
single mongos router per client



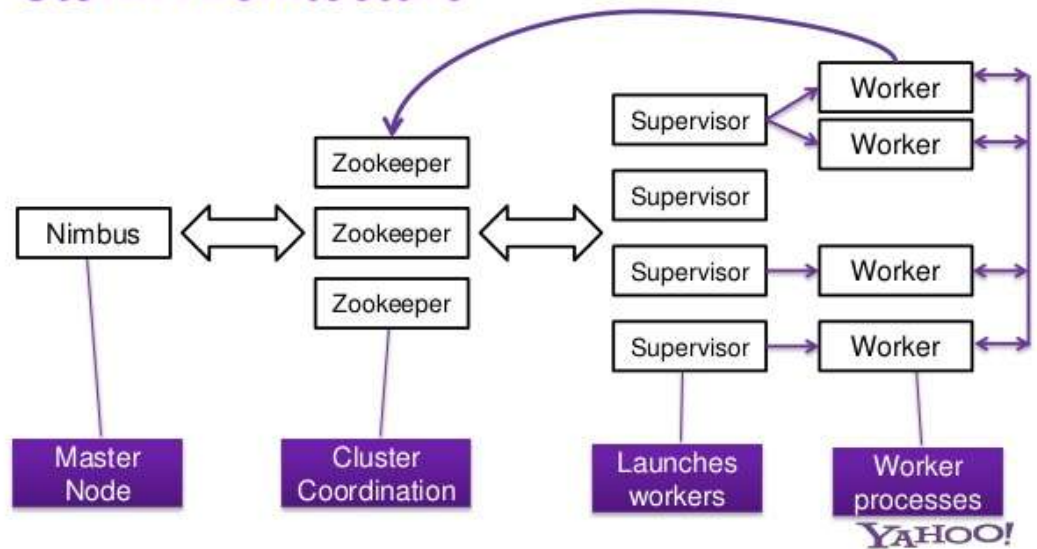
Jump Box



Rôzne prístupy ku clustrom



Storm Architecture



CAP theorem

ACID

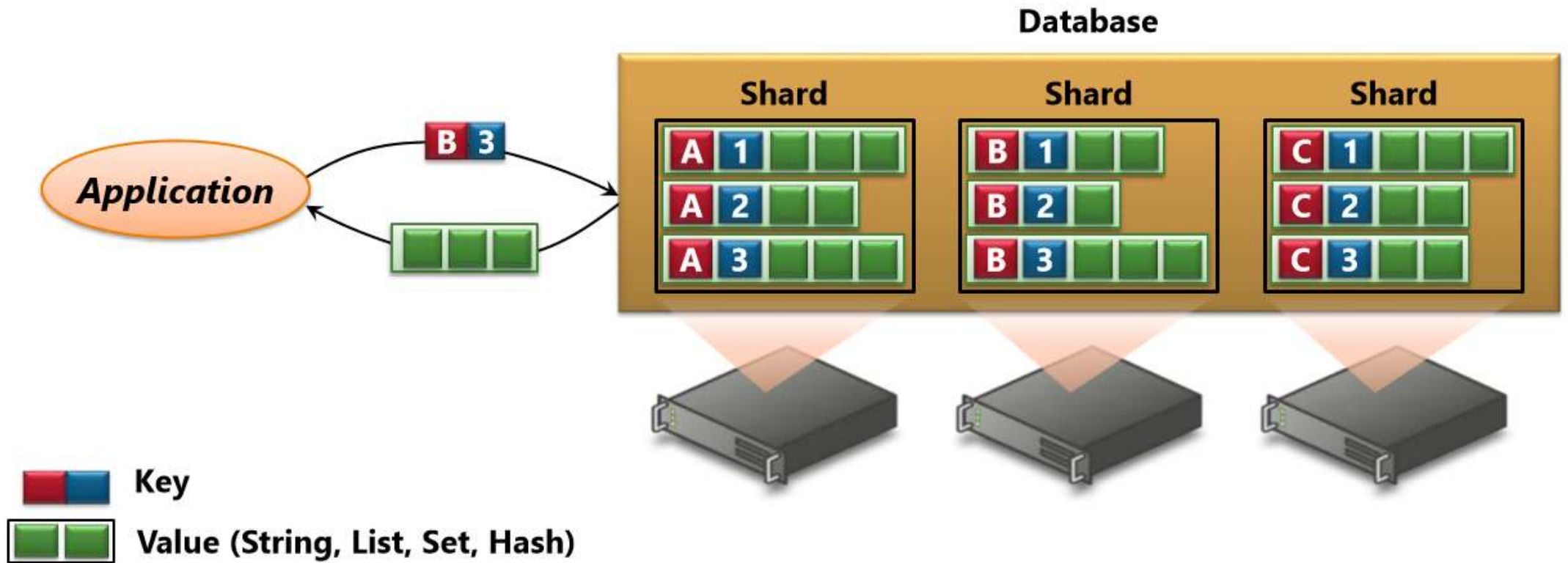
- **Atomic**
 - All operations in a transaction succeed or every operation is rolled back
- **Consistent**
 - On the completion of a transaction, the database is structurally sound
- **Isolated**
 - Transactions do not contend with one another.
- **Durable**
 - The results of applying a transaction are permanent, even in the presence of failures

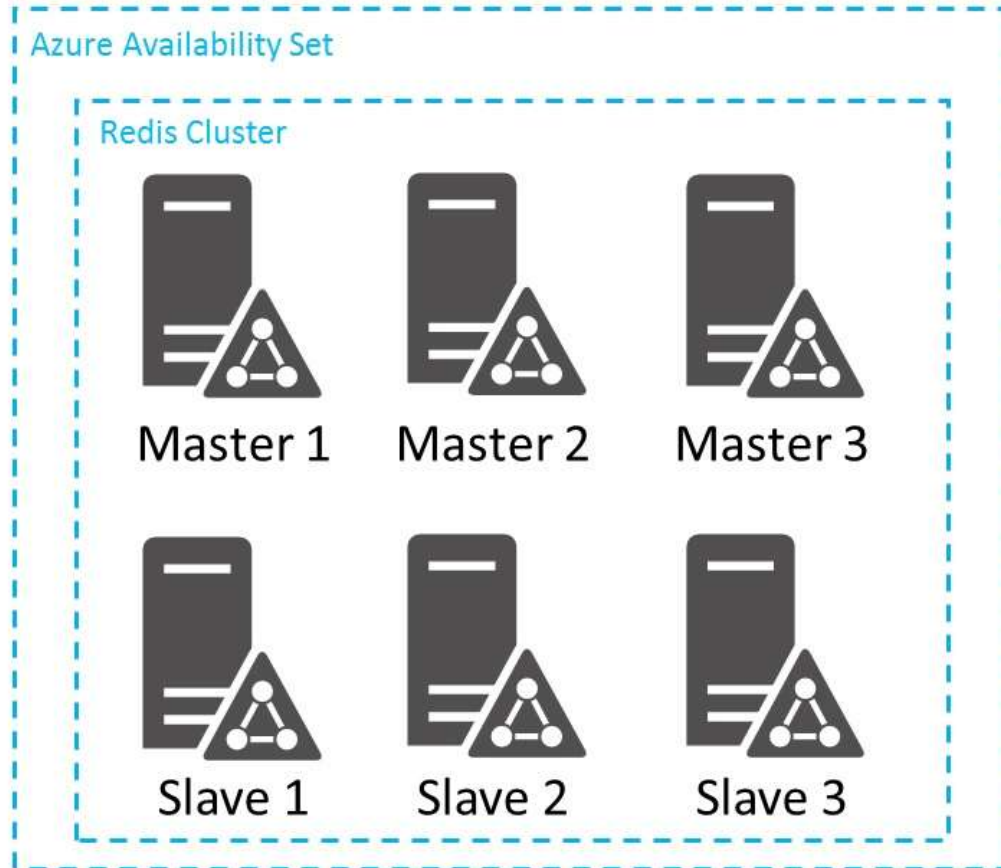
BASE

- **Basic Availability**
 - The database appears to work most of the time
- **Soft-state**
 - Stores don't have to be write-consistent, nor do different replicas have to be mutually consistent all the time
- **Eventual consistency**
 - Stores exhibit consistency at some later point

Key/value stores

Azure Tables, Redis...



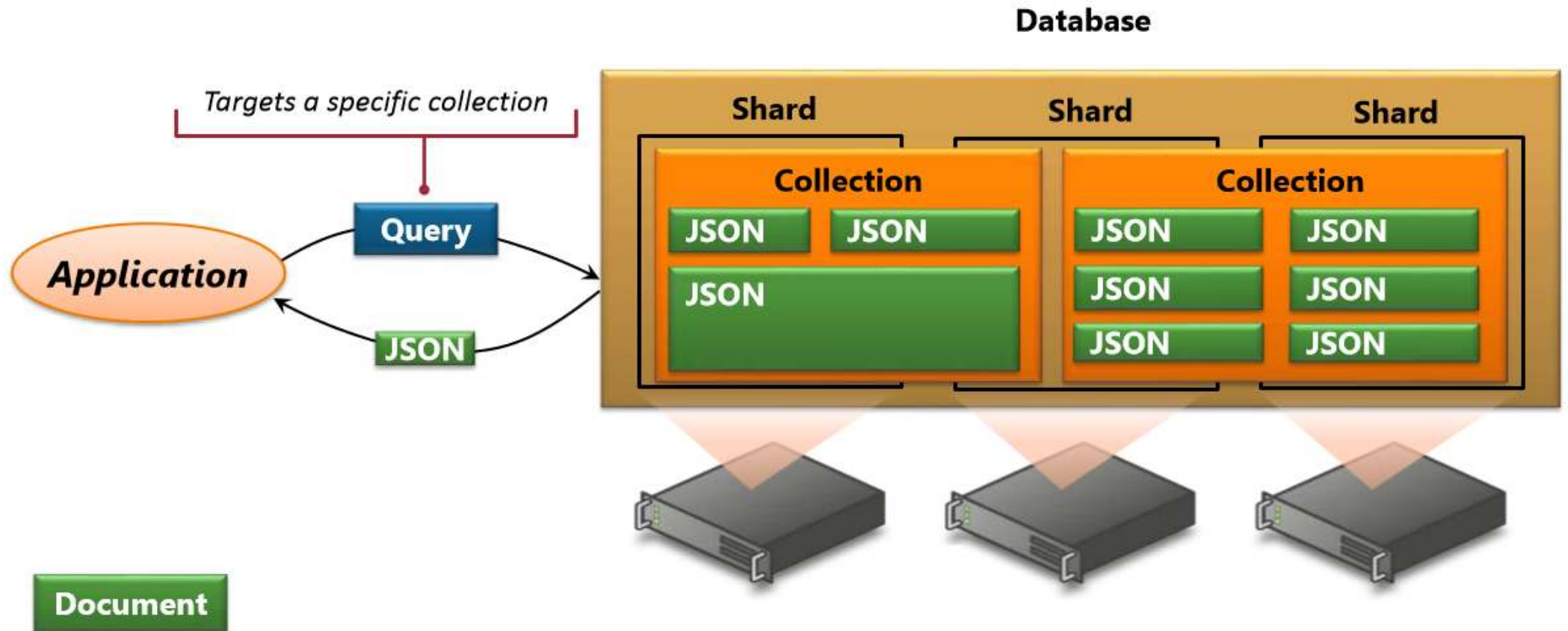


Redis

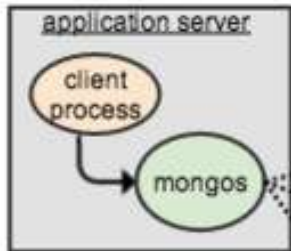
- [Redis playground](#)
- useful for storing session information, user profiles, preferences, shopping cart data
- avoid using when we need to query by data, have relationships between the data being stored or we need to operate on multiple keys at the same time

Document databases

MongoDB, DocumentDB, CouchBase, RavenDB...



Client Application



MongDB Cluster

Replica Set



Replica Set

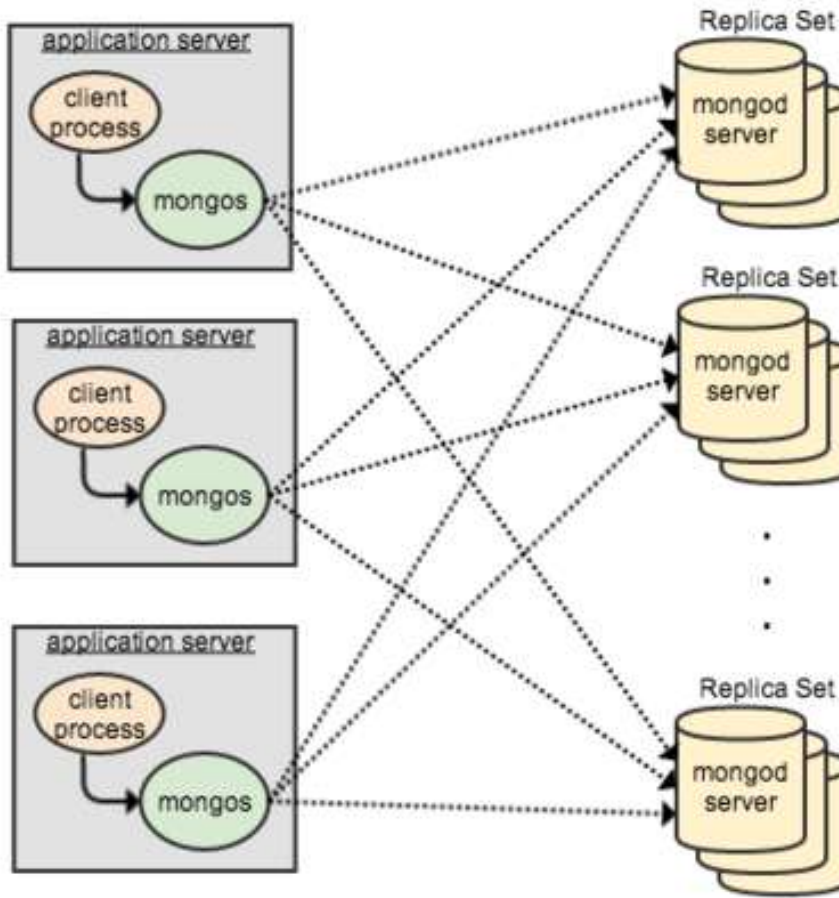


...

Replica Set



single mongos router per client

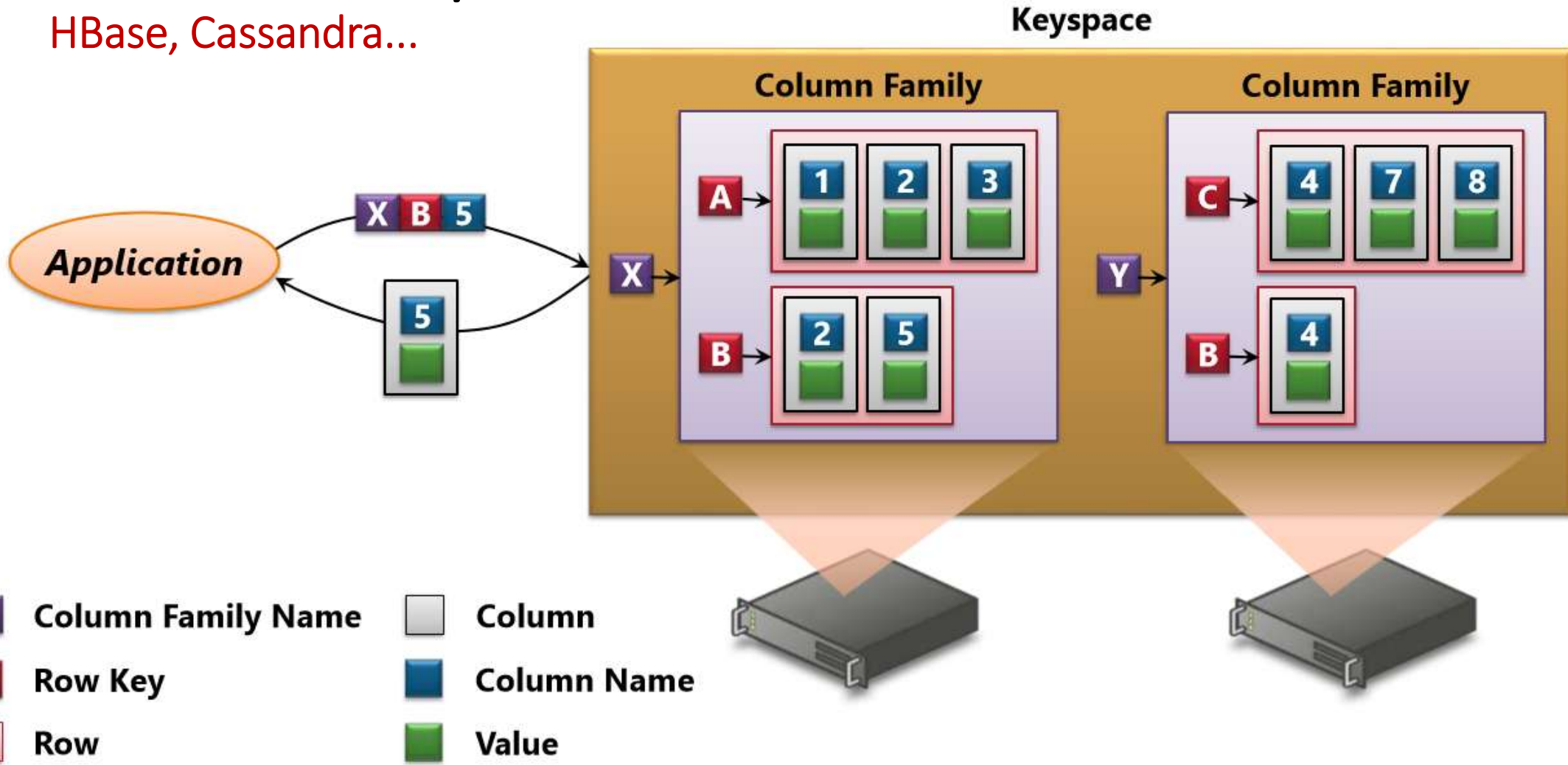


DocumentDB

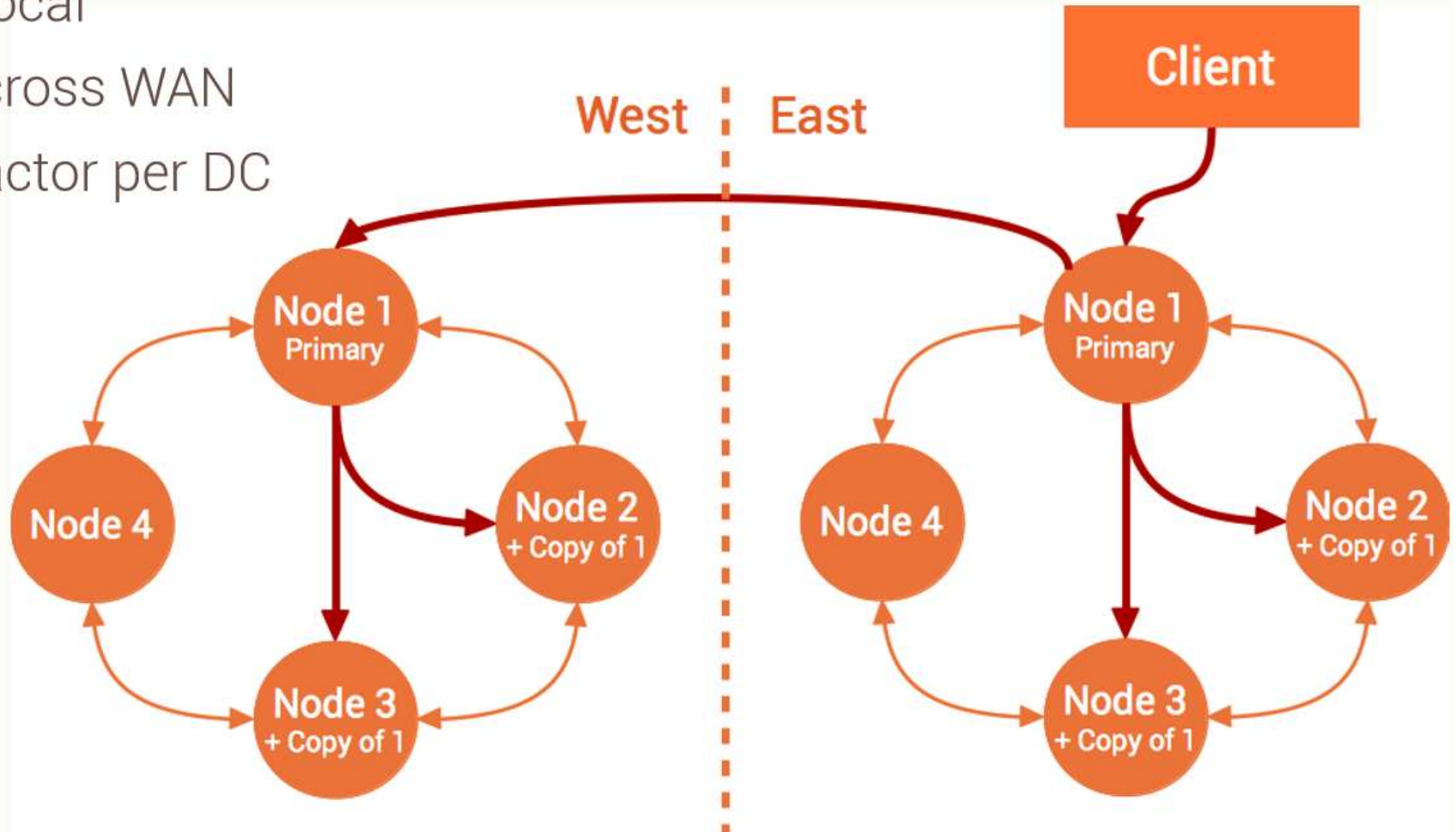
- [DocumentDB demo](#)
- useful for content management systems, blogging platforms, web analytics, real-time analytics, ecommerce-applications
- avoid using document databases for systems that need complex transactions spanning multiple operations or queries against varying aggregate structures

Column family stores

HBase, Cassandra...



- Client writes local
- Data syncs across WAN
- Replication Factor per DC

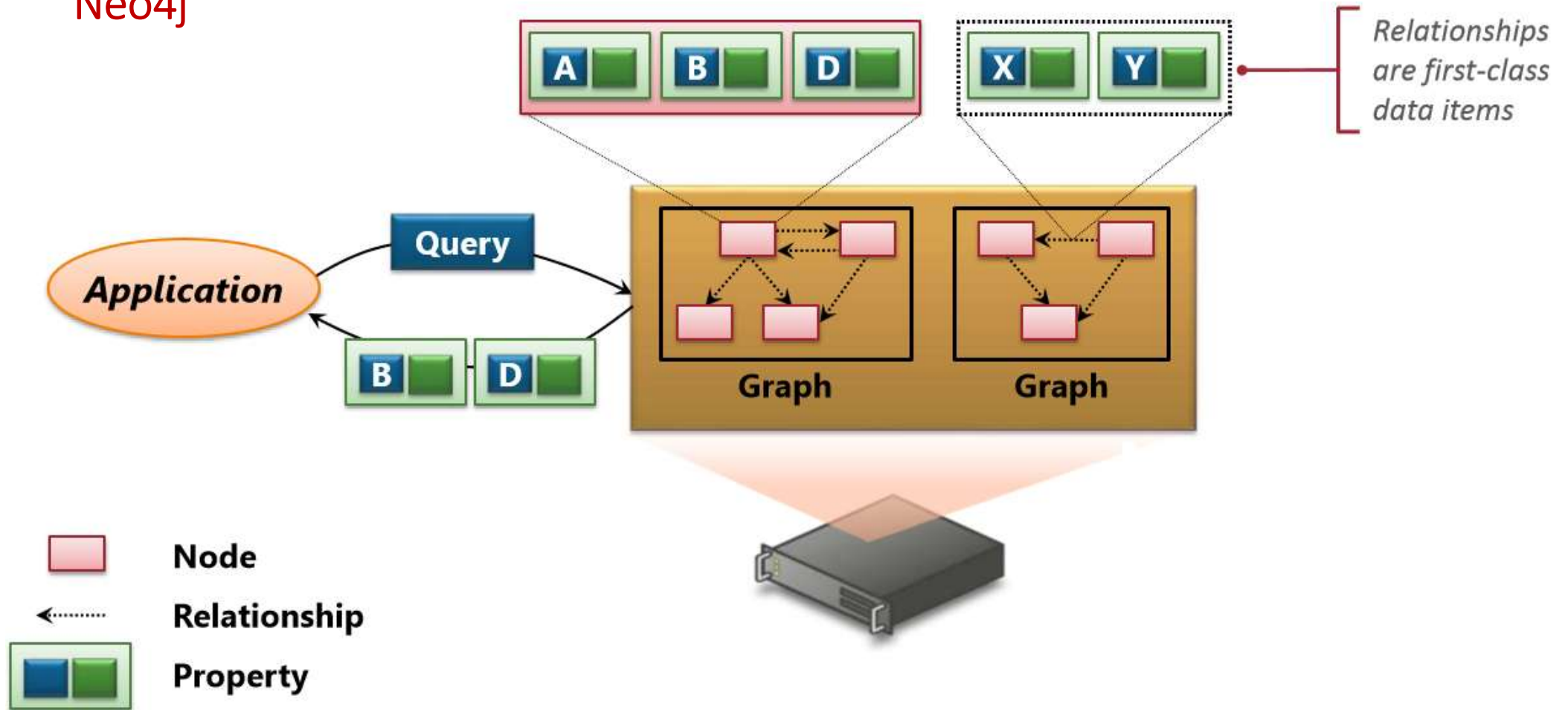


Cassandra

- [Datastax in Azure tutorial](#)
- useful for content management systems, blogging platforms, maintaining counters, expiring usage, heavy write volume such as log aggregation
- avoid using column family databases for systems that are in early development, changing query patterns

Graph databases

Neo4j



Neo4j

- [Neo4j use cases](#)
- suited to problem spaces where we have connected data, such as social networks, spatial data, routing information for goods and money, recommendation engines

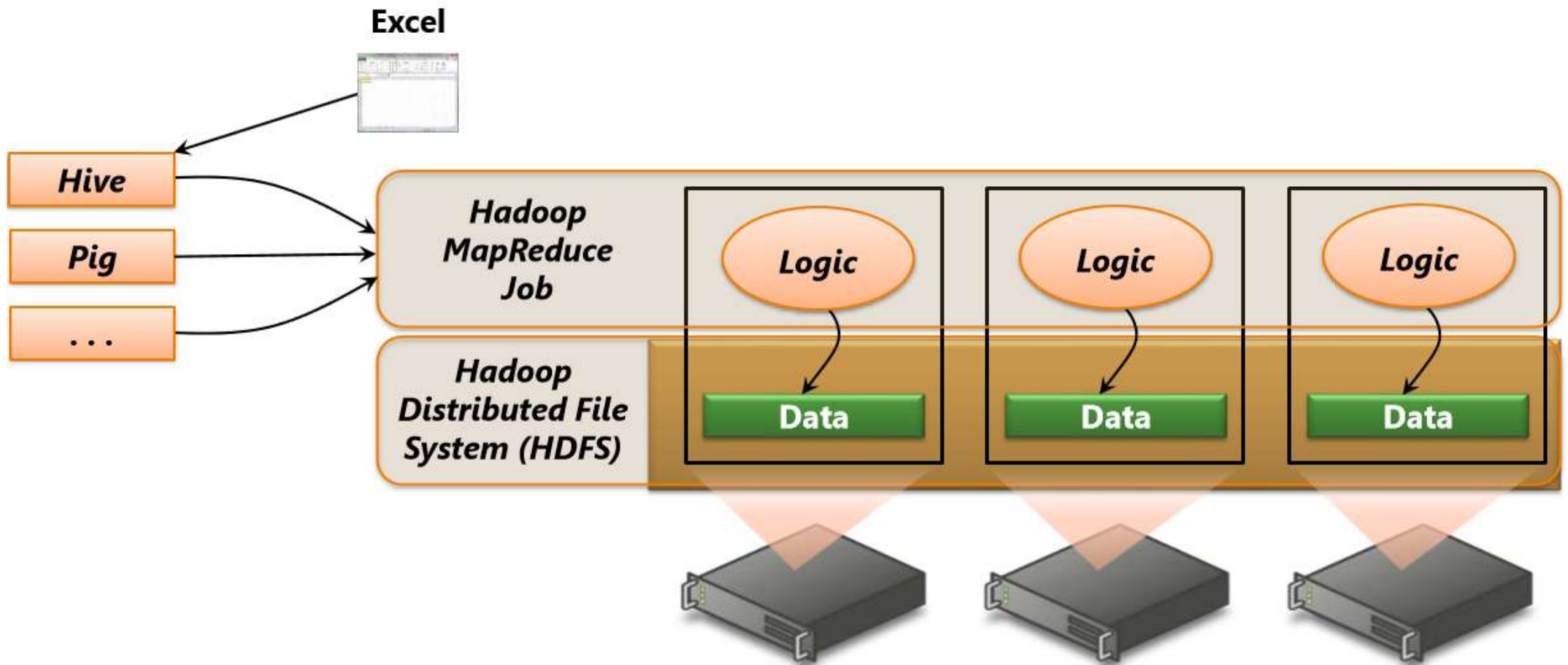
Dátové úložiska v cloude

	Transakčné	Analytické	
noSQL	<p>Key/Value Store (Tables, Redis, ...)</p> <p>Document Store (DocumentDB, MongoDB, RavenDB, CouchBase, ...)</p> <p>Column Family Store (HBase, Cassandra, ...)</p> <p>Graph Database (Neo4j, ...)</p>		
SQL			<p>Poskytované MS Azure</p> <p>Komerčné Pay-as-you-go</p> <p>Open Source</p>

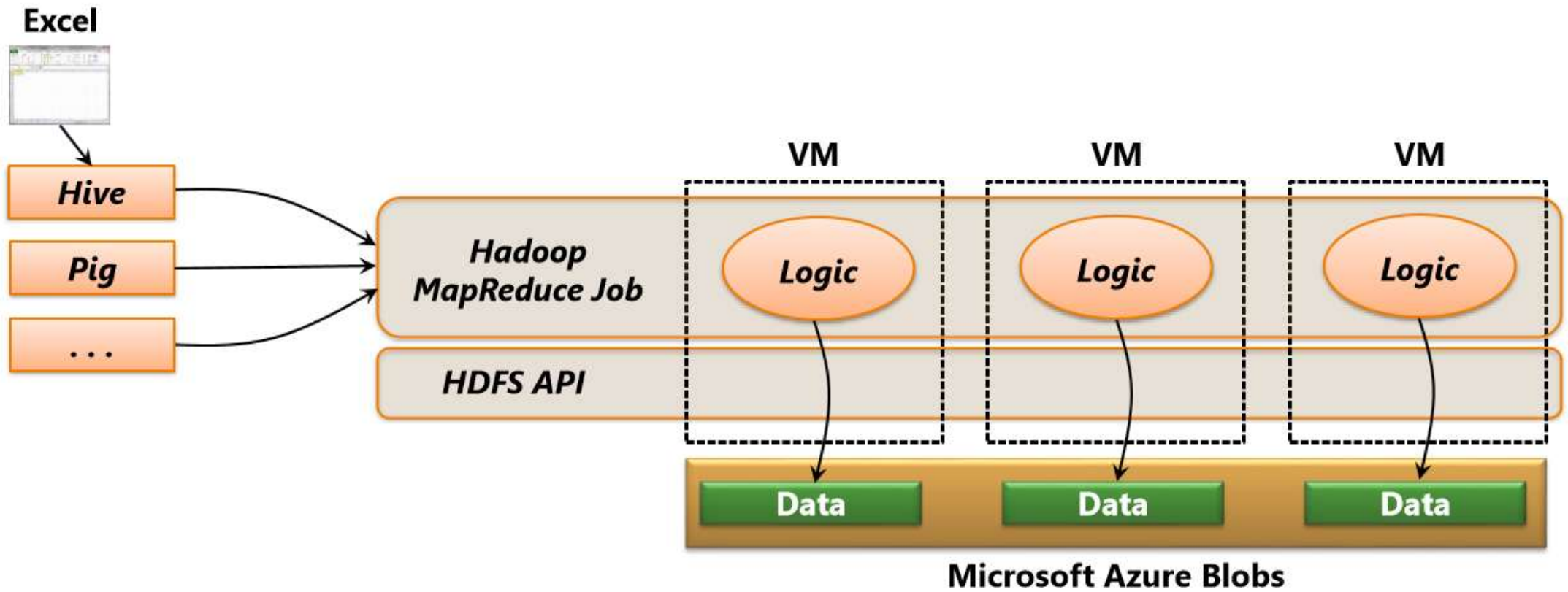
Dátové úložiska v cloude

	Transakčné	Analytické	
noSQL	<p>Key/Value Store (Tables, Redis, ...)</p> <p>Document Store (DocumentDB, MongoDB, RavenDB, CouchBase, ...)</p> <p>Column Family Store (HBase, Cassandra, ...)</p> <p>Graph Database (Neo4j, ...)</p>	<p>Big Data Analytics (HDInsight, Hadoop)</p>	<p>Poskytované MS Azure</p>
SQL			<p>Komerčné Pay-as-you-go</p> <p>Open Source</p>

Big Data – Hadoop

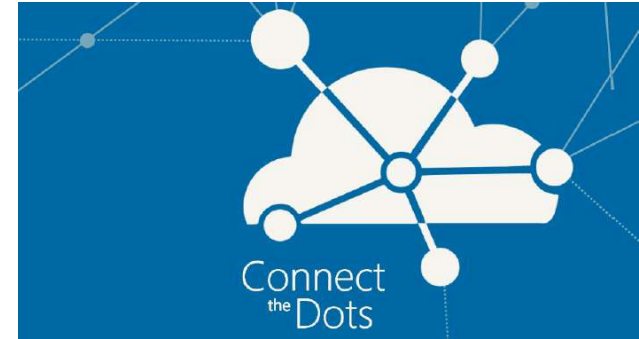


Big Data – HDInsight



Ďalšie zdroje

- [NoSQL Databases: An Overview](#)
- [ACID vs. BASE](#)
- [“Joining” two collections in MongoDB using MapReduce](#)
- [Get started using Hadoop in HDInsight](#)



Q&A