



Building Enterprise-Grade Blockchain Applications with Azure Cosmos DB

February 2019

Executive Summary

Blockchain and distributed ledgers are frequently mentioned in digital transformation discussions across every industry. The core technical problem the blockchain technology tries to address is *trustless multi-party collaboration*. When a blockchain is appropriate for a particular scenario, it is typically complemented by a mature database technology.

In addition to these, there are a number of scenarios where an organization may believe they need a blockchain, but where, in fact, it is not required. We have seen many high-value applications and business processes that can benefit from Azure Cosmos DB to implement these scenarios. Azure Cosmos DB implements the core characteristics that are sought after in a distributed ledger technology: global distribution, elastically scalable storage and throughput, guaranteed low latency with rich querying capabilities, and integration with familiar tools.

This paper provides an overview of both blockchain and Azure Cosmos DB and highlights how Azure Cosmos DB can be used to deliver value in many of the same scenarios.

Blockchains and Smart Contracts

Blockchains are shared, secure, ledgers of transactions distributed among a network of computers. At its core, a blockchain is a data structure, similar in many ways to a distributed database system, but with a cryptographically immutable ledger or log. The result is a shared, immutable record, which enables people and organizations to collaborate more efficiently. As a shared source of trust, it extends the scope of digital transformation from a single company to the processes and assets it shares with its suppliers, customers, and partners.

Many blockchains support *smart contracts* that enable organizations to write and execute code that defines the state, behaviors, and actions of business processes or assets. With smart contracts, organizations can model a business process and have multiple parties interact with it on an immutable ledger.

Public and Consortium Blockchain Networks

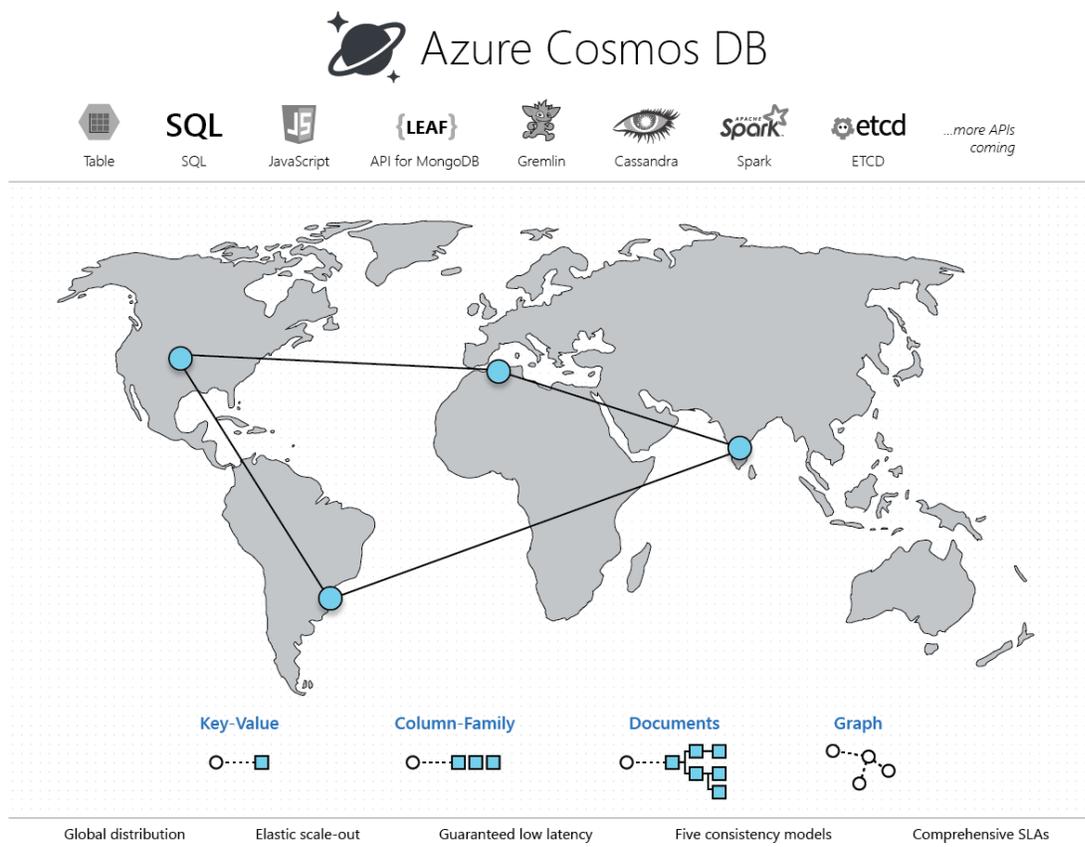
While public networks, e.g., *Bitcoin* and *Ethereum*, are most well-known, enterprise customers also deploy private networks for use within consortiums. Depending on the level of trust within a consortium and the blockchain technology stack selected, different consensus algorithms can be used to deliver increased throughput in consortium blockchain networks. Consortium blockchains are also different in that they often support confidentiality, allowing transactions and contracts to be shared only between appropriate parties.

While very high throughput scenarios historically disqualified blockchain as an option and not all ledgers support confidentiality, Microsoft has developed the [Confidential Consortium Blockchain Framework](#) (CCBF) which can be integrated into ledgers to improve performance, to enable confidentiality, and to enable consortium governance.

Azure Cosmos DB

[Azure Cosmos DB](#) is Microsoft's globally distributed, fully-decentralized, multi-mastered database service for mission-critical applications. Azure Cosmos DB provides turnkey global distribution with multi-master replication, [elastic scaling of throughput and storage](#) worldwide, **single-digit millisecond** write and read latencies at the 99th percentile, [five well-defined consistency models](#), and guaranteed **99.999** high availability, all backed by [industry-leading comprehensive SLAs](#).

Azure Cosmos DB [automatically indexes all data](#) without requiring customers to deal with schema or index management. It is a multi-model service and natively supports document, key-value, graph, and column-family data models through the familiar APIs and syntax of SQL, MongoDB, Cassandra, Tables, or Gremlin. As a natively born in the cloud service, Azure Cosmos DB is carefully engineered with global distribution, horizontal scalability and multi-tenancy from the ground up, and can scale throughput and storage across regions, support a high rate of data ingestion and distribute that data to where your users are.

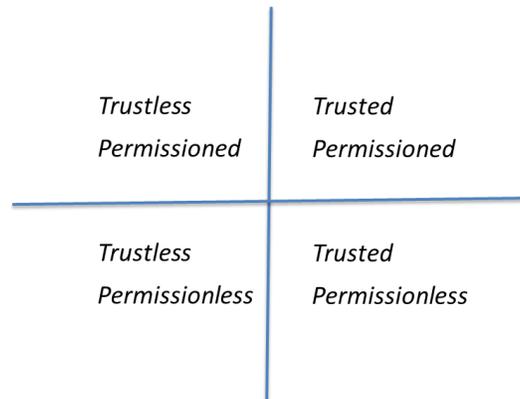


Unlike centralized “single-site” databases, Azure Cosmos DB is fully decentralized by virtue of horizontal partitioning and global distribution. The multi-master replication enables shifting the computation away from a centrally located database to a datacenter or an edge device closest/local to the user, wherever the user may be.

The Azure Cosmos DB multi-master replication protocol is designed to scale across 100s of datacenters and trillions of edge devices. The architecture treats an Azure region or an edge device as equals – both are capable of hosting Azure Cosmos DB replicas and participate as true-peers in the multi-master replication protocol. Azure Cosmos DB also has rich audit log functionality and can support confidentiality from parties via multiple containers and encryption.

Trust and Permissions

To help determine whether a blockchain is appropriate for a solution, the combination of requirements for *trust* and *permissions* should be evaluated.



- **Trusted and permissioned** collaboration involves a set of participants well known to each other and where each participant trusts the data provided by the others. An example scenario includes the tracking of changes to employee profiles in a human resource (HR) system. This may be shared across geographies and systems and may even be shared across organizations, but there is single trusted provider of the information and centralized access control.

While having an immutable ledger that tracks the changes to profiles over time is important, this scenario does not require a blockchain. Azure Cosmos DB can deliver the core capabilities in this scenario and do so in a way that is compatible with existing reporting and analytics tools. In addition, Azure Cosmos DB Gremlin API could also enable developers to build robust graph queries that could unlock additional business value.

- **Trustless and permissioned** scenarios involve federations of companies, or consortiums, where multiple parties may not trust one another, but participate in a shared workflow. In these scenarios, the blockchain is often used to track state changes to a physical, financial or digital asset. The consortiums build shared applications on a blockchain to monitor the movement of those assets as they move across organizational trust boundaries and govern state changes with “on chain” logic in permissioned smart contracts. Use cases cover a range of scenarios that are asset and workflow-centric, from supply chain provenance and traceability to financial securities trading and insurance.

Here, blockchain-powered solutions are typically complemented with one or more additional database technologies. Azure Cosmos DB provides global distribution, global replication of blockchain data with low latency, robust storage of the data, as well as providing the types of sophisticated querying capabilities blockchain solutions lack.

When *augmenting a blockchain with a database technology*, it's important to understand which transactions are required, where they can be found, and what data should be distributed to what members in the "off chain" world. When trying to operationalize blockchains, a database solution like Azure Cosmos DB becomes crucial and serves essential functionality as the underlying data management layer to handle the following considerations:

- **Confidentiality:** If confidentiality is employed on the blockchain, Company A and B may have confidential transactions that are not delivered to the node(s) of Company C. This impacts where data must be retrieved from and to whom it should be made available
 - **Compliance:** In many cases, there is data that is inappropriate to be in a blockchain, such as personally identifiable information (PII) or data governed by the legislation such as the General Data Protection Regulation (GDPR) in Europe.
 - **Developer Experience:** Rationalizing the skills of an organization and their ability to readily use a given technology is also important. Azure Cosmos DB supports SQL interfaces which address much of the market, as well as APIs for MongoDB, Cassandra, and Gremlin and syntax that may be more familiar to certain developers.
 - **User Experience:** An emerging theme is the use of search, such as Azure Search or Elasticsearch, when building client applications for blockchain-powered solutions. These technologies deliver capabilities such as [facets](#) that can enable an excellent data navigation experience in a customer user interface. When used, Search is used in conjunction with Azure Cosmos DB.
- **Trustless and permissionless** scenarios that are most well-known include cryptocurrencies and exchanges. Enterprise applications are nascent, due in part to low throughput, high latency, and high costs associated with trustless and permissionless solutions.

[Xbox Royalties – Blockchain + Azure Cosmos DB case study](#)

The recent work by [Xbox and its game publishers](#) show how blockchain-based smart contracts can be used to codify complex royalty payment agreements. Metadata on millions of Xbox Live purchases are sent to the ledger network as a tamper-proof transactional log. The solution augments the blockchain by copying purchase data to Azure Cosmos DB.

Game publishers, who previously had to wait 45 days for reconciled royalty information, now have access to an off-chain copy of immutable purchase data that is updated in near real-time and can be queried/visualized or connected to consuming accounting applications.

Azure Cosmos DB was the off-chain database of choice for Xbox, because of its turn-key global distribution capability, single-digit millisecond guaranteed latency, elastic scalability of storage and throughput and multi-master write capabilities. Azure Cosmos DB ensured responsive, real-time data access to royalties data throughout the world.

Conclusion

Whether providing an alternative to blockchain for many categories of solutions or complementing a blockchain as an off-chain data store, Azure Cosmos DB can provide a massive scale, globally distributed multi-model database service that can meet the needs of both.

There have been several analyses of the capability and the use-cases of blockchains that present simple flowcharts to define the real use-cases for blockchains. These analyses make the point that globally distributed databases are in fact the correct solution to use for many perceived applications of blockchains as shown below.

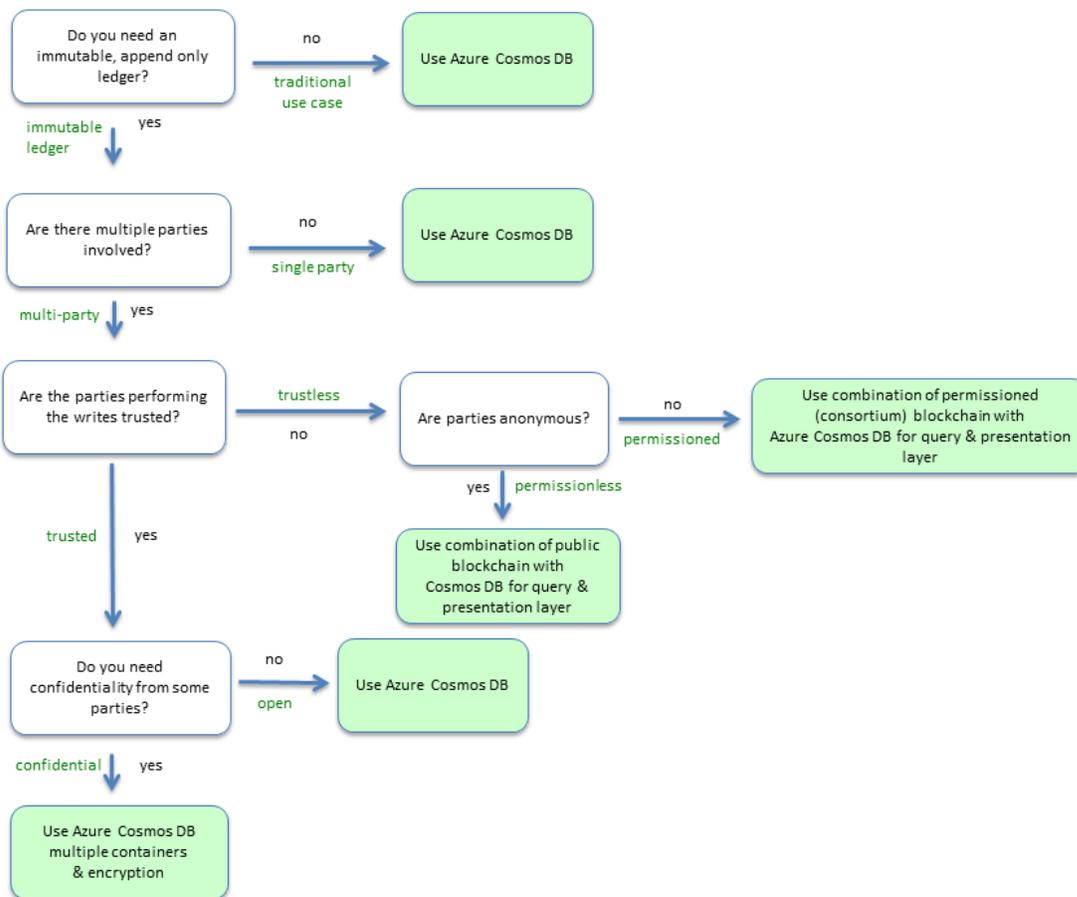


Figure 1: Blockchain workflow with Azure Cosmos DB

We can help

An increasing number of enterprises across all industry sectors are now exploring how they can use blockchain technology to remove friction from business processes and build systems of trust for value exchange. A fully-managed, enterprise-grade, globally-distributed, scalable and secure database service such as Azure Cosmos DB are core to unlocking the potential.

Contact the Azure Cosmos DB team (AskCosmosDB@microsoft.com), and we can schedule a workshop with you to explore how to build a blockchain application with Azure Cosmos DB.

Resources

- [Azure Cosmos DB overview and case studies](#)
- [Azure Cosmos DB documentation](#)
- [Blockchain on Azure](#)
- [Video – Azure Blockchain Basics](#)
- [Azure Blockchain Workbench](#) – the quickest way to get started with blockchain on Azure
- [Block Talk](#) – A video series on blockchain on Azure