Optimizing performance and ROI with Azure Cosmos DB

Best practices for data modeling, partitioning, data distribution and cost management

AzureCosmosDB.com
Speed, performance, and cost management with Azure Cosmos DB

Azure Cosmos DB is a fully managed NoSQL database for modern app development. It offers SLA-backed single-digit millisecond speed and 99.999% availability, automatic and instant scalability, and open source APIs for MongoDB, Cassandra, and other NoSQL engines.

You can optimize your database performance and manage costs through proper data modeling, partitioning, and data distribution, and understanding how database operations and storage are billed. This guide will walk you through the key concepts, tips, and tricks you need to know to get the most out of Azure Cosmos DB.

Data modeling

You can significantly lower latency for common queries by optimizing your data model. We recommend spending time thinking about your workload: How will data be queried? What is the anticipated read/write ratio? Answering these questions before designing your data model will help you speed up and lower costs of your most frequent queries.

To learn about data modeling in Azure Cosmos DB:

1. Watch: data modeling for low-latency
2. Read documentation on data modeling

Partitioning

For horizontal scalability, Azure Cosmos DB stores your data in smaller partitions. Developers specify a partition key for each Azure Cosmos DB container at the time of creation. This critical design decision defines how your data will be stored in Azure Cosmos DB, as modifications are unavailable after initial creation, and non-optimized partition keys can lead to rate errors/throttling, misallocation of throughput, and unnecessary costs.

To effectively manage partition keys:

1. Watch: choosing the right partition key
2. Watch: avoiding “hot” partitions
3. Read best practices for choosing a partition key
4. Read this real world example of how to model and partition data

If you need to change your partition key later, create a new container with your desired key and transfer data using Azure Data Factory. You can also use change feed to do a live migration into another partitioned collection.

Data distribution and multi-master

Azure Cosmos DB can replicate and distribute your data to any Azure region (geo-redundancy), and also supports multi-region writes (multi-master). Distributing data and enabling multi-master brings data physically closer to your users and delivers enhanced SLAs including 99.999% availability and single-digit...
millisecond latency.

To manage data distribution settings, performance, and costs:

1. Watch: one-click data distribution to any Azure region
2. Check your account settings to see selected regions and monitor activity (you can add or remove regions at any time) and if multi-region writes (“multi-master”) is on or off
   a. Note: It is possible to have an Azure Cosmos DB account with geo-redundancy across multiple Azure regions without enabling multi-region writes (multi-master)
3. Watch: overview of multi-region writes (multi-master)
4. Understand how to optimize costs for multi-region workloads

Performance & cost optimization
Your account will be billed each month for database operations and storage across every Azure region your database is replicated and distributed across. There are many things you can do to manage costs such as sharing throughput across the database, setting custom index policies, and monitoring usage in the Azure portal.

To manage costs:

1. Watch: overview of Azure Cosmos DB pricing model with optimization and cost-saving tips
2. Watch: choose the right partition key to lower cost and improve query speed
3. Use the Azure Cosmos DB capacity calculator to estimate and plan throughput needs
4. Discover ways to plan and manage costs, and scale throughput with Azure Functions timer triggers
5. Consider reserved capacity 1- or 3-year term to save up to 65% on provisioned throughput

Get started
Now that you know what to look for, visit the Azure portal and update your location, geo-redundancy, multi-region write, and other account settings to ensure they’re set the way you want.

For more updates, tips, and documentation, visit our website and blog, or find us on Twitter and YouTube.