February 23, 2018

Azure Cosmos DB and GDPR

Contents

[Disclaimer 3](#_Toc513308546)

[Abstract 4](#_Toc513308547)

[Introduction: The GDPR and its implications 5](#_Toc513308548)

[Microsoft is committed to helping customers successfully master GDPR compliance requirements 5](#_Toc513308549)

[Getting started with GDPR 7](#_Toc513308550)

[Discover 9](#_Toc513308551)

[Search for and identify personal data 9](#_Toc513308552)

[How does Azure Cosmos DB help you in the discovery phase? 9](#_Toc513308553)

[Facilitate data classification 10](#_Toc513308554)

[Manage 11](#_Toc513308555)

[Enable data governance practices and processes 11](#_Toc513308556)

[Network security – IP Firewall 11](#_Toc513308557)

[Virtual Network (VNet) service endpoints 12](#_Toc513308558)

[Authentication & Authorization 12](#_Toc513308559)

[Geofencing 14](#_Toc513308560)

[Encryption at Rest 14](#_Toc513308561)

[Protected facilities 15](#_Toc513308562)

[Division of responsibility 15](#_Toc513308563)

[Security and data protection certifications 16](#_Toc513308564)

[Provide detailed notice of processing activities to data subjects 16](#_Toc513308565)

[Discontinue processing on the request 16](#_Toc513308566)

[Collect unambiguous, granular consent from data subjects 16](#_Toc513308567)

[Facilitate requests for the rectification, erasure, or transfer of personal data 17](#_Toc513308568)

[Rectify inaccurate or incomplete personal data regarding data subjects 17](#_Toc513308569)

[Erase personal data regarding a data subject 17](#_Toc513308570)

[Backup and Restore 18](#_Toc513308571)

[TTL 18](#_Toc513308572)

[Provide data subject with their personal data in a common, structured format 19](#_Toc513308573)

[Restrict the processing of personal data 19](#_Toc513308574)

[Protect 20](#_Toc513308575)

[Data protection and privacy by design and default 20](#_Toc513308576)

[Secure personal data through encryption 20](#_Toc513308577)

[Transport Layer Security 20](#_Toc513308578)

[Secure personal data by leveraging security controls that ensure the confidentiality, integrity, and availability of personal data. 21](#_Toc513308579)

[Detect and respond to data breaches. 21](#_Toc513308580)

[Logging 21](#_Toc513308581)

[What is logged? 21](#_Toc513308582)

[Managing your logs 23](#_Toc513308583)

[Activity log Alerts 23](#_Toc513308584)

[Diagnostic Log and Alert using EventHub 25](#_Toc513308585)

[Facilitate regular testing of security measures 25](#_Toc513308586)

[Report 27](#_Toc513308587)

[Maintain audit trails to show GDPR compliance 27](#_Toc513308588)

[Track and record flows of personal data into and out of the EU 27](#_Toc513308589)

[Policy-based geo-fencing 27](#_Toc513308590)

[Track and record flows of personal data to third-party service providers 28](#_Toc513308591)

[Facilitate Data Protection Impact assessments 28](#_Toc513308592)

[Certifications & Attestations 29](#_Toc513308593)

[Conclusion 30](#_Toc513308594)

[Further Reading 30](#_Toc513308595)

# Disclaimer

This white paper is a commentary on the General Data Protection Regulation (GDPR), as Microsoft interprets it, as of the date of publication. We’ve spent a lot of time with GDPR and like to think we’ve been thoughtful about its intent and meaning. But the application of GDPR is highly fact-specific, and not all aspects and interpretations of GDPR are well-settled. As a result, this white paper is provided for informational purposes only and should not be relied upon as legal advice or to determine how GDPR might apply to you and your organization. We encourage you to work with a legally qualified professional to discuss GDPR, how it applies specifically to your organization, and how best to ensure compliance.

MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, AS TO THE INFORMATION IN THIS WHITE PAPER. This white paper is provided “as-is”. Information and views expressed in this white paper, including URL and other Internet website references, may change without notice.

This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this white paper for your internal, reference purposes only.

Published January 2018

Version 1.0

© 2018 Microsoft. All rights reserved.

# Abstract

In this age of digital transformation, protecting privacy and enhancing security are top of mind. The European Union (EU) General Data Protection Regulation (GDPR) sets a new bar for privacy rights, security, and compliance. The GDPR mandates many requirements and obligations on organizations across the globe. Complying with this regulation necessitates significant investments in data handling and data protection for many organizations.

Microsoft Azure Cosmos DB customers who are subject to GDPR, whether they’re managing cloud-based or on-premises databases or both, need to ensure that personal data in their database systems is handled and protected according to GDPR principles. Many customers will need to review or modify their database management and data handling procedures, especially focusing on implementing or strengthening the security of data processing as required by GDPR.

Microsoft has considerable experience in successfully addressing data privacy principles and complying with complex regulations. Microsoft is helping customers to successfully address the privacy requirements of the GDPR. With a comprehensive set of compliance and security offerings and a vast partner ecosystem, Microsoft is prepared to support customers’ privacy, compliance, and security initiatives around GDPR and other important regulations.

# Introduction: The GDPR and its implications

On May 25, 2018, a new [European privacy regulation](http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32016R0679) is due to take effect. It sets a new global bar for privacy rights. The General Data Protection Regulation, or GDPR, is fundamentally about enabling and protecting the privacy rights of individuals. GDPR establishes strict global data privacy requirements that govern how personal data is managed and protected, while respecting individual’s choices.

To achieve its objectives, GDPR features several specific requirements related to the rights of individuals, such as the right to access their personal data, correct inaccuracies, erase data, object to data processing, and obtain a copy of their data. GDPR also seeks to ensure that personal data is protected no matter where it is sent, processed, or stored. As a result, data protection and security are key components in addressing the principles of GDPR.

Personal data in scope of GDPR can include but is not limited to:

* Name
* Identification number
* Email address
* Online user identifier
* Social media posts
* Physical, physiological, or genetic information
* Medical information
* Location
* Bank details
* IP address
* Cookies
* Cultural identity

GDPR introduces several obligations related to controls and security around the handling of personal data. The regulation obligates the data controller (the entity that determines the purposes and means of the processing of personal data) as well as the data processor (the entity that processes personal data on behalf of the controller) to “implement appropriate technical and organizational measures” to address these obligations.

Note: In terms of GDPR, an organization that is using Microsoft Cloud Services such as Azure Cosmos DB is a controller, while Microsoft acts as a processor.

## Microsoft is committed to helping customers successfully master GDPR compliance requirements

Microsoft has outlined its commitment to the GDPR and its support for customers in the compliance process within these blog posts:

* [Get GDPR compliant with the Microsoft Cloud](https://blogs.microsoft.com/on-the-issues/2017/02/15/get-gdpr-compliant-with-the-microsoft-cloud/) by Brendon Lynch, Microsoft Chief Privacy Officer
* [Earning your trust with contractual commitments to the General Data Protection Regulation](https://blogs.microsoft.com/on-the-issues/2017/04/17/earning-trust-contractual-commitments-general-data-protection-regulation/) by Rich Sauer, Microsoft Corporate Vice President & Deputy General Counsel

Moreover, Microsoft provides customers with a contractual commitment regarding GDPR obligations in the Microsoft [Online Services Terms](https://aka.ms/online-services-terms). For additional information about the GDPR, Microsoft’s commitments, and Microsoft’s recommendations for beginning the compliance journey, visit the [GDPR section of the Microsoft Trust Center](https://www.microsoft.com/en-us/TrustCenter/Privacy/gdpr/default.aspx).

# Getting started with GDPR

GDPR compliance can be broken into four key steps:



The following table lists the recommended tasks associated with each step and provides links to supporting documentation for each recommended task.

|  |  |  |  |
| --- | --- | --- | --- |
| Discover | Manage | Protect | Report |
| 1. [Search for and identify personal data](#_Search_for_and)  2. [Facilitate data classification](#_Facilitate_data_classification) | 1[. Enable data governance practices and processes](#_Enable_data_governance)  2. [Provide detailed notice of processing activities to data subjects](#_Provide_detailed_notice)  3. [Discontinue processing on the request](#_Discontinue_processing_on)  4. [Collect unambiguous, granular consent from data subjects](#_Collect_unambiguous,_granular)  5. [Facilitate requests for the rectification, erasure, or transfer of personal data](#_Facilitate_requests_for)  6. [Rectify inaccurate or incomplete personal data regarding data subjects](#_Rectify_inaccurate_or)  7. [Erase personal data regarding a data subject](#_Erase_personal_data)  8[. Provide data subject with their personal data in a common, structured format](#_Provide_data_subject)  9[. Restrict the processing of personal data](#_Restrict_the_processing) | 1. [Data protection and privacy by design and default](#_Data_protection_and)  2. [Secure personal data through encryption](#_Secure_personal_data)  3. [Secure personal data by leveraging security controls that ensure the confidentiality, integrity, and availability of personal data](#_Secure_personal_data_1)  4. [Detect and respond to data breaches](#_Detect_and_respond)  5. [Facilitate regular testing of security measures](#_Facilitate_regular_testing) | 1. [Maintain audit trails to show GDPR compliance](#_Maintain_audit_trails)  2. [Track and record flows of personal data into and out of the EU](#_Track_and_record)  3. [Track and record flows of personal data to third-party service providers](#_Track_and_record_1)  4. [Facilitate Data Protection Impact assessments](#_Facilitate_Data_Protection) |

# Discover

## Search for and identify personal data

***A critical first step to addressing GDPR requirements is to identify all personal data managed by your organization, so that you can adequately protect it and respond to data subject requests, such as erasure, rectification, and data portability. Microsoft business products and services offer a number of ways to identify personal information***

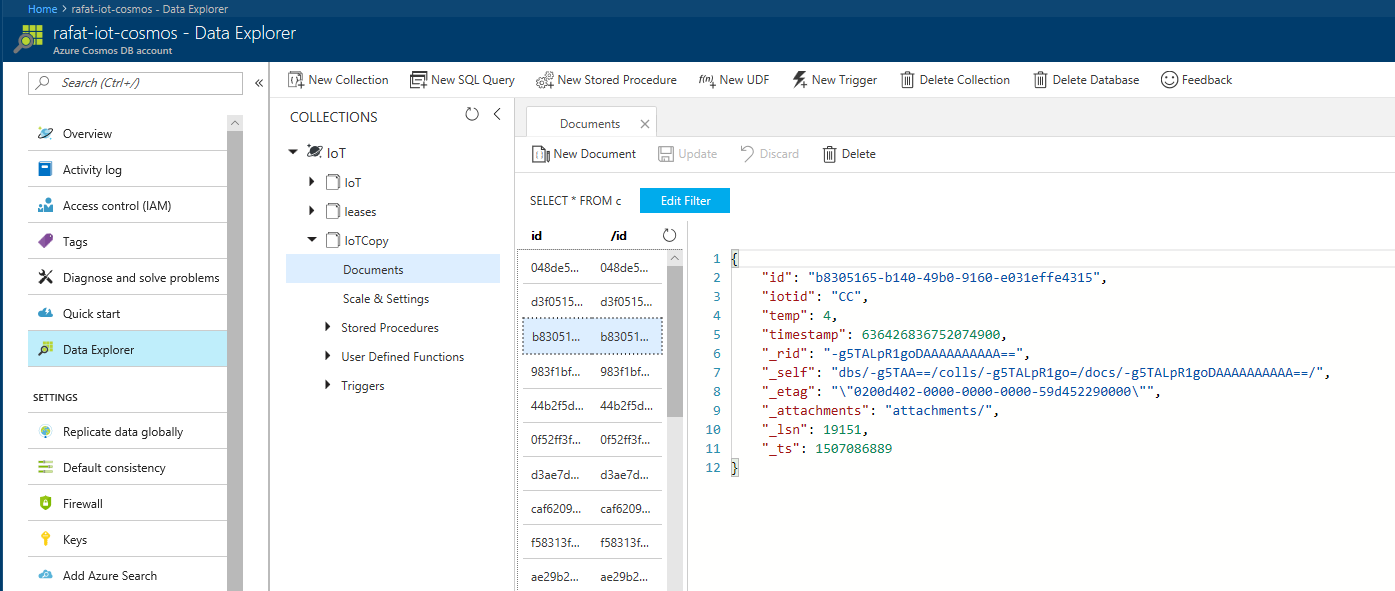
Azure Cosmos DB provides mechanisms to help identify personal data. Azure Cosmos DB is a schema-less database; thus, it is the application designer's responsibility to define the schema with obvious attributions (such as **Birthdate**, **Name**, and **address**) and determine whether those values hold sensitive data.

### How does Azure Cosmos DB help you in the discovery phase?

In Azure Cosmos DB, you have multiple options to explore the data. By default, all Azure Cosmos DB data is indexed and therefore queryable. You can use one of the following options to query data in Azure Cosmos DB:

1. Query for data directly in the Azure portal
2. Use one of the Azure Cosmos DB client APIs to query the data

In the Azure portal, you can quickly explore the data by pointing and clicking. Data is enriched by the last modified time to add context to personally identifiable information. You can also perform queries in the Azure portal (by replacing SELECT \* FROM c in the screenshot below) to query data.



Ensure that you review attachments if they are being used, as they do not directly display in the Azure portal UI.

For client APIs, you can use [SQL API](https://docs.microsoft.com/en-us/azure/cosmos-db/sql-api-index), [MongoDB API](https://docs.microsoft.com/en-us/azure/cosmos-db/mongodb-index), [Table API](https://docs.microsoft.com/en-us/azure/cosmos-db/table-powershell), [Graph API](https://docs.microsoft.com/en-us/azure/cosmos-db/gremlin-support), and [Cassandra API](https://docs.microsoft.com/en-us/azure/cosmos-db/cassandra-introduction) to investigate and query for personal data.

## Facilitate data classification

Organizations must have in place a mature data classification process and effective supporting technology that will enable them to comply with data subject requests and meet other GDPR requirements. Microsoft business products and services offer a number of ways to classify personal information.

In Azure Cosmos DB, a document can be tagged with extra properties as needed to help classify personal data. This is one of the benefits of schema agnostic data store; documents can have extra properties added and removed as needed.

There are many ways personal data can be tagged. Here are few examples:

Example 1: Have a GDPR section in the document:

{   
   **"**GRPR":{   
      "FirstName":"a",  
      "LastName":"b"  
   }  
}

Example 2: Have a property named GDPR that contains other properties about GDPR:

{

“FirstName”:”a”,

“LastName”:”b”,

“GDPR”: “FirstName, LastName”

}

Example 3: Add GDPR to the property name:

{

“GDPR\_FirstName”:”a”

“GDPR\_LastName”:”b”

}

# Manage

## Enable data governance practices and processes

To satisfy GDPR requirements, organizations must understand what types of personal data they process, how, and for what purposes. To help ensure that data handling processes comply with the GDPR, organizations must implement a data governance program.

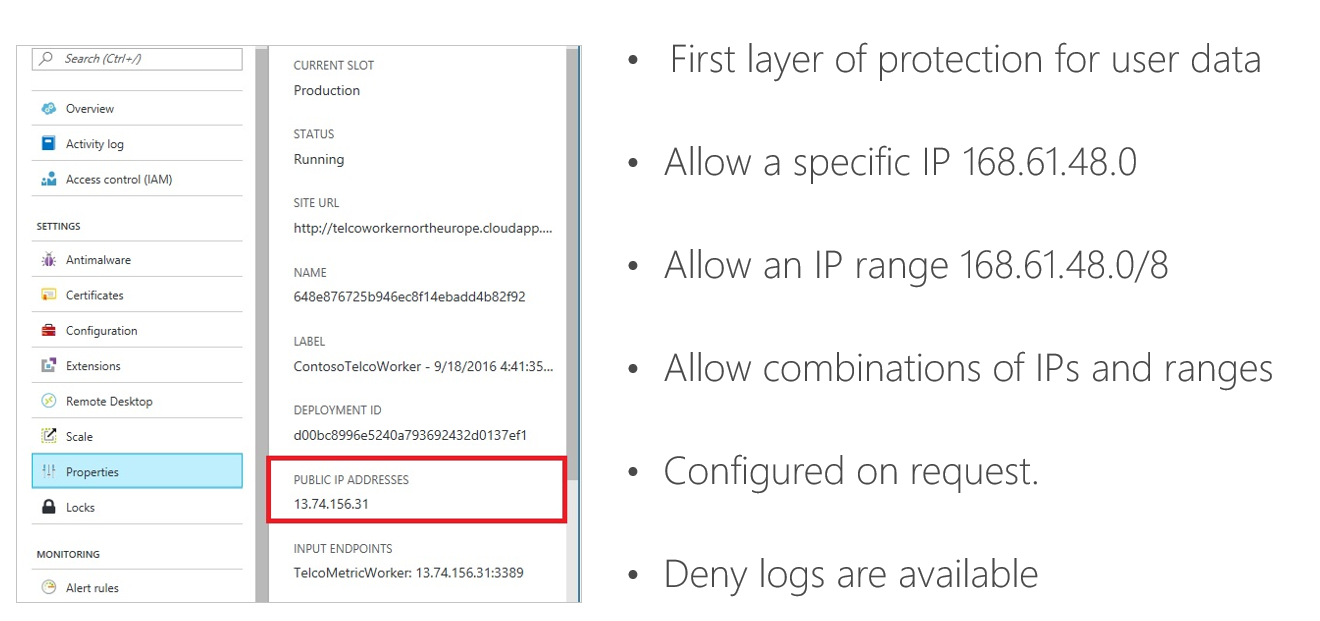
Azure Cosmos DB offers several methods to secure a database and the data stored therein, which are described in this section.

### Network security – IP Firewall

Using an IP firewall is the first layer of protection to help secure your database. Azure Cosmos DB supports policy-driven, IP-based access controls for inbound firewall support. The IP-based access controls are similar to the firewall rules used by traditional database systems, but they’re expanded so that an Azure Cosmos DB database account is accessible from only an approved set of machines or cloud services.

With Azure Cosmos DB, you can enable a specific IP address (168.61.48.0), an IP range (168.61.48.0/8), and combinations of IPs and ranges.

Azure Cosmos DB blocks all user requests that originate from machines outside this allowed list. Requests from approved machines and cloud services then must complete the authentication process to be given access control to the resources. However, this excludes the management back-plane addresses for Azure system and Cosmos DB internal gateway calls. The firewall check is performed only on the request path.



Learn more about [Azure Cosmos DB firewall support](https://docs.microsoft.com/en-us/azure/cosmos-db/firewall-support).

### Virtual Network (VNet) service endpoints

Azure CosmosDB accounts can be configured to allow access only from specific Azure Virtual Network’s subnet. By enabling a [Service Endpoint](https://review.docs.microsoft.com/en-us/azure/virtual-network/virtual-network-service-endpoints-overview) for Azure CosmosDB from a Virtual Network and its subnet, traffic is ensured an optimal and secure route to the Azure Cosmos DB.

Once an Azure Cosmos DB account is configured with a virtual network service endpoint, it can be accessed only from the specified subnet, all public/internet access is removed. To learn in detailed about service endpoints, refer to the Azure [Virtual network service endpoints overview](https://review.docs.microsoft.com/en-us/azure/virtual-network/virtual-network-service-endpoints-overview) article.

[Learn](https://aka.ms/cosmosdb-vnet-blog) how to configure VNet for Azure Cosmos DB.

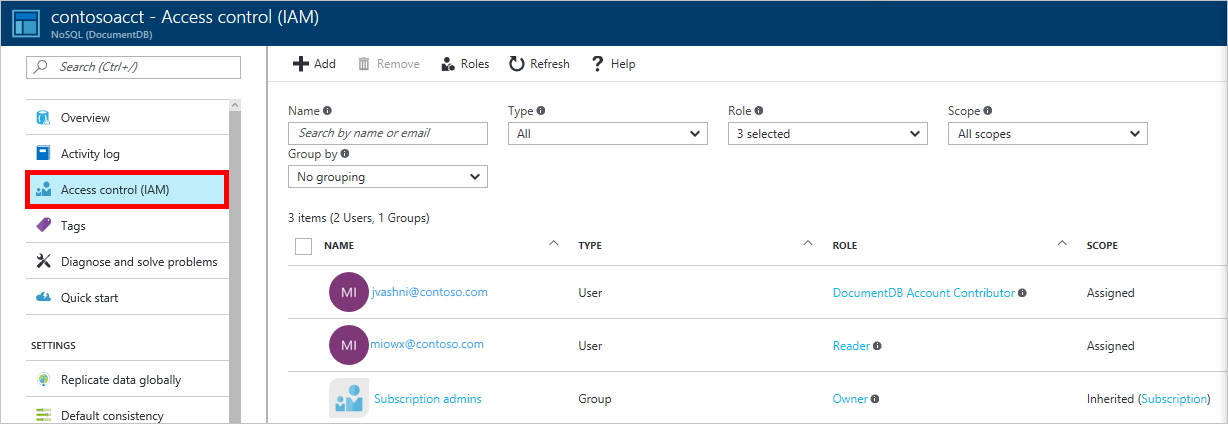
### Authentication & Authorization

Azure Cosmos DB uses a hash-based message authentication code (HMAC) for authorization.   
Each request is hashed using the secret account key, and the subsequent base-64 encoded hash is sent with each call to Azure Cosmos DB. To validate the request, the Azure Cosmos DB service uses the correct secret key and properties to generate a hash, then it compares the value with the one in the request. If the two values match, the operation is authorized successfully, and the request is processed. Otherwise there is an authorization failure, and the request is rejected.  
You can use either a [master key](https://docs.microsoft.com/en-us/azure/cosmos-db/secure-access-to-data#master-keys) or a [resource token](https://docs.microsoft.com/en-us/azure/cosmos-db/secure-access-to-data#resource-tokens) allowing fine-grained access to a resource such as a document.  
You can also provide access to the database account using Access control (IAM) in the Azure portal. IAM provides role-based access control and integrates with Active Directory. You can use built-in roles or custom roles for individuals and groups as shown in the following image.

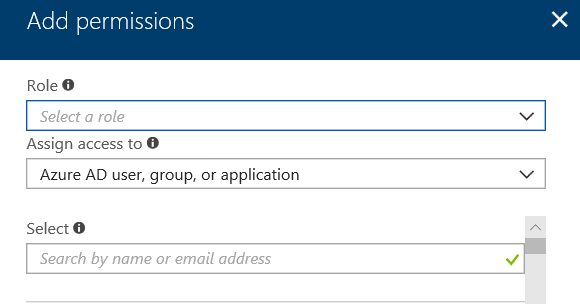
All access to the Azure Cosmos DB portal is authenticated and authorized by Azure and Azure Active Directory (AAD). Any user who needs to access the Azure Cosmos DB portal must have explicit permission to access the resources. The user must belong to one of the pre-defined roles:

* Owner
* Contributor
* Reader
* Cosmos DB Account Reader Role
* DocumentDB Account Contributor
* etc.

These roles are set in the **Access control (IAM)** pane in the Azure Cosmos DB portal, as shown in the following image. New roles will be added in future.



New permissions are set by clicking **+ Add** in the **Access control (IAM)** pane and selecting the role and assignee.



Other than portal access control, Azure Cosmos DB uses two types of keys to authenticate users and provide access to its data and resources.

| Key type | Resources |
| --- | --- |
| [Master keys](https://docs.microsoft.com/en-us/azure/cosmos-db/secure-access-to-data#master-keys) | Used for administrative resources: database accounts, databases, users, and permissions |
| [Resource tokens](https://docs.microsoft.com/en-us/azure/cosmos-db/secure-access-to-data#resource-tokens) | Used for application resources: collections, documents, attachments, stored procedures, triggers, and UDFs |

#### Master keys

Master keys provide access to the all the administrative resources for the database account. Master keys:

* Provide access to accounts, databases, users, and permissions.
* Cannot be used to provide granular access to collections and documents.
* Are created during the creation of an account.
* Can be regenerated at any time.

Each account consists of two Master keys: a primary key and secondary key. The purpose of dual keys is so that you can regenerate, or roll keys, providing continuous access to your account and data.

In addition to the two master keys for the Azure Cosmos DB account, there are two read-only keys. These read-only keys only allow read operations on the account. Read-only keys do not provide access to read permissions resources but only to the data.

#### Resource tokens

Azure Cosmos DB resource tokens provide a safe alternative that enables clients to read, write, and delete resources in your Azure Cosmos DB account according to the permissions you've granted, and without the need for either a master or read-only key.

You can use a resource token (by creating Azure Cosmos DB users and permissions) when you want to provide access to resources in your Azure Cosmos DB account to a client that cannot be trusted with the master key.

Learn more about [Securing access to Azure Cosmos DB data](https://docs.microsoft.com/en-us/azure/cosmos-db/secure-access-to-data).

### Geofencing

Azure Cosmos DB is designed to have policy-based geo-fencing capabilities. Geo-fencing is an important component to ensure data governance and compliance restrictions and may prevent associating a specific region with your account. Examples of geo-fencing include (but are not restricted to), scoping global distribution to the regions within a sovereign cloud (for example, China, and Germany), or within a government taxation boundary (for example, Australia). The policies are controlled using the metadata of your Azure subscription.

Learn more about [How to distribute data globally with Azure Cosmos DB](https://docs.microsoft.com/en-us/azure/cosmos-db/distribute-data-globally).

### Encryption at Rest

All data stored in Azure Cosmos DB is encrypted at rest and in transport, you do not have to take any action. Another way to phrase this is that encryption at rest and on flight is "ON" by default. There are no controls to turn it off or on. Azure Cosmos DB provides this feature while continuing to meet its availability and performance SLAs.

Learn more about [Azure Cosmos DB database encryption at rest](https://docs.microsoft.com/en-us/azure/cosmos-db/database-encryption-at-rest).

### Protected facilities

Data in Azure Cosmos DB is stored in protected Azure datacenters. Microsoft’s state-of-the-art datacenters safeguard your data in facilities that are protected by industry-leading physical security and compliant with a comprehensive portfolio of standards and regulations.

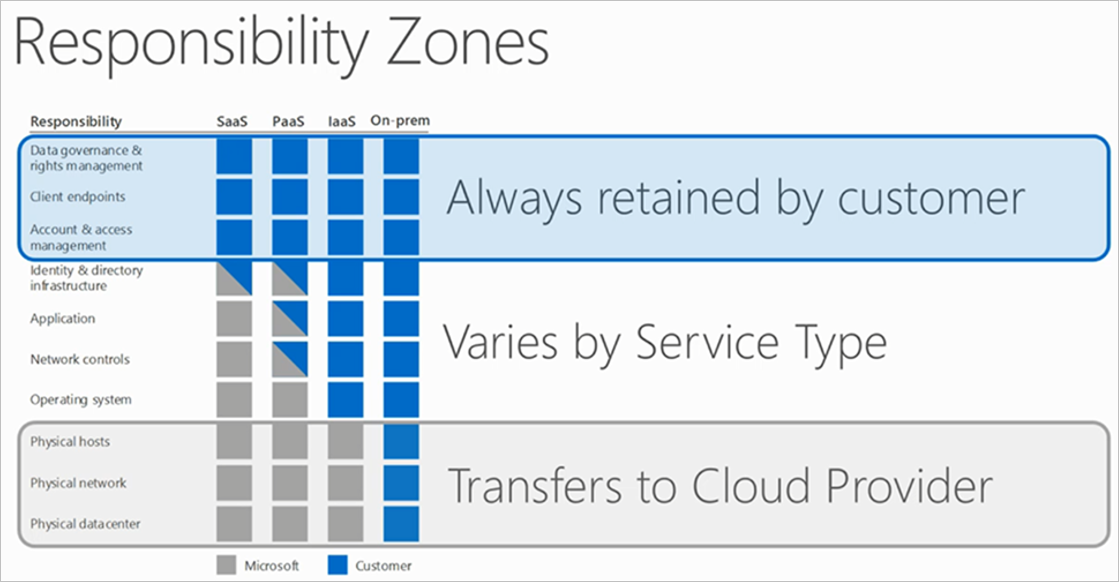
[Watch a video about Microsoft datacenter security](https://www.youtube.com/watch?time_continue=6&v=r1cyTL8JqRg).

[Find out more about the Microsoft holistic security approach.](https://www.microsoft.com/en-US/security/default.aspx)

Learn more about [Global datacenters](https://www.microsoft.com/en-us/cloud-platform/global-datacenters).

### Division of responsibility

It’s important to understand the division of responsibility between you (the customer) and Microsoft. On-premises, you are responsible for the whole stack, but as you move to the cloud some responsibilities transfer to Microsoft. The following responsibility matrix shows what areas of the stack in a SaaS, PaaS, and IaaS deployment that you are responsible for and what areas Microsoft is responsible for.



For all cloud deployment types, you own your data and identities. You are responsible for protecting the security of your data and identities, on-premises resources, and the cloud components you control (which vary by service type).

Responsibilities that are always retained by you, regardless of the type of deployment, are:

* Data
* Endpoints
* Account
* Access management

Learn more about security in [Securing PaaS deployments](https://docs.microsoft.com/en-us/azure/security/security-paas-deployments) and [Global datacenters](https://www.microsoft.com/en-us/cloud-platform/global-datacenters).

### Security and data protection certifications

Azure Cosmos DB has achieved several information security and data privacy certifications and attestations, including:

ISO 27001, ISO 27018, EUMC, HIPAA, PCI

SOC1 and SOC2 (Audit complete, Certification in Q2 2017)

FedRAMP, IRS 1075, UK Official (IL2) (Q2 2017)

HITRUST (H2 2017)

Learn more in [Overview of Microsoft Azure compliance](https://aka.ms/AzureCompliance).

## Provide detailed notice of processing activities to data subjects

The GDPR requires transparency from organizations regarding the collection, use, and processing of personal data. Notice to data subjects should be easily accessible and easy to understand and use clear and plain language.

It is the organization’s responsibility to provide notice of processing activities to data subjects as this functionality is generally built into the organization’s public-facing services.

## Discontinue processing on the request

Under the GDPR, organizations must enable data subject rights such as the right to object to the processing of their personal data.

Azure Cosmos DB provides you with the design options to discontinue processing the data subjects. For example, any record in a collection can be appended with a custom property and a “Discontinued” data tag to support application logic to prevent the processing of the associated personal data. Appropriately tagging content and implementing the logic to stop processing that data is the responsibility of the organization.

## Collect unambiguous, granular consent from data subjects

Before processing personal data, organizations must have a legal basis for processing, such as consent. Such consent must be freely given, specific, informed, and unambiguous. Consent can be withdrawn at any time.

Azure Cosmos DB can be used as a storage and processing technology to help manage the process of documenting consent for processing activities. Using other supporting technologies, Microsoft customers can dynamically populate a database that documents the affirmative consent of a data subject obtained by a customer application.

While Azure Cosmos DB can provide a platform capable of hosting external-facing privacy notices, it is the responsibility of the customer to ensure that the specific language of the notice meets their obligations under the GDPR.

## Facilitate requests for the rectification, erasure, or transfer of personal data

The GDPR requires that an organization processing personal data must give data subjects a means to submit requests to rectify, erase, or transfer their personal data. Controllers must also give the requesting data subject information on action taken within 30 days of receipt of the request absent an extension.

Azure Cosmos DB can be used as a storage and processing technology to document the requests of data subjects and the actions taken against requests. Requests to rectify, erase, or transfer personal data, and meeting the 30 day follow up requirement are the responsibility of the customer.

## Rectify inaccurate or incomplete personal data regarding data subjects

The GDPR requires organizations that process personal data to enable data subjects to request rectification of “inaccurate personal data” and the completion of “incomplete personal data.

Azure Cosmos DB enables you to search, select, modify and delete any personal data located in collections using any of the provided API’s ([SQL API](https://docs.microsoft.com/en-us/azure/cosmos-db/sql-api-index), [MongoDB API](https://docs.microsoft.com/en-us/azure/cosmos-db/mongodb-index), [Table API](https://docs.microsoft.com/en-us/azure/cosmos-db/table-powershell), [Graph API](https://docs.microsoft.com/en-us/azure/cosmos-db/gremlin-support), and [Cassandra API](https://docs.microsoft.com/en-us/azure/cosmos-db/cassandra-introduction) ). Finding and updating inaccurate or incomplete personal data is the responsibility of the customer.

## Erase personal data regarding a data subject

Under the GDPR, all data subjects have the right to request the erasure of their personal data by organizations. The target data must be identified and erased without undue delay except in certain circumstances.

Azure Cosmos DB enables you to search for and erase personal data located in collections using general queries against character-based data in collections. You can also use any other Delete APIs to erase personal data.

You can also enable the time to live (TTL) feature to delete data automatically after a specified period, without incurring any additional cost. See the [TTL](#_TTL) section in this document for additional information. However, please be aware that today TTL does not provide the audit log of deleted records.

Please note, that TTL works on the modified time, so if you are using TTL to delete the record then make sure it is not be modified again, else TTL will start the new count down.

Once data is deleted from the system, it is not accessible by the application.

Azure Cosmos DB provides methods to help erase data; however, erasing personal data is the responsibility of the customer.

Learn more about ChangeFeed [Working with the change feed support in Azure Cosmos DB](https://docs.microsoft.com/azure/cosmos-db/change-feed).

### Backup and Restore

Azure Cosmos DB automatically takes backups of all data, except attachments, at regular intervals. The automatic backups don’t affect the performance or availability of your database operations. All your backups are stored separately in another storage service, and those backups are globally replicated for resiliency against regional disasters.

Please note if your data is in sovereign cloud then your backup will also stay in sovereign cloud. All backups are encrypted.

The automatic backups are intended for scenarios when you accidentally delete your Azure Cosmos DB container.

Note that attachments are not currently backed up, so this section does not apply to them.

With Azure Cosmos DB, the backups of your non-attachment data are made highly redundant and resilient to regional disasters. These automated backups are currently taken approximately every four hours, and the latest two backups are stored at all times. If the data is accidentally deleted or modified, please [contact Azure support](https://azure.microsoft.com/support/options/) within eight hours.

Automatic backups are stored in Azure Blob storage. To help guarantee low-latency and efficient upload, the snapshot of your backup is uploaded to an instance of Azure Blob storage in the same region as the current write region of your Azure Cosmos DB database account. For resiliency against a regional disaster, each snapshot of your backup data in Azure Blob storage is again replicated via geo-redundant storage to another region.

As described earlier, Azure Cosmos DB takes snapshots of your data every four hours at the partition level. At any time, only the last two snapshots are retained. However, if the collection/database is deleted, we retain the existing snapshots for all the deleted partitions within the collection/database for 30 days.

If you want to maintain your own snapshots, you can use the export-to-JSON option in the Azure Cosmos DB [Data Migration tool](https://docs.microsoft.com/en-us/azure/cosmos-db/import-data) to schedule additional backups.

In normal functioning, if any data is deleted from a container, it will be deleted from the backups after 8 hours automatically. However, if customers have requested a longer-term backup say for 30 days, then they must restore the data manually and delete the record. Today, at the time of writing this document, customers must open a support ticket to restore data.

Learn more about [Automatic online backup and restore with Azure Cosmos DB](https://docs.microsoft.com/en-us/azure/cosmos-db/online-backup-and-restore).

### TTL

With "time to live" or TTL, Azure Cosmos DB provides the ability to have documents automatically purged from the database after a specified period. You can set the default TTL at the container level, and you can override it on a per-document basis. After TTL is set, Azure Cosmos DB automatically removes documents that exist after that period (in seconds) since they were last modified.

Time to live in Azure Cosmos DB uses an offset against when the document was last modified. To do this, it uses the \_ts field that exists in every document. The \_ts field is a Unix-style epoch timestamp that represents the date and time. The \_ts field is updated every time a document is modified.

The following information applies to the container’s **DefaultTTL** property:

* If missing (or set to null), documents are not deleted automatically.
* If present and the value is "-1" = infinite, documents don’t expire by default.
* If present and the value is some number ("n"), documents expire “n” seconds after the last modification.

The following information applies for the documents’ TTL property:

* It’s applicable only if **DefaultTTL** is present for the parent collection.
* It overrides the **DefaultTTL** value for the parent collection.

Learn more about TTL in [Expire data in Azure Cosmos DB collections automatically with time to live](https://docs.microsoft.com/en-us/azure/cosmos-db/time-to-live).

## Provide data subject with their personal data in a common, structured format

Under the GDPR, data subjects have the right to request and receive their personal data from an organization in a structured, commonly used, and machine-readable format.

Azure Cosmos DB provides several APIs to export personal data in a common structured format. To identify the target personal data, you can use general queries against character-based data in collections to locate the target personal data. Client applications can obtain the data [using SELECT statements](https://docs.microsoft.com/en-us/azure/cosmos-db/sql-api-introduction) and then they can write it to any desired format.

## Restrict the processing of personal data

Data subjects may request a temporary restriction of processing activities utilizing their personal data in certain circumstances, for example if a data subject objects to the processing of their personal data, but the controller has a legal requirement to retain it. Organizations may need to employ technical means to prevent a specific data subject’s personal data from undergoing certain processing activities

Azure Cosmos DB does provide you with the design options to discontinue processing the data subjects. Any record in a collection can be appended with a custom property and tag data as “Discontinued” to support application logic to prevent the processing of the associated personal data, as described in [Discontinue processing on request](#_Discontinue_processing_on).

# Protect

## Data protection and privacy by design and default

The GDPR requires organizations that collect or process personal data to ensure that their activities and supporting technology are built to support data privacy and data protection principles. To support customers’ efforts to protect personal data, Microsoft services are developed using the Microsoft Security Development Lifecycle, which incorporates privacy-by-design and privacy-by-default methodologies. In addition, Microsoft business cloud services are managed in accordance with Microsoft privacy policies, which allow Microsoft personnel to access customer data only for purposes compatible with providing the contracted services, such as troubleshooting and improving features.

As described in the [Manage](#_Manage), Azure Cosmos DB offers several methods to protect a database and the data stored therein.

1. [IP Firewall](#_Network_security_–)
2. [Authentication and Authorization](#_Authentication_&_Authorization)
3. [User Permission](#_Users_and_permissions)
4. [Geofencing](#_Geofencing)

Additional tasks you can use to protect data are described in this section.

## Secure personal data through encryption

The GDPR requires organizations that process personal data to maintain a high standard of security to ensure the confidentiality, integrity, and availability of personal data. Though organizations should explore how to structure their processes and supporting technologies to meet their specific needs, the GDPR identifies encryption as one tool that may be appropriate given the risk to support this requirement.

### Transport Layer Security

Azure Cosmos DB always uses connections secured with Transport Layer Security (TLS). This ensures that data is encrypted in transit to and from the database, and it reduces susceptibility to “man-in-the-middle” attacks.

Azure Cosmos DB has TLS 1.2 support enabled, and this is the recommended protocol to use for highly secure communication. For Azure Cosmos DB, all connections require encryption at all times.

Encryption at rest is a phrase that commonly refers to the encryption of data on nonvolatile storage devices, such as solid-state drives (SSDs) and hard disk drives (HDDs). Azure Cosmos DB stores its primary databases on SSDs. Its media attachments and backups are stored in Azure Blob storage, which is backed up by HDDs. With the release of encryption at rest for Azure Cosmos DB, all your databases, media attachments, and backups are encrypted. Your data is now encrypted in transit (over the network) and at rest (nonvolatile storage), giving you end-to-end encryption.

As a PaaS service, Azure Cosmos DB is very easy to use. Because all user data stored in Azure Cosmos DB is encrypted at rest and in transport, you don't have to take any action. Another way to put this is that encryption at rest is "on" by default. There are no controls to turn it off or on. We provide this feature while we continue to meet our availability and performance SLAs.

Learn more about [Azure Cosmos DB database encryption at rest](https://docs.microsoft.com/en-us/azure/cosmos-db/database-encryption-at-rest).

## Secure personal data by leveraging security controls that ensure the confidentiality, integrity, and availability of personal data.

Critical to meeting GDPR requirements is the implementation of technical security measures that limit the risk of unauthorized access, use, or disclosure of personal data. Though the GDPR does not prescribe specific control requirements, organizations should select technologies that provide adequate protection.

Data in Azure Cosmos DB is protected by in transit (over the network) and at rest (nonvolatile storage), giving you end-to-end encryption as described in the [Transport layer security](#_Transport_Layer_Security) section.

Azure Cosmos DB offers several methods to secure a database and the data stored therein, as described in the [Enable data governance practices and processes](#_Enable_data_governance) section.

## Detect and respond to data breaches.

The GDPR requires organizations to maintain appropriate technologies to secure personal data and defend against data breaches. If a data breach does occur, the controller may be required to notify regulators and may also need to notify affected data subjects quickly.

Azure Cosmos DB provides a powerful set of built-in logging capabilities that safeguard data and help you identify potential breaches. Database logging enables customers to understand ongoing database activities and analyze and investigate potential threats or suspected abuse.

### Logging

You can monitor how and when your databases are accessed by using diagnostic logging in Azure Cosmos DB. By enabling diagnostic logging, you can send logs to [Azure Storage](https://azure.microsoft.com/services/storage/), stream them to [Azure Event Hubs](https://azure.microsoft.com/services/event-hubs/), and export them to [Azure Log Analytics](https://azure.microsoft.com/services/log-analytics/), which is part of [Microsoft Operations Management Suite](https://www.microsoft.com/cloud-platform/operations-management-suite). For a detailed description of how to enable diagnostic logging on your container, see [Azure Cosmos DB diagnostic logging](https://docs.microsoft.com/en-us/azure/cosmos-db/logging).

### What is logged?

* All authenticated REST Azure Cosmos DB (SQL) API requests, including failed requests because of access permissions, system errors, or bad requests. Support for MongoDB, Graph, Table, and Cassandra APIs is not currently available.
* Operations on the database itself, including CRUD operations on all documents, containers, and databases.
* Operations on account keys, such as creating, modifying, or deleting these keys.
* Unauthenticated requests that result in a 401 response. Examples include requests that do not have a bearer token, or are malformed or expired or have an invalid token.

The following table describes the content of each log entry.

|  |  |  |
| --- | --- | --- |
| **Azure Storage Field or Property** | **Log Analytics property** | **Description** |
| time | TimeGenerated | The date and time (UTC) when the operation occurred. |
| resourceId | Resource | The Azure Cosmos DB account for which logs are enabled. |
| category | Category | For Azure Cosmos DB logs, DataPlaneRequests is the only available value. |
| operationName | OperationName | Name of the operation. This value can be any of the following operations: Create, Update, Read, ReadFeed, Delete, Replace, Execute, SqlQuery, Query, JSQuery, Head, HeadFeed, or Upsert. |
| properties | n/a | The contents of this field are described in the following rows. |
| activityId | activityId\_g | The unique GUID for the logged operation. |
| userAgent | userAgent\_s | A string that specifies the client user agent performing the request. The format is {user agent name}/{version}. |
| resourceType | ResourceType | The type of the resource accessed. This value can be any of the following resource types: Database, Collection, Document, Attachment, User, Permission, StoredProcedure, Trigger, UserDefinedFunction, or Offer. |
| statusCode | statusCode\_s | The response status of the operation. |
| requestResourceId | ResourceId | The resourceId pertaining to the request, may point to databaseRid, collectionRid or documentRid depending on the operation performed. |
| clientIpAddress | clientIpAddress\_s | The client's IP address. |
| requestCharge | requestCharge\_s | The number of RUs used by the operation |
| collectionRid | collectionId\_s | The unique ID for the collection. |
| duration | duration\_s | The duration of operation, in ticks. |
| requestLength | requestLength\_s | The length of the request, in bytes. |
| responseLength | responseLength\_s | The length of the response, in bytes. |
| resourceTokenUserRid | resourceTokenUserRid\_s | This is non-empty when [resource tokens](https://docs.microsoft.com/azure/cosmos-db/secure-access-to-data#resource-tokens) are used for authentication and points to resource ID of the user. |

Learn more about [Azure Cosmos DB diagnostic logging](https://docs.microsoft.com/en-us/azure/cosmos-db/logging).

### Managing your logs

Logs are made available in your account two hours after the Azure Cosmos DB operation is made. It's up to you to manage your logs in your storage account. Specifically:

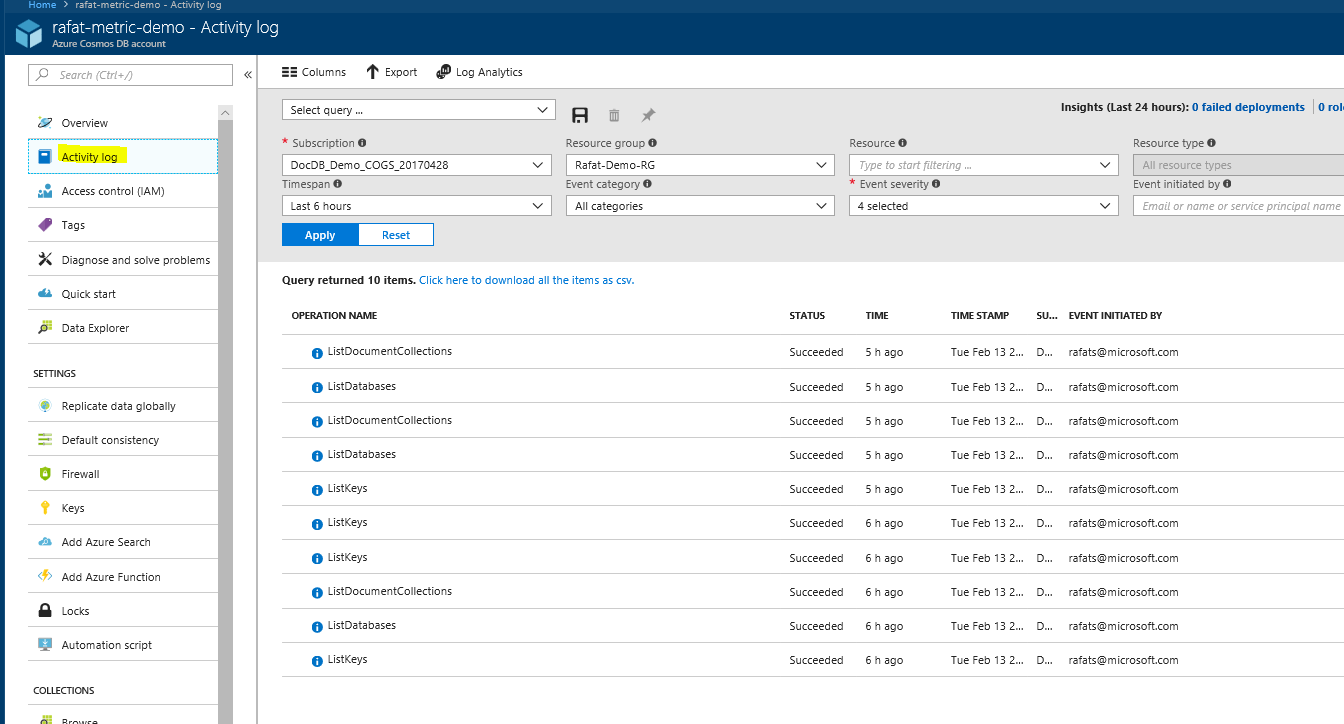
* Use standard Azure access control methods to secure your logs by restricting who can access them.
* Delete logs that you no longer want to keep in your storage account.

The retention period for data plane requests archived to a storage account is configured in the portal when **Log DataPlaneRequests** is selected. To change that setting, see [Turn on logging in the Azure portal](https://docs.microsoft.com/en-us/azure/cosmos-db/logging#turn-on-logging-in-the-azure-portal).

Learn more about the metrics and log categories that the various Azure services support in [Overview of metrics in Microsoft Azure](https://docs.microsoft.com/en-us/azure/monitoring-and-diagnostics/monitoring-overview-metrics) and [Collect and consume log data from your Azure resources](https://docs.microsoft.com/en-us/azure/monitoring-and-diagnostics/monitoring-overview-of-diagnostic-logs).

### Activity log Alerts

All control plane activities can be logged, and alerts can be created on the activities. The following image shows the **Activity Log** pane in the Azure portal.

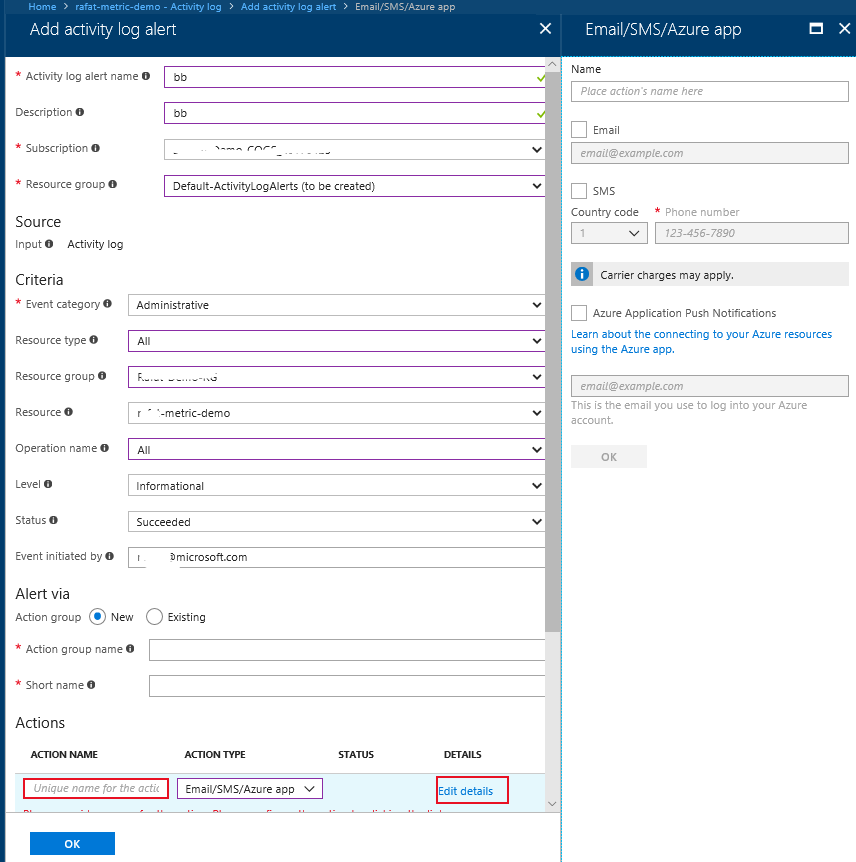


With Activity Logs you can view who is accessing your keys and make changes to the account.

Alerts can be created on following operations:

1. Delete Database account
2. Read Database account
3. Update database account
4. List Key
5. Read database account read-only key
6. Rotate key
7. Get Connection String
8. Change Resource group
9. Manual Failover

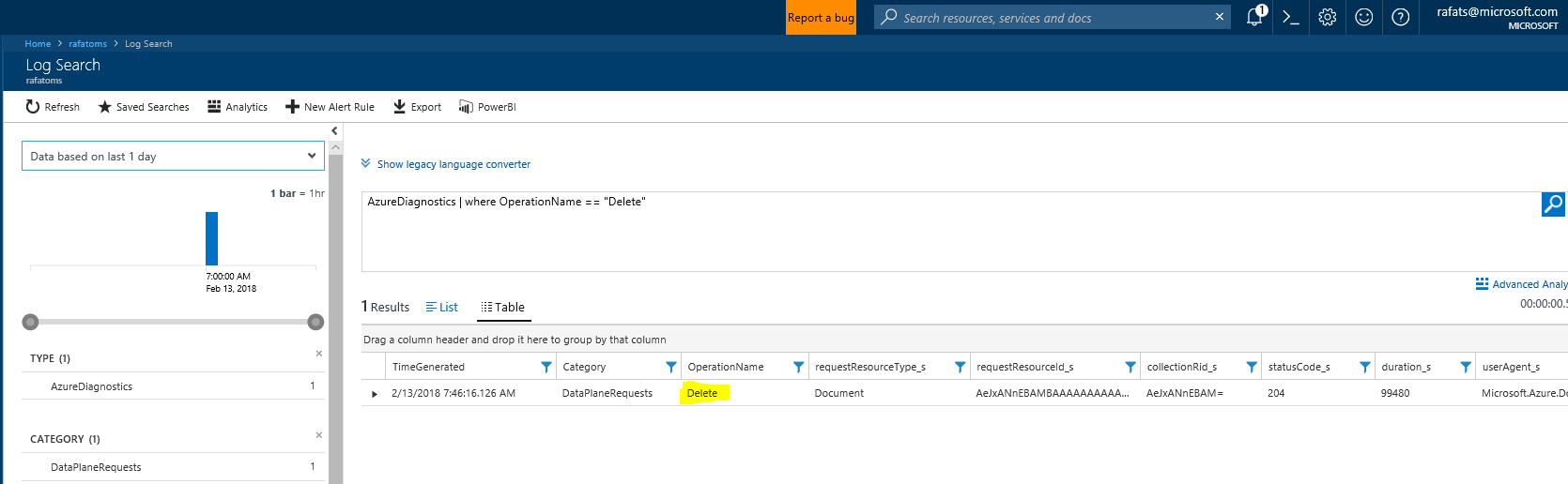
You can also create alerts when specific activities occur, as shown in the following image:



Using OMS, you can query your log for a specific action. For example, to look for a delete operation on a document, you can write the query as follows:

AzureDiagnostics | where OperationName == "Delete"

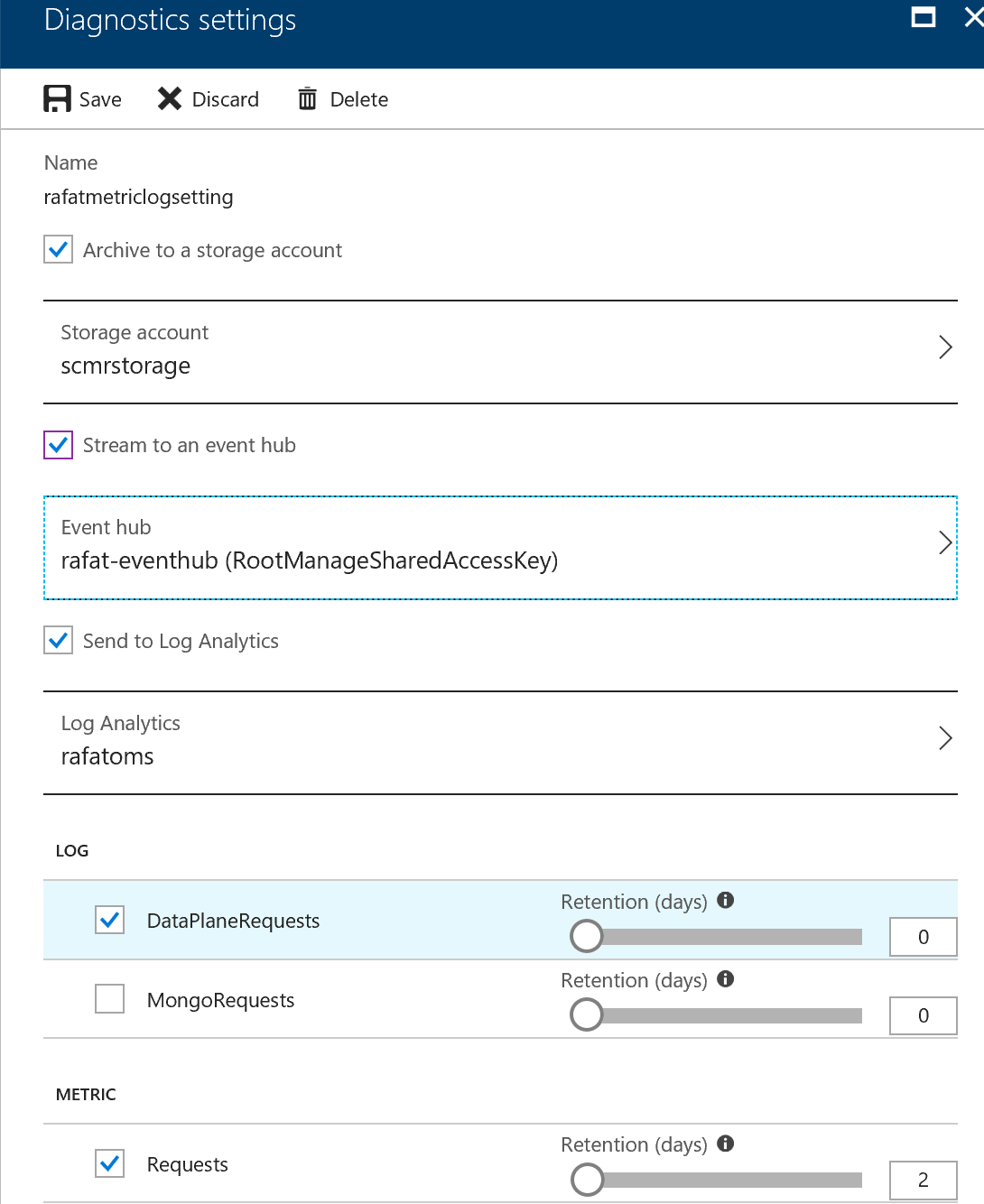
The following image shows the query being performed in the Azure portal and the result of the query:



The record on the table can be expanded for more information. Learn more about [How to interpret your logs](https://docs.microsoft.com/en-us/azure/cosmos-db/logging#interpret).

### Diagnostic Log and Alert using EventHub

Another alternate method is to let Diagnostic log flow in [EventHub](https://azure.microsoft.com/en-us/services/event-hubs/) and connects a [Logic App](https://azure.microsoft.com/en-us/services/logic-apps/) to EventHub to send instant alerts in many different formats, e.g. Mail, SMS, etc.



Above picture showing how can a Diagnostic log can be connected to EventHub.

## Facilitate regular testing of security measures

To ensure that the technical and organizational measures in place to protect personal data are adequate and appropriated for the risk, organizations should regularly test, assess, and evaluate their effectiveness.

Microsoft also conducts ongoing monitoring and testing of Azure security measures. These include ongoing threat modeling, code review and security testing, live site penetration testing, and centralized security logging and monitoring.

# Report

## Maintain audit trails to show GDPR compliance

Organizations must retain records of data subjects’ requests to demonstrate compliance with GDPR requirements. Records should contain both the nature of each request—for example, to view or rectify personal data—and their resolution.

Azure Cosmos DB provides a powerful set of built-in logging capabilities. Operations on the database itself, including CRUD operations on all documents, containers, and databases are logged.

Today, you can see many things using Azure Log Analytics, e.g. resource name, actions, collection name, call results, HTTP status codes, user agent, etc., all of which will help you maintain audit trails. You can see the IP address of the caller, but the last octal is hidden. For control pane operations, you can see the user information (email address). Data plane operations currently do not provide an email address but are on our roadmap. See the [Logging](#_Logging) section in this document to learn more.

Learn more about ChangeFeed in [Working with the change feed support in Azure Cosmos DB](https://docs.microsoft.com/en-us/azure/cosmos-db/change-feed).

## Track and record flows of personal data into and out of the EU

The GDPR reinforces controls on the transfer of personal data outside of the EU. In addition to a robust process to evaluate the appropriateness of cross-border transfers of personal data, organizations should consider technical solutions to maintain data within approved geographies, and to track and record flows of personal data into and out of the EU.

Azure Cosmos DB lets you reduce the need for transfer of personal data outside of the EU by enabling you to select a region or a national cloud during the initial setup of services and to store your data in any of more than 40+ regions around the globe. These choices include multiple regional choices within Europe as well as the German sovereign data storage region.

### Policy-based geo-fencing

Azure Cosmos DB is designed to have policy-based geo-fencing capabilities. Geo-fencing is an important component to ensure data governance and compliance restrictions and may prevent associating a specific region with your account. Examples of geo-fencing include (but are not restricted to), scoping global distribution to the regions within a sovereign cloud (for example, China, and Germany), or within a government taxation boundary (for example, Australia). The policies are controlled using the metadata of your Azure subscription.

Additionally, Microsoft has made several contractual commitments related to Azure that enable the appropriate flow of personal data within the Microsoft ecosystem. Microsoft has implemented EU Model Clauses and is certified to the EU-US Privacy Shield framework.

## Track and record flows of personal data to third-party service providers

The GDPR places responsibility for the safeguarding and appropriate processing of personal data on the organization. This responsibility continues even if that data is sent to third-party service providers for additional processing.

You can use Azure Cosmos DB as a storage and processing technology to document personal data that is transferred to additional processors. This may help customers trace personal data distributed to down-stream processors. Azure Cosmos DB customers acting as controllers are responsible for tracking the distribution of personal data to third parties by their custom services and applications.

## Facilitate Data Protection Impact assessments

Under the GDPR, controllers that process personal data may be required to conduct Data Protection Impact assessments (DPIAs). A DPIA requires the organization to identify and analyze the impact of a proposed processing activity on the protection of personal data.

Azure Cosmos DB provides detailed [logging](https://docs.microsoft.com/en-us/azure/cosmos-db/logging) on data and control pane operations, which can help customers seeking information that may be useful in performing a DPIA.

# Certifications & Attestations

To help you comply with national, regional, and industry-specific requirements governing the collection and use of individuals’ data, Microsoft offers the most comprehensive set of compliance offerings of any cloud service provider. Please see the detail list of all the compliance offered [here](https://www.microsoft.com/en-us/trustcenter/compliance/complianceofferings).

# Conclusion

Organizations will need to invest significantly to ensure the GDPR principles are effectively implemented and sustained in their environments. Microsoft itself is going through the substantial process of validating that all its data systems meet GDPR compliance readiness—and is approaching this task methodically and meticulously across the company.

The Azure Cosmos DB platform provides many built-in capabilities to help meet various requirements of the GDPR and helps to significantly ease this process. Ranging from granular controls that can be defined, to integration with centralized authentication management services and industry-leading methods to protect and maintain the availability of data, Azure Cosmos DB - offer a wide set of powerful capabilities to address data privacy principles in the data platform tier.

## Further Reading

* [A technical overview](https://na01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fazure.microsoft.com%2Fen-us%2Fblog%2Fa-technical-overview-of-azure-cosmos-db%2F&data=02%7C01%7Cdharmas%40microsoft.com%7C41077c9f554f4920e44508d4cd2b7c77%7C72f988bf86f141af91ab2d7cd011db47%7C1%7C0%7C636359034044841051&sdata=qZeLmgFoxl4xZ6lL%2FKjkM3p2sb%2BWg2fwFnUhD%2BOVfDI%3D&reserved=0) of Azure Cosmos DB
* [Global distribution with Azure Cosmos DB.](https://docs.microsoft.com/en-us/azure/cosmos-db/distribute-data-globally)
* [Automatic regional failover for business continuity in Azure Cosmos DB](https://docs.microsoft.com/en-us/azure/cosmos-db/regional-failover)
* [How to partition and scale in Azure Cosmos DB](https://docs.microsoft.com/en-us/azure/cosmos-db/partition-data)
* [Tunable data consistency levels in Azure Cosmos DB](https://docs.microsoft.com/en-us/azure/cosmos-db/consistency-levels)
* [Performance and scale testing with Azure Cosmos DB](https://docs.microsoft.com/en-us/azure/cosmos-db/performance-testing)
* [Schema-agnostic indexing in Cosmos DB](http://www.vldb.org/pvldb/vol8/p1668-shukla.pdf)
* [How does Azure Cosmos DB index data](https://docs.microsoft.com/en-us/azure/cosmos-db/indexing-policies)
* [Request Units in Azure Cosmos DB](https://docs.microsoft.com/en-us/azure/cosmos-db/request-units)
* [Expire data in Azure Cosmos DB collections automatically with time to live](https://docs.microsoft.com/en-us/azure/cosmos-db/time-to-live)
* [Automatic online backup and restore with Azure Cosmos DB](https://docs.microsoft.com/en-us/azure/cosmos-db/online-backup-and-restore)
* [Integration of Cosmos DB with Spark](https://docs.microsoft.com/en-us/azure/cosmos-db/spark-connector)
* [Azure Cosmos DB SLA](https://azure.microsoft.com/en-us/support/legal/sla/cosmos-db/v1_0/)
* [Use cases](https://docs.microsoft.com/en-us/azure/cosmos-db/use-cases)

Need More Help?

Please send your questions to [askcosmosdb@microsoft.com](mailto:askcosmosdb@microsoft.com).