**CORE SERVICE DESIGN:**

**Cosmos DB**

atabricks

|  |  |
| --- | --- |
| IT Owner Details | |
| **Department** | DTS |
| **Contact Name** | Dominic Panzera |
| **Address** | 375 Manningham Road, Doncaster, Victoria 3108 |

|  |  |
| --- | --- |
| Document Control | |
| Title | Ambulance Victoria – Cosmos DB Core Service Design |
| **File Name** | Ambulance Victoria – Cosmos DB Core Service Design v1.0.docx |
| **Version** | 1.0 |
| **Status** | Released |
| **Release Date** | 8/03/2024 |

|  |  |  |  |
| --- | --- | --- | --- |
| Preparation |  |  |  |
| **Prepared** | Daniela Nikolic |  |  |
| **Authorised** | Dileep Pradeep |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Version Tracking | | | |  |
| Version | Remarks | Change Requested | Pages Affected | Release Date |
| 1.0 | Initial Release | N/A | All | 8/03/2024 |

Table of Contents

[1. Overview 5](#_Toc160795456)

[1.1 Purpose and Audience 5](#_Toc160795457)

[1.2 Scope and Key Deliverables 5](#_Toc160795458)

[1.3 Glossary and Definitions 6](#_Toc160795459)

[2. Executive Summary 7](#_Toc160795460)

[3. Resource Cost 8](#_Toc160795461)

[4. WAF and Security Control Alignment 8](#_Toc160795462)

[4.1 Reliability 9](#_Toc160795463)

[4.1.1 Overview 9](#_Toc160795464)

[4.1.2 Cosmos DB Reliability Checklist 9](#_Toc160795465)

[4.2 Cost Optimisation 10](#_Toc160795466)

[4.2.1 Overview 10](#_Toc160795467)

[4.2.2 Cosmos DB Cost Optimisation Checklist 10](#_Toc160795468)

[4.3 Operational Excellence 12](#_Toc160795469)

[4.3.1 Overview 12](#_Toc160795470)

[4.3.2 Cosmos DB Operational Excellence Checklist 12](#_Toc160795471)

[4.4 Performance Efficiency 14](#_Toc160795472)

[4.4.1 Overview 14](#_Toc160795473)

[4.4.2 Cosmos DB Performance Efficiency Checklist 14](#_Toc160795474)

[4.5 Security 16](#_Toc160795475)

[4.5.1 Overview 16](#_Toc160795476)

[5. Architecture Summary 17](#_Toc160795477)

[5.1 Resource Overview 17](#_Toc160795478)

[5.1.1 Throughput 18](#_Toc160795479)

[5.2 RBAC 19](#_Toc160795480)

[5.3 Design Decisions and Justifications 19](#_Toc160795481)

[5.3.1 Availability 19](#_Toc160795482)

[5.3.2 Regions 20](#_Toc160795483)

[5.3.3 Failover 20](#_Toc160795484)

[5.3.4 Monitoring 20](#_Toc160795485)

[5.3.5 Throughput 20](#_Toc160795486)

[5.3.6 Consistency 21](#_Toc160795487)

[5.3.7 Alerts 21](#_Toc160795488)

[5.3.8 Connectivity 21](#_Toc160795489)

[5.3.9 Managed Identities 21](#_Toc160795490)

[5.3.10 Encryption 22](#_Toc160795491)

[5.3.11 Threat Detection 22](#_Toc160795492)

[5.3.12 Logging 22](#_Toc160795493)

[6. Azure Policies 22](#_Toc160795494)

[7. Configuration Templates 23](#_Toc160795495)

[7.1 Primary Region Production Cosmos NoSQL Database 23](#_Toc160795496)

[7.2 Secondary Region Production Cosmos NoSQL Database 24](#_Toc160795497)

[7.3 Secondary Region Non-Production Cosmos NoSQL Database 25](#_Toc160795498)

[8. Acceptance 26](#_Toc160795499)

# Overview

This document covers the baseline design for the Cosmos DB core service. The intention of this document is to define the overall resource design in isolation from a specific application. It is aimed to highlight the general process and requirements for building a Cosmos DB in a repeatable fashion with consistent configurations. Design decisions and justifications have been included in the Architecture section, and this document can be used as a reference for new builds that require a Cosmos DB.

This design caters to a Level 2 design which covers both Microsoft’s WAF (Well Architected Framework)[[1]](#footnote-2) and the Department of Health Control list.

Any deviations required to the standards defined in this document will require separate exemption and approval from the Cloud Governance Forum if they are required for any reason for a specific build.

## Purpose and Audience

This document will outline the standard design and configuration of this Azure service in Ambulance Victoria’s Azure tenancy as a baseline for any application infrastructure deployments.

This design is intended to:

* Meet Microsoft WAF standards.
* Meet the controls stipulated by the Department of Health.
* Define the baseline required for the deployment of the resource.

The audience for this document is those involved in the planning, designing, and implementing of the Application/Data infrastructure. This includes:

* + Ambulance Victoria IT staff

It is assumed that the reader knows and is familiar with Azure Cloud concepts and related topics.

## Scope and Key Deliverables

The scope of this core service design is to define the baseline deployment requirements and standards for the Cosmos DB core service.

The key deliverables for this are:

* This design to outline the service definition Level 2 baseline standards.
* A technical configuration document that defines the deployment of this resource for each of the Service Tiers, or for any other logical standard such as size
* IaC templates for repeatable deployment of this core service

## Glossary and Definitions

|  |  |
| --- | --- |
| **Term** | **Definition** |
| **AV** | Ambulance Victoria |
| **WAF** | Well Architected Framework |
| **CAF** | Cloud Adoption Framework |
| **Level 1** | Refers to a resource that has been designed to a CAF standard |
| **Level 2** | Refers to a resource that has been designed to a WAF standard with Department of Health controls overlayed |
| **AZ 2** | Refers to Ambulance Victoria’s legacy Azure Landing Zone still in use in some regards |
| **AZ 3** | Refers to Ambulance Victoria’s current Azure Landing Zone, also referred to as the Enterprise landing zone. This is the target state for migrations. |
| **SLA** | Service Level Agreement as defined by Microsoft |
| **DH** | Department of Health |
| **IaC** | Infrastructure as Code |
| **NSG** | Network Security Groups |

Table : Glossary and definitions

# Executive Summary

This design covers the baseline standards for the Cosmos DB Core Service. This service has been assessed against the five pillars of WAF as well as the Department of Health Security Controls.

This section contains a summary of the major design decisions that have been made for defining the baseline of this resource as an outcome of the WAF and Security analysis detailed throughout this document.

Of the five WAF Pillars, it was found that Reliability Performance Efficiency Operational Excellence Cost Optimisation and Security were relevant.

For this service the main baseline configurations include:

* Public connectivity will be disabled, and private endpoints will be used for connectivity.
* Managed Identities will be enabled.
* Minimum TLS will be set to 1.2.
* Defender for Cosmos DB will be enabled.
* Monitoring will be enabled via Application Insights.

There are some notable differences across the service tier configurations for this service:

* Production workloads will have Service Managed Failover enabled. This will not be enabled for Non-Production.
* Multi-region writes will be enabled for Production services across Australia Southeast and Australia East.
* For Non-production dev and test, the Cosmos DB emulator will be used where possible, else the serverless throughput will be used.
* Backups will be set as Continuous – with 30 days for Production and 7 days for Non-Production

# Resource Cost

Azure Cosmos DB bills across three usages:

* Compute
* Storage
* Bandwidth.

The following assumes the NoSQL selection with autoscale throughput:

|  |  |  |
| --- | --- | --- |
| Autoscale provisioned throughput | Total RU/s\* per hour | Price per 100 RU/s |
| Single-region write account | 100 RU/s x 1.5 x 1 region | **$0.015**/hour |
| Single-region write account with data distributed across multiple regions (with or without availability zones\*) | 100 RU/s x 1.5 x N regions | **$0.015**/hour |
| Multi-region write (formerly called multi-master) account distributed across multiple regions\*\* | 100 RU/s x N regions | **$0.029**/hour |

Table : Pricing Construct for Request Units

*\*RUs are Request Units.*

|  |  |  |
| --- | --- | --- |
| Consumed Storage | Total GB | Price |
| Transactional storage (row-oriented) | 1 GB x N regions | **$0.440**/month |
| Analytical storage (column-oriented) | 1 GB x N regions | **$0.046**/month |

Table : Pricing construct for consumed storage

# WAF and Security Control Alignment

The following are the five pillars of the Microsoft Well Architected Framework:

* [Reliability](https://learn.microsoft.com/en-us/azure/well-architected/#reliability)
* [Cost optimization](https://learn.microsoft.com/en-us/azure/well-architected/#cost-optimization)
* [Operational excellence](https://learn.microsoft.com/en-us/azure/well-architected/#operational-excellence)
* [Performance efficiency](https://learn.microsoft.com/en-us/azure/well-architected/#performance-efficiency)
* [Security](https://learn.microsoft.com/en-us/azure/well-architected/#security)

For this design, the security section will also cover the Department of Health Controls in addition with any Microsoft Security Best Practices. Each of these sections will detail relevant controls or baseline requirements for this core service that will be put in place.

## Reliability

### Overview

The term reliability refers to the availability of the system and its ability to recover from failure[[2]](#footnote-3). Resiliency strategies must be built into each element of the architecture. The pillars of reliability include:

* Design for business requirements
* Design for failure
* Observe application health
* Drive Automation

### Cosmos DB Reliability Checklist

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | Checklist Item | Applicable to AV | Built Into Design | Enforcement Option | Applicability |
| **R1** | Distribute your Azure Cosmos DB account across availability zones (when available). | No | No | N/A | N/A |
| **R2** | Configure your Azure Cosmos DB account to span at least two regions. | Yes | Yes | IaC | At deployment |
| **R3** | Enable service-managed failover for your account. | Yes | Yes | IaC | At deployment |
| **R4** | Validate availability by testing failover manually with service-managed failover temporarily disabled. | Yes | No | Governance | Operational - during application deployment |

Table : WAF Reliability checklist summary

## Cost Optimisation

### Overview

The cost optimisation pillar is structured to support creating cost-effective workloads in the cloud[[3]](#footnote-4). It looks at removal of unnecessary spend and improving operational efficiency. The principles of cost optimisation revolve around:

* Choosing the correct resources
* Setting up budgets and maintaining cost constraints
* Dynamically allocate and deallocate resources
* Optimising workloads whilst aiming for scalable costs
* Continuously monitoring and cost managing

### Cosmos DB Cost Optimisation Checklist

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | Checklist Item | Applicable to AV | Built Into Template | Enforcement Option | Applicability |
| **CO1** | Monitor RU/s utilization and patterns. | Yes | Yes | Governance  IaC – Application Insights | At deployment, Operational – review as required |
| **CO2** | Customize your indexing policy to map to your workload. | Yes | No | Governance | Operational – during application design and deployment |
| **CO3** | Select partition key[s] that are ideal for your workload. | Yes | No | Governance | Operational – during application design and deployment |
| **CO4** | Use serverless or provisioned throughput | Yes | Yes | IaC | At deployment |
| **CO5** | Configure the default consistency level for your application. When appropriate, downgrade the default consistency level in client sessions. | Yes | Yes | IaC | At deployment |
| **CO6** | For dev/test workloads, use the Azure Cosmos DB emulator. | Yes | No | Governance | Operational – during application design/non-production deployment |
| **CO7** | Use transactional batch operations | Yes | No | Governance | Operational – during application design and deployment |
| **CO8** | Use projection to reduce throughput costs of large query result sets. | Yes | No | Governance | Operational – during application design and deployment |
| **CO9** | Avoid using unbounded cross-partition queries. | Yes | No | Governance | Operational – during application design and deployment |
| **C10** | Implement time-to-live (TTL) to remove unused items. | Yes | No | Governance | Operational – during application design and deployment |
| **C11** | Consider an analytical store for heavy aggregations. | Yes | No | Governance | Operational – during application design and deployment |

Table : WAF Cost Optimisation checklist summary

## Operational Excellence

### Overview

Operational Excellence aims to ensure that once the architecture is built, the ongoing operations are flawless. This includes repeatable and reliable deployments, automating to eliminate human error. To do this the following must be considered:

* Optimise the build and release process (including CI/CD and IaC)
* Understand Operational Health
* Test recovery and failure
* Focus on continuous improvement
* Use loosely coupled architecture

### Cosmos DB Operational Excellence Checklist

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | Checklist Item | Applicable to AV | Built Into Template | Enforcement Option | Applicability |
| **OE1** | Ensure application developers are using the latest version of the developer SDK. | Yes | No | Governance | Operational – during application design |
| **OE2** | Create identifiers in the client application to differentiate workloads. | Yes | No | Governance | Operational – during application design |
| **OE3** | Capture supplemental diagnostics using the developer SDK. | Yes | No | Governance | Operational – during application design and deployment |
| **OE4** | Create alerts associated with host machine resources. | Yes | No | Governance | Operational – during application design and deployment |
| **OE5** | Use the bulk features of client SDKs for large operations. | Yes | No | Governance | Operational – during application design and deployment |
| **OE6** | Create alerts for throughput throttling. | Yes | Yes | IaC | At deployment |
| **OE7** | Track query performance using metrics. | Yes | No | Governance | Operational – during application design and deployment |
| **OE8** | Use templates to automatically deploy account resources. | Yes | Yes | IaC | At deployment |
| **OE9** | Track key metrics to identify common problems in your workload. | Yes | No | Governance | Operational – during application deployment |

Table : WAF Operational Excellence checklist summary

## Performance Efficiency

### Overview

Performance Efficiency refers to the ability of your systems and applications to meet user demands without breaking or creating a negative user experience[[4]](#footnote-5). This covers capacity and scalability:

* Design for horizontal scaling
* Run stress and performance tests
* Continuously monitor performances, particularly in Production systems

### Cosmos DB Performance Efficiency Checklist

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | Checklist Item | Applicable to AV | Built Into Template | Enforcement Option | Applicability |
| **PE1** | Configure your throughput based on your performance baseline. | Yes | No | Governance | Operational – during application design and deployment |
| **PE2** | Use optimization techniques on the client and server sides when appropriate. | Yes | No | Governance | Operational – during application design and deployment |
| **PE3** | Deploy Azure Cosmos DB for NoSQL to regions closest to your end users. | Yes | Yes | IaC | At deployment |
| **PE4** | Configure the SDK for Direct mode. | Yes | No | Governance | Operational – during application design and deployment |
| **PE5** | Disable indexing for bulk operations. | Yes | No | Governance | Operational – during application design and deployment |
| **PE6** | Create composite indexes for fields that are used in complex operations. | Yes | No | Governance | Operational – during application design and deployment |
| **PE7** | Optimize host client machines for the SDKs. | Yes | No | Governance | Operational – during application design and deployment |
| **PE8** | Use the singleton pattern for the CosmosClient class in most SDKs. | Yes | No | Governance | Operational – during application design and deployment |
| **PE9** | Keep item sizes less than 100 KB in size. | Yes | No | Governance | Operational – during application design and deployment |
| **PE10** | Use subqueries strategically to optimize queries that join large data sets. | Yes | No | Governance | Operational – during application design and deployment |
| **PE11** | Use analytical workloads for the most complex queries. | Yes | No | Governance | Operational – during application design and deployment |

Table : WAF Performance Efficiency checklist summary

## Security

### Overview

Security refers to the ability of the environment to resist and manage threats.

This section covers both Microsoft Best Practices as well as relevant security controls provided by the Department of Health. With respect to the Microsoft WAF, Security is underpinned by the following[[5]](#footnote-6):

* Plan resources and how to harden them
* Automate and use least privilege
* Classify and encrypt data
* Monitor system security, plan incident response
* Identify and protect endpoints
* Protect against code-level vulnerabilities
* Model and test against potential threats

In addition to the Microsoft controls, the Department of Health has mandated security posture to Ambulance Victoria. Note there may be duplication between the Microsoft Security Best Practices and the Department of Health controls.

The following Microsoft Security Benchmarks are applicable:

* NS-1: Establish network segmentation boundaries
* NS-2: Secure cloud services with network controls
* IM-3: Manage application identities securely and automatically
* DP-2: Monitor anomalies and threats targeting sensitive data
* DP-3: Encrypt sensitive data in transit
* DP-4: Enable data at rest encryption by default
* LT-1: Enable threat detection capabilities
* LT-4: Enable logging for security investigation
* BR-1: Ensure regular automated backups

# Architecture Summary

## Resource Overview

Azure Cosmos DB is a fully managed NoSQL relational database that can support development operations for AI, commerce, IoT, and many other types of solutions[[6]](#footnote-7). It fully manages the underlying infrastructure. Whilst it is well positioned as a database for AI, it can also be used to support Web Applications, mobile, gaming, and other services with the capability to handle large amounts of data, reads, and writes with near-real time response time.

It supports various APIs including MongoDB, PostgreSQL, Apache Cassandra, Apache Gremlin and Table:

A diagram of a software development process

Description automatically generated

Figure : Supported APIs under Cosmos DB

Azure Cosmos DB is often used as the middle portion of an overall application or analytics capability as shown below:

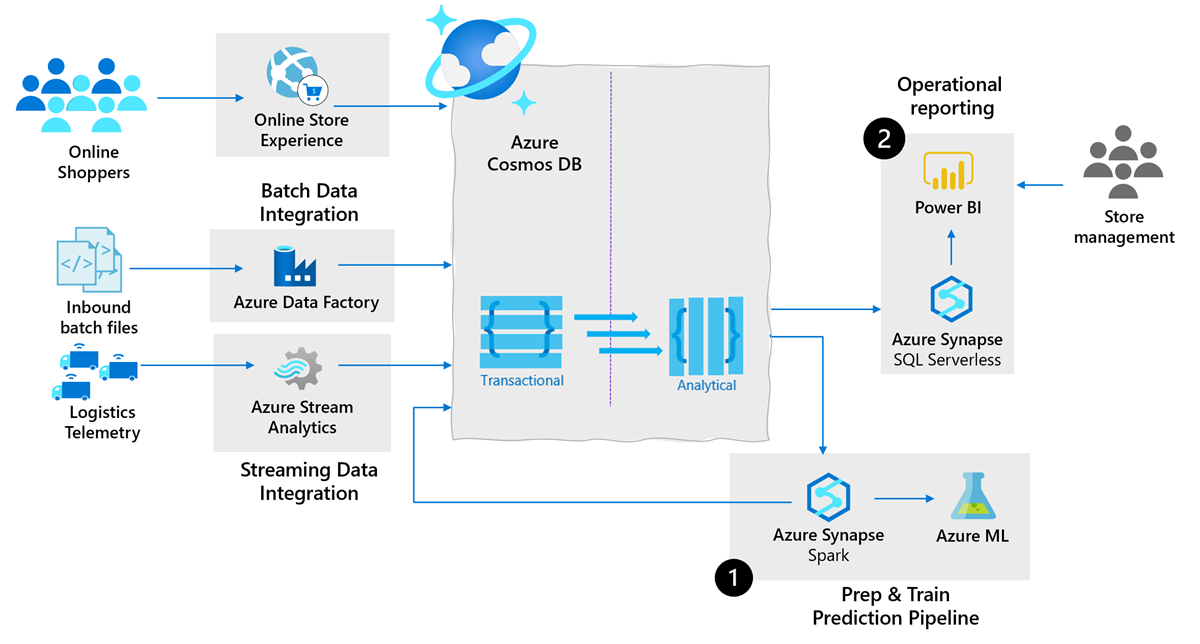


Figure : Cosmos DB for Transactions and Analytics

### Throughput

Cosmos DB offers two capacity modes: serverless and provisioned throughput. Both modes will perform the same database operations, but the pricing is different. The following outlines the feature differences:

|  |  |  |
| --- | --- | --- |
| Criteria | Provisioned throughput | Serverless |
| Best suited for | Workloads with sustained traffic requiring predictable performance | Workloads with intermittent or unpredictable traffic and low average-to-peak traffic ratio |
| How it works | For each of your containers, you configure some amount of provisioned throughput expressed in [Request Units (RUs)](https://learn.microsoft.com/en-us/azure/cosmos-db/request-units) per second. Every second, this quantity of Request Units is available for your database operations. Provisioned throughput can be updated manually or adjusted automatically with [autoscale](https://learn.microsoft.com/en-us/azure/cosmos-db/provision-throughput-autoscale). | You run your database operations against your containers without having to configure any previously provisioned capacity. |
| Geo-distribution | Available (unlimited number of Azure regions) | Unavailable (serverless accounts can only run in a single Azure region) |
| Maximum storage per container | Unlimited | 1 TB1 |
| Performance | < 10-ms latency for point-reads and writes covered by SLA | < 10-ms latency for point-reads and < 30 ms for writes covered by SLO |
| Billing model | Billing is done on a per-hour basis for the RU/s provisioned, regardless of how many RUs were consumed. | Billing is done on a per-hour basis for the number of RUs consumed by your database operations. |

Table : Feature comparison of throughput modes

## RBAC

The following roles are applicable to Cosmos DB[[7]](#footnote-8):

|  |  |
| --- | --- |
| Role Name | Description |
| Cosmos DB Account Reader Role | Can read Azure Cosmos DB account data. See DocumentDB Account Contributor for managing Azure Cosmos DB accounts. |
| Cosmos DB Operator | Lets you manage Azure Cosmos DB accounts, but not access data in them. Prevents access to account keys and connection strings. |
| CosmosBackupOperator | Can submit restore request for a Cosmos DB database or a container for an account |
| CosmosRestoreOperator | Can perform restore action for Cosmos DB database account with continuous backup mode |
| DocumentDB Account Contributor | Can manage Azure Cosmos DB accounts. Azure Cosmos DB is formerly known as DocumentDB. |

Table : RBAC roles relevant for this core service

## Design Decisions and Justifications

This section covers the design decisions and justifications that reflect the findings of the WAF and Security alignment. This will form the baseline requirements for the Cosmos DB core service and will be captured in the accompanying Configuration Template with a set of pre-approved deployment settings for this resource. Any changes, modifications or removals to the pre-approved deployments must have specific approval from the Cloud Governance Forum prior to deployment.

### Availability

**Design Reference:** Table 4 – [R1](#_Cosmos_DB_Reliability)

**Design Decision**: The Cosmos DB account will not be distributed across availability zones as they are not available in the Primary region. If they become available, then the account should be distributed across Availability Zones.

**Design Justification**: Availability Zones are currently not available in Australia Southeast so this is not applicable to the Primary region. If they do become available in this region in future, then they should be reviewed and applied for Cosmos DB.

### Regions

**Design Reference:** Table 4 – [R2](#_Cosmos_DB_Reliability), Table 7 – [PE3](#_Cosmos_DB_Performance)

**Design Decision:** The Cosmos DB account will be configured to span Australia Southeast and Australia East. Multi-region writes will also be enabled for Platinum services.

**Design Justification:** Australia Southeast and Australia East are the architecturally defined regions for Azure deployment for Ambulance Victoria.

### Failover

**Design Reference:** Table 4 – [R3](#_Cosmos_DB_Reliability)

**Design Decision:** Service Managed Failover will be enabled for Production. Non-Production does not need to have Failover configured.

**Design Justification:** Service Managed Failover is preferred to manual failover as the base setting as it removes a step that may be missed in a disaster recovery scenario by an administrator. It also makes the failover more instantaneous resulting in less potential downtime for the Database.

### Monitoring

**Design Reference:** Table 5 – [CO1](#_Cosmos_DB_Cost)

**Design Decision:** Application Insights will be deployed to enable monitoring of Request Units.

**Design Justification:** Application Insights is the native method that captures Throughput, Requests, Storage, Availability, Latency, System (metadata), and Management Operations[[8]](#footnote-9). Request Units should be monitored through here.

### Throughput

**Design Reference:** Table 5 – [CO4](#_Cosmos_DB_Cost)

**Design Decision:** Provisioned throughput will be used for Production and mission-critical workloads. Serverless can be used for Non-Production.

**Design Justification:** Serverless throughput can only be run in a single Azure region, which means that Production workloads are not suitable for this. It also has higher performance with respect to latency.

### Consistency

**Design Reference:** Table 5 – [CO5](#_Cosmos_DB_Cost)

**Design Decision:** Session Consistency will be set as the default.

**Design Justification:** Session consistency is the most widely used setting for both single region and multi-region deployments. It offers a fairly optimal trade-off between availability and latency and strongest consistency[[9]](#footnote-10).

### Alerts

**Design Reference:** Table 6 – [OE6](#_Cosmos_DB_Operational)

**Design Decision:** Alerts will be set for Throughput throttling. It will be set for the Normalized RU consumption reaching greater than 80%.

**Design Justification:** This is considered a critical metric for Cosmos DB and so an alert will be set to ensure that an action can be taken to reduce throughput or evaluate the performance of the application up or down stream.

### Connectivity

**Design Reference:** Microsoft Security Benchmark [NS-1, NS-2](#_Overview)

**Design Decision:** Public access will be disabled and private endpoints will be deployed.

**Design Justification:** Private endpoints are the most secure form of connectivity. Public access must be disabled to meet this security requirement.

### Managed Identities

**Design Reference: M**icrosoft Security Benchmark [IM-3](#_Overview)

**Design Decision:** Managed Identities will be used for application authorisation to Cosmos DB.

**Design Justification:**

### Encryption

**Design Reference**: Microsoft Security Benchmark [DP-3, DP-4](#_Overview)

**Design Decision:** The minimum TLS will be set to 1.2

**Design Justification:** TLS is currently the most secure level of this protocol and will be set as the baseline.

### Threat Detection

**Design Reference:** Microsoft Security Benchmark [LT-1](#_Overview)

**Design Decision:** Defender for Cosmos DB will be enabled.

**Design Justification:** Defender for Cosmos DB offers a native capability to detect attempts to exploit Cosmos DB databases and protects against common attacks such as SQL injection. It will also monitor for suspicious accesses.

### Logging

**Design Reference:** Microsoft Security Benchmark [LT-4](#_Overview)

**Design Decision:** Diagnostic logs will be collected and sent to the central log analytics workspace in that region.

**Design Justification:** The collection of logs will assist in diagnosing any potential issues or anomalies in the operation of the service.

# Azure Policies

There are no additional Azure Policies required for this service.

# Configuration Templates

## Primary Region Production Cosmos NoSQL Database

|  |  |  |
| --- | --- | --- |
| Configuration item | Configuration Template | |
| **Account Name for NoSQL** | cosnoavaccount |  |
| **Database Name** | cosmosprdause[appname]01 |  |
| **Subscription** | AV ALZ [Subscription Name] | |
| **Region** | Australia Southeast | |
| **Availability Zones** | N/A | |
| **Capacity Mode** | Provisioned throughput | |
| **Apply free tier discount** | Apply | |
| **Geo-redundancy** | Enabled | |
| **Multi-region writes** | Enabled | |
| **Connectivity Method** | Private Endpoint | |
| **Minimum TLS** | 1.2 | |
| ***Private Endpoint Settings*** |  | |
| **Allow access from Azure Portal** | Allow | |
| **Allow access from my IP** | Deny | |
| **Allow Public Network Access** | Deny | |
| **PE Name** | pe-cosmos-prd-ause-[appname]-01 | |
| **CosmosDB Sub-resource** | Azure Cosmos DB for NoSQL | |
| **Private DNS Zone** | privatelink.documents.azure.com | |
| ***Backup Settings*** |  | |
| **Backup Type** | Continuous (30 days) | |
| **Encryption** | Service-managed keys | |
| ***Diagnostic Settings*** |  | |
| **Logs** | allLogs | |
| **Metrics** | Requests | |
| **Send to** | Log Analytics Workspace (log-prd-ause-mgmt-01) | |

## Secondary Region Production Cosmos NoSQL Database

|  |  |  |
| --- | --- | --- |
| Configuration item | Configuration Template | |
| **Account Name for NoSQL** | cosnoavaccount |  |
| **Database Name** | cosmosprdauea[appname]01 |  |
| **Subscription** | AV ALZ [Subscription Name] | |
| **Region** | Australia East | |
| **Availability Zones** | N/A | |
| **Capacity Mode** | Provisioned throughput | |
| **Apply free tier discount** | Apply | |
| **Geo-redundancy** | Enabled | |
| **Multi-region writes** | Enabled | |
| **Connectivity Method** | Private Endpoint | |
| **Minimum TLS** | 1.2 | |
| ***Private Endpoint Settings*** |  | |
| **Allow access from Azure Portal** | Allow | |
| **Allow access from my IP** | Deny | |
| **Allow Public Network Access** | Deny | |
| **PE Name** | pe-cosmos-prd-auea-[appname]-01 | |
| **CosmosDB Sub-resource** | Azure Cosmos DB for NoSQL | |
| **Private DNS Zone** | privatelink.documents.azure.com | |
| ***Backup Settings*** |  | |
| **Backup Type** | Continuous (30 days) | |
| **Encryption** | Service-managed keys | |
| ***Diagnostic Settings*** |  | |
| **Logs** | allLogs | |
| **Metrics** | Requests | |
| **Send to** | Log Analytics Workspace (log-prd-auea-mgmt-01) | |

## Secondary Region Non-Production Cosmos NoSQL Database

|  |  |  |
| --- | --- | --- |
| Configuration item | Configuration Template | |
| **Account Name for NoSQL** | cosnoavaccount |  |
| **Database Name** | Cosmos[env]ause[appname]01 |  |
| **Subscription** | AV ALZ [Subscription Name] | |
| **Region** | Australia Southeast | |
| **Availability Zones** | N/A | |
| **Capacity Mode** | Serverless | |
| **Apply free tier discount** | Apply | |
| **Geo-redundancy** | Disabled | |
| **Multi-region writes** | Disabled | |
| **Connectivity Method** | Private Endpoint | |
| **Minimum TLS** | 1.2 | |
| ***Private Endpoint Settings*** |  | |
| **Allow access from Azure Portal** | Allow | |
| **Allow access from my IP** | Deny | |
| **Allow Public Network Access** | Deny | |
| **PE Name** | pe-cosmos-[env]-ause-[appname]-01 | |
| **CosmosDB Sub-resource** | Azure Cosmos DB for NoSQL | |
| **Private DNS Zone** | privatelink.documents.azure.com | |
| ***Backup Settings*** |  | |
| **Backup Type** | Continuous (7 days) | |
| **Encryption** | Service-managed keys | |
| ***Diagnostic Settings*** |  | |
| **Logs** | allLogs | |
| **Metrics** | Requests | |
| **Send to** | Log Analytics Workspace (log-prd-ause-mgmt-01) | |

# Acceptance

Signature of this page by appropriately delegated representatives of ​Ambulance Victoria​ signifies acceptance of this design document.

Logicalis will commence build and implementation work once it receives a signed copy of this design document.

Signature of this page by appropriately delegated representatives of Ambulance Victoria signifies acceptance of this design document.

Logicalis will commence build and implementation work once it receives a signed copy of this design document.

|  |  |
| --- | --- |
| Project | Core Services |
| Document Version | 1.0 |

**Signed on behalf of Ambulance Victoria**

|  |  |
| --- | --- |
| Name | Dan Howarth |
| Position |  |
| Signature |  |
| Date signed |  |

**Signed on behalf of Logicalis Australia**

|  |  |
| --- | --- |
| Name | Daniela Nikolic |
| Position | Senior Cloud Engineer |
| Signature |  |
| Date signed |  |

1. https://learn.microsoft.com/en-us/azure/well-architected/ [↑](#footnote-ref-2)
2. https://learn.microsoft.com/en-us/azure/well-architected/resiliency/overview [↑](#footnote-ref-3)
3. https://learn.microsoft.com/en-us/azure/well-architected/cost/overview [↑](#footnote-ref-4)
4. https://learn.microsoft.com/en-us/azure/well-architected/scalability/overview [↑](#footnote-ref-5)
5. https://learn.microsoft.com/en-us/azure/well-architected/security/security-principles [↑](#footnote-ref-6)
6. https://learn.microsoft.com/en-us/azure/cosmos-db/introduction [↑](#footnote-ref-7)
7. https://learn.microsoft.com/en-us/azure/role-based-access-control/built-in-roles [↑](#footnote-ref-8)
8. https://learn.microsoft.com/en-us/azure/cosmos-db/monitor-normalized-request-units [↑](#footnote-ref-9)
9. https://learn.microsoft.com/en-us/azure/cosmos-db/consistency-levels?WT.mc\_id=Portal-Microsoft\_Azure\_DocumentDB#strong-consistency [↑](#footnote-ref-10)