Elastic search Backup and Restoration Test

|  |  |
| --- | --- |
| Author | Gowtham / Pradeep |
| Issue Date | 4th Aug 2023 |
| Version | 1 |
| Distribution list | Platform engineering |

Contents

[Introduction 2](#_Toc142079666)

[Objective 3](#_Toc142079667)

[Full Back up snapshot process 3](#_Toc142079668)

[Restore snapshots (DR site) 7](#_Toc142079669)

# Introduction

This document describes the procedures for creating backup snapshots and conducting tests for the maya.ai product data stored in Elastic Search. Regular snapshot creation and testing are vital to ensure the effectiveness of our backup strategies, providing assurance in our ability to recover data in the event of a data loss incident.

# Objective

The primary objective of these procedures is to confirm the integrity of the backup snapshots for maaya.ai and the effectiveness of the restoration process. Through successful execution of these tests, we can ensure that the data backup and restore processes for maya.ai meet the required Recovery Point Objective (RPO) and Recovery Time Objective (RTO).

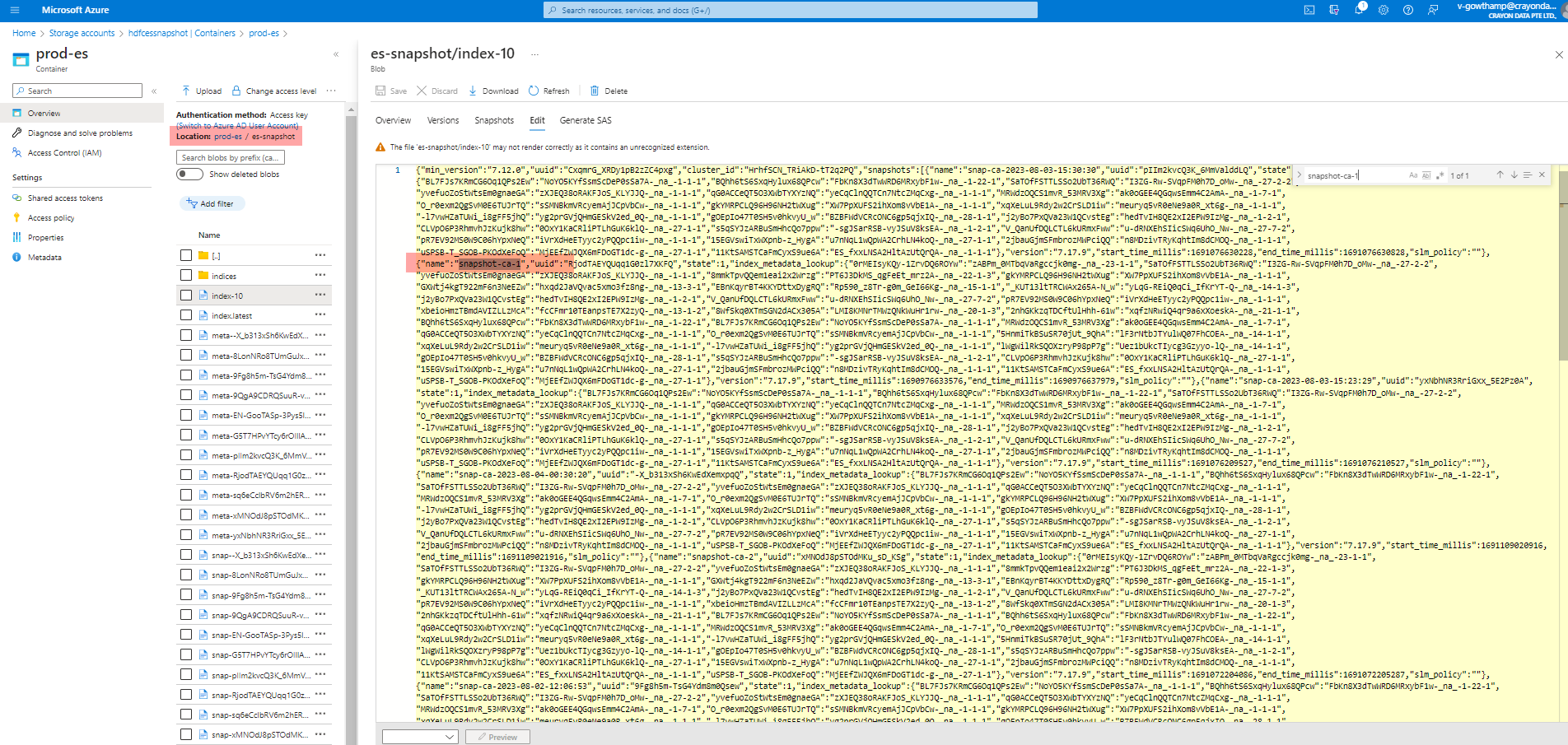
# Full Back up snapshot process

Snapshots will be taken every 6 hours and then stored at the production blob storage located in Western India (Mumbai). The data will then be immediately replicated to the blob storage at the DR site located in Central India (Pune). In the event of any failures, restoration will occur at the DR sites

To facilitate this process, the account name and key are added to the Elasticsearch keystore. Following this, we proceed to create the repository.

After the repository is created, we perform a test to ensure that the snapshots are being accurately sent to the storage account. This test validates the proper functioning of the backup process and verifies that the snapshots are correctly stored and accessible in the storage account.

Please refer to the accompanying screenshot, which provides evidence that the snapshot is securely stored in the 'hdfcessnapshot' storage account. Within this account, the snapshot is located in the 'prod-es' container, with the base path designated as 'es-snapshot'.



Now that we are Automating the Snapshot backup every 6 hours using the Azure Automation runbook, we have a Python Script.

A screenshot of a computer

Description automatically generated

For a 6-hour schedule, we need to create a schedule in the automation runbook.

A screenshot of a computer

Description automatically generated

This Schedule will take a backup every 6 hours. A snapshot will be created with the name snap-snap-ca-%Y-%m-%d-%H:%M:%S ."Time will be available in UTC.

A screenshot of a computer

Description automatically generated

Once the Snapshot has been taken,

We have to Create a Sync Prod Storage account to a DR Storage Account, where we have another Automation Script that Will perform this action.

A screenshot of a computer

Description automatically generated

We are enabling the Sync for this automation script every 6 hours.

A screenshot of a computer

Description automatically generated

The production snapshot was taken at 12:01 PM.

A screenshot of a computer

Description automatically generated

DR Sync happened at 12:15 PM

A screenshot of a computer

Description automatically generated

# Restore snapshots (DR site)

Restore the Snapshot using the below Curl command

A black screen with white text

Description automatically generated

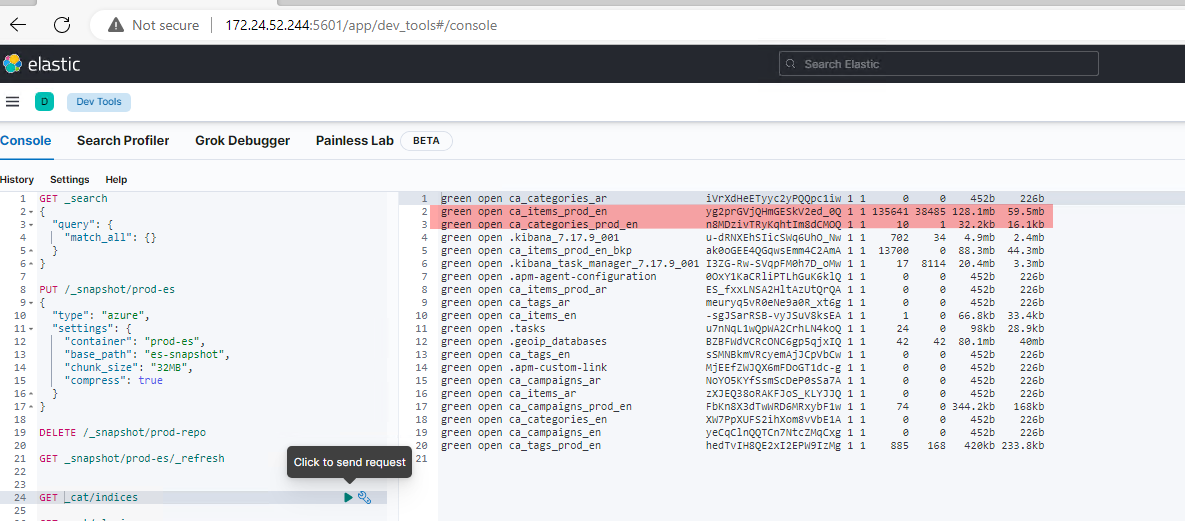
Run the below to check if all the indies have been restored.

A screen shot of a computer

Description automatically generated

Once the restoration is completed, it needs to be tested.

Prod indices data screenshot



DR indices data screenshot

A screenshot of a computer

Description automatically generated

Based on the Above screenshots, the Data and the size should be the same.