

Backup Elasticsearch Snapshot Data to Dell EMC Isilon

Backup and restore Elasticsearch data to Dell EMC Isilon using the Snapshot API

Abstract

This document describes how to configure, backup and restore Elasticsearch data to Dell EMC™ Isilon™.

April 2020

Revisions

Date	Description
April 2020	Initial release

Acknowledgments

This paper was produced by the Unstructured Technical Marketing Engineering and Solution Architects team.

Author: Rich Paulson

The information in this publication is provided “as is.” Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software described in this publication requires an applicable software license.

This document may contain certain words that are not consistent with Dell's current language guidelines. Dell plans to update the document over subsequent future releases to revise these words accordingly.

This document may contain language from third party content that is not under Dell's control and is not consistent with Dell's current guidelines for Dell's own content. When such third party content is updated by the relevant third parties, this document will be revised accordingly.

Copyright © 2021 Dell Inc. or its subsidiaries. All Rights Reserved. Dell Technologies, Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners. [3/16/2021] [Configuration and Deployment] [H18274]

Table of contents

Revisions.....	2
Acknowledgements.....	2
Table of contents	3
Executive summary.....	4
Objectives	4
Audience	4
1 Solution overview	5
1.1 Solution architecture.....	5
1.2 Key components.....	5
2 Solution implementation	6
2.1 Implementation workflow	6
2.2 Installation and configuration steps	6
2.2.1 Step 1: Prepare the Isilon cluster	6
2.2.2 Step 2: Prepare the Elasticsearch nodes	7
2.2.3 Step 3: Register Isilon as a repository.....	8
2.2.4 Step 4: Take a snapshot (backup) of an Index.....	9
2.2.5 Step 5: Restore a snapshot to the same cluster	11
3 Best Practices.....	13
A Technical support and resources	14
A.1 Related resources	14

Executive summary

Elasticsearch is a distributed, RESTful search and analytics engine capable of storing your data so you can discover the expected and uncover the unexpected. Dell EMC™ Isilon™ is the industry's #1 family of scale-out network-attached storage systems, designed for demanding enterprise file workloads.

Elasticsearch provides a smart solution to backup single indices or entire clusters to remote storage that can be a shared filesystem such as Isilon, object storage such as ECS or even HDFS. Snapshots can be used to backup indices for retention purposes or as a point in time recovery solution. Backing up the data to Dell EMC storage allows customers to have peace of mind that their data is protected.

Objectives

This document illustrates how configure Elasticsearch to store snapshot data in Dell EMC Isilon using the Elasticsearch backup and restore API.

Audience

This document is intended for administrators who manage Elasticsearch deployments. This guide assumes a level of technical knowledge for the devices and technologies described herein.

1 Solution overview

This section provides an overview of the Dell EMC Isilon integration with Elasticsearch including the key technologies used.

1.1 Solution architecture

The below figure illustrates the architectural workflow used in this guide. Server system metrics are collected by Metricbeat and forwarded to the Elasticsearch nodes. When a snapshot of the index is initiated, it is written to and restored from Isilon.

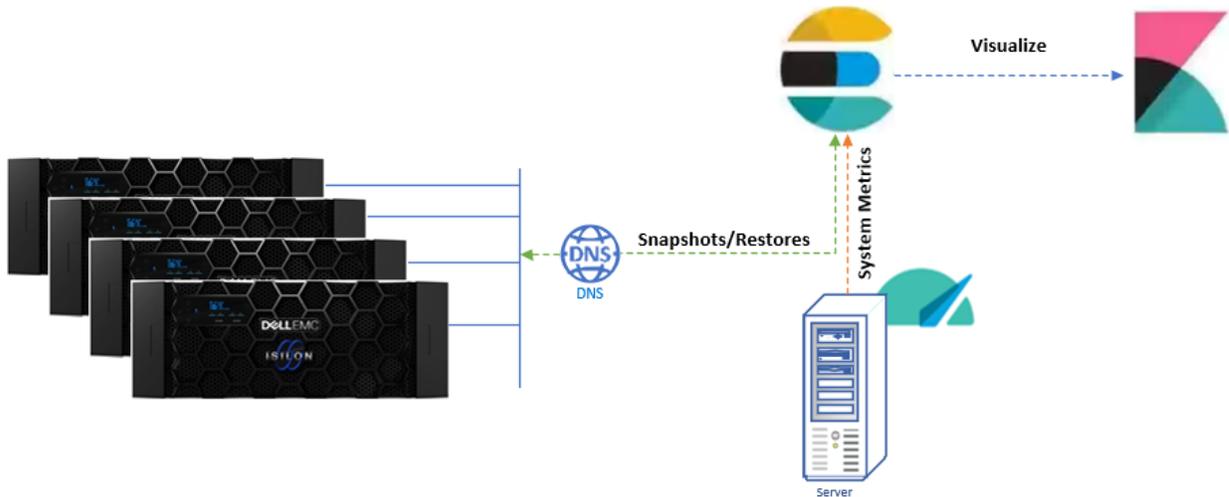


Figure 1 Architectural Workflow

1.2 Key components

The following components and versions were used for the examples in this guide.

Table 1 Dell EMC Components

Component	Version	Description
Isilon	OneFS 8.2	Physical 4-node H500

Table 2 Elasticsearch Components

Component	Version	Description
Elasticsearch	7.5.1	Three node virtualized environment
Kibana	7.5.1	Single virtualized machine
Metricbeats	7.5.1	Single virtualized machine

2 Solution implementation

This section describes the high-level steps required to configure Elasticsearch to store snapshot data to Dell EMC Isilon. Note that this guide assumes Elasticsearch and Kibana (optional) have been installed and are functioning.

Note that in this example, we are using Metricbeat which is a lightweight shipper to collect various system metrics, the Metricbeat index is what we'll snapshot to and restore from Dell EMC Isilon.

2.1 Implementation workflow

The below outlines the steps this document will use to implement Isilon as a snapshot repository.



Figure 2 Implementation Steps

2.2 Installation and configuration steps

The following steps outline the steps to prepare OneFS to store Elasticsearch snapshots, setup the snapshot repository and initiate a snapshot of an index as well as a restore.

2.2.1 Step 1: Prepare the Isilon cluster

In this section we'll prepare the Isilon cluster to store Elasticsearch snapshots by creating a user, group, directory and NFS export.

- a. Log in to one of the Isilon nodes as the root account with SSH.
- b. Create a Splunk user and group in the access zone of your choice. The NFS client uses this user to access Isilon.

```
isi auth groups create elasticsearch --provider local --gid <Elastic group ID> -
-zone System
hop-isi-x-4#
hop-isi-x-4# isi auth groups create elasticsearch --provider local --gid 115 --zone system
hop-isi-x-4#
```

```
isi auth users create elasticsearch --password <Elastic user password> --
primary-group elasticsearch --provider local --uid <Elastic user ID> --enabled
yes --zone system
```

```
hop-isi-x-4#
hop-isi-x-4# isi auth users create elasticsearch --password isilon-123 --primary-group
elasticsearch --provider local --uid 111 --enabled yes --zone system
hop-isi-x-4#
```

Note: The user ID and group ID of the elasticsearch user in OneFS must be the same as the one Elasticsearch uses. A user named 'elasticsearch' is normally the default user on *nix systems.

- c. Create a base directory `/ifs/data/elasticsearch/snapshots` to store snapshots. This base directory will be configured as an Isilon NFS export later.

```
mkdir -p /ifs/data/elasticsearch/snapshots
```

- d. Grant read/write permissions to the elasticsearch user on the directory.

```
chown -R 111:115 /ifs/data/elasticsearch/snapshots
chmod -R 755 /ifs/data/elasticsearch/snapshots
```

- e. Create an NFS export to allow the Elasticsearch node to access the Isilon zone.

```
isi nfs exports create /ifs/data/elasticsearch/snapshots --all-dirs=yes --zone
system
```

```
hop-isi-x-4#
hop-isi-x-4# isi nfs exports create /ifs/data/elasticsearch/snapshots --all-dirs=yes --
zone system
hop-isi-x-4#
```

2.2.2 Step 2: Prepare the Elasticsearch nodes

Mount the Isilon NFS export to the Linux File System on the Elasticsearch data node to allow Elasticsearch to store the snapshot data on the Isilon storage. To mount the Isilon NFS export:

Note: It is highly recommended to use SmartConnect so that traffic is evenly distributed to each node in the Isilon cluster. Install the NFS package for your distribution if it is not already installed on the system and start the related services. This package provides NFS-related utilities that allows mounting of the NFS share.

The shared file system repository ("type": "fs") uses the shared file system to store snapshots. In order to register the shared file system repository, it is necessary to mount the same shared filesystem to the same location on all Elasticsearch nodes. This location (or one of its parent directories) must be registered in the `path.repo` setting on all Elasticsearch nodes.

- a. Assuming that the OneFS filesystem is mounted to `/mnt/elasticsearch/snapshots`, the following setting should be added to `elasticsearch.yml` file on each Elasticsearch node:

```
path.repo: ["/mnt/elasticsearch/snapshots"]
```

- b. Create a mount point for the Isilon export and mount it. Ensure that the elasticsearch user owns the mount point.

```
mkdir -p /mnt/elasticsearch/snapshots
```

```
sudo mount <SmartConnect Zone name>:/ifs/data/elasticsearch/snapshots -o
nfsvers=4,tcp,rw,hard,intr,timeo=600,retrans=2 /mnt/elasticsearch/snapshots/
elastic-01:
elastic-01: sudo mount hop-isi-x.solarch.lab.emc.com:/ifs/data/elasticsearch/snapshots
-o nfsvers=4,tcp,rw,hard,intr,timeo=600,retrans=2 /mnt/elasticsearch/snapshots
elastic-01:
```

- c. Add the Isilon NFS Export to the file systems table **/etc/fstab** to ensure it is automatically mounted on system startup.

```
<SmartConnect Zone Name>:/ifs/data/elasticsearch/snapshots
/mnt/elasticsearch/snapshots    nfs    defaults    0    0
```

2.2.3 Step 3: Register Isilon as a repository

Elasticsearch provides REST APIs to register the repository. The below example registers Isilon as a repository using the `_snapshot` API. Note that this can be accomplished using either **curl** or the **Dev Tools** tab in Kibana.

Note: There are several supported settings which can be used when using a shared filesystem to store snapshots. Reference the [Shared file system repository](#) section for more details.

- a. Create the snapshot repository. In this example the repository is named `isilon` and we specified the location as the mount point which was created on each Elasticsearch node.

```
curl -X PUT "10.246.156.199:9200/_snapshot/isilon" -H 'Content-Type:
application/json' -d'
{
  "type": "fs",
  "settings": {
    "location": "/mnt/elasticsearch/snapshots/",
    "compress": true
  }
}
```

```
Elastic-01: curl -X PUT "10.246.156.199:9200/_snapshot/isilon" -H 'Content-Type: application/json' -d'
> {
> "type": "fs",
> "settings": {
>   "location": "/mnt/elasticsearch/snapshots/",
>   "compress": true
> }
> }
> '
{"acknowledged":true}Elastic-01:
Elastic-01:
```

Figure 3 Create the Isilon Snapshot Repository

- b. Verify that the repository was successfully created by using the 'POST /_snapshot/isilon/_verify' call.

```
Elastic-01: curl -X POST 'http://10.246.156.199:9200/_snapshot/isilon/_verify?format=json&pretty'
{
  "nodes" : {
    "XbxUE9ciT9eY9RLZ-FQzIQ" : {
      "name" : "aries01"
    },
    "TuE79wDBQ565S03bEjbslw" : {
      "name" : "aries03"
    },
    "1JAsqqb0QrqlHfEa37tNAw" : {
      "name" : "aries02"
    }
  }
}
```

Figure 4 S3 Repository Verification

2.2.4 Step 4: Take a snapshot (backup) of an Index

Next, we'll take a snapshot of an index which contains the Metricbeat data from the illustration in Section 1.1.

- a. List the Metricbeat indices and select one to snapshot using the GET /_cat/indices API

```
Elastic-01: curl -X GET '10.246.156.199:9200/_cat/indices/metric*?v&s=index&pretty'
health status index          uuid                                pri rep docs.count docs.del
eted store.size pri.store.size
green open  metricbeat-7.5.1-2020.02.27-000001 UggwTyRPQAi6c06iyfs-mQ  1  1  7237603
0 3.8gb 1.9gb
green open  metricbeat-7.5.1-2020.03.28-000002 7AG8qiEZTxeemDIZjVZYZA  1  1  4053364
0 2.2gb 1.1gb
```

Figure 5 Metricbeat indices

- b. Take a snapshot of the `metricbeat-7.5.1-2020.02.27-000001` index using the `PUT /_snapshot` API

```
Elastic-01: curl -X PUT "10.246.156.199:9200/_snapshot/isilon/snapshot_1?wait_for_completion=true&pretty" -H 'Content-Type: application/json' -d'
> {
>   "indices": "metricbeat-7.5.1-2020.02.27-000001",
>   "ignore_unavailable": true,
>   "include_global_state": false
> }
> }
> '
{
  "snapshot" : {
    "snapshot" : "snapshot_1",
    "uuid" : "6t8BK-TIR5GdZzGH-YYeBw",
    "version_id" : 7050199,
    "version" : "7.5.1",
    "indices" : [
      "metricbeat-7.5.1-2020.02.27-000001"
    ],
    "include_global_state" : false,
    "state" : "SUCCESS",
    "start_time" : "2020-04-14T17:45:17.724Z",
    "start_time_in_millis" : 1586886317724,
    "end_time" : "2020-04-14T17:46:16.995Z",
    "end_time_in_millis" : 1586886376995,
    "duration_in_millis" : 59271,
    "failures" : [ ],
    "shards" : {
      "total" : 1,
      "failed" : 0,
      "successful" : 1
    }
  }
}
```

Figure 6 Perform a snapshot

2.2.5 Step 5: Restore a snapshot to the same cluster

In this example we'll restore one of our snapshots using the `POST /_snapshot/isilon/snapshot_x/_restore` API.

Since an existing index can be only restored if it's closed and has the same number of shards as the index in the snapshot, we'll restore our snapshot to a new index with a different name.

Note: Reference the Elasticsearch [Snapshot and Restore](#) documentation for detailed information on restoring indices.

- a. List the snapshots in the from the Isilon repository using the `GET _cat/snapshots` API.

```
Elastic-01:
Elastic-01: curl -X GET "10.246.156.199:9200/_cat/snapshots/isilon?v&s=id"
id      status start_epoch start_time end_epoch end_time duration indices successful_shards failed_shards total_shards
snapshot_1 SUCCESS 1586886317 17:45:17 1586886376 17:46:16 59.2s      1              1              0              1
Elastic-01:
```

Figure 7 List repository snapshots

- b. Restore `snapshot_1` to a new index using the prefix

Verify that the snapshot was restored to a new index using the prefix `restored_index` and the `POST /_snapshot/isilon/snapshot_1/_restore` API.

```
Curl -X POST "10.246.156.199:9200/_snapshot/isilon/snapshot_1/_restore?pretty" -
H 'Content-Type: application/json' -d
```

```
Elastic-01: curl -X POST "10.246.156.199:9200/_snapshot/isilon/snapshot_1/_restore?pretty" -H 'Content-
Type: application/json' -d'
> {
>   "indices": "metricbeat-7.5.1-2020.02.27-000001",
>   "ignore_unavailable": true,
>   "include_global_state": true,
>   "rename_pattern": "metricbeat-(.+)",
>   "rename_replacement": "restored_index_$1"
> }
> '
{
  "accepted" : true
}
```

Figure 8 Restore a snapshot stored in isilon

c. Verify that the index was restored by listing the index using the GET `/_cat/indices` API

```
Curl -X GET "10.246.156.199:9200/_cat/indices/restored_*?v&s=index&pretty"
```

```
Elastic-01: curl -X GET "10.246.156.199:9200/_cat/indices/restored_*?v&s=index&pretty"
health status index                                uuid                                pri rep docs.count docs.deleted store.s
ize pri.store.size
green open  restored_index_7.5.1-2020.02.27-000001 yRkjpavwS2eQx0hAhXmcVA  1  1  7237603          0  3.
8gb          1.9gb
```

3 Best Practices

The following configuration and tuning changes are recommended.

Table 3 Configuration and tuning recommendations

Description	Detail
Use SmartConnect to distribute traffic to the isilon nodes	Through a single host name, SmartConnect enables client-connection load balancing and dynamic network file system (NFS) failover and failback of client connections across storage nodes to provide optimal utilization of the cluster resources.
Mount the shared filesystem to the same location on all Elasticsearch nodes	The shared file system repository ("type": "fs") uses the shared file system to store snapshots. In order to register the shared file system repository, it is necessary to mount the same shared filesystem to the same location on all Elasticsearch nodes. This location (or one of its parent directories) must be registered in the <code>path.repo</code> setting on all Elasticsearch nodes.
Snapshot retention	Automatic deletion of older snapshots is an optional feature of snapshot lifecycle management (SLM). https://www.elastic.co/guide/en/elasticsearch/reference/current/slm-retention.html
Elasticsearch Curator	Consider using curator to automate and manage your Elasticsearch snapshots https://www.elastic.co/guide/en/elasticsearch/client/curator/current/index.html

A Technical support and resources

[Dell.com/support](https://dell.com/support) is focused on meeting customer needs with proven services and support.

[Storage technical documents and videos](#) provide expertise that helps to ensure customer success on Dell EMC storage platforms.

A.1 Related resources

Isilon product documentation

- Dell EMC Isilon Info Hub
 - <https://community.emc.com/docs/DOC-75133>

Elasticsearch product documentation

- Introduction to Elasticsearch
 - <https://www.elastic.co/guide/en/elasticsearch/reference/current/elasticsearch-intro.html>
- Elasticsearch Snapshot and Restore
 - <https://www.elastic.co/guide/en/elasticsearch/reference/7.6/snapshot-restore.html>