



# Predix Search



# Contents

<b>Predix Search Overview</b>	<b>1</b>
About Predix Search	1
Predix Search Clusters	1
<b>Get Started with Predix Search</b>	<b>3</b>
Predix Search Service Setup	3
Create a Predix Search Service Instance	3
Use the cf CLI to Create a Predix Search Service Instance	4
Bind an Application to a Predix Search Service Instance	5
Unbind a Predix Search Service Instance	7
Update a Predix Search Service Instance	7
Delete a Predix Search Service Instance	8
Authorities and Scopes Required for Predix Search	8
<b>Use a Predix Search Service Instance</b>	<b>9</b>
Index and Query the Predix Search Cluster	9
Use Snapshot and Restore	10
API Documentation	10
<b>Troubleshoot Predix Search</b>	<b>12</b>
Authentication Errors	12
Cloud Foundry CLI Errors	13
 <b>Chapter 2: Release Notes</b>	 <b>14</b>
Fourth Quarter of 2018	15

# Copyright GE Digital

© 2020 General Electric Company.

GE, the GE Monogram, and Predix are either registered trademarks or trademarks of General Electric Company. All other trademarks are the property of their respective owners.

This document may contain Confidential/Proprietary information of General Electric Company and/or its suppliers or vendors. Distribution or reproduction is prohibited without permission.

THIS DOCUMENT AND ITS CONTENTS ARE PROVIDED "AS IS," WITH NO REPRESENTATION OR WARRANTIES OF ANY KIND, WHETHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO WARRANTIES OF DESIGN, MERCHANTABILITY, OR FITNESS FOR A PARTICULAR PURPOSE. ALL OTHER LIABILITY ARISING FROM RELIANCE UPON ANY INFORMATION CONTAINED HEREIN IS EXPRESSLY DISCLAIMED.

Access to and use of the software described in this document is conditioned on acceptance of the End User License Agreement and compliance with its terms.



# Predix Search Overview

## About Predix Search

Predix Search adds powerful search capabilities to your applications.

Predix Search stores and searches big data in near real time. Powered by Elasticsearch (version 6.2), Predix Search exposes an extensive REST API accessible through a variety of programming languages, including Java, Python, PHP, .NET, and SQL.

The API enables you to:

- Index documents formatted in JSON
- Search the contents of indices
- Perform advanced search operations such as sorting, filtering, scripting, pagination, and aggregation
- Update and delete indices
- Check cluster, node, and index status and statistics

## Predix Search Clusters

Predix Search clusters deliver high performance by enabling federated indexing and searching across the individual nodes that make up the server cluster.

### Dedicated Master Clusters

Predix Search includes the option of adding a dedicated master cluster to increase overall system stability. Dedicated master clusters offload cluster management tasks to master nodes. Because master nodes do not take part in indexing and search-related data processing, system stability is enhanced by creating a separation of responsibilities between the master nodes and data nodes in the system. At least one dedicated master cluster is mandatory if the total number of nodes is more than 10.

Dedicated master nodes manage clusters by:

- Tracking all nodes in the cluster
- Tracking the number of indices in the cluster
- Tracking the number of Elasticsearch shards belonging to each index
- Maintaining routing information for nodes in the cluster
- Updating the cluster state after state changes, such as the creation of an index or addition or deletion of nodes in the cluster
- Replicating changes to the cluster state across all nodes in the cluster
- Monitoring the health of nodes by sending heartbeat signals, periodic signals that monitor the availability of the data nodes in the cluster

### Multiple Availability Zones

Predix Search ensures data availability and durability by enabling cluster deployment across availability zones (AZs). AZs are physically separate computing infrastructures designed for reliability. In the event of an infrastructure failure in the primary AZ, Predix Search automatically fails over to the standby AZ.

Multiple availability zone support (Multi-AZ) is an optional feature of Predix Search. Multi-AZ requires an even number of data nodes; half of the data nodes in the primary AZ, half in the standby. In Multi-AZ implementations, Elasticsearch shards are stored on nodes in the primary AZ, replicas on nodes in the standby.

**Additional Features**

- Data nodes scale to a maximum of 10 for the Dedicated-E10 plan, 20 for the Dedicated-E20 to Dedicated-E80 plans.
- Daily backups are archived for up to 14 days.

# Get Started with Predix Search

## Predix Search Service Setup

Before you can use Predix Search, you must create an instance of the service, bind the instance to your application, and enable authentication and authorization of the service instance.

Authentication and authorization are controlled by the User Account and Authentication (UAA) web service. You must set up a UAA service instance as a designated trusted issuer before creating your Predix Search service instance. For information about authentication and authorization in Predix services, see the [UAA Service Overview](#).

### Task Roadmap

**Note:** If you have used other services from the Predix catalog, you probably already have UAA services set up. If so, you can start with Task 4 below to create your Predix Search service instance.

Task	References
1 Configure your proxy settings.	Depending on your location and network configuration, you may need to configure your proxy settings to access remote resources.  See <a href="#">Defining Proxy Connections to Remote Resources</a> .
2 Deploy your application to Cloud Foundry.	See <a href="#">Creating and Deploying a Simple Web App to Cloud Foundry</a> .
3 Create a UAA service instance.	See <a href="#">Creating a UAA Service Instance</a> .
4 Create a Predix Search service instance.	See <a href="#">Create a Predix Search Service Instance</a> on page 3.  See also: <ul style="list-style-type: none"><li>• <a href="#">Update a Predix Search Service Instance</a> on page 7</li><li>• <a href="#">Delete a Predix Search Service Instance</a> on page 8</li></ul>
5 Bind your application to the Predix Search service instance.	See <a href="#">Bind an Application to a Predix Search Service Instance</a> on page 5.  See also <a href="#">Unbind a Predix Search Service Instance</a> on page 7.
6 Create an OAuth2 client for the Predix Search service instance.	The OAuth2 client enables the Predix Search service to make protected resource requests on behalf of the resource owner with the authorization of the resource owner.  See <a href="#">Creating an OAuth2 Client</a> .  See also: <ul style="list-style-type: none"><li>• <a href="#">Updating an OAuth2 Client for Services</a></li><li>• <a href="#">Authorities and Scopes Required for Predix Search</a> on page 8</li></ul>
7 Use the Predix Search service instance.	See <a href="#">Index and Query the Predix Search Cluster</a> on page 9.

## Create a Predix Search Service Instance

A Predix Search service instance gives your application indexing and search capabilities.

### Before You Begin

You must have a Predix account. See [Registering for a Predix Account](#).

## Procedure

1. Sign into your Predix account at <https://www.predix.io>.
2. In the main navigation bar, select **Catalog**.
3. Select the **Predix Cloud Services** tab.
4. Select the **Predix Cloud Connect** tile.
5. Select **Subscribe**.
6. Complete the entries on the **New Service Instance** page:

Field	Description
ORG	Select your organization (if you belong to more than one org).
SPACE	Select the space for your application.
USER ACCOUNT & AUTHENTICATION (UAA)	Choose an existing UAA instance or create a new instance of UAA. See <a href="#">Creating a UAA Service Instance</a> .
SERVICE INSTANCE NAME	Specify a unique name for your instance.
SERVICE PLAN	Select a plan.

7. Select **Subscribe to service**.

## Next Steps

Bind the Predix Search service instance to your application, after which you can retrieve service instance details from the `VCAP_SERVICES` environment variable for your application. See [Bind an Application to a Predix Search Service Instance](#) on page 5.

# Use the cf CLI to Create a Predix Search Service Instance

The Cloud Foundry command line interface (cf CLI) enables you to create a Predix Search service instance and examine its settings.

## Before You Begin

Bind the instance of your trusted issuer (UAA service) to your application. See [Connecting Your Application to a Platform Service Instance](#).

The binding enables you to obtain the UAA service instance details, including `issuerId`, from the `VCAP_SERVICES` environment variable for your application.

## About This Task

**Note:** Cloud Foundry CLI syntax can differ between Windows, Linux, and Macintosh operating systems. See the Cloud Foundry Help for the appropriate syntax for your operating system; for example, to see help for the `create-service`, or `cs`, command, run the command `cf help cs`.

## Procedure

1. Use the cf CLI to log into Cloud Foundry.

```
cf login
```

Enter your username and password when prompted by Cloud Foundry.



## 2. Create a Predix Search service instance.

```
cf cs predix-search <plan_name> <instance_name> -c
'{"trusted_issuer_ids":["<trusted_issuer_id>"],
"enable_dedicated_master":"<enable_dedicated_master>",
"enable_multi_az":"<enable_multi_az>",
"data_node_count":"<data_node_count>"}
```

- `<plan_name>` — Specify the service plan, such as **Dedicated-E10**.
- `<instance_name>` — Provide a unique name for the service instance.
- `<trusted_issuer_id>` — Substitute the `issuerID` of your trusted issuer (UAA instance); for example, `https://bc199ceb-7ef6-4783-a0c6-2d8aa7ce3300.predix-uaa.run.aws-usw02-pr.ice.predix.io/oauth/token`. Use a comma-separated list to specify multiple trusted issuers. You can obtain the `issuerId` from the `VCAP_SERVICES` environment variable for your application (if you have bound a UAA instance to your application).
- `<enable_dedicated_master>` — Enter `true` or `false` to specify whether your service instance includes a dedicated master cluster. You can omit the `enable_dedicated_master` parameter, in which case the default is `false`.

**Note:** A dedicated master cluster is mandatory if the number of data nodes is more than 10. See [Predix Search Clusters](#) on page 1.

- `<enable_multi_az>` — Enter `true` or `false` to specify whether your service instance includes multiple availability zones. You can omit the `enable_multi_az` parameter, in which case the default is `false`.
- `<data_node_count>` — Enter an integer value to specify the number of nodes in your Predix Search service cluster. If Multi-AZ is enabled, the data node count must be an even number. You can omit the `data_node_count` parameter, in which case the default is 1, unless Multi-AZ is enabled, in which case the default is 2.

**Note:** Predix Search supports 10 data nodes for the Dedicated-E10 plan and up to 20 nodes for all other plans. See [Predix Search Clusters](#) on page 1.

### Next Steps

Bind the Predix Search service instance to your application, after which you can retrieve service instance details from the `VCAP_SERVICES` environment variable for your application. See [Bind an Application to a Predix Search Service Instance](#) on page 5 below.

## Bind an Application to a Predix Search Service Instance

Binding your application to a Predix Search service instance adds the service connection details to the `VCAP_SERVICES` environment variable.

### About This Task

The Cloud Foundry runtime uses the `VCAP_SERVICES` environment variable to communicate with a deployed application about its environment. You can retrieve the following Predix Search service instance details from the `VCAP_SERVICES` environment variable:

- Authorization credentials
- HTTP header information for the service instance
- URI of the service instance

## Before You Begin

Your application must be available in Cloud Foundry. To learn about deploying your application to Cloud Foundry, see [Creating and Deploying a Simple Web App to Cloud Foundry](#).

## Procedure

1. Bind your application to a Predix Search service instance.

```
cf bind-service <app-name> <instance-name>
```

### For example:

```
$ cf bind-service sample_app_name test_instance_name
Binding service test_instance_name to app sample_app_name in org
predix-search / space search as predix.search@ge.com...
OK
TIP: Use 'cf restage sample_app_name' to ensure your env variable
changes take effect
```

2. Verify the binding and retrieve the Predix Search service instance credentials.

```
cf env <app-name>
```

### For example:

```
$ cf env sample_app_name
Getting env variables for app sample_app_name in org predix-search /
space search as predix.search@ge.com...
OK
"VCAP_SERVICES": {
  "predix-search": [
    {
      "credentials": {
        "data_node_count": 1,
        "dedicated_master_enabled": false,
        "zone_id": "8a350557-3f87-4222-84f9-b016cbe36fb1",
        "multi_az_enabled": false,
        "url": "https://predix-search-auth-app-prefectorial-dt.run.aws-
usw02-dev.ice.predix.io"
      },
      "label": "predix-search",
      "name": "test_instance_name",
      "plan": "es-small",
      "provider": null,
      "syslog_drain_url": null,
      "tags": [
        "Predix-Search",
        "ES"
      ],
      "volume_mounts": []
    }
  ]
}
```

## Unbind a Predix Search Service Instance

Unbind a Predix Search service instance to dissociate it from your application.

### Procedure

- To unbind a Predix Search service instance from an application, run the following command:

```
cf unbind-service <app-name> <instance-name>
```

## Update a Predix Search Service Instance

Update a Predix Search service instance to modify the specifications for the dedicated master cluster, availability zones, and number of data nodes.

### Procedure

1. Modify the dedicated master cluster specification.

```
cf update-service <instance-name> -c  
'{"enable_dedicated_master":"<true or false>"}'
```

**Note:** A dedicated master cluster is mandatory if the number of data nodes is more than 10.

2. Modify the availability zone specification.

```
cf update-service <instance-name> -c '{"enable_multi_az":"<true or  
false>"}'
```

3. Modify the data node count.

```
cf update-service <instance-name> -c  
'{"data_node_count":"<node_count>"}'
```

The data node count must be an integer value. If Multi-AZ is enabled, the data node count must be an even number.

### Note:

- You can increase but not decrease the `data_node_count` value for a Predix Search service instance.
- The maximum value for `data_node_count` for the Dedicated-E10 plan is 10; for all other plans, 20.

### Related Concepts

[Predix Search Clusters](#) on page 1

Predix Search clusters deliver high performance by enabling federated indexing and searching across the individual nodes that make up the server cluster.

## Delete a Predix Search Service Instance

Delete a Predix Search service instance when it is no longer needed.

### Procedure

To delete a Predix Search service instance, run the following command:

```
cf delete-service <instance-name>
```

## Authorities and Scopes Required for Predix Search

The OAuth2 client for your Predix Search service instance requires special parameters, including authorities and scopes, to enable the client to access the Predix Search cluster.

Required parameters:

- **Authorized Grant Types** — `refresh_token`, `client_credentials`, `password`
- **Scopes** — `uaa.none`
- **Authorities** — `uaa.none`, `predix-search.zones.<zone_id>.user`
  - `<zone_id>` — Provide the value of `predix-search[0].credentials.zone_id` from the `VCAP_SERVICES` environment variable for your application. See [Bind an Application to a Predix Search Service Instance](#) on page 5.
- **Allowed Providers** — `uaa`



## Related Information

[Elasticsearch API Documentation](#)

# Use Snapshot and Restore

Predix Search enables you to take snapshots of indices, view information about snapshots, and restore from snapshots any indices that have been closed or deleted in the Predix Search cluster.

## Snapshot

Predix Search takes an automated snapshot of all the indices in the cluster once every 24 hours.

To view all available automated snapshots, use the following endpoint:

```
GET /_snapshot/cs-automated/_all
```

... where `cs-automated` is the repository where automated snapshots are stored.

To take a snapshot of all open and started indices in the cluster, use the following endpoint:

```
PUT /_snapshot/<repository_name>/<snapshot_name>
```

- `<repository_name>` — Specify the name of the repository where the snapshot will be stored.
- `<snapshot_name>` — Provide a unique name for the snapshot.

## Restore

To restore all of the indices contained in a snapshot, use the following endpoint:

```
POST /_snapshot/<snapshot_repository>/<snapshot_name>/_restore
```

For example, to restore the snapshot named `daily-20170101` from the `cs_automated` repository:

```
POST /_snapshot/cs-automated/daily-20170101/_restore
```

You can also restore individual indices from a snapshot by specifying the names of the indices in the body of the POST request. For more information, see [Elasticsearch Snapshot and Restore](#).

# API Documentation

Predix Search supports an assortment of Elasticsearch operations.

## Supported API Endpoints — Elasticsearch 6.2

<ul style="list-style-type: none"><li>• <code>/_alias</code></li><li>• <code>/_aliases</code></li><li>• <code>/_all</code></li><li>• <code>/_analyze</code></li><li>• <code>/_bulk</code></li><li>• <code>/_cache/clear</code> (Index only)</li></ul>	<ul style="list-style-type: none"><li>• <code>/_delete_by_query</code></li><li>• <code>/_explain</code></li><li>• <code>/_field_stats</code></li><li>• <code>/_flush</code></li><li>• <code>/_forcemerge</code> (Index only)</li></ul>	<ul style="list-style-type: none"><li>• <code>/_scripts</code></li><li>• <code>/_search</code></li><li>• <code>/_search profile</code></li><li>• <code>/_segments</code> (Index only)</li><li>• <code>/_shard_stores</code></li><li>• <code>/_shrink</code></li></ul>
---	--	---

<ul style="list-style-type: none"> <li>• <code>/_cat</code></li> <li>• <code>/_cluster/allocation/explain</code></li> <li>• <code>/_cluster/health</code></li> <li>• <code>/_cluster/pending_tasks</code></li> <li>• <code>/_cluster/settings</code> for several properties: <ul style="list-style-type: none"> <li>◦ <code>action.auto_create_index</code></li> <li>◦ <code>action.search.shard_count.limit</code></li> <li>◦ <code>indices.breaker.field.data.limit</code></li> <li>◦ <code>indices.breaker.request.limit</code></li> <li>◦ <code>indices.breaker.total.limit</code></li> </ul> </li> <li>• <code>/_cluster/state</code></li> <li>• <code>/_cluster/stats</code></li> <li>• <code>/_count</code></li> </ul>	<ul style="list-style-type: none"> <li>• <code>/_ingest</code></li> <li>• <code>/_mapping</code></li> <li>• <code>/_mget</code></li> <li>• <code>/_msearch</code></li> <li>• <code>/_mtermvectors</code></li> <li>• <code>/_nodes</code></li> <li>• <code>/_plugin/kibana</code></li> <li>• <code>/_rank_eval</code></li> <li>• <code>/_recovery</code> (Index only)</li> <li>• <code>/_refresh</code></li> <li>• <code>/_reindex</code></li> <li>• <code>/_rollover</code></li> </ul>	<ul style="list-style-type: none"> <li>• <code>/_snapshot</code></li> <li>• <code>/_split</code></li> <li>• <code>/_stats</code></li> <li>• <code>/_status</code></li> <li>• <code>/_tasks</code></li> <li>• <code>/_template</code></li> <li>• <code>/_termvectors</code> (Index only)</li> <li>• <code>/update</code></li> <li>• <code>/update_by_query</code></li> <li>• <code>/_validate</code></li> </ul>
---	--	--

## Related Information

[Supported Elasticsearch Operations](#)

# Troubleshoot Predix Search

## Authentication Errors

You may encounter the following authentication error conditions in Predix Search:

### Token Expired

```
{
  "error": "invalid_token",
  "error_description": "Token is expired. Expiration date is [Thu Sep
14 02:19:35 UTC 2017].
                                Current date is [Thu Sep 14 06:31:55 UTC 2017]"
}
```

### Token Missing

```
{
  "error": "unauthorized",
  "error_description": "Full authentication is required to access this
resource"
}
```

### Token Incorrect

```
{
  "error": "invalid_token",
  "error_description": "Malformed Access Token"
}
```

### Zone ID Missing

```
{
  "error": "invalid_request",
  "error_description": "No zone specified for zone specific request: /
cluster/health"
}
```

### Zone Authorities Access Not Set Up in UAA

```
{
  "error": "invalid_token",
  "error_description": "Unauthorized zone access by principal: admin
for zone: 8a350557-3f87-4222-84f9-b016cbe36fb1"
}
```



## Cloud Foundry CLI Errors

Cloud Foundry CLI commands related to Predix Search may produce the following error conditions:

### **Data Node Count Exceeds the Maximum Allowed Limit**

- Error: `Server error`
- Error code: `10001`
- Status code: `502`,
- Message: `Service broker error: data_node_count: X is outside the minimum and maximum limit`

### **Enable Dedicated Master Set to False for Data Node Count > 10**

- Error: `Server error`
- Error code: `10001`
- Status code: `502`
- Message: `Service broker error: data_node_count: X is > 10 and masterNode disabled`

### **Odd Value Specified for Data Node Count and Multi-AZ is Enabled**

- Error: `Server error`
- Error code: `10001`
- Status code: `502`
- Message: `Service broker error: data_node_count: X is odd and multiAZ enabled`

# Chapter 2

---

## Release Notes

### Topics:

- [Fourth Quarter of 2018](#)

## Fourth Quarter of 2018

### Predix Search

The first release of Predix Search enables you to add full-featured search capabilities to your applications.

**Release Date: December 14, 2018**

In this release, Predix Search establishes an Elasticsearch implementation on Amazon Web Services, including the following features:

- Optional three-node dedicated master cluster
- Optional multiple availability zones
- Support for clusters having from one to 20 nodes

**Note:** This release does not include Kibana.