

**IBM Watson**

# **Watson Discovery on Managed CP4D in IBM Cloud**

**March 30, 2020**  
**Yoshinori Tahara**



# Overview

---

- Introduction of deployment procedure of Watson Discovery on managed OpenShift v3.11 with Cloud Pak for Data v2.5 in IBM Cloud
- Watson Discovery version is v2.1.2 (2020 March fixpack)

NOTE: This is currently only for evaluation in IBM

# Deployment Steps

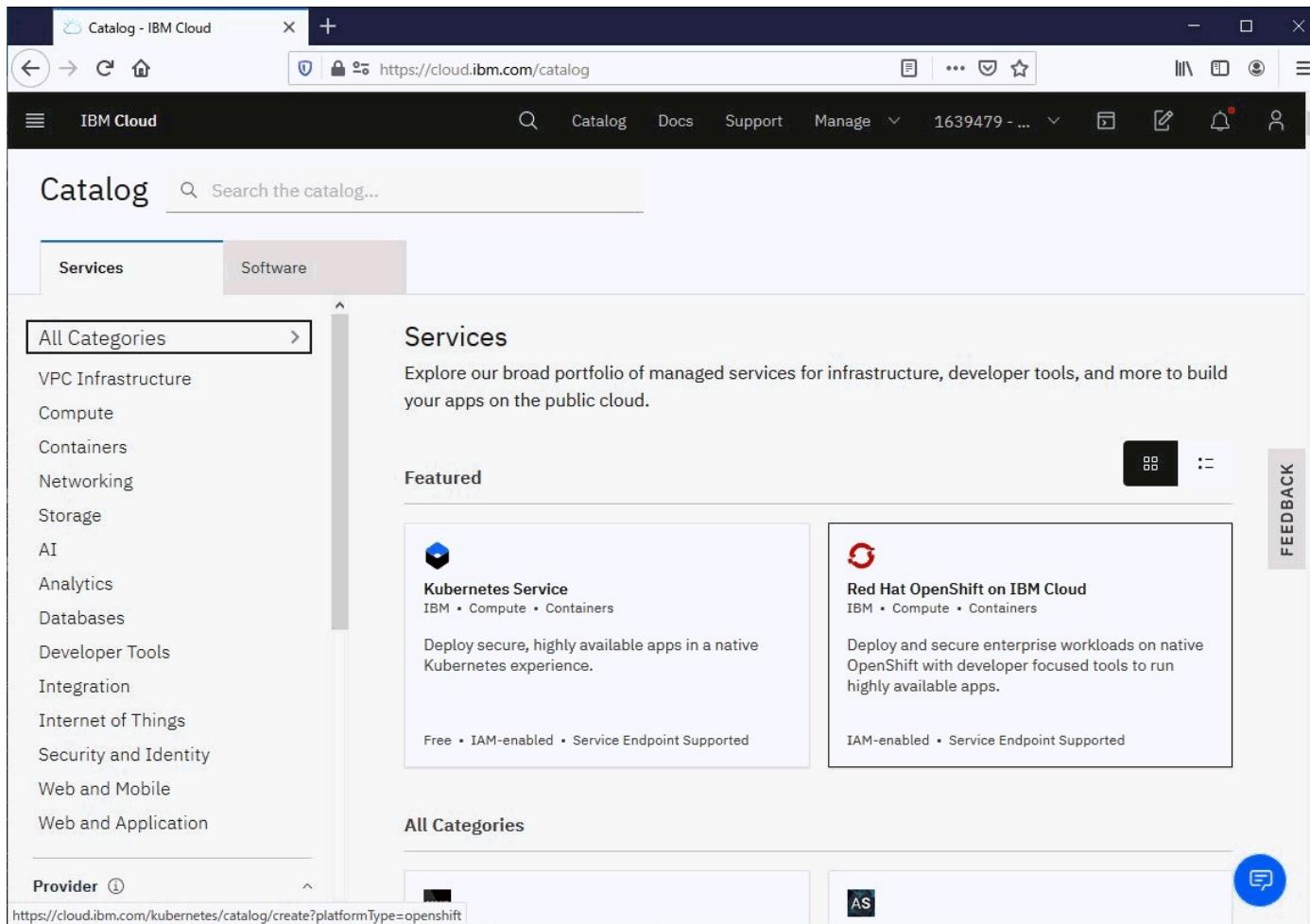
---

1. Create Managed OpenShift v3.11 cluster
2. Create CP4D v2.5 for the created OpenShift cluster
3. Create File Storage for Watson Discovery
4. Gather parameter values
5. Create NFS storage class and provisioner
6. Patch Watson Discovery v2.1.2 installation charts
7. Deploy the patched Watson Discovery v2.1.2
8. Launch Watson Discovery

# **1. Create Managed OpenShift v3.11 cluster**

---

# Select OpenShift in Catalog



The screenshot shows the IBM Cloud Catalog interface. On the left, there is a sidebar with a search bar and a list of categories under 'All Categories': VPC Infrastructure, Compute, Containers, Networking, Storage, AI, Analytics, Databases, Developer Tools, Integration, Internet of Things, Security and Identity, Web and Mobile, and Web and Application. Below this is a 'Provider' dropdown. The main area has tabs for 'Services' and 'Software', with 'Services' selected. A 'Featured' section displays two services: 'Kubernetes Service' and 'Red Hat OpenShift on IBM Cloud'. The 'Red Hat OpenShift on IBM Cloud' card includes a brief description, deployment requirements, and status indicators. At the bottom, there is another 'All Categories' section and a feedback button.

1. Login to IBM Cloud
2. Click Catalog and find OpenShift
3. Click “Red Hat OpenShift on IBM Cloud” tile

# Create OpenShift cluster (1 of 5)

The screenshot shows the 'Red Hat OpenShift Cluster' creation page on the IBM Cloud platform. The left sidebar has 'View All / Red Hat OpenShift Cluster' selected. The main area shows a 'Create' button and an 'About' button. Under 'Cluster type and version', there's a section for 'OpenShift' with a red circular logo and a dropdown menu showing '3.11.170 (Stable, Default)'. Below this, 'Cluster name' is set to 'mycluster' and 'Resource group' is 'default'. Under 'Tags', there's a note about examples: 'Examples: env:dev, version-1'. In the bottom right corner of the main form, there's a blue 'Create' button.

Order summary

b3c.4x16 - 4 vCPUs 16GB RAM	
9 worker nodes	\$2.42 / hr
Multizone load balancer ⓘ	\$0.02 / hr
IP allocation	
OCP license fee ⓘ	\$3,600.00 / month
Total*	\$5,357.52 / month estimated

\*Actual monthly total will vary with tiered pricing for the hourly worker nodes and the 30-day fixed OCP license fee.  
Additional charges for bandwidth might apply.  
[Learn more.](#)

Add to estimate

1. Set cluster name (default: "mycluster")

# Create OpenShift cluster (2 of 5)

Red Hat OpenShift Cluster

IBM Cloud

Location

Availability

Single zone Multizone

Geography

North America

Worker zone

Dallas 10  
 Dallas 12  
 Dallas 13  
 Mexico 01  
 Montreal 01  
 San Jose 03    No VLANs exist: VLANs will be created for you.

b3c.4x16 - 4 vCPUs 16GB RAM

3 worker nodes    **\$0.81 / hr**

OCP license fee    **\$1,200.00 / month**

Total\*    **\$1,781.04 / month** estimated

\*Actual monthly total will vary with tiered pricing for the hourly worker nodes and the 30-day fixed OCP license fee.  
Additional charges for bandwidth might apply.  
[Learn more.](#)

Add to estimate

Cancel Create

1. Select “Single Zone” in “Availability”
2. Select preferred Geo in “Geography”
3. Select preferred zone in “Worker zone”

NOTE: You will need to select same zone in same Geography for CP4D and File Storage.

As of March 29, 2020, CP4D deployment failed in Dallas 10 zone. CP4D deployment succeeded in San Jose 03 and Tokyo zone.

# Create OpenShift cluster (3 of 5)

Red Hat OpenShift Cluster

https://cloud.ibm.com/kubernetes/catalog/create?platformType=openshift

IBM Cloud

Default worker pool

Configure a set of worker nodes of the same flavor to create a default worker pool. Later, you can resize your worker pool to add or remove worker nodes. If you want a different flavor of worker node, you can create a new worker pool.

Filter

Flavor

Machine

- Bare Metal (0)
- Virtual - shared (9)
- Virtual - dedicated (9)

Use cases

- Balanced Cores and RAM (10)
- Extra local storage for SDS (0)
- GPU (0)
- RAM-intensive (4)
- Data-intensive (0)

Size

- Medium (10)
- Large (8)

4 vCPUs 16GB RAM

Virtual - shared  
b3c.4x16  
RHEL  
25GB SSD primary disk  
100GB SSD secondary disk  
1Gbps network speed

\$0.27 / hr

8 vCPUs 32GB RAM

Virtual - shared  
b3c.8x32  
RHEL  
25GB SSD primary disk  
100GB SSD secondary disk  
1Gbps network speed

\$0.50 / hr

16 vCPUs 16GB RAM

Virtual - shared  
c3c.16x16  
RHEL  
25GB SSD primary disk  
100GB SSD secondary disk  
1Gbps network speed

\$0.66 / hr

Order summary

b3c.16x16 - 16 vCPUs 64GB RAM

3 worker nodes      \$2.85 / hr

OCP license fee      \$4,800.00 / month

Total\*      \$6,852.00 / month

estimated

\*Actual monthly total will vary with tiered pricing for the hourly worker nodes and the 30-day fixed OCP license fee.

Additional charges for bandwidth might apply.  
[Learn more.](#)

Add to estimate

Cancel      Create

# Create OpenShift cluster (4 of 5)

The screenshot shows the IBM Cloud Catalog interface for creating a Red Hat OpenShift Cluster. On the left, there are four cluster configuration options listed:

- 16 vCPUs 32GB RAM**: Virtual - shared, c3c.16x32, RHEL, 25GB SSD primary disk, 100GB SSD secondary disk, 1Gbps network speed. Price: \$0.77 / hr.
- 16 vCPUs 64GB RAM**: Virtual - shared, b3c.16x64, RHEL, 25GB SSD primary disk, 100GB SSD secondary disk, 1Gbps network speed. Price: \$0.95 / hr. This option is highlighted with a blue border and a checkmark icon.
- 32 vCPUs 32GB RAM**: Virtual - shared, c3c.32x32, RHEL, 25GB SSD primary disk, 100GB SSD secondary disk, 1Gbps network speed. Price: \$1.23 / hr.
- 32 vCPUs 64GB RAM**: Virtual - shared, c3c.32x64, RHEL, 25GB SSD primary disk, 100GB SSD secondary disk, 1Gbps network speed. Price: \$1.50 / hr.

On the right, the **Order summary** section details the selection:

- b3c.16x64 - 16 vCPUs 64GB RAM
- 3 worker nodes      \$2.85 / hr
- OCP license fee ⓘ    \$4,800.00 / month
- Total\***            **\$6,852.00 / month**      estimated

Below the summary, there is a note about tiered pricing and bandwidth charges, with a "Learn more" link.

At the bottom, there are "Cancel" and "Create" buttons, with the "Create" button being blue.

1. Select “16 vCPUs 64GB RAM – Virtual shared”

# Create OpenShift cluster (5 of 5)

The screenshot shows the IBM Cloud interface for creating a Red Hat OpenShift Cluster. On the left, there's a sidebar with 'IBM Cloud' navigation and a main panel for 'Red Hat OpenShift Cluster'. The main panel includes sections for 'Worker nodes' (set to 3), 'Entitlements' (with a note about Cloud Pak OCP license), and an 'Infrastructure permissions checker' (showing 'Permission requirements and suggestions satisfied'). On the right, the 'Order summary' section details the configuration:

b3c.16x64 - 16 vCPUs 64GB RAM	\$1.87 / hr	
3 worker nodes	\$2.85 / hr	
IP allocation		
OCP license fee ⓘ	\$4,800.00 / month	
Cloud Pak entitlement applied	- \$4,800.00 / month	
Total*	\$2,052.00 / month	<small>estimated</small>

\*Actual monthly total will vary with tiered pricing for the hourly worker nodes and the 30-day fixed OCP license fee.

Additional charges for bandwidth might apply. [Learn more.](#)

At the bottom right, there are 'Cancel' and 'Create' buttons.

1. (Optional) Set “Encrypt local disk” off
2. Set number of “Worker nodes”. “3 nodes” is default.
3. Select “OCP entitlement”
4. Click “Create” button at bottom-right.

# OpenShift cluster creation has started

The screenshot shows the IBM Cloud interface for creating an OpenShift cluster named "mycluster". The status bar at the top indicates "Preparing master, workers...". On the left, a sidebar lists cluster management options like Access, Overview, Worker Nodes, Worker Pools, Add-ons, and DevOps (marked as New). The main content area is titled "Before your cluster provisions, set up your CLI tools" and provides two steps: 1. Run a PowerShell command to download and install CLI tools, and 2. Download the OpenShift CLI (oc) matching the local operating system and cluster version. It also includes instructions for logging in via the OpenShift web console or OAuth token request page. A terminal window at the bottom shows the command "oc version". A note at the bottom encourages setting up an IAM API key for automation.

1. Wait until creation completes

# OpenShift cluster is now ready for use

The screenshot shows the IBM Cloud web interface with the 'mycluster Cluster - IBM Cloud' tab selected. The main content area displays the 'Access' section for the 'mycluster' cluster. It includes instructions for setting up CLI tools and provides a PowerShell command for step 1. Step 2 instructs users to download the OpenShift CLI (oc) matching their local operating system and cluster version, with a link to the documentation. Below these steps, there's a section titled 'After your cluster provisions, gain access' with instructions for logging in using the 'OpenShift web console' button or the 'Copy Login Command' option. A terminal window at the bottom shows the command 'oc version' being run. A note at the bottom encourages users to set up an IAM API key for automation.

mycluster Cluster - IBM Cloud

https://cloud.ibm.com/kubernetes/clusters/bpvvl27d0hc2lsh0ov0/access?p

Clusters / mycluster • Normal Add tags

OpenShift web console Actions...

Access

Before your cluster provisions, set up your CLI tools

- Run this command with your PowerShell to download and install a few CLI tools and plugins.

```
Set-ExecutionPolicy Unrestricted; iex(New-Object Net.WebClient).DownloadString('https://mirror.openshift.com/pub/openshift-v4/clients/ocp/latest/install.ps1')
```
- Download the OpenShift CLI (oc) that matches your local operating system and cluster version. For information about how to install the CLI, [see the docs](#).

After your cluster provisions, gain access

- When your cluster is fully deployed, log in to it with your IBMid by using one of the following methods:
  - Browse to the OpenShift console by using the button below. From the dropdown menu in the upper right of the page, click Copy Login Command. Paste the copied command in your local terminal.
  - Browse to the [oauth token request page](#). Follow the instructions on the page.
- Verify you can communicate with your cluster.  
oc version

If you want to automate logging in to your cluster, [see the docs](#) to set up an IAM API key.

1. After cluster is ready ("Normal"), launch OpenShift web console. Click "OpenShift web console" blue button.

# OpenShift web console

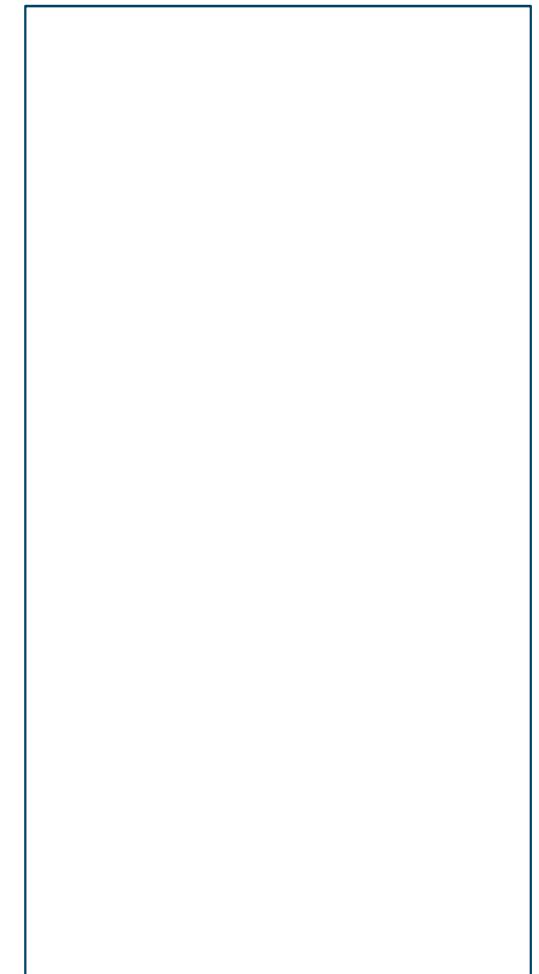
The screenshot shows the OpenShift Web Console interface. At the top, there's a header bar with tabs for "mycluster Cluster - IBM Cloud" and "OpenShift Web Console". The URL in the address bar is <https://c103-e.us-south.containers.cloud.ibm.com:30063/console/catalog>. Below the header, the main content area has a dark header bar with "OPENShift CONTAINER PLATFORM" and "Service Catalog" buttons.

The left side features a "Browse Catalog" section with a search bar and buttons for "Deploy Image", "Import YAML / JSON", and "Select from Project". It includes filters for "All", "Languages", "Databases", "Middleware", and "Other", with "All" currently selected. A "Filter" dropdown shows "40 Items". Below this, there are four categories of items:

- .NET**: .NET Core, .NET Core + PostgreSQL (Persistent), .NET Core Example, .NET Core Runtime Example. Icons include a 3scale-gateway icon (orange circle with arrows), an amp-apicast-wildcard-router icon (two overlapping squares), an amp-pvc icon (two overlapping squares), and an Apache HTTP Server icon (a feather).
- Apache HTTP Server**: (httpd) icon (feather), CakePHP + MySQL icon (php logo), CakePHP + MySQL (Ephemeral) icon (php logo), Dancer + MySQL icon (camel).
- PHP**: CakePHP + MySQL icon (php logo), CakePHP + MySQL (Ephemeral) icon (php logo).
- Apache HTTP Server**: (httpd) icon (feather).

The right side shows the "My Projects" section with a "Create Project" button. It lists 5 of 15 projects, each with a creation timestamp:

- default (created 24 minutes ago)
- ibm-cert-store (created 19 minutes ago)
- ibm-system (created 24 minutes ago)
- kube-proxy-and-dns (created 20 minutes ago)
- kube-public (created 24 minutes ago)



## **2. Create CP4D for the created OpenShift cluster**

---

# Create Cloud Pak for Data (1 of 5)

Catalog - IBM Cloud

https://cloud.ibm.com/catalog?search=label%3Acloud\_pak#software

IBM Cloud

Catalog Docs Support Manage 1639479 - ...

Catalog

Search the catalog...

Services Software (5)

All Categories (5) >

- Compute (1)
- Networking
- Storage
- AI
- Analytics (1)
- Databases (1)
- Developer Tools (2)
- Integration (1)
- Security and Identity
- Web and Application

Offering type

- Cloud Paks (5)
- Helm charts (67)
- Terraform (4)
- Starter kits (12)

Cloud Pak for Data  
IBM • Analytics • Databases  
Move from monolithic to microservices with DevOps methodologies and modernization toolkits

Cloud Paks • Red Hat OpenShift

Cloud Pak for Integration  
IBM • Integration  
Delivers application modernization and cloud-native application development for multicloud deployments.

Cloud Paks • Red Hat OpenShift

Cloud Pak for Multicloud Management  
IBM • Developer Tools  
Accelerate innovation with a next-generation hybrid multicloud platform.

Cloud Paks • Red Hat OpenShift

Feedback

https://cloud.ibm.com/catalog/content/ibm-cp-datacore-6825cc5d-dbfb-4ba2-ad98-690e6f221701-global

1. Click “Catalog”
2. Click “Software” tab
3. Find “Cloud Pak for Data” Tile and click it

# Create Cloud Pak for Data (2 of 5)

The screenshot shows the IBM Cloud Catalog interface for creating a Cloud Pak for Data. The URL is https://cloud.ibm.com/catalog/content/ibm-cp-datacore-6825cc5d-... . The page title is "Cloud Pak for Data".

**Summary:**

- Cloud Pak for Data
- Installation environment: Red Hat OpenShift
- Red Hat OpenShift cluster: mycluster (us-south)
- Project: zen
- Workspace: ibm-cp-datacore-03-29-2020
- Resource group: default

**IBM Passport Advantage:**

- Name: IBM Cloud Pak for Data Enterprise Edition per Virtual Processor Core License + SW
- Subscription & Support 12 Months
- Owner ID: IBMid-060000HVYV
- Part number: D1YGZLL
- Effective dates: 03/22/2020 - 12/31/9999

**Agreement:**

I have read and agree to the following license agreements:

- Cloud Pak for Data License
- Apache 2.0

**Preinstallation required:** You must run the preinstallation script on mycluster (us-south) or [refresh](#) to check the preinstallation script status.

**Configuration:**

- Red Hat OpenShift** (selected)
- View the existing installations**
- Red Hat OpenShift cluster:** mycluster (us-south)
- Project:** zen
- Configure your workspace**

1. Confirm your created managed OpenShift cluster is shown as “Red Hat OpenShift cluster”
2. Type “zen” in “Project”

# Create Cloud Pak for Data (3 of 5)

The screenshot shows the IBM Cloud interface for creating a Cloud Pak for Data workspace. The left sidebar has a 'Cloud Pak for Data - IBM Cloud X' tab. The main content area is titled 'Configure your workspace'.

**Name:** ibm-cp-datacore-03-29-2020

**Resource group:** default

**Tags:** Examples: env:dev, version-1

**Complete the preinstallation**

This task must be completed by a Red Hat OpenShift cluster administrator. Specifically, the administrator must have an [access](#) policy in IBM Cloud Identity and Access Management that has an Operator role or higher. The script makes the following changes to your Red Hat OpenShift cluster:

- Increases the size of the Docker registry volume to 200 GB.  
This change increases the cost of your Red Hat OpenShift cluster.

**Summary**

**Cloud Pak for Data**

- Installation environment: Red Hat OpenShift
- Red Hat OpenShift cluster: mycluster (us-south)
- Project: zen
- Workspace: ibm-cp-datacore-03-29-2020
- Resource group: default

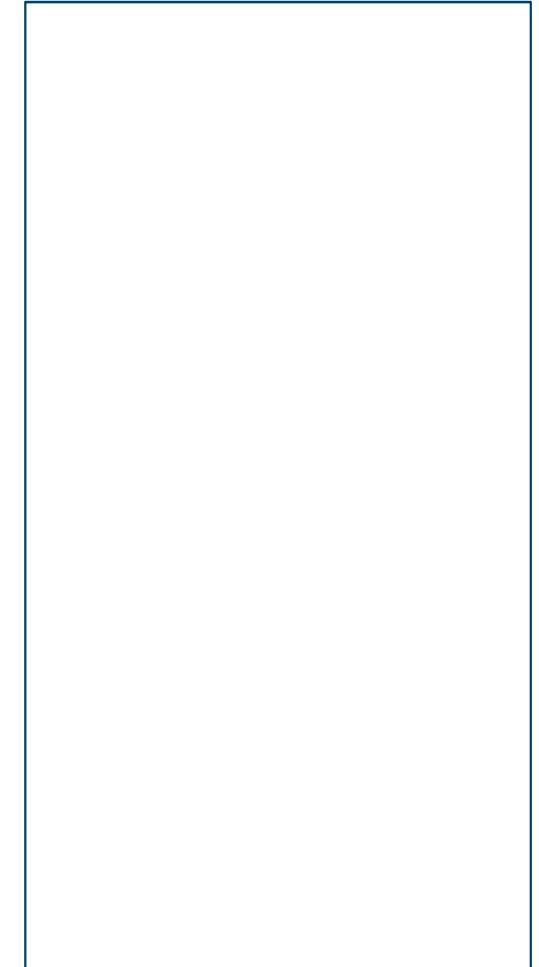
**IBM Passport Advantage**

- Name: IBM Cloud Pak for Data Enterprise Edition per Virtual Processor Core License + SW
- Subscription & Support 12 Months
- Owner ID: IBMid-060000HVYV
- Part number: D1YGZLL
- Effective dates: 03/22/2020 - 12/31/9999

I have read and agree to the following license agreements:

[Cloud Pak for Data License](#)  
[Apache 2.0](#)

**Preinstallation required:** You must run the preinstallation script on mycluster (us-south) or [refresh](#) to check the preinstallation script status.



# Create Cloud Pak for Data (4 of 5)

The screenshot shows the IBM Cloud Catalog interface for creating a Cloud Pak for Data. At the top, there is a message: "Preinstallation required: You must run the preinstallation script on mycluster (us-south) or [refresh](#) to check the preinstallation script status." Below this are two buttons: "Share link" and "Run script".

On the left, there is a section titled "Set the deployment values" with a sub-section "Parameters with default values". It lists four parameters:

Parameter	Description	Value
aiopenscale	Deploy Watson Open Scale. the default value is false.	false
dv	Deploy Data Virtualization. the default value is false.	false
wkc	Deploy Watson Knowledge Catalog. the default value is false.	false
wml	Deploy Watson Machine Learning. the default value is false.	false

At the bottom right of the catalog page, there is a "Feedback" button.

On the right side of the screen, there is a "Summary" panel for the Cloud Pak for Data instance. It includes the following details:

- Cloud Pak for Data**
- Installation environment: Red Hat OpenShift
- Red Hat OpenShift cluster: mycluster (us-south)
- Project: zen
- Workspace: ibm-cp-datacore-03-29-2020
- Resource group: default

Below the summary, there is a section for **IBM Passport Advantage** with the following information:

- Name: IBM Cloud Pak for Data Enterprise Edition per Virtual Processor Core License + SW
- Subscription & Support 12 Months
- Owner ID: IBMid-060000HVYV
- Part number: D1YGZLL
- Effective dates: 03/22/2020 - 12/31/9999

A checkbox is checked next to the statement: "I have read and agree to the following license agreements:" followed by links to "Cloud Pak for Data License" and "Apache 2.0".

At the bottom of the summary panel, there is another message: "Preinstallation required: You must run the preinstallation script on mycluster (us-south) or [refresh](#) to check the preinstallation script status."

1. Click “Run script” button to expand your created OpenShift cluster resources.

# Create Cloud Pak for Data (5 of 5)

Cloud Pak for Data - IBM Cloud X

Preinstallation complete: The script ran successfully on mycluster (us-south). View the [logs](#) to check the status of the script.

Share link Run script

Set the deployment values

Parameters with default values

A default value is set for each parameter. Review and accept the defaults, or you can update with customized values.

Parameter	Description	Value
aiopenscale	Deploy Watson Open Scale. the default value is false.	false
dv	Deploy Data Virtualization. the default value is false.	false
wkc	Deploy Watson Knowledge Catalog. the default value is false.	false
wml	Deploy Watson Machine Learning. the default value is false.	false

Summary

Cloud Pak for Data

Installation environment: Red Hat OpenShift  
Red Hat OpenShift cluster: mycluster (us-south)  
Project: zen  
Workspace: ibm-cp-datacore-03-29-2020  
Resource group: default

IBM Passport Advantage

Name: IBM Cloud Pak for Data Enterprise Edition per Virtual Processor Core License + SW  
Subscription & Support 12 Months  
Owner ID: IBMid-060000HVY  
Part number: D1YGZLL  
Effective dates: 03/22/2020 - 12/31/9999

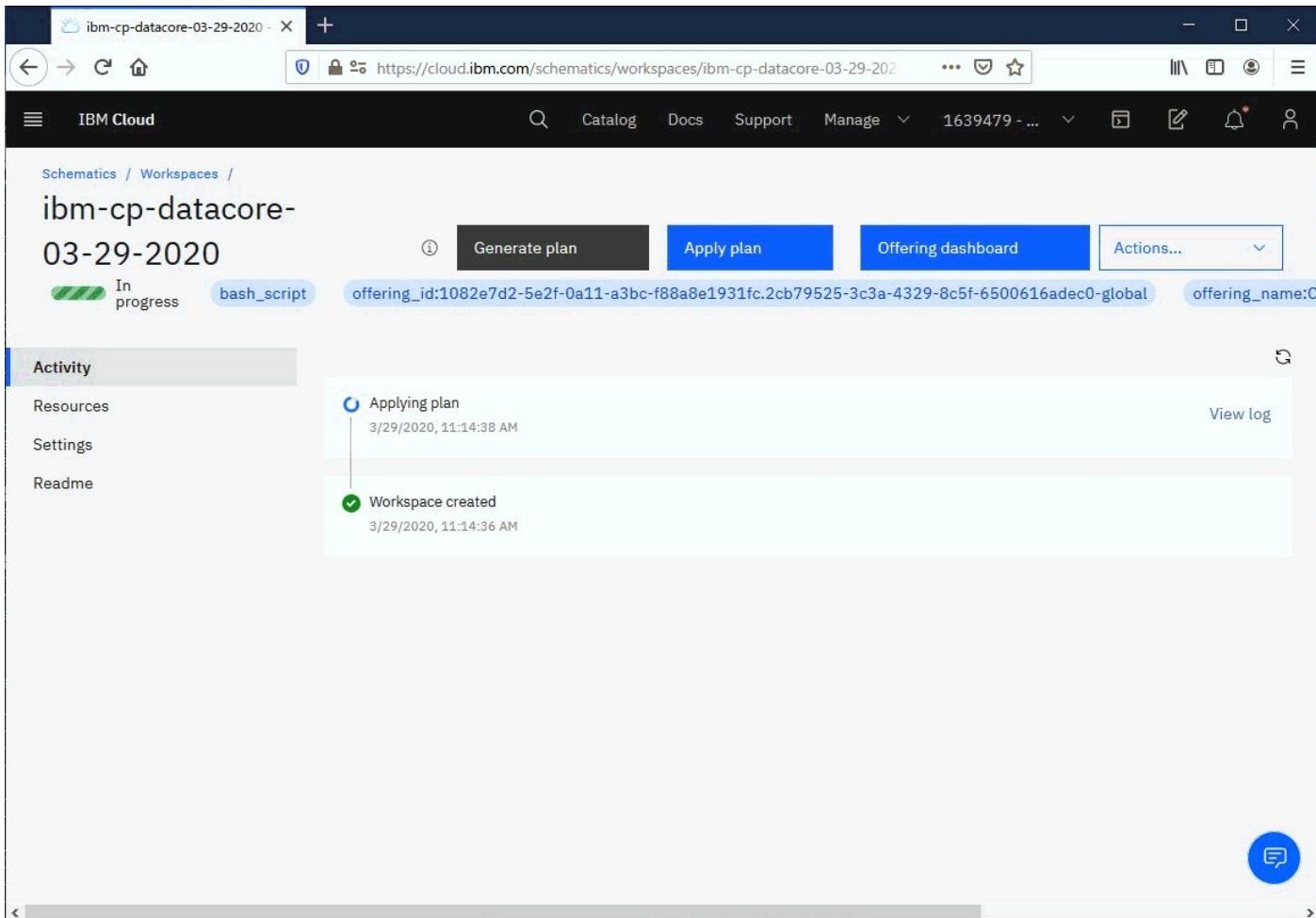
I have read and agree to the following license agreements:

Cloud Pak for Data License [\[link\]](#)  
Apache 2.0 [\[link\]](#)

Install

1. After completion of “Run Script”, you can see “Preinstallation complete”.
2. (Optional) Select additional CP4D services to be deployed with CP4D control plane. Leave “false” if only control plane (lite) is needed. (More cluster resources may be required for adding the services)
3. Check “I have read and agree to the....” check box.
4. Click “Install” blue button at bottom-right

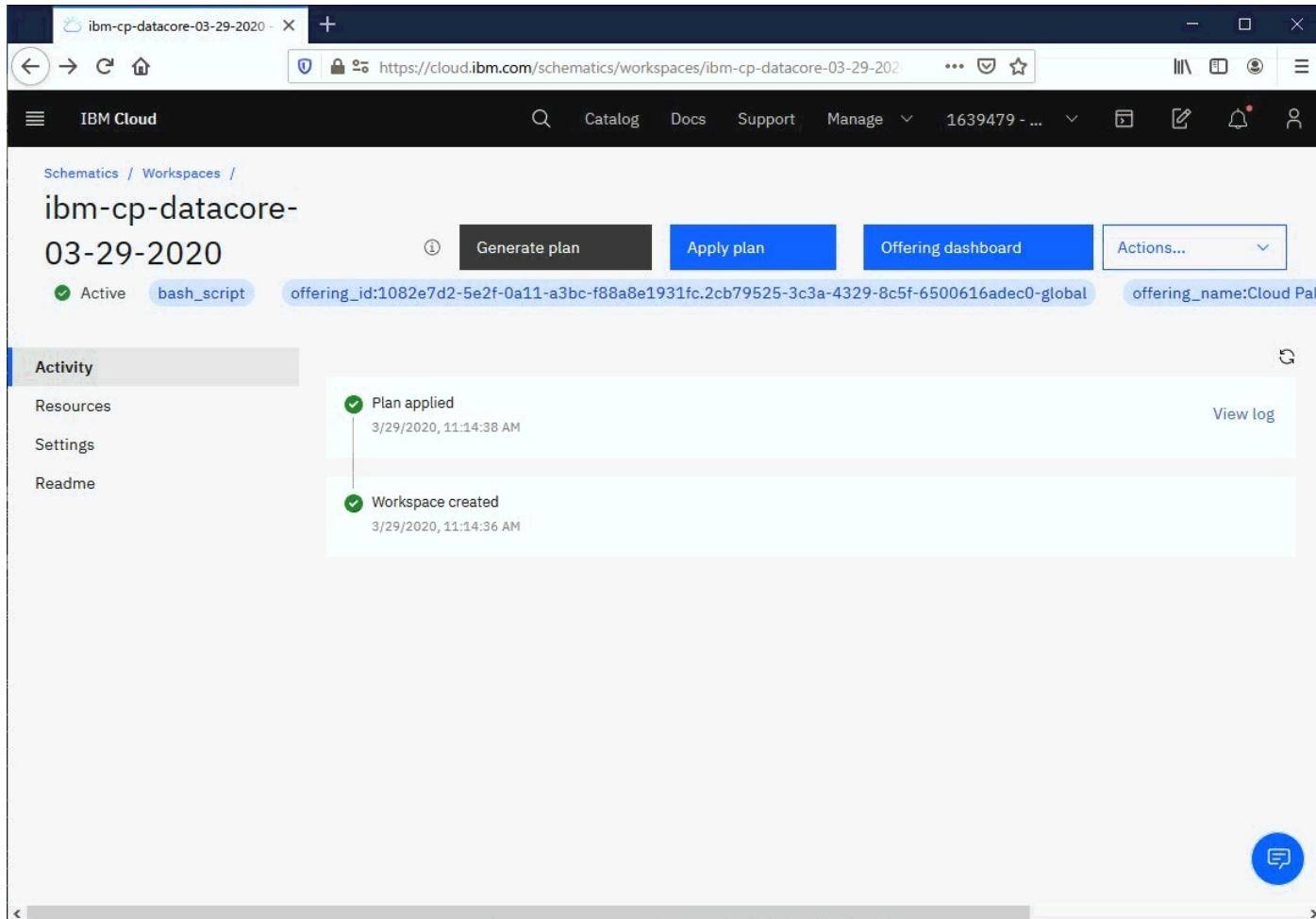
# CP4D deployment has started



1. Wait until “In progress” becomes “Active”.

NOTE: By some (unknown) reason, this deployment sometimes failed. If failed, try to re-run the plan by clicking “Apply plan” blue button. However, in some zone/region, CP4D deployment always failed. In the case, try to create OpenShift cluster in other region or zone. (or, open support ticket)

# CP4D is now ready for use

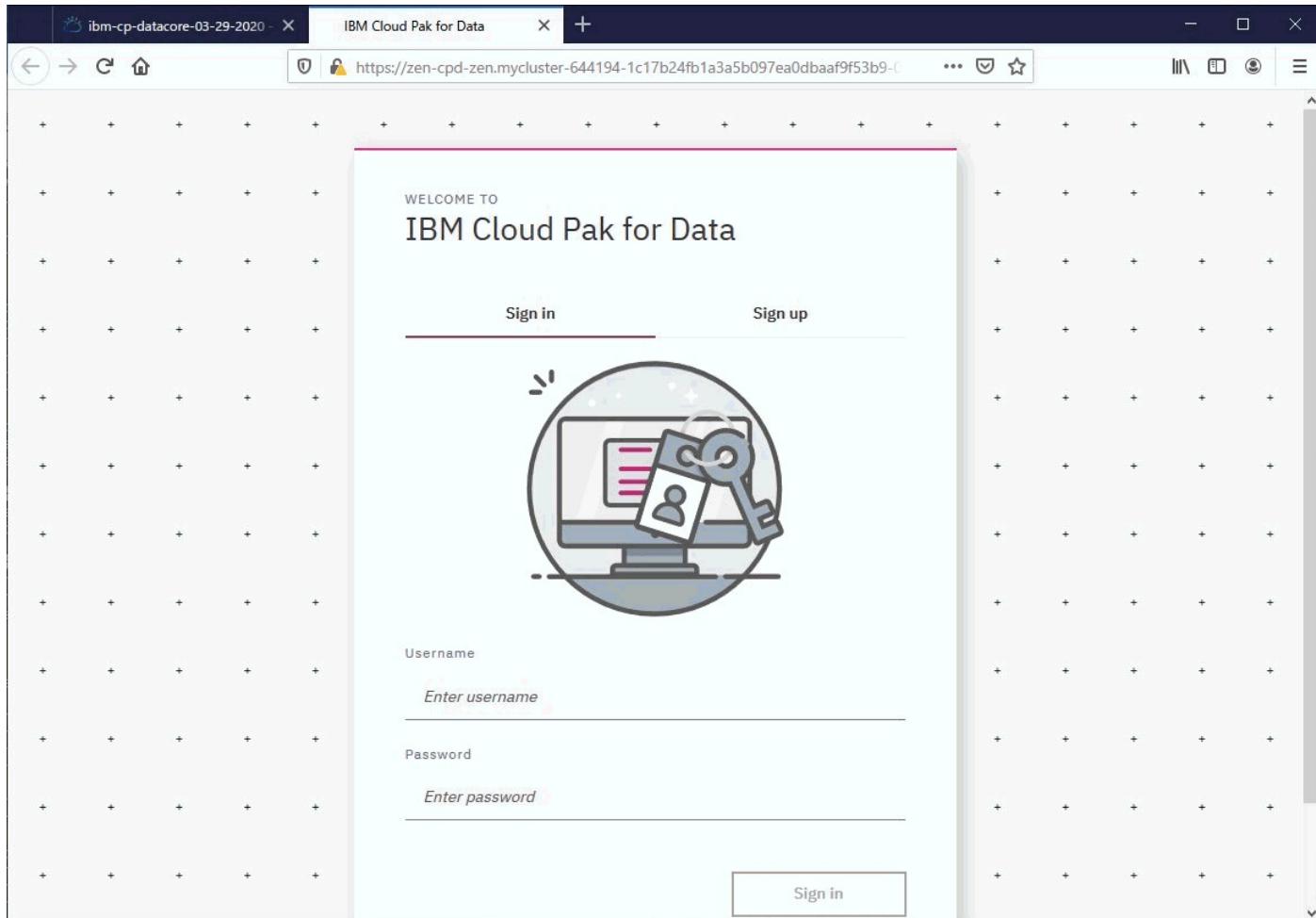


# Launch CP4D web console

The screenshot shows the IBM Cloud Schematics workspace interface for the workspace "ibm-cp-datacore-03-29-2020". The left sidebar has "Resources" selected. The main area displays the "mycluster" resource, which is an OpenShift cluster hosting workspace resources. Below it, the "Cloud Pak for Data" section shows a resource created by Schematics named "Cloud Pak for Data". The "URL" field contains a long, complex string starting with <https://zen-cpd-zen.mycluster-644194-1c17b24fb1a3a5b097ea0dbaaf9f53b9-0001.sjc03.containers.appdomain.cloud>.

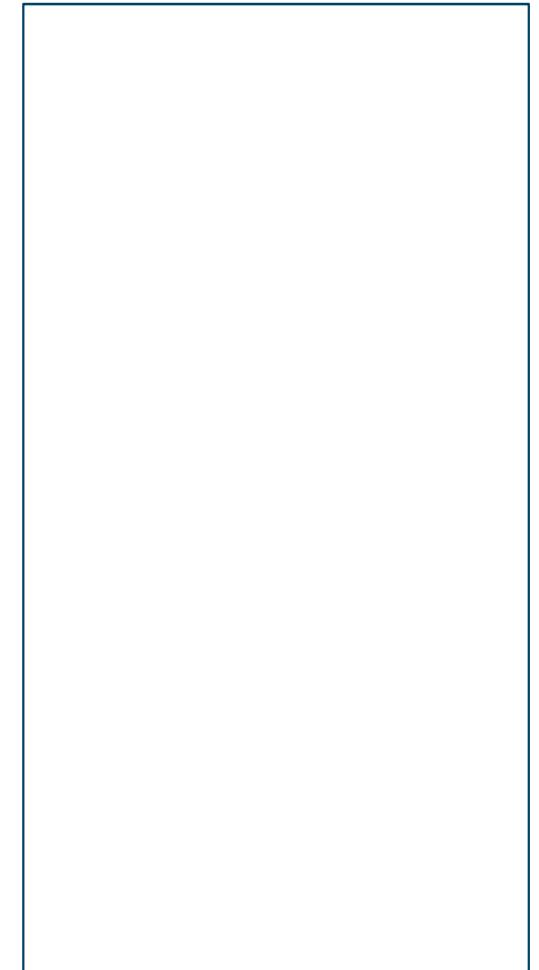
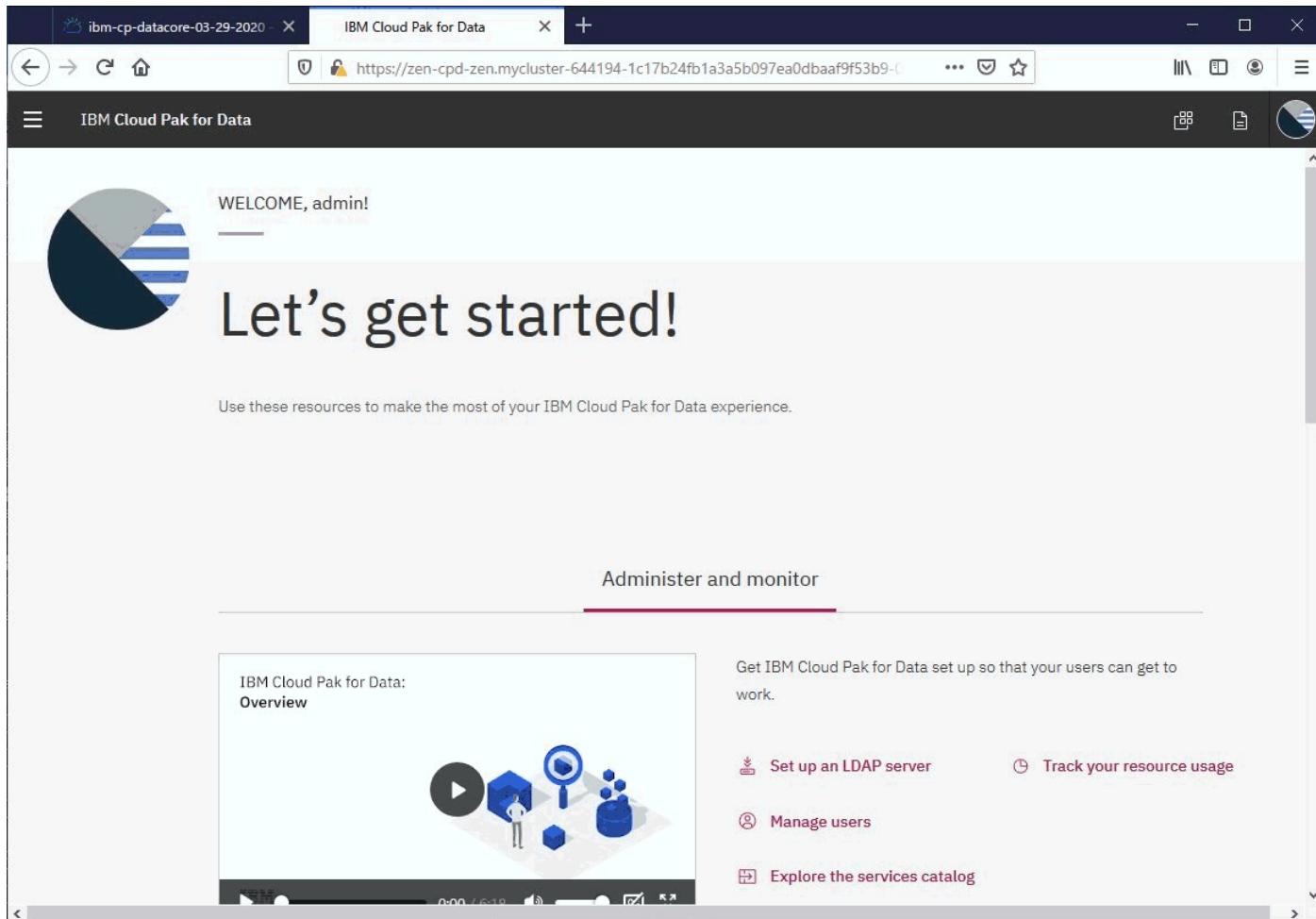
1. Click “Resources” in left menu
2. Click “URL” link in “Cloud Pak for Data”. Clicking the hyper link launches CP4D web console

# CP4D v2.5 web console has started



1. This is usual CP4D web console login dialog.  
"admin" for Username and  
"password" for  
"Password", then sign in

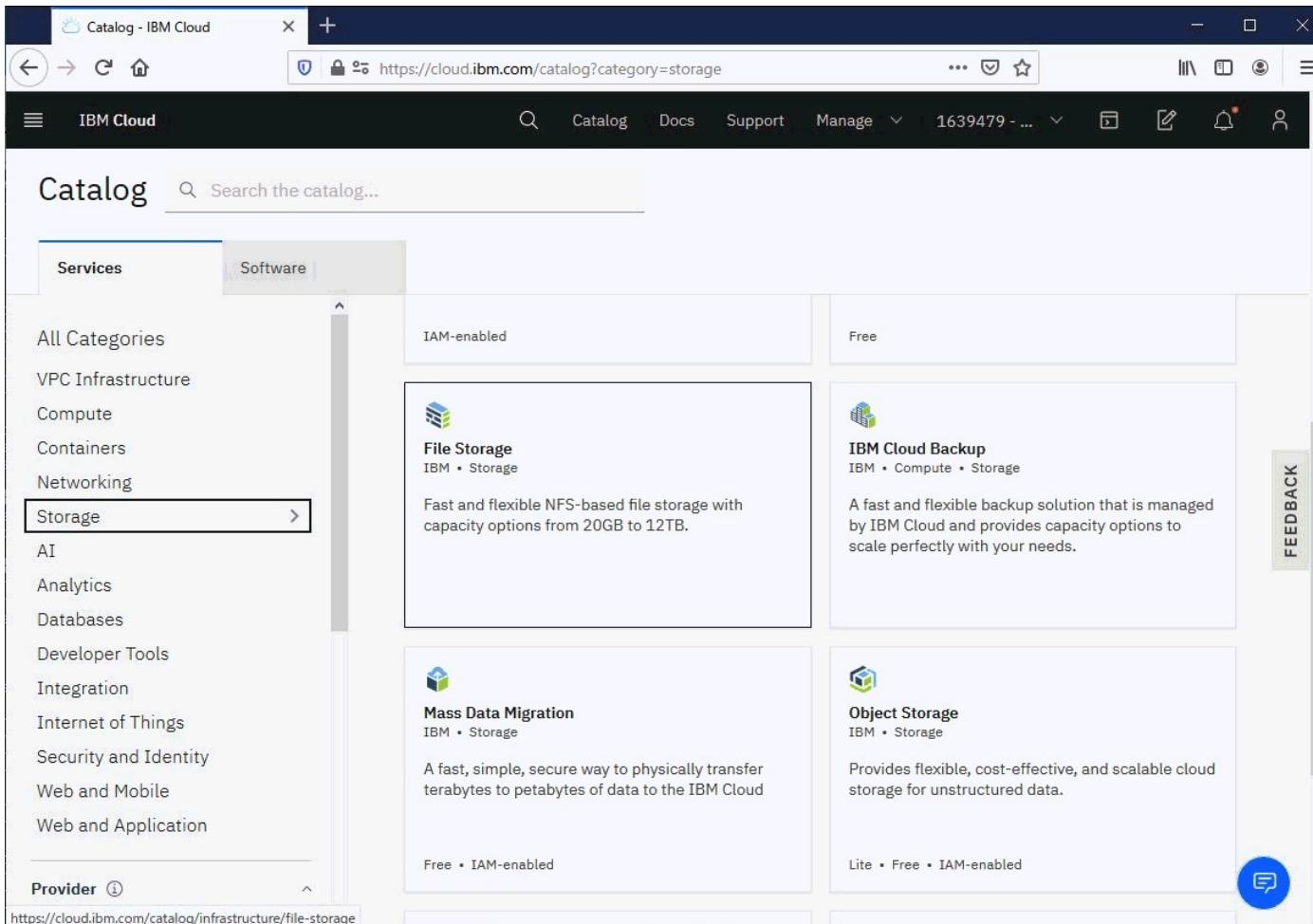
# CP4D v2.5 welcome page



### **3. Create File Storage for Watson Discovery**

---

# Create File Storage (1 of 5)

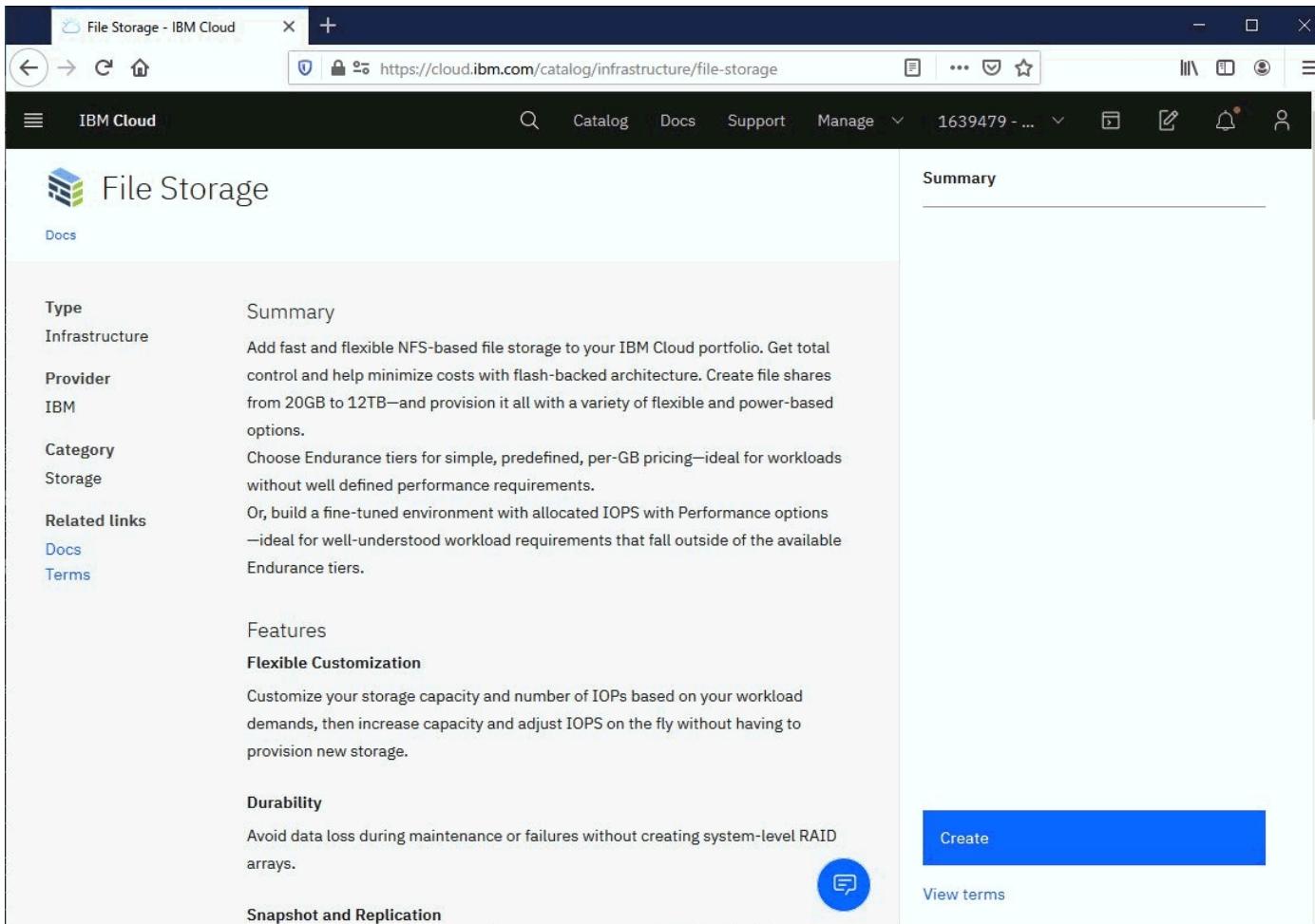


The screenshot shows the IBM Cloud Catalog interface. On the left, there is a sidebar with a navigation menu. The 'Storage' option under the 'Services' section is highlighted with a blue border. The main content area displays a grid of service tiles. The first tile in the top-left corner is 'File Storage' by IBM, Storage, described as a fast and flexible NFS-based file storage solution. To its right is 'IBM Cloud Backup' by IBM, Compute, Storage, which is a backup solution. Below these are 'Mass Data Migration' and 'Object Storage', both also by IBM, Storage. The URL in the browser bar is https://cloud.ibm.com/catalog?category=storage.

**NOTE:** There may be many ways to add storages for managed OpenShift cluster. This section introduces how you can add File Storage in IBM Cloud as one of example.

1. Click “Catalog” and click “Storage” in Services menu.
2. Find “File Storage” tile and click it

# Create File Storage (2 of 5)



The screenshot shows the IBM Cloud File Storage catalog page. On the left, there's a sidebar with navigation links: Docs, Type (Infrastructure), Provider (IBM), Category (Storage), Related links, Docs, and Terms. The main content area has two sections: 'Summary' and 'Features'. The 'Summary' section contains text about adding fast and flexible NFS-based file storage. The 'Features' section includes 'Flexible Customization' (describing capacity and IOPs customization) and 'Durability' (describing data loss avoidance). At the bottom right of the main content area is a large blue 'Create' button. Below it are 'View terms' and a feedback icon.

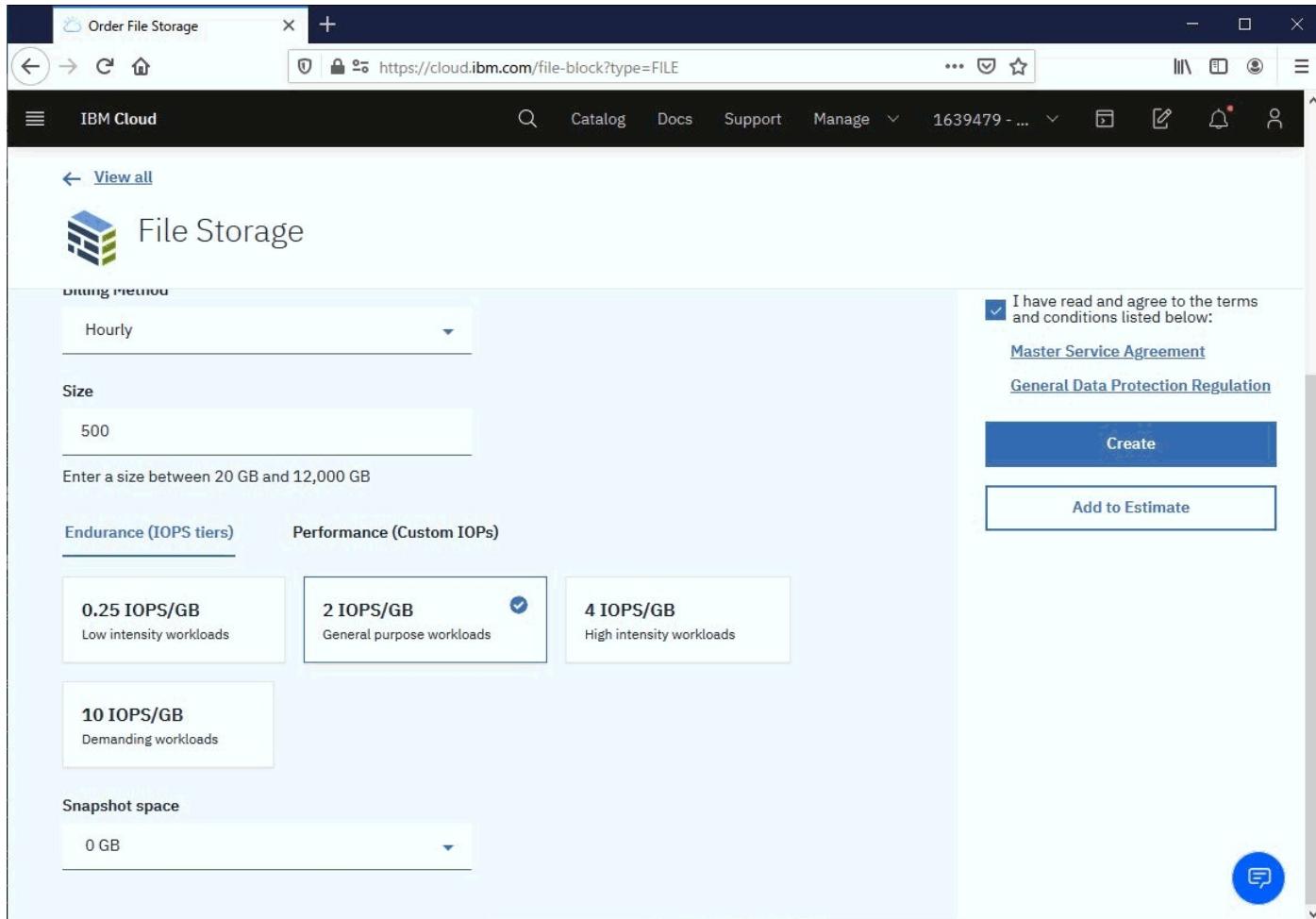
1. Click “Create” blue button at bottom-right

# Create File Storage (3 of 5)

The screenshot shows the 'Order File Storage' interface on the IBM Cloud platform. The 'Location' section is open, displaying six data center options: US West (SJC03 - San Jose \*), US South (DAL13 - Dallas \*), US East (WDC07 - Washington, DC \*), South America (SAO01 - Sao Paulo \*), Europe (LON06 - London \*), and Asia-Pacific (SYD04 - Sydney \*). The 'US West' option is selected. The 'Order Summary' section shows a 'File Storage Volume' of 500 GB storage space at \$0.065 per hour, with 2 IOPS/GB and 0 GB snapshot space. The 'Total due per hour' is listed as \$0.065 estimated. Below the summary, there's a link to 'Apply promo code'. A note states '\* Price does not include any tax'. Under 'Billing Method', 'Hourly' is selected. In the 'Size' section, '500' is entered, with a note: 'Enter a size between 20 GB and 12,000 GB'. At the bottom, there are links for 'Master Service Agreement' and 'General Data Protection Regulation', and two buttons: 'Create' and 'Add to Estimate'.

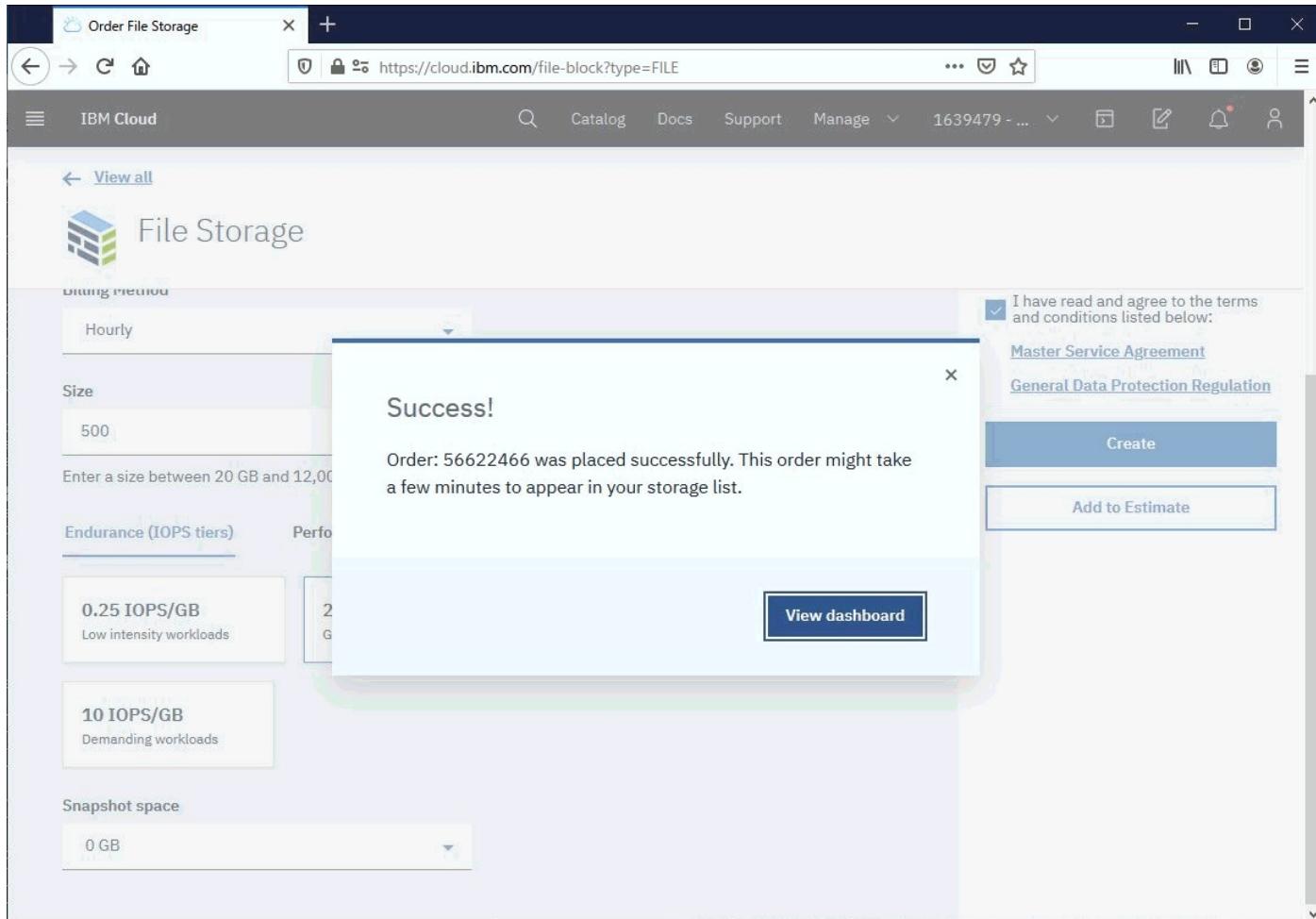
1. Select “Location”. Highly recommend creating File Storage at the same cluster zone where OpenShift for CP4D is created.
2. (Optional) Select “Billing Method”.
3. Set “Size”.

# Create File Storage (4 of 5)



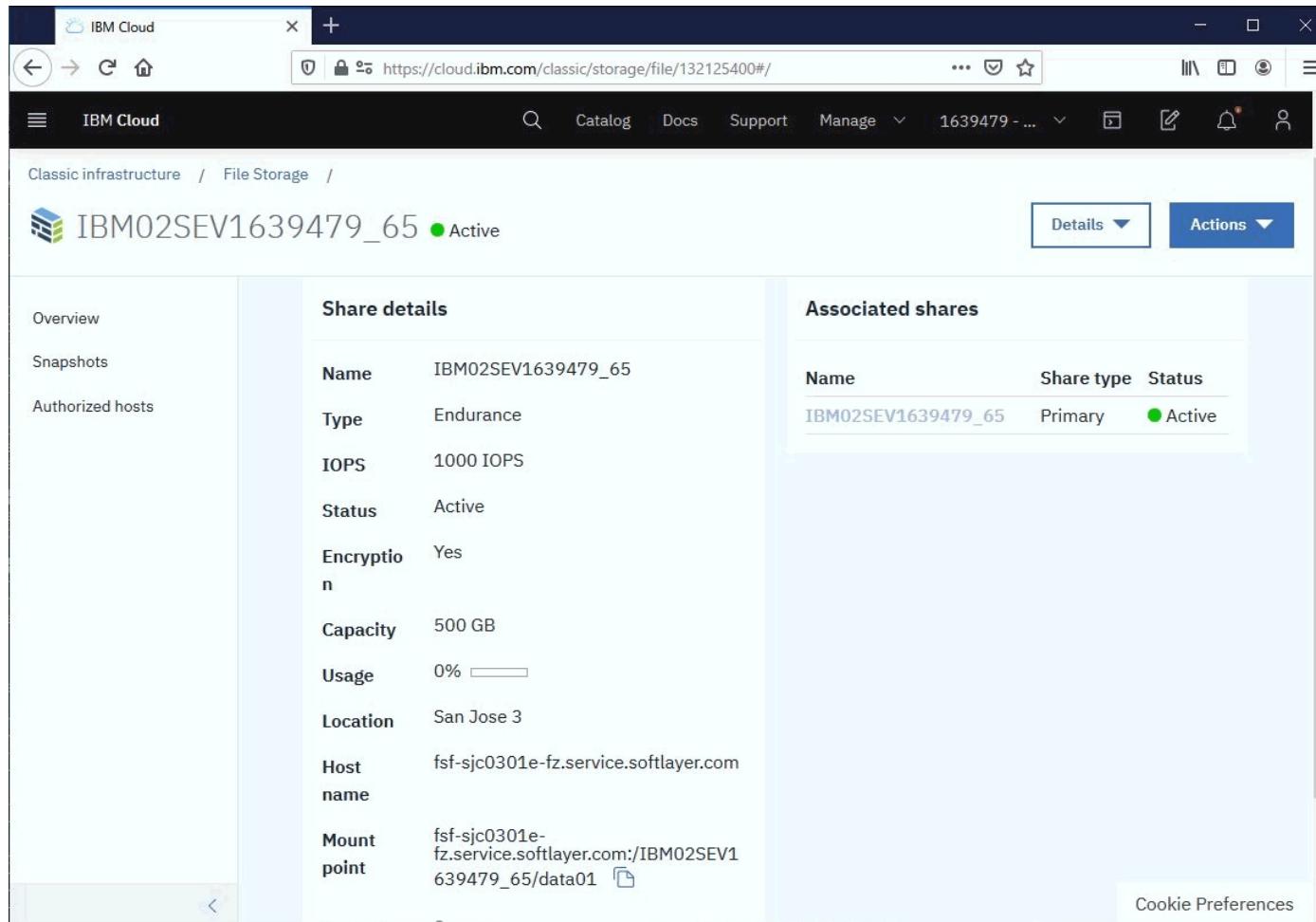
1. (Optional) Select “Performance (Custom IOPs)” if you need better performance. (2 IOPS/GB is default.)
2. Check “I have read and agree to the ....” check box.
3. Click “Create” blue button

# Create File Storage (5 of 5)



1. Order is placed. (But, the storage is not ready yet)
2. Click “View dashboard” button.

# Authorize hosts for File Storage (1 of 3)



The screenshot shows the IBM Cloud File Storage dashboard. On the left, there's a sidebar with 'Overview', 'Snapshots', and 'Authorized hosts'. The main area displays 'Share details' for a share named 'IBM02SEV1639479\_65'. The share has the following details:

Name	IBM02SEV1639479_65
Type	Endurance
IOPS	1000 IOPS
Status	Active
Encryption	Yes
Capacity	500 GB
Usage	0% <span style="width: 0%;"> </span>
Location	San Jose 3
Host name	ssf-sjc0301e-fz.service.softlayer.com
Mount point	ssf-sjc0301e-fz.service.softlayer.com:/IBM02SEV1639479_65/data01 <a href="#">File</a>

On the right, under 'Associated shares', there is one entry:

Name	Share type	Status
IBM02SEV1639479_65	Primary	Active

1. Select the newly created File Storage in File Storage dashboard. (You may need to wait until the File Storage becomes Active after order.)

# Authorize hosts for File Storage (2 of 3)

The screenshot shows the 'Authorize hosts' dialog box from the IBM Cloud classic infrastructure File Storage interface. The dialog has two main sections: 'Available Hosts for IBM02SEV1639479\_65' and 'Selected hosts'. In the 'Available Hosts' section, there is a note: 'Select a host type and then choose a host from the dropdown that is available for the volume.' Below this, there are three radio buttons: 'Devices' (selected), 'Subnets', and 'IP Address'. A 'Device Type' dropdown is set to 'Virtual Server'. Under 'Virtual Guest', there is a list with three items: 'kube-bpvl27d0hc2lsht0ov0-mycluster-default-00000140.iks.ibm', 'kube-bpvl27d0hc2lsht0ov0-mycluster-default-000003df.iks.ibm', and 'kube-bpvl27d0hc2lsht0ov0-mycluster-default-00000291.iks.ibm'. A 'Remove Selected' button is located at the bottom of this list. At the very bottom of the dialog are 'Close' and 'Save' buttons.

1. Click “Actions” pull down menu at top-right, and select “Authorize hosts”
2. Select “Devices” radio button.
3. Select “Virtual Server” in “Device Type” list box
4. Select servers in created managed OpenShift cluster in “Virtual Guest” list box.  
All servers (Devices) of your managed OpenShift have to be authorized.

# Authorize hosts for File Storage (3 of 3)

The screenshot shows the 'Authorize hosts' dialog box from the IBM Cloud interface. On the left, the sidebar shows 'IBM02SE' selected under 'Authorized hosts'. The main area has a title 'Authorize hosts' and two sections: 'Available Hosts for IBM02SEV1639479\_65' and 'Selected hosts'. Under 'Available Hosts', there are radio buttons for 'Devices', 'Subnets', and 'IP Address', with 'Devices' selected. A dropdown menu for 'Device Type' is set to 'Virtual Server'. A dropdown for 'Virtual Guest' is labeled 'Click to Select'. The 'Selected hosts' section contains a list of host names: 'kube-bpvvl27d0hc2lsht0ov0-mycluster-default-i...', 'kube-bpvvl27d0hc2lsht0ov0-mycluster-default-i...', and 'kube-bpvvl27d0hc2lsht0ov0-mycluster-default-i...'. Below this list is a red-bordered button labeled 'Remove Selected'. At the bottom right of the dialog are 'Close' and 'Save' buttons.

1. Push “Save” blue button at bottom-right

## 4. Gather Parameter values

---

# Required Parameters

---

- CLUSTER\_HOST
- DOCKER\_REGISTRY
- NFS\_HOST & NFS\_PATH
- OC\_LOGIN

# Parameter: CLUSTER\_HOST

The screenshot shows the IBM Cloud Kubernetes cluster overview for a cluster named 'mycluster'. The 'Overview' tab is selected in the left sidebar. Key details shown include:

- Cluster ID:** bpvvl27d0hc2lsh0ov0
- Master status:** Ready
- Version:** 3.11.170\_1544
- Zones:** sjc03
- Created:** 3/29/2020, 10:31 AM
- Ingress subdomain:** mycluster-644194-1c17b24fb1  
a3a5b097ea0dbaa9f53b9-000  
0.sjc03.containers.appdomain.  
cloud
- Resource group:** default
- Key management service:** Enable
- Image pull secrets:** Enabled

On the right side, there are two circular dashboards: one showing 'Worker Nodes' at 100% Normal and another showing 'Cluster Insights' at 2.2% Total CPU Usage.

1. Value for “Ingress subdomain” is copied as CLUSTER\_HOST parameter value

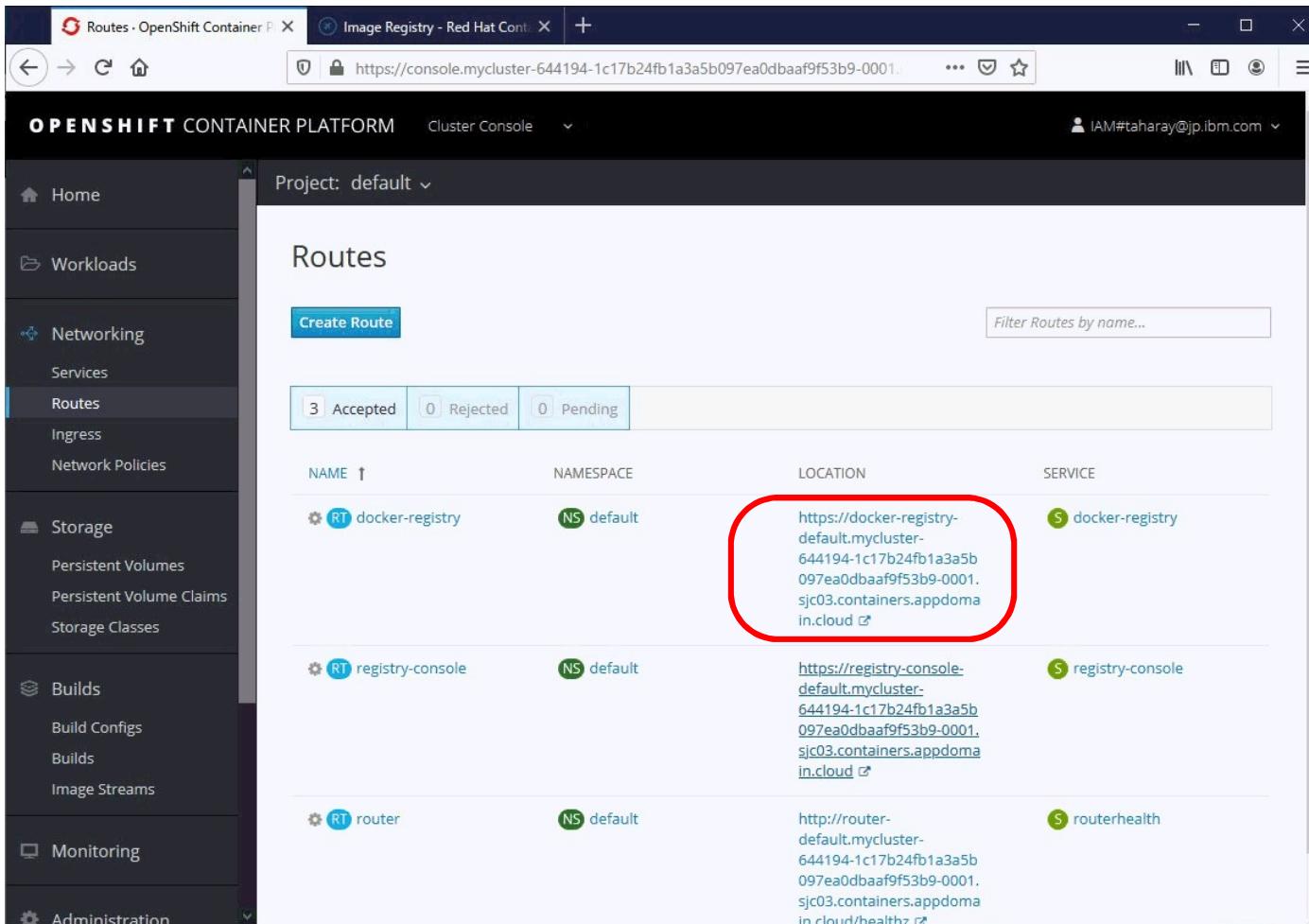
# Parameter: DOCKER\_REGISTRY (1 of 2)

The screenshot shows the OpenShift Service Catalog interface. At the top, there's a navigation bar with tabs for 'mycluster Cluster - IBM Cloud' and 'OpenShift Web Console'. The URL is https://c103-e.us-south.containers.cloud.ibm.com:30063/console/catalog. Below the navigation is a search bar labeled 'Search Catalog'. The main area is titled 'Browse Catalog' with a 'My Projects' sidebar on the right showing 5 of 15 projects: default, ibm-cert-store, ibm-system, kube-proxy-and-dns, and kube-public, all created within the last 24 minutes. The catalog lists various service offerings under categories like '.NET', 'Apache HTTP Server', 'PHP', etc. Each listing includes a small icon, the service name, and a brief description.

Category	Service Name	Description
.NET	.NET Core	
	.NET Core + PostgreSQL (Persistent)	
	.NET Core Example	
	.NET Core Runtime Example	
Apache HTTP Server	3scale-gateway	
	amp-apicast-wildcard-router	
	amp-pvc	
	Apache HTTP Server (httpd)	
PHP	Apache HTTP Server (httpd)	
	CakePHP + MySQL	
	CakePHP + MySQL (Ephemeral)	
	Dancer + MySQL	

1. Open OpenShift web console
2. Click “Service Catalog” as pull down menu, and select “Cluster Console”. UI navigates you with some additional logins (e.g. “ocadmin/ocadmin”), and Cluster Console web page is shown (see next slide)

# Parameter: DOCKER\_REGISTRY (2 of 2)

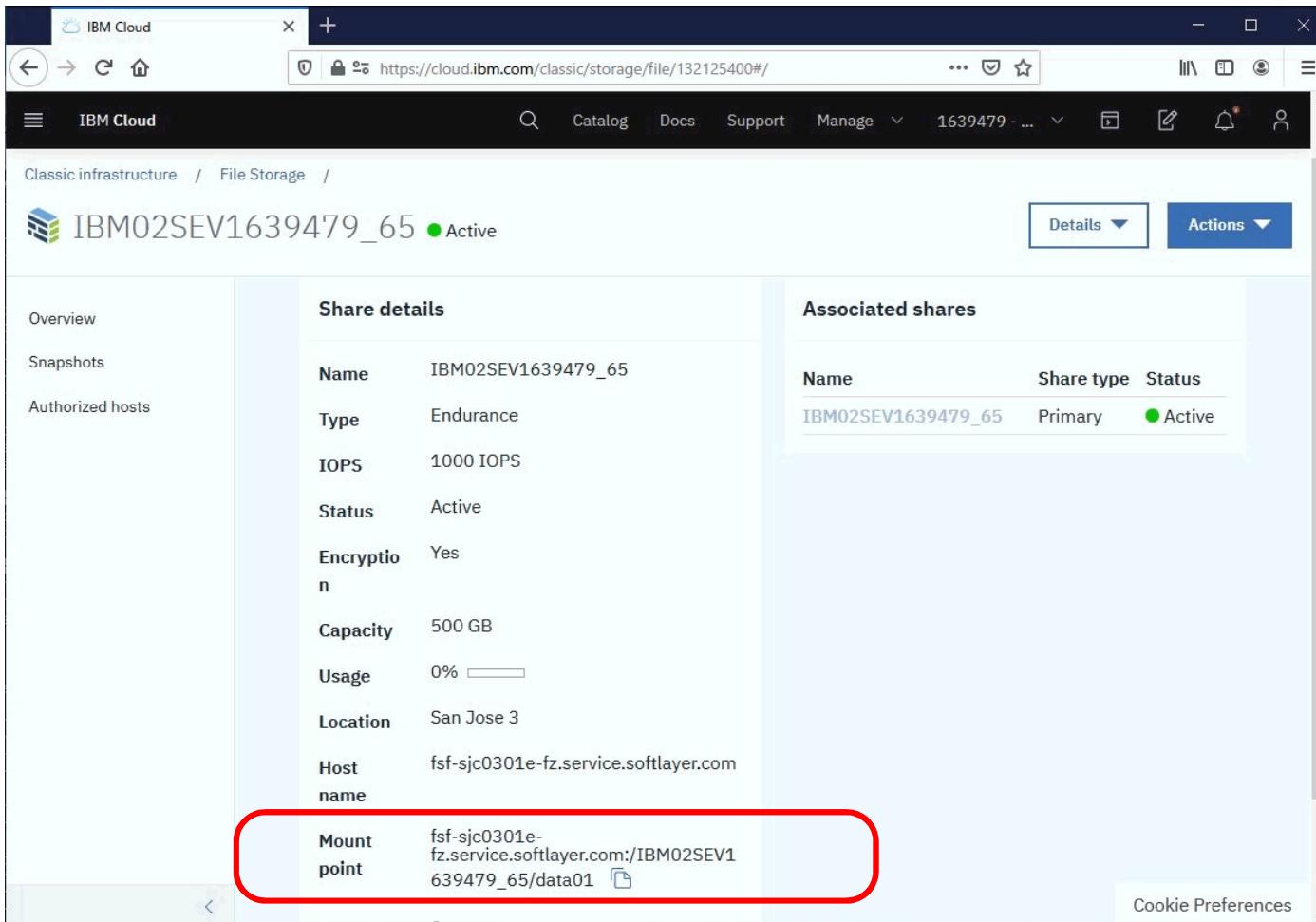


The screenshot shows the OpenShift Container Platform Cluster Console with the 'Networking' menu open. The 'Routes' section is selected. A table lists three routes: 'docker-registry', 'registry-console', and 'router'. The 'docker-registry' row is highlighted with a red circle around its 'LOCATION' column, which contains the URL: <https://docker-registry-default.mycluster-644194-1c17b24fb1a3a5b097ea0dbaa9f53b9-0001.sjc03.containers.appdomain.cloud>.

NAME	NAMESPACE	LOCATION	SERVICE
docker-registry	default	<a href="https://docker-registry-default.mycluster-644194-1c17b24fb1a3a5b097ea0dbaa9f53b9-0001.sjc03.containers.appdomain.cloud">https://docker-registry-default.mycluster-644194-1c17b24fb1a3a5b097ea0dbaa9f53b9-0001.sjc03.containers.appdomain.cloud</a>	docker-registry
registry-console	default	<a href="https://registry-console-default.mycluster-644194-1c17b24fb1a3a5b097ea0dbaa9f53b9-0001.sjc03.containers.appdomain.cloud">https://registry-console-default.mycluster-644194-1c17b24fb1a3a5b097ea0dbaa9f53b9-0001.sjc03.containers.appdomain.cloud</a>	registry-console
router	default	<a href="http://router-default.mycluster-644194-1c17b24fb1a3a5b097ea0dbaa9f53b9-0001.sjc03.containers.appdomain.cloud/healthz">http://router-default.mycluster-644194-1c17b24fb1a3a5b097ea0dbaa9f53b9-0001.sjc03.containers.appdomain.cloud/healthz</a>	routerhealth

1. In Cluster console, select “Routes” in “Networking” in left menu.
2. Find “docker-registry” as route.
3. Get the value of “LOCATION” for the “docker-registry” route. Remove [“https://”](https://) from the value and set the remained string as DOCKER\_REGISTRY

# Parameter: NFS\_HOST and NFS\_PATH



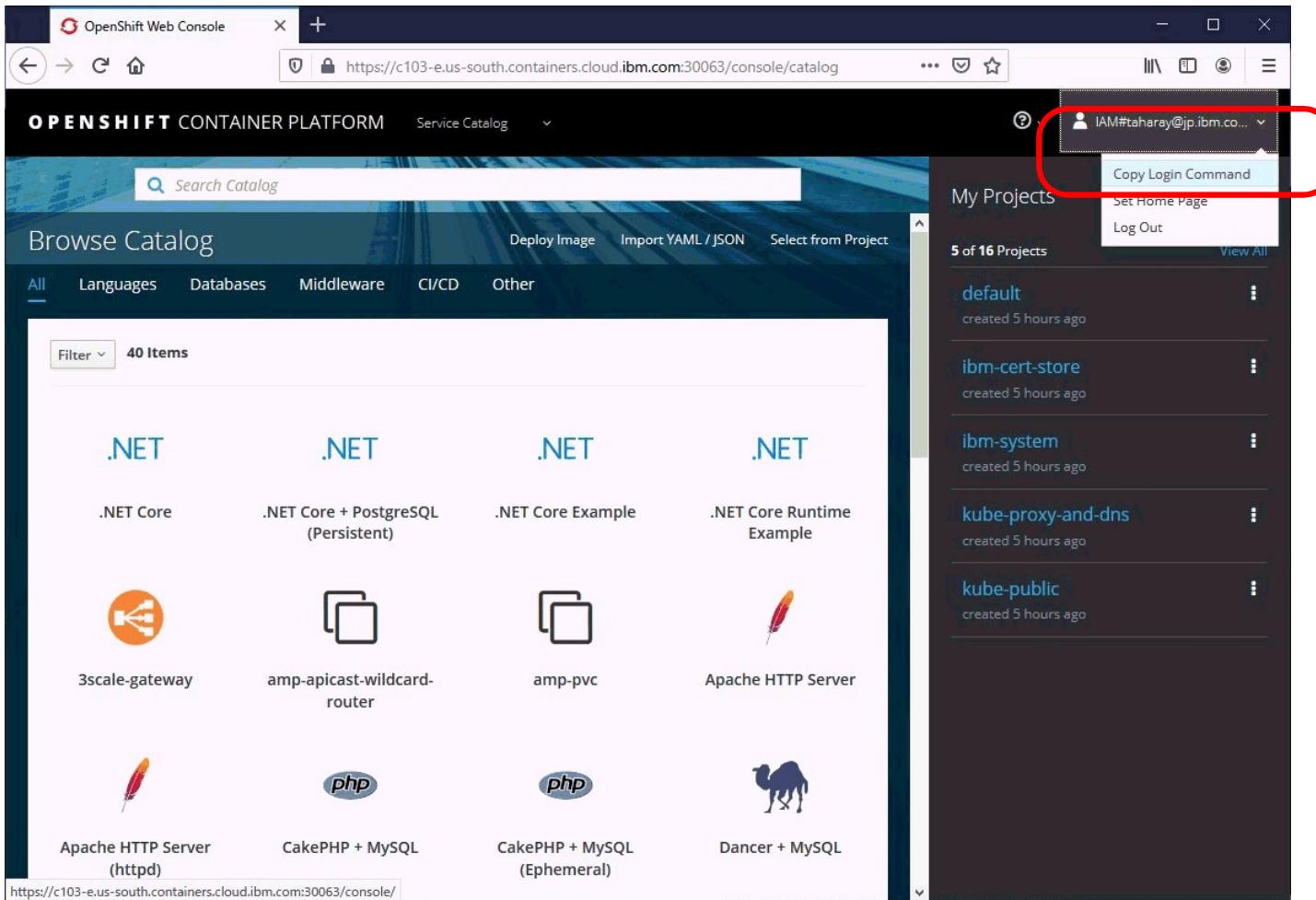
The screenshot shows the IBM Cloud File Storage Overview page for a specific file share. The share is named "IBM02SEV1639479\_65" and is marked as "Active". The "Share details" section provides the following information:

Name	IBM02SEV1639479_65
Type	Endurance
IOPS	1000 IOPS
Status	Active
Encryption	Yes
Capacity	500 GB
Usage	0% <span style="width: 0%;"> </span>
Location	San Jose 3
Host name	ssf-sjc0301e-fz.service.softlayer.com
Mount point	ssf-sjc0301e-fz.service.softlayer.com:/IBM02SEV1639479_65/data01 

The "Associated shares" section lists one primary share with the same name and status.

1. Find your create File Storage and open "Overview"
2. Split "Mount point" value for hostname and path.  
In the left example, "ssf-sjc...softlayer.com" should be set to **NFS\_HOST**, and "/IBM02SEV1639479\_65/data01" should be set to **NFS\_PATH**

# Parameter: OC\_LOGIN



1. Click your account at top-right, and pull down menu
2. Select “Copy Login Command”
3. Paste the copied login command into OC\_LOGIN (OC\_LOGIN value starts from “oc login https://” with hostname and login token for your IBM Cloud account.

## **5. Create NFS storage class and provisioner**

---

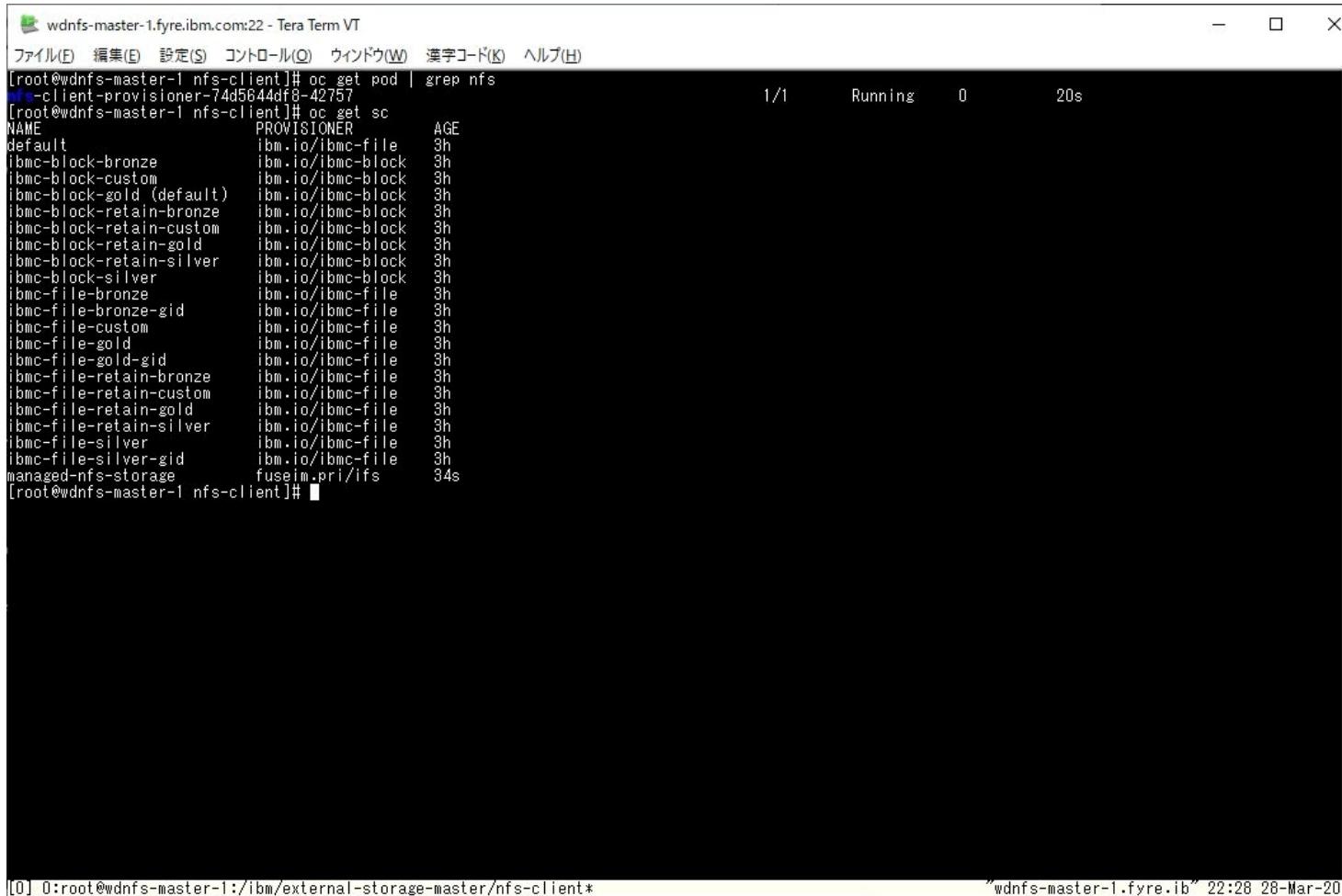
# Create NFS storage class and provisioner

---

Reference: <https://medium.com/faun/openshift-dynamic-nfs-persistent-volume-using-nfs-client-provisioner-fcbb8c9344e>

```
# You need to login to some Linux server where you can run oc cli and docker command remotely
export NFSNS=zen
export NFSHOST=${NFS_HOST}
export NFSPATH=${NFS_PATH}
Export OCLOGIN=${OC_LOGIN}
mkdir ~/wd212
cd ~/wd212
curl -L -o kubernetes-incubator.zip https://github.com/kubernetes-incubator/external-storage/archive/master.zip
unzip kubernetes-incubator.zip
cd ~/wd212/external-storage-master/nfs-client
sed -i"" "s/namespace:.*/namespace: $NFSNS/g" ./deploy/rbac.yaml
sed -i"" "s/namespace:.*/namespace: $NFSNS/g" ./deploy/deployment.yaml
sed -i"" "s/value: 10.10.10.60/value: $NFSHOST/g" ./deploy/deployment.yaml
sed -i"" "s/server: 10.10.10.60/server: $NFSHOST/g" ./deploy/deployment.yaml
sed -i"" "s|value: /ifs/kubernetes|value: $NFSPATH|g" ./deploy/deployment.yaml
sed -i"" "s|path: /ifs/kubernetes|path: $NFSPATH|g" ./deploy/deployment.yaml
${OCLOGIN}
oc new-project $NFSNS
oc create -f ./deploy/rbac.yaml
oc adm policy add-scc-to-user hostmount-anyuid system:serviceaccount:$NFSNS:nfs-client-provisioner
oc create -f ./deploy/class.yaml
oc create -f ./deploy/deployment.yaml
```

# Confirm NFS storage class and provisioner



The screenshot shows a terminal window titled "wdnfs-master-1.fyre.ibm.com:22 - Tera Term VT". The window displays two command outputs from a root shell:

```
[root@wdnfs-master-1 nfs-client]# oc get pod | grep nfs
nfs-client-provisioner-74d5b44df8-42757
[root@wdnfs-master-1 nfs-client]# oc get sc
NAME           PROVISIONER          AGE
default        ibm.io/ibmc-file   3h
ibmc-block-bronze ibm.io/ibmc-block 3h
ibmc-block-custom ibm.io/ibmc-block 3h
ibmc-block-gold (default) ibm.io/ibmc-block 3h
ibmc-block-retain-bronze ibm.io/ibmc-block 3h
ibmc-block-retain-custom ibm.io/ibmc-block 3h
ibmc-block-retain-gold ibm.io/ibmc-block 3h
ibmc-block-retain-silver ibm.io/ibmc-block 3h
ibmc-block-silver ibm.io/ibmc-block 3h
ibmc-file-bronze ibm.io/ibmc-file 3h
ibmc-file-bronze-gid ibm.io/ibmc-file 3h
ibmc-file-custom ibm.io/ibmc-file 3h
ibmc-file-gold ibm.io/ibmc-file 3h
ibmc-file-gold-gid ibm.io/ibmc-file 3h
ibmc-file-retain-bronze ibm.io/ibmc-file 3h
ibmc-file-retain-custom ibm.io/ibmc-file 3h
ibmc-file-retain-gold ibm.io/ibmc-file 3h
ibmc-file-retain-silver ibm.io/ibmc-file 3h
ibmc-file-silver ibm.io/ibmc-file 3h
ibmc-file-silver-gid ibm.io/ibmc-file 3h
managed-nfs-storage fuseim.pri/ifs 34s
[root@wdnfs-master-1 nfs-client]#
```

The terminal window has a dark background and light-colored text. It includes standard window controls (minimize, maximize, close) at the top right. At the bottom, it shows the session identifier [0], the current directory (0:/root@wdnfs-master-1:/ibm/external-storage-master/nfs-client\*), and the host and timestamp ("wdnfs-master-1.fyre.ib" 22:28 28-Mar-20).

## **6. Patch Watson Discovery installation charts**

---

# Apply patch to Watson Discovery helm charts

---

Prerequisite: Obtain Watson Discovery v2.1.2 installation image and patch zip file. Copy these files under ~/wd212

```
# You need to login to some Linux server where you can run oc cli and docker command remotely

cd ~/wd212
tar xvf <install image of Watson Discovery v2.1.2>
cd ~/wd212/lib/charts
tar zxvf ibm-discovery-bedrock-1.0.0.tgz
tar zxvf ibm-discovery-core-1.0.0.tgz
tar zxvf ibm-discovery-substrate-1.0.0.tgz
unzip -o ~/wd212/patch.managed_cp4d.wd212.zip
rm -f ibm-discovery-bedrock-1.0.0.tgz
rm -f ibm-discovery-core-1.0.0.tgz
rm -f ibm-discovery-substrate-1.0.0.tgz
tar zcvf ibm-discovery-bedrock-1.0.0.tgz ibm-discovery-bedrock/*
tar zcvf ibm-discovery-core-1.0.0.tgz ibm-discovery-core/*
tar zcvf ibm-discovery-substrate-1.0.0.tgz ibm-discovery-substrate/*
```

## **7. Deploy the patched Watson Discovery**

---

# Deploy the patched Watson Discovery

---

Installation procedure is exactly same as deployment into On-premise servers but all operations have to be done remotely

```
# You need to login to some Linux server where you can run oc cli and docker command remotely.

cd ~/wd212/bin
export IBMC_ACCOUNT=<your login account id> # format is like "IAM#xxxxx@ibm.com"
export WDSC=managed-nfs-storage
export DOCKERREG=${DOCKER_REGISTRY}
export CLUSTERHOST=${CLUSTER_HOST}
export OCLOGIN=${OC_LOGIN}

${OCLOGIN}
export PRIV_IP=`oc get node | grep compute | awk '{ print $1 }' | head -n 1`

docker login $(oc get routes docker-registry -n default -o template={{.spec.host}})/zen:5000 -u ${IBMC_ACCOUNT} -p $(oc whoami -t)

./loadImages.sh -r ${DOCKERREG}

./installDiscovery.sh -a ${CLUSTERHOST} -I ${PRIV_IP} -n zen -s ${WDSC} -S ${WDSC} --cluster-pull-prefix docker-registry.default.svc:5000/zen
```

## 8. Launch Watson Discovery

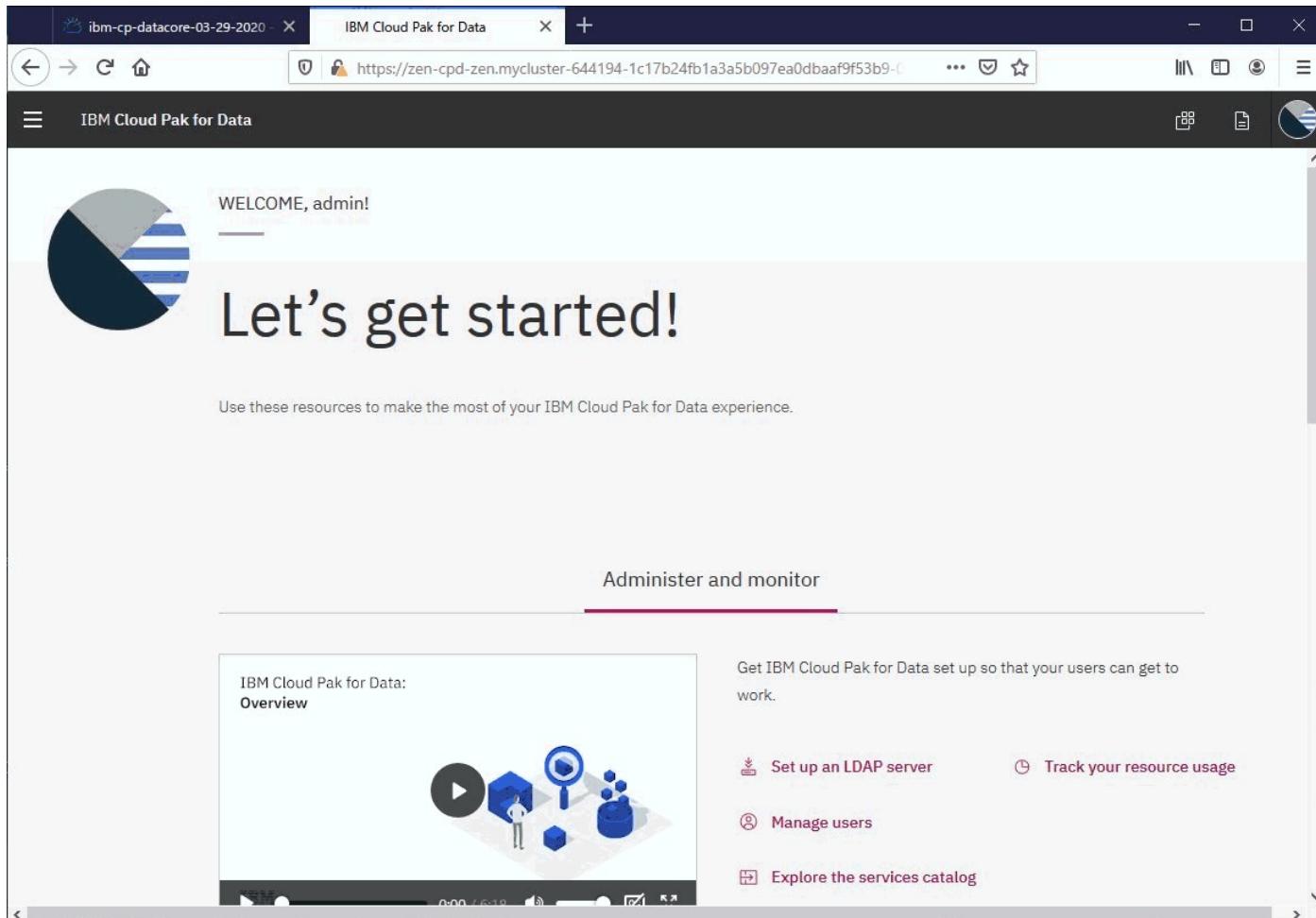
---

# Create Watson Discovery Instance (1 of 5)

The screenshot shows the IBM Cloud Schematics workspace interface for the workspace 'ibm-cp-datacore-03-29-2020'. The 'Resources' tab is selected in the sidebar. A cluster named 'mycluster' is listed, described as a 'Cluster hosting workspace resources' of type 'Openshift cluster'. A 'View cluster' button is present. Below this, a section for 'Cloud Pak for Data' is shown, stating 'Resource created by Schematics'. It lists a single resource named 'Cloud Pak for Data' with the URL: <https://zen-cpd-zen.mycluster-644194-1c17b24fb1a3a5b097ea0dbaaf9f53b9-0001.sjc03.containers.appdomain.cloud>. A blue speech bubble icon is located in the bottom right corner of the main content area.

1. Click URL hyperlink for CP4D instance

# Create Watson Discovery Instance (2 of 5)

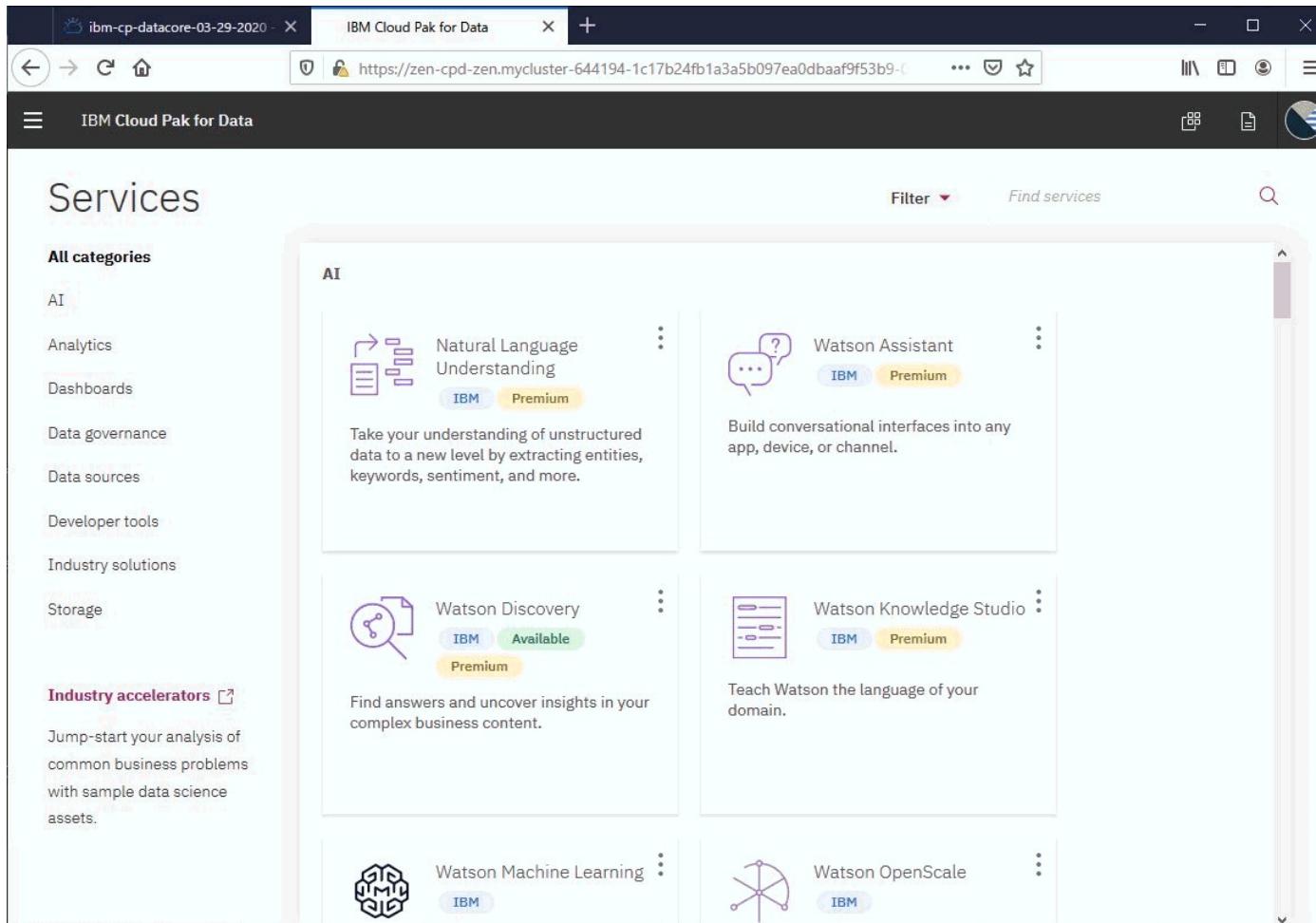


The screenshot shows the IBM Cloud Pak for Data welcome screen. At the top, there's a navigation bar with tabs for 'ibm-cp-datacore-03-29-2020' and 'IBM Cloud Pak for Data'. Below the navigation bar, the main content area has a 'WELCOME, admin!' message and a large 'Let's get started!' button. A pie chart icon is on the left. Below the button, there's a section titled 'Use these resources to make the most of your IBM Cloud Pak for Data experience.' It includes links for 'IBM Cloud Pak for Data: Overview', 'Set up an LDAP server', 'Track your resource usage', 'Manage users', and 'Explore the services catalog'. The 'Administer and monitor' tab is currently selected.

1. Click Service catalog icon.



# Create Watson Discovery Instance (3 of 5)



The screenshot shows the IBM Cloud Pak for Data service catalog. On the left, there's a sidebar with 'All categories' listed under 'AI'. The main area displays several AI services in a grid:

- Natural Language Understanding**: IBM Premium. Description: Take your understanding of unstructured data to a new level by extracting entities, keywords, sentiment, and more.
- Watson Assistant**: IBM Premium. Description: Build conversational interfaces into any app, device, or channel.
- Watson Discovery**: IBM Available. Premium. Description: Find answers and uncover insights in your complex business content.
- Watson Knowledge Studio**: IBM Premium. Description: Teach Watson the language of your domain.
- Watson Machine Learning**: IBM. Description: (not visible in the screenshot)
- Watson OpenScale**: IBM. Description: (not visible in the screenshot)

1. Click “Watson Discovery” tile in AI services.

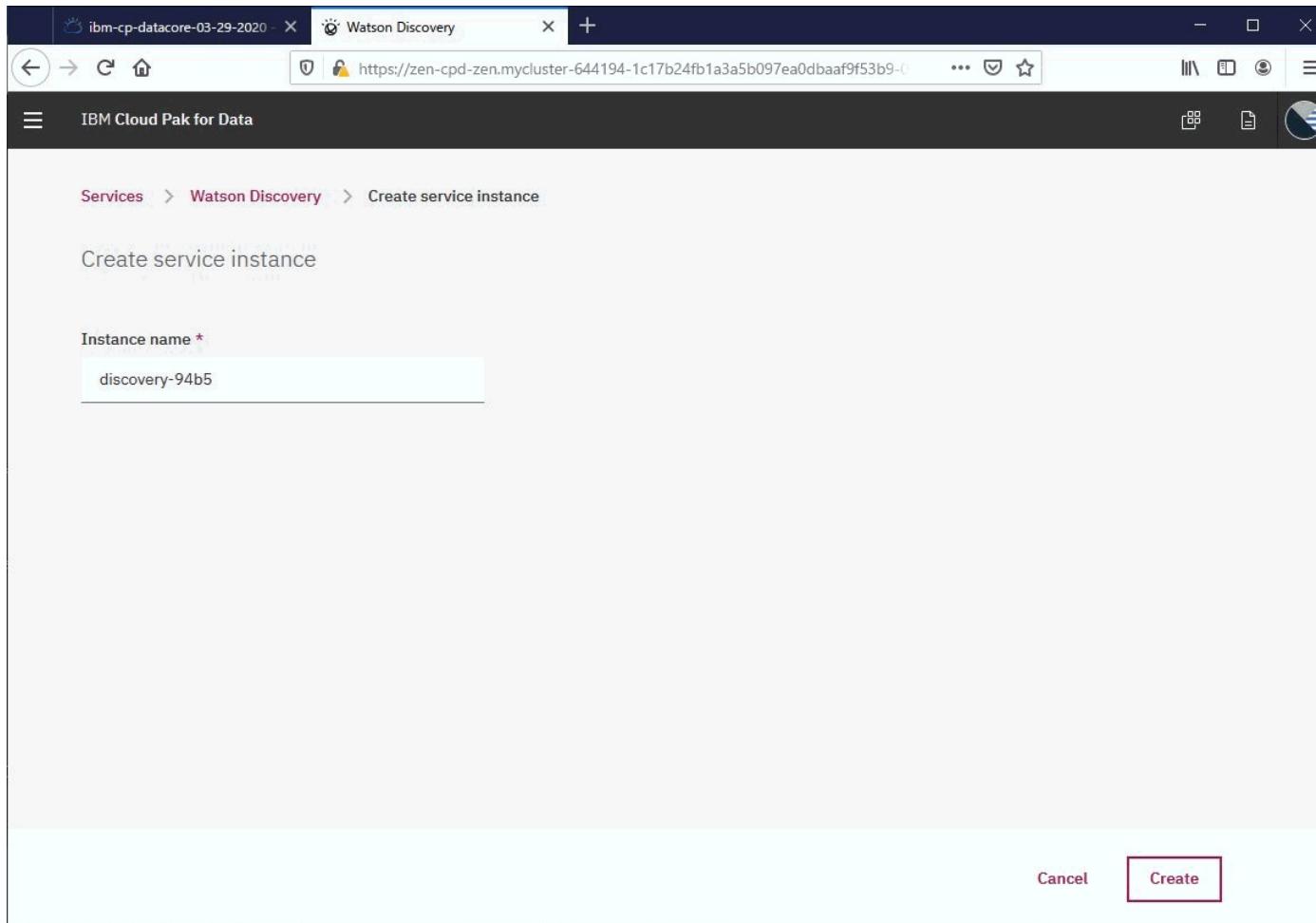
# Create Watson Discovery Instance (4 of 5)

The screenshot shows the IBM Cloud Pak for Data interface with the URL <https://zen-cpd-zen.mycluster-644194-1c17b24fb1a3a5b097ea0dbaaf9f53b9-0.ibmcloud.com/>. The page displays the Watson Discovery service details:

- Watson Discovery** icon
- IBM Available Premium Version: 2.3.1**
- Provision instance** button (highlighted with a red box)
- Description:** The IBM Watson Discovery add-on is an AI-powered search and content analytics engine that enables you to find answers and uncover insights that hide in your complex business content. With the Smart Document Understanding training interface, Watson Discovery can learn where answers live in your content based on a visual understanding of your documents.
- Want to do more? These Watson add-ons can help:**
  - Enhance Watson Discovery's ability to understand domain-specific language with Watson Knowledge Studio.
  - Surface answers to users through a conversational dialog driven by Watson
- Sample Query Results:** Shows three document snippets from "Banking\_Lending.pdf" and "Banking\_LoanSample.pdf". Each snippet includes a "View Document" link and "Relevant" or "Not relevant" buttons.

1. Click “Provision Instance” button

# Create Watson Discovery Instance (5 of 5)



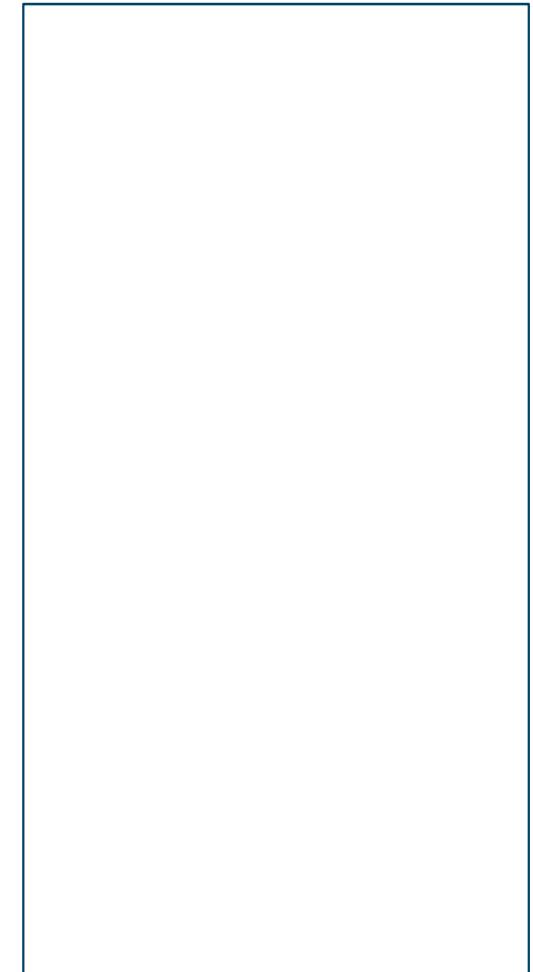
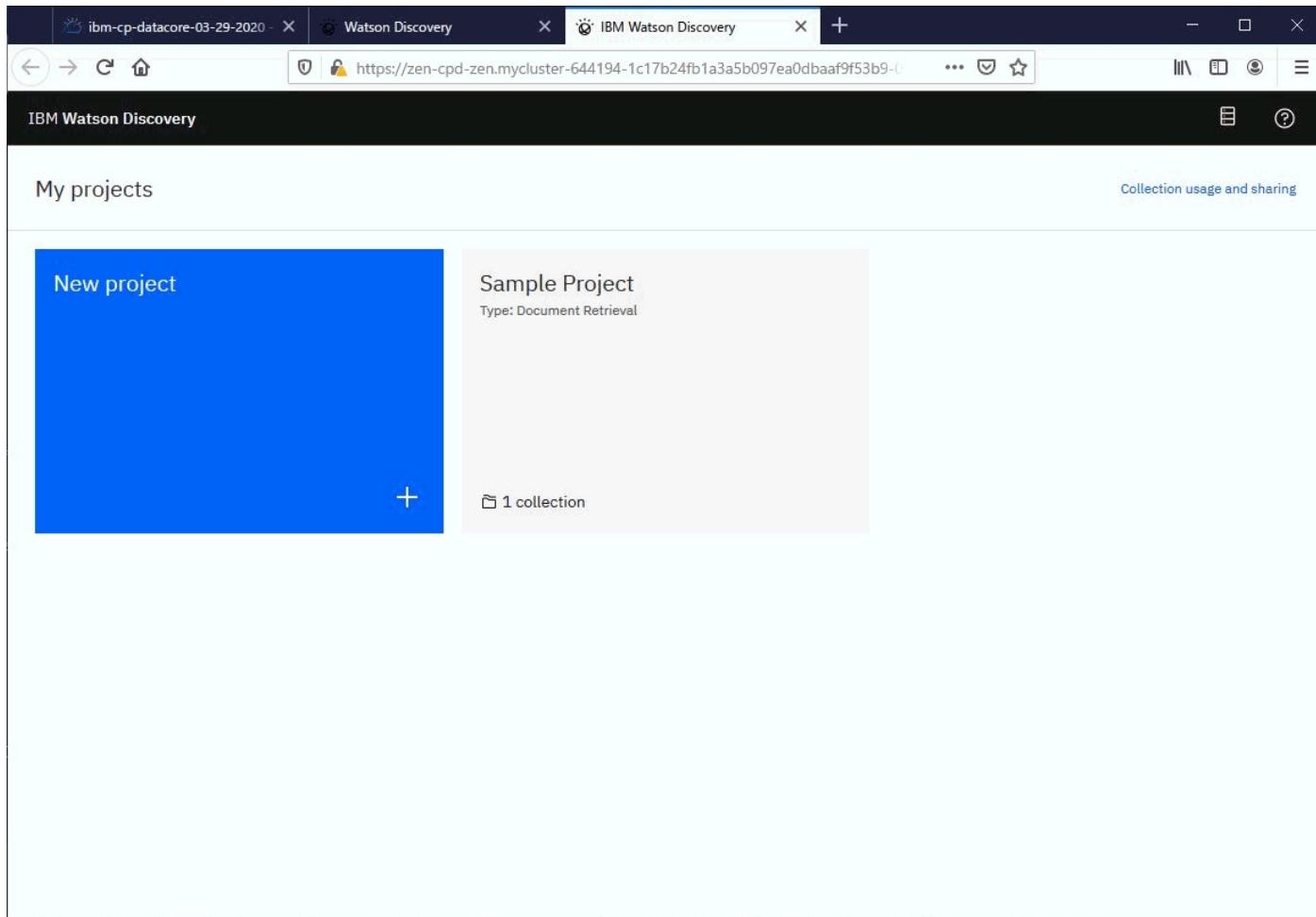
1. Set instance name and click “Create” button

# Launch Watson Discovery

The screenshot shows a web browser window with the title bar "ibm-cp-datacore-03-29-2020" and "Watson Discovery". The address bar displays the URL <https://zen-cpd-zen.mycluster-644194-1c17b24fb1a3a5b097ea0dbaaf9f53b9-0001.sjc03.cloud.ibm.com>. The main content area is titled "Service instance details" under "My instances > discovery-94b5". A "Menu" dropdown is open. The page includes a "Get started by reading through the [documentation](#) and [API reference](#)." button, a "Connection details" section with a "Download" link, and a "About this instance" section. The "About this instance" table has columns "Item" and "Description".

Item	Description
URL	<a href="https://zen-cpd-zen.mycluster-644194-1c17b24fb1a3a5b097ea0dbaaf9f53b9-0001.sjc03.cloud.ibm.com">https://zen-cpd-zen.mycluster-644194-1c17b24fb1a3a5b097ea0dbaaf9f53b9-0001.sjc03.cloud.ibm.com</a>
Bearer token	.....

# Watson Discovery



# Watson Discovery – SDU

The screenshot shows the IBM Watson Discovery interface. The top navigation bar includes a logo, a search bar with the URL <https://zen-cpd-zen.mycluster-644194-1c17b24fb1a3a5b097ea0dbaaf9f53b9-0>, and a 'My projects' section. Below the header, the main area is titled 'Sample Collection'. A horizontal navigation bar at the top of the collection view includes tabs for 'Activity', 'Identify fields' (which is selected and highlighted in blue), 'Manage fields', 'Enrichments', 'Processing settings', and 'CSV settings'. The 'Identify fields' tab displays a document titled 'Working with an integrated IBM Db2 Event Store database'. The document contains several sections: 'Working with an integrated IBM Db2 Event Store database', 'Accessing the database', 'Connecting to the database from another application', and 'Accessing sample notebooks'. On the right side of the collection view, there is a sidebar titled 'Field labels' which lists various document elements with corresponding color-coded squares: answer (yellow), author (purple), footer (green), header (blue), question (dark blue), subtitle (pink), table\_of\_contents (light blue), text (orange), title (red), table (brown), and image (teal). At the bottom of the collection view, there is a 'Submit page' button.

Text