Machine Learning 1.5.0

Release Notes

Date published: 2020-07-16
Date modified: 2023-01-31

CLOUDERA

Legal Notice

Contents

What's New

Major features and updates for the Cloudera Machine Learning service on Private Cloud.

January 24, 2023

CML on Private Cloud, version 1.5.0, has the following updates.

New features and updates

- Kubernetes - Kubernetes supports ECS 1.23 and OCP 4.10.
- ECS - External Image Registry & Cloudera-Default-Docker Registry are now supported on new installations of Private Cloud 1.5.0, but not on workloads upgraded from previous Private Cloud versions.
- External Image registry - The use of an external customer ImageDocker registry for reading images is now supported. For information on registry options, see Docker repository access for OCP or for ECS. CML supports these options with some known limitations, see here for more information.
- Cloudera Data Science Workbench (CDSW) 1.10.0 or later to CML migration (Technical Preview) - You can easily move your CDSW workloads to CML using this new migration software. The CDSW to CML migration software is in technical preview in CDP 1.5.0. Cloudera recommends that you use this process in test and development environments. It is not recommended for production deployments. For more information, see "Migrating Data Science Workbench to Machine Learning".
- ML Runtimes - Support for Custom Runtime Addons and PBJ Runtimes added.
- ML Runtime Addons - In Site Administration, Administrators can now change the status of Spark addons to "AVAILABLE", "DEPRECATED" or "DISABLED".
- Data Science and Machine Learning - Added support for Experiments v2 with MLflow, Monitoring for Applications and Models, and Model Registry (Technical Preview).
- Spark pushdown (Technical Preview) - Spark 2 pushdown to CDP Base functionality is supported.
- Model Registry - The Model Registry stores and manages machine learning models and associated metadata, such as the model's version, dependencies, and performance. The registry enables MLOps and facilitates the development, deployment, and maintenance of machine learning models in a production environment.

  Note: Model Registry is not supported with R models.
- Internal NFS Storage options on OCP: CephFS is used as the underlying storage provisioner for any new workspace using internal NFS on PVC 1.5.0. A storage class named "ocs-storagecluster-cephfs" with csi driver set to "openshift-storage.cephfs.csi.ceph.com" must exist in the cluster for new internal workspaces to get provisioned. Each workspace will have separate 1 TB internal storage.
- Internal NFS Storage options on ECS: Any new workspace using internal NFS on 1.5.0 will use Longhorn as the underlying storage provisioner. A storage class named "longhorn" with csi driver set to "driver.longhorn.io" must exist in the cluster for new internal workspaces to get provisioned. Each workspace will have separate 1 TB internal storage.
- Upgrades from 1.4.x with workspaces using internal NFS: On either ECS or OCP, internal workspaces running on PVC 1.4.0/1.4.1 use NFS server provisioner as the storage provisioner. These workspaces when upgraded to 1.5.0 will continue to run with the same NFS server Provisioner. However, NFS server provisioner is deprecated now and will not be supported from the 1.5.1 release onwards.

  Existing workspaces in 1.4.0/1.4.1 can be upgraded to 1.5.0 from PVC UI. After this, you can do one of the following:

  - Migrate the 1.5.0 upgraded workspace from NFS server provisioner to Longhorn (ECS) / Cephfs (OCP) if you want to continue using the same workspace in PVC 1.5.1 as well
  - Create a new 1.5.0 workspace and migrate the existing workloads to that before 1.5.1 release.

    Note: There is no change in the underlying storage of external NFS backed workspaces and these can be simply upgraded to 1.5.0.

- CML Team to CDP Group Sync - This synchronizes groups from the CDP management console to the ML workspace. See Creating a Team for more information.

November 18, 2022

CML on Private Cloud, version 1.4.1, has the following updates.

New features and updates

- Legacy CDSW Cluster Detected - After the upgrade to Private Cloud 1.4.1, if the base cluster contains a CDSW installation, you will see a message recommending you to upgrade the cluster. Do NOT click Upgrade as this feature is not yet GA.
- HDFS transparent encryption - Encryption at rest (HDFS transparent encryption) is supported in CML.

June 22, 2022

CML on Private Cloud, version 1.4.0, has the following updates.

New features and updates

- Model metrics visualization - This feature allows Data Scientists and Machine Learning Engineers to monitor technical metrics relating to their running models, such as resource consumption and request throughput, within Cloudera Machine Learning.

Fixed issues

- DSE-19937 - Fixed an issue where the pagination widget on the Session list page may not function as expected.
- DSE-20085 - Fixed a bug where Job report recipients who subscribed to notification emails when their jobs terminated, may receive notification emails for termination statuses that they did not subscribe to.
- DSE-19751Fixed a bug where projects may not be sorted correctly on the project list page when using the Created By field for sorting.

April 11, 2022

CML on Private Cloud, version 1.3.4, has the following updates.

New features and updates

- ML Runtimes - ML Runtimes are now supported in CML Private Cloud. For more information, see Managing ML Runtimes.
- Cloudera Data Visualization - Cloudera Data Visualization is now available in the default runtime.
- GPU Taint support - GPU taints, which affect node scheduling, are now supported for both OCP and ECS clusters. For more information, see GPU node setup.

January 13, 2022

CML on Private Cloud, version 1.3.3, has the following updates.

New features and updates

- Business User Experience - A new user role, ML BusinessUser, provides restricted access to view Applications created in CML.
- API v2 - A new API for operations on projects, jobs, models, and applications is now generally available.

Installation notes

- Upgrade - It is not possible to upgrade an existing ML workspace on an ECS cluster. You have to provision a new workspace.
- Upgrade - After upgrading the ECS control plane, model and experiment building in an ML workspace might fail. See the Known Issue for more information.

- Engines - Engine version 15-cml-2021.09-2 has been patched for CVE-2021-44228, the Apache Log4j2 vulnerability. You should use engine version 15-cml-2021.09-2 instead of version 15-cml-2021.09-1 wherever possible.

October 29, 2021

CML on Private Cloud, version 1.3.2, has the following updates.

Installation notes

- Upgrade - Fixed an issue so that upgrading a CML workspace with ML Governance enabled works.

October 4, 2021

CML on Private Cloud, version 1.3.1, has the following new features and updates.

New features and updates

- Embedded Container Service (ECS) is now supported.

Installation notes

- Installation - If ECS is installed using Cloudera Docker Registries, then CML Workspace Model and Experiment building is not supported.
- Upgrade - Upgrading a CML workspace with ML Governance enabled fails.

April 27, 2021

CML on Private Cloud, version 1.2, has the following new features and updates.

New features and updates

- Support for OCP 4.6 and upgrading from PVC 1.1.
- Improved non-transparent proxy support for air-gapped environments.
- Introduced Applied ML Prototypes (AMPs).
- Added NFS support:

    - NFS versions v3 and v4.x are supported.
    - External NFS security improvements - no_root_squash export option has been removed.
- Support added for custom service principals (Beta).
- Monitoring now uses CDP centralized Grafana. Added database metrics and improved alerts.

Bug fixes

- DSE-12037 - Fixed an issue with the seamless login for Grafana.
- DSE-14891 - Fixed an issue with broken Engine and Session log links.
- Various security fixes.

December 16, 2020

CML on Private Cloud, version 1.1, has the following new features and updates.

- MLOPS-216 - Production ML Support

    Model Metrics track machine model serving performance metrics. Model Governance use Apache Atlas to track builds, experiments and deployment of machine learning models.
- DSE-10777 - UMS Integration

    MLUser and MLAdmin resource roles are now available and assignable through Environment settings.
- DSE-12955 - Self Signed Private CA certs For custom container registries

    Customers can now use Container registries that are using self signed or private CA signed certificates. There is an option to upload the self signed or private CA signed certificates certificate during Private Cloud installation.

- DSE-10759 - GPU support

   The OpenShift Nvidia operator is now supported for use with CML workloads.

August 17, 2020

This is the first release of CML on Private Cloud, version 1.0.

CML on Private Cloud lets you:

- Run Machine Learning workloads on OpenShift clusters in your own data center.
- Easily onboard a new tenant and provision an ML workspace in a shared OpenShift environment.
- Enable data scientists to access shared data on CDP Private Cloud Base and CDW.
- Leverage Spark-on-K8s to spin up and down Spark clusters on demand.
- Take advantage of most CML features on public cloud, including Teams, Projects, Experiments, Models, and Applications.

Related Information
Known Issues and Limitations

Known Issues and Limitations

You might run into some known issues while using Cloudera Machine Learning on Private Cloud.

Do not use backtick characters in environment variable names

Avoid using backtick characters (`) in environment variable names, as this will cause sessions to fail with exit code 2.

Model Registry is not supported on R models

Model Registry is not supported on R models.

Model Registry model name cannot exceed 19 characters

The Model Registry model name cannot exceed 19 characters.

After upgrading a workspace, the pod evaluator is in CrashLoopBackOff state

This typically happens after a workspace upgrade. The solution is to upgrade the workspace as soon as the Control Plane is upgraded.

Migration from CDSW fails if Enable Model Metrics is not selected

CDSW to CML migration fails if you do not select the Enable Model Metrics option.

Workaround: In Cloudera Machine Learning Workspaces, Provision Workspace (CDSW Migration), select the Enable Model Metrics option:
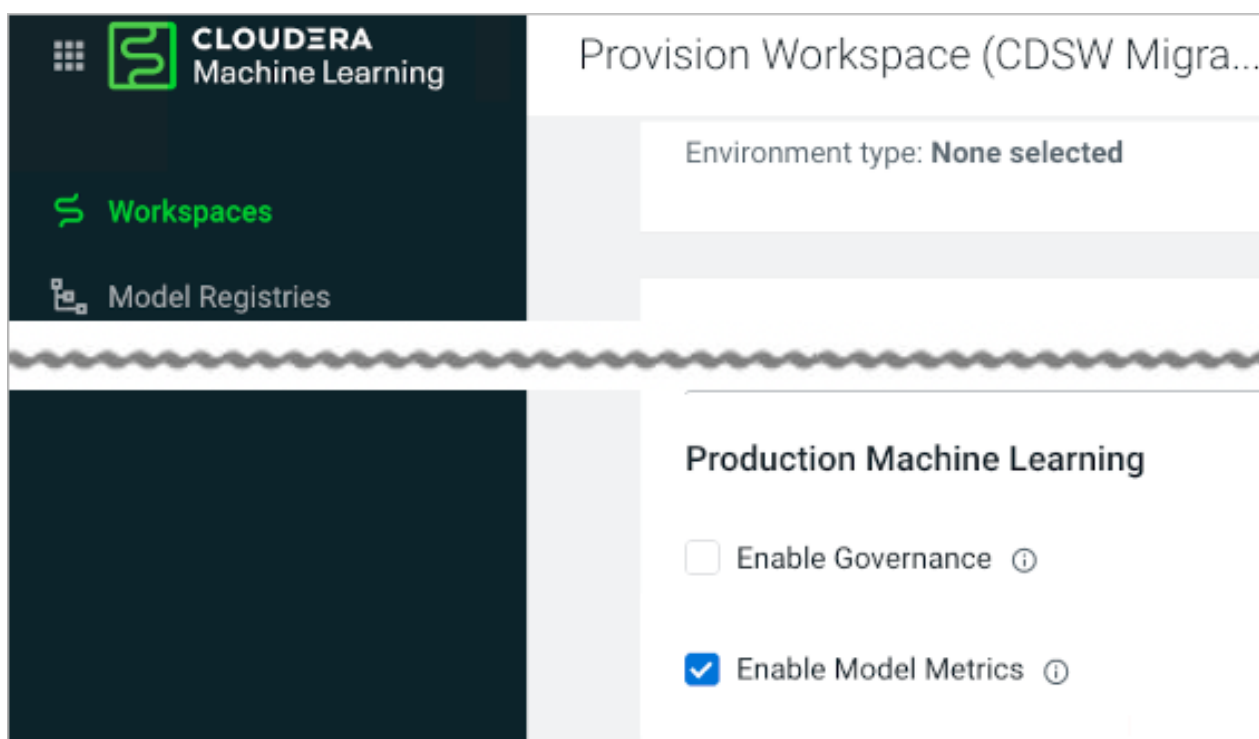
DSE-24690 CDSW to CML migration fails with the Cloudera default docker repository

During CDP Private Cloud installation, you are prompted to configure the Docker repository that Cloudera uses to deliver CDP Private Cloud Services. Choosing Use Cloudera's default Docker Repository can cause the Cloudera Data Science Workbench (CDSW) to CML migration to fail.

Workaround: During installation, in Configure Docker Repository, choose either one of the following options:

* Use an embedded Docker Repository
* Use a custom Docker Repository

The mlflow.log_model registered model files might not be available on NFS Server

When using mlflow.log_model, registered model files might not be available on the NFS server due to NFS server settings or network connections. This could cause the model to remain in the registering status.

Workaround:

* Re-register the model. It will register as an additional version, but it should correct the problem.
* Add the ARTIFACET_SYNC_PERIOD environment variable to ds-vfs Kubernetes deployment and set it to an integer value. This will set the model registry retry operation to twice the number of seconds specified by the artifact sync period integer value. If the ARTIFACET_SYNC_PERIOD is set to 30 seconds then model registry will retry for 60 seconds. The default value is 10 and model registry retries for 20 seconds. For example: -name: ARTIFACT_SYNC_PERIOD value: "30".

DSE-23329: Spark API fails when reading or writing to Ozone filesystem

CML with Spark as a runtime addon does not work with the Ozone file system.

DSE-33636: Workloads unable to start up after changing default hadoopCLI addon

Changing the default Hadoop CLI Runtime Addon causes jobs, models, and application workloads to be unable to start up.

Workarounds:

1. Open affected workload settings.
2. Update the workload (this updates the Hadoop CLI Addon associated with the workload to the default one.)
3. For Jobs: update.
4. For Applications: update and restart.
5. For Models: deploy a new build.

Please see Machine Learning Site administration>Disable Addons for related tasks.

DSE-20923: User search requires lower case

When searching for a user name in Site Administration Users , you must enter only lower-case letters for the name you are searching for. Lower-case letters will match upper-case letters in the target name.

DSE-19036: Model and experiment building fails after ECS upgrade

After ECS Control Plane upgrade, your ML workspace may fail to perform Model and Experiment building with a message like: Cannot connect to the Docker daemon at   unix:///var/run/docker.sock. Is the docker daemon running?

To resolve this problem, perform these steps:

1. Access Cloudera Manager
2. Navigate to the Embedded Cluster ECS Web UI: Clusters Your Embedded Cluster ECS Web UI ECS Web UI
3. Select the namespace of your ML workspace on the top left dropdown.
4. In Workloads Deployments , locate s2i-builder in the list.
5. In the Action menu for s2i-builder, select Restart.

DSE-13117: Container Image Registries assuming mutual TLS for authentication are not supported

If Private Cloud images are hosted in an image registry assuming mutual TLS for authentication, this will cause Model deployments and Experiments to fail. Mutual TLS registries are not supported.

DSE-12541: Self Signed Certificates for Container Registry cause Models and Experiments to fail

If you are using self-signed or Private CA signed certificates for Container image registry authentication, model deployements and experiments will fail with an error similar to: Error initializing source docker://<registryIP>:5000/alpine:latest: error   pinging docker registry <registryIP>:5000: Get https://<registryIP>:5000/v2/: x509:   certificate signed by unknown authority

As a workaround, create a ConfigMap in the namespace where the CML workspace is installed.

1. Create a ConfigMap as shown in this example. Here, <namespace> indicates the workspace where the CML workspace is installed.

```
kind: ConfigMap
apiVersion: v1
metadata:
  name: <external-registry-name>
  Namespace: <namespace>
data:
  registry.crt: |
    -----BEGIN CERTIFICATE-----
    < certificate content goes here >
    -----END CERTIFICATE-----
    -----BEGIN CERTIFICATE-----
    < certificate content goes here >
    -----END CERTIFICATE-----
```

2. Mount the ConfigMap to the s2i-builder deployment as shown here. Add the following mountPath in the volu meMounts section for the s2i-builder pod:

```
- mountPath: /etc/docker/certs.d/<registry>[:portnum]
```

```
  name: external-registry
```

Under the volumes section, add the ConfigMap reference:

```
- configMap:
     defaultMode: 420
     name: <external-registry-name>
  name: external-registry
```

3. Run the following command and check the output. Note in particular the mountPath and configMap specifications at the end.

```
# kubectl get deployment s2i-builder -n <namespace> -o yaml

apiVersion: extensions/v1beta1
kind: Deployment
metadata:
  [...]
  name: s2i-builder
  namespace: <namespace/workspacename>
  [...]
spec:
  [...]
  template:
    [...]
    spec:
    [...]
    containers:
    - name: s2i-builder
    [...]
    volumeMounts:
          [...]
          - mountPath:/etc/docker/certs.d/<registry>[:portnum]
            name: external-registry
      [...]
  volumes:
  [...]
  - configMap:
     defaultMode: 420
     name: <external-registry-name>
    name: external-registry
status:
[...]
```

DSE-12367: s2i-queue pod goes into CrashLoop Failure causing ML workspace installation to fail

CML workspace install can fail because the s2i-queue pod may be stuck in a CrashLoop Failure. The error in the logs might look similar to: Failed to create thread: Resource temporarily unavailable (11)   /usr/lib/rabbitmq/bin/rabbitmq-server: line 182: 45 Aborted (core dumped) start_rabbitmq_server "$@"   Only root or rabbitmq can run rabbitmq-server

To fix this, apply the following workaround: # kubectl set env statefulset/s2i-queue RABBITMQ_IO_THREAD_POOL_SIZE="50" -n <namespace>

DSE-12329: Email invitation feature

The feature to invite new users by email does not work in Public or Private cloud, but it still appears in the UI.

DSE-12289: Airgap support: Proxies are not supported in CML Private Cloud 1.0

Use of a proxy server, for example for external internet connectivity for an airgap cluster, is not supported. Transparent proxies, however, should work normally.

DSE-12238: Create Project request takes longer than timeout

If a Create Project request takes longer than a certain timeout, a second request might be submitted. If this happens, multiple projects with similar names might be created.

As a workaround, create an empty project, create a session inside the project, then git clone your project inside a workbench terminal. Additionally, you can upload a zip file or a folder using the file preview table.

If multiple forks are created, delete the extra ones.

DSE-12090: User displays as unknown in Event History

In the Event History on the workspace Events tab, a user may display as unknown if they are authenticated by LDAP.

Fix: The user needs to be assigned the IamViewer role to view these details.

DSE-11979: Certificate failure when pulling images from the S2I container registry

During Model or Experiment deployment, a certificate failure similar Failed to pull image x509: certificate signed by unknown authority can occur. This is due to a Red Hat issue with OpenShift Container Platform 4.3.x where the image registry cluster operator configuration must be set to Managed.

To set the configuration, first apply a patch using this command: # oc patch configs.imageregistry.operator.openshift. io cluster --patch '{"spec":{"managementState":"Managed"}}'

Next, run the following command: # oc get config cluster -o yaml

The managementState is now set to Managed.

DSE-11870: Hung File, Stale File, and Fork issues with NFS

Hung File Operations: Certain file operations, such as `stat(2)` or `stat(1)` might stop responding, and if the file operation was performed through the CML web UI, the web operation might timeout. This indicates an NFS server that is not reachable for some reason. The error might manifest itself on the web UI when you try to open an ML project as an HTTP error, code 500. Check the logs for error messages similar to the following:

```
 2020-07-13 22:42:23.914 1 ERROR AppServer.Lib.Utils Finish grpc, failed dat
a =
        [{"rpc":"1","service":"2","reqId":"3","err":"4"},"stat","VFS","18a07
980-c55a-11ea-9bb9-a35829b422d9",{"message":"
        5","stack":"6","code":4,"metadata":"7","details":"8","futureStack"
:"6"},"4
        DEADLINE_EXCEEDED: Deadline Exceeded","Error: 4 DEADLINE_EXCEEDED:
Deadline Exceeded\n at
        Object.exports.createStatusError (/home/cdswint/services
        /web/node_modules/grpc/src/common.js:91:15)\n at Object.onReceiveSta
tus
        (/home/cdswint/services/web/node_modules/grpc/src/client_interceptor
s.js:1209:28)\n at
        InterceptingListener._callNext (/home/cdswint/services/web/
        node_modules/grpc/src/client_interceptors.js:568:42)\n at
        InterceptingListener.onReceiveStatus
        (/home/cdswint/services/web/node_modules/grpc/src/client_intercept
ors.js:618:8)\n at
        callback (/home/cdswint/services/web/n
        ode_modules/grpc/src/client_interceptors.js:847:24)",{"_internal_r
epr":"9","flags":0},"Deadline
        Exceeded",{}]
```

Solution: Check your NFS server and make sure it is running. You will need to restart the NFS clients in your ML workspace's namespace. These are the "ds-vfs" and "s2i-client" pods. Simply delete the Kubernetes pods whose names start with "ds-vfs" and "s2i-client".

Stale File Handles: When opening a project from the ML web UI, an error message like "NFS: Stale file handle" shows up on the UI.

Solution: This is indicative of an NFS server and a client being out of sync, probably caused by a server restart along with file system content change on the server that the client is not aware of. You should restart NFS client pods in your ML workspace's namespace. The are the "ds-vfs", "s2i-client", and any user sessions that are affected by the "Stale file handle" error.

Project Fork Creating Multiple Copies: When creating a new project from an existing project using the "Fork" feature, you might see the operation seemingly fail on the UI, but it still ends up creating multiple copies of the source project.

Solution: This issue happens when forking a project takes longer than the idle connection timeout set on the external load balancer, as well as in HA Proxy policy settings on OpenShift. Increase the idle connection timeout to at least 5 minutes. Depending on the performance of the NFS server, a higher timeout may be necessary.

DSE-11837: Timeout limitation for Project API

If you create a project in the UI using git clone, you may get the error message Whoops, there was an unexpected error. If you create a project using the API, a timeout may occur.

Prerequisites for CML in Private Cloud:

•   Set any external load balancer server timeout to 5 min.

For a TLS Enabled Workspace:

•   Set the annotation haproxy.router.openshift.io/timeout=300 on each route in a deployed CML workspace namespace:

```
oc annotate route --all=true --overwrite=true -n
            <cml-namespace> haproxy.router.openshift.io/timeout=300s
```

For non-TLS Enabled Workspaces, this setting is made automatically.

Workaround: Even though an error message displays, project creation still occurs. Check the Projects page after a few minutes; project creation should be complete.

DSE-9549: TLS enabled workspaces require manual configuration

To provision a TLS-enabled workspace, the customer needs to perform several manual steps. This procedure is described in Deploy an ML Workspace with Support for TLS.

OPSAPS-58019: CML workspace installation failure due to includedir in krb5.conf file

If the /etc/krb5.conf on the Cloudera Manager host contains include or includedir directives, Kerberos-related failures may occur.

As a workaround, comment out the include and includedir lines in /etc/krb5.conf on the Cloudera Manager host. If configuration in those files and directories are needed, add them directly to /etc/krb5.conf.

DSE-21768: Spark3 runtime addons are not supported with Python 3.8 Runtimes

If the /etc/krb5.conf on the Cloudera Manager host contains include or includedir directives, Kerberos-related failures may occur.

Spark3 runtime addons are currently not supported with Python 3.8 Runtimes.

Related Information