Machine Learning 1.5.4

Release Notes

Date published: 2020-07-16
Date modified: 2024-07-31

CLOUDERA

Legal Notice

Contents

What's New

CDP Private Cloud 1.5.4 includes the following features for Cloudera Machine Learning.

New features

CML Service Accounts are available in CML Private Cloud

In Cloudera Machine Learning (CML), the Kerberos principal for the Service Account may not be the same as your login information. Therefore, ensure you provide the Kerberos identity when you sign in to the Service Account. For more information, see Authenticating Hadoop for CML Service Accounts.

Model Registry is available in CDP Private Cloud

Model Registry is now generally available (GA) in CDP Private Cloud. Model Registry in CDP Private Cloud uses Apache Ozone to store model artifacts. For creating a Model Registry you need the Ozone S3 gateway endpoint, the Ozone access key, and the Ozone secret key.

If you deploy Model Registry in an environment that contains one or more CML workspaces, you must synchronize the Model Registry with the workspaces. For more information, see Prerequisites for creating Model Registry and Synchronizing the model registry with a workspace.

Heterogeneous GPU usage

When using heterogeneous GPU clusters to run sessions and jobs, the available GPU accelerator labels need to be selected during workload creation. For more information, see Heterogeneous GPU clusters.

Data connections without auto discovery

Cloudera Machine Learning is a flexible, open platform, supporting connections to many data sources. The provided code samples demonstrate how to access local data for CML workloads. For more information, see Connecting to CDW.

Spark Log4j Configuration

Cloudera Machine Learning allows you to update Spark's internal logging configuration on a per-project basis. Spark logging properties can be customized for every session, and job with a default file path found at the root of your project. You can also specify a custom location with a custom environment variable. For more information, see Spark Log4j Configuration.

ML Metrics Collector service

The Metrics Collector service gathers data about how users and groups use resource quota, like how much CPU, Memory and GPU capacity (if any) is allocated, and what the users or groups utilize from that. The Metrics Collector service is running by default, but to collect data about resource quota metrics, you need to enable the Quota Management feature. For more information, see ML Metrics Collector Service overview.

Quota Management for group level

Quota Management Technical Preview (TP) release enables you to control how resources are allocated within your CML workspace on user and on group level. Yunikorn Gang Scheduling is also available, which is the default scheduling mechanism in Cloudera Machine Learning. For more information, see Quota Management overview and Yunikorn Gang Scheduling.

Restarting a failed AMP setup

You can now retry failed AMP deployment steps and continue the AMP setup to handle intermittent and configuration issues. For more information, see Restarting a failed AMP setup.

New Hadoop CLI Runtime Addon versions are available

The HadoopCLI 7.1.8.3-601 Runtime Addon is released for the Private Cloud.

Known Issues

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

Using dollar character in environment variables in Cloudera Machine Learning

Environment variables with the dollar ($) character are not parsed correctly by Cloudera Machine Learning (CML). For example, if you set PASSWORD="pass$123" in the project environment variables, and then try to read it using the echo command, the following output will be displayed: pass23

Workaround: Use one of the following commands to print the $ sign:

```
echo 24 | xxd -r -p
or
echo JAo= | base64 -d
```

Insert the value of the environment variable by wrapping it in the command substitution using $() or ``. For example, if you want to set the environment variable to ABC$123, specify:

```
ABC$(echo 24 | xxd -r -p)123
or
ABC`echo 24 | xxd -r -p`123
```

DSE-37827: Jupyter's RTC extension throws an error and notebooks become unusable

In certain cases, Jupyter's RTC (Real Time Collaboration) extension may cause errors claiming either that other sessions are active, or that other processes have accessed the notebook files. After these errors, the notebook becomes unusable due to the error messages and the CML session needs to be restarted.

Workaround:

You must disable the Jupyter RTC extension by performing the following tasks:

1. Create a Session.
2. Open the terminal.
3. Enter `nano /home/cdsw/.jupyter/labconfig/page_config.json`.
4. Add the following lines to the file:

```
{
  "disabledExtensions": {
    "@jupyter/collaboration-extension": true
  },
  "lockedExtensions": {
    "@jupyter/collaboration-extension": true
  }
}
```

5. Save and close the file.

DSE-36718: Disable auto synchronization feature for users and teams

The automated team and user synchronization feature is disabled. Newly installed or upgraded workspaces do not have the automatic synchronization option in the Cloudera Machine Learning UI.

Workaround: none

DSE-36759: AMPs and Feature Announcement sections do not work in NTP setups

Clouder Machine Learning (CML) Private Cloud setups with Non Transparent Proxy do not function properly, that affects Accelerators for ML Projects (AMPs) and Feature Announcements. The home page freezes, the feature announcement displays error message, and the AMPs do not load.

Workaround:

To avoid the home page freeze copy the following environment variables from the web deployment, and add them to the environment section of the API deployments:

- HTTP_PROXY
- HTTPS_PROXY
- NO_PROXY
- http_proxy
- https_proxy
- no_proxy

Note: The values are the same for both the lower and upper case variants.

DSE-32943: Enabling Service Accounts

Teams in the CML workspace can only run workloads within team projects with the Run as option for service accounts if they have previously manually added service accounts as a collaborator to the team.

DSE-35013: First CML workspace creation fails

On RHEL 8.8, during the first CML workspace installation on GPU with ECS external registry, pods might get stuck in the init or CrashLoop state.

First-time workspace installation is expected to fail. Consider this as a test workspace, and apply the following manual workaround for creating subsequent workspaces:

1. Restart or delete the pods which are in init or CrashLoop state in the test workspace.
2. Once all pods are in the running state, create new workspaces as needed.
3. Delete the test workspace from the CML UI if no longer needed.

OPSX-4603: Buildkit in ECS in CML PrivateCloud

Issue: BuildKit was introduced in ECS for building images of models and experiments. BuildKit is a replacement for Docker, which was previously used to build images of CML's models and experiments in ECS. Buildkit is only for OS RHEL8.x and CentOS 8.x.

Buildkit in CML Private Cloud 1.5.2 is a Technical Preview feature. Hence, having Docker installed on the nodes/hosts is still mandatory for models and experiments to work smoothly. Upcoming release will be completely eliminating the dependency of Docker on the nodes.

Workaround: None.

DSE-32285: Migration: Migrated models are failing due to image pull errors

Issue: After CDSW to CML migration (on-premises) via full-fledged migration tool, migrated models on CML Private Cloud Workspace fails on initial deployment. This is because the initial model deployment tries to pull images from on-premises's registry.

Workaround: Redeploy the migrated model. As this involves the build and deploy process, the image will be built, pushed to the CML Private Cloud Workspace's configured registry, and then the same image will be consumed for further usage.

DSE-28768: Spark Pushdown is not working with Scala 2.11 runtime

Issue: Scala and R are not supported for Spark Pushdown.

Workaround: None.

DSE-32304: On CML Private Cloud ECS terminal and ssh connections can terminate

Issue: In Private Cloud ECS, CML Terminal and SSH connections can terminate after an uncertain amount of time, usually after 4-10 minutes. This issue affects the usage of local IDEs to work with CML, as well as any customer application using a websocket connection.

Workaround: None.

DSE- 35251: Web pod crashes if a project forking takes more than 60 minutes

The web pod crashes if a project forking takes more than 60 minutes. This is because the timeout is set to 60 minutes using the `grpc_git_clone_timeout_minutes` property. The following error is displayed after the web pod crash:

```
2024-04-23 22:52:36.384    1737     ERROR       AppServer.VFS.grpc
                   crossCopy grpc error    data = [{"error":"1"},
{"code":4,"details":"2","metadata":"3"},"Deadline exceeded",{}]
           ["Error: 4 DEADLINE_EXCEEDED: Deadline exceeded\n
at callErrorFromStatus (/home/cdswint/services/web/node_modules/
@grpc/grpc-js/build/src/call.js:31:19)\n    at Object.onReceiveS
tatus (/home/cdswint/services/web/node_modules/@grpc/grpc-js/bui
ld/src/client.js:192:76)\n    at Object.onReceiveStatus (/home/c
dswint/services/web/node_modules/@grpc/grpc-js/build/src/client-
interceptors.js:360:141)\n    at Object.onReceiveStatus (/home/c
dswint/services/web/node_modules/@grpc/grpc-js/build/src/client-
interceptors.js:323:181)\n    at /home/cdswint/services/web/node
_modules/@grpc/grpc-js/build/src/resolving-call.js:94:78\n    at
 process.processTicksAndRejections (node:internal/process/task_q
ueues:77:11)\nfor call at\n    at ServiceClientImpl.makeUnaryReq
uest (/home/cdswint/services/web/node_modules/@grpc/grpc-js/buil
d/src/client.js:160:34)\n    at ServiceClientImpl.crossCopy (/ho
me/cdswint/services/web/node_modules/@grpc/grpc-js/build/src/mak
e-client.js:105:19)\n    at /home/cdswint/services/web/server-di
st/grpc/vfs-client.js:235:19\n    at new Promise (<anonymous>)\n
    at Object.crossCopy (/home/cdswint/services/web/server-dist/
grpc/vfs-client.js:234:12)\n    at Object.crossCopy (/home/cdswi
nt/services/web/server-dist/models/vfs.js:280:38)\n    at projec
tForkAsyncWrapper (/home/cdswint/services/web/server-dist/models/
projects/projects-create.js:229:19)"]
           node:internal/process/promises:288
           triggerUncaughtException(err, true /* fromPromise */);
           ^Error: 4 DEADLINE_EXCEEDED: Deadline exceeded
           at callErrorFromStatus (/home/cdswint/services/web/n
ode_modules/@grpc/grpc-js/build/src/call.js:31:19)
           at Object.onReceiveStatus (/home/cdswint/services/web/n
ode_modules/@grpc/grpc-js/build/src/client.js:192:76)
           at Object.onReceiveStatus (/home/cdswint/services/we
b/node_modules/@grpc/grpc-js/build/src/client-interceptors.js:36
0:141)
           at Object.onReceiveStatus (/home/cdswint/services/web/n
ode_modules/@grpc/grpc-js/build/src/client-interceptors.js:323:1
81)
           at /home/cdswint/services/web/node_modules/@grpc/grpc-
js/build/src/resolving-call.js:94:78
           at process.processTicksAndRejections (node:internal/p
rocess/task_queues:77:11)
           for call at
           at ServiceClientImpl.makeUnaryRequest (/home/cdswint/s
ervices/web/node_modules/@grpc/grpc-js/build/src/client.js:160:3
4)
```

```
        at ServiceClientImpl.crossCopy (/home/cdswint/services/
web/node_modules/@grpc/grpc-js/build/src/make-client.js:105:19)
        at /home/cdswint/services/web/server-dist/grpc/vfs-cli
ent.js:235:19
        at new Promise (<anonymous>)
        at Object.crossCopy (/home/cdswint/services/web/server-
dist/grpc/vfs-client.js:234:12)
        at Object.crossCopy (/home/cdswint/services/web/server-
dist/models/vfs.js:280:38)
        at projectForkAsyncWrapper (/home/cdswint/services/web/
server-dist/models/projects/projects-create.js:229:19) {
        code: 4,
        details: 'Deadline exceeded',
        metadata: Metadata { internalRepr: Map(0) {}, options:
  {} }
        }
```

Workaround: Increase the timeout limit, for example, to 120 minutes, using the
`grpc_git_clone_timeout_minutes` property.

```
UPDATE site_config SET grpc_git_clone_timeout_minutes = <new
 value>;
```

Fixed Issues

This section lists issues fixed in this release of Cloudera Machine Learning (CML) on Private Cloud.

DSE-24333: Embedded Spark UI fails to load in Spark pushdown mode for Spark 3

When Spark is run on yarn, the configuration attempts to disable proxying of the Spark UI through the Yarn Resource Manager. This setting was not respected by Spark 3 binaries and resulted in the Could not determine the proxy server for redirection error message when attempting to reach the embedded Spark UI on the driver.

This behavior works again in Spark 3.2.3 with CDH 7.1.7 and Spark 3.3.2 with CDH 7.1.9.

DSE-32887: Runtime updates cannot be enabled from the Cloudera Manager UI Site Administration dashboard

Runtime updates can now be enabled from the CML UI Site Administration dashboard.

DSE-29976: Job Arguments fail in PBJ runtimes

PBJ (Powered by Jupyter) Runtimes were not able to access Job arguments as standard command line arguments. To mitigate this, a new environment variable was added to Jobs, called "JOB_ARGUMENTS". This environment variable can be used to access arguments specified in the Job Creation or Job Update UI.

DSE-27285: The calculation of user resources was incorrect

The User Resources tab in the workspace UI calculated with the used and available resources for a cluster that also included the resources for the master node. As the workload cannot be scheduled in the master nodes, the master node resources shall not be considered in these calculations. This issue has been solved, the resources for the master are no longer included in the calculations for the User Resources.

Cumulative Hotfixes: ML Private Cloud

Review the cumulative hotfixes for Cloudera Machine Learning Private Cloud 1.5.4.

ML Private Cloud 1.5.4 CHF1

Review the features, fixes, and known issues in the Cloudera Machine Learning (CML) 1.5.4 Cumulative hotfix 1 release.

Fixed issues in 1.5.4-CHF1

Review the fixed issues in the Cloudera Machine Learning (CML) 1.5.4 Cumulative hotfix 1 release.

DSE-37209: Upgrade of CML workspace with external NFS fails

For setups with an external network file system (NFS), the Cloudera Machine Learning (CML) workspace upgrade flow was different from setups with internal NFSs. There was an issue with the persistent volume claim (PVC) size in external NFS flows which caused failures during the upgrade path to 1.5.4. Now, the issue has been fixed. For more details, see Cloudera Customer Advisory 2024-778: CML Private Cloud Data Service upgrade to 1.5.4 is failing with external NFS Technical Service Bulletin.

DSE-36078: Spark workload histories are not logged in Spark history server

The Spark history server did not log Spark workload history so far. This issue has been fixed and the Spark workload history is now logged in the base cluster Spark history server.

DSE-37192: Time zone discrepancy for usage API

Some data with respect to the workload run were stored in the local time zone, as a result of which the API returned empty results. By setting all the columns of the dashboard database to UTC time zone, there is no longer time zone discrepancy for the usage API. Note that the data stored prior to the upgrade remains in the previously set time zone.

DSE-36489: Opening existing JupyterLab session redirected to project page

After having logged out from the CML UI and having logged back, the opening of an existing JupyterLab session led to the project page instead of opening the JupyterLab UI. This issue has been fixed and after a new login the JupyterLab UI can be opened.

DSE-32204: Improved error messaging during failed Kerberos authentication in session startup

If there was an issue with Kerberos credential authentication, CML session startup hung until the regular session timed out and no warning message was displayed for the user. Now, a warning message with additional details on the status of the Kerberos credential authentication is displayed for the user along with a startup failure message after 1 minute.

DSE-36666: The migration readiness check is disrupted when engine readiness check is run

The shell command executor on Cloudera Data Science Workbench (CDSW) could not handle the CDSW to CML migration readiness timeout. This issue has been fixed, the timeout has been handled. Additionally, any SQL query-based readiness checks that remain unhandled will be displayed for manual validation.

DSE-36759: AMPs and Feature Announcement sections do not work in NTP setups

Cloudera Machine Learning (CML) Private Cloud setups with non-transparent proxy (NTP) did not function properly, affecting Accelerators for ML Projects (AMPs) and feature announcements. The home page froze, the feature announcement displayed an error message, and the AMPs did not load. This issue has been fixed with the enablement of non-transparent proxy in a Private Cloud setup.

DSE-37454: Possible CML terminal disconnections

The Cloudera Machine Learning (CML) terminal tended to disconnect on Embedded Container Service (ECS) because the proxy_read_timeout value on workload pods was set to a default value of 60 seconds. The proxy_read_timeout value has been increased to 121 seconds and the issue of the CML terminal disconnecting connections has been resolved.

DSE-37610: Due to version mismatch of the jackson-module-scala_3 the mlops governance pod image failed

Earlier, the version of the jackson-module-scala_3 was updated in the mlopsgovernance/pom.xml file, but it was not updated in the parent/pom.xml file. This issue has been solved and there is no longer version mismatch for the jackson-module-scala_3.

DSE-36748: File upload issue fixed using APIv2

Earlier, when uploading a file to a sub-folder in a project, the file was added to the root folder. This issue has been fixed and now when uploading a file to a subfolder it is added to the project's sub folder.

Known issues in 1.5.4-CHF1

Review the known issues in the Cloudera Machine Learning (CML) 1.5.4 Cumulative hotfix 1 release.

DSE-36967: Namespace Termination issue when using Portworx storage

There is an issue with Portworx version lower than 3.1.1, as the namespace deletion gets stuck in terminating state. Portworx is not able to cleanly unmount and clean up the underlying resources.

Workaround: The issues is fixed with Portworx version 3.1.1. Upgrade to Portworx version 3.1.1 or to later versions.

DSE-37002: Download API using v2 python swagger client does not work

The download_project_file method does not function as expected. The v2 python swagger only allows content to be downloaded as a string and not in bytes, as described in Binary string response encoding with Python 3.

Workaround: Use download_project_file_v2 as an alternate method.