

Machine Learning

CML Metrics Collector Service

Date published: 2020-07-16

Date modified: 2024-05-30

CLOUDBERA

<https://docs.cloudera.com/>

Legal Notice

© Cloudera Inc. 2024. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

ML Metrics Collector Service overview.....	4
Visualization in CML quota management.....	4
CML Quota Management dashboard.....	4
Troubleshooting for CML Metrics Collector Service.....	6

ML Metrics Collector Service overview

The Metrics Collector service shows metrics data for Cloudera Machine Learning (CML).

It sends this data to a server called Prometheus, which is part of the Cloudera Data Platform (CDP) control plane. Grafana, the visualization tool, connects to Prometheus to visualize these metrics. With the CML Quota Management, you can see various metrics from different CML workspaces under the CDP control plane.

Currently, 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.

Visualization in CML quota management

The Grafana dashboard, called Cloudera Machine Learning (CML) Quota Management, gives details about user and group resource quota limits.

The CML Quota Management shows to what extent the user or the group uses the available resources. However, currently it provides no information on which workload is utilized.

When the CML Quota Management feature is enabled, the Metrics Collector monitors to what extent the users and groups use their resources. It collects data when new quota limits are set or when new workloads are created, so that it can keep track of how resource quotas are being used over time.

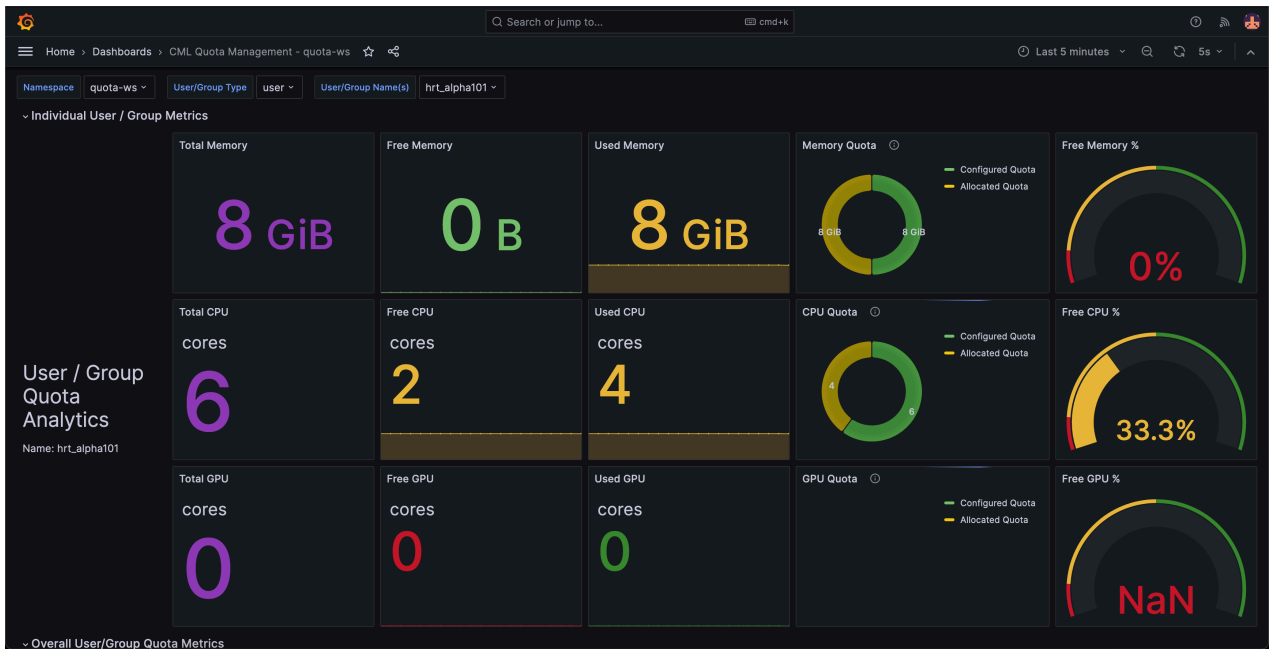
CML Quota Management dashboard

The Cloudera Machine Learning Quota Management Dashboard provides an overview of resource quota allocation and utilization of the users or groups within the Cloudera Machine Learning (CML) environment. Users can filter the dashboard by selecting the namespace for the CML workspace, choosing the user or group type, and specifying the user or group name.

Individual user or group metrics

- Total Memory, CPU, and GPU: It displays the total configured resources.
- Free Memory, CPU, and GPU: It shows the current available resources for the user or group.
- Used Memory, CPU, and GPU: It indicates the current utilized resources.
- Memory, CPU, and GPU Quota: It presents pie charts comparing configured versus allocated quotas.
- Free Memory %, CPU %, and GPU %: It displays the percentage of available free resources in gauge charts.

Figure 1: Individual user or group metrics



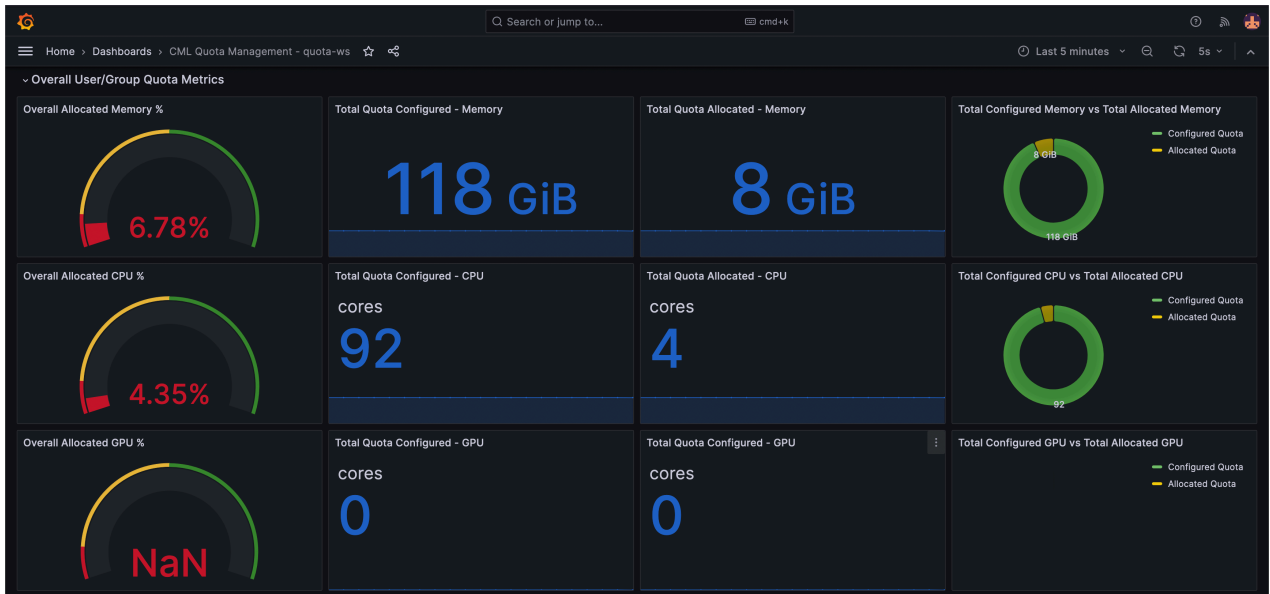
Note:

The panels are tailored specifically to the selections made in the filter, including the CML workspace, user or group type, and the chosen user or group name.

Overall user or group metrics

- Overall Allocated Memory %, CPU %, and GPU %: It shows the total allocated resources as a percentage of all users/groups.
- Total Quota Configured - Memory, CPU, and GPU: It displays the total configured resources of all users/groups.
- Total Quota Allocated - Memory, CPU, and GPU: It indicates the total allocated resources of all users/groups.
- Total Configured Memory vs Total Allocated Memory, CPU vs CPU, GPU vs GPU: It presents pie charts comparing the total configured resources versus the total allocated resources.

Figure 2: Overall user or group metrics



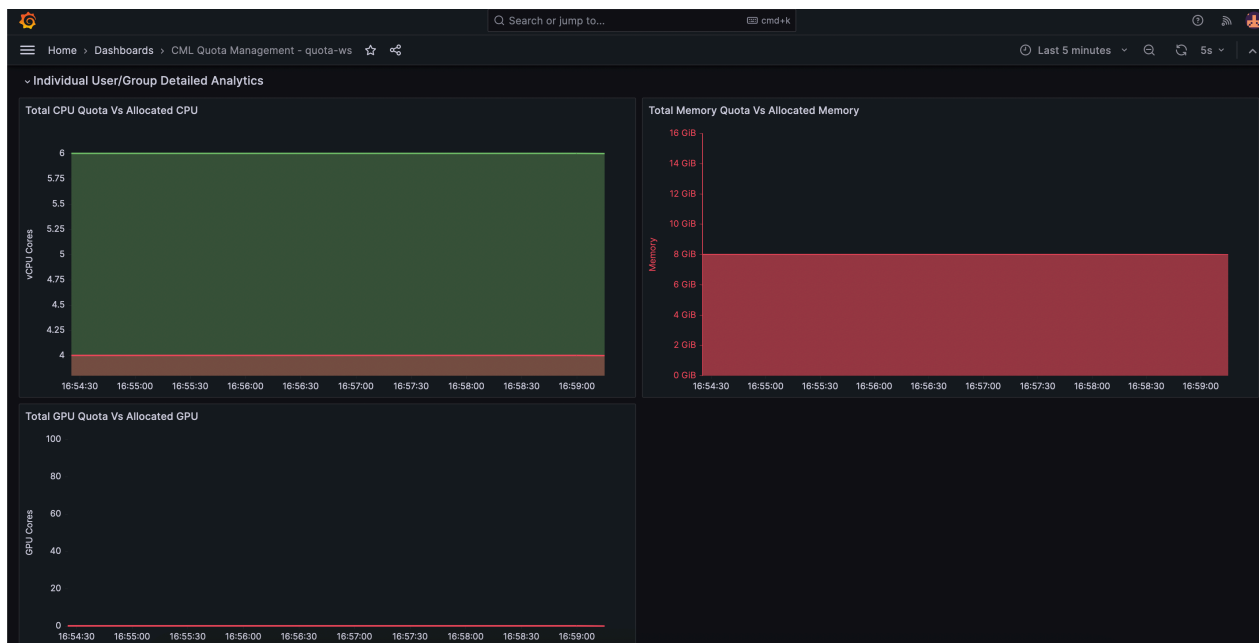


Note: The examples exclusively showcase panels that belong to the selected CML workspace in the filter.

Individual user or group detailed analytics

The examples showcase the configured Memory, CPU, and GPU resources compared to the allocated resources for the specified user or group.

Figure 3: individual user or group detailed analytics



Note:

The panels are tailored specifically to the selections made in the filter, including the CML workspace, user or group type, and the chosen user or group name.

Troubleshooting for CML Metrics Collector Service

Restarting the CML Metrics Collector Service might lead to multiple entries appearing in the dashboard.

Condition

If the Metrics Collector pod restarts, you might observe multiple entries appearing within the panels of the Cloudera Machine Learning (CML) Quota Management for some time.

Cause

This happens because Prometheus scrapes data from the metrics collector at the pod level, which includes retrieving information from the restarted pod during the scraping process.

Procedure

- No action is required.