

Managing Cloud Costs in Cloudera Data Engineering

Date published: 2020-07-30

Date modified: 2024-02-26



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

Cloudera Data Engineering Spot Instances.....	4
--	----------

Cloudera Data Engineering Spot Instances

Cloudera Data Engineering (CDE) supports spot instances to facilitate cloud cost savings for workloads that are not SLA-bound. If you have enabled spot instances for your CDE service, when you create a virtual cluster, you can specify whether drivers and executors run on spot instances or on-demand instances. The default setting uses on-demand instances for drivers, and spot instances for executors.



Note: Spot Instances are not currently supported for CDE on Microsoft Azure.

Spot instances are discounted instances that can be reclaimed at any time by the cloud provider. For fault-tolerant workloads such as Apache Spark, utilizing spot instances can provide significant cost savings. Because spot instances can disappear, impacting job performance, Cloudera recommends only using them for workloads without strict SLA requirements.

For development, testing, and other workloads that have more flexible requirements, Cloudera recommends using on-demand instances for drivers, and spot instances for executors. This reduces the impact of losing a spot instance, because executor failure recovery is more efficient than recovery from a driver failure.

To use spot instances, you must enable them on the CDE service. For instructions, see [Enabling Cloudera Data Engineering](#). You cannot enable spot instances on an existing CDE service. To use spot instances, create a new service.

After you have enabled spot instances for the CDE service, you can configure individual virtual clusters to use on-demand instances only, spot instances only, or the default hybrid model (on-demand for drivers, spot for executors).

For more information on spot instances, see the cloud provider documentation:

Amazon AWS

[Amazon EC2 Spot Instances](#)