

cloudera[®]

**CDK 4.1.x Powered By
Apache Kafka[®]**

Important Notice

© 2010-2021 Cloudera, Inc. All rights reserved.

Cloudera, the Cloudera logo, and any other product or service names or slogans contained in this document are trademarks of Cloudera and its suppliers or licensors, and may not be copied, imitated or used, in whole or in part, without the prior written permission of Cloudera or the applicable trademark holder. If this documentation includes code, including but not limited to, code examples, Cloudera makes this available to you under the terms of the Apache License, Version 2.0, including any required notices. A copy of the Apache License Version 2.0, including any notices, is included herein. A copy of the Apache License Version 2.0 can also be found here: <https://opensource.org/licenses/Apache-2.0>

Hadoop and the Hadoop elephant logo are trademarks of the Apache Software Foundation. All other trademarks, registered trademarks, product names and company names or logos mentioned in this document are the property of their respective owners. Reference to any products, services, processes or other information, by trade name, trademark, manufacturer, supplier or otherwise does not constitute or imply endorsement, sponsorship or recommendation thereof by us.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Cloudera.

Cloudera may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Cloudera, the furnishing of this document does not give you any license to these patents, trademarks copyrights, or other intellectual property. For information about patents covering Cloudera products, see <http://tiny.cloudera.com/patents>.

The information in this document is subject to change without notice. Cloudera shall not be liable for any damages resulting from technical errors or omissions which may be present in this document, or from use of this document.

Cloudera, Inc.

**395 Page Mill Road
Palo Alto, CA 94306
info@cloudera.com
US: 1-888-789-1488
Intl: 1-650-362-0488
www.cloudera.com**

Release Information

Version: CDK 3.1.x Powered By Apache Kafka®
Date: July 20, 2021

Table of Contents

| | |
|---|-----------|
| CDK Powered By Apache Kafka® | 6 |
| Understanding Kafka Terminology..... | 6 |
| | |
| CDK Powered By Apache Kafka® Release Notes..... | 11 |
| What's New in CDK Powered By Apache Kafka?..... | 11 |
| CDK Powered By Apache Kafka Requirements and Supported Versions..... | 14 |
| <i>Supported CDH and Cloudera Manager Releases.....</i> | <i>14</i> |
| <i>Supported Operating Systems.....</i> | <i>15</i> |
| <i>Supported JDK Versions.....</i> | <i>15</i> |
| <i>Ports Used by Kafka.....</i> | <i>15</i> |
| Issues Fixed in CDK Powered By Apache Kafka..... | 15 |
| <i>Issues Fixed in CDK 4.1.0 Powered By Apache Kafka.....</i> | <i>16</i> |
| <i>Issues Fixed in CDK 4.0.0 Powered By Apache Kafka.....</i> | <i>19</i> |
| <i>Issues Fixed in CDK 3.1.1 Powered By Apache Kafka.....</i> | <i>22</i> |
| <i>Issues Fixed in CDK 3.1.0 Powered By Apache Kafka.....</i> | <i>23</i> |
| <i>Issues Fixed in CDK 3.0.0 Powered By Apache Kafka.....</i> | <i>24</i> |
| <i>Issues Fixed in CDK 2.2.0 Powered by Apache Kafka.....</i> | <i>24</i> |
| <i>Issues Fixed in CDK 2.1.2 Powered By Apache Kafka.....</i> | <i>25</i> |
| <i>Issues Fixed in CDK 2.1.1 Powered By Apache Kafka.....</i> | <i>25</i> |
| <i>Issues Fixed in CDK 2.1.0 Powered By Apache Kafka.....</i> | <i>25</i> |
| <i>Issues Fixed in CDK 2.0.2 Powered By Apache Kafka.....</i> | <i>26</i> |
| <i>Issues Fixed in CDK 2.0.1 Powered By Apache Kafka.....</i> | <i>26</i> |
| <i>Issues Fixed in CDK 2.0.0 Powered By Apache Kafka.....</i> | <i>26</i> |
| <i>Issues Fixed in CDK 1.4.0 Powered By Apache Kafka.....</i> | <i>27</i> |
| <i>Issues Fixed in CDK 1.3.2 Powered By Apache Kafka.....</i> | <i>27</i> |
| <i>Issues Fixed in CDK 1.3.1 Powered By Apache Kafka.....</i> | <i>27</i> |
| <i>Issues Fixed in CDK 1.3.0 Powered By Apache Kafka.....</i> | <i>28</i> |
| <i>Issues Fixed in CDK 1.2.0 Powered By Apache Kafka.....</i> | <i>28</i> |
| Known Issues in CDK Powered By Apache Kafka..... | 29 |
| CDK Powered By Apache Kafka Incompatible Changes and Limitations..... | 34 |
| <i>Incompatible Changes and Limitations in CDK 4.1.0 Powered By Apache Kafka.....</i> | <i>34</i> |
| <i>Incompatible Changes and Limitations in CDK 4.0.0 Powered By Apache Kafka.....</i> | <i>34</i> |
| <i>Incompatible Changes and Limitations in CDK 3.1.0 Powered By Apache Kafka.....</i> | <i>35</i> |
| <i>Incompatible Changes and Limitations in CDK 3.0.0 Powered By Apache Kafka.....</i> | <i>35</i> |
| <i>Incompatible Changes and Limitations in CDK 2.0.0 Powered By Apache Kafka.....</i> | <i>35</i> |

CDK Powered By Apache Kafka® Version and Packaging Information.....37

Examples of CDK Powered By Apache Kafka Versions.....37
CDK Powered By Apache Kafka Versions.....37
Maven Artifacts for CDK Powered By Apache Kafka.....39

Installing, Migrating and Upgrading CDK Powered By Apache Kafka.....48

Installing or Upgrading CDK Powered By Apache Kafka®48
General Information Regarding Installation and Upgrade48
Rolling Upgrade to CDK 4.1.x Powered By Apache Kafka®48
Graceful Shutdown of Kafka Brokers.....50
Disks and Filesystem.....51
Installing or Upgrading Kafka from a Parcel.....51
Installing or Upgrading Kafka from a Package.....51
Migrating from Apache Kafka to CDK Powered By Apache Kafka.....52
Steps for Migrating from Apache Kafka to Cloudera Distribution of Apache Kafka.....53

Using Apache Kafka.....55

Using Apache Kafka Command-line Tools.....55
kafka-topics.....55
kafka-console-consumer.....55
kafka-console-producer.....55
kafka-consumer-groups.....55
kafka-reassign-partitions.....56
kafka-log-dirs.....60
Using Apache Kafka with Apache Spark Streaming.....61
Using Apache Kafka with Apache Flume.....61
Additional Considerations When Using Apache Kafka.....64

Apache Kafka Administration.....65

Configuring Apache Kafka Security.....65
Deploying SSL for Kafka.....65
Using Kafka Supported Protocols.....68
Enabling Kerberos Authentication.....69
Enabling Encryption at Rest.....70
Using Kafka with Sentry Authorization.....70
Configuring High Availability and Consistency for Apache Kafka.....73
Configuring Apache Kafka for Performance and Resource Management.....75
Partitions and Memory Usage.....75
Garbage Collection.....76
Handling Large Messages.....76

- Tuning Kafka for Optimal Performance*.....78
- Configuring JMX Ephemeral Ports*79
- Quotas*.....79
- Setting User Limits for Kafka*.....80
- Disk Management.....80
- Monitoring.....80
- Handling Disk Failures.....80
- Reassigning Replicas Between Log Directories*82
- Retrieving Log Directory Replica Assignment Information*.....82
- JBOD.....82
- JBOD Setup and Migration*.....82
- Viewing Apache Kafka Metrics.....85
- Working with Apache Kafka Logs.....85

Kafka Streams.....87

Kafka Frequently Asked Questions.....88

- Basics.....88
- Use Cases.....90
- References.....96

Appendix: Apache License, Version 2.0.....97

CDK Powered By Apache Kafka®

CDK Powered By Apache Kafka® is a distributed commit log service. Kafka functions much like a publish/subscribe messaging system, but with better throughput, built-in partitioning, replication, and fault tolerance. Kafka is a good solution for large scale message processing applications. It is often used in tandem with Apache Hadoop, Apache Storm, and Spark Streaming.

You might think of a log as a time-sorted file or data table. Newer entries are appended to the log over time, from left to right. The log entry number is a convenient replacement for a timestamp.

Kafka integrates this unique abstraction with traditional publish/subscribe messaging concepts (such as producers, consumers, and brokers), parallelism, and enterprise features for improved performance and fault tolerance.

The original use case for Kafka was to track user behavior on websites. Site activity (page views, searches, or other actions users might take) is published to central topics, with one topic per activity type.

Kafka can be used to monitor operational data, aggregating statistics from distributed applications to produce centralized data feeds. It also works well for log aggregation, with low latency and convenient support for multiple data sources.

Kafka provides the following:

- Persistent messaging with $O(1)$ disk structures, meaning that the execution time of Kafka's algorithms is independent of the size of the input. Execution time is constant, even with terabytes of stored messages.
- High throughput, supporting hundreds of thousands of messages per second, even with modest hardware.
- Explicit support for partitioning messages over Kafka servers. It distributes consumption over a cluster of consumer machines while maintaining the order of the message stream.
- Support for parallel data load into Hadoop.

Understanding Kafka Terminology

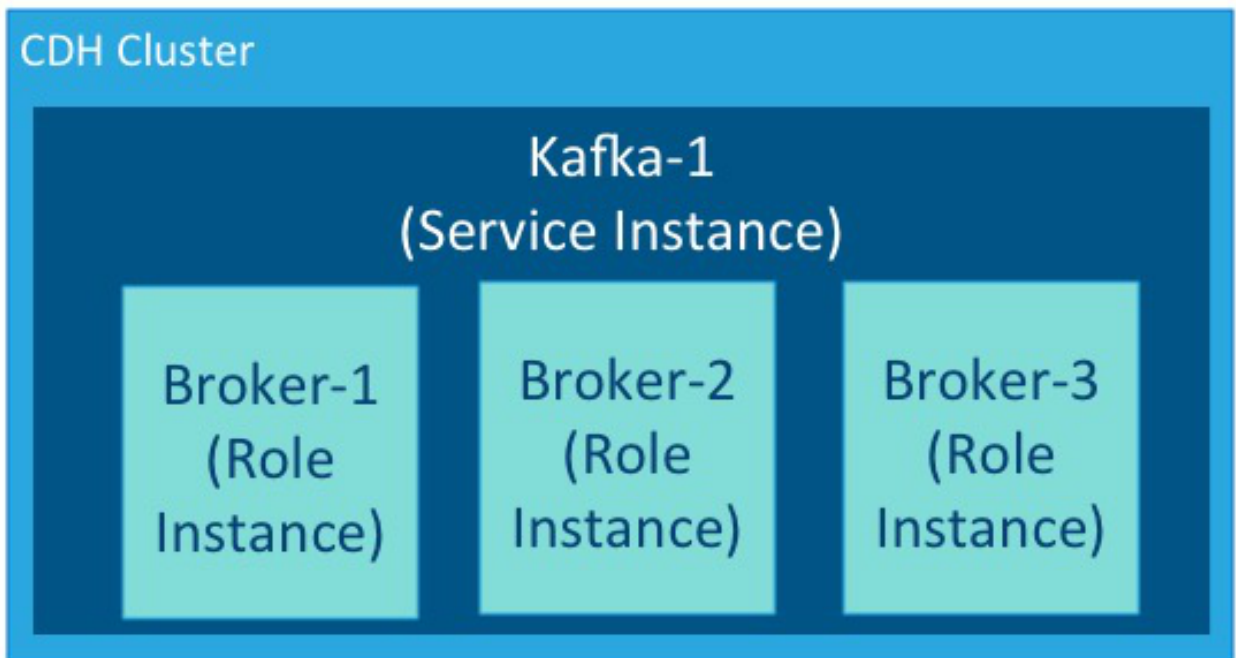
Kafka and Cloudera Manager use terms in ways that might vary from other technologies. This topic provides definitions for how these terms are used in Kafka with Cloudera Manager.

A *service* is an application that runs in a CDH cluster. Kafka is a service. ZooKeeper is a service that runs within a Kafka cluster. Other services include MapReduce, HDFS, YARN, Flume, and Spark.

A *role* is a feature of a service. A *broker* is a role in a Kafka service.

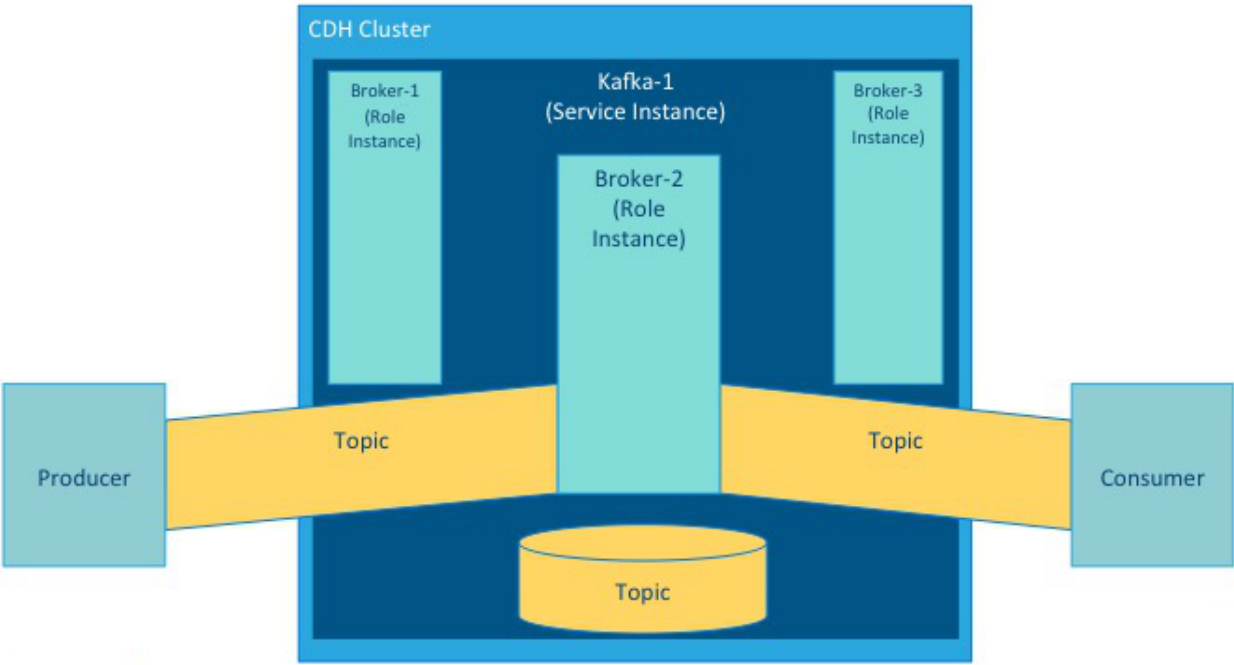


An *instance* is a deployed and configured software component. A cluster can include multiple roles and multiple instances of the same role. A *service instance* might be `Kafka-1`. `Kafka-1` might host the *role instances* `Broker-1`, `Broker-2`, and `Broker-3`.

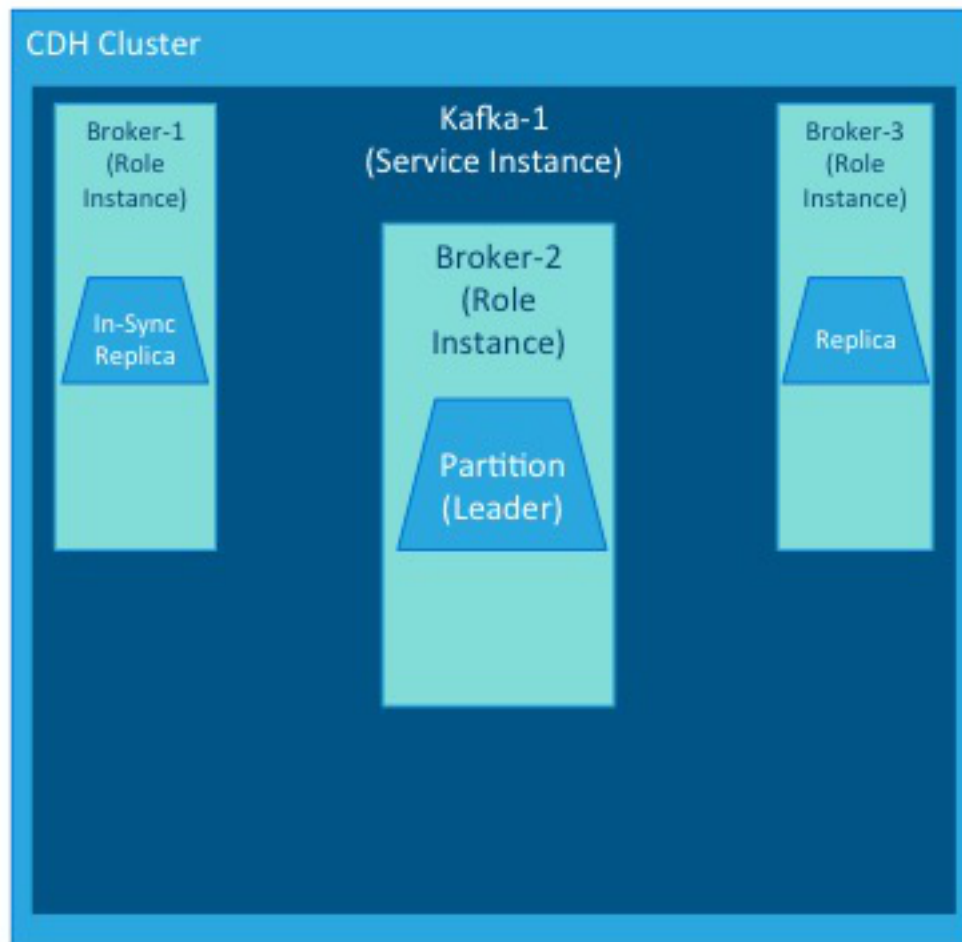


Kafka brokers process *records* organized into *topics*. A topic is a category of records that share similar characteristics. For example, a topic might consist of instant messages from social media or navigation information for users on a web site. Each topic has a unique corresponding table in data storage.

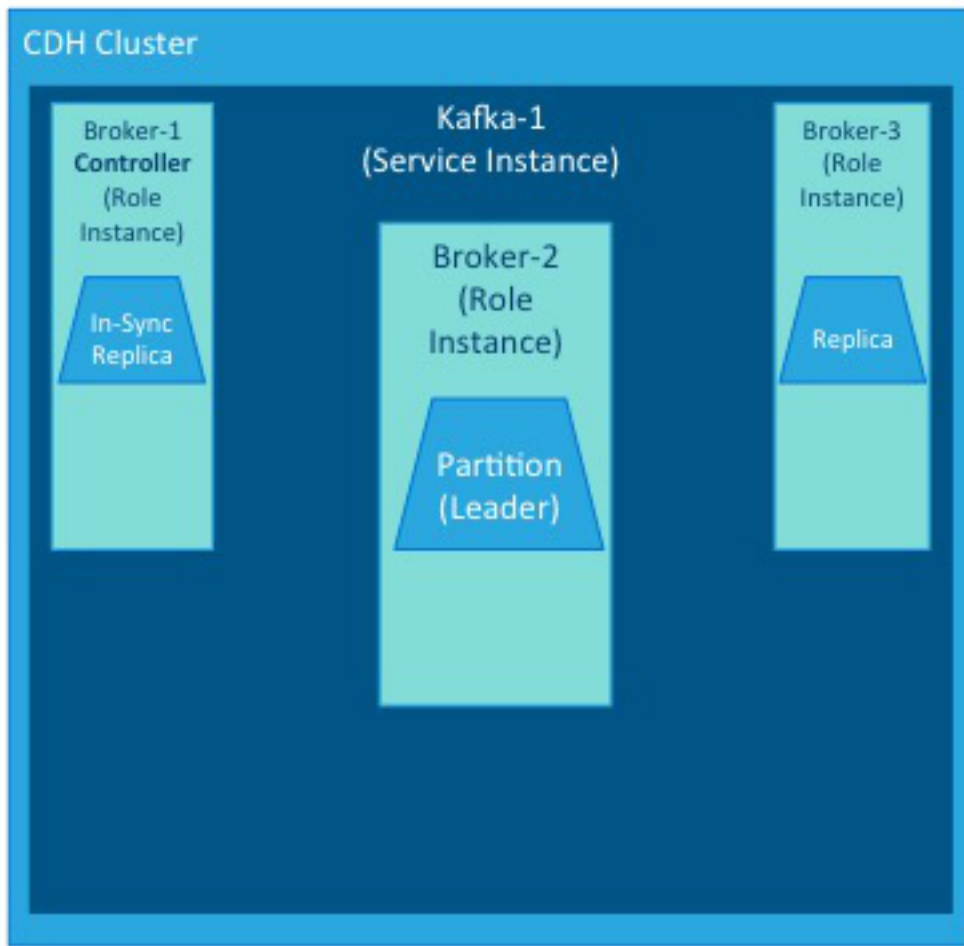
A *producer* is an external process that sends records to a Kafka topic. A *consumer* is an external process that receives topic streams from a Kafka cluster.



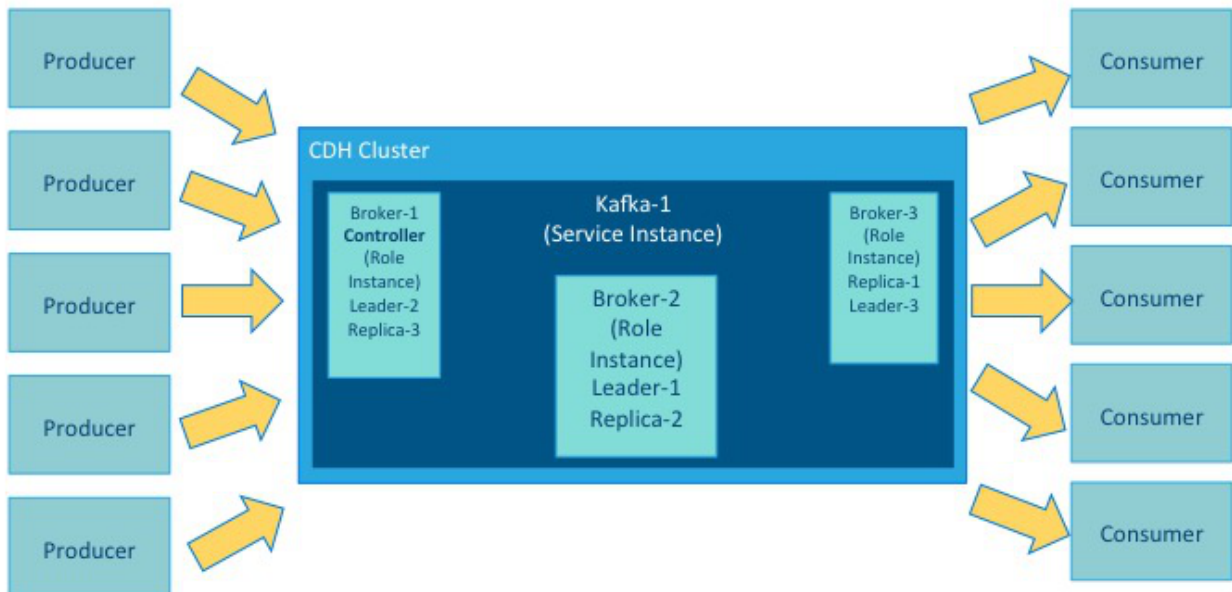
Brokers process topics in *partitions*. A partition on one broker in a cluster is the *leader*. The same partition is mirrored on one or more other brokers in the cluster as *replicas*. When a leader goes offline, a replica automatically takes its place and becomes the new leader for the topic. An *in-sync replica* is a replica that is completely up-to-date with the leader.



Each Kafka cluster has one broker that also acts as the *controller*. The controller is responsible for managing the states of partitions and replicas. It also performs administrative tasks, such as reassigning partitions.



While these illustrations show single instances of the components of a Kafka implementation, Kafka brokers typically host multiple partitions and replicas, with any number of producers and consumers, up to the requirements and limits of the deployed system.



CDK Powered By Apache Kafka® Release Notes

CDK Powered By Apache Kafka® provides a release guide that contains release and download information for installers and administrators. It includes release notes as well as information about versions and downloads. The guide also provides a release matrix that shows which major and minor release version of a product is supported with which release version of CDK Powered By Apache Kafka.

The Release Guide is comprised of topics including:

What's New in CDK Powered By Apache Kafka?

This section lists new features in CDK Powered By Apache Kafka. The following links provide detailed information for each release:

New Features in CDK 4.1.0 Powered By Apache Kafka

- **Rebase on Kafka 2.2.1**

CDK 4.1.0 Powered By Apache Kafka is a minor release based on Apache Kafka 2.2.1. For upstream release notes, see Apache Kafka version [2.2.0](#) and [2.2.1](#) release notes.

- **Kafka Topics Tool Able to Connect Directly to Brokers**

The `kafka-topics` command line tool is now able to connect directly to brokers with the `--bootstrap-server` option instead of `zookeeper`. The old `--zookeeper` option is still available for now. For more information, see [KIP-377](#).

New Features in CDK 4.0.0 Powered By Apache Kafka

- **Rebase on Kafka 2.1.0**

CDK 4.0.0 Powered By Apache Kafka is a major release based on Apache Kafka 2.1.0. For upstream release notes, see Apache Kafka version [1.0.2](#), [1.1.0](#), [1.1.1](#), [2.0.0](#), [2.0.1](#) and [2.1.0](#) release notes.



Warning: CDK Powered by Apache Kafka 4.0.0 is based on Apache Kafka 2.1.0, which contains a change to the internal schema used to store consumer offsets. As a result of this change, downgrading CDK Powered by Apache Kafka to a version lower than CDK 4.0.0 is **NOT** possible once Kafka has been upgraded to CDK 4.0.0 or higher.

- **JBOD Support**

As of CDK 4.0.0, Cloudera officially supports Kafka clusters with nodes using JBOD configurations. JBOD support introduces a new command line tool and improves an existing tool:

- A new tool, `kafka-log-dirs`, is added. The tool allows users to query partition assignment information.
- The `kafka-reassign-partitions` tool is expanded with a new functionality that allows users to reassign partitions between log directories. Users can move partitions to a different log directory on the same broker as well as to log directories on other brokers.

- **Kafka Streams**

Starting with CDK 4.0.0, Cloudera officially supports Kafka Streams. You can access the Apache Kafka website for information about how to use Kafka Streams.

- Read the [Kafka Streams Introduction](#) for an overview of the feature and an introductory video.
- Get familiar with Kafka Streams [Core Concepts](#).
- Understand Kafka Streams [Architecture](#).

- Access the [Quick Start](#) documentation to run a demonstration Kafka Streams Application.
- Use the [Tutorial](#) to write your first Kafka Streams Application.

- **Exactly Once Semantics**

Starting with CDK 4.0.0, Cloudera officially supports idempotent and transactional capabilities in the producer. This feature ensures that messages are delivered exactly once to a particular topic partition during the lifetime of a single producer.



Important: The configuration properties related to the idempotent and transactional capabilities of the producer are **NOT** available for configuration via Cloudera Manager. These properties must be set through safety valves. For a list of the properties, see [Incompatible Changes and Limitations in CDK 4.0.0 Powered By Apache Kafka](#) on page 34.

New Features in CDK 3.1.0 Powered By Apache Kafka

- **Rebase on Kafka 1.0.1**

CDK 3.1.0 Powered By Apache Kafka is a minor release based on Apache Kafka 1.0.1.

For upstream release notes, see Apache Kafka version [1.0.0](#) and [1.0.1](#) release notes.

- **Kafka uses HA-capable Sentry client**

This functionality enables automatic failover in the event that the primary Sentry host goes down or is unavailable.

- **Wildcard usage for Kafka-Sentry components**

You can specify an asterisk (*) in a Kafka-Sentry command for the TOPIC component of a privilege to refer to any topic in the privilege. Supported with CDH 5.14.2.

You can also use an asterisk (*) in a Kafka-Sentry command for the CONSUMERGROUPS component of a privilege to refer to any consumer groups in the privilege. This is useful when used with Spark Streaming, where a generated group.id may be needed. Supported with CDH 5.14.2.

- **Health Tests in Cloudera Manager**

Two new Kafka Broker Health Tests have been added to Cloudera Manager:

- Kafka Broker Swap Memory Usage
- Kafka Broker Unexpected Exits

These health tests are available when Kafka is managed by Cloudera Manager version 5.14 and later. For details, see [Kafka Broker Health Tests](#).

New Features in CDK 3.0.0 Powered By Apache Kafka

- **Rebase on Kafka 0.11.0.0**

CDK 3.0.0 Powered By Apache Kafka is a major release based on Apache Kafka 0.11.0.0. For upstream release notes, see Apache Kafka version [0.11.0.0](#) release notes.

- **Health test for offline and lagging partitions**

New health tests set the controller broker's health to BAD if the broker hosts at least one offline partition and the leader broker's health to CONCERNING if it hosts any lagging partitions. Supported with Cloudera Manager 5.14.0.

New Features in CDK 2.2.0 Powered By Apache Kafka

- **Rebase on Kafka 0.10.2**

CDK 2.2.0 Powered By Apache Kafka is rebased on Apache Kafka 0.10.2. For upstream release notes, see Apache Kafka version [0.10.2](#) release notes.

New Features in CDK 2.1.0 Powered By Apache Kafka

- **Rebase on Kafka 0.10**

Cloudera Distribution of Apache Kafka 2.1.0 is rebased on Apache Kafka 0.10. For upstream release notes, see Apache Kafka version [0.10](#) release notes.

- **Sentry Authentication**

Apache Sentry includes Kafka binding you can use to enable authorization in Kafka with Sentry. See [Configuring Kafka to Use Sentry Authorization](#) on page 70.

New Features in Cloudera Distribution CDK 2.0.0 Powered By Apache Kafka

- **Rebase on Kafka 0.9**

CDK 2.0.0 Powered By Apache Kafka is rebased on Apache Kafka 0.9. For upstream release notes, see Apache Kafka version [0.9](#) release notes.

- **Kerberos**

CDK 2.0.0 Powered By Apache Kafka supports Kerberos authentication of connections from clients and other brokers, including to ZooKeeper.

- **SSL**

CDK 2.0.0 Powered By Apache Kafka supports wire encryption of communications from clients and other brokers using SSL.

- **New Consumer API**

CDK 2.0.0 Powered By Apache Kafka includes a new Java API for consumers.

- **MirrorMaker**

MirrorMaker is enhanced to help prevent data loss and improve reliability of cross-data center replication.

- **Quotas**

You can use per-user quotas to throttle producer and consumer throughput in a multitenant cluster. See [Quotas](#) on page 79.

New Features in CDK 1.4.0 Powered By Apache Kafka

- CDK 1.4.0 Powered By Apache Kafka is distributed as a package as well as a parcel. See [CDK Powered By Apache Kafka® Version and Packaging Information](#) on page 37.

New Features in CDK 1.3.2 Powered By Apache Kafka

- **RHEL 7.1**

Kafka 1.3.2 supports RHEL 7.1. See [Supported Operating Systems](#) on page 15

New features in CDK 1.3.0 Powered By Apache Kafka

- **Metrics Reporter**

Cloudera Manager now displays Kafka metrics. Use the values to identify current performance issues and plan enhancements to handle anticipated changes in workload. See [Viewing Apache Kafka Metrics](#) on page 85.

- **MirrorMaker configuration**

Cloudera Manager allows you to configure the Kafka MirrorMaker cross-cluster replication service. You can add a MirrorMaker role and use it to replicate to a machine in another cluster. See [Kafka MirrorMaker](#).

New Features in CDK 1.1.0 Powered By Apache Kafka

- **New producer**

The producer added in CDK 1.1.0 Powered By Apache Kafka combines features of the existing synchronous and asynchronous producers. Send requests are batched, allowing the new producer to perform as well as the asynchronous producer under load. Every send request returns a response object that can be used to retrieve status and exceptions.

- **Ability to delete topics**

You can now delete topics using the `kafka-topics --delete` command.

- **Offset management**

In previous versions, consumers that wanted to keep track of which messages were consumed did so by updating the offset of the last consumed message in ZooKeeper. With this new feature, Kafka itself tracks the offsets. Using offset management can significantly improve consumer performance.

- **Automatic leader rebalancing**

Each partition starts with a randomly selected leader replica that handles requests for that partition. When a cluster first starts, the leaders are evenly balanced among hosts. When a broker restarts, leaders from that broker are distributed to other brokers, which results in an unbalanced distribution. With this feature enabled, leaders are assigned to the original replica after a restart.

- **Connection quotas**

Kafka administrators can limit the number of connections allowed from a single IP address. By default, this limit is 10 connections per IP address. This prevents misconfigured or malicious clients from destabilizing a Kafka broker by opening a large number of connections and using all available file handles.

CDK Powered By Apache Kafka Requirements and Supported Versions

The following sections describe software requirements and supported versions of complementary software for CDK Powered By Apache Kafka:

Supported CDH and Cloudera Manager Releases

For the list of supported releases of CDH and Cloudera Manager, see [CDH and Cloudera Manager Supported Versions](#).

Supported Integrations

Flume and Spark Connectors to Kafka

Flume and Spark connectors to Kafka are included with CDH 5.7.x and higher and only work with CDK 2.0.x Powered By Apache Kafka and higher.

Sentry Integration with Kafka

Sentry authorization integration with Kafka is available with CDK Powered By Apache Kafka version 2.1.x and higher on CDH 5.9.x and higher.

CDH 5.13.x introduces Sentry-HA; CDK 3.1 provides a Sentry-HA aware client. You can connect with an older version, but it won't be in high-availability mode: if the Sentry server that serves Kafka goes away, Kafka will not have access to authorization information.

Supported Operating Systems

For the list of supported operating systems, see [CDH and Cloudera Manager Supported Operating Systems](#).

SUSE Linux Enterprise Server (SLES)

Unlike CentOS, SLES limits virtual memory by default. Changing this default requires adding the following entries to the `/etc/security/limits.conf` file:

```
* hard as unlimited
* soft as unlimited
```

Supported JDK Versions

CDK 3.0 and higher Powered By Apache Kafka require JDK 8, and do not support JDK 7.

For a list of supported and tested JDK versions for CDK 2.2.x and below Powered By Apache Kafka, see [CDH and Cloudera Manager Supported JDK Versions](#).



Note: If you decide to use the G1 garbage collector and you use JDK 1.7, make sure you use u51 or newer.

Ports Used by Kafka

Kafka uses the TCP ports listed in the following table. Before deploying Kafka, ensure that these ports are open on each system.

| Component | Service | Port | Access Requirement | Comment |
|-------------|-------------------------|-------|--------------------|---|
| Broker | TCP Port | 9092 | External/Internal | The primary communication port used by producers and consumers; also used for inter-broker communication. |
| Broker | TLS/SSL Port | 9093 | External/Internal | A secured communication port used by producers and consumers; also used for inter-broker communication. |
| Broker | JMX Port | 9393 | Internal | Internal use only. Used for administration via JMX. |
| MirrorMaker | JMX Port | 9394 | Internal | Internal use only. Used to administer the producer and consumer of the MirrorMaker. |
| Broker | HTTP Metric Report Port | 24042 | Internal | Internal use only. This is the port via which the HTTP metric reporter listens. It is used to retrieve metrics through HTTP instead of JMX. |

Issues Fixed in CDK Powered By Apache Kafka

The following upstream issues are fixed in each release of CDK Powered By Apache Kafka:

Issues Fixed in CDK 4.1.0 Powered By Apache Kafka

Potential to bypass transaction and idempotent ACL checks in Apache Kafka

It is possible to manually craft a Produce request which bypasses transaction and idempotent ACL validation. Only authenticated clients with Write permission on the respective topics are able to exploit this vulnerability.

Products affected:

- CDH
- CDK Powered by Apache Kafka

Releases affected:

- CDH versions 6.0.x, 6.1.x, 6.2.0
- CDK versions 3.0.x, 3.1.x, 4.0.x

Users affected: All users who run Kafka in CDH and CDK.

Date/time of detection: September, 2018

Severity (Low/Medium/High):7.1 (High) ([CVSS:3.0/AV:N/AC:L/PR:L/UI:N/S:U/C:L/I:H/A:H](#))

Impact: Attackers can exploit this issue to bypass certain security restrictions to perform unauthorized actions. This can aid in further attacks.

CVE: CVE-2018-17196

Immediate action required: Update to a version of CDH containing the fix.

Addressed in release/refresh/patch:

- CDH 6.2.1, 6.3.2
- CDK 4.1.0

Upstream Issues Fixed

The following upstream issues are fixed in CDK 4.1.0:

- [\[KAFKA-1120\]](#) - Controller could miss a broker state change
- [\[KAFKA-2334\]](#) - Prevent HW from going back during leader failover
- [\[KAFKA-3832\]](#) - Kafka Connect's JSON Converter never outputs a null value
- [\[KAFKA-3832\]](#) - Kafka Connect's JSON Converter never outputs a null value
- [\[KAFKA-4217\]](#) - KStream.transform equivalent of flatMap
- [\[KAFKA-4453\]](#) - add request prioritization
- [\[KAFKA-4468\]](#) - Correctly calculate the window end timestamp after read from state stores
- [\[KAFKA-4850\]](#) - RocksDb cannot use Bloom Filters
- [\[KAFKA-5117\]](#) - Kafka Connect REST endpoints reveal Password typed values
- [\[KAFKA-5272\]](#) - Improve validation for Alter Configs (KIP-133)
- [\[KAFKA-5462\]](#) - Add a configuration for users to specify a template for building a custom principal name
- [\[KAFKA-5503\]](#) - Idempotent producer ignores shutdown while fetching ProducerId
- [\[KAFKA-5503\]](#) - Idempotent producer ignores shutdown while fetching ProducerId
- [\[KAFKA-5692\]](#) - Refactor PreferredReplicaLeaderElectionCommand to use AdminClient
- [\[KAFKA-5950\]](#) - AdminClient should retry based on returned error codes
- [\[KAFKA-5994\]](#) - Improve transparency of broker user ACL misconfigurations
- [\[KAFKA-6036\]](#) - Enable logical materialization to physical materialization
- [\[KAFKA-6388\]](#) - Error while trying to roll a segment that already exists
- [\[KAFKA-6431\]](#) - Lock contention in Purgatory
- [\[KAFKA-6478\]](#) - kafka-run-class.bat fails if CLASSPATH contains spaces

- [\[KAFKA-6567\]](#) - KStreamWindowReduce can be replaced by KStreamWindowAggregate
- [\[KAFKA-6774\]](#) - Improve default groupId behavior in consumer
- [\[KAFKA-6833\]](#) - KafkaProducer throws "Invalid partition given with record" exception
- [\[KAFKA-6928\]](#) - StreamsPartitionAssignor is double retrying within InternalTopicManager
- [\[KAFKA-6970\]](#) - Kafka streams lets the user call init() and close() on a state store, when inside Processors
- [\[KAFKA-7024\]](#) - Rocksdb state directory should be created before opening the DB
- [\[KAFKA-7037\]](#) - delete topic command replaces '+' from the topic name which leads incorrect topic deletion
- [\[KAFKA-7051\]](#) - Improve the efficiency of the ReplicaManager when there are many partitions
- [\[KAFKA-7054\]](#) - Kafka describe command should throw topic doesn't exist exception.
- [\[KAFKA-7096\]](#) - Consumer should drop the data for unassigned topic partitions
- [\[KAFKA-7110\]](#) - Windowed changelog keys not deserialized properly by TimeWindowedSerde
- [\[KAFKA-7165\]](#) - Error while creating ephemeral at /brokers/ids/BROKER_ID
- [\[KAFKA-7181\]](#) - Kafka Streams State stuck in rebalancing after one of the StreamThread encounters IllegalStateException
- [\[KAFKA-7235\]](#) - Use brokerZkNodeVersion to prevent broker from processing outdated controller request
- [\[KAFKA-7253\]](#) - The connector type responded by worker is always null when creating connector
- [\[KAFKA-7259\]](#) - Remove deprecated ZKUtils usage from ZkSecurityMigrator
- [\[KAFKA-7312\]](#) - Transient failure in kafka.api.AdminClientIntegrationTest.testMinimumRequestTimeouts
- [\[KAFKA-7352\]](#) - KIP-368: Allow SASL Connections to Periodically Re-Authenticate
- [\[KAFKA-7367\]](#) - Streams should not create state store directories unless they are needed
- [\[KAFKA-7391\]](#) - Introduce close(Duration) to Producer and AdminClient instead of close(long, TimeUnit)
- [\[KAFKA-7402\]](#) - Kafka Streams should implement AutoCloseable where appropriate
- [\[KAFKA-7412\]](#) - Bug prone response from producer.send(ProducerRecord, Callback) if Kafka broker is not running
- [\[KAFKA-7418\]](#) - Add '--help' option to all available Kafka CLI commands (KIP-374)
- [\[KAFKA-7420\]](#) - Global stores should be guarded as read-only for regular tasks
- [\[KAFKA-7433\]](#) - Introduce broker options in TopicCommand to use AdminClient
- [\[KAFKA-7443\]](#) - OffsetOutOfRangeException in restoring state store from changelog topic when start offset of local checkpoint is smaller than that of changelog topic
- [\[KAFKA-7446\]](#) - Better error message to explain the upper limit of TimeWindow
- [\[KAFKA-7449\]](#) - Kafka console consumer is not sending topic to deserializer
- [\[KAFKA-7461\]](#) - Connect Values converter should have coverage of logical types
- [\[KAFKA-7476\]](#) - SchemaProjector is not properly handling Date-based logical types
- [\[KAFKA-7478\]](#) - Reduce OAuthBearerLoginModule verbosity
- [\[KAFKA-7510\]](#) - KStreams RecordCollectorImpl leaks data to logs on error
- [\[KAFKA-7518\]](#) - FutureRecordMetadata.get deadline calculation from timeout is not using timeunit
- [\[KAFKA-7528\]](#) - Standardize on Min/Avg/Max Kafka metrics' default value
- [\[KAFKA-7532\]](#) - Missing stringification in error message "Removed ArrayBuffer() from list of shutting down brokers. (kafka.controller.KafkaController)"
- [\[KAFKA-7536\]](#) - TopologyTestDriver cannot pre-populate KTable or GlobalKTable
- [\[KAFKA-7537\]](#) - Only include live brokers in the UpdateMetadataRequest sent to existing brokers if there is no change in the partition states
- [\[KAFKA-7549\]](#) - Old ProduceRequest with zstd compression does not return error to client
- [\[KAFKA-7551\]](#) - Refactor to create both producer & consumer in Worker
- [\[KAFKA-7557\]](#) - optimize LogManager.truncateFullyAndStartAt()
- [\[KAFKA-7560\]](#) - PushHttpMetricsReporter should not convert metric value to double
- [\[KAFKA-7567\]](#) - Clean up internal metadata usage for consistency and extensibility
- [\[KAFKA-7568\]](#) - Return leader epoch in ListOffsets responses
- [\[KAFKA-7576\]](#) - Dynamic update of replica fetcher threads may fail to start/close fetchers
- [\[KAFKA-7584\]](#) - StreamsConfig throws ClassCastException if max.in.flight.request.per.connect is specified as String
- [\[KAFKA-7607\]](#) - NetworkClientUtils.sendAndReceive can take a long time to return during shutdown

- [\[KAFKA-7610\]](#) - Detect consumer failures in initial JoinGroup
- [\[KAFKA-7612\]](#) - Fix javac warnings and enable warnings as errors
- [\[KAFKA-7616\]](#) - MockConsumer can return ConsumerRecords objects with a non-empty map but no records
- [\[KAFKA-7620\]](#) - ConfigProvider is broken for KafkaConnect when TTL is not null
- [\[KAFKA-7641\]](#) - Add `consumer.group.max.size` to cap consumer metadata size on broker
- [\[KAFKA-7655\]](#) - Metadata spamming requests from Kafka Streams under some circumstances, potential DOS
- [\[KAFKA-7657\]](#) - Invalid reporting of stream state in Kafka streams application
- [\[KAFKA-7660\]](#) - Stream Metrics - Memory Analysis
- [\[KAFKA-7671\]](#) - A KStream/GlobalKTable join shouldn't reset the repartition flag
- [\[KAFKA-7672\]](#) - The local state not fully restored after KafkaStream rebalanced, resulting in data loss
- [\[KAFKA-7673\]](#) - Upgrade RocksDB to include fix for WinEnvIO::GetSectorSize
- [\[KAFKA-7678\]](#) - Failed to close producer due to java.lang.NullPointerException
- [\[KAFKA-7687\]](#) - Print batch level information in DumpLogSegments when deep iterating
- [\[KAFKA-7692\]](#) - updateFirstUnstableOffset NPE due to sequenceId overflow in ProducerStateManager.append
- [\[KAFKA-7693\]](#) - "IllegalArgumentException: Invalid negative sequence number used" in Kafka Client
- [\[KAFKA-7697\]](#) - Possible deadlock in kafka.cluster.Partition
- [\[KAFKA-7702\]](#) - Prefixed ACLs don't work with single character prefix
- [\[KAFKA-7704\]](#) - kafka.server.ReplicaFetchManager.MaxLag.Replica metric is reported incorrectly
- [\[KAFKA-7709\]](#) - ConcurrentModificationException occurs when iterating through multiple partitions in Sender.getExpiredInflightBatches
- [\[KAFKA-7712\]](#) - Handle exceptions from immediately connected channels in Selector
- [\[KAFKA-7719\]](#) - Improve fairness in SocketServer processors
- [\[KAFKA-7734\]](#) - Metrics tags should use LinkedHashMap to guarantee ordering
- [\[KAFKA-7738\]](#) - Track partition leader epochs in client metadata
- [\[KAFKA-7741\]](#) - Bad dependency via SBT
- [\[KAFKA-7742\]](#) - DelegationTokenCache#hmaclDCache entry is not cleared when a token is removed using removeToken(String tokenId) API.
- [\[KAFKA-7752\]](#) - zookeeper-security-migration.sh does not remove ACL on kafka-acl-extended
- [\[KAFKA-7755\]](#) - Kubernetes - Kafka clients are resolving DNS entries only one time
- [\[KAFKA-7758\]](#) - When Naming a Repartition Topic with Aggregations Reuse Repartition Graph Node for Multiple Operations
- [\[KAFKA-7759\]](#) - Disable WADL output in Connect REST API
- [\[KAFKA-7762\]](#) - KafkaConsumer uses old API in the javadocs
- [\[KAFKA-7781\]](#) - Add validation check for Topic retention.ms property
- [\[KAFKA-7786\]](#) - Fast update of leader epoch may stall partition fetching due to FENCED_LEADER_EPOCH
- [\[KAFKA-7798\]](#) - Expose embedded client context from KafkaStreams threadMetadata
- [\[KAFKA-7808\]](#) - AdminClient#describeTopics should not throw InvalidTopicException if topic name is not found
- [\[KAFKA-7824\]](#) - Require member.id for initial join group request
- [\[KAFKA-7837\]](#) - maybeShrinkIsr may not reflect OfflinePartitions immediately
- [\[KAFKA-7838\]](#) - improve logging in Partition.maybeShrinkIsr()
- [\[KAFKA-7873\]](#) - KafkaBasedLog's consumer should always seek to beginning when starting
- [\[KAFKA-7880\]](#) - KafkaConnect should standardize worker thread name
- [\[KAFKA-7890\]](#) - Invalidate ClusterConnectionState cache for a broker if the hostname of the broker changes.
- [\[KAFKA-7895\]](#) - Ktable suppress operator emitting more than one record for the same key per window
- [\[KAFKA-7897\]](#) - Invalid use of epoch cache with old message format versions
- [\[KAFKA-7902\]](#) - SASL/OAUTHBEARER can become unable to connect: javax.security.sasl.SaslException: Unable to find OAuth Bearer token in Subject's private credentials (size=2)
- [\[KAFKA-7909\]](#) - Ensure timely rebalance completion after pending members rejoin or fail
- [\[KAFKA-7915\]](#) - SASL authentication failures may return sensitive data to client
- [\[KAFKA-7916\]](#) - Streams store cleanup: unify wrapping!!

- [\[KAFKA-7920\]](#) - Do not permit zstd use until `inter.broker.protocol.version` is updated to 2.1
- [\[KAFKA-7935\]](#) - `UNSUPPORTED_COMPRESSION_TYPE` if `ReplicaManager.getLogConfig` returns `None`
- [\[KAFKA-7945\]](#) - `ExpiringCredentialRefreshingLogin` - `timeout` value is negative
- [\[KAFKA-8002\]](#) - Replica reassignment to new log dir may not complete if future and current replicas segment files have different base offsets
- [\[KAFKA-8012\]](#) - `NullPointerException` while truncating at high watermark can crash replica fetcher thread
- [\[KAFKA-8040\]](#) - Streams needs to handle timeout in `initTransactions`
- [\[KAFKA-8061\]](#) - Handle concurrent `ProducerId` reset and call to `Sender` thread shutdown
- [\[KAFKA-8065\]](#) - Forwarding modified timestamps does not reset timestamp correctly
- [\[KAFKA-8069\]](#) - Committed offsets get cleaned up right after the coordinator loading them back from `__consumer__offsets` in broker with old inter-broker protocol version (< 2.2)

Issues Fixed in CDK 4.0.0 Powered By Apache Kafka

CDK 4.0.0 Powered By Apache Kafka fixes the following issues:

Kafka stuck with under-replicated partitions after Zookeeper session expires

This problem might occur when your Kafka cluster includes a large number of under-replicated Kafka partitions. One or more broker logs include messages such as the following:

```
[2016-01-17 03:36:00,888] INFO Partition [__samza_checkpoint_event-creation_1,3] on
broker 3: Shrinking ISR for partition [__samza_checkpoint_event-creation_1,3] from 6,5
to 5 (kafka.cluster.Partition)
[2016-01-17 03:36:00,891] INFO Partition [__samza_checkpoint_event-creation_1,3] on
broker 3: Cached zkVersion [66] not equal to that in zookeeper, skip updating ISR
(kafka.cluster.Partition)
```

There will also be an indication of the ZooKeeper session expiring in one or more Kafka broker logs around the same time as the previous errors:

```
INFO zookeeper state changed (Expired) (org.I0Itec.zkclient.ZkClient)
```

The log is typically in `/var/log/kafka` on each host where a Kafka broker is running. The location is set by the property `kafka.log4j.dir` in Cloudera Manager. The log name is `kafka-broker-hostname.log`. In diagnostic bundles, the log is under `logs/hostname-ip-address/`.

Affected Versions: CDK 1.4.x, 2.0.x, 2.1.x, 2.2.x Powered By Apache Kafka

Fixed Versions:

- **Full Fix:** CDK 4.0.0 and higher Powered By Apache Kafka
- **Partial Fix:** CDK 3.0.0 and higher Powered By Apache Kafka are less likely to encounter this issue.

Workaround: To move forward after seeing this problem, restart the Kafka brokers affected. You can restart individual brokers from the **Instances** tab in the Kafka service page in Cloudera Manager.



Note: If restarting the brokers does not resolve the problem, you might not have this issue; see [KAFKA-3083 A soft failure in controller may leave a topic partition in an inconsistent state](#). This problem also involves the ZooKeeper session expiring, but will not involve the error message with `Cached zkVersion [XX] not equal to that in zookeeper`.

To reduce the chances of this issue happening again, do what you can to make sure ZooKeeper sessions do not expire:

- Reduce the potential for long garbage collection pauses by brokers:

- Use a better garbage collection mechanism in the JVM, such as G1GC. You can do this by adding `-XX:+UseG1GC` in the `broker_java_opts`.
- Increase broker heap size if it is too small (`broker_max_heap_size`) (be careful that you don't choose a heap size that can cause out-of-memory problems given all the services running on the node).
- Increase the ZooKeeper session timeout configuration on brokers (`zookeeper.session.timeout.ms`), to reduce the likelihood that sessions expire.
- Ensure ZooKeeper itself is well resourced and not overwhelmed, so it can respond. For example, it is highly recommended to locate the ZooKeeper log directory is on its own disk.

Cloudera JIRA: CDH-42514

Apache JIRA: KAFKA-2729

Upstream Issues Fixed

The following upstream issues are fixed in CDH 4.0.0:

- [KAFKA-3514](#) - Stream timestamp computation needs some further thoughts
- [KAFKA-4514](#) - Add Codec for ZStandard Compression
- [KAFKA-4932](#) - Add support for UUID serialization and deserialization
- [KAFKA-5066](#) - Add KafkaMetricsConfig (Yammer metrics reporters) props to documentation
- [KAFKA-5690](#) - Add support to list ACLs for a given principal
- [KAFKA-5891](#) - Proper handling of LogicalTypes in Cast
- [KAFKA-5975](#) - No response when deleting topics and `delete.topic.enable=false`
- [KAFKA-6123](#) - Give client MetricsReporter auto-generated client.id
- [KAFKA-6195](#) - Resolve DNS aliases in `bootstrap.server` (KIP-235)
- [KAFKA-6684](#) - Support casting Connect values with bytes schema to string
- [KAFKA-6753](#) - Updating the OfflinePartitions count only when necessary
- [KAFKA-6761](#) - Reduce streams footprint part IV add optimization
- [KAFKA-6835](#) - Enable topic unclean leader election to be enabled without controller change
- [KAFKA-6859](#) - Do not send LeaderEpochRequest for undefined leader epochs
- [KAFKA-6863](#) - Kafka clients should try to use multiple DNS resolved IP
- [KAFKA-6914](#) - Set parent classloader of DelegatingClassLoader same as the worker's
- [KAFKA-6923](#) - Refactor Serializer/Deserializer for KIP-336
- [KAFKA-6926](#) - Simplified some logic to eliminate some suppressions of NPath complexity checks
- [KAFKA-6950](#) - Delay response to failed client authentication to prevent potential DoS issues (KIP-306)
- [KAFKA-6998](#) - Disable Caching when `max.cache.bytes` are zero.
- [KAFKA-7019](#) - Make reading metadata lock-free by maintaining an atomically-updated read snapshot
- [KAFKA-7044](#) - Fix `Fetcher.fetchOffsetsByTimes` and NPE in describe consumer group
- [KAFKA-7080](#) - and KAFKA-7222: Cleanup overlapping KIP changes
- [KAFKA-7096](#) - Clear buffered data for partitions that are explicitly unassigned by user
- [KAFKA-7117](#) - Support AdminClient API in `AclCommand` (KIP-332)
- [KAFKA-7128](#) - ; Follower has to catch up to offset within current leader epoch to join ISR
- [KAFKA-7134](#) - `KafkaLog4jAppender` exception handling with `ignoreExceptions`
- [KAFKA-7139](#) - Support option to exclude the internal topics in `kafka-topics.sh`
- [KAFKA-7169](#) - Validate SASL extensions through callback on server side
- [KAFKA-7196](#) - Remove heartbeat delayed operation for those removed consumers at the end of each rebalance
- [KAFKA-7211](#) - ; MM should handle `TimeoutException` in `commitSync`
- [KAFKA-7215](#) - Improve LogCleaner Error Handling
- [KAFKA-7216](#) - Ignore unknown `ResourceTypes` while loading acl cache
- [KAFKA-7222](#) - Add Windows grace period
- [KAFKA-7223](#) - In-Memory Suppression Buffering
- [KAFKA-7240](#) - -total metrics in Streams are incorrect

- [KAFKA-7242](#) - Reverse xform configs before saving
- [KAFKA-7277](#) - Migrate Streams API to Duration instead of longMs times
- [KAFKA-7278](#) - replaceSegments() should not call asyncDeleteSegment() for segments which have been removed from segments list
- [KAFKA-7280](#) - Synchronize consumer fetch request/response handling
- [KAFKA-7285](#) - Create new producer on each rebalance if EOS enabled
- [KAFKA-7286](#) - Avoid getting stuck loading large metadata records
- [KAFKA-7287](#) - Set open ACL for old consumer znode path
- [KAFKA-7288](#) - Fix for SslSelectorTest.testCloseConnectionInClosingState
- [KAFKA-7296](#) - Handle coordinator loading error in TxnOffsetCommit
- [KAFKA-7298](#) - Raise UnknownProducerIdException if next sequence number is unknown
- [KAFKA-7299](#) - Batch LeaderAndIsr requests for AutoLeaderRebalance
- [KAFKA-7301](#) - Fix streams Scala join ambiguous overload
- [KAFKA-7311](#) - Reset next batch expiry time on each poll loop
- [KAFKA-7313](#) - StopReplicaRequest should attempt to remove future replica for the partition only if future replica exists
- [KAFKA-7316](#) - Fix Streams Scala filter recursive call #5538
- [KAFKA-7322](#) - Fix race condition between log cleaner thread and log retention thread when topic cleanup policy is updated
- [KAFKA-7324](#) - NPE due to lack of SASLExtensions in SASL/OAUTHBEARER
- [KAFKA-7326](#) - KStream.print() should flush on each line for PrintStream
- [KAFKA-7332](#) - Update CORRUPT_MESSAGE exception message description
- [KAFKA-7333](#) - Protocol changes for KIP-320
- [KAFKA-7338](#) - Specify AES128 default encryption type for Kerberos tests
- [KAFKA-7347](#) - Return not leader error for OffsetsForLeaderEpoch requests to non-replicas
- [KAFKA-7353](#) - Connect logs 'this' for anonymous inner classes
- [KAFKA-7354](#) - Fix IdlePercent and NetworkProcessorAvgIdlePercent metric
- [KAFKA-7366](#) - Make topic configs segment.bytes and segment.ms to take effect immediately
- [KAFKA-7369](#) - Handle retrievable errors in AdminClient list groups API
- [KAFKA-7379](#) - [streams] send.buffer.bytes should be allowed to set -1 in KafkaStreams
- [KAFKA-7385](#) - Fix log cleaner behavior when only empty batches are retained
- [KAFKA-7386](#) - streams-scala should not cache serdes
- [KAFKA-7388](#) - equal sign in property value for password
- [KAFKA-7394](#) - OffsetsForLeaderEpoch supports topic describe access
- [KAFKA-7395](#) - Add fencing to replication protocol (KIP-320)
- [KAFKA-7396](#) - Materialized, Serialized, Joined, Consumed and Produced with implicit Serdes
- [KAFKA-7399](#) - KIP-366, Make FunctionConversions deprecated
- [KAFKA-7400](#) - Compacted topic segments that precede the log start offse...
- [KAFKA-7403](#) - Use default timestamp if no expire timestamp set in offset commit value
- [KAFKA-7406](#) - Name join group repartition topics
- [KAFKA-7409](#) - Validate message format version before creating topics or altering configs
- [KAFKA-7414](#) - Out of range errors should never be fatal for follower
- [KAFKA-7415](#) - Persist leader epoch and start offset on becoming a leader
- [KAFKA-7428](#) - ConnectionStressSpec: add "action", allow multiple clients
- [KAFKA-7429](#) - Enable key/truststore update with same filename/password
- [KAFKA-7434](#) - Fix NPE in DeadLetterQueueReporter
- [KAFKA-7437](#) - Persist leader epoch in offset commit metadata
- [KAFKA-7439](#) - Replace EasyMock and PowerMock with Mockito in clients module
- [KAFKA-7441](#) - Allow LogCleanerManager.resumeCleaning() to be used concurrently
- [KAFKA-7453](#) - Expire registered channels not selected within idle timeout

- [KAFKA-7454](#) - Use lazy allocation for SslTransportLayer buffers and null them on close
- [KAFKA-7456](#) - Serde Inheritance in DSL
- [KAFKA-7459](#) - Use thread-safe Pool for RequestMetrics.requestRateInternal
- [KAFKA-7460](#) - Fix Connect Values converter date format pattern
- [KAFKA-7462](#) - Make token optional for OAuthBearerLoginModule
- [KAFKA-7467](#) - NoSuchElementException is raised because controlBatch is empty
- [KAFKA-7475](#) - capture remote address on connection authentication errors, and log it
- [KAFKA-7476](#) - Fix Date-based types in SchemaProjector
- [KAFKA-7477](#) - Improve Streams close timeout semantics
- [KAFKA-7481](#) - Add upgrade/downgrade notes for 2.1.x
- [KAFKA-7482](#) - LeaderAndIsrRequest should be sent to the shutting down broker
- [KAFKA-7483](#) - Allow streams to pass headers through Serializer.
- [KAFKA-7496](#) - Handle invalid filters gracefully in KafkaAdminClient#describeAcls
- [KAFKA-7498](#) - Remove references from `common.requests` to `clients`
- [KAFKA-7501](#) - Fix producer batch double deallocation when receiving message too large error on expired batch
- [KAFKA-7505](#) - Process incoming bytes on write error to report SSL failures
- [KAFKA-7519](#) - Clear pending transaction state when expiration fails
- [KAFKA-7532](#) - Clean-up controller log when shutting down brokers
- [KAFKA-7534](#) - Error in flush calling close may prevent underlying store from closing
- [KAFKA-7535](#) - KafkaConsumer doesn't report records-lag if isolation.level is read_committed
- [KAFKA-7560](#) - PushHttpMetricsReporter should not convert metric value to double
- [KAFKA-7742](#) - Fixed removing hmac entry for a token being removed from DelegationTokenCache

Issues Fixed in CDK 3.1.1 Powered By Apache Kafka

CDK 3.1.1 Powered By Apache Kafka fixes the following issues:

Shell wrapper script for `kafka-configs` fails to execute

Running the `kafka-configs` tool returns a `No such file or directory` error message.

Workaround: Call the script directly with the following command:

```
/opt/cloudera/parcels/KAFKA/lib/kafka/bin/kafka-configs.sh
```

Affected Versions: CDK 3.0.0 and 3.1.0 Powered by Apache Kafka

Fixed Versions: CDK 3.1.1 Powered by Apache Kafka

Cloudera Issue: CDH-61121

Upstream Issues Fixed

The following upstream issues are fixed in CDH 3.1.1:

- [KAFKA-3978](#) - Ensure high watermark is always positive
- [KAFKA-6593](#) - Fix livelock with consumer heartbeat thread in `commitSync`
- [KAFKA-6857](#) - Leader should reply with undefined offset if undefined leader epoch requested
- [KAFKA-6917](#) - Process txn completion asynchronously to avoid deadlock
- [KAFKA-6975](#) - Fix replica fetching from non-batch-aligned log start offset
- [KAFKA-7012](#) - Don't process SSL channels without data to process
- [KAFKA-7104](#) - More consistent leader's state in fetch response
- [KAFKA-7278](#) - `replaceSegments()` should not call `asyncDeleteSegment()` for segments which have been removed from segments list

Issues Fixed in CDK 3.1.0 Powered By Apache Kafka

CDK 3.1.0 Powered By Apache Kafka fixes the following issues:

Authenticated Kafka clients may impersonate other users

Authenticated Kafka clients may impersonate any other user via a manually crafted protocol message with SASL/PLAIN or SASL/SCRAM authentication when using the built-in PLAIN or SCRAM server implementations in Apache Kafka.

Note that the SASL authentication mechanisms that apply to this issue are neither recommended nor supported by Cloudera. In Cloudera Manager (CM) there are [four choices](#): PLAINTEXT, SSL, SASL_PLAINTEXT, and SASL_SSL. The SASL/PLAIN option described in this issue is not the same as SASL_PLAINTEXT option in CM. That option uses Kerberos and is not affected. As a result it is highly unlikely that Kafka is susceptible to this issue when managed by CM unless the authentication protocol is overridden by an Advanced Configuration Snippet (Safety Valve).

Products affected: CDK Powered by Apache Kafka

Releases affected: CDK 2.1.0 to 2.2.0, CDK 3.0

Users affected: All users

Detected by: Rajini Sivaram (rsivaram@apache.org)

Severity (Low/Medium/High):8.3 (High) ([CVSS:3.0/AV:N/AC:L/PR:L/UI:N/S:U/C:L/I:H/A:H](#))

Impact:Privilege escalation.

CVE:CVE-2017-12610

Immediate action required: Upgrade to a newer version of CDK Powered by Apache Kafka where the issue has been fixed.

Addressed in release/refresh/patch: CDK 3.1, CDH 6.0 and higher

Knowledge article: For the latest update on this issue see the corresponding Knowledge article: [TSB 2018-332: Two Kafka Security Vulnerabilities: Authenticated Kafka clients may impersonate other users and and may interfere with data replication](#)

Authenticated clients may interfere with data replication

Authenticated Kafka users may perform an action reserved for the Broker via a manually created fetch request interfering with data replication, resulting in data loss.

Products affected: CDK Powered by Apache Kafka

Releases affected: CDK 2.0.0 to 2.2.0, CDK 3.0.0

Users affected: All users

Detected by: Rajini Sivaram (rsivaram@apache.org)

Severity (Low/Medium/High):6.3 (Medium) ([CVSS:3.0/AV:N/AC:L/PR:L/UI:N/S:U/C:L/I:L/A:L](#))

Impact:Potential data loss due to improper replication.

CVE:CVE-2018-1288

Immediate action required: Upgrade to a newer version of CDK Powered by Apache Kafka where the issue has been fixed.

Addressed in release/refresh/patch: CDK 3.1, CDH 6.0 and higher

Knowledge article: For the latest update on this issue see the corresponding Knowledge article: [TSB 2018-332: Two Kafka Security Vulnerabilities: Authenticated Kafka clients may impersonate other users and and may interfere with data replication](#)

Upstream Issues Fixed

The following upstream issues are fixed in CDK 3.1.0 Powered By Apache Kafka:

- [KAFKA-6739](#): Down-conversion fails for records with headers.
- [KAFKA-6185](#): Selector memory leak with high likelihood of OOM in case of down conversion.
- [KAFKA-6134](#): High memory usage on controller during partition reassignment.
- [KAFKA-6119](#): Silent Data Loss in Kafka011 Transactional Producer.
- [KAFKA-6116](#): Major performance issue due to excessive logging during leader election.
- [KAFKA-6093](#): Replica dir not deleted after topic deletion.
- [KAFKA-6042](#): Kafka Request Handler deadlocks and brings down the cluster..
- [KAFKA-6026](#): KafkaFuture timeout fails to fire if a narrow race condition is hit.
- [KAFKA-6015](#): NPE in RecordAccumulator.
- [KAFKA-6012](#): NoSuchElementException in markErrorMeter during TransactionsBounceTest.
- [KAFKA-6004](#): Enable custom authentication plugins to return error messages to clients.
- [KAFKA-6003](#): Replication Fetcher thread for a partition with no data fails to start.
- [KAFKA-5987](#): Kafka metrics templates used in document generation should maintain order of tags.
- [KAFKA-5970](#): Deadlock due to locking of DelayedProduce and group.
- [KAFKA-5960](#): Producer uses unsupported ProduceRequest version against older brokers.
- [KAFKA-5959](#): NPE in NetworkClient.
- [KAFKA-5957](#): Producer IllegalStateException due to second deallocate after aborting a batch.
- [KAFKA-5879](#): Controller should read the latest IsrChangeNotification znodes when handling IsrChangeNotification event.
- [KAFKA-5829](#): Speedup broker startup after unclean shutdown by reducing unnecessary snapshot files deletion.
- [KAFKA-5790](#): SocketServer.processNewResponses should not skip a response if exception is thrown.
- [KAFKA-5767](#): Kafka server should halt if IBP < 1.0.0 and there is log directory failure.
- [KAFKA-5752](#): Delete topic and re-create topic immediate will delete the new topic's timeindex.
- [KAFKA-5708](#): Update Jackson dependencies (from 2.8.5 to 2.9.x).
- [KAFKA-5630](#): Consumer poll loop over the same record after a CorruptRecordException.
- [KAFKA-5610](#): KafkaApis.handleWriteTxnMarkerRequest can return UNSUPPORTED_FOR_MESSAGE_FORMAT error on partition emigration.
- [KAFKA-5600](#): Group loading regression causing stale metadata/offsets cache.
- [KAFKA-5556](#): KafkaConsumer.commitSync throws IllegalStateException: Attempt to retrieve exception from future which hasn't failed.
- [KAFKA-5417](#): Clients get inconsistent connection states when SASL/SSL connection is marked CONNECTED and DISCONNECTED at the same time.
- [KAFKA-4669](#): KafkaProducer.flush hangs when NetworkClient.handleCompletedReceives throws exception.

Issues Fixed in CDK 3.0.0 Powered By Apache Kafka

- [KAFKA-5506](#): Add system test for connector failure/restartFix NPE in OffsetFetchRequest.toString and logging improvements.
- [KAFKA-5522](#): ListOffsets should bound timestamp search by LSO in read_committed.
- [KAFKA-5556](#): Fix IllegalStateException in KafkaConsumer.commitSync due to missing future completion check.
- [KAFKA-5584](#): Fix integer overflow in Log.size.
- [KAFKA-5611](#): AbstractCoordinator should handle wakeup raised from onJoinComplete.
- [KAFKA-5630](#): Consumer should block on corrupt records and keep throwing an exception.
- [KAFKA-5634](#): Do not allow segment deletion beyond high watermark.
- [KAFKA-5658](#): Fix AdminClient request timeout handling bug resulting in continual BrokerNotAvailableExceptions.
- [KAFKA-5700](#): Producer should not drop header information when splitting batches.
- [KAFKA-5737](#): KafkaAdminClient thread should be daemon.
- [KAFKA-5752](#): Update index files correctly during async delete.

Issues Fixed in CDK 2.2.0 Powered by Apache Kafka

- [KAFKA-4525](#): Kafka should not require SSL truststore password

- [KAFKA-4811](#); ReplicaFetchThread may fail to create due to existing metric
- [KAFKA-4741](#); Fix potential buffer leak in RecordAccumulator in case of exception
- [KAFKA-4735](#); Fix deadlock issue during MM shutdown
- [KAFKA-4636](#); Per listener security settings overrides (KIP-103)
- [KAFKA-5150](#); Reduce lz4 decompression overhead
- [KAFKA-5316](#); LogCleaner should account for larger record sets after cleaning
- [KAFKA-5097](#); Fix regression in consumer caused by unsafe access to potentially unassigned partitions
- [KAFKA-4959](#); Remove controller concurrent access to non-threadsafe NetworkClient, Selector, and SSLEngine
- [KAFKA-4631](#); Request metadata in consumer if topic/partitions unavailable

Issues Fixed in CDK 2.1.2 Powered By Apache Kafka

- [KAFKA-3863](#): Add system test for connector failure/restart.
- [KAFKA-3994](#): Deadlock between consumer heartbeat expiration and offset commit.

Issues Fixed in CDK 2.1.1 Powered By Apache Kafka

- [KAFKA-724](#): Allow automatic socket.send.buffer from operating system in SocketServer
- [KAFKA-2684](#): Add force option to topic / config command so they can be called programatically
- [KAFKA-2720](#): Expire group metadata when all offsets have expired
- [KAFKA-2948](#): Remove unused topics from producer metadata set
- [KAFKA-3111](#): Fix ConsumerPerformance reporting to use time-based instead of message-based intervals
- [KAFKA-3158](#): ConsumerGroupCommand should tell whether group is actually dead
- [KAFKA-3175](#): Topic not accessible after deletion even when delete.topic.enable is disabled
- [KAFKA-3501](#): Console consumer process hangs on exit
- [KAFKA-3562](#): Handle topic deletion during a send
- [KAFKA-3645](#): Fix ConsumerGroupCommand and ConsumerOffsetChecker to correctly read endpoint info from ZK
- [KAFKA-3716](#): Validate all timestamps are not negative
- [KAFKA-3719](#): Allow underscores in hostname
- [KAFKA-3748](#): Add consumer-property to console tools consumer
- [KAFKA-3774](#): Make 'time' an optional argument of GetOffsetShell
- [KAFKA-3810](#): replication of internal topics should not be limited by replica.fetch.max.bytes
- [KAFKA-3934](#): Start scripts enable GC by default with no way to disable
- [KAFKA-3965](#): MirrorMaker should not commit offset when exception is thrown from producer.send
- [KAFKA-4158](#): Reset quota to default value if quota override is deleted
- [KAFKA-4229](#): Controller can't start after several zk expired event
- [KAFKA-4319](#): Parallelize shutdown of fetchers in AbstractFetcherManager to speedup shutdown
- [KAFKA-4360](#): Controller may deadLock when autoLeaderRebalance encounter ZK expired
- [KAFKA-4428](#): Kafka does not exit if port is already bound

Issues Fixed in CDK 2.1.0 Powered By Apache Kafka

- [KAFKA-3787](#): Preserve the message timestamp in MirrorMaker
- [KAFKA-3789](#): Upgrade Snappy to fix Snappy decompression errors
- [KAFKA-3802](#): Log mtimes reset on broker restart or shutdown
- [KAFKA-3894](#): Log cleaner can partially clean a segment
- [KAFKA-3915](#): Do not convert messages from v0 to v1 during log compaction
- [KAFKA-3933](#): Always fully read deepliterator
- [KAFKA-3950](#): Only throw authorization exception if pattern subscription matches topic
- [KAFKA-3977](#): Defer fetch parsing for space efficiency and to ensure exceptions are raised to the user
- [KAFKA-4050](#): Allow configuration of the PRNG used for SSL
- [KAFKA-4073](#): MirrorMaker should handle messages without timestamp correctly

Cloudera Distribution of Apache Kafka 2.1.0 is rebased on Apache Kafka 0.10. For a complete list of fixed issues, see https://www.apache.org/dyn/closer.cgi?path=/kafka/0.10.0.0/RELEASE_NOTES.html.

Issues Fixed in CDK 2.0.2 Powered By Apache Kafka

- [KAFKA-3495](#): `NetworkClient.blockingSendAndReceive` should rely on `requestTimeout`.
- [KAFKA-2998](#): Log warnings when client is disconnected from bootstrap brokers.
- [KAFKA-3488](#): Avoid failing of unsent requests in consumer where possible.
- [KAFKA-3528](#): Handle wakeups while rebalancing more gracefully.
- [KAFKA-3594](#): After calling `MemoryRecords.close()` method, `hasRoomFor()` method should return `false`.
- [KAFKA-3602](#): Rename `RecordAccumulator dequeFor()` and ensure proper usage.
- [KAFKA-3789](#): Upgrade Snappy to fix Snappy decompression errors.
- [KAFKA-3830](#): `getTGT()` debug logging exposes confidential information.
- [KAFKA-3840](#): Allow clients default OS buffer sizes.
- [KAFKA-3691](#): Confusing logging during metadata update timeout.
- [KAFKA-3810](#): Replication of internal topics should not be limited by `replica.fetch.max.bytes`.
- [KAFKA-3854](#): Fix issues with new consumer's subsequent regex (pattern) subscriptions.

Issues Fixed in CDK 2.0.1 Powered By Apache Kafka

- [KAFKA-3409](#): `MirrorMaker` hangs indefinitely due to commit.
- [KAFKA-3378](#): Client blocks forever if `SocketChannel` connects instantly.
- [KAFKA-3426](#): Improve protocol type errors when invalid sizes are received.
- [KAFKA-3330](#): Truncate log cleaner offset checkpoint if the log is truncated.
- [KAFKA-3463](#): Change default receive buffer size for consumer to 64K.
- [KAFKA-1148](#): Delayed fetch/producer requests should be satisfied on a leader change.
- [KAFKA-3352](#): Avoid DNS reverse lookups.
- [KAFKA-3341](#): Improve error handling on invalid requests.
- [KAFKA-3310](#): Fix for NPEs observed when throttling clients.
- [KAFKA-2784](#): swallow exceptions when `MirrorMaker` exits.
- [KAFKA-3243](#): Fix Kafka basic ops documentation for `MirrorMaker`, blacklist is not supported for new consumers.
- [KAFKA-3235](#): Unclosed stream in `AppInfoParser` static block.
- [KAFKA-3147](#): Memory records is not writable in `MirrorMaker`.
- [KAFKA-3088](#): Broker crash on receipt of produce request with empty client ID.
- [KAFKA-3159](#): Kafka consumer client poll is very CPU intensive under certain conditions.
- [KAFKA-3189](#): Kafka server returns `UnknownServerException` for inherited exceptions.
- [KAFKA-3157](#): `MirrorMaker` does not commit offset with low traffic.
- [KAFKA-3179](#): Kafka consumer delivers message whose offset is earlier than sought offset.
- [KAFKA-3198](#): Ticket Renewal Thread exits prematurely due to inverted comparison.

Issues Fixed in CDK 2.0.0 Powered By Apache Kafka

- [KAFKA-2799](#): `WakeupException` thrown in the followup poll() could lead to data loss.
- [KAFKA-2878](#): Kafka broker throws `OutOfMemory` exception with invalid join group request.
- [KAFKA-2880](#): `Fetcher.getTopicMetadata` `NullPointerException` when broker cannot be reached.
- [KAFKA-2882](#): Add constructor cache for Snappy and LZ4 Output/Input streams in `Compressor.java`
- [KAFKA-2913](#): `GroupMetadataManager` unloads all groups in `removeGroupsForPartitions`.
- [KAFKA-2942](#): Inadvertent auto-commit when pre-fetching can cause message loss.
- [KAFKA-2950](#): Fix performance regression in the producer.
- [KAFKA-2973](#): Fix leak of child sensors on remove.
- [KAFKA-2978](#): Consumer stops fetching when consumed and fetch positions get out of sync.
- [KAFKA-2988](#): Change default configuration of the log cleaner.

- [KAFKA-3012](#): Avoid reserved.broker.max.id collisions on upgrade.

Issues Fixed in CDK 1.4.0 Powered By Apache Kafka

- [KAFKA-1664](#): Kafka does not properly parse multiple ZK nodes with non-root chroot.
- [KAFKA-1994](#): Evaluate performance effect of chroot check on Topic creation.
- [KAFKA-2002](#): It does not work when kafka_mx4jenable is false.
- [KAFKA-2024](#): Cleaner can generate unindexable log segments.
- [KAFKA-2048](#): java.lang.IllegalMonitorStateException thrown in AbstractFetcherThread when handling error returned from simpleConsumer.
- [KAFKA-2050](#): Avoid calling .size() on java.util.ConcurrentLinkedQueue.
- [KAFKA-2088](#): kafka-console-consumer.sh should not create zookeeper path when no brokers found and chroot was set in zookeeper.connect.
- [KAFKA-2118](#): Cleaner cannot clean after shutdown during replaceSegments.
- [KAFKA-2477](#): Fix a race condition between log append and fetch that causes OffsetOutOfRangeException.
- [KAFKA-2633](#): Default logging from tools to Stderr.

Issues Fixed in CDK 1.3.2 Powered By Apache Kafka

- [KAFKA-1057](#): Trim whitespaces from user specified configs
- [KAFKA-1641](#): Log cleaner exits if last cleaned offset is lower than earliest offset
- [KAFKA-1702](#): Messages silently lost by the (old) producer
- [KAFKA-1758](#): corrupt recovery file prevents startup
- [KAFKA-1836](#): metadata.fetch.timeout.ms set to zero blocks forever
- [KAFKA-1866](#): LogStartOffset gauge throws exceptions after log.delete()
- [KAFKA-1883](#): NullPointerException in RequestSendThread
- [KAFKA-1896](#): Record size function of record in mirror maker hit NPE when the message value is null.
- [KAFKA-2012](#): Broker should automatically handle corrupt index files
- [KAFKA-2096](#): Enable keepalive socket option for broker to prevent socket leak
- [KAFKA-2114](#): Unable to set default min.insync.replicas
- [KAFKA-2189](#): Snappy compression of message batches less efficient in 0.8.2.1
- [KAFKA-2234](#): Partition reassignment of a nonexistent topic prevents future reassignments
- [KAFKA-2235](#): LogCleaner offset map overflow
- [KAFKA-2336](#): Changing offsets.topic.num.partitions after the offset topic is created breaks consumer group partition assignment
- [KAFKA-2393](#): Correctly Handle InvalidTopicException in KafkaApis.getTopicMetadata()
- [KAFKA-2406](#): ISR propagation should be throttled to avoid overwhelming controller
- [KAFKA-2407](#): Only create a log directory when it will be used
- [KAFKA-2437](#): Fix ZookeeperLeaderElector to handle node deletion correctly
- [KAFKA-2468](#): SIGINT during Kafka server startup can leave server deadlocked
- [KAFKA-2504](#): Stop logging WARN when client disconnects

Issues Fixed in CDK 1.3.1 Powered By Apache Kafka

- [KAFKA-972](#) - MetadataRequest returns stale list of brokers
- [KAFKA-1367](#) - Broker topic metadata not kept in sync with ZooKeeper
- [KAFKA-1867](#) - liveBroker list not updated on a cluster with no topics
- [KAFKA-2308](#) - New producer + Snappy face un-compression errors after broker restart
- [KAFKA-2317](#) - De-register isrChangeNotificationListener on controller resignation
- [KAFKA-2337](#) - Verify that metric names will not collide when creating new topics

Issues Fixed in CDK 1.3.0 Powered By Apache Kafka

- [KAFKA-2009](#) - Fix UncheckedOffset.removeOffset synchronization and trace logging issue in mirror maker
- [KAFKA-1984](#) - Java producer may miss an available partition
- [KAFKA-1971](#) - Starting a broker with a conflicting id will delete the previous broker registration
- [KAFKA-1952](#) - High CPU Usage in 0.8.2 release
- [KAFKA-1919](#) - Metadata request issued with no backoff in new producer if there are no topics

Issues Fixed in CDK 1.2.0 Powered By Apache Kafka

- [KAFKA-1642](#) - [Java New Producer Kafka Trunk] CPU Usage Spike to 100% when network connection is lost
- [KAFKA-1650](#) - avoid data loss when mirror maker shutdown uncleanly
- [KAFKA-1797](#) - add the serializer/deserializer api to the new java client -
- [KAFKA-1667](#) - topic-level configuration not validated
- [KAFKA-1815](#) - ServerShutdownTest fails in trunk
- [KAFKA-1861](#) - Publishing kafka-client:test in order to utilize the helper utils in TestUtils
- [KAFKA-1729](#) - Add constructor to javaapi to allow constructing explicitly versioned offset commit requests
- [KAFKA-1902](#) - fix MetricName so that Yammer reporter can work correctly
- [KAFKA-1890](#) - Fix bug preventing Mirror Maker from successful rebalance
- [KAFKA-1891](#) - MirrorMaker hides consumer exception - making troubleshooting challenging
- [KAFKA-1706](#) - Add a byte bounded blocking queue utility
- [KAFKA-1879](#) - Log warning when receiving produce requests with acks > 1
- [KAFKA-1876](#) - pom file for scala 2.11 should reference a specific version
- [KAFKA-1761](#) - num.partitions documented default is 1 while actual default is 2
- [KAFKA-1210](#) - Windows Bat files are not working properly
- [KAFKA-1864](#) - Revisit defaults for the internal offsets topic
- [KAFKA-1870](#) - Cannot commit with simpleConsumer on Zookeeper only with Java API
- [KAFKA-1868](#) - ConsoleConsumer shouldn't override dual.commit.enabled to false if not explicitly set
- [KAFKA-1841](#) - OffsetCommitRequest API - timestamp field is not versioned
- [KAFKA-1723](#) - make the metrics name in new producer more standard
- [KAFKA-1819](#) - Cleaner gets confused about deleted and re-created topics
- [KAFKA-1851](#) - OffsetFetchRequest returns extra partitions when input only contains unknown partitions
- [KAFKA-1512](#) - Fixes for limit the maximum number of connections per ip address
- [KAFKA-1624](#) - bump up default scala version to 2.11.4 to compile with java 8
- [KAFKA-742](#) - Existing directories under the Kafka data directory without any data cause process to not start
- [KAFKA-1698](#) - Validator.ensureValid() only validates default config value
- [KAFKA-1799](#) - ProducerConfig.METRIC_REPORTER_CLASSES_CONFIG doesn't work
- [KAFKA-1743](#) - ConsumerConnector.commitOffsets in 0.8.2 is not backward compatible
- [KAFKA-1769](#) - javadoc should only include client facing packages
- [KAFKA-1481](#) - Stop using dashes AND underscores as separators in MBean names
- [KAFKA-1721](#) - Snappy compressor is not thread safe
- [KAFKA-1764](#) - ZookeeperConsumerConnector should not put multiple shutdown commands to the same data chunk queue
- [KAFKA-1733](#) - Producer.send will block indeterminately when broker is unavailable
- [KAFKA-1742](#) - ControllerContext removeTopic does not correctly update state
- [KAFKA-1738](#) - Partitions for topic not created after restart from forced shutdown
- [KAFKA-1647](#) - Replication offset checkpoints (high water marks) can be lost on hard kills and restarts
- [KAFKA-1732](#) - DumpLogSegments tool fails when path has a '.'

Known Issues in CDK Powered By Apache Kafka

The following sections describe known issues in CDK Powered By Apache Kafka:

Unsupported features

- CDK Powered by Apache Kafka supports Java based clients only. Clients developed with C, C++, Python, .NET and other languages are currently not supported.
- Kafka Connect is included with CDK 2.0.0 and higher Powered By Apache Kafka, but is not supported at this time. Instead, we recommend Flume and Sqoop as proven solutions for batch and real time data loading that complement Kafka's message broker capability. See [Flafka: Apache Flume Meets Apache Kafka for Event Processing](#) and for more information.

In addition, Spark and Spark Streaming can be used to get the functionality of "Kafka Streaming" and have a fully functional ETL or stream processing pipeline. See [Using Apache Kafka with Apache Spark Streaming](#) on page 61 for more information.

- The Kafka default authorizer is included with CDK 2.0.0 and higher Powered By Apache Kafka, but is not supported at this time. This includes setting ACLs and all related APIs, broker functionality, and command-line tools.

Kafka does not work with Apache Sentry HA

For CDK 3.0.0 and earlier Powered By Apache Kafka, you cannot use Sentry high availability with Kafka. If Sentry HA is enabled, Kafka might intermittently lose the connection to Sentry and you won't be able to authorize users.

Affected Versions: All versions of CDK Powered By Apache Kafka with CDH 5.13.x and 5.14.x

Fixed Versions: CDK 3.1.0

Cloudera JIRA: CDH-56519

Topics created with the kafka-topics tool may not be secured

Topics that are created and deleted via Kafka are secured (for example, auto created topics). However, most topic creation and deletion is done via the kafka-topics tool, which talks directly to Zookeeper or some other third-party tool that talks directly to Zookeeper. Since this is the responsibility of Zookeeper authorization and authentication, Kafka cannot prevent users from making Zookeeper changes. Anyone with access to Zookeeper can create and delete topics. Note that they will not be able to describe, read, or write to the topics even if they can create them.

The following commands talk directly to Zookeeper and therefore are not secured via Kafka:

- kafka-topics.sh
- kafka-configs.sh
- kafka-preferred-replica-election.sh
- kafka-reassign-partitions.sh

Kafka Garbage Collection Logs are Written to the Process Directory

By default Kafka garbage collection logs are written to the CDH process directory. Changing the default path for these log files is currently unsupported.

Workaround: N/A

Affected Versions:All

Fixed Versions: N/A

Cloudera Issue: OPSAPS-43236

Replication Factor in Kafka Streams is set to 1 by Default

In Kafka Streams the `replication.factor` Streams configuration parameter is set to 1 by default. In other words, the internal topics that the Streams application creates are not replicated by default. Without replication, even a single broker failure can prevent progress of the stream processing application which in turn can lead to data being lost. .

Workaround: Set the `replication.factor` Streams configuration parameter to a value higher than 1. Cloudera recommends that the replication factor set for the Streams application should be identical to the replication factor of the source topics. For more information, see [Configuring a Streams Application](#)

Affected Versions: CDK 4.0.0 or higher

Fixed Versions: N/A

Cloudera Issue: N/A

Apache Issue: N/A

Topic-level metrics do not display in Cloudera Manager for topics that contain a period (.) in the topic name

If you have Kafka topics that contain a period (.) in the topic name, Cloudera Manager might not display the topic-level metrics for those topics in the Charts Library. Only topic-level metrics are affected.

Affected Versions: CDK 3.0.0 Powered By Apache Kafka

Fixed Versions: CDK 3.1.0

Cloudera JIRA: CDH-64370

`offsets.topic.replication.factor` must be less than or equal to the number of live brokers (CDK 3.0.0 Powered By Apache Kafka)

In CDK 3.0.0 Powered By Apache Kafka, the `offsets.topic.replication.factor` broker config is now enforced upon auto topic creation. Internal auto topic creation will fail with a `GROUP_COORDINATOR_NOT_AVAILABLE` error until the cluster size meets this replication factor requirement.

Kafka stuck with under-replicated partitions after ZooKeeper session expires

This problem might occur when your Kafka cluster includes a large number of under-replicated Kafka partitions. One or more broker logs include messages such as the following:

```
[2016-01-17 03:36:00,888] INFO Partition [__samza_checkpoint_event-creation_1,3] on broker 3: Shrinking ISR for partition [__samza_checkpoint_event-creation_1,3] from 6,5 to 5 (kafka.cluster.Partition)
[2016-01-17 03:36:00,891] INFO Partition [__samza_checkpoint_event-creation_1,3] on broker 3: Cached zkVersion [66] not equal to that in zookeeper, skip updating ISR (kafka.cluster.Partition)
```

There will also be an indication of the ZooKeeper session expiring in one or more Kafka broker logs around the same time as the previous errors:

```
INFO zookeeper state changed (Expired) (org.I0Itec.zkclient.ZkClient)
```

The log is typically in `/var/log/kafka` on each host where a Kafka broker is running. The location is set by the property `kafka.log4j.dir` in Cloudera Manager. The log name is `kafka-broker-hostname.log`. In diagnostic bundles, the log is under `logs/hostname-ip-address/`.

Affected Versions: CDK 1.4.x, 2.0.x, 2.1.x, 2.2.x Powered By Apache Kafka

Fixed Versions:

- **Full Fix:** CDK 4.0.0 and higher Powered By Apache Kafka
- **Partial Fix:** CDK 3.0.0 and higher Powered By Apache Kafka are less likely to encounter this issue.

Workaround: To move forward after seeing this problem, restart the Kafka brokers affected. You can restart individual brokers from the **Instances** tab in the Kafka service page in Cloudera Manager.



Note: If restarting the brokers does not resolve the problem, you might not have this issue; see [KAFKA-3083 A soft failure in controller may leave a topic partition in an inconsistent state](#). This problem also involves the ZooKeeper session expiring, but will not involve the error message with `Cached zkVersion [XX] not equal to that in zookeeper`.

To reduce the chances of this issue happening again, do what you can to make sure ZooKeeper sessions do not expire:

- Reduce the potential for long garbage collection pauses by brokers:
 - Use a better garbage collection mechanism in the JVM, such as G1GC. You can do this by adding `-XX:+UseG1GC` in the `broker_java_opts`.
 - Increase broker heap size if it is too small (`broker_max_heap_size`) (be careful that you don't choose a heap size that can cause out-of-memory problems given all the services running on the node).
- Increase the ZooKeeper session timeout configuration on brokers (`zookeeper.session.timeout.ms`), to reduce the likelihood that sessions expire.
- Ensure ZooKeeper itself is well resourced and not overwhelmed, so it can respond. For example, it is highly recommended to locate the ZooKeeper log directory is on its own disk.

Cloudera JIRA: CDH-42514

Apache JIRA: KAFKA-2729

Kafka client jars included in CDH might not match the newest Kafka parcel jar

The Kafka client jars included in CDH may not match the newest Kafka parcel jar that is released. This is done to maintain compatibility across CDH 5.7 and higher for integrations such as Spark and Flume.

The Flume and Spark connectors to Kafka shipped with CDH 5.7 and higher only work with CDK 2.x Powered By Apache Kafka

Use CDK 2.x and higher Powered By Apache Kafka to be compatible with the Flume and Spark connectors included with CDH 5.7.x.

Only new Java clients support authentication and authorization

The legacy Scala clients (producer and consumer) that are under the `kafka.producer.*` and `kafka.consumer.*` package do not support authentication.

Workaround: Migrate to the new Java producer and consumer APIs.

Requests fail when sending to a nonexistent topic with `auto.create.topics.enable` set to `true`

The first few produce requests fail when sending to a nonexistent topic with `auto.create.topics.enable` set to `true`.

Affected Versions: All

Workaround: Increase the number of `retries` in the Producer configuration settings.

Custom Kerberos principal names must not be used for Kerberized ZooKeeper and Kafka instances

When using ZooKeeper authentication and a custom Kerberos principal, Kerberos-enabled Kafka does not start.

Affected Versions: CDK 2.0.0 and higher Powered By Apache Kafka

Workaround: None. You must disable ZooKeeper authentication for Kafka or use the default Kerberos principals for ZooKeeper and Kafka.

Performance degradation when SSL is enabled

Significant performance degradation can occur when SSL is enabled. The impact varies, depending on your CPU type and JVM version. The reduction is generally in the range 20-50%. Consumers are typically more affected than producers.

Affected Versions: CDK 2.0.0 and higher Powered By Apache Kafka

Workaround for CDK 2.1.0 and higher Powered By Apache Kafka: Configure brokers and clients with `ssl.secure.random.implementation = SHA1PRNG`. It often reduces this degradation drastically, but its effect is CPU and JVM dependent.

AdminUtils is not binary-compatible between CDK 1.x and 2.x Powered By Apache Kafka

The AdminUtils APIs have changed between CDK 1.x and 2.x Powered By Apache Kafka. If your application uses AdminUtils APIs, you must modify your application code to use the new APIs before you compile your application against CDK 2.x Powered By Apache Kafka.



Note: AdminUtils APIs are not part of the publicly supported CDK Powered By Apache Kafka API.

Source cluster not definable in CDK 1.x Powered By Apache Kafka

In CDK 1.x Powered By Apache Kafka, the source cluster is assumed to be the cluster that MirrorMaker is running on. In CDK 2.0 Powered By Apache Kafka, you can define a custom source and target cluster.

Monitoring is not supported in Cloudera Manager 5.4

If you use CDK 1.2 Powered By Apache Kafka with Cloudera Manager 5.4, you must disable monitoring.

MirrorMaker does not start when Sentry is enabled

When MirrorMaker is used in conjunction with Sentry, MirrorMaker reports an authorization issue and does not start. This is due to Sentry being unable to authorize the `kafka_mirror_maker` principal which is automatically created.

Workaround: Complete the following steps prior to enabling Sentry:

1. Create the `kafka_mirror_maker` Linux user ID and the `kafka_mirror_maker` Linux group ID on the MirrorMaker hosts. Use the following command:

```
useradd kafka_mirror_maker
```

2. Create the necessary Sentry rules for the `kafka_mirror_maker` group.



Note: Alternatively, you can add the `kafka_mirror_maker` user to `super.users`, this bypasses authorization.

Affected Versions: CDK 2.1.1 and higher Powered by Apache Kafka

Fixed Versions: N/A

Apache JIRA: N/A

Cloudera JIRA: CDH-53706

Authenticated Kafka clients may impersonate other users

Authenticated Kafka clients may impersonate any other user via a manually crafted protocol message with SASL/PLAIN or SASL/SCRAM authentication when using the built-in PLAIN or SCRAM server implementations in Apache Kafka. Note that the SASL authentication mechanisms that apply to this issue are neither recommended nor supported by Cloudera. In Cloudera Manager (CM) there are [four choices](#): PLAINTEXT, SSL, SASL_PLAINTEXT, and SASL_SSL. The SASL/PLAIN option described in this issue is not the same as SASL_PLAINTEXT option in CM. That option uses Kerberos and is not affected. As a result it is highly unlikely that Kafka is susceptible to this issue when managed by CM unless the authentication protocol is overridden by an Advanced Configuration Snippet (Safety Valve).

Products affected: CDK Powered by Apache Kafka

Releases affected: CDK 2.1.0 to 2.2.0, CDK 3.0

Users affected: All users

Detected by: Rajini Sivaram (rsivaram@apache.org)

Severity (Low/Medium/High): 8.3 (High) (CVSS:3.0/AV:N/AC:L/PR:L/UI:N/S:U/C:L/I:H/A:H)

Impact: Privilege escalation.

CVE: CVE-2017-12610

Immediate action required: Upgrade to a newer version of CDK Powered by Apache Kafka where the issue has been fixed.

Addressed in release/refresh/patch: CDK 3.1, CDH 6.0 and higher

Knowledge article: For the latest update on this issue see the corresponding Knowledge article: [TSB 2018-332: Two Kafka Security Vulnerabilities: Authenticated Kafka clients may impersonate other users and and may interfere with data replication](#)

Authenticated clients may interfere with data replication

Authenticated Kafka users may perform an action reserved for the Broker via a manually created fetch request interfering with data replication, resulting in data loss.

Products affected: CDK Powered by Apache Kafka

Releases affected: CDK 2.0.0 to 2.2.0, CDK 3.0.0

Users affected: All users

Detected by: Rajini Sivaram (rsivaram@apache.org)

Severity (Low/Medium/High): 6.3 (Medium) (CVSS:3.0/AV:N/AC:L/PR:L/UI:N/S:U/C:L/I:L/A:L)

Impact: Potential data loss due to improper replication.

CVE: CVE-2018-1288

Immediate action required: Upgrade to a newer version of CDK Powered by Apache Kafka where the issue has been fixed.

Addressed in release/refresh/patch: CDK 3.1, CDH 6.0 and higher

Knowledge article: For the latest update on this issue see the corresponding Knowledge article: [TSB 2018-332: Two Kafka Security Vulnerabilities: Authenticated Kafka clients may impersonate other users and and may interfere with data replication](#)

The Idempotent and Transactional Capabilities of Kafka are Incompatible with Sentry

The idempotent and transactional capabilities of Kafka are not compatible with Sentry. The issue is due to Sentry being unable to handle authorization policies for Kafka transactions. As a result, users cannot use Kafka transactions in combination with Sentry.

Workaround: Either disable Sentry, or use the Kafka superuser in applications where idempotent producing is a requirement.

Affected Versions: CDK 4.0.0 and later

Fixed Versions: N/A

Apache Issue: N/A

Cloudera Issue: CDH-80606

CDK Powered By Apache Kafka Incompatible Changes and Limitations

This section describes incompatible changes and limitations:



Warning: The open file handlers of CDK 2.2.0 Powered By Apache Kafka will increase by roughly 33% because of the addition of time index files for each segment.

Incompatible Changes and Limitations in CDK 4.1.0 Powered By Apache Kafka

Default Consumer Group ID Change

The default consumer group ID has been changed from the empty string (" ") to `null`. Consumers that use the new default group ID will not be able to subscribe to topics, and fetch or commit offsets. The empty string as consumer group ID is deprecated but will be supported until a future major release. Old clients that rely on the empty string group id will now have to explicitly provide it as part of their consumer configuration. For more information, see [KIP-289](#).

Incompatible Changes and Limitations in CDK 4.0.0 Powered By Apache Kafka



Warning: CDK Powered by Apache Kafka 4.0.0 is based on Apache Kafka 2.1.0, which contains a change to the internal schema used to store consumer offsets. As a result of this change, downgrading CDK Powered by Apache Kafka to a version lower than CDK 4.0.0 is **NOT** possible once Kafka has been upgraded to CDK 4.0.0 or higher.

Scala-based Clients API Removed

Scala-based clients were deprecated in a previous release and are removed as of CDK 4.0.0. The following Scala-based client implementations from package `kafka.*` (known as 'old clients') are affected:

- `kafka.consumer.*`
- `kafka.producer.*`
- `kafka.admin.*`

Client applications making use of these implementations must be migrated to corresponding Java clients available in `org.apache.kafka.*` (the 'new clients') package. Existing command line options and tools now use the new clients package.

Properties for Exactly Once Semantics Not Available in Cloudera Manager

The configuration properties related the idempotent and transactional capabilities of the producer are not available for configuration via Cloudera Manager. These properties must be set through the **Kafka Broker Advanced Configuration Snippet (Safety Valve) for `kafka.properties`** safety valve. For more information regarding configuration using safety valves, see [Custom Configuration](#).

The following are the configuration properties related to the idempotent and transactional capabilities of the producer:

- **Broker Properties**

- `transactional.id.expiration.ms`
- `transaction.max.timeout.ms`
- `transaction.state.log.replication.factor`
- `transaction.state.log.num.partitions`
- `transaction.state.log.min.isr`
- `transaction.state.log.segment.bytes`

- **Producer Properties**

- `enable.idempotence`
- `transaction.timeout.ms`
- `transactional.id`

- **Consumer Properties**

- `isolation.level`

For more information, see the [upstream Apache Kafka documentation](#).

Default Behaviour Changes in CDK 4.0.0 Powered by Apache Kafka

Kafka CDK 4.0.0. Introduces the following default behaviour changes:

- Unclean leader election is automatically enabled by the controller when `unclean.leader.election.enable.config` is dynamically updated by using per-topic config override.
- The default value for `request.timeout.ms` is decreased to 30 seconds. In addition, a new logic is added that makes the `JoinGroup` requests ignore this timeout.

Incompatible Changes and Limitations in CDK 3.1.0 Powered By Apache Kafka

Scala-based Clients API Deprecated

Scala-based clients are deprecated in this release and will be removed in an upcoming release. The following Scala-based client implementations from package `kafka.*` (known as 'old clients') are affected:

- `kafka.consumer.*`
- `kafka.producer.*`
- `kafka.admin.*`

Client applications making use of these implementations must be migrated to corresponding Java clients available in `org.apache.kafka.*` (the 'new clients') package. Existing command line options and tools now use the new clients package.

Incompatible Changes and Limitations in CDK 3.0.0 Powered By Apache Kafka

CDK 3.0 Requires CDH 5.13 when Co-located

Using version 3.0 and later of CDK Powered by Apache Kafka requires a newer version of Cloudera Manager and/or CDH when Kafka and CDH are in the same logical cluster in Cloudera Manager. For more information on compatibilities among versions, see [Product Compatibility Matrix for CDK Powered By Apache Kafka](#).

Incompatible Changes and Limitations in CDK 2.0.0 Powered By Apache Kafka

Flume shipped with CDH 5.7 and lower can only send data to CDK 2.0 and higher Powered By Apache Kafka via unsecured transport.

Security additions to CDK 2.0 Powered By Apache Kafka are not supported by Flume in CDH 5.7 (or lower versions).

Topic Blacklist Removed

The MirrorMaker **Topic blacklist** setting has been removed in CDK 2.0 and higher Powered By Apache Kafka.

Avoid Data Loss Option Removed

The **Avoid Data Loss** option from earlier releases has been removed in CDK 2.0 Powered By Apache Kafka in favor of automatically setting the following properties.

1. Producer settings

- `acks=all`
- `retries=max integer`
- `max.block.ms=max long`

2. Consumer setting

- `auto.commit.enable=false`

3. MirrorMaker setting

- `abort.on.send.failute=true`

CDK Powered By Apache Kafka® Version and Packaging Information

This section describes naming conventions for CDK Powered By Apache Kafka® package versions, lists versions and where to download components.

For installation instructions, see [Installing, Migrating and Upgrading CDK Powered By Apache Kafka](#) on page 48.

Examples of CDK Powered By Apache Kafka Versions

Cloudera packages are designed to be transparent and easy to understand. CDK Powered By Apache Kafka package versions are labeled using the following format:

```
base_version+cloudera_version+patch_level
```

where:

- `base_version` is the version of the open-source component included in the Cloudera package.
- `cloudera_version` is the version of the Cloudera package.
- `patch_level` is the number of source commits applied on top of the base version forked from the Apache Kafka branch. Note that the number of commits does not indicate the number of functional changes or bug fixes in the release. For example, a commit can be used to amend a version number or make other non-functional changes.

CDK Powered By Apache Kafka Versions

Table 1: CDK Powered By Apache Kafka Version Information

| CDK Powered By Apache Kafka Version | Component | Version | Release Date | Release Notes | Parcel Repository |
|-------------------------------------|--------------|-------------------------|-------------------|-------------------------------|---|
| 4.1.0 | Apache Kafka | 2.2.1-kafka4.1.0+14 | July 11, 2019 | Release notes | CDK Powered By Apache Kafka 4.1.0 Parcel Repository |
| 4.0.0 | Apache Kafka | 2.1.0+kafka4.0.0+17 | April 8, 2019 | Release notes | CDK Powered By Apache Kafka 4.0.0 Parcel Repository |
| 3.1.1 | Apache Kafka | 1.0.1+kafka3.1.1+3 | November 19, 2018 | Release notes | CDK Powered By Apache Kafka 3.1.1 Parcel Repository |
| 3.1.0 | Apache Kafka | 1.0.1+kafka3.1.0+35 | June 7, 2018 | Release notes | CDK Powered By Apache Kafka 3.1.0 Parcel Repository |
| 3.0.0 | Apache Kafka | 0.11.0+kafka3.0.0+50 | October 16, 2017 | Release notes | CDK Powered By Apache Kafka 3.0.0 Parcel Repository |
| 2.2.0 | Apache Kafka | 0.10.2.0+kafka2.2.0+110 | July 13, 2017 | Release notes | CDK Powered By Apache Kafka 2.2.0 Parcel Repository |
| 2.1.2 | Apache Kafka | 0.10.0.1+kafka2.1.2+6 | October 4, 2017 | Release notes | CDK Powered By Apache Kafka 2.1.2 Parcel Repository |
| 2.1.1 | Apache Kafka | 0.10.0.0+kafka2.1.1+21 | March 31, 2017 | Release notes | CDK Powered By Apache Kafka 2.1.1 Parcel Repository |
| 2.1.0 | Apache Kafka | 0.10.0.0+kafka2.1.0+63 | January 31, 2017 | Release notes | CDK Powered By Apache Kafka 2.1.0 Parcel Repository |

| CDK Powered By Apache Kafka Version | Component | Version | Release Date | Release Notes | Parcel Repository |
|-------------------------------------|--------------|------------------------|-------------------|-------------------------------|---|
| 2.0.2 | Apache Kafka | 0.9.0.0+kafka2.0.2+305 | July 22, 2016 | Release notes | CDK Powered By Apache Kafka 2.0.2 Parcel Repository |
| 2.0.1 | Apache Kafka | 0.9.0.0+kafka2.0.1+283 | April 7, 2016 | Release notes | CDK Powered By Apache Kafka 2.0.1 Parcel Repository |
| 2.0.0 | Apache Kafka | 0.9.0.0+kafka2.0.0+188 | February 19, 2016 | Release notes | CDK Powered By Apache Kafka 2.0.0 Parcel Repository |
| 1.4.0 | Apache Kafka | 0.8.2.0+kafka1.4.0+127 | December 10, 2015 | Release notes | CDK Powered By Apache Kafka 1.4.0 Parcel Repository |
| 1.3.2 | Apache Kafka | 0.8.2.0+kafka1.3.2+116 | October 8, 2015 | Release notes | CDK Powered By Apache Kafka 1.3.2 Parcel Repository |
| 1.3.1 | Apache Kafka | 0.8.2.0+kafka1.3.1+80 | August 3, 2015 | Release notes | CDK Powered By Apache Kafka 1.3.1 Parcel Repository |
| 1.3.0 | Apache Kafka | 0.8.2.0+kafka1.3.0+72 | April 23, 2015 | Release notes | CDK Powered By Apache Kafka 1.3.0 Parcel Repository |
| 1.2.0 | Apache Kafka | 0.8.2.0+kafka1.2.0+57 | February 18, 2015 | Release notes | CDK Powered By Apache Kafka 1.2.0 Parcel Repository |

Table 2: Compatible Release Versions for CDK Powered By Apache Kafka 2.1.0

| Type | Location | Parcel File |
|--------------------------|---|---|
| yum RHEL/CentOS/Oracle 7 | http://archive.cloudera.com/kafka/redhat/7/x86_64/kafka/ | http://archive.cloudera.com/kafka/parcels/2.1.0/KAFKA-2.1.0-1.2.1.0.p0.115-el7.parcel |
| yum RHEL/CentOS/Oracle 6 | http://archive.cloudera.com/kafka/redhat/6/x86_64/kafka/ | http://archive.cloudera.com/kafka/parcels/2.1.0/KAFKA-2.1.0-1.2.1.0.p0.115-el6.parcel |
| yum RHEL/CentOS/Oracle 5 | http://archive.cloudera.com/kafka/redhat/5/x86_64/kafka/ | http://archive.cloudera.com/kafka/parcels/2.1.0/KAFKA-2.1.0-1.2.1.0.p0.115-el5.parcel |
| Debian Jesse 8.2 | https://archive.cloudera.com/kafka/debian/jessie/amd64/kafka/dists | http://archive.cloudera.com/kafka/parcels/2.1.0/KAFKA-2.1.0-1.2.1.0.p0.115-jessie.parcel |
| apt Debian Wheezy 7.0 | https://archive.cloudera.com/kafka/debian/wheezy/amd64/kafka/dists | http://archive.cloudera.com/kafka/parcels/2.1.0/KAFKA-2.1.0-1.2.1.0.p0.115-wheezy.parcel |
| zypper/YaST SLES 12 | https://archive.cloudera.com/kafka/sles/12/x86_64/kafka/ | http://archive.cloudera.com/kafka/parcels/2.1.0/KAFKA-2.1.0-1.2.1.0.p0.115-sles12.parcel |
| zypper/YaST SLES 11 | https://archive.cloudera.com/kafka/sles/11/x86_64/kafka/ | http://archive.cloudera.com/kafka/parcels/2.1.0/KAFKA-2.1.0-1.2.1.0.p0.115-sles11.parcel |
| Ubuntu Trusty 14.04 | https://archive.cloudera.com/kafka/ubuntu/trusty/amd64/kafka/dists | http://archive.cloudera.com/kafka/parcels/2.1.0/KAFKA-2.1.0-1.2.1.0.p0.115-trusty.parcel |

| Type | Location | Parcel File |
|----------------------|---|---|
| Ubuntu Precise 12.04 | https://archive.cloudera.com/kafka/ubuntu/precise/amd64/kafka/dists | http://archive.cloudera.com/kafka/parcels/2.1.0/KAFKA-2.1.0-1.2.1.0.p0.115-precise.parcel |

Table 3: Compatible Release Versions for CDK Powered By Apache Kafka 1.4.0

| Type | Location | Parcel File |
|--------------------------|---|---|
| yum RHEL/CentOS/Oracle 7 | http://archive.cloudera.com/kafka/redhat/7/x86_64/kafka/ | http://archive.cloudera.com/kafka/parcels/1.4.0/KAFKA-0.8.2.0-1.kafka1.4.0.p0.56-el5.parcel |
| yum RHEL/CentOS/Oracle 6 | http://archive.cloudera.com/kafka/redhat/6/x86_64/kafka/ | http://archive.cloudera.com/kafka/parcels/1.4.0/KAFKA-0.8.2.0-1.kafka1.4.0.p0.56-el6.parcel |
| yum RHEL/CentOS/Oracle 5 | http://archive.cloudera.com/kafka/redhat/5/x86_64/kafka/ | http://archive.cloudera.com/kafka/parcels/1.4.0/KAFKA-0.8.2.0-1.kafka1.4.0.p0.56-el7.parcel |
| apt Debian Wheezy 7.0 | https://archive.cloudera.com/kafka/debian/wheezy/amd64/kafka/ | http://archive.cloudera.com/kafka/parcels/1.4.0/KAFKA-0.8.2.0-1.kafka1.4.0.p0.56-wheezy.parcel |
| zypper/YaST SLES | https://archive.cloudera.com/kafka/sles/11/x86_64/kafka/ | http://archive.cloudera.com/kafka/parcels/1.4.0/KAFKA-0.8.2.0-1.kafka1.4.0.p0.56-sles11.parcel |
| Ubuntu Trusty 14.04 | https://archive.cloudera.com/kafka/ubuntu/trusty/amd64/kafka/ | http://archive.cloudera.com/kafka/parcels/1.4.0/KAFKA-0.8.2.0-1.kafka1.4.0.p0.56-trusty.parcel |
| Ubuntu Precise 12.04 | https://archive.cloudera.com/kafka/ubuntu/precise/amd64/kafka/ | http://archive.cloudera.com/kafka/parcels/1.4.0/KAFKA-0.8.2.0-1.kafka1.4.0.p0.56-precise.parcel |

Maven Artifacts for CDK Powered By Apache Kafka

The following tables lists the project name, groupId, artifactId, and version required to access each Kafka artifact from a Maven POM. For information on how to use Kafka Maven artifacts, see [Using the CDH 5 Maven Repository](#).

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 4.1.0 artifact.

Table 4: Maven Artifacts for CDK Powered By Apache Kafka 4.1.0

| Project | groupId | artifactId | version |
|--------------|------------------|------------------------------|-------------------|
| Apache Kafka | org.apache.kafka | connect-api | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | connect-basic-auth-extension | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | connect-file | 2.2.1-kafka-4.1.0 |

| Project | groupId | artifactId | version |
|---------|------------------|--------------------------|-------------------|
| | org.apache.kafka | connect-json | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | connect-runtime | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | connect-transforms | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka-clients | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka-examples | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka-log4j-appender | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka-streams | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka-streams-examples | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka-streams-scala_2.11 | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka-streams-scala_2.12 | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka-streams-test-utils | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka-tools | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka_2.11 | 2.2.1-kafka-4.1.0 |
| | org.apache.kafka | kafka_2.12 | 2.2.1-kafka-4.1.0 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 4.0.0 artifact.

Table 5: Maven Artifacts for CDK Powered By Apache Kafka 4.0.0

| Project | groupId | artifactId | version |
|--------------|------------------|------------------------------|-------------------|
| Apache Kafka | org.apache.kafka | connect-api | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | connect-basic-auth-extension | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | connect-file | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | connect-json | 2.1.0-kafka-4.0.0 |

| Project | groupId | artifactId | version |
|---------|------------------|--------------------------|-------------------|
| | org.apache.kafka | connect-runtime | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | connect-transforms | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka-clients | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka-examples | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka-log4j-appender | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka-streams | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka-streams-examples | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka-streams-scala_2.11 | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka-streams-scala_2.12 | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka-streams-test-utils | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka-tools | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka_2.11 | 2.1.0-kafka-4.0.0 |
| | org.apache.kafka | kafka_2.12 | 2.1.0-kafka-4.0.0 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 3.1.1 artifact.

Table 6: Maven Artifacts for CDK Powered By Apache Kafka 3.1.1

| Project | groupId | artifactId | version |
|--------------|------------------|--------------------|-------------------|
| Apache Kafka | org.apache.kafka | connect-api | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | connect-file | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | connect-json | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | connect-runtime | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | connect-transforms | 1.0.1-kafka-3.1.1 |

| Project | groupId | artifactId | version |
|---------|------------------|------------------------|-------------------|
| | org.apache.kafka | kafka-clients | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | kafka-examples | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | kafka-log4j-appender | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | kafka-streams | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | kafka-streams-examples | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | kafka-tools | 1.0.1-kafka-3.1.1 |
| | org.apache.kafka | kafka_2.11 | 1.0.1-kafka-3.1.1 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 3.1.0 artifact.

Table 7: Maven Artifacts for CDK Powered By Apache Kafka 3.1.0

| Project | groupId | artifactId | version |
|---------|------------------|------------------------|-------------------|
| Kafka | org.apache.kafka | connect-api | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | connect-file | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | connect-json | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | connect-runtime | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | connect-transforms | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | kafka-clients | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | kafka-examples | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | kafka-log4j-appender | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | kafka-streams | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | kafka-streams-examples | 1.0.1-kafka-3.1.0 |
| | org.apache.kafka | kafka-tools | 1.0.1-kafka-3.1.0 |

| Project | groupId | artifactId | version |
|---------|------------------|------------|-------------------|
| | org.apache.kafka | kafka_2.11 | 1.0.1-kafka-3.1.0 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 3.0.0 artifact.

Table 8: Maven Artifacts for CDK Powered By Apache Kafka 3.0.0

| Project | groupId | artifactId | version |
|---------|------------------|------------------------|--------------------|
| Kafka | org.apache.kafka | connect | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | connect-api | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | connect-file | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | connect-json | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | connect-runtime | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | connect-transforms | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | kafka-clients | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | kafka-examples | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | kafka-log4j-appender | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | kafka-streams | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | kafka-streams-examples | 0.11.0-kafka-3.0.0 |
| | org.apache.kafka | kafka-tools | 0.11.0-kafka-3.0.0 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 2.2.0 artifact.

Table 9: Maven Artifacts for CDK Powered By Apache Kafka 2.2.0

| Project | groupId | artifactId | version |
|---------|------------------|--------------|--------------------|
| Kafka | org.apache.kafka | connect | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | connect-api | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | connect-file | 0.10.2-kafka-2.2.0 |

| Project | groupId | artifactId | version |
|---------|------------------|------------------------|--------------------|
| | org.apache.kafka | connect-json | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | connect-runtime | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | connect-transforms | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | kafka-clients | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | kafka-examples | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | kafka-log4j-appender | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | kafka-streams | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | kafka-streams-examples | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | kafka-tools | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | kafka_2.10 | 0.10.2-kafka-2.2.0 |
| | org.apache.kafka | kafka_2.11 | 0.10.2-kafka-2.2.0 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 2.1.2 artifact.

Table 10: Maven Artifacts for CDK Powered By Apache Kafka 2.1.2

| Project | groupId | artifactId | version |
|---------|------------------|-----------------|--------------------|
| Kafka | org.apache.kafka | connect | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | connect-api | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | connect-file | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | connect-json | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | connect-runtime | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | kafka-clients | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | kafka-examples | 0.10.0-kafka-2.1.2 |

| Project | groupId | artifactId | version |
|---------|------------------|------------------------|--------------------|
| Kafka | org.apache.kafka | kafka-log4j-appender | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | kafka-streams | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | kafka-streams-examples | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | kafka-tools | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | kafka_210 | 0.10.0-kafka-2.1.2 |
| Kafka | org.apache.kafka | kafka_2.11 | 0.10.0-kafka-2.1.2 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 2.1.1 artifact.

Table 11: Maven Artifacts for CDK Powered By Apache Kafka 2.1.1

| Project | groupId | artifactId | version |
|---------|------------------|------------------------|--------------------|
| Kafka | org.apache.kafka | connect | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | connect-api | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | connect-file | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | connect-json | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | connect-runtime | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | kafka-clients | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | kafka-examples | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | kafka-log4j-appender | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | kafka-streams | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | kafka-streams-examples | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | kafka-tools | 0.10.0-kafka-2.1.1 |
| Kafka | org.apache.kafka | kafka_2.10 | 0.10.0-kafka-2.1.1 |

| Project | groupId | artifactId | version |
|---------|------------------|------------|--------------------|
| Kafka | org.apache.kafka | kafka_2.11 | 0.10.0-kafka-2.1.1 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 2.1.0 artifact.

Table 12: Maven Artifacts for CDK Powered By Apache Kafka 2.1.0

| Project | groupId | artifactId | version |
|---------|------------------|----------------------|--------------------|
| Kafka | org.apache.kafka | connect | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | connect-api | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | connect-file | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | connect-json | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | connect-runtime | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | kafka-clients | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | kafka-examples | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | kafka-log4j-appender | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | kafka-tools | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | kafka_2.10 | 0.10.0-kafka-2.1.0 |
| Kafka | org.apache.kafka | kafka_2.11 | 0.10.0-kafka-2.1.0 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 2.0.0 artifact.

Table 13: Maven Artifacts for CDK Powered By Apache Kafka 2.0.0

| Project | groupId | artifactId | version |
|---------|------------------|-----------------|-------------------|
| Kafka | org.apache.kafka | connect | 0.9.0-kafka-2.0.0 |
| | org.apache.kafka | connect-api | 0.9.0-kafka-2.0.0 |
| | org.apache.kafka | connect-file | 0.9.0-kafka-2.0.0 |
| | org.apache.kafka | connect-json | 0.9.0-kafka-2.0.0 |
| | org.apache.kafka | connect-runtime | 0.9.0-kafka-2.0.0 |
| | org.apache.kafka | kafka-clients | 0.9.0-kafka-2.0.0 |

| Project | groupId | artifactId | version |
|---------|------------------|----------------------|-------------------|
| | org.apache.kafka | kafka-examples | 0.9.0-kafka-2.0.0 |
| | org.apache.kafka | kafka-log4j-appender | 0.9.0-kafka-2.0.0 |
| | org.apache.kafka | kafka-tools | 0.9.0-kafka-2.0.0 |
| | org.apache.kafka | kafka_2.10 | 0.9.0-kafka-2.0.0 |
| | org.apache.kafka | kafka_2.11 | 0.9.0-kafka-2.0.0 |

The following table lists the project name, groupId, artifactId, and version required to access each CDK Powered By Apache Kafka 1.4.0 artifact.

Table 14: Maven Artifacts for CDK Powered By Apache Kafka 1.4.0

| Project | groupId | artifactId | version |
|---------|------------------|----------------|---------------------|
| Kafka | org.apache.kafka | kafka-clients | 0.8.2.0-kafka-1.4.0 |
| Kafka | org.apache.kafka | kafka-examples | 0.8.2.0-kafka-1.4.0 |
| Kafka | org.apache.kafka | kafka_2.10 | 0.8.2.0-kafka-1.4.0 |

Installing, Migrating and Upgrading CDK Powered By Apache Kafka

Minimum Required Role: **Cluster Administrator** (also provided by **Full Administrator**)



Important: As of February 1, 2021, all downloads of CDK, CDH, and Cloudera Manager require a username and password and use a modified URL. You must use the modified URL, including the username and password when downloading the repository contents described below. You may need to upgrade Cloudera Manager to a newer version that uses the modified URLs.

This can affect new installations, upgrades, adding new hosts to a cluster, downloading a new parcel, and adding a new cluster.

For more information, see [Updating an existing CDH/Cloudera Manager deployment to access downloads with authentication.](#)

The steps required to install or upgrade Kafka vary based on the version of Cloudera Manager you are using. This section describes several possible installation and upgrade scenarios. Before you install, review the Release Notes, particularly:

- [What's New in CDK Powered By Apache Kafka?](#) on page 11
- [CDK Powered By Apache Kafka Requirements and Supported Versions](#) on page 14
- [Known Issues in CDK Powered By Apache Kafka](#) on page 29

In addition, make sure to also review the [Kafka Product Compatibility Matrix](#).

Installing or Upgrading CDK Powered By Apache Kafka®

Minimum Required Role: **Cluster Administrator** (also provided by **Full Administrator**)

Kafka is distributed as a parcel, separate from the CDH parcel. It is also distributed as a package. The steps to install Kafka vary, depending on whether you choose to install from a parcel or a package.

General Information Regarding Installation and Upgrade

Cloudera Manager 5.4 and higher includes the Kafka service. To install, download Kafka using Cloudera Manager, distribute Kafka to the cluster, activate the new parcel, and add the service to the cluster. For a list of available parcels and packages, see [CDK Powered By Apache Kafka® Version and Packaging Information](#) on page 37

Colocation of Kafka and ZooKeeper services on the same host is possible. However, for optimal performance, Cloudera recommends the usage of dedicated hosts. This is especially true for larger, production environments.



Note: Upgrade instructions assume you want to upgrade parcel-based Kafka with parcels or package-based Kafka with packages. If you want to switch to using parcel-based Kafka using a Kafka package, you first must uninstall parcel-based Kafka. See [Uninstalling an Add-on Service](#).



Warning: CDK Powered by Apache Kafka 4.0.0 is based on Apache Kafka 2.1.0, which contains a change to the internal schema used to store consumer offsets. As a result of this change, downgrading CDK Powered by Apache Kafka to a version lower than CDK 4.0.0 is **NOT** possible once Kafka has been upgraded to CDK 4.0.0 or higher.

Rolling Upgrade to CDK 4.1.x Powered By Apache Kafka®

Before upgrading from CDK 2.x.x or CDK 3.x.x to CDK 4.1.x, ensure that you set `inter.broker.protocol.version` and `log.message.format.version` to the current Kafka version, and then unset them after the upgrade. This is a

good practice because the newer broker versions might write log entries that the older brokers will not be able to read. And if you need to rollback to the older version, and you have not set `inter.broker.protocol.version` and `log.message.format.version`, data loss might occur.

Based on the current version of Kafka, use the following three-digit values to set `inter.broker.protocol.version` and `log.message.format.version`:

- To upgrade from CDK 2.0.x Powered By Apache Kafka, use 0.9.0
- To upgrade from CDK 2.1.x Powered By Apache Kafka, use 0.10.0
- To upgrade from CDK 2.2.x Powered By Apache Kafka, use 0.10.2
- To upgrade from CDK 3.0.x Powered By Apache Kafka, use 0.11.0
- To upgrade from CDK 3.1.x Powered By Apache Kafka, use 1.0.1
- To upgrade from CDK 4.0.x Powered By Apache Kafka, use 2.1.0

From the Cloudera Manager Admin Console:

1. Upgrade Kafka brokers to CDK 4.1.x Powered By Apache Kafka.

a. Update `server.properties` file on all brokers with the following properties:

`inter.broker.protocol.version = <current_Kafka_version>` and
`log.message.format.version = <current_Kafka_version>`, as follows:

b. From the **Clusters** menu, select the Kafka cluster.

c. Click the **Configuration** tab.

d. Use the Search field to find the **Kafka Broker Advanced Configuration Snippet (Safety Valve)** configuration property.

e. Add the following properties to the **Kafka Broker Advanced Configuration Snippet (Safety Valve)** for `kafka.properties`:

To upgrade from CDK 2.0.x to CDK 4.1.x, enter:

```
inter.broker.protocol.version=0.9.0
```

```
log.message.format.version=0.9.0
```

To upgrade from CDK 2.1.x to CDK 4.1.x, enter:

```
inter.broker.protocol.version=0.10.0
```

```
log.message.format.version=0.10.0
```

To upgrade from CDK 2.2.x to CDK 4.1.x, enter:

```
inter.broker.protocol.version=0.10.2
```

```
log.message.format.version=0.10.2
```

To upgrade from CDK 3.0.x to CDK 4.1.x, enter:

```
inter.broker.protocol.version=0.11.0
```

```
log.message.format.version=0.11.0
```

To upgrade from CDK 3.1.x to CDK 4.1.x, enter

```
inter.broker.protocol.version=1.0.1
```

```
log.message.format.version=1.0.1
```

To upgrade from CDK 4.0.x to CDK 4.1.x, enter

```
inter.broker.protocol.version=2.1.0
```

```
log.message.format.version=2.1.0
```

Make sure you enter full version identifier including trailing zeros (for example, 0.11.0). Otherwise, the following error will occur:

```
2017-12-14 14:25:47,818 FATAL kafka.Kafka$:
java.lang.IllegalArgumentException: Version `0.11` is not a valid version
    at kafka.api.ApiVersion$$anonfun$apply$1.apply(ApiVersion.scala:72)
    at kafka.api.ApiVersion$$anonfun$apply$1.apply(ApiVersion.scala:72)
    at scala.collection.MapLike$class.getOrElse(MapLike.scala:128)
```

f. Save your changes.

2. Download, distribute, and activate the new parcel. Do not restart the Kafka service, select **Activate Only** and click **OK**.
3. Perform a rolling restart. Select **Rolling Restart** or **Restart** based on the downtime that can be afforded.
4. Upgrade all CDK 2.x.x clients and CDK 3.x.x clients to CDK 4.1.x.

Upgrading Kafka clients does not have to happen at one time or even immediately. The `inter.broker.protocol.version` should remain set until all clients are upgraded.

5. After all clients are upgraded and the cluster restart is successful, remove the above settings and restart the cluster again.

Graceful Shutdown of Kafka Brokers

If the Kafka brokers do not shut down gracefully, subsequent restarts may take longer than expected. This can happen when the brokers take longer than 30 seconds to clear their backlog while stopping the Kafka service, stopping the Kafka Broker role, or stopping a cluster where the Kafka service is running. The Kafka brokers are also shut down as part of performing an upgrade. There are two configuration properties you can set to control whether Cloudera Manager waits for the brokers to shut down gracefully:

Table 15: Kafka Shutdown Properties

| Property | Description | Default Value |
|-----------------------------------|---|------------------------------------|
| Enable Controlled Shutdown | Enables controlled shutdown of the broker. If enabled, the broker moves all leaders on it to other brokers before shutting itself down. This reduces the unavailability window during shutdown. | Enabled |
| Graceful Shutdown Timeout | The timeout in milliseconds to wait for graceful shutdown to complete. | 30000 milliseconds (30 seconds) |

To configure these properties, go to **Clusters > Kafka Service > Configuration** and search for "shutdown".

If Kafka is taking a long time for controlled shutdown to complete, consider increasing the value of **Graceful Shutdown Timeout**. When this timeout is reached, Cloudera Manager issues a forced shutdown, which interrupts the controlled shutdown and could cause subsequent restarts to take longer than expected.

Disks and Filesystem

Cloudera recommends that you use multiple drives to get good throughput. To ensure good latency, do not share the same drives used for Kafka data with application logs or other OS filesystem activity. You can either use RAID to combine these drives into a single volume, or format and mount each drive as its own directory. Since Kafka has replication, RAID can also provide redundancy at the application level. This choice has several tradeoffs.

If you configure multiple data directories, partitions are assigned round-robin to data directories. Each partition is stored entirely in one of the data directories. This can lead to load imbalance between disks if data is not well balanced among partitions.

RAID can potentially do a better job of balancing load between disks because it balances load at a lower level. The primary downside of RAID is that it is usually a big performance hit for write throughput, and it reduces the available disk space.

Both RAID and JBOD setups have the ability to tolerate single disk failures. However, rebuilding the RAID array is so I/O intensive that it can effectively disable the server. As a result, improvement in availability is not as significant.

The following table summarizes these pros and cons for RAID10 versus JBOD.

| RAID10 | JBOD |
|---------------------------------|---|
| Can survive single disk failure | Can survive single disk failure |
| Single log directory | More available disk space |
| Lower total I/O | Higher write throughput |
| | Broker is not smart about balancing partitions across disk. |

Installing or Upgrading Kafka from a Parcel

Minimum Required Role: **Cluster Administrator** (also provided by **Full Administrator**)

- In Cloudera Manager, select **Hosts > Parcels**.
- If you do not see Kafka in the list of parcels, you can add the parcel to the list.
 - Find the parcel for the version of Kafka you want to use on [CDK Powered By Apache Kafka Versions](#) on page 37.
 - Copy the parcel repository link.
 - On the Cloudera Manager **Parcels** page, click **Configuration**.
 - In the field **Remote Parcel Repository URLs**, click + next to an existing parcel URL to add a new field.
 - Paste the parcel repository link.
 - Save your changes.
- On the Cloudera Manager **Parcels** page, download the Kafka parcel, distribute the parcel to the hosts in your cluster, and then activate the parcel. See [Managing Parcels](#). After you activate the Kafka parcel, Cloudera Manager prompts you to restart the cluster. You *do not* need to restart the cluster after installing Kafka. Click **Close** to ignore this prompt.
- Add the Kafka service to your cluster. See [Adding a Service](#).

Installing or Upgrading Kafka from a Package

Minimum Required Role: **Cluster Administrator** (also provided by **Full Administrator**)

You install the Kafka package from the command line.

- Navigate to the `/etc/repos.d` directory.
- Use `wget` to download the Kafka repository. See [CDK Powered By Apache Kafka® Version and Packaging Information](#) on page 37.
- Install Kafka using the appropriate commands for your operating system.

Table 16: Kafka Installation Commands

| Operating System | Commands |
|------------------|--|
| RHEL-compatible | <pre>\$ sudo yum clean all \$ sudo yum install kafka \$ sudo yum install kafka-server</pre> |
| SLES | <pre>\$ sudo zypper clean --all \$ sudo zypper install kafka \$ sudo zypper install kafka-server</pre> |
| Ubuntu or Debian | <pre>\$ sudo apt-get update \$ sudo apt-get install kafka \$ sudo apt-get install kafka-server</pre> |

4. Edit `/etc/kafka/conf/server.properties` to ensure that the `broker.id` is unique for each node and broker in Kafka cluster, and `zookeeper.connect` points to same ZooKeeper for all nodes and brokers.
5. Start the Kafka server with the following command:

```
$ sudo service kafka-server start.
```

To verify all nodes are correctly registered to the same ZooKeeper, connect to ZooKeeper using `zookeeper-client`.

```
$ zookeeper-client
$ ls /brokers/ids
```

You should see all of the IDs for the brokers you have registered in your Kafka cluster.

To discover to which node a particular ID is assigned, use the following command:

```
$ get /brokers/ids/<ID>
```

This command returns the host name of node assigned the ID you specify.

Migrating from Apache Kafka to CDK Powered By Apache Kafka

Minimum Required Role: **Cluster Administrator** (also provided by **Full Administrator**)

This topic describes the required steps to migrate an existing Apache Kafka instance to CDK Powered By Apache Kafka.

Assumptions

- You are migrating to a Kafka cluster managed by Cloudera Manager.
- You can plan a maintenance window for your migration.
- You are migrating from a compatible release version, as shown in the table below:

Table 17: Compatible Release Versions

| From Apache Kafka | To CDK Powered By Apache Kafka |
|-------------------|--------------------------------|
| 0.8.x | 1.x |

| From Apache Kafka | To CDK Powered By Apache Kafka |
|-------------------|--------------------------------|
| 0.9.x | 2.0.x |
| 0.10.0 | 2.1.x |
| 0.10.2 | 2.2.x |
| 0.11.0 | 3.0.x |
| 1.0.1 | 3.1.x |
| 2.1.0 | 4.0.x |



Note: Migration from Apache Kafka 0.7.x is not supported. If running Apache Kafka 0.7.x or earlier, you must first migrate to Apache Kafka 0.8.x or higher.

Steps for Migrating from Apache Kafka to Cloudera Distribution of Apache Kafka

Cloudera recommends the following migration procedure. You must migrate brokers first, and then clients.

Before You Begin

1. Shut down all existing producers, consumers, MirrorMaker instances, and Kafka brokers.
2. If not already installed, install Cloudera Manager. See [Installing Cloudera Manager, CDH, and Managed Services](#).
 - a. Add the CDH and Kafka parcels at installation time.
 - b. Do not add any services yet. Skip the install page by clicking the **Cloudera Manager** icon in the top navigation bar.

Step 1. Migrating Zookeeper

Kafka stores its metadata in ZooKeeper. When migrating to Cloudera Distribution of Kafka, you must also migrate your ZooKeeper instance to the supported version included with CDH.

1. Shut down your existing ZooKeeper cluster.
2. Back up your `dataDir` and `dataLogDir` by copying them to another location or machine.
3. Add the ZooKeeper service to the cluster where you will run Cloudera Kafka. See [Adding a Service](#).
4. Add the ZooKeeper role to all machines that were running ZooKeeper.
5. Set any custom configuration from your old `zoo.cfg` file in Cloudera Manager.
6. Make sure `dataDir` and `dataLogDir` match your old configuration. This is important because this is where all your data is stored.
7. Make sure the `zookeeper` user owns the files in the `dataDir` and `dataLogDir`. For example:

```
ex: chown -R zookeeper /var/lib/zookeeper
```

8. Start the new ZooKeeper service.
9. Use the `zookeeper-client` CLI to validate that data exists. You should see nodes such as *brokers*, *consumers*, and *configs*. You might need to adjust your `chroot`. For example:

```
zookeeper-client -server hostname:port
ls /
```

Step 2. Migrating Kafka Brokers

All producers, consumers, and Kafka brokers should still be shut down.

Installing, Migrating and Upgrading CDK Powered By Apache Kafka

1. Back up your `log.dirs` from the old broker machines by copying them to another location or machine.
2. Add the Kafka service to the cluster where you migrated ZooKeeper. See [Adding a Service](#).
3. Add the `broker` role to all machines that were running brokers.
4. Make sure the `kafka` user owns the `log.dirs` files. For example:

```
chown -R kafka /var/local/Kafka/data
```

5. Set any custom configuration from your old `server.properties` file in Cloudera Manager.
 - Make sure to override the `broker.id` on each node to match the configured value in your old configurations. If these values do not match, Kafka treats your brokers as new brokers and not your existing ones.
 - Make sure `log.dirs` and `zookeeper.chroot` match your old configuration. All of your data and state information is stored here.
6. Start the Kafka brokers using Cloudera Manager.

Step 3. Migrating MirrorMaker

These are the steps for migrating the MirrorMaker role. To avoid compatibility issues, migrate downstream clusters first.

1. Add the MirrorMaker role to all machines that were running MirrorMaker before.
2. Set any custom configuration from your old `producer.properties` and `consumer.properties` files in Cloudera Manager.
3. Start the MirrorMaker instances using Cloudera Manager.

Step 4. Migrating Kafka Clients

Although Kafka might function with your existing clients, you must also upgrade all of your producers and consumers to have all Cloudera patches and bug fixes, and to have a fully supported system.

Migration requires that you change your Kafka dependencies from the Apache versions to the Cloudera versions, recompile your classes, and redeploy them. Use the Maven repository locations as described in [Maven Artifacts for CDK Powered By Apache Kafka](#) on page 39.

Using Apache Kafka

This section describes ways you can use Apache Kafka tools to capture data for analysis.

Using Apache Kafka Command-line Tools

kafka-topics

The `kafka-topics` tool can be used to create, alter, list, and describe topics. For example:

```
$ kafka-topics --zookeeper zk01.example.com:2181 --list
sink1
t1
t2
$ kafka-topics --create --zookeeper hostname:2181/kafka --replication-factor 2
--partitions 4 --topic topicname
```

kafka-console-consumer

The `kafka-console-consumer` tool can be used to read data from a Kafka topic and write it to standard output. For example:

```
$ kafka-console-consumer --zookeeper zk01.example.com:2181 --topic t1
```

kafka-console-producer

The `kafka-console-producer` tool can be used to read data from standard output and write it to a Kafka topic. For example:

```
$ kafka-console-producer --broker-list kafka02.example.com:9092,kafka03.example.com:9092
--topic t1
```

kafka-consumer-groups

The `kafka-consumer-groups` tool can be used to list all consumer groups, describe a consumer group, delete consumer group info, or reset consumer group offsets.

This tool is primarily used for describing consumer groups and debugging any consumer offset issues. The output from the tool shows the log and consumer offsets for each partition connected to the consumer group that is being described. You can see at a glance which consumers are current with their partition and which ones are behind. From there, you can determine which partitions (and likely the corresponding brokers) are slow.

Using the tool on secure and unsecure clusters differs slightly. On secure clusters, you have use the `command-config` option together with an appropriate property file.

Viewing offsets on an unsecure cluster

Use the following command to view offsets committed to Kafka:

```
kafka-consumer-groups --new-consumer --bootstrap-server broker01.example.com:9092
--describe --group flume
```

Output Example:

```
GROUP      TOPIC      PARTITION  CURRENT-OFFSET  LOG-END-OFFSET  LAG      OWNER
```

```
flume t1 0 1 3 2
test-consumer-group_postamac.local-1456198719410-29ccd54f-0
```



Note: The `--new-consumer` option is **NOT** required if you are using CDK Powered by Apache Kafka 2.2.0 or later.

Viewing offsets on a secure cluster

In order to view offsets on a secure Kafka cluster, the `consumer-groups` tool has to be run with the `command-config` option. This option specifies the property file that contains the necessary configurations to run the tool on a secure cluster. The process to create property file is identical to the client configuration process detailed in [Enabling Kerberos Authentication](#) on page 69 and Step 5 in [Deploying SSL for Kafka](#) on page 65. Which process you need to follow depends on the security configuration of the cluster..

To view offsets do the following:

1. Pass the `jaas.conf` file location as a JVM parameter.

```
export KAFKA_OPTS='-Djava.security.auth.login.config=path/to/jaas.conf'
```

2. Run the tool with the `command-config` option.

```
kafka-consumer-groups --bootstrap-server host.server.com:9093 --describe --command-config client.properties
```

The `command-config` option specifies the property file that contains the necessary configurations to run the tool on a secure cluster. Which properties are configured in this file is dependent on the protocols being used.

Example `client.properties` file:

```
exclude.internal.topics=false
security.protocol = SASL_SSL
sasl.kerberos.service.name = kafka
ssl.truststore.location = /var/private/ssl/kafka.client.truststore.jks
ssl.truststore.password = test1234
```

This example shows what properties you have to set for the `consumer-groups` when both Kerberos and TLS/SSL are configured on your cluster.

kafka-reassign-partitions

This tool provides substantial control over partitions in a Kafka cluster. It is mainly used to balance storage loads across brokers through the following reassignment actions:

- Change the ordering of the partition assignment list. Used to control leader imbalances between brokers.
- Reassign partitions from one broker to another. Used to expand existing clusters.
- Reassign partitions between log directories on the same broker. Used to resolve storage load imbalance among available disks in the broker.
- Reassign partitions between log directories across multiple brokers. Used to resolve storage load imbalance across multiple brokers.

The tool uses two JSON files for input. Both of these are created by the user. The two files are the following:

- [Topics-to-Move JSON](#) on page 56
- [Reassignment Configuration JSON](#) on page 57

Topics-to-Move JSON

This JSON file specifies the topics that you want to reassign. This is a simple file that tells the `kafka-reassign-partitions` tool which partitions it should look at when generating a proposal for the reassignment configuration. The user has to create the topics-to-move JSON file from scratch.

The format of the file is the following:

```
{ "topics": [ { "topic": "mytopic1" },
               { "topic": "mytopic2" } ],
  "version": 1
}
```

Reassignment Configuration JSON

This JSON file is a configuration file that contains the parameters used in the reassignment process. This file is created by the user, however, a proposal for its contents is generated by the tool. When the `kafka-reassign-partitions` tool is executed with the `--generate` option, it generates a proposed configuration which can be fine-tuned and saved as a JSON file. The file created this way is the reassignment configuration JSON. To generate a proposal, the tool requires a `topics-to-move` file as input.

The format of the file is the following:

```
{ "version": 1,
  "partitions": [
    { "topic": "mytopic1", "partition": 3, "replicas": [4, 5], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 1, "replicas": [5, 4], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic2", "partition": 2, "replicas": [6, 5], "log_dirs": [ "any", "any" ] }
  ]
}
```

The reassignment configuration contains multiple properties that each control and specify an aspect of the configuration. The Reassignment Configuration Properties table lists each property and its description.

Table 18: Reassignment Configuration Properties

| Property | Description |
|-----------|--|
| topic | Specifies the topic. |
| partition | Specifies the partition. |
| replicas | Specifies the brokers that the selected partition is assigned to. The brokers are listed in order, which means that the first broker in the list is always the leader for that partition. Change the order of brokers to resolve any leader balancing issues among brokers. Change the broker IDs to reassign partitions to different brokers. |
| log_dirs | Specifies the log directory of the brokers. The log directories are listed in the same order as the brokers. By default <code>any</code> is specified as the log directory, which means that the broker is free to choose where it places the replica. By default, the current broker implementation selects the log directory using a round-robin algorithm. An absolute path beginning with a <code>/</code> can be used to explicitly set where to store the partition replica. |

Notes and Recommendations:

- Cloudera recommends that you minimize the volume of replica changes per command instance. Instead of moving 10 replicas with a single command, move two at a time in order to save cluster resources.
- This tool cannot be used to make an out-of-sync replica into the leader partition.
- Use this tool only when all brokers and topics are healthy.
- Anticipate system growth. Redistribute the load when the system is at 70% capacity. Waiting until redistribution becomes necessary due to reaching resource limits can make the redistribution process extremely time consuming.

Tool Usage

To reassign partitions, complete the following steps:

1. Create a topics-to-move JSON file that specifies the topics you want to reassign. Use the following format:

```
{ "topics": [ { "topic": "mytopic1" },
              { "topic": "mytopic2" } ],
  "version": 1
}
```

2. Generate the content for the reassignment configuration JSON with the following command:

```
kafka-reassign-partitions --zookeeper hostname:port --topics-to-move-json-file topics
to move.json --broker-list broker 1, broker 2 --generate
```

Running the command lists the distribution of partition replicas on your current brokers followed by a proposed partition reassignment configuration.

Example output:

```
Current partition replica assignment
{"version":1,
 "partitions":
  [ { "topic": "mytopic2", "partition": 1, "replicas": [2,3], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 0, "replicas": [1,2], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic2", "partition": 0, "replicas": [1,2], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 2, "replicas": [3,1], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 1, "replicas": [2,3], "log_dirs": [ "any", "any" ] }
  ]
}

Proposed partition reassignment configuration
{"version":1,
 "partitions":
  [ { "topic": "mytopic1", "partition": 0, "replicas": [4,5], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 2, "replicas": [4,5], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic2", "partition": 1, "replicas": [4,5], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 1, "replicas": [5,4], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic2", "partition": 0, "replicas": [5,4], "log_dirs": [ "any", "any" ] }
  ]
}
```

In this example, the tool proposed a configuration which reassigns existing partitions on broker 1, 2, and 3 to brokers 4 and 5.

3. Copy and paste the proposed partition reassignment configuration into an empty JSON file.
4. Review, and if required, modify the suggested reassignment configuration.
5. Save the file.
6. Start the redistribution process with the following command:

```
kafka-reassign-partitions --zookeeper hostname:port --reassignment-json-file reassignment
configuration.json --bootstrap-server hostname:port --execute
```



Note: Specifying a bootstrap server with the `--bootstrap-server` option is only required when an absolute log directory path is specified for a replica in the reassignment configuration JSON file.

The tool prints a list containing the original replica assignment and a message that reassignment has started.

Example output:

```
Current partition replica assignment
```

```
{
  "version":1,
  "partitions":
  [
    { "topic": "mytopic2", "partition":1, "replicas": [2,3], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition":0, "replicas": [1,2], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic2", "partition":0, "replicas": [1,2], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition":2, "replicas": [3,1], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition":1, "replicas": [2,3], "log_dirs": [ "any", "any" ] }
  ]
}
```

Save this to use as the `--reassignment-json-file` option during rollback. Successfully started reassignment of partitions.

7. Verify the status of the reassignment with the following command:

```
kafka-reassign-partitions --zookeeper hostname:port --reassignment-json-file reassignment
configuration.json --bootstrap-server hostname:port --verify
```

The tool prints the reassignment status of all partitions. Example output:

```
Status of partition reassignment:
Reassignment of partition mytopic2-1 completed successfully
Reassignment of partition mytopic1-0 completed successfully
Reassignment of partition mytopic2-0 completed successfully
Reassignment of partition mytopic1-2 completed successfully
Reassignment of partition mytopic1-1 completed successfully
```

Examples

There are multiple ways to modify the configuration file. The following list of examples shows how a user can modify a proposed configuration and what these changes do. Changes to the original example are marked in bold.

Suppose that the `kafka-reassign-partitions` tool generated the following proposed reassignment configuration:

```
{
  "version":1,
  "partitions":
  [
    { "topic": "mytopic1", "partition":0, "replicas": [1,2], "log_dirs": [ "any", "any" ] }
  ]
}
```

Reassign partitions between brokers

To reassign partitions from one broker to another, change the broker ID specified in `replicas`. For example:

```
{ "topic": "mytopic1", "partition":0, "replicas": [5,2], "log_dirs": [ "any", "any" ] }
```

This reassignment configuration moves partition `mytopic1-0` from broker 1 to broker 5.

Reassign partitions to another log directory on the same broker

To reassign partitions between log directories on the same broker, change the appropriate `any` entry to an absolute path. For example:

```
{ "topic": "mytopic1", "partition":0, "replicas": [1,2], "log_dirs": [ "/log/directory1", "any" ] }
```

This reassignment configuration moves partition `mytopic1-0` to the `/log/directory1` log directory.

Reassign partitions between log directories across multiple brokers

To reassign partitions between log directories across multiple brokers, change the broker ID specified in `replicas` and the appropriate `any` entry to an absolute path. For example:

```
{ "topic": "mytopic1", "partition":0, "replicas": [5,2], "log_dirs": [ "/log/directory1", "any" ] }
```

This reassignment configuration moves partition `mytopic1-0` to `/log/directory1` on broker 5.

Change partition assignment order (elect a new leader)

To change the ordering of the partition assignment list, change the order of the brokers in `replicas`. For example:

```
{ "topic": "mytopic1", "partition": 0, "replicas": [2,1], "log_dirs": [ "any", "any" ] }
```

This reassignment configuration elects broker 2 as the new leader.

kafka-log-dirs

The `kafka-log-dirs` tool allows user to query a list of replicas per log directory on a broker. The tool provides information that is required for optimizing replica assignment across brokers.

On successful execution, the tool prints a list of partitions per log directory for the specified topics and brokers. The list contains information on topic partition, size, offset lag, and reassignment state. Example output:

```
{
  "brokers": [
    {
      "broker": 86,
      "logDirs": [
        {
          "error": null,
          "logDir": "/var/local/kafka/data",
          "partitions": [
            {
              "isFuture": false,
              "offsetLag": 0,
              "partition": "mytopic1-2",
              "size": 0
            }
          ]
        }
      ]
    },
    ...
  ],
  "version": 1
}
```

The Contents of the `kafka-log-dirs` Output table gives an overview of the information provided by the `kafka-log-dirs` tool.

Table 19: Contents of the `kafka-log-dirs` Output

| Property | Description |
|-----------|--|
| broker | Displays the ID of the broker. |
| error | Indicates if there is a problem with the disk that hosts the topic partition. If an error is detected, <code>org.apache.kafka.common.errors.KafkaStorageException</code> is displayed. If no error is detected, the value is null. |
| logDir | Specifies the location of the log directory. Returns an absolute path. |
| isfuture | The reassignment state of the partition. This property shows whether there is currently replica movement underway between the log directories. |
| offsetLag | Displays the offset lag of the partition. |
| partition | Displays the name of the partition. |
| size | Displays the size of the partition in bytes. |

Tool Usage

To retrieve replica assignment information, run the following command:

```
kafka-log-dirs --describe --bootstrap-server hostname:port --broker-list broker 1, broker 2 --topic-list topic 1, topic 2
```



Important: On secure clusters the admin client config property file has to be specified with the `--command-config` option. Otherwise, the tool fails to execute.

If no topic is specified with the `--topic-list` option, then all topics are queried. If no broker is specified with the `--broker-list` option, then all brokers are queried. If a log directory is offline, the log directory will be marked offline in the script output. Error example:

```
"error": "org.apache.kafka.common.errors.KafkaStorageException"
```

Using Apache Kafka with Apache Spark Streaming

For information on how to configure Apache Spark Streaming to receive data from Apache Kafka, see the appropriate version of the Spark Streaming + Kafka Integration Guide: [1.6.0](#) or [2.3.0](#).

In CDH 5.7 and higher, the Spark connector to Kafka only works with Kafka 2.0 and higher.

Validating Kafka Integration with Spark Streaming

To validate your Kafka integration with Spark Streaming, run the `KafkaWordCount` example.

If you installed Spark using parcels, use the following command:

```
/opt/cloudera/parcels/CDH/lib/spark/bin/run-example streaming.KafkaWordCount <zkQuorum> <group> <topics> <numThreads>
```

If you installed Spark using packages, use the following command:

```
/usr/lib/spark/bin/run-example streaming.KafkaWordCount <zkQuorum> <group> <topics><numThreads>
```

Replace the variables as follows:

- `<zkQuorum>` - ZooKeeper quorum URI used by Kafka (for example, `zk01.example.com:2181, zk02.example.com:2181, zk03.example.com:2181`).
- `<group>` - Consumer group used by the application.
- `<topic>` - Kafka topic containing the data for the application.
- `<numThreads>` - Number of consumer threads reading the data. If this is higher than the number of partitions in the Kafka topic, some threads will be idle.



Note: If multiple applications use the same group and topic, each application receives a subset of the data.

Using Apache Kafka with Apache Flume

In CDH 5.2 and higher, Apache Flume contains an Apache Kafka source and sink. Use these to stream data from Kafka to Hadoop or from any Flume source to Kafka.

In CDH 5.7 and higher, the Flume connector to Kafka only works with Kafka 2.0 and higher.



Important: Do not configure a Kafka source to send data to a Kafka sink. If you do, the Kafka source sets the topic in the event header, overriding the sink configuration and creating an infinite loop, sending messages back and forth between the source and sink. If you need to use both a source and a sink, use an interceptor to modify the event header and set a different topic.

For information on configuring Kafka to securely communicate with Flume, see [Configuring Flume Security with Kafka](#).

Kafka Source

Use the Kafka source to stream data in Kafka topics to Hadoop. The Kafka source can be combined with any Flume sink, making it easy to write Kafka data to HDFS, HBase, and Solr.

The following Flume configuration example uses a Kafka source to send data to an HDFS sink:

```
tier1.sources = source1
tier1.channels = channel1
tier1.sinks = sink1

tier1.sources.source1.type = org.apache.flume.source.kafka.KafkaSource
tier1.sources.source1.kafka.bootstrap.servers = kafka-broker01.example.com:9092
tier1.sources.source1.kafka.topics = weblog
tier1.sources.source1.kafka.consumer.group.id = flume
tier1.sources.source1.channels = channel1
tier1.sources.source1.interceptors = i1
tier1.sources.source1.interceptors.i1.type = timestamp
tier1.sources.source1.kafka.consumer.timeout.ms = 100

tier1.channels.channel1.type = memory
tier1.channels.channel1.capacity = 10000
tier1.channels.channel1.transactionCapacity = 1000

tier1.sinks.sink1.type = hdfs
tier1.sinks.sink1.hdfs.path = /tmp/kafka/{topic}/%y-%m-%d
tier1.sinks.sink1.hdfs.rollInterval = 5
tier1.sinks.sink1.hdfs.rollSize = 0
tier1.sinks.sink1.hdfs.rollCount = 0
tier1.sinks.sink1.hdfs.fileType = DataStream
tier1.sinks.sink1.channel = channel1
```

For higher throughput, configure multiple Kafka sources to read from the same topic. If you configure all the sources with the same `kafka.consumer.group.id`, and the topic contains multiple partitions, each source reads data from a different set of partitions, improving the ingest rate.

For the list of Kafka Source properties, see [Kafka Source Properties](#).

For the full list of Kafka consumer properties, see the [Kafka documentation](#).

Tuning Notes

The Kafka source overrides two Kafka consumer parameters:

1. `auto.commit.enable` is set to `false` by the source, and every batch is committed. For improved performance, set this to `true` using the `kafka.auto.commit.enable` setting. This can lead to data loss if the source goes down before committing.
2. `consumer.timeout.ms` is set to 10, so when Flume polls Kafka for new data, it waits no more than 10 ms for the data to be available. Setting this to a higher value can reduce CPU utilization due to less frequent polling, but introduces latency in writing batches to the channel.

Kafka Sink

Use the Kafka sink to send data to Kafka from a Flume source. You can use the Kafka sink in addition to Flume sinks such as HBase or HDFS.

The following Flume configuration example uses a Kafka sink with an `exec` source:

```
tier1.sources = source1
tier1.channels = channel1
tier1.sinks = sink1

tier1.sources.source1.type = exec
tier1.sources.source1.command = /usr/bin/vmstat 1
tier1.sources.source1.channels = channel1

tier1.channels.channel1.type = memory
tier1.channels.channel1.capacity = 10000
tier1.channels.channel1.transactionCapacity = 1000

tier1.sinks.sink1.type = org.apache.flume.sink.kafka.KafkaSink
tier1.sinks.sink1.topic = sink1
tier1.sinks.sink1.brokerList = kafka01.example.com:9092,kafka02.example.com:9092
tier1.sinks.sink1.channel = channel1
tier1.sinks.sink1.batchSize = 20
```

For the list of Kafka Sink properties, see [Kafka Sink Properties](#).

For the full list of Kafka producer properties, see the [Kafka documentation](#).

The Kafka sink uses the `topic` and `key` properties from the `FlumeEvent` headers to determine where to send events in Kafka. If the header contains the `topic` property, that event is sent to the designated topic, overriding the configured topic. If the header contains the `key` property, that key is used to partition events within the topic. Events with the same key are sent to the same partition. If the `key` parameter is not specified, events are distributed randomly to partitions. Use these properties to control the topics and partitions to which events are sent through the Flume source or interceptor.

Kafka Channel

CDH 5.3 and higher includes a Kafka channel to Flume in addition to the existing memory and file channels. You can use the Kafka channel:

- To write to Hadoop directly from Kafka without using a source.
- To write to Kafka directly from Flume sources without additional buffering.
- As a reliable and highly available channel for any source/sink combination.

The following Flume configuration uses a Kafka channel with an `exec` source and `hdfs` sink:

```
tier1.sources = source1
tier1.channels = channel1
tier1.sinks = sink1

tier1.sources.source1.type = exec
tier1.sources.source1.command = /usr/bin/vmstat 1
tier1.sources.source1.channels = channel1

tier1.channels.channel1.type = org.apache.flume.channel.kafka.KafkaChannel
tier1.channels.channel1.capacity = 10000
tier1.channels.channel1.zookeeperConnect = zk01.example.com:2181
tier1.channels.channel1.parseAsFlumeEvent = false
tier1.channels.channel1.kafka.topic = channel2
tier1.channels.channel1.kafka.consumer.group.id = channel2-grp
tier1.channels.channel1.kafka.consumer.auto.offset.reset = earliest
tier1.channels.channel1.kafka.bootstrap.servers =
kafka02.example.com:9092,kafka03.example.com:9092
tier1.channels.channel1.transactionCapacity = 1000
tier1.channels.channel1.kafka.consumer.max.partition.fetch.bytes=2097152

tier1.sinks.sink1.type = hdfs
tier1.sinks.sink1.hdfs.path = /tmp/kafka/channel
tier1.sinks.sink1.hdfs.rollInterval = 5
tier1.sinks.sink1.hdfs.rollSize = 0
tier1.sinks.sink1.hdfs.rollCount = 0
```

```
tier1.sinks.sink1.hdfs.fileType = DataStream
tier1.sinks.sink1.channel = channel1
```

For the list of Kafka Channel properties, see [Kafka Channel Properties](#).

For the full list of Kafka producer properties, see the [Kafka documentation](#).

Additional Considerations When Using Apache Kafka

When using Apache Kafka, consider the following:

- Use Cloudera Manager to start and stop Kafka and ZooKeeper services. Do not use the `kafka-server-start`, `kafka-server-stop`, `zookeeper-server-start`, and `zookeeper-server-stop` commands.
- All Kafka command-line tools are located in `/opt/cloudera/parcels/KAFKA/lib/kafka/bin/`.
- Ensure that the `JAVA_HOME` environment variable is set to your JDK installation directory before using the command-line tools. For example:

```
export JAVA_HOME=/usr/java/jdk1.7.0_55-cloudera
```

See the [Apache Kafka documentation](#).

See the [Apache Kafka FAQ](#).

See [Kafka code examples](#).

Apache Kafka Administration

This section describes ways to configure and manage Apache Kafka, including performance tuning and high availability considerations.

Configuring Apache Kafka Security

This topic describes additional steps you can take to ensure the safety and integrity of your data stored in Apache Kafka, with features available in CDK 2.0.0 and higher Powered By Apache Kafka:

Deploying SSL for Kafka

Kafka allows clients to connect over SSL. By default, SSL is disabled, but can be turned on as needed.

Step 1. Generating Keys and Certificates for Kafka Brokers

First, generate the key and the certificate for each machine in the cluster using the Java keytool utility. See [Creating Certificates](#).

keystore is the keystore file that stores your certificate. *validity* is the valid time of the certificate in days.

```
$ keytool -keystore {tmp.server.keystore.jks} -alias localhost -validity {validity}
-genkey
```

Make sure that the common name (CN) matches the fully qualified domain name (FQDN) of your server. The client compares the CN with the DNS domain name to ensure that it is connecting to the correct server.

Step 2. Creating Your Own Certificate Authority

You have generated a public-private key pair for each machine, and a certificate to identify the machine. However, the certificate is unsigned, so an attacker can create a certificate and pretend to be any machine. Sign certificates for each machine in the cluster to prevent unauthorized access.

A Certificate Authority (CA) is responsible for signing certificates. A CA is similar to a government that issues passports. A government stamps (signs) each passport so that the passport becomes difficult to forge. Similarly, the CA signs the certificates, and the cryptography guarantees that a signed certificate is computationally difficult to forge. If the CA is a genuine and trusted authority, the clients have high assurance that they are connecting to the authentic machines.

```
openssl req -new -x509 -keyout ca-key -out ca-cert -days 365
```

The generated CA is a public-private key pair and certificate used to sign other certificates.

Add the generated CA to the client truststores so that clients can trust this CA:

```
keytool -keystore {client.truststore.jks} -alias CARoot -import -file {ca-cert}
```



Note: If you configure Kafka brokers to require client authentication by setting `ssl.client.auth` to be requested or required on the [Kafka brokers config](#), you must provide a truststore for the Kafka brokers as well. The truststore should have all the CA certificates by which the clients keys are signed.

The keystore created in step 1 stores each machine's own identity. In contrast, the truststore of a client stores all the certificates that the client should trust. Importing a certificate into a truststore means trusting all certificates that are signed by that certificate. This attribute is called the *chain of trust*. It is particularly useful when deploying SSL on a large Kafka cluster. You can sign all certificates in the cluster with a single CA, and have all machines share the same truststore that trusts the CA. That way, all machines can authenticate all other machines.

Step 3. Signing the certificate

Now you can sign all certificates generated by step 1 with the CA generated in step 2.

1. Export the certificate from the keystore:

```
keytool -keystore server.keystore.jks -alias localhost -certreq -file cert-file
```

2. Sign it with the CA:

```
openssl x509 -req -CA ca-cert -CAkey ca-key -in cert-file -out cert-signed -days {validity} -CAcreateserial -passin pass:{ca-password}
```

3. Import both the certificate of the CA and the signed certificate into the keystore:

```
keytool -keystore server.keystore.jks -alias CARoot -import -file ca-cert keytool -keystore server.keystore.jks -alias localhost -import -file cert-signed
```

The definitions of the variables are as follows:

- *keystore*: the location of the keystore
- *ca-cert*: the certificate of the CA
- *ca-key*: the private key of the CA
- *ca-password*: the passphrase of the CA
- *cert-file*: the exported, unsigned certificate of the server
- *cert-signed*: the signed certificate of the server

The following Bash script demonstrates the steps described above. One of the commands assumes a password of test1234, so either use that password or edit the command before running it.

```
#!/bin/bash
#Step 1
keytool -keystore server.keystore.jks -alias localhost -validity 365 -genkey
#Step 2
openssl req -new -x509 -keyout ca-key -out ca-cert -days 365
keytool -keystore server.truststore.jks -alias CARoot -import -file ca-cert
keytool -keystore client.truststore.jks -alias CARoot -import -file ca-cert
#Step 3
keytool -keystore server.keystore.jks -alias localhost -certreq -file cert-file
openssl x509 -req -CA ca-cert -CAkey ca-key -in cert-file -out cert-signed -days 365 -CAcreateserial -passin pass:test1234
keytool -keystore server.keystore.jks -alias CARoot -import -file ca-cert
keytool -keystore server.keystore.jks -alias localhost -import -file cert-signed
```

Step 4. Configuring Kafka Brokers

Kafka Brokers support listening for connections on multiple ports. If SSL is enabled for inter-broker communication (see below for how to enable it), both PLAINTEXT and SSL ports are required.

To configure the listeners from Cloudera Manager, perform the following steps:

1. In Cloudera Manager, click Kafka > Instances, and then click on "Kafka Broker" > Configurations > Kafka Broker Advanced Configuration Snippet (Safety Valve) for kafka.properties. Enter the following information:

```
listeners=PLAINTEXT://<kafka-broker-host-name>:9092,SSL://<kafka-broker-host-name>:9093
advertised.listeners=PLAINTEXT://<kafka-broker-host-name>:9092,SSL://<kafka-broker-host-name>:9093
```

where `kafka-broker-host-name` is the FQDN of the broker that you selected from the Instances page Cloudera Manager. In the above sample configurations we used PLAINTEXT and SSL protocols for the SSL enabled brokers. For information about other supported security protocols, see [Using Kafka Supported Protocols](#)

2. Repeat the above step for all the brokers. The `advertised.listeners` configuration above is needed to connect the brokers from external clients.

3. Deploy the above client configurations and rolling restart the Kafka service from Cloudera Manager.

Kafka CSD auto-generates listeners for Kafka brokers, depending on your SSL and Kerberos configuration. To enable SSL for Kafka installations, do the following:

1. Turn on SSL for the Kafka service by turning on the `ssl_enabled` configuration for the Kafka CSD.
2. Set `security.inter.broker.protocol` as `SSL`, if Kerberos is disabled; otherwise, set it as `SASL_SSL`.

The following SSL configurations are required on each broker. Each of these values can be set in Cloudera Manager. See [Modifying Configuration Properties Using Cloudera Manager](#):

```
ssl.keystore.location=/var/private/ssl/kafka.server.keystore.jks
ssl.keystore.password=test1234
ssl.key.password=test1234
ssl.truststore.location=/var/private/ssl/kafka.server.truststore.jks
ssl.truststore.password=test1234
```

Other configuration settings might also be needed, depending on your requirements:

- `ssl.client.auth=none`: Other options for client authentication are `required`, or `requested`, where clients without certificates can still connect. The use of `requested` is discouraged, as it provides a false sense of security and misconfigured clients can still connect.
- `ssl.cipher.suites`: A cipher suite is a named combination of authentication, encryption, MAC, and a key exchange algorithm used to negotiate the security settings for a network connection using TLS or SSL network protocol. This list is empty by default.
- `ssl.enabled.protocols=TLSv1.2,TLSv1.1,TLSv1`: Provide a list of SSL protocols that your brokers accept from clients.
- `ssl.keystore.type=JKS`
- `ssl.truststore.type=JKS`

To enable SSL for inter-broker communication, add the following line to the broker properties file. The default value is `PLAINTEXT`. See [Using Kafka Supported Protocols](#) on page 68.

```
security.inter.broker.protocol=SSL
```

Due to import regulations in some countries, the Oracle implementation limits the strength of cryptographic algorithms available by default. If you need stronger algorithms (for example, AES with 256-bit keys), you must obtain the [JCE Unlimited Strength Jurisdiction Policy Files](#) and install them in the JDK/JRE. For more information, see the [JCA Providers Documentation](#).

Once you start the broker, you should see the following message in the server.log:

```
with addresses: PLAINTEXT -> EndPoint(192.168.64.1,9092,PLAINTEXT),SSL ->
EndPoint(192.168.64.1,9093,SSL)
```

To check whether the server keystore and truststore are set up properly, run the following command:

```
openssl s_client -debug -connect localhost:9093 -tls1
```



Note: TLSv1 should be listed under `ssl.enabled.protocols`.

In the output of this command, you should see the server certificate:

```
-----BEGIN CERTIFICATE-----
{variable sized random bytes}
-----END CERTIFICATE-----
subject=/C=US/ST=CA/L=Santa Clara/O=org/OU=org/CN=John Smith
issuer=/C=US/ST=CA/L=Santa Clara/O=org/OU=org/CN=kafka/emailAddress=test@test.com
```

If the certificate does not appear, or if there are any other error messages, your keystore is not set up properly.

Step 5. Configuring Kafka Clients

SSL is supported only for the new Kafka Producer and Consumer APIs. The configurations for SSL are the same for both the producer and consumer.

If client authentication is not required in the broker, the following shows a minimal configuration example:

```
security.protocol=SSL
ssl.truststore.location=/var/private/ssl/kafka.client.truststore.jks
ssl.truststore.password=test1234
```

If client authentication is required, a keystore must be created as in step 1, and you must also configure the following properties:

```
ssl.keystore.location=/var/private/ssl/kafka.client.keystore.jks
ssl.keystore.password=test1234
ssl.key.password=test1234
```

Other configuration settings might also be needed, depending on your requirements and the broker configuration:

- `ssl.provider` (Optional). The name of the security provider used for SSL connections. Default is the default security provider of the JVM.
- `ssl.cipher.suites` (Optional). A cipher suite is a named combination of authentication, encryption, MAC, and a key exchange algorithm used to negotiate the security settings for a network connection using TLS or SSL network protocol.
- `ssl.enabled.protocols`=`TLSv1.2`,`TLSv1.1`,`TLSv1`. This property should list at least one of the protocols configured on the broker side
- `ssl.truststore.type`=`JKS`
- `ssl.keystore.type`=`JKS`

Using Kafka Supported Protocols

Kafka can expose multiple communication endpoints, each supporting a different protocol. Supporting multiple communication endpoints enables you to use different communication protocols for client-to-broker communications and broker-to-broker communications. Set the Kafka inter-broker communication protocol using the `security.inter.broker.protocol` property. Use this property primarily for the following scenarios:

- Enabling SSL encryption for client-broker communication but keeping broker-broker communication as PLAINTEXT. Because SSL has performance overhead, you might want to keep inter-broker communication as PLAINTEXT if your Kafka brokers are behind a firewall and not susceptible to network snooping.
- Migrating from a non-secure Kafka configuration to a secure Kafka configuration without requiring downtime. Use a rolling restart and keep `security.inter.broker.protocol` set to a protocol that is supported by all brokers until all brokers are updated to support the new protocol.

For example, if you have a Kafka cluster that needs to be configured to enable Kerberos without downtime, follow these steps:

1. Set `security.inter.broker.protocol` to `PLAINTEXT`.
2. Update the Kafka service configuration to enable Kerberos.
3. Perform a rolling restart.
4. Set `security.inter.broker.protocol` to `SASL_PLAINTEXT`.

CDK 2.0 and higher Powered By Apache Kafka supports the following combinations of protocols.

| | SSL | Kerberos |
|-----------|-----|----------|
| PLAINTEXT | No | No |
| SSL | Yes | No |

| | SSL | Kerberos |
|----------------|-----|----------|
| SASL_PLAINTEXT | No | Yes |
| SASL_SSL | Yes | Yes |

These protocols can be defined for broker-to-client interaction and for broker-to-broker interaction.

`security.inter.broker.protocol` allows the broker-to-broker communication protocol to be different than the broker-to-client protocol. It was added to ease the upgrade from non-secure to secure clusters while allowing rolling upgrades.

In most cases, set `security.inter.broker.protocol` to the protocol you are using for broker-to-client communication. Set `security.inter.broker.protocol` to a protocol different than the broker-to-client protocol only when you are performing a rolling upgrade from a non-secure to a secure Kafka cluster.

Enabling Kerberos Authentication

CDK 2.0 and higher Powered By Apache Kafka supports Kerberos authentication, but it is supported only for the new Kafka Producer and Consumer APIs. If you already have a Kerberos server, you can add Kafka to your current configuration. If you do not have a Kerberos server, install it before proceeding. See [Enabling Kerberos Authentication Using the Wizard](#).

If you already have configured the mapping from Kerberos principals to short names using the `hadoop.security.auth_to_local` HDFS configuration property, configure the same rules for Kafka by adding the `sasl.kerberos.principal.to.local.rules` property to the Advanced Configuration Snippet for Kafka Broker Advanced Configuration Snippet using Cloudera Manager. Specify the rules as a comma separated list.

To enable Kerberos authentication for Kafka:

1. From Cloudera Manager, navigate to **Kafka > Configurations**. Set SSL client authentication to none. Set **Inter Broker Protocol** to `SASL_PLAINTEXT`.
2. Click **Save Changes**.
3. Restart the Kafka service.
4. Make sure that `listeners = SASL_PLAINTEXT` is present in the Kafka broker logs
`/var/log/kafka/server.log`.
5. Create a `jaas.conf` file with the following contents to use with cached Kerberos credentials (you can modify this to use keytab files instead of cached credentials. To generate keytabs, see [Step 6: Get or Create a Kerberos Principal for Each User Account](#)).

If you use kinit first, use this configuration.

```
KafkaClient {
  com.sun.security.auth.module.Krb5LoginModule required
  useTicketCache=true;
};
```

If you use keytab, use this configuration:

```
KafkaClient {
  com.sun.security.auth.module.Krb5LoginModule required
  useKeyTab=true
  keyTab="/etc/security/keytabs/kafka_server.keytab"
  principal="kafka/kafkal.hostname.com@EXAMPLE.COM";
};
```

6. Create the `client.properties` file containing the following properties.

```
security.protocol=SASL_PLAINTEXT
sasl.kerberos.service.name=kafka
```

7. Test with the Kafka console producer and consumer. To obtain a Kerberos ticket-granting ticket (TGT):

```
$ kinit <user>
```

8. Verify that your topic exists. (This does not use security features, but it is a best practice.)

```
$ kafka-topics --list --zookeeper <zookeeper>:2181
```

9. Verify that the `jaas.conf` file is used by setting the environment.

```
$ export KAFKA_OPTS="-Djava.security.auth.login.config=/home/user/jaas.conf"
```

10. Run a Kafka console producer.

```
$ kafka-console-producer --broker-list <anybroker>:9092 --topic test1  
--producer.config client.properties
```

11. Run a Kafka console consumer.

```
$ kafka-console-consumer --new-consumer --topic test1 --from-beginning  
--bootstrap-server <anybroker>:9092 --consumer.config client.properties
```

Enabling Encryption at Rest

Data encryption is increasingly recognized as an optimal method for protecting [data at rest](#).

Perform the following steps to encrypt Kafka data that is not in active use.

1. [Stop the Kafka service](#).
2. Archive the Kafka data to an alternate location, using TAR or another archive tool.
3. Unmount the affected drives.
4. Install and configure [Navigator Encrypt](#).
5. Expand the TAR archive into the encrypted directories.

Using Kafka with Sentry Authorization

Starting with CDK 2.1.x on CDH 5.9.x and higher Powered By Apache Kafka, Apache Sentry includes Kafka binding you can use to enable authorization in Kafka with Sentry. For more information, see [Authorization With Apache Sentry](#).

Configuring Kafka to Use Sentry Authorization

The following steps describe how to configure Kafka to use Sentry authorization. These steps assume you have installed Kafka and Sentry on your cluster.

Sentry requires that your cluster include HDFS. After you install and start Sentry with the correct configuration, you can stop the HDFS service.



Note: CDK Powered By Apache Kafka can make use of LDAP-based user groups when the LDAP directory is synchronized to Linux via tools such as SSSD. CDK Powered By Apache Kafka does not support direct integration with LDAP, either through direct Kafka LDAP authentication, or via Hadoop's group mapping (when `hadoop.group.mapping` is set to `LdapGroupMapping`). For more information, see [Configuring LDAP Group Mappings](#).

For more information, see [Installing or Upgrading CDK Powered By Apache Kafka®](#) on page 48 and [Installing and Upgrading the Sentry Service](#).

To configure Sentry authentication for Kafka:

1. Go to **Kafka > Configuration**.

2. Select the checkbox **Enable Kerberos Authentication**.
3. Select a Sentry service in the Kafka service configuration.
4. Add **Super users**. Super users can perform any action on any resource in the Kafka cluster. The `kafka` user is added as a super user by default. Super user requests are authorized without going through Sentry, which provides enhanced performance.
5. Select the checkbox **Enable Sentry Privileges Caching** to enhance performance.

Authorizable Resources

Authorizable resources are resources or entities in a Kafka cluster that require special permissions for a user to be able to perform actions on them. Kafka has four authorizable resources.

- **Cluster**, which controls who can perform cluster-level operations such as creating or deleting a topic. This can only have one value, `kafka-cluster`, as one Kafka cluster cannot have more than one cluster resource.
- **Topic**, which controls who can perform topic-level operations such as producing and consuming topics. Its value must match exactly the topic name in the Kafka cluster. With CDK 3.1.0 and CDH 5.14.2 and later, wildcards (*) can be used to refer to any topic in the privilege.
- **Consumergroup**, which controls who can perform consumergroup-level operations such as joining or describing a consumergroup. Its value must exactly match the `group.id` of a consumergroup. With CDK 3.1.0 and CDH 5.14.2 and later, you can use a wildcard (*) to refer to any consumer groups in the privilege. This is useful when used with Spark Streaming, where a generated `group.id` may be needed.
- **Host**, which controls from where specific operations can be performed. Think of this as a way to achieve IP filtering in Kafka. You can set the value of this resource to the wildcard (*), which represents all hosts.



Note: Only IP addresses should be specified in the host component of Kafka Sentry privileges, hostnames are not supported.

Authorized Actions

You can perform multiple actions on each resource. The following operations are supported by Kafka, though not all actions are valid on all resources.

- ALL, this is a wildcard action, and represents all possible actions on a resource.
- read
- write
- create
- delete
- alter
- describe
- clusteraction

Authorizing Privileges

Privileges define what actions are allowed on a resource. A privilege is represented as a string in Sentry. The following rules apply to a valid privilege.

- Can have at most one Host resource. If you do not specify a Host resource in your privilege string, `Host=*` is assumed.
- Must have exactly one non-Host resource.
- Must have exactly one action specified at the end of the privilege string.

For example, the following are valid privilege strings:

```
Host=*->Topic=myTopic->action=ALL
Topic=test->action=ALL
```

Granting Privileges to a Role

The following examples grant privileges to the role `test`, so that users in `testGroup` can create a topic named `testTopic` and produce to it.

The user executing these commands must be added to the Sentry parameter `sentry.service.allow.connect` and also be a member of a group defined in `sentry.service.admin.group`.

Before you can assign the `test` role, you must first create it. To create the `test` role:

```
$kafka-sentry -cr -r test
```

To confirm that the role was created, list the roles:

```
$ kafka-sentry -lr
```

If Sentry privileges caching is enabled, as recommended, the new privileges you assign take some time to appear in the system. The `time` is the time-to-live interval of the Sentry privileges cache, which is set using `sentry.kafka.caching.ttl.ms`. By default, this interval is set to 30 seconds. For test clusters, it is beneficial to have changes appear within the system as fast as possible, therefore, Cloudera recommends that you either use a lower time interval, or disable caching with `sentry.kafka.caching.enable`.

- Allow users in `testGroup` to write to `testTopic` from localhost, which allows users to produce to `testTopic`. They need both write and describe permissions.

```
$ kafka-sentry -gpr -r test -p "Host=127.0.0.1->Topic=testTopic->action=write"  
$ kafka-sentry -gpr -r test -p "Host=127.0.0.1->Topic=testTopic->action=describe"
```

- Assign the test role to the group `testGroup`:

```
kafka-sentry -arg -r test -g testGroup
```

- Verify that the test role is part of the group `testGroup`:

```
kafka-sentry -lr -g testGroup
```

- Create `testTopic`.

```
$ kafka-topics --create --zookeeper localhost:2181 --replication-factor 1 \  
  --partitions 1 --topic testTopic  
$ kafka-topics --list --zookeeper localhost:2181 testTopic
```

- Produce to `testTopic`. Note that you have to pass a configuration file, `producer.properties`, with information on JAAS configuration and other Kerberos authentication related information. See [SASL Configuration for Kafka Clients](#).

```
$ kafka-console-producer --broker-list localhost:9092 --topic testTopic \  
  --producer.config producer.properties  
This is a message  
This is another message
```

- Grant the `create` privilege to the test role.

```
$ kafka-sentry -gpr -r test -p "Host=127.0.0.1->Cluster=kafka-cluster->action=create"
```

- Allow users in `testGroup` to describe `testTopic` from localhost, which the user creates and uses.

```
$ kafka-sentry -gpr -r test -p "Host=127.0.0.1->Topic=testTopic->action=describe"
```


- Grant the `describe` privilege to the test role.

```
$ kafka-sentry -gpr -r test -p
"Host=127.0.0.1->ConsumerGroup=testconsumergroup->action=describe"
```

- Allow users in `testGroup` to read from a consumer group, `testconsumergroup`, that it will start and join.

```
$ kafka-sentry -gpr -r test -p
"Host=127.0.0.1->ConsumerGroup=testconsumergroup->action=read"
```

- Allow users in `testGroup` to read from `testTopic` from `localhost` and to consume from `testTopic`.

```
$ kafka-sentry -gpr -r test -p "Host=127.0.0.1->Topic=testTopic->action=read"
```

- Consume from `testTopic`. Note that you have to pass a configuration file, `consumer.properties`, with information on JAAS configuration and other Kerberos authentication related information. The configuration file must also specify `group.id` as `testconsumergroup`.

```
kafka-console-consumer --new-consumer --topic test1 --from-beginning --bootstrap-server
<anybroker>:9092 --consumer.config consumer.properties
This is a message
This is another message
```

Troubleshooting

If Kafka requests are failing due to authorization, the following steps can provide insight into the error:

- Make sure you are kinit'd as a user who has privileges to perform an operation.
- Identify which broker is hosting leader of the partition you are trying to produce to or consume from, as this leader is going to authorize your request against Sentry. One easy way of debugging is to just have one Kafka broker. Change log level of the Kafka broker to debug and restart the broker.
- Run the Kafka client or Kafka CLI with required arguments and capture the Kafka log, which should be something like `/var/log/kafka/kafka-broker-<HOST_ID>.log` on kafka broker's host.
- There will be many Jetty logs, and filtering that out usually helps in reducing noise. Look for log messages from `org.apache.sentry`.
- Look for following information in the filtered logs:
 - Groups the user Kafka client or CLI is running as.
 - Required privileges for the operation.
 - Retrieved privileges from Sentry.
 - Required and retrieved privileges comparison result.

This log information can provide insight into which privilege is not assigned to a user, causing a particular operation to fail.

Configuring High Availability and Consistency for Apache Kafka

To achieve high availability and consistency targets, adjust the following parameters to meet your requirements:

Replication Factor

The default replication factor for new topics is one. For high availability production systems, Cloudera recommends setting the replication factor to at least three. This requires at least three Kafka brokers.

To change the replication factor, navigate to **Kafka Service > Configuration > Service-Wide**. Set **Replication factor** to 3, click **Save Changes**, and restart the Kafka service.

Preferred Leader Election

Kafka is designed with failure in mind. At some point in time, web communications or storage resources fail. When a broker goes offline, one of the replicas becomes the new leader for the partition. When the broker comes back online, it has no leader partitions. Kafka keeps track of which machine is configured to be the leader. Once the original broker is back up and in a good state, Kafka restores the information it missed in the interim and makes it the partition leader once more.

Preferred Leader Election is enabled by default, and should occur automatically unless you actively disable the feature. Typically, the leader is restored within five minutes of coming back online. If the preferred leader is offline for a very long time, though, it might need additional time to restore its required information from the replica.

There is a small possibility that some messages might be lost when switching back to the preferred leader. You can minimize the chance of lost data by setting the `acks` property on the Producer to `all`. See [Acknowledgements](#) on page 74.

Unclean Leader Election

Enable unclean leader election to allow an out-of-sync replica to become the leader and preserve the availability of the partition. With unclean leader election, messages that were not synced to the new leader are lost. This provides balance between consistency (guaranteed message delivery) and availability. With unclean leader election disabled, if a broker containing the leader replica for a partition becomes unavailable, and no in-sync replica exists to replace it, the partition becomes unavailable until the leader replica or another in-sync replica is back online.

To enable unclean leader election, navigate to **Kafka Service > Configuration > Service-Wide**. Check the box labeled **Enable unclean leader election**, click **Save Changes**, and restart the Kafka service.

Acknowledgements

When writing or configuring a Kafka producer, you can choose how many replicas commit a new message before the message is acknowledged using the `acks` property.

Set `acks` to `0` (immediately acknowledge the message without waiting for any brokers to commit), `1` (acknowledge after the leader commits the message), or `all` (acknowledge after all in-sync replicas are committed) according to your requirements. Setting `acks` to `all` provides the highest consistency guarantee at the expense of slower writes to the cluster.

Minimum In-sync Replicas

You can set the minimum number of in-sync replicas (ISRs) that must be available for the producer to successfully send messages to a partition using the `min.insync.replicas` setting. If `min.insync.replicas` is set to `2` and `acks` is set to `all`, each message must be written successfully to at least two replicas. This guarantees that the message is not lost unless both hosts crash.

It also means that if one of the hosts crashes, the partition is no longer available for writes. Similar to the unclean leader election configuration, setting `min.insync.replicas` is a balance between higher consistency (requiring writes to more than one broker) and higher availability (allowing writes when fewer brokers are available).

The leader is considered one of the in-sync replicas. It is included in the count of total `min.insync.replicas`. However, leaders are special, in that producers and consumers can only interact with leaders in a Kafka cluster.

To configure `min.insync.replicas` at the cluster level, navigate to **Kafka Service > Configuration > Service-Wide**. Set **Minimum number of replicas in ISR** to the desired value, click **Save Changes**, and restart the Kafka service.

To set this parameter on a per-topic basis, navigate to **Kafka Service > Configuration > Kafka broker Default Group > Advanced**, and add the following to the **Kafka Broker Advanced Configuration Snippet (Safety Valve) for kafka.properties**:

```
min.insync.replicas.per.topic=topic_name_1:value,topic_name_2:value
```

Replace `topic_name_n` with the topic names, and replace `value` with the desired minimum number of in-sync replicas.

You can also set this parameter using the `/usr/bin/kafka-topics --alter` command for each topic. For example:

```
/usr/bin/kafka-topics --alter --zookeeper zk01.example.com:2181 --topic topicname \
--config min.insync.replicas=2
```

Kafka MirrorMaker

Kafka mirroring enables maintaining a replica of an existing Kafka cluster. You can configure MirrorMaker directly in Cloudera Manager 5.4 and higher.

The most important configuration setting is **Destination broker list**. This is a list of brokers on the destination cluster. You should list more than one, to support high availability, but you do not need to list all brokers.

MirrorMaker requires that you specify a **Topic whitelist** that represents the exclusive set of topics to replicate. The **Topic blacklist** setting has been removed in CDK 2.0 and higher Powered By Apache Kafka.



Note: The **Avoid Data Loss** option from earlier releases has been removed in favor of automatically setting the following properties. Also note that MirrorMaker starts correctly if you enter the numeric values in the configuration snippet (rather than using "max integer" for `retries` and "max long" for `max.block.ms`).

1. Producer settings

- `acks=all`
- `retries=2147483647`
- `max.block.ms=9223372036854775807`

2. Consumer setting

- `auto.commit.enable=false`

3. MirrorMaker setting

- `abort.on.send.failure=true`

Configuring Apache Kafka for Performance and Resource Management

Apache Kafka is optimized for small messages. [According to benchmarks](#), the best performance occurs with 1 KB messages. Larger messages (for example, 10 MB to 100 MB) can decrease throughput and significantly impact operations.

This topic describes options that can improve performance and reliability in your Kafka cluster:

Partitions and Memory Usage

For a quick video introduction to load balancing, see [tl;dr: Balancing Apache Kafka Clusters](#).

Brokers allocate a buffer the size of `replica.fetch.max.bytes` for each partition they replicate. If `replica.fetch.max.bytes` is set to 1 MiB, and you have 1000 partitions, about 1 GiB of RAM is required. Ensure that the number of partitions multiplied by the size of the largest message does not exceed available memory.

The same consideration applies for the consumer `fetch.message.max.bytes` setting. Ensure that you have enough memory for the largest message for each partition the consumer replicates. With larger messages, you might need to use fewer partitions or provide more RAM.

Partition Reassignment

At some point you will likely exceed configured resources on your system. If you add a Kafka broker to your cluster to handle increased demand, new partitions are allocated to it (the same as any other broker), but it does not automatically

share the load of existing partitions on other brokers. To redistribute the existing load among brokers, you must manually reassign partitions. You can do so using `bin/kafka-reassign-partitions.sh` script utilities.

To reassign partitions:

1. Create a list of topics you want to move.

```
topics-to-move.json
{"topics": [{"topic": "foo1"},
            {"topic": "foo2"}],
  "version":1
}
```

2. Use the `--generate` option in `kafka-reassign-partitions.sh` to list the distribution of partitions and replicas on your current brokers, followed by a list of suggested locations for partitions on your new broker.

```
bin/kafka-reassign-partitions.sh --zookeeper localhost:2181
  --topics-to-move-json-file topics-to-move.json
  --broker-list "4"
  --generate

Current partition replica assignment

{"version":1,
 "partitions":[{"topic":"foo1","partition":2,"replicas":[1,2]},
              {"topic":"foo1","partition":0,"replicas":[3,1]},
              {"topic":"foo2","partition":2,"replicas":[1,2]},
              {"topic":"foo2","partition":0,"replicas":[3,2]},
              {"topic":"foo1","partition":1,"replicas":[2,3]},
              {"topic":"foo2","partition":1,"replicas":[2,3]}]}

{"version":1,
 "partitions":[{"topic":"foo1","partition":3,"replicas":[4]},
              {"topic":"foo1","partition":1,"replicas":[4]},
              {"topic":"foo2","partition":2,"replicas":[4]}]}

}
```

3. Revise the suggested list if required, and then save it as a JSON file.
4. Use the `--execute` option in `kafka-reassign-partitions.sh` to start the redistribution process, which can take several hours in some cases.

```
> bin/kafka-reassign-partitions.sh \
  --zookeeper localhost:2181 \
  --reassignment-json-file expand-cluster-reassignment.json
  --execute
```

5. Use the `--verify` option in `kafka-reassign-partitions.sh` to check the status of your partitions.

Although reassigning partitions is labor-intensive, you should anticipate system growth and redistribute the load when your system is at 70% capacity. If you wait until you are forced to redistribute because you have reached the limit of your resources, the redistribution process can be extremely slow.

Garbage Collection

Large messages can cause longer garbage collection (GC) pauses as brokers allocate large chunks. Monitor the GC log and the server log. If long GC pauses cause Kafka to abandon the ZooKeeper session, you may need to configure longer timeout values for `zookeeper.session.timeout.ms`.

Handling Large Messages

Before configuring Kafka to handle large messages, first consider the following options to reduce message size:

- The Kafka producer can compress messages. For example, if the original message is a text-based format (such as XML), in most cases the compressed message will be sufficiently small.

Use the `compression.codec` and `compressed.topics` producer configuration parameters to enable compression. Gzip and Snappy are supported.

- If shared storage (such as NAS, HDFS, or S3) is available, consider placing large files on the shared storage and using Kafka to send a message with the file location. In many cases, this can be much faster than using Kafka to send the large file itself.
- Split large messages into 1 KB segments with the producing client, using partition keys to ensure that all segments are sent to the same Kafka partition in the correct order. The consuming client can then reconstruct the original large message.

If you still need to send large messages with Kafka, modify the configuration parameters presented in the following sections to match your requirements.

Broker Configuration

Table 20: Broker Configuration Properties

| Property | Default Value | Description |
|--------------------------------------|-----------------------|--|
| <code>message.max.bytes</code> | 1000000 (1 MB) | Maximum message size the broker accepts. When using the old consumer, this property must be lower than the consumer <code>fetch.message.max.bytes</code> , or the consumer cannot consume the message. |
| <code>log.segment.bytes</code> | 1073741824 (1 GiB) | Size of a Kafka data file. Must be larger than any single message. |
| <code>replica.fetch.max.bytes</code> | 1048576 (1 MiB) | Maximum message size a broker can replicate. Must be larger than <code>message.max.bytes</code> , or a broker can accept messages it cannot replicate, potentially resulting in data loss. |

Consumer Configuration

Kafka offers two separate consumer implementations, the old consumer and the new consumer. The old consumer is the `Consumer` class written in Scala. The new consumer is the `KafkaConsumer` class written in Java. When configuring Kafka to handle large messages, different properties have to be configured for each consumer implementation.

Old Consumer

Table 21: Old Consumer Configuration Properties

| Property | Default Value | Description |
|--------------------------------------|----------------------|--|
| <code>fetch.message.max.bytes</code> | 52428800 (50 MiB) | The maximum amount of data the server should return for a fetch request. This is a hard limit. If a message batch is larger than this limit, the consumer will not be able to consume the message or any subsequent messages in a given partition. |

New Consumer

Table 22: New Consumer Configuration Properties

| Property | Default Value | Description |
|--|---------------------|--|
| <code>max.partition.fetch.bytes</code> | 1048576 (10 MiB) | The maximum amount of data per-partition the server will return. |

| Property | Default Value | Description |
|------------------------------|----------------------|--|
| <code>fetch.max.bytes</code> | 52428800 (50 MiB) | The maximum amount of data the server should return for a fetch request. |



Note: The new consumer is able to consume a message batch that is larger than the default value of the `max.partition.fetch.bytes` or `fetch.max.bytes` property. However, the batch will be sent alone, which can cause performance degradation.

Tuning Kafka for Optimal Performance

For a quick video introduction to tuning Kafka, see [tldr: Tuning Your Apache Kafka Cluster](#).

Performance tuning involves two important metrics: *Latency* measures how long it takes to process one event, and *throughput* measures how many events arrive within a specific amount of time. Most systems are optimized for either latency or throughput. Kafka is balanced for both. A well tuned Kafka system has just enough brokers to handle topic throughput, given the latency required to process information as it is received.

Tuning your producers, brokers, and consumers to send, process, and receive the largest possible batches within a manageable amount of time results in the best balance of latency and throughput for your Kafka cluster.

Tuning Kafka Producers

Kafka uses an asynchronous publish/subscribe model. When your producer calls the `send()` command, the result returned is a *future*. The future provides methods to let you check the status of the information in process. When the batch is ready, the producer sends it to the broker. The Kafka broker waits for an event, receives the result, and then responds that the transaction is complete.

If you do not use a future, you could get just one record, wait for the result, and then send a response. Latency is very low, but so is throughput. If each transaction takes 5 ms, throughput is 200 events per second.—slower than the expected 100,000 events per second.

When you use `Producer.send()`, you fill up buffers on the producer. When a buffer is full, the producer sends the buffer to the Kafka broker and begins to refill the buffer.

Two parameters are particularly important for latency and throughput: batch size and linger time.

Batch Size

`batch.size` measures batch size in total bytes instead of the number of messages. It controls how many bytes of data to collect before sending messages to the Kafka broker. Set this as high as possible, without exceeding available memory. The default value is 16384.

If you increase the size of your buffer, it might never get full. The Producer sends the information eventually, based on other triggers, such as `linger.time` in milliseconds. Although you can impair memory usage by setting the buffer batch size too high, this does not impact latency.

If your producer is sending all the time, you are probably getting the best throughput possible. If the producer is often idle, you might not be writing enough data to warrant the current allocation of resources.

Linger Time

`linger.ms` sets the maximum time to buffer data in asynchronous mode. For example, a setting of 100 batches 100ms of messages to send at once. This improves throughput, but the buffering adds message delivery latency.

By default, the producer does not wait. It sends the buffer any time data is available.

Instead of sending immediately, you can set `linger.ms` to 5 and send more messages in one batch. This would reduce the number of requests sent, but would add up to 5 milliseconds of latency to records sent, even if the load on the system does not warrant the delay.

The farther away the broker is from the producer, the more overhead required to send messages. Increase `linger.ms` for higher latency and higher throughput in your producer.

Tuning Kafka Brokers

Topics are divided into partitions. Each partition has a leader. Most partitions are written into leaders with multiple replicas. When the leaders are not balanced properly, one might be overworked, compared to others. For more information on load balancing, see [Partitions and Memory Usage](#).

Depending on your system and how critical your data is, you want to be sure that you have sufficient replication sets to preserve your data. Cloudera recommends starting with one partition per physical storage disk and one consumer per partition.

Tuning Kafka Consumers

Consumers can create throughput issues on the other side of the pipeline. The maximum number of consumers for a topic is equal to the number of partitions. You need enough partitions to handle all the consumers needed to keep up with the producers.

Consumers in the same consumer group split the partitions among them. Adding more consumers to a group can enhance performance. Adding more consumer groups does not affect performance.

How you use the `replica.high.watermark.checkpoint.interval.ms` property can affect throughput. When reading from a partition, you can mark the last point where you read information. That way, if you have to go back and locate missing data, you have a checkpoint from which to move forward without having to reread prior data. If you set the checkpoint watermark for every event, you will never lose a message, but it significantly impacts performance. If, instead, you set it to check the offset every hundred messages, you have a margin of safety with much less impact on throughput.

Configuring JMX Ephemeral Ports

Kafka uses two high-numbered [ephemeral ports](#) for JMX. These ports are listed when you view `netstat -anp` information for the Kafka Broker process.

You can change the number for the first port by adding a command similar to `-Dcom.sun.management.jmxremote.rmi.port=<port number>` to the field **Additional Broker Java Options (broker_java_opts)** in Cloudera Manager. The `JMX_PORT` configuration maps to `com.sun.management.jmxremote.port` by default.

The second ephemeral port used for JMX communication is implemented for the JRMP protocol and cannot be changed.

Quotas

For a quick video introduction to quotas, see [tl;dr: Quotas](#).

In CDK 2.0 and higher Powered By Apache Kafka, Kafka can enforce quotas on produce and fetch requests. Producers and consumers can use very high volumes of data. This can monopolize broker resources, cause network saturation, and generally deny service to other clients and the brokers themselves. *Quotas* protect against these issues and are important for large, multi-tenant clusters where a small set of clients using high volumes of data can degrade the user experience.

Quotas are byte-rate thresholds, defined per client ID. A client ID logically identifies an application making a request. A single client ID can span multiple producer and consumer instances. The quota is applied for all instances as a single entity: For example, if a client ID has a produce quota of 10 MB/s, that quota is shared across all instances with that same ID.

When running Kafka as a service, quotas can enforce API limits. By default, each unique client ID receives a fixed quota in bytes per second, as configured by the cluster (`quota.producer.default`, `quota.consumer.default`). This quota is defined on a per-broker basis. Each client can publish or fetch a maximum of *X* bytes per second per broker before it gets throttled.

The broker does not return an error when a client exceeds its quota, but instead attempts to slow the client down. The broker computes the amount of delay needed to bring a client under its quota and delays the response for that

amount of time. This approach keeps the quota violation transparent to clients (outside of client-side metrics). This also prevents clients from having to implement special backoff and retry behavior.

Setting Quotas

You can override the default quota for client IDs that need a higher or lower quota. The mechanism is similar to per-topic log configuration overrides. Write your client ID overrides to ZooKeeper under `/config/clients`. All brokers read the overrides, which are effective immediately. You can change quotas without having to do a rolling restart of the entire cluster.

By default, each client ID receives an unlimited quota. The following configuration sets the default quota per producer and consumer client ID to 10 MB/s.

```
quota.producer.default=10485760
quota.consumer.default=10485760
```

To set quotas using Cloudera Manager, open the Kafka **Configuration** page and search for *Quota*. Use the fields provided to set the **Default Consumer Quota** or **Default Producer Quota**. For more information, see [Modifying Configuration Properties Using Cloudera Manager](#).

Setting User Limits for Kafka

Kafka opens many files at the same time. The default setting of 1024 for the maximum number of open files on most Unix-like systems is insufficient. Any significant load can result in failures and cause error messages such as `java.io.IOException...(Too many open files)` to be logged in the Kafka or HDFS log files. You might also notice errors such as this:

```
ERROR Error in acceptor (kafka.network.Acceptor)
java.io.IOException: Too many open files
```

Cloudera recommends setting the value to a relatively high starting point, such as 100,000.

You can monitor the number of file descriptors in use on the Kafka Broker dashboard. In Cloudera Manager:

1. Go to the Kafka service.
2. Select a Kafka Broker.
3. Open **Charts Library** > **Process Resources** and scroll down to the **File Descriptors** chart.

See http://www.cloudera.com/documentation/enterprise/latest/topics/cm_dg_view_charts.html.

Disk Management

Monitoring

Cloudera recommends that administrators continuously monitor the following on a cluster:

Replication Status

Monitor replication status using Cloudera Manager Health Tests. Cloudera Manager automatically and continuously monitors both the `OfflineLogDirectoryCount` and `OfflineReplicaCount` metrics. Alerts are raised when failures are detected. For more information, see [Cloudera Manager Health Tests](#).

Disk Capacity

Monitor free space on mounted disks and open file descriptors. Reassign partitions or move log files around if necessary. For more information, see [kafka-reassign-partitions](#) on page 56.

Handling Disk Failures

Cloudera Manager has built in monitoring functionalities that automatically trigger alerts when disk failures are detected. When a log directory fails, Kafka also detects the failure and takes the partitions stored in that directory offline.



Important: If there are no healthy log directories present in the system, the broker stops working.

The cause of disk failures can be analyzed with the help of the [kafka-log-dirs](#) on page 60 tool, or by reviewing the error messages of `KafkaStorageException` entries in the Kafka broker log file.

To view the Kafka broker log file, complete the following steps:

1. In Cloudera Manager go to the Kafka service, select **Instances** and select the broker.
2. Go to **Log Files > Role Log File**.

In case of a disk failure, a Kafka administrator can carry out either of the following actions. The action taken depends on the failure type and system environment:

- Replace the faulty disk with a new one.
- Remove the disk and redistribute data across remaining disks to restore the desired replication factor.



Note: Disk replacement and disk removal both require stopping the broker. Therefore, Cloudera recommends that you perform these actions during a maintenance window.

Disk Replacement

To replace a disk, complete the following steps:

1. Stop the broker that has a faulty disk.
 - a. In Cloudera Manager, go to the Kafka service, select **Instances** and select the broker.
 - b. Go to **Actions > Gracefully stop this Kafka Broker**.
2. Replace the disk.
3. Mount the disk.
4. Set up the directory structure on the new disk the same way as it was set up on the previous disk.



Note: You can find the directory paths for the old disk in the **Data Directories** property of the broker.

5. Start the broker.
 - a. In Cloudera Manager go to the Kafka service, select **Instances** and select the broker.
 - b. Go to **Actions > Start this Kafka Broker**.

The Kafka broker re-creates topic partitions in the same directory by replicating data from other brokers.

Disk Removal

To remove a disk from the configuration, complete the following steps:

1. Stop the broker that has a faulty disk.
 - a. In Cloudera Manager, go to the Kafka service, select **Instances** and select the broker.
 - b. Go to **Actions > Gracefully stop this Kafka Broker**.
2. Remove the log directories on the faulty disk from the broker.
 - a. Go to **Configuration** and find the **Data Directories** property.
 - b. Remove the affected log directories with the **Remove** button.
 - c. Enter a **Reason for change**, and then click **Save Changes** to commit the changes.
3. Start the broker.

- a. In Cloudera Manager go to the Kafka service, select **Instances** and select the broker.
- b. Go to **Actions > Start this Kafka Broker**.

The Kafka broker redistributes data across the cluster.

Reassigning Replicas Between Log Directories

Reassigning replicas between log directories can prove useful when you have multiple disks available, but one or more of them is nearing capacity. Moving a replica from one disk to another ensures that the service will not go down due to disks reaching capacity. To balance storage loads, the Kafka administrator has to continuously monitor the system and reassign replicas between log directories on the same broker or across different brokers. These actions can be carried out with the `kafka-reassign-partitions` tool.

For more information on tool usage, see the documentation for the [kafka-reassign-partitions](#) on page 56 tool.

Retrieving Log Directory Replica Assignment Information

To optimize replica assignment across log directories, the list of partitions per log directory and the size of each partition is required. This information can be exposed with the `kafka-log-dirs` tool.

For more information on tool usage, see the documentation for the [kafka-log-dirs](#) on page 60 tool.

JBOD

As of CDH 6.1.0, Kafka clusters with nodes using JBOD configurations are supported by Cloudera.

JBOD refers to a system configuration where disks are used independently rather than organizing them into redundant arrays (RAID). Using RAID usually results in more reliable hard disk configurations even if the individual disks are not reliable. RAID setups like these are common in large scale big data environments built on top of commodity hardware. RAID enabled configurations are more expensive and more complicated to set up. In a large number of environments, JBOD configurations are preferred for the following reasons:

- **Reduced storage cost:** RAID-10 is recommended to protect against disk failures. However, scaling RAID-10 configurations can become excessively expensive. Storing the data redundantly on each node means that storage space requirements have to be multiplied because the data is also replicated across nodes.
- **Improved performance:** Just like HDFS, the slowest disk in RAID-10 configuration limits overall throughput. Writes need to go through a RAID controller. On the other hand, when using JBOD, IO performance is increased as a result of isolated writes across disks without a controller.

JBOD Setup and Migration

Consider the following before using JBOD support in Kafka:

- **Manual operation and administration:** Monitoring offline directories and JBOD related metrics is done through Cloudera Manager. However, identifying failed disks and rebalancing partitions between disks is done manually.
- **Manual load balancing between disks:** Unlike with RAID-10, JBOD does not automatically distribute data across disks. The process is fully manual.

To provide robust JBOD support in Kafka, changes in the Kafka protocol have been made. When performing an upgrade to a new version of Kafka, make sure that you follow the recommended rolling upgrade process.

For more information, see [Upgrading the CDH Cluster](#).

For more information regarding the JBOD related Kafka protocol changes, see [KIP-112](#) and [KIP-113](#).

Setup

To set up JBOD in your Kafka environment, perform the following steps:

1. Mount the required number of disks on your system.
2. In Cloudera Manager, set up log directories for all Kafka brokers.

- a. Go to the Kafka service, select **Instances** and select the broker.
- b. Go to **Configuration** and find the **Data Directories** property.
- c. Modify the path of the log directories so that they correspond with the newly mounted disks.



Note: Depending on your setup you may need to add or remove multiple data directories.

- d. Enter a **Reason for change**, and then click **Save Changes** to commit the changes.

3. Go to the Kafka service and select **Configuration**.
4. Find and configure the following properties depending on your system and use case.
 - **Number of I/O Threads**
 - **Number of Replica Fetchers**
 - **Minimum Number of Replicas in ISR**

Additionally you also have to configure the number of network threads, `num.network.threads`. However, in Cloudera Manager 5.x.x, this property can only be configured via the **Kafka Broker Advanced Configuration Snippet (Safety Valve) for kafka.properties** safety valve. For more information regarding configuration using safety valves, see [Custom Configuration](#)

5. Set replication factor to at least 3.



Important: If you set replication factor to less than 3, your data will be at risk. In addition, in case of a disk failure, disk maintenance cannot be carried out without system downtime.

6. Restart the service.
 - a. Return to the Home page by clicking the Cloudera Manager logo.
 - b. Go to the Kafka service and select **Actions > Rolling Restart**.
 - c. Check the **Restart roles with stale configurations only** checkbox and click **Rolling restart**.
 - d. Click **Close** when the restart has finished.

Migration

Migrating data from one disk to another is achieved with the `kafka-reassign-partitions` tool. The following instructions focus on migrating existing Kafka partitions to JBOD configured disks. For a full tool description, see [kafka-reassign-partitions](#) on page 56.



Note: Cloudera recommends that you minimize the volume of replica changes per command instance. Instead of moving 10 replicas with a single command, move two at a time in order to save cluster resources.

Prerequisites

- Set up JBOD in your Kafka environment. For more information, see [Setup](#) on page 82.
- Collect the log directory paths on the JBOD disks where you want to migrate existing data.
- Collect the broker IDs of the brokers you want to migrate data to.
- Collect the name of the topics you want to migrate partitions from.

Steps



Note: Output examples in these instructions are cleaned and formatted to make them easily readable.

To migrate data to JBOD configured disks, perform the following steps:

1. Create a topics-to-move JSON file that specifies the topics you want to reassign. Use the following format:

```
{ "topics": [ { "topic": "mytopic1" },
               { "topic": "mytopic2" } ],
  "version": 1
}
```

2. Generate the content for the reassignment configuration JSON with the following command:

```
kafka-reassign-partitions --zookeeper hostname:port --topics-to-move-json-file topics
to move.json --broker-list broker 1, broker 2 --generate
```

Running the command lists the distribution of partition replicas on your current brokers followed by a proposed partition reassignment configuration.

Example output:

```
Current partition replica assignment
{"version":1,
 "partitions":
  [ { "topic": "mytopic2", "partition": 1, "replicas": [2,3], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 0, "replicas": [1,2], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic2", "partition": 0, "replicas": [1,2], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 2, "replicas": [3,1], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 1, "replicas": [2,3], "log_dirs": [ "any", "any" ] }
  ]
}

Proposed partition reassignment configuration
{"version":1,
 "partitions":
  [ { "topic": "mytopic1", "partition": 0, "replicas": [4,5], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 2, "replicas": [4,5], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic2", "partition": 1, "replicas": [4,5], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 1, "replicas": [5,4], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic2", "partition": 0, "replicas": [5,4], "log_dirs": [ "any", "any" ] }
  ]
}
```

In this example, the tool proposed a configuration which reassigns existing partitions on broker 1, 2, and 3 to brokers 4 and 5.

3. Copy and paste the proposed partition reassignment configuration into an empty JSON file.
4. Modify the suggested reassignment configuration.

When migrating data you have two choices. You can move partitions to a different log directory on the same broker, or move it to a different log directory on another broker.

- a. To reassign partitions between log directories on the same broker, change the appropriate `any` entry to an absolute path. For example:

```
{ "topic": "mytopic1", "partition": 0, "replicas": [4,5], "log_dirs": [ "/JBOD-disk/directory1", "any" ] }
```

- b. To reassign partitions between log directories across different brokers, change the broker ID specified in `replicas` and the appropriate `any` entry to an absolute path. For example:

```
{ "topic": "mytopic1", "partition": 0, "replicas": [6,5], "log_dirs": [ "/JBOD-disk/directory1", "any" ] }
```

5. Save the file.
6. Start the redistribution process with the following command:



Important: The bootstrap server has to be specified with the `--bootstrap-server` option if an absolute log directory path is specified for a replica in the reassignment configuration JSON file.

The tool prints a list containing the original replica assignment and a message that reassignment has started. Example output:

Current partition replica assignment

```
{ "version": 1,
  "partitions": [
    { "topic": "mytopic2", "partition": 1, "replicas": [2, 3], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 0, "replicas": [1, 2], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic2", "partition": 0, "replicas": [1, 2], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 2, "replicas": [3, 1], "log_dirs": [ "any", "any" ] },
    { "topic": "mytopic1", "partition": 1, "replicas": [2, 3], "log_dirs": [ "any", "any" ] }
  ]
}
```

Save this to use as the `--reassignment-json-file` option during rollback
Successfully started reassignment of partitions.

7. Verify the status of the reassignment with the following command:

```
kafka-reassign-partitions --zookeeper hostname:port --reassignment-json-file reassignment_configuration.json --bootstrap-server hostname:port --verify
```

The tool prints the reassignment status of all partitions. Example output:

```
Status of partition reassignment:
Reassignment of partition mytopic2-1 completed successfully
Reassignment of partition mytopic1-0 completed successfully
Reassignment of partition mytopic2-0 completed successfully
Reassignment of partition mytopic1-2 completed successfully
Reassignment of partition mytopic1-1 completed successfully
```

Viewing Apache Kafka Metrics

For the complete list of Apache Kafka metrics, see:

- [Kafka Metrics](#)
- [Kafka Broker Metrics](#)
- [Kafka Broker Topic Metrics](#)
- [Kafka MirrorMaker Metrics](#)
- [Kafka Replica Metrics](#)

Working with Apache Kafka Logs

By default, the CDK Powered By Apache Kafka parcel is configured to log all Kafka log messages to one file per host, or broker, in the following location:

```
${log.dir}/kafka-broker-<your_hostname>.log
```

The default log directory is `/var/log/kafka`. You can view, filter, and search the logs using Cloudera Manager. See [Logs](#) for more information about viewing logs in Cloudera Manager.

You can view, filter, and search this log using Cloudera Manager. See [Logs](#).

For debugging purposes, you can create a separate file with `TRACE` level logs of a specific component (such as the controller) or the state changes.

For example, to restore the default Apache Kafka log4j configuration, do the following:

1. Navigate to the Kafka Broker Configuration page.
2. Search for the **Kafka Broker Logging Advanced Configuration Snippet (Safety Valve)** field.
3. Copy and paste the configuration snippet from the [Apache Kafka log4j.properties file](#).

Alternatively, you can add only the appenders you need.

For more information on setting properties, see [Modifying Configuration Properties Using Cloudera Manager](#).

Kafka Streams

Starting with CDK 4.0.0, Cloudera officially supports Kafka Streams. You can access the Apache Kafka website for information about how to use Kafka Streams.

- Read the [Kafka Streams Introduction](#) for an overview of the feature and an introductory video.
- Get familiar with Kafka Streams [Core Concepts](#).
- Understand Kafka Streams [Architecture](#).
- Access the [Quick Start](#) documentation to run a demonstration Kafka Streams Application.
- Use the [Tutorial](#) to write your first Kafka Streams Application.

Kafka Frequently Asked Questions

This is intended to be an easy to understand FAQ on the topic of Kafka. One part is for beginners, one for advanced users and use cases. We hope you find it fruitful. If you are missing a question, please send it to your favorite Cloudera representative and we'll populate this FAQ over time.

Basics

What is Kafka?

Kafka is a streaming message platform. Breaking it down a bit further:

“Streaming”: Lots of messages (think tens or hundreds of thousands) being sent frequently by publishers ("producers"). Message polling occurring frequently by lots of subscribers ("consumers").

“Message”: From a technical standpoint, a key value pair. From a non-technical standpoint, a relatively small number of bytes (think hundreds to a few thousand bytes).

If the above isn't your planned use case, Kafka *may not* be the solution you are looking for. Contact your favorite Cloudera representative to discuss and find out. It is better to understand what you can and cannot do upfront than to go ahead based on some enthusiastic arbitrary vendor message with a solution that will not meet your expectations in the end.

What is Kafka designed for?

Kafka was designed at LinkedIn to be a horizontally scaling publish-subscribe system. It offers a great deal of configurability at the system- and message-level to achieve these performance goals. There are well documented cases ([Uber](#) and [LinkedIn](#)) that showcase how well Kafka can scale when everything is done right.

What is Kafka not well fitted for (or what are the tradeoffs)?

It's very easy to get caught up in all the things that Kafka can be used for without considering the tradeoffs. Kafka configuration is also not automatic. You need to understand each of your use cases to determine which configuration properties can be used to tune (and retune!) Kafka for each use case.

Some more specific examples where you need to be deeply knowledgeable and careful when configuring are:

- *Using Kafka as your microservices communication hub*

Kafka can replace both the message queue and the services discovery part of your software infrastructure. However, this is generally at the cost of some added latency as well as the need to monitor a new complex system (i.e. your Kafka cluster).

- *Using Kafka as long-term storage*

While Kafka does have a way to configure message retention, it's primarily designed for low latency message delivery. Kafka does not have any support for the features that are usually associated with filesystems (such as metadata or backups). As such, using some form of long-term ingestion, such as HDFS, is recommended instead.

- *Using Kafka as an end-to-end solution*

Kafka is only part of a solution. There are a lot of best practices to follow and support tools to build before you can get the most out of it (see this wise [LinkedIn post](#)).

- *Deploying Kafka without the right support*

Uber has given some numbers for their engineering organization. These numbers could help give you an idea what it takes to reach that kind of scale: [1300 microservers](#), [2000 engineers](#).

Where can I get a general Kafka overview?

The first four sections (Introduction, Setup, Client Basics, Broker Details) of the [CDH 6 Kafka Documentation](#) cover the basics and design of Kafka. This should serve as a good starting point. If you have any remaining questions after reading that documentation, come to this FAQ or talk to your favorite Cloudera representative about training or a best practices deep dive.

Where does Kafka fit well into an Analytic DB solution?

ADB deployments benefit from Kafka by utilizing it for data ingest. Data can then populate tables for various analytics workloads. For ad hoc BI the real-time aspect is less critical, but the ability to utilize the same data used in real time applications, in BI and analytics as well, is a benefit that Cloudera's platform provides, as you will have Kafka for both purposes, already integrated, secured, governed and centrally managed.

Where does Kafka fit well into an Operational DB solution?

Kafka is commonly used in the real-time, mission-critical world of Operational DB deployments. It is used to ingest data and allow immediate serving to other applications and services via Kudu or HBase. The benefit of utilizing Kafka in Cloudera's platform for ODB is the integration, security, governance and central management. You avoid the risks and costs of siloed architecture and "yet another solution" to support.

What is a Kafka consumer?

If Kafka is the system that stores messages, then a consumer is the part of your system that reads those messages from Kafka.

While Kafka does come with a command line tool that can act as a consumer, practically speaking, you will most likely write Java code using the KafkaConsumer API for your production system.

What is a Kafka producer?

While consumers read from a Kafka cluster, producers write to a Kafka cluster.

Like the consumer (see previous question paragraph), your producer will also be custom Java code for your particular use case.

Your producer may need some tuning for write performance and SLA guarantees, but will generally be simpler (fewer error cases) to tune than your consumer.

What functionality can I call in my Kafka Java code?

The best way to get more information on what functionality you can call in your Kafka Java code is to look at the Java docs. And read very carefully!

What's a good size of a Kafka record if I care about performance and stability?

There is an older blog post from 2014 from LinkedIn titled: [Benchmarking Apache Kafka: 2 Million Writes Per Second \(On Three Cheap Machines\)](#). In the "Effect of Message Size" section, you can see two charts which indicate that Kafka throughput starts being affected at a record size of 100 bytes through 1000 bytes and bottoming out around 10000 bytes. In general, keeping topics specific and keeping message sizes deliberately small will help you get the most out of Kafka.

Excerpting from [Deploying Apache Kafka: A Practical FAQ](#):

How to send large messages or payloads through Kafka?

Cloudera benchmarks indicate that Kafka reaches maximum throughput with message sizes of around 10KB. Larger messages will show decreased throughput. However, in certain cases, users need to send messages much larger than 10K.

If the message payload sizes are in the order of 100s of MB, we recommend exploring the following alternatives:

Kafka Frequently Asked Questions

- If shared storage is available (HDFS, S3, NAS), place the large payload on shared storage and use Kafka just to send a message with the payload location.
- Handle large messages by chopping them into smaller parts before writing into Kafka, using a message key to make sure all the parts are written to the same partition so that they are consumed by the same Consumer, and re-assembling the large message from its parts when consuming.

Where can I get Kafka training?

You have many options. Cloudera provides training as listed in the next two questions. You can also ask your Resident Solution Architect to do a deep dive on Kafka architecture and best practices. And you could always engage in the community to get insight and expertise on specific topics.

Where can I get basic Kafka training?

Cloudera training offers [a basic on-demand training for Kafka](#)¹.

This covers basics of Kafka architecture, messages, ordering, and a few slides (code examples) of (to my knowledge) an older version of the Java API. It also covers using Flume + Kafka.

Where can I get Kafka developer training?

Kafka developer training is included in Cloudera's [Developer Training for Apache Spark and Hadoop](#)².

Use Cases

Like most Open Source projects, Kafka provides a lot of configuration options to maximize performance. In some cases, it is not obvious how best to map your specific use case to those configuration options. We attempt to address some of those situations below.

What can I do to ensure that I never lose a Kafka event?

This is a simple question which has lots of far-reaching implications for your entire Kafka setup. A complete answer includes the next few related FAQs and their answers.

What is the recommended node hardware for best reliability?

Operationally, you need to make sure your Kafka cluster meets the following hardware setup:

1. Have a 3 or 5 node cluster only running Zookeeper (higher only necessary at largest scales).
2. Have at least a 3 node cluster only running Kafka.
3. Have the disks on the Kafka cluster running in RAID 10. (Required for resiliency against disk failure.)
4. Have sufficient memory for both the Kafka and Zookeeper roles in the cluster. (Recommended: 4GB for the broker, the rest of memory automatically used by the kernel as file cache.)
5. Have sufficient disk space on the Kafka cluster.
6. Have a sufficient number of disks to handle the bandwidth requirements for Kafka and Zookeeper.
7. You need a number of nodes greater than or equal to the highest replication factor you expect to use.

What are the network requirements for best reliability?

Kafka expects a reliable, low-latency connection between the brokers and the Zookeeper nodes.

1. The number of network hops between the Kafka cluster and the Zookeeper cluster is relatively low.
2. Have highly reliable network services (such as DNS).

What are the system software requirements for best reliability?

Assuming you're following the recommendations of the previous two questions, the actual system outside of Kafka must be configured properly.

1. The kernel must be configured for maximum I/O usage that Kafka requires.
 - a. Large page cache
 - b. Maximum file descriptions
 - c. Maximum file memory map limits
2. Kafka JVM configuration settings:
 - a. Brokers generally don't need more than 4GB-8GB of heap space.
 - b. Run with the +G1GC garbage collection using Java 8 or later.

How can I configure Kafka to ensure that events are stored reliably?

The following recommendations for Kafka configuration settings make it extremely difficult for data loss to occur.

1. Producer
 - a. `block.on.buffer.full=true`
 - b. `retries=Long.MAX_VALUE`
 - c. `acks=all`
 - d. `max.in.flight.requests.per.connections=1`
 - e. Remember to close the producer when it is finished or when there is a long pause.
2. Broker
 - a. `Topic.replication.factor >= 3`
 - b. `Min.insync.replicas = 2`
 - c. Disable unclean leader election
3. Consumer
 - a. Disable `auto.offset.commit`
 - b. Commit offsets after messages are processed by your consumer client(s).

If you have more than 3 hosts, you can increase the broker settings appropriately on topics that need more protection against data loss.

Once I've followed all the previous recommendations, my cluster should never lose data, right?

Kafka does **not ensure** that data loss never occurs. There will always be the following tradeoffs:

1. Throughput vs. reliability. For example, the higher the replication factor, the more resilient your setup will be against data loss. However, to make those extra copies takes time and can affect throughput.
2. Reliability vs. free disk space. Extra copies due to replication use up disk space that would otherwise be used for storing events.

Beyond the above design tradeoffs, there are also the following issues:

- To ensure events are consumed you need to monitor your Kafka brokers and topics to verify sufficient consumption rates are sustained to meet your ingestion requirements.
- Ensure that replication is enabled on any topic that requires consumption guarantees. This protects against Kafka broker failure and host failure.
- Kafka is designed to store events for a defined duration after which the events will be deleted. You can increase the duration that events will be retained up to the amount of supporting storage space.
- You will always run out of disk space unless you add more nodes to the cluster.

My Kafka events must be processed in order. How can I accomplish this?

Once your topic is configured with partitions, Kafka will send each record (based on key/value pair) to a particular partition based on key. So, any given key, the corresponding records will be “in order” within a partition.

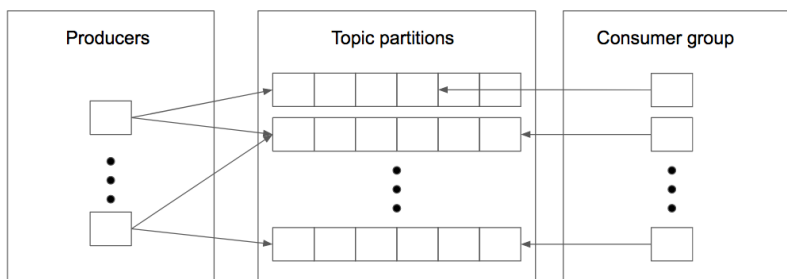
For global ordering, you have two options:

1. Your topic must consist of one partition (but a higher replication factor could be useful for redundancy and failover). However, this will result in very limited message throughput.
2. You configure your topic with a small number of partitions and perform the ordering after the consumer has pulled data. This does not result in **guaranteed** ordering, but, given a large enough time window, will likely be equivalent.

Conversely, it is best to take Kafka’s partitioning design into consideration when designing your Kafka setup rather than rely on global ordering of events.

How do I size my topic? *Alternatively: What is the “right” number of partitions for a topic?*

Choosing the proper number of partitions for a topic is the key to achieving a high degree of parallelism with respect to writes to and reads and to distribute load. Evenly distributed load over partitions is a key factor to have good throughput (avoid hot spots). Making a good decision requires estimation based on the desired throughput of producers and consumers per partition.



For example, if you want to be able to read 1 GB / sec, but your consumer is only able process 50 MB / sec, then you will need at least 20 partitions and 20 consumers in the consumer group. Similarly, if you want to achieve the same for producers, and 1 producer can only write at 100 MB / sec, you will need 10 partitions. In this case, if you have 20 partitions, you can maintain 1 GB /s for producing and consuming messages. You should adjust the exact number of partitions to number of consumers or producers, so that each consumer and producer achieve their target throughput.

So a simple formula could be:

$$\#Partitions = \max(N_P, N_C)$$

where:

- N_P is the number of required producers determined by calculating: T_T/T_P
- N_C is the number of required consumers determined by calculating: T_T/T_C
- T_T is the total expected throughput for our system
- T_P is the max throughput of a single producer to a single partition
- T_C is the max throughput of a single consumer from a single partition

This calculation gives you a rough indication of the number of partitions. It's a good place to start. Keep in mind the following considerations for improving the number of partitions after you have your system in place:

- The number of partitions can be specified at topic creation time or later.
- Increasing the number of partitions also affects the number of open file descriptors. So make sure you set file descriptor limit properly.
- Reassigning partitions can be very expensive, and therefore it’s better to over- than under-provision.
- Changing the number of partitions that are based on keys is challenging and involves manual copying (see [Apache Kafka Administration](#) on page 65).

- Reducing the number of partitions is not currently supported. Instead, create a new a topic with a lower number of partitions and copy over existing data.
- Metadata about partitions are stored in ZooKeeper in the form of znodes. Having a large number of partitions has effects on ZooKeeper and on client resources:
 - Unneeded partitions put extra pressure on ZooKeeper (more network requests), and might introduce delay in controller and/or partition leader election if a broker goes down,
 - Producer and consumer clients need more memory, because they need to keep track of more partitions and also buffer data for all partitions.
- As guideline for optimal performance, you should not have more than 4000 partitions per broker and not more than 200,000 partitions in a cluster.

Make sure consumers don't lag behind producers by monitoring consumer lag. To check consumers' position in a consumer group (that is, how far behind the end of the log they are), use the following command:

```
$ kafka-consumer-groups --bootstrap-server BROKER_ADDRESS --describe --group
CONSUMER_GROUP --new-consumer
```

How can I scale a topic when it's already deployed in production?

Recall the following facts about Kafka:

1. When you create a topic, you set the number of partitions. The higher the partition count, the better the parallelism and the better the events are spread somewhat evenly through the cluster.
2. In most cases, as events go to the Kafka cluster, events with the same key go to the same partition. This is a consequence of using a hash function to determine which key goes to which partition.

Now, you might assume that scaling means increasing the number of partitions in a topic. However, due to the way hashing works, simply increasing the number of partitions means that you will lose the “events with the same key go to the same partition” fact.

Given that, there are two options:

1. Your cluster may not be scaling well because the partition loads are not balanced properly (for example, one broker has 4 very active partitions, while another has 0). In those cases, you can use the `kafka-reassign-partitions` script to manually do some partition balancing.
2. Create a new topic with more partitions, pause the producers, copy data over from the old topic, and then move the producers and consumers over to the new topic. This can be a bit tricky operationally.

How do I rebalance my Kafka cluster?

This one comes up when a customer adds new nodes or disks to existing nodes. Partitions are not automatically balanced. If a topic already has a number of nodes equal to the replication factor (typically 3), then adding disks will not help with rebalancing.

Using the `kafka-reassign-partitions` command after adding new hosts is the recommended method.

Caveats

There are several caveats to using this command:

- It is highly recommended that you minimize the volume of replica changes. Say, instead of moving ten replicas with a single command, move two at a time to keep the cluster healthy.
- It is not possible to use this command to make an out-of-sync replica into the leader partition.
- If too many replicas are moved, then there could be serious performance impact on the cluster. It is recommended that when using the `kafka-reassign-partitions` command that you look at the partition counts and sizes. From there, you can test various partition sizes along with the `--throttle` flag to determine what volume of data can be copied without affecting broker performance significantly.
- Given the earlier restrictions, it is best to use this command only when all brokers and topics are healthy.

Kafka Frequently Asked Questions

How do I monitor my Kafka cluster?

As of Cloudera Enterprise 5.14, Cloudera Manager has monitoring for a Kafka cluster.

Currently, there are three GitHub projects as well that provide additional monitoring functionality:

- [Doctor Kafka](#)³ (Pinterest, Apache 2.0 License)
- [Kafka Manager](#)⁴ (Yahoo, Apache 2.0 License)
- [Cruise Control](#)⁵ (LinkedIn, BSD 2-clause License)

All of these projects are Apache-compatible licensed, but are not Open Source (no community, bug filing, or transparency).

What are the best practices concerning consumer `group.id`?

The `group.id` is just a string that helps Kafka track which consumers are related (by having the same group id).

- In general, timestamps as part of `group.id` are not useful. Since `group.ids` will correspond to multiple consumers, you cannot have a unique timestamp for each consumer.
- Add any helpful identifiers. This could be related to a group (for example, transactions, marketing), purpose (fraud, alerts), or technology (flume, spark).

How do I monitor consumer group lag?

This is typically done using the `kafka-consumer-groups` command line tool. Copying directly from the [upstream documentation](#)⁶, we have this example output (reformatted for readability):

```
$ bin/kafka-consumer-groups.sh --bootstrap-server localhost:9092 --describe --group my-group
TOPIC PARTITION CURRENT-OFFSET LOG-END-OFFSET LAG CONSUMER-ID HOST
CLIENT-ID
my-topic 0 2 4 2 consumer-1-69d6 /127.0.0.1
consumer-1
my-topic 1 2 3 1 consumer-1-69d6 /127.0.0.1
consumer-1
my-topic 2 2 3 1 consumer-2-9bb2 /127.0.0.1
consumer-2
```

In general, if everything is going well with a particular topic, each consumer's `CURRENT-OFFSET` should be up-to-date or nearly up-to-date with the `LOG-END-OFFSET`. From this command, you can determine whether a particular host or a particular partition is having issues keeping up with the data rate.

How do I reset the consumer offset to an arbitrary value?

This is also done using the `kafka-consumer-groups` command line tool. This is generally an administration feature used to get around corrupted records, data loss, or recovering from failure of the broker or host. Aside from those special cases, using the command line tool for this purpose is not recommended.

By using the `--execute --reset-offsets` flags, you can change the consumer offsets for a consumer group (or even all groups) to a specific setting based on each partitions log's beginning/end or a fixed timestamp. Typing the `kafka-consumer-groups` command with no arguments will give you the complete help output.

How do I configure MirrorMaker for bi-directional replication across DCs?

Mirror Maker is a one way copy of one or more topics from a Source Kafka Cluster to a Destination Kafka Cluster. Given this restriction on Mirror Maker, you need to run two instances, one to copy from A to B and another to copy from B to A.

You may also wish to pay attention to the following:

- Cloudera recommends using the "pull" model for Mirror Maker, meaning that the Mirror Maker instance writing to the destination is running on a host "near" the destination cluster.

- The topics must be unique across the two clusters being copied.
- On secure clusters, the source cluster and destination cluster must be in the same Kerberos realm.

How does the consumer max retries vs timeout work?

With the newer versions of Kafka, consumers have two ways they communicate with brokers.

- Retries: This is generally related to reading data. When a consumer reads from a brokers, it's possible for that attempt to fail due to problems such as intermittent network outages or I/O issues on the broker. To improve reliability, the consumer will retry (up to the configured `max.retries` value) before actually failing to read a log offset.
- Timeout. This term is a bit vague, since there are two timeouts related to consumers:
 - Poll Timeout: This is the timeout between calls to `KafkaConsumer.poll()`. This timeout is set based on whatever read latency requirements your particular use case needs.
 - Heartbeat Timeout: The newer consumer has a "heartbeat thread" which will heartbeat to the broker (actually the GroupCoordinator within a broker) to let the broker know that the consumer is still alive. This happens on a regular basis and if the broker doesn't receive at least one heartbeat within the timeout period, it will assume the consumer is dead and disconnect it.

How do I size my Kafka cluster?

There are several considerations for sizing your Kafka cluster.

- *Disk space*

Disk space will primarily consist of your Kafka data and broker logs. When in debug mode, the broker logs can get quite large (10s to 100s of GB), so reserving a significant amount of space could save you some future headaches.

For Kafka data, you need to perform estimates on message size, number of topics, and redundancy. Also remember that you will be using RAID10 for Kafka's data, so half your hard drives will go towards redundancy. From there, you can calculate how many drives will be needed.

In general, you will want to have more hosts than the minimum suggested by the number of drives. This leaves room for growth and some scalability headroom.

- *Zookeeper nodes*

One node is fine for a test cluster. Three is standard for most Kafka clusters. At really large scale, five nodes is fairly common for reliability.

- *Looking at leader partition count/bandwidth usage*

This is likely the metric with the highest variability. Any Kafka broker will get overloaded if it has too many leader partitions. In the worst cases, each leader partition requires high bandwidth, high message rates, or both. For other topics, leader partitions will be a tiny fraction of what a broker can handle (limited by software and hardware). We recommend that you group topics by partition data throughput requirements and try to figure out a good average (such as 2 high bandwidth data partitions, 4 medium bandwidth data partitions, 20 small bandwidth data partitions) that works on a per-host basis. From there, you can determine how many hosts are needed.

How can I combine Kafka with Flume to ingest into HDFS?

We have two blog posts on using Kafka with Flume:

- The original post: [Flafka: Apache Flume Meets Apache Kafka for Event Processing](#)
- This updated version for CDH 5.8/Apache Kafka 0.9/Apache Flume 1.7: [New in Cloudera Enterprise 5.8: Flafka Improvements for Real-Time Data Ingest](#)

How can I build a Spark streaming application that consumes data from Kafka?

You will need to set up your development environment to use both Spark libraries and Kafka libraries:

Kafka Frequently Asked Questions

- [Building Spark Applications](#)
- The [kafka-examples](#) directory on Cloudera's public GitHub has an example pom.xml.

From there, you should be able to read data using the `KafkaConsumer` class and using Spark libraries for real-time data processing. The blog post [Reading data securely from Apache Kafka to Apache Spark](#) has a pointer to a GitHub repo which contains a word count example.

For further background, read the blog post [Architectural Patterns for Near Real-Time Data Processing with Apache Hadoop](#).

References

1. Kafka basic training: <https://ondemand.cloudera.com/courses/course-v1:Cloudera+Kafka+201601/info>
2. Kafka developer training: <https://ondemand.cloudera.com/courses/course-v1:Cloudera+DevSH+201709/info>
3. Doctor Kafka: <http://github.com/pinterest/doctorkafka>
4. Kafka manager: <http://github.com/yahoo/kafka-manager>
5. Cruise control: <http://github.com/linkedin/cruise-control>
6. Upstream documentation: http://kafka.apache.org/documentation/#basic_ops_consumer_lag

Appendix: Apache License, Version 2.0

SPDX short identifier: Apache-2.0

Apache License
Version 2.0, January 2004
<http://www.apache.org/licenses/>

TERMS AND CONDITIONS FOR USE, REPRODUCTION, AND DISTRIBUTION

1. Definitions.

"License" shall mean the terms and conditions for use, reproduction, and distribution as defined by Sections 1 through 9 of this document.

"Licensor" shall mean the copyright owner or entity authorized by the copyright owner that is granting the License.

"Legal Entity" shall mean the union of the acting entity and all other entities that control, are controlled by, or are under common control with that entity. For the purposes of this definition, "control" means (i) the power, direct or indirect, to cause the direction or management of such entity, whether by contract or otherwise, or (ii) ownership of fifty percent (50%) or more of the outstanding shares, or (iii) beneficial ownership of such entity.

"You" (or "Your") shall mean an individual or Legal Entity exercising permissions granted by this License.

"Source" form shall mean the preferred form for making modifications, including but not limited to software source code, documentation source, and configuration files.

"Object" form shall mean any form resulting from mechanical transformation or translation of a Source form, including but not limited to compiled object code, generated documentation, and conversions to other media types.

"Work" shall mean the work of authorship, whether in Source or Object form, made available under the License, as indicated by a copyright notice that is included in or attached to the work (an example is provided in the Appendix below).

"Derivative Works" shall mean any work, whether in Source or Object form, that is based on (or derived from) the Work and for which the editorial revisions, annotations, elaborations, or other modifications represent, as a whole, an original work of authorship. For the purposes of this License, Derivative Works shall not include works that remain separable from, or merely link (or bind by name) to the interfaces of, the Work and Derivative Works thereof.

"Contribution" shall mean any work of authorship, including the original version of the Work and any modifications or additions to that Work or Derivative Works thereof, that is intentionally submitted to Licensor for inclusion in the Work by the copyright owner or by an individual or Legal Entity authorized to submit on behalf of the copyright owner. For the purposes of this definition, "submitted" means any form of electronic, verbal, or written communication sent to the Licensor or its representatives, including but not limited to communication on electronic mailing lists, source code control systems, and issue tracking systems that are managed by, or on behalf of, the Licensor for the purpose of discussing and improving the Work, but excluding communication that is conspicuously marked or otherwise designated in writing by the copyright owner as "Not a Contribution."

"Contributor" shall mean Licensor and any individual or Legal Entity on behalf of whom a Contribution has been received by Licensor and subsequently incorporated within the Work.

2. Grant of Copyright License.

Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable copyright license to reproduce, prepare Derivative Works of, publicly display, publicly perform, sublicense, and distribute the Work and such Derivative Works in Source or Object form.

3. Grant of Patent License.

Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable (except as stated in this section) patent license to make, have made, use, offer to sell, sell, import, and otherwise transfer the Work, where such license applies only to those patent claims

licensable by such Contributor that are necessarily infringed by their Contribution(s) alone or by combination of their Contribution(s) with the Work to which such Contribution(s) was submitted. If You institute patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Work or a Contribution incorporated within the Work constitutes direct or contributory patent infringement, then any patent licenses granted to You under this License for that Work shall terminate as of the date such litigation is filed.

4. Redistribution.

You may reproduce and distribute copies of the Work or Derivative Works thereof in any medium, with or without modifications, and in Source or Object form, provided that You meet the following conditions:

1. You must give any other recipients of the Work or Derivative Works a copy of this License; and
2. You must cause any modified files to carry prominent notices stating that You changed the files; and
3. You must retain, in the Source form of any Derivative Works that You distribute, all copyright, patent, trademark, and attribution notices from the Source form of the Work, excluding those notices that do not pertain to any part of the Derivative Works; and
4. If the Work includes a "NOTICE" text file as part of its distribution, then any Derivative Works that You distribute must include a readable copy of the attribution notices contained within such NOTICE file, excluding those notices that do not pertain to any part of the Derivative Works, in at least one of the following places: within a NOTICE text file distributed as part of the Derivative Works; within the Source form or documentation, if provided along with the Derivative Works; or, within a display generated by the Derivative Works, if and wherever such third-party notices normally appear. The contents of the NOTICE file are for informational purposes only and do not modify the License. You may add Your own attribution notices within Derivative Works that You distribute, alongside or as an addendum to the NOTICE text from the Work, provided that such additional attribution notices cannot be construed as modifying the License.

You may add Your own copyright statement to Your modifications and may provide additional or different license terms and conditions for use, reproduction, or distribution of Your modifications, or for any such Derivative Works as a whole, provided Your use, reproduction, and distribution of the Work otherwise complies with the conditions stated in this License.

5. Submission of Contributions.

Unless You explicitly state otherwise, any Contribution intentionally submitted for inclusion in the Work by You to the Licensor shall be under the terms and conditions of this License, without any additional terms or conditions. Notwithstanding the above, nothing herein shall supersede or modify the terms of any separate license agreement you may have executed with Licensor regarding such Contributions.

6. Trademarks.

This License does not grant permission to use the trade names, trademarks, service marks, or product names of the Licensor, except as required for reasonable and customary use in describing the origin of the Work and reproducing the content of the NOTICE file.

7. Disclaimer of Warranty.

Unless required by applicable law or agreed to in writing, Licensor provides the Work (and each Contributor provides its Contributions) on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied, including, without limitation, any warranties or conditions of TITLE, NON-INFRINGEMENT, MERCHANTABILITY, or FITNESS FOR A PARTICULAR PURPOSE. You are solely responsible for determining the appropriateness of using or redistributing the Work and assume any risks associated with Your exercise of permissions under this License.

8. Limitation of Liability.

In no event and under no legal theory, whether in tort (including negligence), contract, or otherwise, unless required by applicable law (such as deliberate and grossly negligent acts) or agreed to in writing, shall any Contributor be liable to You for damages, including any direct, indirect, special, incidental, or consequential damages of any character arising as a result of this License or out of the use or inability to use the Work (including but not limited to damages for loss of goodwill, work stoppage, computer failure or malfunction, or any and all other commercial damages or losses), even if such Contributor has been advised of the possibility of such damages.

9. Accepting Warranty or Additional Liability.

While redistributing the Work or Derivative Works thereof, You may choose to offer, and charge a fee for, acceptance of support, warranty, indemnity, or other liability obligations and/or rights consistent with this License. However, in accepting such obligations, You may act only on Your own behalf and on Your sole responsibility, not on behalf of any other Contributor, and only if You agree to indemnify, defend, and hold each Contributor harmless for any liability incurred by, or claims asserted against, such Contributor by reason of your accepting any such warranty or additional liability.

END OF TERMS AND CONDITIONS

APPENDIX: How to apply the Apache License to your work

To apply the Apache License to your work, attach the following boilerplate notice, with the fields enclosed by brackets "[]" replaced with your own identifying information. (Don't include the brackets!) The text should be enclosed in the appropriate comment syntax for the file format. We also recommend that a file or class name and description of purpose be included on the same "printed page" as the copyright notice for easier identification within third-party archives.

```
Copyright [yyyy] [name of copyright owner]

Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License.
```