

KIP-1009: Add Broker-level Throttle Configurations

- [Status](#)
- [Motivation](#)
- [Public Interfaces](#)
- [Compatibility, Deprecation, and Migration Plan](#)
- [Test Plan](#)
- [Rejected Alternatives](#)

Status

Current state: *Under Discussion*

Discussion thread: [here](#)

JIRA: [KAFKA-7983](#)

Motivation

For each partition of a Kafka topic, a set of in-sync replicas is maintained. If a replica becomes out-of-sync, it continually fetches missing data from the partition leader until it is back in sync.

Occasionally, we notice that large amounts of data need to be replicated to maintain the ISR set. This can happen if a broker is offline for hours, perhaps due to machine maintenance or hardware failures, or when a large number of partitions are reassigned.

Without any replication quotas, we observe significantly increased end-to-end latency on the cluster, measured by [Xinfra Monitor](#). In the worst cases, tail latency jumps from 10s of milliseconds to multiple seconds.

To address this, replication quotas were introduced in [KIP-73](#), and [KIP-542](#) improved upon replication throttling by applying it to out-of-sync replicas only. However, these throttles can only be configured for a specific topics and partitions. While this satisfies the need to throttle out-of-sync replication traffic while reassigning partitions, this solution is not sufficient at a larger scale, when an *entire* broker's traffic needs to be throttled.

Currently, replication throttles can only be set for a specific topic and/or partitions, and not as a blanket configuration for the whole broker. With the current configuration options for replication throttles, the only way to throttle an entire broker's traffic is to do the following:

1. Set `leader.replication.throttled.rate` and `follower.replication.throttled.rate` at the broker level.
2. For each topic on the broker, configure `leader.replication.throttled.replicas` and `follower.replication.throttled.replicas`.

Depending on how large the broker is, this process can be unnecessarily time consuming and expensive.

Opening this KIP to go along with the linked JIRA.

Public Interfaces

Two new dynamic broker configurations will be introduced to indicate if replication throttling is enabled for the entire broker.

New configurations:

- `leader.replication.throttled` - boolean representing if leader replication traffic is throttled on the whole broker
- `follower.replication.throttled` - boolean representing if follower replication traffic is throttled on the whole broker

This is an example of how leader replication throttling can be enabled for the whole broker:

```
bin/kafka-configs ... --alter
--add-config 'leader.replication.throttled=true'
--entity-type broker
--entity-name brokerId
```

Proposed Changes

Currently, setting replication throttles at the topic and partition levels is supported. This KIP proposes changing the replication logic to also support configuring throttles at the broker level - meaning all replication traffic to/from any partition on a specified broker will be throttled.

This will be done by introducing two new dynamic broker configurations - `leader.replication.throttled` and `follower.replication.throttled`, which are boolean values representing if leader/follower replication traffic is throttled on the whole broker. This also involves adding the configuration options, updating those values accordingly in the ConfigHandler, and updating the `isThrottled` function in the ReplicationQuotaManager.

With these proposed changes, an admin can throttle an entire broker's traffic by doing the following:

1. Set `leader.replication.throttled.rate` and `follower.replication.throttled.rate` at the broker level.

2. Set *leader.replication.throttled=true* and *follower.replication.throttled=true* at the broker level.

Compatibility, Deprecation, and Migration Plan

As the only exposed changes are configuration options, there is no impact to existing users.

Test Plan

New tests will be added to the `ReplicationQuotaManagerTest` suite.

Rejected Alternatives

This is not a rejected alternative to this KIP, but somewhat related and worth mentioning.

[KAFKA-10190](#) proposes a change to apply *follower.replication.throttled.rate*, *leader.replication.throttled.rate* and *replica.alter.log.dirs.io.max.bytes.per.second* configs at the broker level, but seems to have been abandoned. Unlike this KIP, KAFKA-10190 does not enable applying throttles at the broker level, just setting the throttle rate.

However, this could be a useful feature to include in this KIP as well.