

KIP-186: Increase offsets retention default to 7 days

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

Status

Current state: Accepted

Discussion thread: [here](#)

JIRA:

 Unable to render Jira issues macro, execution error.

PR: <https://github.com/apache/kafka/pull/4648>

Released: 2.0

Please keep the discussion on the mailing list rather than commenting on the wiki (wiki discussions get unwieldy fast).

Motivation

Users, especially new users, are frequently confused when the offsets for their application disappear after 24 hours. This differs from the default retention for topics so a user doing testing against a stock config can see the data is still there, but if their application is offline for more than 24h, it either starts reprocessing all the data if `auto.offset.reset=earliest` or skips some data if `auto.offset.reset=latest`.

There are at least a couple of cases this trips people up:

- Development against a test cluster where some development or debugging takes longer than 24h
- Periodic execution – despite being a streaming platform, some people have batch processes that run nightly or only on business days. The former might manage to just barely work, the latter loses offsets over the weekend.
- Topics with infrequent new data that may happen less than every 24h.

The commonly stated solution for this problem – commit offsets for all partitions periodically – only addresses the first issue and not the second. It is also unintuitive to users in the third case, where the assumption is that you only need to commit offsets when you process new data, although that could

potentially be addressed in other ways (see

 Unable to render Jira issues macro, execution error.

). Given how many people this trips

up, we should try to fix the problem.

Public Interfaces

The default value for offsets retention will be increased to 7 days:

`offsets.retention.minutes=10080`

Proposed Changes

The default value for offsets retention will be increased to 7 days. This matches the default retention for regular topics `log.retention.ms|minutes|hours`. This aligns better with user's expectations.

The only impact is that the default configuration will a) hold onto offsets for longer and b) hold onto offsets for more consumer groups for longer. (a) shouldn't be an issue since the offsets topic is compacted. According to Jun, (b) was one of the reasons for selecting the 24h retention and is potentially more of a concern since it increases the storage required for the offsets topic and the amount of memory required for the offset cache in the broker. However, the improved user experience outweighs this cost. The cost should only be large if users automatically create large numbers of ephemeral consumer groups. The most common case for this is probably use of the console consumer, which has offset commit enabled by default (despite most use being for ephemeral groups for debugging/inspection).

Compatibility, Deprecation, and Migration Plan

Users that have not overridden the value in their default configs will see increased log size for the offsets topic if they also use many ephemeral consumer groups.

In contrast, many users who have discovered they need to increase this default will no longer need to.

No migration tool is required, but we'll want to include a mention in the upgrade notes.

Test Plan

This is a simple change to an already tested configuration. Adjustments to existing unit tests will sufficiently cover testing this.

Rejected Alternatives

None