KIP-943: Add independent “offset.storage.segment.bytes” for connect-distributed.properties

Status

Current state: *Under Discussion*

Discussion thread: *here*

JIRA: *here*

Motivation

As the config 'segment.bytes' for internal topic related connect cluster(offset.storage.topic), if following the default configuration of the broker or set it larger, then when the connect cluster runs many and complicated tasks(for example, Involving a lot of topic and partition replication), especially the log volume of the topic 'offset.storage.topic' is very large, it will affect the restart speed of the connect workers. The actual impact can be seen in the jira link above.

After investigation, the reason is that a consumer needs to be started to read the data of ‘offset.storage.topic’ at startup. Although this topic is set to compact, if the 'segment size' is set to a large value, such as the default value of 1GB, then this topic may have tens of gigabytes of data(if following default 25 partitions) that cannot be compacted and has to be read from the earliest (because the active segment cannot be cleaned), which will consume a lot of time and caused the worker to be unable to start and execute tasks for a long time.

Therefore, I want to extract the “segment.bytes” settings for “offset.storage.topic” separately, just like "offsets.topic.segment.bytes" and "transaction.state.log.segment.bytes", where the size is set by the user, and if there is no explicit setting, give the default value such as 50MB. In this way to avoid receiving interference from kafka broker configuration. As for "config.storage.topic" and "status.storage.topic", It is difficult to write a large amount of data in practical use, and it may not be necessary to take similar measures.

Public Interfaces

Add a config in connect-distributed.properties:

- config name: offset.storage.segment.bytes
- config definition: Represents the “log.segment.bytes” size of the “offset.storage.topic”

Proposed Changes

For new connect cluster, if “offset.storage.segment.bytes” is explicitly set in connect-distributed.properties, the size is set as KIP-605, It all depends on the users themselves. Otherwise, the default size (50MB) is taken. Compared to the previous behavior, this setting has nothing to do with the configuration of kafka broker, although it is possible that the configuration value of kafka broker is smaller and better than the default value 50MB.

For existed connect cluster, if “offset.storage.segment.bytes” is explicitly set in connect-distributed.properties, we will update topic config by admin client. Otherwise, the default size (50MB) is taken to add or update topic config by admin client.

Compatibility, Deprecation, and Migration Plan

- There is no impact on existing users if just upgrade the new version.
- Just upgrade new version and do not need to do another thing.

Test Plan

It is only necessary to verify that the configuration is successfully set up as expected through a unit test.

Rejected Alternatives

1. Set a upper limit segment size for "offset.storage.topic", but it will override the value that user set, which means user has no right to decide.
2. Expand to consumer concurrency, if network transfer between the worker and the Kafka cluster is the limiting factor, then it won't have any impact.