

KIP-215: Add topic regex support for Connect sinks

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

Status

Current state: Adopted

Discussion thread: [msg81955](#)

Voting thread: [msg82225](#)

JIRA: [KAFKA-3073](#)

PR: [kafka#4151](#)

Released: 1.1.0

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

Motivation

Gwen Shapira describe the motivation well in the JIRA:

*KafkaConsumer supports both a list of topics or a pattern when subscribing.
KafkaConnect only supports a list of topics, which is not just more of a hassle to configure - it also requires more maintenance.*

We should provide a configuration option for Connect sinks to specify a regular expression instead of an explicit topic list.

Public Interfaces

This KIP introduces a new 'topics.regex' configuration option for Kafka Connect sinks that expects a string compatible with Java's regex Pattern class. Users may specify only one of 'topics' or 'topics.regex'.

Proposed Changes

If 'topics.regex' is specified, its string value will be used to instantiate a regex Pattern which will be passed to `subscribe` instead of a list of strings.

If both 'topics.regex' and 'topics' are specified, a `ConfigException` will be thrown to prevent startup.

Compatibility, Deprecation, and Migration Plan

'topics.regex' will default to the empty string "" so that existing configurations will not be affected by the additional option.

Rejected Alternatives

There is some concern that this solution leans on the Java-specific regular expression syntax. We could choose instead to support a less language-specific regular expression specification such as PCRE or [re2](#).

Kafka Connect, however, already relies on Java-specific Patterns in configurations such as the `RegexRouter` transform, so the proposal in this KIP stays consistent with that.

Taking on evaluation of a new regular expression specification would significantly enlarge the scope of this change and seems more appropriate for a separate effort that would add that support consistently across Kafka Connect configurations.

Additional regular expression specifications could be accommodated in the future by adding an additional configuration option 'topics.regex.type'.