

KIP-190: Handle client-ids consistently between clients and brokers

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

Status

Current state: Accepted

Discussion thread: [here](#)

JIRA: [KAFKA-5735](#)

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

Motivation

At the moment, client-ids are not handled consistently between clients and brokers. For example, the Java clients and tools prevent client-ids to contain a number of characters. These restrictions are not enforced by the brokers and are simply not mentioned in the documentation. Moreover, many third party clients, including `librdkafka`, don't have these restrictions. This leads to issues creating metrics and handling quotas like

 Unable to render Jira issues macro, execution error.

Public Interfaces

The current client-id characters restrictions will be removed from the Java clients (Producer, Consumer, Connect, Admin, Streams) allowing to use any string as a client-id. In addition, `kafka-configs.sh` will also be updated to allow setting quotas on any client-id.

Some metrics containing a client-id (for example: `kafka.server:type={Produce|Fetch},user={[-.\\w]+},client-id={[-.\\w]+}`) will change as the client-id will now be URL-encoded.

Proposed Changes

- Client-ids will be sanitized by the brokers using URL-encoding similar to the encoding used for the user Principal. This leaves metrics for client-ids using the current limited set of characters as-is, but will allow arbitrary characters in encoded form.
- Quota metrics and sensors as well as ZooKeeper quota configuration paths should use sanitized ids for client-ids.
- Remove client-id validation in `kafka-configs.sh` and in Java clients and allow any characters for client-id similar to usernames, URL-encoding the names to generate ZK path

Compatibility, Deprecation, and Migration Plan

- The only change of behaviour will be on JMX metrics containing a client-id. The client-id will now be URL-encoded, only affecting client-ids containing special characters.

Rejected Alternatives

Instead of properly URL-encoding client-ids, we first looked into simply using `ObjectName.quote()` when using the client-id in metrics. This would change all existing metrics as we would be quoted. Also it feels cleaner to use the same sanitizing logic for both client-ids and user principals.