

KIP-855: Add schema.namespace parameter to SetSchemaMetadata SMT in Kafka Connect

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

Status

Current state: Under Discussion

Discussion thread: <https://lists.apache.org/thread/3hkd9lljobf9rl56ogjpcbo4ldoxcz5n>

JIRA: [KAFKA-7883](#)

Pull request changes: <https://github.com/apache/kafka/pull/11442/files>

Motivation

SetSchemaMetadata SMT doesn't support transformation of the source schema namespace separately from the schema name transformation. When using Kafka Connect and Avro Converter with Schema Registry you may have to transform your source schema namespace independently from the schema name(s). Another use-case as mentioned in [KAFKA-7883](#), when use JDBC connector with multiple tables in the source database you may have to preserve the original table names but only specify some non-empty namespace for them.

Public Interfaces

Adding `schemaNamespace` attribute to the class `org.apache.kafka.connect.transforms.SetSchemaMetadata`.

From the perspective of Kafka Connect SMT configuration it looks like adding of a new optional attribute `"transforms.transformschema.schema.namespace"`:

```
{
  "transforms" : "TransformSchema",
  "transforms.TransformSchema.type" : "org.apache.kafka.connect.transforms.SetSchemaMetadata$Value",
  "transforms.TransformSchema.schema.namespace" : "my.new.namespace",
  "transforms.transformschema.schema.name" : "NewSchemaName",
  "transforms.TransformSchema.schema.version" : 42
}
```

Proposed Changes

Adding of a new `schema.namespace` attribute to `SetSchemaMetadata` SMT letting developer to decide either to transform namespace only or transform both schema name and schema namespace by using either `schema.namespace` together with `schema.name` or just single `schema.name` attribute. In `SetSchemaMetadata` class we should concatenate values from `schema.namespace` and `schema.name` while we also should keep backward compatibility and leave all `schema.namespace`, `schema.name` and `schema.version` parameters optional.

Compatibility, Deprecation, and Migration Plan

The added `"transforms.transformschema.schema.namespace"` option's default value should be `null`, to be compatible with the old behavior.

Rejected Alternatives

Currently `SetSchemaMetadata` SMT provides a way to transform the namespace by specifying it directly in the `schema.name` parameter in concatenation with the target schema name like this:

```
{
  "transforms" : "TransformSchema",
  "transforms.TransformSchema.type" : "org.apache.kafka.connect.transforms.SetSchemaMetadata$Value",
  "transforms.TransformSchema.schema.name" : "my.new.namespace.NewSchemaName"
}
```

But with this way we're unable to preserve the original schema name because the new schema name gets hardcoded in the `schema.name` parameter. To be able to transform namespace only and leave the schema name unchanged we have to provide a way to specify schema namespace separately from the schema name.