

Para que Medium funcione, registramos los datos del usuario. Al utilizar Medium , acepta nuestra [Política de privacidad](#) , incluida la política de cookies.

[Seguir](#)

590K Seguidores



Tiene **1** historia sólo para miembros libre que queda de este mes.

[Regístrese en Medium y obtenga uno extra](#)

# Kafka ya no requiere ZooKeeper

La versión 2.8.0 le brinda acceso anticipado a Kafka sin guardián del zoológico



Giorgos Myrianthous · 27 abr · 5 min de lectura ★

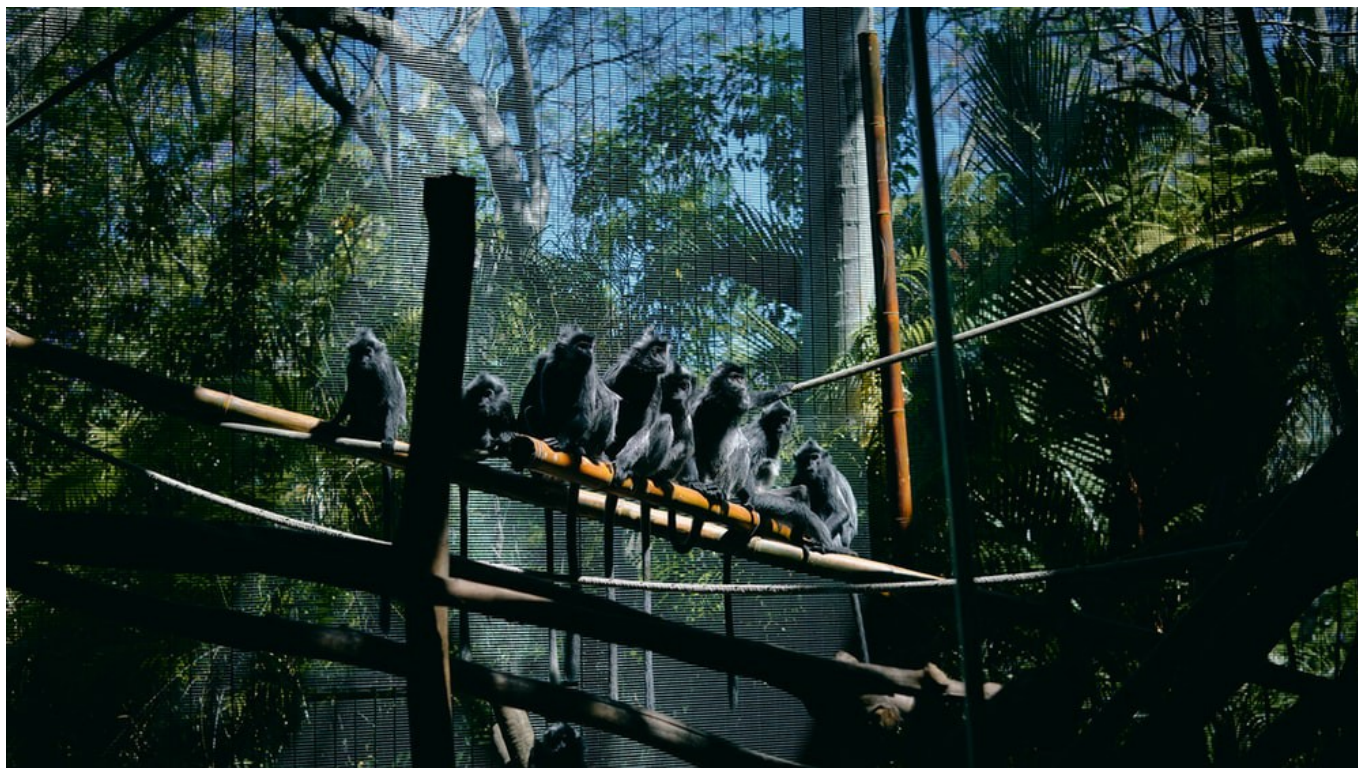


Foto de [Christian Lambert](#) en [Unsplash](#)

## Introducción

Para que Medium funcione, registramos los datos del usuario. Al utilizar Medium , acepta nuestra [Política de privacidad](#) , incluida la política de cookies.



basa en un quórum interno de Raft que se puede activar a través del **modo de metadatos de Kafka Raft** . La nueva función simplifica la administración de la infraestructura y la administración del clúster y marca una nueva era para Kafka.

## Kafka sin cuidador

En este artículo vamos a discutir por qué t h ERE era una necesidad para la eliminación de la dependencia ZooKeeper en el primer lugar. Además, discutiremos cómo ZooKeeper ha sido reemplazado por a `KRaft mode` partir de la versión 2.8.0 y también exploraremos el impacto de eliminar esta dependencia y cómo Kafka se beneficiará de esta mejora.

### ¿Qué es el KPI-500?

Hasta ahora, Kafka utilizaba Apache ZooKeeper como almacén de metadatos. Los metadatos para las particiones y los corredores se almacenaron en el quórum de ZooKeeper, que también era responsable de la elección del controlador de Kafka.

ZooKeeper es un sistema externo a Kafka, lo que significa que viene con su propia sintaxis de configuración, herramientas de administración y mejores prácticas. Por lo tanto, si quisiera implementar un clúster de Kafka, también tendría que administrar, implementar y monitorear Zookeeper. Dado que estos dos sistemas distribuidos requieren una configuración diferente, el nivel de complejidad aumenta y, por lo tanto, es mucho más fácil para los administradores de sistemas cometer errores. Tener dos sistemas también conduce a la duplicación del trabajo; por ejemplo, para habilitar las funciones de seguridad, tendría que aplicar la configuración relevante a ambos servicios.

Tener un almacén de metadatos externo es ineficaz en términos de recursos, ya que necesita ejecutar procesos adicionales. Además, esto limita la escalabilidad del propio Kafka. Cada vez que se inicia el clúster, el controlador de Kafka debe cargar el estado del clúster desde ZooKeeper.

Para que Medium funcione, registramos los datos del usuario. Al utilizar Medium , acepta nuestra [Política de privacidad](#) , incluida la política de cookies. ×

vuelve más ineficiente con el tiempo, lo que lleva a procesos de alta carga y, por lo tanto, limita la cantidad de particiones que el clúster puede almacenar.

KPI son las siglas de Kafka Improvement Proposals y [KPI-500](#) presenta la arquitectura fundamental de Kafka sin ZooKeeper.

## Presentamos el modo de metadatos de Kafka Raft

La versión 2.8.0 introduce un acceso anticipado a Zookeeper-Less Kafka como parte de KPI-500. Sin embargo, tenga en cuenta que la implementación está parcialmente completa y, por lo tanto, no **debe utilizarla en entornos de producción** .

En la última versión, ZooKeeper puede ser reemplazado por un quórum interno de controladores de Raft. Cuando el **modo de metadatos de Kafka Raft** está habilitado, Kafka almacenará sus metadatos y configuraciones en un tema llamado `@metadata` . Este tema interno lo administra el quórum interno y se replica en todo el clúster. Los nodos del clúster ahora pueden servir como intermediarios, controladores o ambos (denominados nodos *combinados* ).

When `KRaft mode` is enabled, only a few selected servers can serve as controllers and these will compose the internal quorum. The controllers can either be in active or standby mode that will eventually take over if the current active controller server fails or is taken down.

Now every *Kafka Server* comes with an additional configuration parameter called `process.roles` . This parameter can take the following values:

- `broker` : The Kafka Server will serve as a broker
- `controller` : The Kafka Server will serve as a controller of the internal Raft quorum
- `broker,controller` : The Kafka Server will serve as both a controller of the quorum and a broker

Para que Medium funcione, registramos los datos del usuario. Al utilizar Medium , acepta nuestra [Política de privacidad](#) , incluida la política de cookies.



the only way you can activate `KRaft mode` .

Additionally, each node is now identified with its `node.id` and must now provide `controller.quorum.voters` configuration parameter which is the equivalent to `zookeeper.connect` in ZooKeeper mode. This parameter is used to identify the controller servers of the internal quorum and the format used for the values is `serverID@host:port, serverID@host:port, ....`

Now let's assume that within our Kafka Cluster which is running in `KRaft mode` we have 7 brokers and 3 controllers. The snippet below demonstrates an example configuration for one of the controller servers in the Raft quorum.

```
process.roles=controller
node.id=1
listeners=CONTROLLER://controller-1-host:9092
controller.quorum.voters=1@controller-1-host:9092,2@controller-2-
host:9092,3@controller-3-host:9092
```

Likewise, the below configuration demonstrates how to setup one of the brokers of the cluster:

```
process.roles=broker
node.id=4
listeners=PLAINTEXT://:9092
controller.quorum.voters=1@controller-1-host:9092,2@controller-2-
host:9092,3@controller-3-host:9092
```

Para que Medium funcione, registramos los datos del usuario. Al utilizar Medium , acepta nuestra [Política de privacidad](#) , incluida la política de cookies. ×

platform. The whole community (and I'd say Confluent in particular) has been working towards this direction for the last few years. The earliest release has been a tremendous effort of the whole Kafka community that is still working on improvements so that ZooKeeper-Less Kafka mode is feature-complete within this year.

## We've been headed this direction for years

— Jason Gustafson @ [Kafka Summit San Francisco 2019](#)

The removal of Apache ZooKeeper dependency simplifies the infrastructure management for Kafka deployments. Kafka and ZooKeeper are two distinct services — now that Kafka has been unified so that it is not dependent on external services to server as metadata stores, the learning curve will be shortened and this will eventually help broaden Kafka's adoption.

Furthermore, this enhancement offers a more scalable and robust overall architecture. As discussed earlier, in ZooKeeper mode Kafka had to store its metadata into ZooKeeper nodes. Every time the cluster was starting up or a controller election was happening, Kafka Controllers had to read the metadata from an external service which was inefficient. By replacing ZooKeeper with this internal Raft quorum, deployments can now support more partitions.

Removing the ZooKeeper dependency also enables the support of clusters with single node. When you want test Kafka as part of a proof of concept you no longer have to start multiple processes.

## Conclusion

The removal of Zookeeper dependency is a huge step forward for Kafka. In fact, the new `KRaft mode` feature will extend scalability capabilities of Apache Kafka and also shorten

Para que Medium funcione, registramos los datos del usuario. Al utilizar Medium , acepta nuestra [Política de privacidad](#) , incluida la política de cookies. ×

In addition to KPI-500, Kafka 2.8.0 comes with numerous improvements and bug fixes so make sure to take a look at the [release notes](#).

**Finally, I want to highlight once again that this is currently an early-access which means it should not be used in production environments.**

Gracias a Anupam Chugh .

---

## Regístrese en The Variable

Por Towards Data Science

Todos los jueves, Variable ofrece lo mejor de Towards Data Science: desde tutoriales prácticos e investigación de vanguardia hasta características originales que no querrá perderse. [Echar un vistazo](#).

Reciba este boletín

Deberá iniciar sesión o crear una cuenta para recibir este boletín.

[Programación](#) [Kafka](#) [Desarrollo de software](#) [Big Data](#) [Ciencia de los datos](#)

[About](#) [Help](#) [Legal](#)

Obtén la aplicación Medium

