

# KIP-723: Add `socket.tcp.no.delay` property to Kafka Config

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

## Status

**Current state:** *Under Discussion*

**Discussion thread:** [here](#)

**JIRA:** [here](#)

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

## Motivation

Now `TCP_NODELAY` socket option is **always** enabled and large number of topic-partitions on one broker causes burst host's packets/sec metric. This can cause service degradation in cases where you are using a cloud traffic shaper or other network controls.

For example, for test cluster with 4 brokers and 30 000 topic-partitions:

	enabled <code>TCP_NODELAY</code>	disabled <code>TCP_NODELAY</code>
idle	~15 000 tcp packets/sec	~300 tcp packets/sec
load	~140 000 tcp packets/sec	~3 000 tcp packets/sec

~99.999% of all packets is inter broker messages. Size of packets 40-160 bytes.

In our production cluster ~27k topic-partitions at 16 brokers generate ~500 000 tcp packets per sec (IN and OUT summary).

More how reproduce, research and fix this issue in description of JIRA ticket.

## Public Interfaces

New server and client property (boolean, default value = true):

`socket.tcp.no.delay`

## Proposed Changes

New boolean property in Kafka Config that forwarded into

- Acceptor class in core project
- Selector class in clients project

## Compatibility, Deprecation, and Migration Plan

New changes has 100% compatibility with all kafka releases and may be backported to any previous release with no changes. No deprecation or migration plan are needed. The default value is the current behaviour.

## Rejected Alternatives

At this moment all new connection are created with enabled `TCP_NODELAY` socket option. The "true" value is hardcoded since 2011 year. There is no way to change this option outside.