

HCP Understanding Profiler 1

## Understanding Profiling

**Date of Publish:** 2019-3-18



<https://docs.hortonworks.com>

# Contents

<b>Understanding Profiling.....</b>	<b>3</b>
-------------------------------------	----------

## Understanding Profiling

A profile describes the behavior of an entity on a network. This feature is typically used by a data scientist and you should coordinate with the data scientist to determine if they need your assistance with customizing the Profiler values.

HCP installs the Profiler which runs as an independent Apache Storm topology. The configuration for the Profiler topology is stored in Apache ZooKeeper at `/metron/topology/profiler`. These properties also exist in the default installation of HCP at `$METRON_HOME/config/zookeeper/profiler.json`. You can change the values on disk and then upload them to ZooKeeper using `$METRON_HOME/bin/zk_load_configs.sh`.



### Note:

The Profiler can persist any serializable object, not just numeric values.

HCP supports the following profiler properties:

<b>profiler.workers</b>	The number of worker processes to create for the topology.
<b>profiler.executors</b>	The number of executors to spawn per component.
<b>profiler.input.topic</b>	The name of the Kafka topic from which to consume data.
<b>profiler.output.topic</b>	The name of the Kafka topic to which profile data is written. Only used with profiles that use the <code>trriage`result field](#result)</code> .
<b>profiler.period.duration</b>	The duration of each profile period. Define this value along with <code>profiler.period.duration.units</code> .
<b>profiler.period.duration.units</b>	The units used to specify the profile period duration. Define this value along with <code>profiler.period.duration</code> .
<b>profiler.ttl</b>	If a message has not been applied to a Profile in this period of time, the Profile is forgotten and its resources cleaned up. Define this value along with <code>profiler.ttl.units</code> .
<b>profiler.ttl.units</b>	The units used to specify <code>profiler.ttl</code> .
<b>profiler.hbase.salt.divisor</b>	A salt is prepended to the row key to help prevent hotspotting. This constant is used to generate the salt. Ideally, this constant should be roughly equal to the number of nodes in the Apache HBase cluster.
<b>profiler.hbase.table</b>	The name of the HBase table that profiles are written to.
<b>profiler.hbase.column.family</b>	The column family used to store profiles.
<b>profiler.hbase.batch</b>	The number of puts written in a single batch.
<b>profiler.hbase.flush.interval.seconds</b>	The maximum number of seconds between batch writes to HBase.