

每种类型的 HBase Sink 都有一些特定的参数。正如前面所讨论的，HBase Sink 有能力写入到安全 HBase 集群。因此，需要在配置中添加安全相关的参数。表 5-6 是 HBase Sink 接受的参数，而 Async HBase Sink 不使用这些参数（与其他组件一样，为 Async HBase Sink 在配置中添加这些参数没有造成危害——只是被忽略）[hbase-security]。

表5-6 HBase Sink安全配置

参数	默认值	描述
kerberosPrincipal	-	用来登录到 KDC 使用的 Kerberos 主体
kerberosKeytab	-	登录到 KDC 的 Kerberos Principal 使用的 keytab 文件的路径

这些参数和 `hdfs.kerberosPrincipal`、`hdfs.kerberosKeytab` 参数有完全相同的含义。

除了两类 Sink 共同的参数之外，Async HBase Sink 还接受一个参数：`timeout`，是整个批处理成功写出到 HBase 的时间周期。如果整个批处理在该时间周期内没有成功写出，那么 Sink 利用 Channel 回滚整个事务，且使这些事件对于该 Sink 和其他 Sink 都是可用的。该参数的默认值是 60000。

例 5-3 中展示了配置有自定义序列化器的 Async HBase Sink 的一个例子，在配置中指定了 `zookeeper quorum`。

例5-3 Async HBase Sink配置的一个例子

```
agent.sinks = asynchbase
agent.channels = memory

agent.sinks.asynchbase.type = asynchbase
agent.sinks.asynchbase.channel = memory
agent.sinks.asynchbase.zookeeperQuorum = zk1.usingflume.com:2181,
zk2.usingflume.com:2181,zk3.usingflume.com:2181
agent.sinks.asynchbase.znodeParent = /hbase
agent.sinks.asynchbase.table = usingFlumeTable
agent.sinks.asynchbase.columnFamily = usingFlumeFamily
agent.sinks.asynchbase.batchSize = 1000
agent.sinks.asynchbase.timeout = 60000
agent.sinks.asynchbase.serializer = usingflume.ch05.AsyncHBaseDirectSerializer

117 agent.channels.memory.type = memory
agent.channels.memory.size = 100000
```

用序列化器将 Flume 事件转换成 HBase Put 和 Increment*

HBase Sink 可以在 HBase 中写数据或增量计数器。就像 HDFS Sink，HBase Sink 允许用