

Hi @nthampi ,

Thanks for the response. Please see the logs and screenshots below.

Starting job elastic at 08:38 25/02/2019.

[statistics] connecting to socket on port 3933

[statistics] connected

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/C:/Apps/ETLTALEND/Talend-Studio-20161216_1026-V6.3.1/workspace/.Java/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/C:/Apps/ETLTALEND/Talend-Studio-20161216_1026-V6.3.1/workspace/.Java/lib/slf4j-log4j12-1.7.16.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/C:/Apps/ETLTALEND/Talend-Studio-20161216_1026-V6.3.1/workspace/.Java/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]

[ERROR]: org.apache.hadoop.util.Shell - Failed to locate the winutils binary in the hadoop binary path

java.io.IOException: Could not locate executable C:\Apps\hadoop-common-2.2.0-bin-master\bin\bin\winutils.exe in the Hadoop binaries.

at org.apache.hadoop.util.Shell.getQualifiedBinPath(Shell.java:356)

at org.apache.hadoop.util.Shell.getWinUtilsPath(Shell.java:371)

at org.apache.hadoop.util.Shell.<clinit>(Shell.java:364)

at org.apache.hadoop.util.StringUtils.<clinit>(StringUtils.java:80)

at org.apache.hadoop.security.SecurityUtil.getAuthenticationMethod(SecurityUtil.java:610)

at org.apache.hadoop.security.UserGroupInformation.initialize(UserGroupInformation.java:272)

at org.apache.hadoop.security.UserGroupInformation.ensureInitialized(UserGroupInformation.java:260)

at org.apache.hadoop.security.UserGroupInformation.loginUserFromSubject(UserGroupInformation.java:790)

at org.apache.hadoop.security.UserGroupInformation.getLoginUser(UserGroupInformation.java:760)

at org.apache.hadoop.security.UserGroupInformation.getCurrentUser(UserGroupInformation.java:633)

at org.apache.spark.util.Utils\$\$anonfun\$getCurrentUserName\$1.apply(Utils.scala:2245)

at org.apache.spark.util.Utils\$\$anonfun\$getCurrentUserName\$1.apply(Utils.scala:2245)

at scala.Option.getOrElse(Option.scala:121)

at org.apache.spark.util.Utils\$.getCurrentUserName(Utils.scala:2245)

at org.apache.spark.SparkContext.<init>(SparkContext.scala:297)

at org.apache.spark.api.java.JavaSparkContext.<init>(JavaSparkContext.scala:58)

at local_project.elastic_1_0.elastic.runJobInTOS(elastic.java:1378)

at local_project.elastic_1_0.elastic.main(elastic.java:1265)

[WARN]: org.apache.hadoop.util.NativeCodeLoader - Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

[WARN]: org.apache.spark.SparkConf - In Spark 1.0 and later spark.local.dir will be overridden by the value set by the cluster manager (via SPARK_LOCAL_DIRS in mesos/standalone and LOCAL_DIRS in YARN).

[WARN]: org.apache.spark.deploy.client.StandaloneAppClient\$ClientEndpoint - Failed to connect to master localhost:7077

org.apache.spark.SparkException: Exception thrown in awaitResult

at org.apache.spark.rpc.RpcTimeout\$\$anonfun\$1.applyOrElse(RpcTimeout.scala:77)

at org.apache.spark.rpc.RpcTimeout\$\$anonfun\$1.applyOrElse(RpcTimeout.scala:75)

at scala.runtime.AbstractPartialFunction.apply(AbstractPartialFunction.scala:36)

at org.apache.spark.rpc.RpcTimeout\$\$anonfun\$addMessageIfTimeout\$1.applyOrElse(RpcTimeout.scala:59)

at org.apache.spark.rpc.RpcTimeout\$\$anonfun\$addMessageIfTimeout\$1.applyOrElse(RpcTimeout.scala:59)

at scala.PartialFunction\$OrElse.apply(PartialFunction.scala:167)

at org.apache.spark.rpc.RpcTimeout.awaitResult(RpcTimeout.scala:83)

at org.apache.spark.rpc.RpcEnv.setupEndpointRefByURI(RpcEnv.scala:88)

at org.apache.spark.rpc.RpcEnv.setupEndpointRef(RpcEnv.scala:96)

at org.apache.spark.deploy.client.StandaloneAppClient\$ClientEndpoint\$\$anonfun\$tryRegisterAllMasters\$1\$\$anon\$1.run(StandaloneAppClient.scala:109)

at java.util.concurrent.Executors\$RunnableAdapter.call(Executors.java:511)

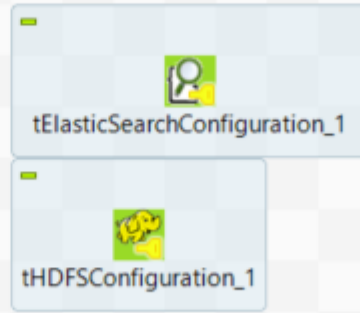
at java.util.concurrent.FutureTask.run(FutureTask.java:266)

```
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748)
Caused by: java.io.IOException: Failed to connect to localhost/127.0.0.1:7077
at org.apache.spark.network.client.TransportClientFactory.createClient(TransportClientFactory.java:228)
at org.apache.spark.network.client.TransportClientFactory.createClient(TransportClientFactory.java:179)
at org.apache.spark.rpc.netty.NettyRpcEnv.createClient(NettyRpcEnv.scala:197)
at org.apache.spark.rpc.netty.Outbox$$anon$1.call(Outbox.scala:191)
at org.apache.spark.rpc.netty.Outbox$$anon$1.call(Outbox.scala:187)
... 4 more
Caused by: java.net.ConnectException: Connection refused: no further information: localhost/127.0.0.1:7077
at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)
at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:717)
at io.netty.channel.socket.nio.NioSocketChannel.doFinishConnect(NioSocketChannel.java:224)
at io.netty.channel.nio.AbstractNioChannel$AbstractNioUnsafe.finishConnect(AbstractNioChannel.java:289)
at io.netty.channel.nio.NioEventLoop.processSelectedKey(NioEventLoop.java:528)
at io.netty.channel.nio.NioEventLoop.processSelectedKeysOptimized(NioEventLoop.java:468)
at io.netty.channel.nio.NioEventLoop.processSelectedKeys(NioEventLoop.java:382)
at io.netty.channel.nio.NioEventLoop.run(NioEventLoop.java:354)
at io.netty.util.concurrent.SingleThreadEventExecutor$2.run(SingleThreadEventExecutor.java:111)
... 1 more
[WARN ]: org.apache.spark.deploy.client.StandaloneAppClient$ClientEndpoint - Failed to connect to master localhost:7077
org.apache.spark.SparkException: Exception thrown in awaitResult
at org.apache.spark.rpc.RpcTimeout$$anonfun$1.applyOrElse(RpcTimeout.scala:77)
at org.apache.spark.rpc.RpcTimeout$$anonfun$1.applyOrElse(RpcTimeout.scala:75)
at scala.runtime.AbstractPartialFunction.apply(AbstractPartialFunction.scala:36)
at org.apache.spark.rpc.RpcTimeout$$anonfun$addMessageIfTimeout$1.applyOrElse(RpcTimeout.scala:59)
at org.apache.spark.rpc.RpcTimeout$$anonfun$addMessageIfTimeout$1.applyOrElse(RpcTimeout.scala:59)
at scala.PartialFunction$OrElse.apply(PartialFunction.scala:167)
at org.apache.spark.rpc.RpcTimeout.awaitResult(RpcTimeout.scala:83)
at org.apache.spark.rpc.RpcEnv.setupEndpointRefByURI(RpcEnv.scala:88)
at org.apache.spark.rpc.RpcEnv.setupEndpointRef(RpcEnv.scala:96)
at org.apache.spark.deploy.client.StandaloneAppClient$ClientEndpoint$$anonfun$tryRegisterAllMasters$1$$anon$1.run(StandaloneAppClient.scala:109)
at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
at java.util.concurrent.FutureTask.run(FutureTask.java:266)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748)
Caused by: java.io.IOException: Failed to connect to localhost/127.0.0.1:7077
at org.apache.spark.network.client.TransportClientFactory.createClient(TransportClientFactory.java:228)
at org.apache.spark.network.client.TransportClientFactory.createClient(TransportClientFactory.java:179)
at org.apache.spark.rpc.netty.NettyRpcEnv.createClient(NettyRpcEnv.scala:197)
at org.apache.spark.rpc.netty.Outbox$$anon$1.call(Outbox.scala:191)
at org.apache.spark.rpc.netty.Outbox$$anon$1.call(Outbox.scala:187)
... 4 more
Caused by: java.net.ConnectException: Connection refused: no further information: localhost/127.0.0.1:7077
at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)
at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:717)
at io.netty.channel.socket.nio.NioSocketChannel.doFinishConnect(NioSocketChannel.java:224)
at io.netty.channel.nio.AbstractNioChannel$AbstractNioUnsafe.finishConnect(AbstractNioChannel.java:289)
at io.netty.channel.nio.NioEventLoop.processSelectedKey(NioEventLoop.java:528)
at io.netty.channel.nio.NioEventLoop.processSelectedKeysOptimized(NioEventLoop.java:468)
at io.netty.channel.nio.NioEventLoop.processSelectedKeys(NioEventLoop.java:382)
```

```
at io.netty.channel.nio.NioEventLoop.run(NioEventLoop.java:354)
at io.netty.util.concurrent.SingleThreadEventExecutor$2.run(SingleThreadEventExecutor.java:111)
... 1 more
[WARN ]: org.apache.spark.deploy.client.StandaloneAppClient$ClientEndpoint - Failed to connect to master localhost:7077
org.apache.spark.SparkException: Exception thrown in awaitResult
at org.apache.spark.rpc.RpcTimeout$$anonfun$1.applyOrElse(RpcTimeout.scala:77)
at org.apache.spark.rpc.RpcTimeout$$anonfun$1.applyOrElse(RpcTimeout.scala:75)
at scala.runtime.AbstractPartialFunction.apply(AbstractPartialFunction.scala:36)
at org.apache.spark.rpc.RpcTimeout$$anonfun$addMessageIfTimeout$1.applyOrElse(RpcTimeout.scala:59)
at org.apache.spark.rpc.RpcTimeout$$anonfun$addMessageIfTimeout$1.applyOrElse(RpcTimeout.scala:59)
at scala.PartialFunction$OrElse.apply(PartialFunction.scala:167)
at org.apache.spark.rpc.RpcTimeout.awaitResult(RpcTimeout.scala:83)
at org.apache.spark.rpc.RpcEnv.setupEndpointRefByURI(RpcEnv.scala:88)
at org.apache.spark.rpc.RpcEnv.setupEndpointRef(RpcEnv.scala:96)
at org.apache.spark.deploy.client.StandaloneAppClient$ClientEndpoint$$anonfun$tryRegisterAllMasters$1$$anon$1.run(StandaloneAppClient.scala:109)
at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
at java.util.concurrent.FutureTask.run(FutureTask.java:266)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748)
Caused by: java.io.IOException: Failed to connect to localhost/127.0.0.1:7077
at org.apache.spark.network.client.TransportClientFactory.createClient(TransportClientFactory.java:228)
at org.apache.spark.network.client.TransportClientFactory.createClient(TransportClientFactory.java:179)
at org.apache.spark.rpc.netty.NettyRpcEnv.createClient(NettyRpcEnv.scala:197)
at org.apache.spark.rpc.netty.Outbox$$anon$1.call(Outbox.scala:191)
at org.apache.spark.rpc.netty.Outbox$$anon$1.call(Outbox.scala:187)
... 4 more
Caused by: java.net.ConnectException: Connection refused: no further information: localhost/127.0.0.1:7077
at sun.nio.ch.SocketChannelImpl.checkConnect(Native Method)
at sun.nio.ch.SocketChannelImpl.finishConnect(SocketChannelImpl.java:717)
at io.netty.channel.socket.nio.NioSocketChannel.doFinishConnect(NioSocketChannel.java:224)
at io.netty.channel.nio.AbstractNioChannel$AbstractNioUnsafe.finishConnect(AbstractNioChannel.java:289)
at io.netty.channel.nio.NioEventLoop.processSelectedKey(NioEventLoop.java:528)
at io.netty.channel.nio.NioEventLoop.processSelectedKeysOptimized(NioEventLoop.java:468)
at io.netty.channel.nio.NioEventLoop.processSelectedKeys(NioEventLoop.java:382)
at io.netty.channel.nio.NioEventLoop.run(NioEventLoop.java:354)
at io.netty.util.concurrent.SingleThreadEventExecutor$2.run(SingleThreadEventExecutor.java:111)
... 1 more
[ERROR]: org.apache.spark.scheduler.cluster.StandaloneSchedulerBackend - Application has been killed. Reason: All masters are unresponsive! Giving up.
[WARN ]: org.apache.spark.scheduler.cluster.StandaloneSchedulerBackend - Application ID is not initialized yet.
[WARN ]: org.apache.spark.deploy.client.StandaloneAppClient$ClientEndpoint - Drop UnregisterApplication(null) because has not yet connected to master
[ERROR]: org.apache.spark.MapOutputTrackerMaster - Error communicating with MapOutputTracker
java.lang.InterruptedException
at java.util.concurrent.locks.AbstractQueuedSynchronizer.tryAcquireSharedNanos(AbstractQueuedSynchronizer.java:1326)
at scala.concurrent.impl.Promise$DefaultPromise.tryAwait(Promise.scala:208)
at scala.concurrent.impl.Promise$DefaultPromise.ready(Promise.scala:218)
at scala.concurrent.impl.Promise$DefaultPromise.result(Promise.scala:223)
at scala.concurrent.Await$$anonfun$result$1.apply(package.scala:190)
at scala.concurrent.BlockContext$DefaultBlockContext$.blockOn(BlockContext.scala:53)
at scala.concurrent.Await$.result(package.scala:190)
at org.apache.spark.rpc.RpcTimeout.awaitResult(RpcTimeout.scala:81)
at org.apache.spark.rpc.RpcEndpointRef.askWithRetry(RpcEndpointRef.scala:102)
at org.apache.spark.rpc.RpcEndpointRef.askWithRetry(RpcEndpointRef.scala:78)
```

```
at org.apache.spark.MapOutputTracker.askTracker(MapOutputTracker.scala:100)
at org.apache.spark.MapOutputTracker.sendTracker(MapOutputTracker.scala:110)
at org.apache.spark.MapOutputTrackerMaster.stop(MapOutputTracker.scala:580)
at org.apache.spark.SparkEnv.stop(SparkEnv.scala:84)
at org.apache.spark.SparkContext$$anonfun$stop$11.apply$mcV$sp(SparkContext.scala:1795)
at org.apache.spark.util.Utills$.tryLogNonFatalError(Utills.scala:1267)
at org.apache.spark.SparkContext.stop(SparkContext.scala:1794)
at org.apache.spark.scheduler.cluster.StandaloneSchedulerBackend.dead(StandaloneSchedulerBackend.scala:140)
at org.apache.spark.deploy.client.StandaloneAppClient$ClientEndpoint.markDead(StandaloneAppClient.scala:263)
at org.apache.spark.deploy.client.StandaloneAppClient$ClientEndpoint$$anon$2.run(StandaloneAppClient.scala:134)
at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
at java.util.concurrent.FutureTask.run(FutureTask.java:266)
at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.access$201(ScheduledThreadPoolExecutor.java:180)
at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.run(ScheduledThreadPoolExecutor.java:293)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748)
[ERROR]: org.apache.spark.util.Utills - Uncaught exception in thread appclient-registration-retry-thread
org.apache.spark.SparkException: Error communicating with MapOutputTracker
at org.apache.spark.MapOutputTracker.askTracker(MapOutputTracker.scala:104)
at org.apache.spark.MapOutputTracker.sendTracker(MapOutputTracker.scala:110)
at org.apache.spark.MapOutputTrackerMaster.stop(MapOutputTracker.scala:580)
at org.apache.spark.SparkEnv.stop(SparkEnv.scala:84)
at org.apache.spark.SparkContext$$anonfun$stop$11.apply$mcV$sp(SparkContext.scala:1795)
at org.apache.spark.util.Utills$.tryLogNonFatalError(Utills.scala:1267)
at org.apache.spark.SparkContext.stop(SparkContext.scala:1794)
at org.apache.spark.scheduler.cluster.StandaloneSchedulerBackend.dead(StandaloneSchedulerBackend.scala:140)
at org.apache.spark.deploy.client.StandaloneAppClient$ClientEndpoint.markDead(StandaloneAppClient.scala:263)
at org.apache.spark.deploy.client.StandaloneAppClient$ClientEndpoint$$anon$2.run(StandaloneAppClient.scala:134)
at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
at java.util.concurrent.FutureTask.run(FutureTask.java:266)
at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.access$201(ScheduledThreadPoolExecutor.java:180)
at java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.run(ScheduledThreadPoolExecutor.java:293)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748)
Caused by: java.lang.InterruptedException
at java.util.concurrent.locks.AbstractQueuedSynchronizer.tryAcquireSharedNanos(AbstractQueuedSynchronizer.java:1326)
at scala.concurrent.impl.Promise$DefaultPromise.tryAwait(Promise.scala:208)
at scala.concurrent.impl.Promise$DefaultPromise.ready(Promise.scala:218)
at scala.concurrent.impl.Promise$DefaultPromise.result(Promise.scala:223)
at scala.concurrent.Await$$anonfun$result$1.apply(package.scala:190)
at scala.concurrent.BlockContext$DefaultBlockContext$.blockOn(BlockContext.scala:53)
at scala.concurrent.Await$.result(package.scala:190)
at org.apache.spark.rpc.RpcTimeout.awaitResult(RpcTimeout.scala:81)
at org.apache.spark.rpc.RpcEndpointRef.askWithRetry(RpcEndpointRef.scala:102)
at org.apache.spark.rpc.RpcEndpointRef.askWithRetry(RpcEndpointRef.scala:78)
at org.apache.spark.MapOutputTracker.askTracker(MapOutputTracker.scala:100)
... 16 more
java.lang.NullPointerException
at org.apache.spark.SparkContext.<init>(SparkContext.scala:549)
at org.apache.spark.api.java.JavaSparkContext.<init>(JavaSparkContext.scala:58)
at local_project.elastic_1_0.elastic.runJobInTOS(elastic.java:1378)
```

```
at local_project.elastic_1_0.elastic.main(elastic.java:1265)
Exception in thread "main" java.lang.RuntimeException: TalendJob: 'elastic' - Failed with exit code: 1.
at local_project.elastic_1_0.elastic.main(elastic.java:1275)
[ERROR]: org.apache.spark.SparkContext - Error initializing SparkContext.
java.lang.NullPointerException
at org.apache.spark.SparkContext.<init>(SparkContext.scala:549)
at org.apache.spark.api.java.JavaSparkContext.<init>(JavaSparkContext.scala:58)
at local_project.elastic_1_0.elastic.runJobInTOS(elastic.java:1378)
at local_project.elastic_1_0.elastic.main(elastic.java:1265)
[ERROR]: local_project.elastic_1_0.elastic - TalendJob: 'elastic' - Failed with exit code: 1.
```



Designer Code Jobsript

Job(elastic 1.0) Component Test Cases Contexts(elastic) Run (Job elastic) Integration Action Modules

Job elastic

- Basic Run
- Spark Configuration**
- Debug Run
- Advanced settings
- Target Exec
- Memory Run

Use local mode

Cluster Version

Property Type Built-In

Distribution Amazon EMR Version EMR 5.0.0 (Apache 2.7.2)

Spark Mode Standalone

Configuration

Spark Host

Spark Home

Define the hadoop home directory

Define the driver hostname or IP address

Tuning

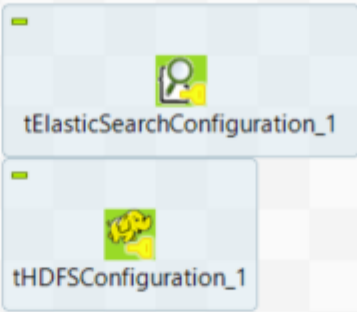
Set tuning properties

Spark "scratch" directory

Activate checkpointing

Advanced properties

Property	Value

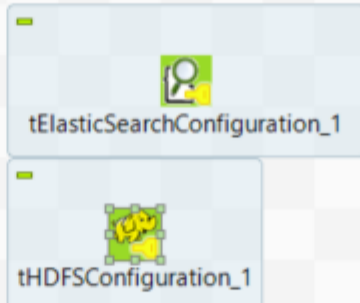


Designer Code Jobscript

Job(elastic 1.0) Component Test Cases Contexts(elastic) Run (Job elastic) Integration Action Modules

tElasticSearchOutput_1

Basic settings	Schema Built-In Edit schema Sync columns
Advanced settings	<input checked="" type="checkbox"/> Use an existing configuration Component List tElasticSearchConfiguration_1
Dynamic settings	Index "person" Type "_doc"
View	Output document JAVABEAN
Documentation	
Validation Rules	



Designer | Code | Jobscript

Job(elastic 1.0) | Component | Test Cases | Contexts(elastic) | Run (Job elastic) | Integration Action | Modules

tHDFSConfiguration_1

- Basic settings
- Advanced settings
- Dynamic settings
- View
- Documentation

Property Type: Built-In

Version: [Empty]

Distribution: Amazon EMR | Version: EMR 5.0.0 (Apache 2.7.2)

NameNode URI: "hdfs://localhost:9000/"

Authentication: [Empty]

Username: ""

Use Datanode Hostname

Temp Folder: "/tmp"

Property	Value

Encryption: Setup HDFS Encryption Configurations

Find:

row1

Column
cn
co

Var

elastic

Expression	Column
row1.cn	cn
row1.co	co

Schema editor Expression editor

row1

Column	K...	Type	<input checked="" type="checkbox"/> N.	Date Pattern (Ctrl+Space ...	Length	Precision	Default	Comment
cn	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		9			
co	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		100			

elastic

Column	K...	Type	<input checked="" type="checkbox"/> N.	Date Pattern (Ctrl+Space ...	Length	Precision	Default	Comment
cn	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		9			
co	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		100			

Apply