

Apache Kafka-0.8.1.1源码编译

经过近一个月时间，终于差不多将之前在Flume 0.9.4上面编写的source、sink等插件迁移到Flume-ng 1.5.0，包括了将Flume 0.9.4上面的TailSource、TailDirSource等插件的迁移（当然，我们加入了许多新的功能，比如故障恢复、日志的断点续传、按块发送日志以及每个一定的时间轮询发送日志而不是等一个日志发送完才发送另外一个日志）。现在我们需要将Flume-ng 1.5.0和最新的Kafka-0.8.1.1进行整合，今天这篇文章主要是说如何编译Kafka-0.8.1.1源码。

在讲述如何编译Kafka-0.8.1.1源码之前，我们先来了解一下什么是Kafka：

Kafka is a distributed, partitioned, replicated commit log service. It provides the functionality of a messaging system, but with a unique design.(Kafka是一个分布式的、可分区的(partitioned)、基于备份的(replicated)和commit-log存储的服务。它提供了类似于messaging system的特性,但是在设计实现上完全不同)。kafka是一种高吞吐量的分布式发布订阅消息系统，它有如下特性：

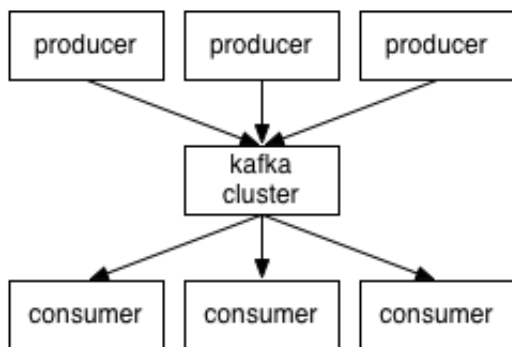
(1)、通过O(1)的磁盘数据结构提供消息的持久化，这种结构对于即使数以TB的消息存储也能够保持长时间的稳定性能。

(2)、高吞吐量：即使是非常普通的硬件kafka也可以支持每秒数十万的消息。

(3)、支持通过kafka服务器和消费机集群来分区消息。

(4)、支持Hadoop并行数据加载。

官方文档中关于kafka分布式订阅架构如下图：



欢迎关注微信公共帐号

好了，更多关于Kafka的介绍可以去<http://kafka.apache.org/>里面查看。现在我们正入正题，说说如何编译 Kafka-0.8.1.1，我们可以用Kafka里面自带的脚本进行编译；我们也可以用sbt进行

编译，sbt编译有点麻烦，我将在文章的后面进行介绍。

一、用Kafka里面自带的脚本进行编译

下载好了Kafka源码，里面自带了一个gradlew的脚本，我们可以利用这个编译Kafka源码：

```
# wget http://mirror.bit.edu.cn/apache/kafka/0.8.1.1/kafka-0.8.1.1-src.tgz
# tar -zxf kafka-0.8.1.1-src.tgz
# cd kafka-0.8.1.1-src
# ./gradlew releaseTarGz
```

运行上面的命令进行编译将会出现以下的异常信息：

```
:core:signArchives FAILED
```

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Execution failed for task ':core:signArchives'.
```

```
> Cannot perform signing task ':core:signArchives' because it
has no configured signatory
```

```
* Try:
```

```
Run with --stacktrace option to get the stack trace. Run with
--info or --debug option to get more log output.
```

```
BUILD FAILED
```

这是一个bug（<https://issues.apache.org/jira/browse/KAFKA-1297>），可以用下面的命令进行编译

```
./gradlew releaseTarGzAll -x signArchives
```

这时候将会编译成功（在编译的过程中将会出现很多的）。在编译的过程中，我们也可以指定对应的Scala版本进行编译：

```
./gradlew -PscalaVersion=2.10.3 releaseTarGz -x signArchives
```

编译完之后将会在core/build/distributions/里面生成kafka_2.10-0.8.1.1.tgz文件，这个和从网上下载的一样，可以直接用。

二、利用sbt进行编译

我们同样可以用sbt来编译Kafka，步骤如下：

```
# git clone https://git-wip-us.apache.org/repos/asf/kafka.git
# cd kafka
# git checkout -b 0.8 remotes/origin/0.8
# ./sbt update
[info] [SUCCESSFUL ] org.eclipse.jdt#core;3.1.1!core.jar (2243ms)
[info] downloading http://repo1.maven.org/maven2/ant/ant/1.6.5/ant-1.6.5.jar ...
[info] [SUCCESSFUL ] ant#ant;1.6.5!ant.jar (1150ms)
[info] Done updating.
[info] Resolving org.apache.hadoop#hadoop-core;0.20.2 ...
[info] Done updating.
[info] Resolving com.yammer.metrics#metrics-annotation;2.2.0 ...
[info] Done updating.
[info] Resolving com.yammer.metrics#metrics-annotation;2.2.0 ...
[info] Done updating.
[success] Total time: 168 s, completed Jun 18, 2014 6:51:38 PM

# ./sbt package
[info] Set current project to Kafka (in build file:/export1/spark/kafka/)
Getting Scala 2.8.0 ...
:: retrieving :: org.scala-sbt#boot-scala
confs: [default]
3 artifacts copied, 0 already retrieved (14544kB/27ms)
[success] Total time: 1 s, completed Jun 18, 2014 6:52:37 PM
```

对于Kafka 0.8及以上版本还需要运行以下的命令：

```
# ./sbt assembly-package-dependency
[info] Loading project definition from /export1/spark/kafka/project
[warn] Multiple resolvers having different access mechanism configured with
same name 'sbt-plugin-releases'. To avoid conflict, Remove duplicate project
resolvers (`resolvers`) or rename publishing resolver (`publishTo`).
[info] Set current project to Kafka (in build file:/export1/spark/kafka/)
```

```
[warn] Credentials file /home/wyp/.m2/.credentials does not exist
[info] Including slf4j-api-1.7.2.jar
[info] Including metrics-annotation-2.2.0.jar
[info] Including scala-compiler.jar
[info] Including scala-library.jar
[info] Including slf4j-simple-1.6.4.jar
[info] Including metrics-core-2.2.0.jar
[info] Including snappy-java-1.0.4.1.jar
[info] Including zookeeper-3.3.4.jar
[info] Including log4j-1.2.15.jar
[info] Including zkclient-0.3.jar
[info] Including jopt-simple-3.2.jar
[warn] Merging 'META-INF/NOTICE' with strategy 'rename'
[warn] Merging 'org/xerial/snappy/native/README' with strategy 'rename'
[warn] Merging 'META-INF/maven/org.xerial.snappy/snappy-java/LICENSE'
with strategy 'rename'
[warn] Merging 'LICENSE.txt' with strategy 'rename'
[warn] Merging 'META-INF/LICENSE' with strategy 'rename'
[warn] Merging 'META-INF/MANIFEST.MF' with strategy 'discard'
[warn] Strategy 'discard' was applied to a file
[warn] Strategy 'rename' was applied to 5 files
[success] Total time: 3 s, completed Jun 18, 2014 6:53:41 PM
```

当然，我们也可以直接在sbt里面指定scala的版本：

```
<!--
User: 过往记忆
Date: 14-6-18
Time: 20:20
bolg:
本文地址：/archives/1044
过往记忆博客，专注于hadoop、hive、spark、shark、flume的技术博客，大量的干货
过往记忆博客微信公共帐号：iteblog_hadoop
-->
sbt "++2.10.3 update"
sbt "++2.10.3 package"
sbt "++2.10.3 assembly-package-dependency"
```

本文地址：[《Apache Kafka-0.8.1.1源码编译》](#)
：/archives/1044，过往记忆博客，专注于hadoop、hive、spark、shark、flume的技术博客，大

量的干货.过往记忆博客微信公共帐号 : iteblog_hadoop

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接: [【】](#)（）