

一个 Flume Agent 的时候，这个名字被传递到 flume-ng 脚本，作为 -n 命令行选项的值。Flume 的配置系统将加载只与特定 Agent 名称相关的配置参数。因为多个 Agent 的配置可以在一个文件中，那么很容易将相同的文件部署到多个层，每一层都有不同的配置。在大多数情况下，因为一层的所有 Agent 与下一层 Agent 的相同集合通信，那层的每个 Agent 可以有一个完全相同的配置。在一层中，所有 Agent 使用相同名称和相同配置文件使得层的部署非常容易实现自动化。配置文件可以使用 -f 命令行选项传递到 Flume Agent。

对于配置目录，可使用 -c 命令行选项将其路径传递给 Flume Agent，它是两个重要文件 *flume-env.sh* 和 *log4j.properties* 的主目录。当启动一个 Flume Agent 时，Agent 初始化脚本将寻找 *flume-env.sh* 脚本。该文件必须包含需要传递给 flume-ng 脚本的任何环境变量。使用 *flume-env.sh* 文件初始化最常用的环境变量都列在表 2-1 中。

表2-1 使用*flume-env.sh*初始化的环境变量

环境变量	描述
FLUME_CLASSPATH	除了 Flume lib 和 <i>plugins.d</i> 目录外，传递到 Flume Agent 的类路径，这是自动添加的
JAVA_OPTS	任何传递给 JVM 的 Java 指定选项，包括 -XX 和 -D 参数，用来控制内存和从命令行传递参数
HADOOP_PREFIX (or) HADOOP_HOME	Hadoop 的安装目录 ( <i>\$HADOOP_PREFIX/bin</i> 包含 Hadoop 可执行文件)
HBASE_HOME	HBase 的安装目录 ( <i>\$HBASE_HOME/bin</i> 包含 HBase 可执行文件)

FLUME\_CLASSPATH 环境变量是 Flume 的 *lib* 和 *plugins.d* 目录之外的目录的列表(以:分隔)，这两个目录被添加到 Flume Agent 的类路径中。*plugins.d* 目录是存放自定义组件 JARs (Java Archive files) 的目录，所以 Flume 可以使用它们来加载自定义组件。

30

JAVA\_OPTS 是直接传递给 JVM 的参数集合。最常使用的是用来修改 JVM 分配的 heap size 的参数：-Xms 和 -Xmx。你可以在 Java 文档的 JAVA\_OPTS 中找到它们和其他的选项 [Java-commandline]。

HADOOP\_PREFIX (在旧的 Hadoop 版本 1 中，等价于 HADOOP\_HOME) 是 Hadoop 的安装目录。如果 hadoop 命令是 PATH，且 HADOOP\_PREFIX 没有设置，Flume 将使用 PATH。对于 HBase，HBASE\_HOME 实现相同的功能。

要运行一个名为 usingFlumeAgent 的 Flume Agent，其使用 Flume 配置 */home/usingflume/flume/flume.conf* 和配置目录 */home/usingflume/flume/conf*，则可使用下面的命令（假设目