Hadoop 本地库

Hadoop是使用Java语言开发的，但是有一些需求和操作并不适合使用Java，所以就引入了本地库（Native Libraries）的概念，通过本地库，Hadoop可以更加高效的执行某些操作。这些组件保存在Hadoop的一个独立的动态链接库（libhadoop.so）。 hadoop本地库仅支持linux及unix平台。

# 1、Native 文件压缩

目前在hadoop中，本地库应用的文件的压缩上，Hadoop现在有以下compression codec本地组件：

[zlib](http://www.zlib.net/)

[gzip](http://www.gzip.org/)

[lzo](http://www.oberhumer.com/opensource/lzo/)

在以上组件中，lzo和gzip压缩编解码器必须使用hadoop本地库才能运行。

Hadoop本地库使用ANSIC编写，使用GNU autotools工具链（autoconf,autoheader,autom

ke,autoscan,libttol）构建，也就是说构建hadoop库平台需要有标识C的编译器和GNU autoto

ols工具链。需要安装以下软件包：

C 编译器 (e.g. [GNU C Compiler](http://gcc.gnu.org/))

GNU Autools 工具链: [autoconf](http://www.gnu.org/software/autoconf/), [automake](http://www.gnu.org/software/automake/), [libtool](http://www.gnu.org/software/libtool/)

zlib开发包 (stable version >= 1.2.0)

lzo开发包 (stable version >= 2.0)

但是本地库的编译，需要通过ant来完成：

$ant -Dcompile.native=true <target>

也可以支持通过mvn命令，编译，但是其实现是通过调用ant来完成：

mvn package -Pdist,native -DskipTests

该功能的源码，所在的项目为：/hadoop-common/src/main/native/../compress目录下