Build and Install Hadoop 2.x or newer on Windows

1. Introduction

Hadoop version 2.2 onwards includes native support for Windows. The official Apache Hadoop releases do not include Windows binaries (yet, as of January 2014). However building a Windows package from the sources is fairly straightforward.

Hadoop is a complex system with many components. Some familiarity at a high level is helpful before attempting to build or install it or the first time. Familiarity with Java is necessary in case you need to troubleshoot.

2. Building Hadoop Core for Windows

2.1. Choose target OS version

The Hadoop developers have used **Windows Server 2008** and**Windows Server 2008 R2** during development and testing.**Windows Vista** and **Windows 7** are also likely to work because of the Win32 API similarities with the respective server SKUs. We have**not** tested on Windows XP or any earlier versions of Windows and these are not likely to work. Any issues reported on Windows XP or earlier will be closed as *Invalid*.

Do not attempt to run the installation from within *Cygwin*. Cygwin is neither required nor supported.

2.2. Choose Java Version and set JAVA_HOME

Oracle JDK versions *1.7* and *1.6* have been tested by the Hadoop developers and are known to work.

Make sure that *JAVA_HOME* is set in your environment and does not contain any spaces. If your default Java installation directory has spaces then you must use the Windows 8.3 Pathname instead e.g. *c:\Progra~1\Java\...* instead of *c:\Program Files\Java\...*.

2.3. Getting Hadoop sources

The current stable release as of August 2014 is 2.5. The source distribution can be retrieved from the ASF download server or using subversion or git.

- From the ASF Hadoop download page or a mirror.
- Subversion URL: *https://svn.apache.org/repos/asf/hadoop/common/branches/branch-2.5*
- Git repository URL: *git://git.apache.org/hadoop-common.git*. After downloading the sources via git, switch to the stable 2.5 using ***git checkout branch-2.5***, or use the

appropriate branch name if you are targeting a newer version.

2.4. Installing Dependencies and Setting up Environment for Building

The BUILDING.txt file in the root of the source tree has detailed information on the list of requirements and how to install them. It also includes information on setting up the environment and a few quirks that are specific to Windows. It is strongly recommended that you read and understand it before proceeding.

2.5. A few words on Native IO support

Hadoop on Linux includes optional Native IO support. However Native IO is mandatory on Windows and without it you will not be able to get your installation working. You must follow all the instructions from BUILDING.txt to ensure that Native IO support is built correctly.

2.6. Build and Copy the Package files

To build a binary distribution run the following command from the root of the source tree.

```
mvn package -Pdist,native-win -DskipTests -Dtar
```

Note that this command must be run from a *Windows SDK command prompt* as documented in BUILDING.txt. A successful build generates a binary hadoop *.tar.gz* package in *hadoop-dist\target*.

The Hadoop version is present in the package file name. If you are targeting a different version then the package name will be different.

2.7. Installation

Pick a target directory for installing the package. We use *c:\deploy* as an example. Extract the tar.gz file (e.g.*hadoop-2.5.0.tar.gz*) under*c:\deploy*. This will yield a directory structure like the following. If installing a multi-node cluster, then repeat this step on every node.

```
C:\deploy>dir
 Volume in drive C has no label.
 Volume Serial Number is 9D1F-7BAC

 Directory of C:\deploy

01/18/2014  08:11 AM    <DIR>          .
01/18/2014  08:11 AM    <DIR>          ..
01/18/2014  08:28 AM    <DIR>          bin
01/18/2014  08:28 AM    <DIR>          etc
01/18/2014  08:28 AM    <DIR>          include
01/18/2014  08:28 AM    <DIR>          libexec
01/18/2014  08:28 AM    <DIR>          sbin
01/18/2014  08:28 AM    <DIR>          share
```

```
               0 File(s)              0 bytes
```