

Hadoop 1.2.1 compilation

Chun-Chen Tu
timtu@umich.edu

Platform

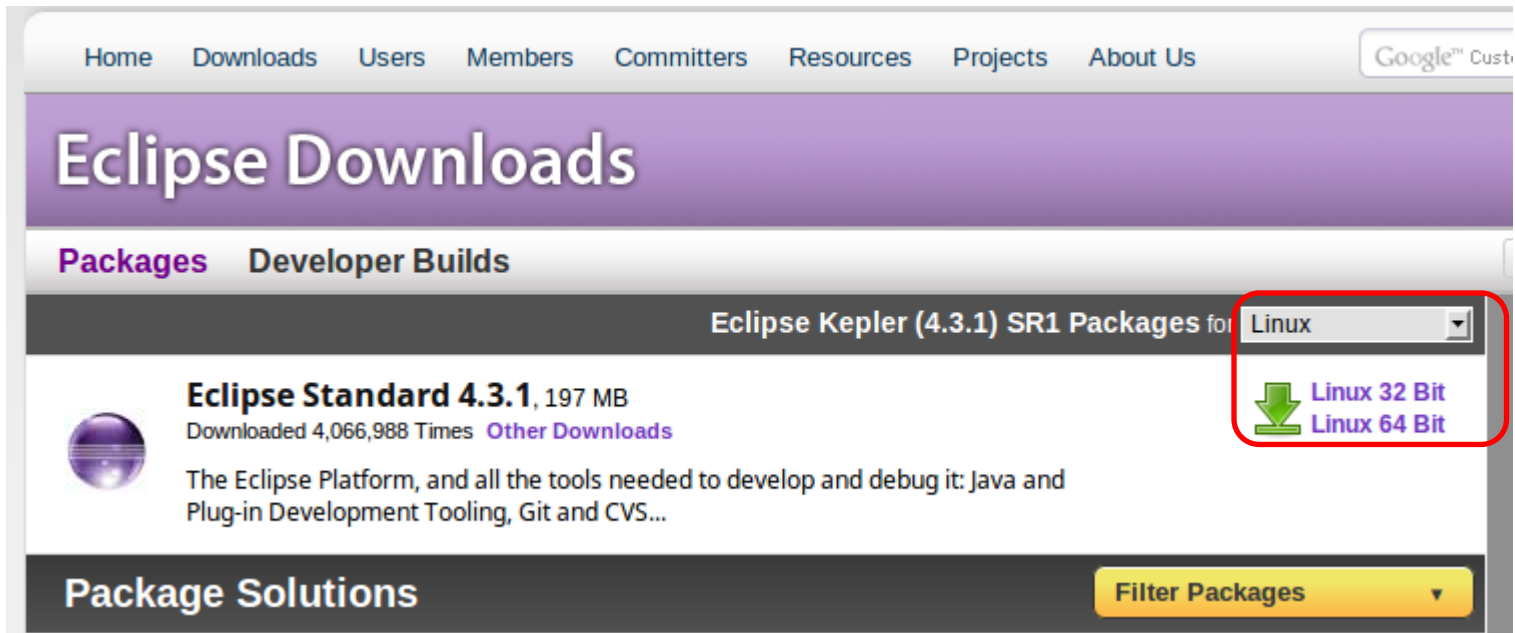
- Virtualbox 4.3.6
 - CPU 1 core 8G RAM
 - Ubuntu 12.04 server 64-bit
- Hadoop 1.2.1

Outline

- Download Eclipse
- Import Hadoop to Eclipse
- Compile changed sources into jar

Download Eclipse

- You should install java jdk first.
- google “eclipse download”
 - <http://www.eclipse.org/downloads/>
 - And download it.



tar Eclipse

In terminal

Find the location you downloaded Eclipse

Usually in ~/Downloads

```
cd ~/Downloads
```

```
tar xvf eclipse*
```

and you'll get a eclipse directory

move it under home

```
mv eclipse ~/
```

Two ways to run Eclipse

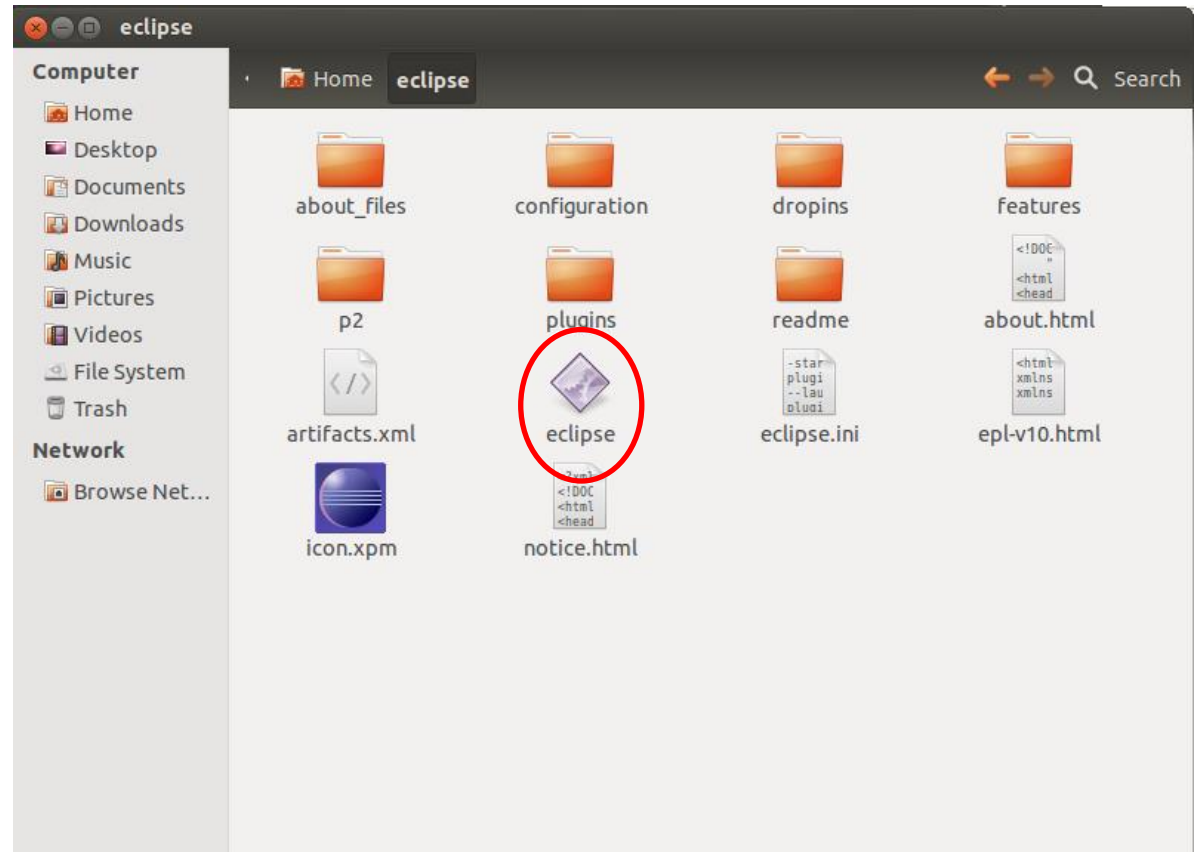
Note: You should install JAVA first

By terminal:

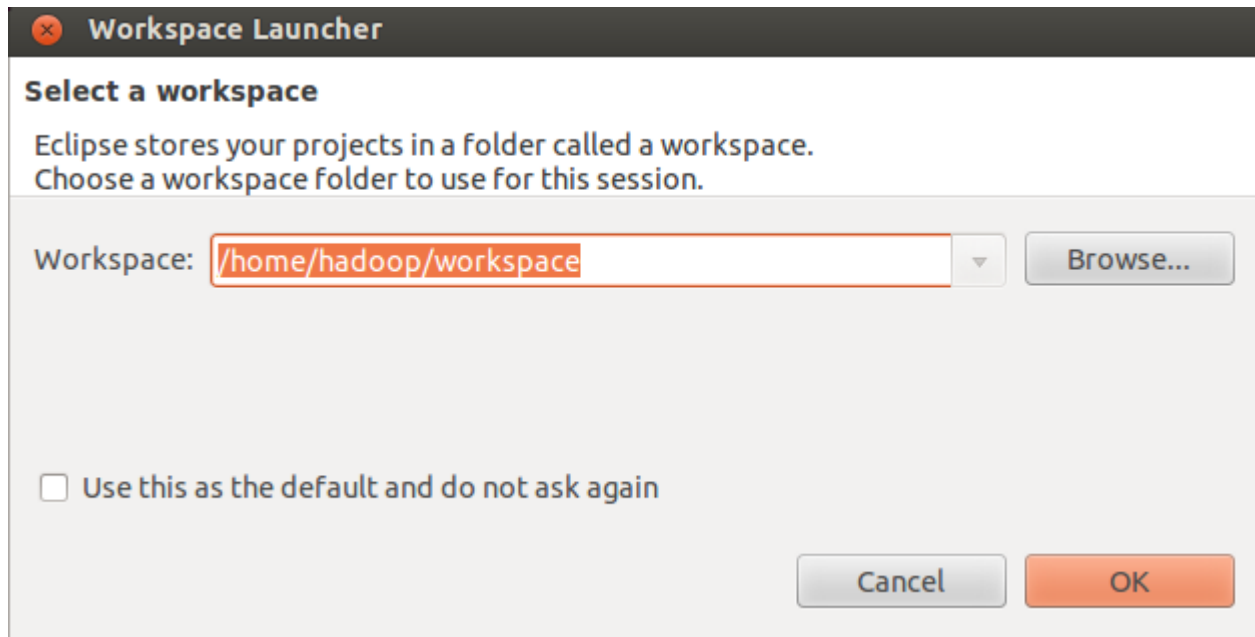
`~/eclipse/eclipse`

By GUI:

double click
the icon



- Selected the workspace of eclipse
- Just use default



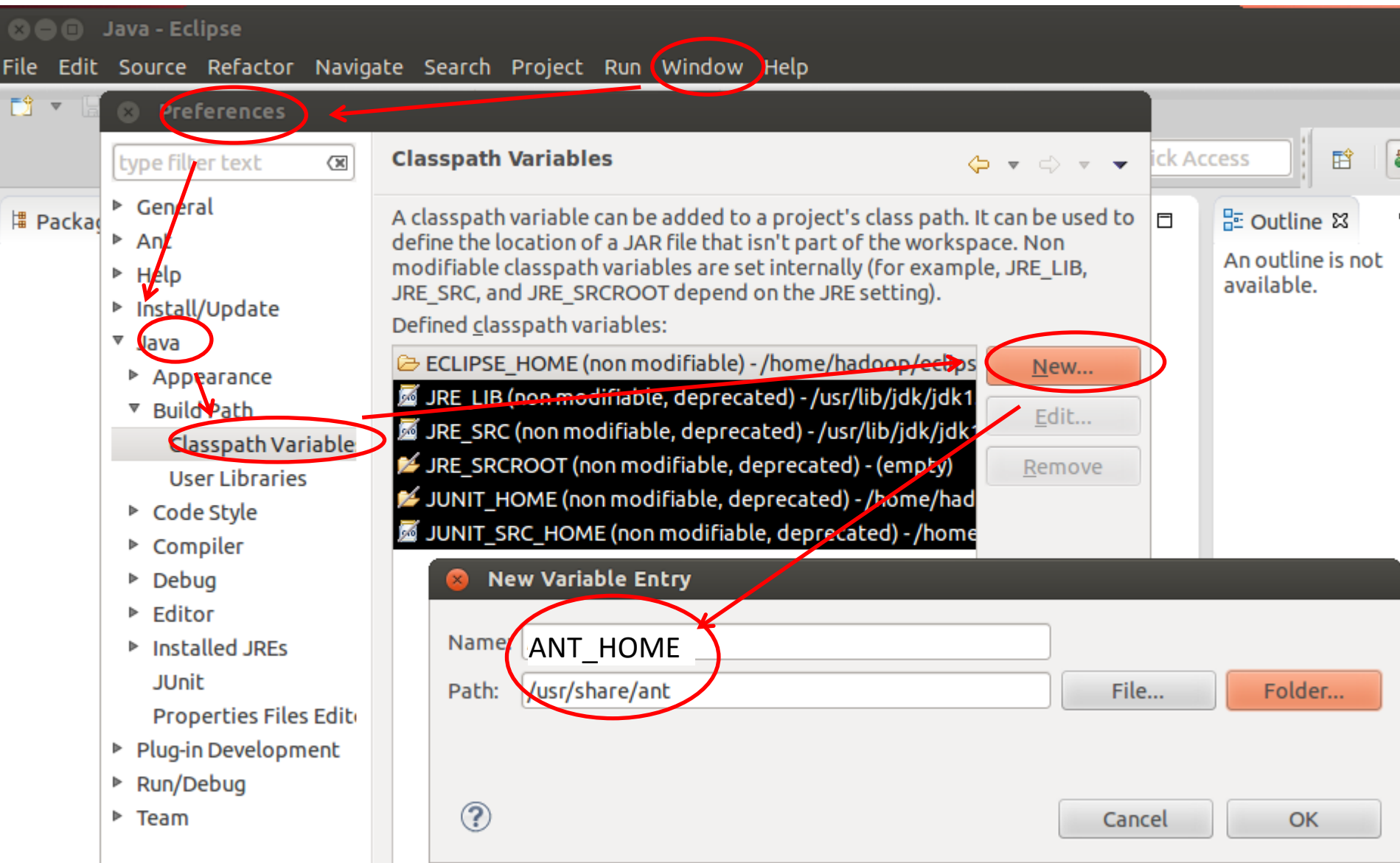
Import Hadoop

- First, get the source
 - <http://ftp.twaren.net/Unix/Web/apache/hadoop/common/hadoop-1.2.1/>
 - ftp://hadoop:hahahadoop@140.113.114.104/hadoop-1.2.1.tar.gz
- tar it and move it under home, named as hadoop

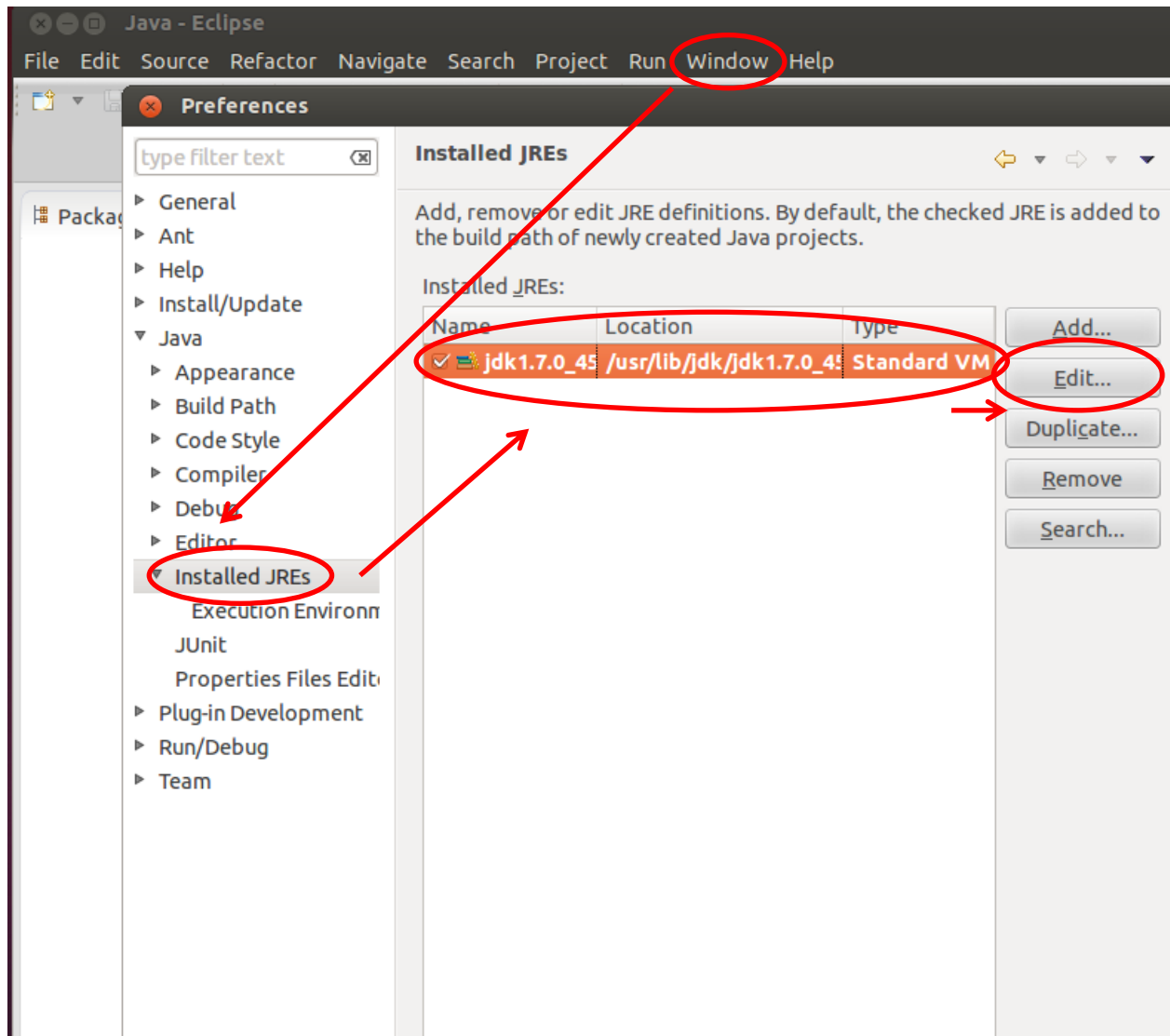
Build hadoop for eclipse

- `sudo apt-get install ant autoconf libtool`
- `sudo aptitude install libtool`
- `cd ~/hadoop`
- `ant`
- `ant eclipse`

Eclipse ant setup



Eclipse include tools.jar



Edit JRE

JRE Definition

Specify attributes for a JRE



JRE home:











Directory...

JRE name:

Default VM arguments:

Variables...

JRE system libraries:

- ▶  /usr/lib/jdk/jdk1.7.0_45/jre/lib/jce.jar
- ▶  /usr/lib/jdk/jdk1.7.0_45/jre/lib/charsets.jar
- ▶  /usr/lib/jdk/jdk1.7.0_45/jre/lib/jfr.jar
- ▶  /usr/lib/jdk/jdk1.7.0_45/jre/lib/ext/localedata.jar
- ▶  /usr/lib/jdk/jdk1.7.0_45/jre/lib/ext/sunpkcs11.jar
- ▶  /usr/lib/jdk/jdk1.7.0_45/jre/lib/ext/dnsns.jar
- ▶  /usr/lib/jdk/jdk1.7.0_45/jre/lib/ext/sunjce_provider.jar
- ▶  /usr/lib/jdk/jdk1.7.0_45/jre/lib/ext/sunec.jar
- ▶  /usr/lib/jdk/jdk1.7.0_45/jre/lib/ext/zipfs.jar
- ▶  /usr/lib/jdk/jdk1.7.0_45/lib/tools.jar

Add External JARs...

Javadoc Location...

Source Attachment...

Remove

Up

Down

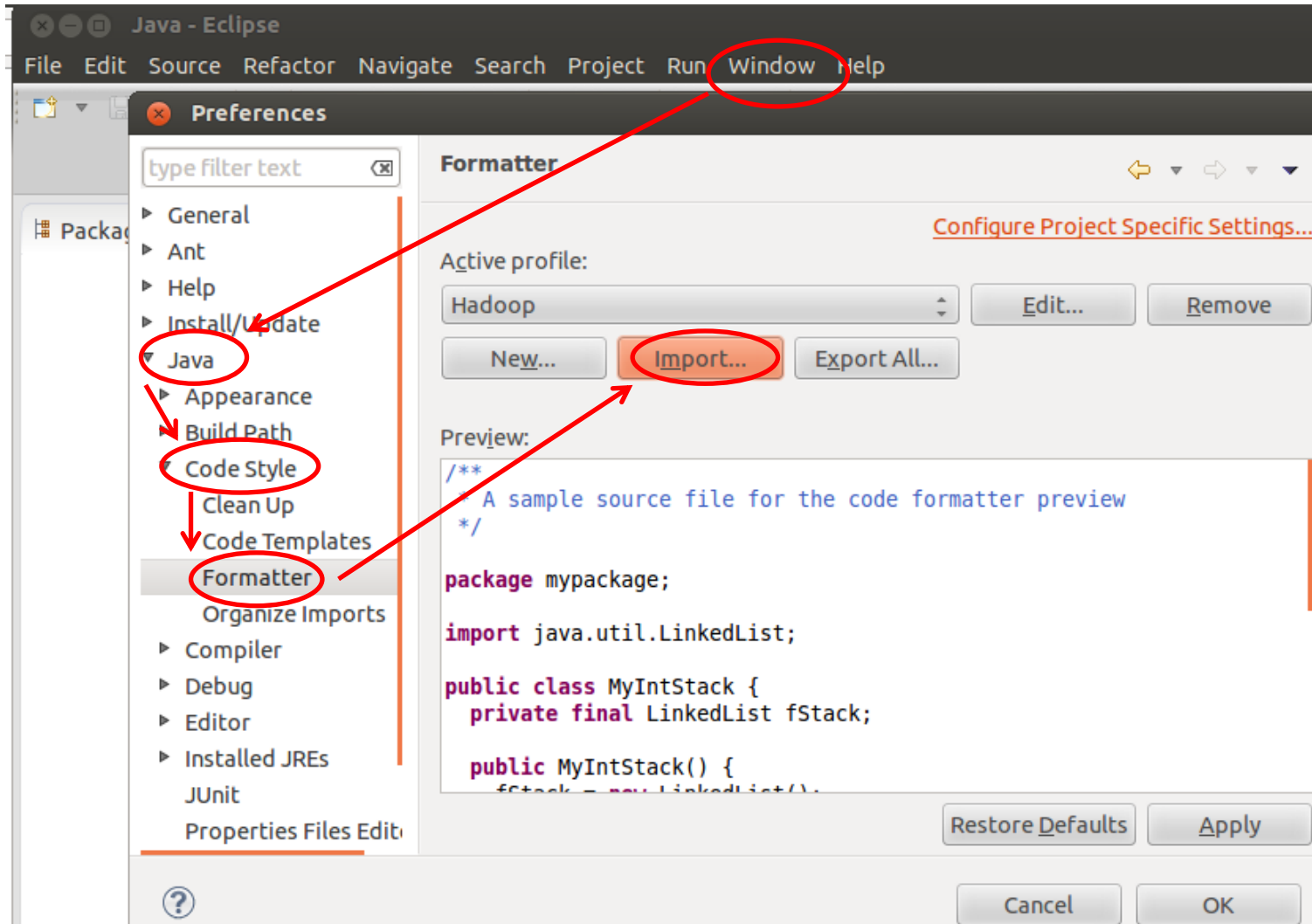
Restore Default



Cancel

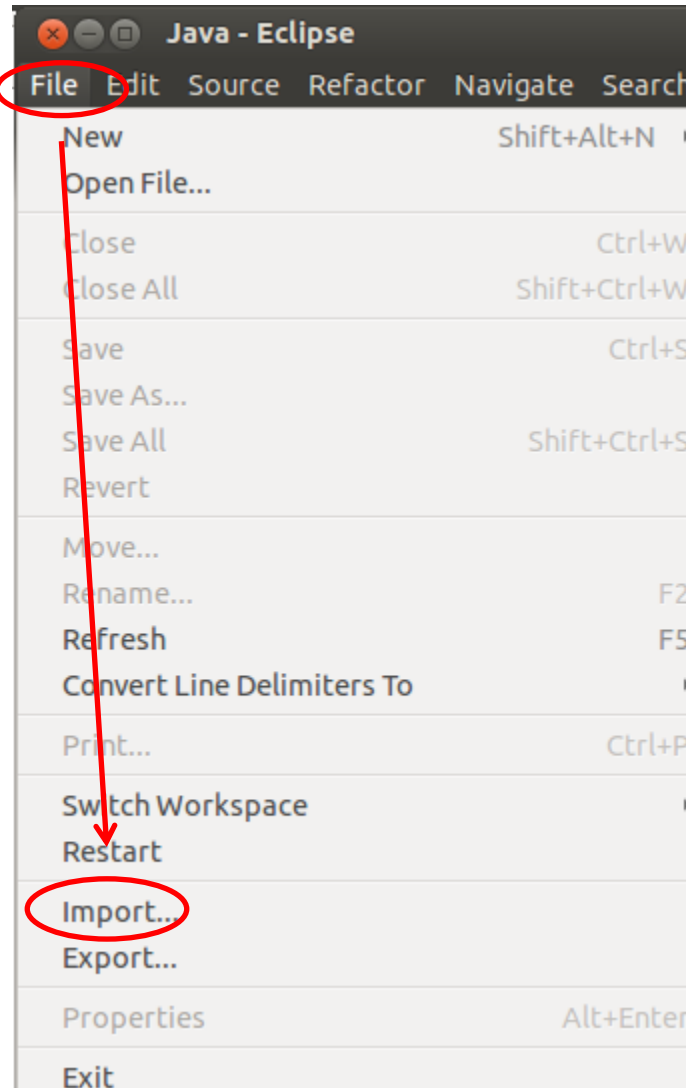
Finish

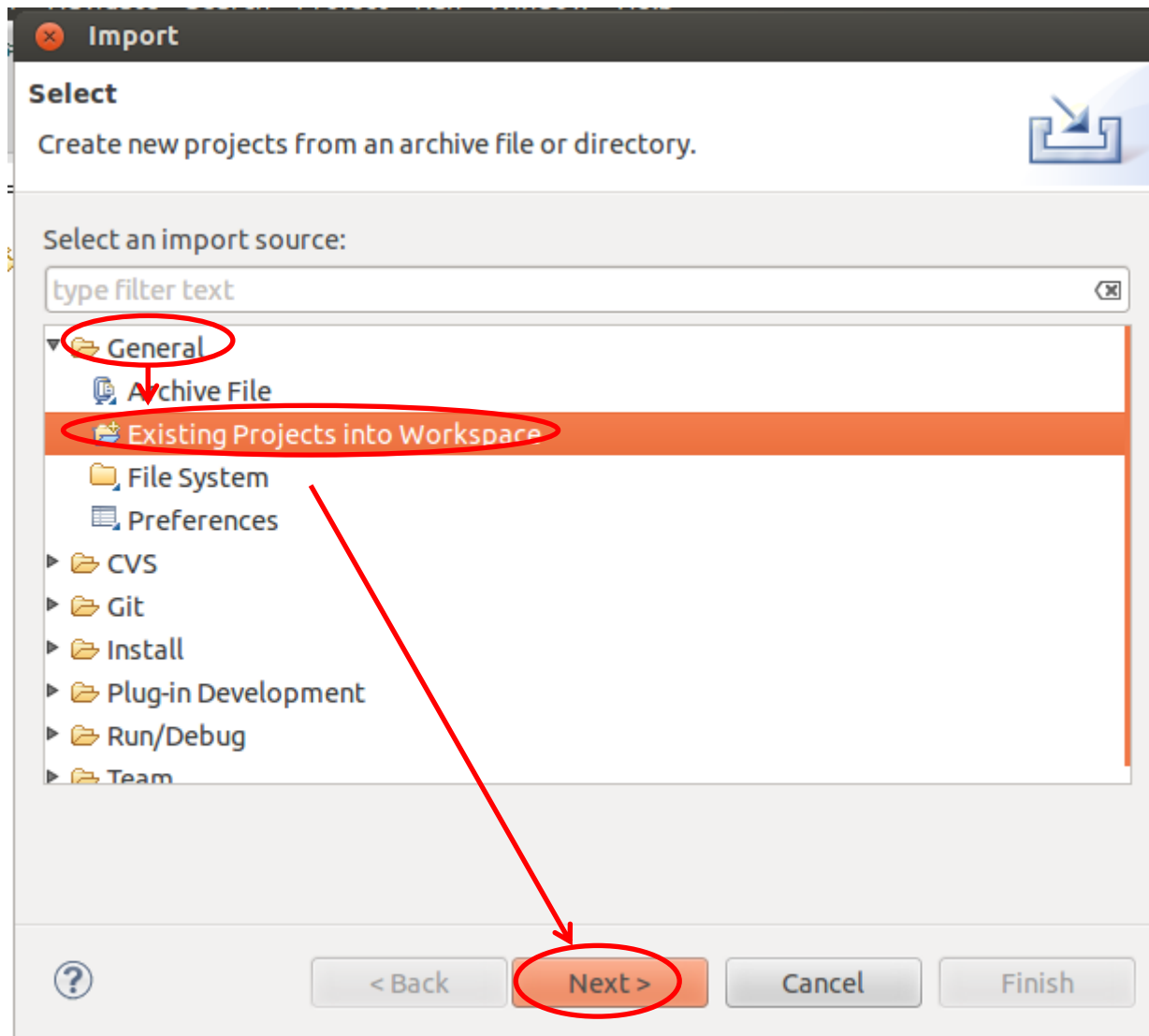
Hadoop code style



Import the formatter from [here](#)

Import hadoop project





Import

Import Projects

Select a directory to search for existing Eclipse projects.

☒ Select root directory:

☐ Select archive file:

Projects:

<input checked="" type="checkbox"/>	hadoop (/home/hadoop/hadoop)
-------------------------------------	------------------------------

Options

☒ Search for nested projects

☒ Copy projects into workspace

Working sets

☐ Add project to working sets

Working sets:

Select the folder after tar which is ~/hadoop for our condition

Compile changed hadoop

- As for example, we want to add message before jar.
- We modified the code
`org.apache.hadoop.util.RunJar.java`

```
RunJar.java
68 }
69 }
70
71 /** Run a Hadoop job jar. If the main class is not in the jar's manifest,
72  * then it must be provided on the command line. */
73 public static void main(String[] args) throws Throwable {
74     String usage = "RunJar jarFile [mainClass] args...";
75
76     System.err.println("Hi Mom, it's me");
77     if (args.length < 1) {
78         System.err.println(usage);
79         System.exit(-1);
80     }
81
82     int firstArg = 0;
83     String fileName = args[firstArg++];
84     File file = new File(fileName);
85     if (!file.exists() || !file.isFile()) {
86         System.err.println("Not a valid JAR: " + file.getCanonicalPath());
87         System.exit(-1);
88     }
89     String mainClassName = null;
90
91     JarFile jarFile;
92     try {
93         jarFile = new JarFile(fileName);
94     } catch (IOException io) {
95         throw new IOException("Error opening job jar: " + fileName)
96             .initCause(io);
97     }
```

Note: It should be System.*err*.println not System.*out*.println

Generate new jar

- By using ant.
- In terminal
- `cd ~/workspace/hadoop`
- `ant jar`
- The new jar appears at `~/workspace/hadoop/build`
- Since `RunJar.java` is a part of `hadoop-core.jar`, we only have to replace the `hadoop-core.jar` with new one.

Hi Mom, it's me

```
hadoop@ubuntu:~/hadoop$ hadoop jar hadoop-examples-1.2.1.jar pi 1 100
Hi Mom, it's me
Number of Maps = 1
Samples per Map = 100
14/02/15 17:05:23 INFO util.NativeCodeLoader: Loaded the native-hadoop library
Wrote input for Map #0
Starting Job
```