

Java 9 Support in Apache Hadoop

May 18, 2017
NTT DATA Corporation
Akira Ajisaka

Self introduction



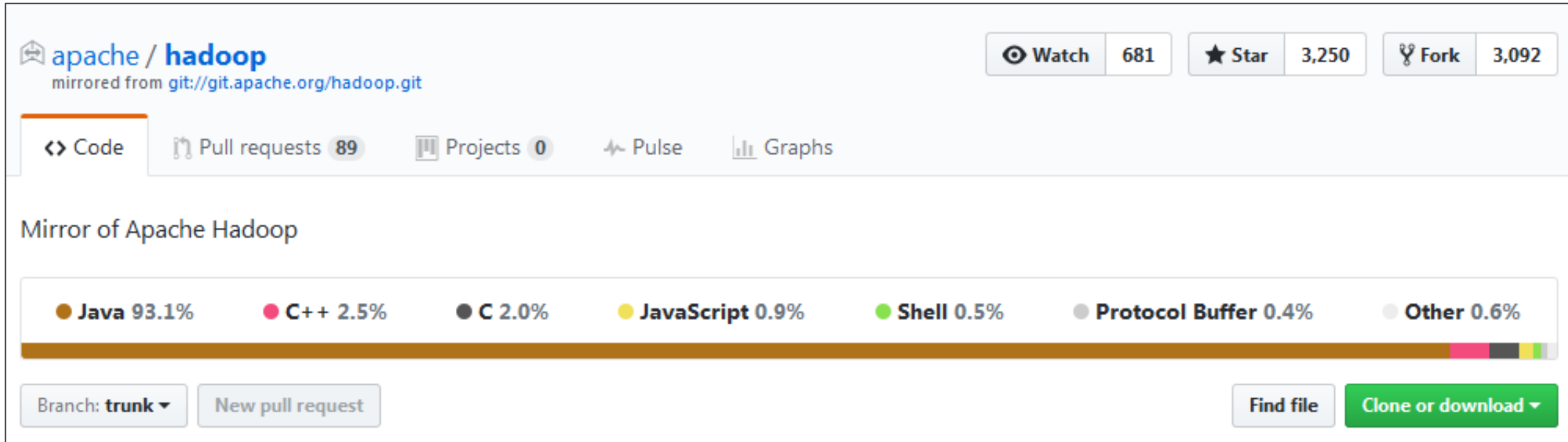
Akira Ajisaka

- Open-Source Software (OSS) Professional Services Team
 - Technical support related to Hadoop/ OSS for our customers
 - Design, integrate, deploy, and operate clusters in the range of 10 - 1200+ servers
- Apache Hadoop Committer & PMC member
 - Fixing test failures
 - Upgrading and managing dependencies
 - Help release process

Agenda

- What is Java 9?
- Why Apache Hadoop does not support Java 9 now?
- Jigsaw
- Classpath isolation in Apache Hadoop

Apache Hadoop and Java



- Hadoop is mainly written in Java (93.1%)
- Hadoop 3.x supports Java 8 only
- Hadoop 2.7+ supports Java 7 and 8

Java 9

- will be released in Jul. 27
- b169 is available as of May 13
- Many new features (e.g. Jigsaw)
- Many incompatible changes

- Meanwhile, Java 8 will be EoL soon
 - Oracle ends public update in Sep 2017
 - Redhat ends public update in Oct 2020
- Need to prepare!

Now Apache Hadoop doesn't work with Java 9

```
$ mvn install -DskipTests
```

```
[INFO] -----  
[INFO] BUILD FAILURE  
[INFO] -----  
[INFO] Total time: 19.719 s  
[INFO] Finished at: 2017-02-20T15:25:07+09:00  
[INFO] Final Memory: 82M/275M  
[INFO] -----  
[ERROR] Failed to execute goal org.apache.maven.plugins:maven-war-plugin:2.4:war  
  (default-war) on project hadoop-auth-examples: Execution default-war of goal or  
g.apache.maven.plugins:maven-war-plugin:2.4:war failed: Unable to load the mojo  
'war' in the plugin 'org.apache.maven.plugins:maven-war-plugin:2.4' due to an AP  
I incompatibility: org.codehaus.plexus.component.repository.exception.ComponentL  
ookupException: null  
[ERROR] -----
```

Cannot compile!

Why compile fails

- 6 problems
 - Encapsulated internal APIs (JEP 260)
 - Banned _ one character identifier (JEP 213)
 - New Version-String Scheme (JEP 223)
 - HTML5 Javadoc (JEP 224)
 - Libraries does not support Java 9
 - Jigsaw
- Work in progress ([HADOOP-11123](#), Umbrella JIRA)

```
/**
 * Unmaps the block from memory. See munmap(2).
 *
 * There isn't any portable way to unmap a memory region in Java.
 * So we use the sun.nio method.
 * Note that unmap if code
 * continues to
 * manually to
 * do it, and
 *
 * @param buffer ... The
 * */
public static void munmap(MappedByteBuffer buffer) {
    if (buffer instanceof sun.nio.ch.DirectBuffer) {
        sun.misc.Cleaner cleaner =
            ((sun.nio.ch.DirectBuffer)buffer).cleaner();
        cleaner.clean();
    }
}
```

NoClassDefFoundError

JEP 260: Encapsulate most internal APIs

- `sun.misc.Cleaner` was moved to `sun.misc.Unsafe::invokeCleaner`
- Usage
 - [Explicit caching mechanism for HDFS](#)
 - Cleaner cleans up off-heap caches
 - equivalent to `munmap(2)`

Support **both** Java 8 and Java 9

- Use Reflection to call methods directly
 - First, call `sun.misc.Unsafe#invokeCleaner`
 - If exception is thrown, then call `sun.misc.Cleaner`
- Apache Lucene hits the same problem
 - Apache Lucene is a pioneer!
 - Fixed in [LUCENE-6989](#)
- Patch available in [HADOOP-12760](#)

HamletSpec.java

```
302 ... *  
303 ... */  
304 public interface _ {}  
305
```

Use of '_' as an identifier might not be supported in releases after Java 8 [more...](#) (⌘F1)

```
307 ... *  
308 ... */  
309 {  
310  
311  
315
```

Banned in Java 9

JEP 213: Milling Project Coin

- Banned `_` one character identifier
- Frequently used by Hamlet, original framework for Hadoop Web UI
 - Inspired from [Hamli](#)

```
.html
... .body()
... .table( selector: "#applications")
... .thead()
... .tr()
... .td()._("ApplicationId")._()
... .td()._("ApplicationState")._()
... .tbody();
```



```
<html>
  <body>
    <table id="applications">
      <thead>
        <tr>
          <td>ApplicationId</td>
          <td>ApplicationState</td>
        </tr>
      </thead>
      <tbody>
```

Be careful with compatibility

- Do not just replace `_` with `__`
 - It affects YARN application
 - e.g. Apache Slider (Incubating)
- How to deal with this problem
 - Create new Hamlet2 package with `__`
 - Deprecate old Hamlet package
 - Replace the usage of `_` with `__`
 - Ignore old Hamlet when compile with Java 9
 - Configure via Maven Compiler Plugin
- Patch available in [HADOOP-11875](#)

```
$ mvn javadoc:javadoc
```

```
[INFO] --- maven-javadoc-plugin:2.8.1:javadoc (default-cli) @ hadoop-maven-plugi
ns ---
[WARNING] Unable to find the javadoc version: Unrecognized version of Javadoc: '
java version "9-ea"
Java(TM) SE Runtime Environment (build 9-ea+154)
Java HotSpot(TM) 64-Bit Server VM (build 9-ea+154, mixed mode)
' near index 37
(?s).*?([0-9]+\.[0-9]+)(\.[0-9]+)?.*
                ^
```

Fail to recognize
Java version

JEP 223: New Version-String Scheme

- Java 8: "1.8.0_xxx"
- Java 9: "9.X.X"

- Affected if regular expression is used to detect Java version
 - Old Maven Javadoc Plugin is affected
 - You must upgrade to 2.10.4+
 - Fixed by [HADOOP-14056](#)

package.html

```
116 <ul><table><tr>
117 <td></td><td></td><td></td><td><tt>3.1415926535.897932
119 ..<td></td><td>=</td><td><tt>3.243F6A8885.A308D31
120
121 11.0010010000.111111
doop.fs
124 ..are.<tt>0010010000</tt>.
125 </p>
126 </td></tr></table>
atadate
127
```

Element table is not allowed here [more...](#) (⌘F1)

Fail with Java 9

JEP 224: HTML5 Javadoc

- Validation for existing html files become strict to support HTML5
- Example
 - `<table>` tag requires summary or caption
 - `'<'` in `<pre>` tag must be rewritten to `<`;
- Fixed by [HADOOP-14057](#)

Update libraries to the version that supports Java 9

- JUnit 3, 4 -> 5
- Mockito 1 -> 2
- Log4J 1 -> 2
- and many more...

\$ mvn install -DskipTests

```
[ERROR] Number of foreign imports: 1
[ERROR] import: Entry[import from realm ClassRealm[project>org.apache.hadoop:hadoop-main:3.0.0-alpha3-SNAPSHOT, parent: ClassRealm[maven.api, parent: null]]
[ERROR]
[ERROR] -----: ExceptionInInitializerError: Unable to make field private final java.util.Comparator java.util.TreeMap.comparator accessible: module java.base does not "opens java.util" to unnamed module @87cb1d8
```



What happened?

The detail

```
[ERROR] Number of foreign imports: 1
[ERROR] import: Entry[import from realm ClassRealm[project>org.apache.hadoop:ha
doo-main:3.0.0-alpha3-SNAPSHOT, parent: ClassRealm[maven.api, parent: null]]
[ERROR]
[ERROR] -----: ExceptionInInitia
lizerError: Unable to make field private final java.util.Comparator java.util.Tr
eeMap.comparator accessible: module java.base does not "opens java.util" to unna
med module @87cb1d8
```

- private fields/methods cannot be accessed from outside
 - Use `Field/Method.setAccessible(true)` to access
 - However, in Java 9, the method can successfully be executed from only the configured 'modules'
- What is 'module'?

Quote from "Java One 2015 keynote"

■ What's this?

```
mbpaa:hadoop$ find . -name "*.jar" | xargs echo
./client/hadoop-client-api-3.0.0-alpha3-SNAPSHOT.jar ./client/hadoop-client-minicluster-3.0.0-alpha3-SNAPSHOT.jar ./client/hadoop-client-runtime-3.0.0-alpha3-SNAPSHOT.jar ./common/hadoop-common-3.0.0-alpha3-SNAPSHOT-tests.jar ./common/hadoop-common-3.0.0-alpha3-SNAPSHOT.jar ./common/hadoop-kms-3.0.0-alpha3-SNAPSHOT.jar ./common/hadoop-nfs-3.0.0-alpha3-SNAPSHOT.jar ./common/lib/avro-1.7.4.jar ./common/lib/commons-beanutils-1.9.3.jar ./common/lib/commons-cli-1.2.jar ./common/lib/commons-codec-1.4.jar ./common/lib/commons-collections-3.2.2.jar ./common/lib/commons-compress-1.4.1.jar ./common/lib/commons-configuration2-2.1.jar ./common/lib/commons-io-2.4.jar ./common/lib/commons-lang-2.6.jar ./common/lib/commons-lang3-3.3.2.jar ./common/lib/commons-logging-1.1.3.jar ./common/lib/commons-math3-3.1.1.jar ./common/lib/commons-net-3.1.jar ./common/lib/curator-client-2.7.1.jar ./common/lib/curator-framework-2.7.1.jar ./common/lib/curator-recipes-2.7.1.jar ./common/lib/gson-2.2.4.jar ./common/lib/guava-11.0.2.jar ./common/lib/hadoop-annotations-3.0.0-alpha3-SNAPSHOT.jar ./common/lib/hadoop-auth-3.0.0-alpha3-SNAPSHOT.jar ./common/lib/hamcrest-core-1.3.jar ./common/lib/htrace-core4-4.1.0-incubating.jar ./common/lib/httpclient-4.5.2.jar ./common/lib/httpcore-4.4.4.jar ./common/lib/jackson-annotations-2.7.8.jar ./common/lib/jackson-core-2.7.8.jar ./common/lib/jackson-core-asl-1.9.13.jar ./common/lib/jackson-databind-2.7.8.jar ./common/lib/jackson-jaxrs-1.9.13.jar ./common/lib/jackson-mapper-asl-1.9.13.jar ./common/lib/jackson-xc-1.9.13.jar ./common/lib/java-xmlbuilder-0.4.jar ./common/lib/javax.servlet-api-3.1.0.jar ./common/lib/jaxb-api-2.2.11.jar ./common/lib/jaxb-impl-2.2.3-1.jar ./common/lib/jcip-annotations-1.0.jar ./common/lib/jersey-core-1.19.jar ./common/lib/jersey-json-1.19.jar ./common/lib/jersey-server-1.19.jar ./common/lib/jersey-servlet-1.19.jar ./common/lib/jets3t-0.9.0.jar ./common/lib/jettison-1.1.jar ./common/lib/jetty-http-9.3.11.v20160721.jar ./common/lib/jetty-io-9.3.11.v20160721.jar ./common/lib/jetty-security-9.3.11.v20160721.jar ./common/lib/jetty-server-9.3.11.v20160721.jar ./common/lib/jetty-servlet-9.3.11.v20160721.jar ./common/lib/jetty-util-9.3.11.v20160721.jar ./common/lib/jetty-webapp-9.3.11.v20160721.jar ./common/lib/jetty-xml-9.3.11.v20160721.jar ./common/lib/jline-0.9.94.jar ./common/lib/jsch-0.1.51.jar ./common/lib/json-smart-1.1.1.jar ./common/lib/jsp-api-2.1.jar ./common/lib/jsr305-3.0.0.jar ./common/lib/jsr311-api-1.1.1.jar ./common/lib/jul-to-slf4j-1.7.10.jar ./common/lib/junit-4.11.jar ./common/lib/kerb-admin-1.0.0-RC2.jar ./common/lib/kerb-client-1.0.0-RC2.jar ./common/lib/kerb-common-1.0.0-RC2.jar ./common/lib/kerb-core-1.0.0-RC2.jar ./common/lib/kerb-crypto-1.0.0-RC2.jar ./common/lib/kerb-identity-1.0.0-RC2.jar ./common/lib/kerb-server-1.0.0-RC2.jar ./common/lib/kerb-simplekdc-1.0.0-RC2.jar ./common/lib/kerb-util-1.0.0-RC2.jar ./common/lib/kerby-asn1-1.0.0-RC2.jar ./common/lib/kerby-config-1.0.0-RC2.jar ./common/lib/kerby-pkix-1.0.0-RC2.jar ./common/lib/kerby-util-1.0.0-RC2.jar ./common/lib/log4j-1.2.17.jar ./common/lib/metrics-core-3.0.1.jar ./common/lib/mockito-all-1.8.5.jar ./common/lib/netty-3.10.5.Final.jar ./common/lib/nimbus-jose-jwt-3.9.jar ./common/lib/paranamer-2.3.jar ./common/lib/protobuf-java-2.5.0.jar ./common/lib/re2j-1.0.jar ./common/lib/slf4j-api-1.7.10.jar ./common/lib/slf4j-log4j12-1.7.10.jar ./common/lib/snappy-java-1.0.4.1.jar ./common/lib/xmlesc-0.52.jar ./common/lib/xz-1.0.jar ./common/lib/zookeeper-3.4.6.jar ./hdfs/hadoop-hdfs-3.0.0-alpha3-SNAPSHOT-tests.jar ./hdfs/hadoop-hdfs-3.0.0-alpha3-SNAPSHOT.jar ./hdfs/hadoop-hdfs-client-3.0.0-alpha3-SNAPSHOT-tests.jar ./hdfs/hadoop-hdfs-native-client-3.0.0-alpha3-SNAPSHOT-tests.jar ./hdfs/hadoop-hdfs-native-client-3.0.0-alpha3-SNAPSHOT.jar ./hdfs/hadoop-hdfs-nfs-3.0.0-alpha3-SNAPSHOT-tests.jar ./hdfs/lib/avro-1.7.4.jar ./hdfs/lib/commons-beanutils-1.9.3.jar ./hdfs/lib/commons-cli-1.2.jar ./hdfs/lib/commons-codec-1.4.jar ./hdfs/lib/commons-collections-3.2.2.jar ./hdfs/lib/commons-compress-1.4.1.jar ./hdfs/lib/commons-configuration2-2.1.jar ./hdfs/lib/commons-daemon-1.0.13.jar ./hdfs/lib/commons-io-2.4.jar ./hdfs/lib/commons-lang-2.6.jar ./hdfs/lib/commons-lang3-3.3.2.jar ./hdfs/lib/commons-logging-1.1.3.jar ./hdfs/lib/commons-math3-3.1.1.jar ./hdfs/lib/commons-net-3.1.jar ./hdfs/lib/curator-client-2.7.1.jar ./hdfs/lib/curator-framework-2.7.1.jar ./hdfs/lib/curator-recipes-2.7.1.jar ./hdfs/lib/gson-2.2.4.jar ./hdfs/lib/guava-11.0.2.jar ./hdfs/lib/hadoop-annotations-3.0.0-alpha3-SNAPSHOT.jar ./hdfs/lib/hadoop-auth-3.0.0-alpha3-SNAPSHOT.jar ./hdfs/lib/hadoop-hdfs-client-3.0.0-alpha3-SNAPSHOT.jar ./hdfs/lib/htrace-core4-4.1.0-incubating.jar ./hdfs/lib/httpclient-4.5.2.jar ./hdfs/lib/httpcore-4.4.4.jar ./hdfs/lib/jackson-annotations-2.7.8.jar ./hdfs/lib/jackson-core-2.7.8.jar ./hdfs/lib/jackson-core-asl-1.9.13.jar ./hdfs/lib/jackson-databind-2.7.8.jar ./hdfs/lib/jackson-jaxrs-1.9.13.jar ./hdfs/lib/jackson-mapper-asl-1.9.13.jar ./hdfs/lib/jackson-xc-1.9.13.jar ./hdfs/lib/javax.servlet-api-3.1.0.jar ./hdfs/lib/jaxb-api-2.2.11.jar ./hdfs/lib/jaxb-impl-2.2.3-1.jar ./hdfs/lib/jcip-annotations-1.0.jar ./hdfs/lib/jersey-core-1.19.jar ./hdfs/lib/jersey-json-1.19.jar ./hdfs/lib/jersey-server-1.19.jar ./hdfs/lib/jersey-servlet-1.19.jar ./hdfs/lib/jettison-1.1.jar ./hdfs/lib/jetty-http-9.3.11.v20160721.jar ./hdfs/lib/jetty-io-9.3.11.v20160721.jar ./hdfs/lib/jetty-security-9.3.11.v20160721.jar ./hdfs/lib/jetty-server-9.3.11.v20160721.jar ./hdfs/lib/jetty-servlet-9.3.11.v20160721.jar ./hdfs/lib/jetty-util-9.3.11.v20160721.jar ./hdfs/lib/jetty-webapp-9.3.11.v20160721.jar ./hdfs/lib/jetty-xml-9.3.11.v20160721.jar ./hdfs/lib/jline-0.9.94.jar ./hdfs/lib/jsch-0.1.51.jar ./hdfs/lib/json-smart-1.1.1.jar ./hdfs/lib/jsp-api-2.1.jar ./hdfs/lib/jsr305-3.0.0.jar ./hdfs/lib/jsr311-api-1.1.1.jar ./hdfs/lib/kerb-admin-1.0.0-RC2.jar ./hdfs/lib/kerb-client-1.0.0-RC2.jar ./hdfs/lib/kerb-common-1.0.0-RC2.jar ./hdfs/lib/kerb-core-1.0.0-RC2.jar ./hdfs/lib/kerb-crypto-1.0.0-RC2.jar ./hdfs/lib/kerb-identity-1.0.0-RC2.jar ./hdfs/lib/kerb-server-1.0.0-RC2.jar ./hdfs/lib/kerb-simplekdc-1.0.0-RC2.jar ./hdfs/lib/kerb-util-1.0.0-RC2.jar ./hdfs/lib/kerby-asn1-1.0.0-RC2.jar ./hdfs/lib/kerby-config-1.0.0-RC2.jar ./hdfs/lib/kerby-pkix-1.0.0-RC2.jar ./hdfs/lib/kerby-util-1.0.0-RC2.jar ./hdfs/lib/leveldbjni-all-1.8.jar ./hdfs/lib/log4j-1.2.17.jar ./hdfs/lib/netty-3.10.5.Final.jar ./hdfs/lib/netty-all-4.1.0.Beta5.jar ./hdfs/lib/nimbus-jose-jwt-3.9.jar ./hdfs/lib/okhttp-2.4.0.jar ./hdfs/lib/okio-1.4.0.jar ./hdfs/lib/paranamer-2.3.jar ./hdfs/lib/protobuf-java-2.5.0.jar ./hdfs/lib/re2j-1.0.jar ./hdfs/lib/snappy-java-1.0.4.1.jar ./hdfs/lib/xercesImpl-2.9.1.jar ./hdfs/lib/xml-apis-1.3.04.jar ./hdfs/lib/xmlesc-0.52.jar ./hdfs/lib/xz-1.0.jar ./mapreduce/hadoop-mapreduce-client-core-3.0.0-alpha3-SNAPSHOT.jar ./mapreduce/hadoop-mapreduce-client-hs-3.0.0-alpha3-SNAPSHOT.jar ./mapreduce/hadoop-mapreduce-client-hs-plugins-3.0.0-alpha3-SNAPSHOT.jar ./mapreduce/hadoop-mapreduce-client-jobclient-3.0.0-alpha3-SNAPSHOT-tests.jar ./mapreduce/hadoop-mapreduce-client-jobclient-3.0.0-alpha3-SNAPSHOT.jar ./mapreduce/hadoop-mapreduce-client-nativetask-3.0.0-alpha3-SNAPSHOT.jar ./mapreduce/hadoop-mapreduce-client-shuffle-3.0.0-alpha3-SNAPSHOT.jar ./mapreduce/hadoop-mapreduce-client-task-3.0.0-alpha3-SNAPSHOT.jar ./mapreduce/hadoop-mapreduce-examples-3.0.0-alpha3-SNAPSHOT.jar ./mapreduce/lib-examples/hsqldb-2.0.0.jar ./tools/lib/alliun-sdk-oss-2.4.1.jar ./tools/lib/aws-java-sdk-bundle-1.11.86.jar ./tools/lib/azure-data-lake-store-sdk-2.1.4.jar ./tools/lib/azure-keyvault-core-0.8.0.jar ./tools/lib/azure-storage-4.2.0.jar ./tools/lib/ezmorph-1.0.6.jar ./tools/lib/hadoop-aliyun-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-archive-logs-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-archives-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-azure-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-azure-data-lake-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-datajoin-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-distcp-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-extras-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-gridmix-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-openstack-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-rumen-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-sls-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/hadoop-streaming-3.0.0-alpha3-SNAPSHOT.jar ./tools/lib/jackson-dataformat-cbor-2.7.8.jar ./tools/lib/jdom-1.1.jar ./tools/lib/joda-time-2.9.4.jar ./tools/lib/json-lib-2.4-jdk15.jar ./tools/lib/json-path-2.2.0.jar ./yarn/hadoop-yarn-api-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-applications-distributedshell-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-applications-unmanaged-am-launcher-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-client-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-common-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-registry-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-applicationhistoryservice-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-common-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-nodemanager-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-resourcemanager-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-sharedcachemanager-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-tests-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-timeline-pluginstorage-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-timelineservice-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-timelineservice-hbase-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-timelineservice-hbase-tests-3.0.0-alpha3-SNAPSHOT.jar ./yarn/hadoop-yarn-server-web-proxy-3.0.0-alpha3-SNAPSHOT.jar ./yarn/lib/aopalliance-1.0.jar ./yarn/lib/commons-csv-1.0.jar ./yarn/lib/commons-el-1.0.jar ./yarn/lib/commons-httpclient-3.1.jar ./yarn/lib/commons-math-2.2.jar ./yarn/lib/curator-test-2.7.1.jar ./yarn/lib/disruptor-3.3.0.jar ./yarn/lib/findbugs-annotations-1.3.9-1.jar ./yarn/lib/fst-2.24.jar ./yarn/lib/guice-4.0.jar ./yarn/lib/hbase-common-1.2.4.jar ./yarn/lib/hbase-client-1.2.4.jar ./yarn/lib/hbase-common-1.2.4.jar ./yarn/lib/hbase-hadoop-compat-1.2.4.jar ./yarn/lib/hbase-hadoop2-compat-1.2.4.jar ./yarn/lib/hbase-prefix-tree-1.2.4.jar ./yarn/lib/hbase-procedure-1.2.4.jar ./yarn/lib/hbase-protocol-1.2.4.jar ./yarn/lib/hbase-server-1.2.4.jar ./yarn/lib/htrace-core-3.1.0-incubating.jar ./yarn/lib/jackson-jaxrs-base-2.7.8.jar ./yarn/lib/jackson-jaxrs-json-provider-2.7.8.jar ./yarn/lib/jackson-module-jaxb-annotations-2.7.8.jar ./yarn/lib/jamon-runtime-2.4.1.jar ./yarn/lib/jasper-compiler-5.5.23.jar ./yarn/lib/jasper-runtime-5.5.23.jar ./yarn/lib/javassist-3.18.1-GA.jar ./yarn/lib/javax.inject-1.jar ./yarn/lib/jcodings-1.0.8.jar ./yarn/lib/jersey-client-1.19.jar ./yarn/lib/jersey-guice-1.19.jar ./yarn/lib/joni-2.1.2.jar ./yarn/lib/jsp-2.1-6.1.14.jar ./yarn/lib/jsp-api-2.1-6.1.14.jar ./yarn/lib/metrics-core-2.2.0.jar ./yarn/lib/metrics-core-3.0.1.jar ./yarn/lib/objenesis-2.1.jar ./yarn/lib/servlet-api-2.5-6.1.14.jar ./yarn/lib/zookeeper-3.4.6-tests.jar
```

Long classpath -> JAR hell

- Answer: Hadoop classpath
- The very long classpath often cause version conflicts of the libraries (JAR hell) between Hadoop and its applications
- Example:
 - Hadoop uses Guava ~~21.0~~
 - Hadoop uses HBase as the backend for YARN Timeline Service v2
 - HBase uses Guava 11.0.2

Now 11.0.2 and the version is now configurable.
(HADOOP-14380, HADOOP-14386)

Jigsaw introduces 'module'

- Create module-info.java for each module to define the dependency between modules

```
$ cat src/com.greetings/module-info.java  
module com.greetings {  
    requires com.astro;  
}
```

Require external
com.astro package

```
$ cat src/com.astro/module-info.java  
module com.astro {  
    exports com.astro;  
}
```

Expose com.astro package
for other module

'module' can enforce the visibility

- Hadoop is using @InterfaceAudience annotation to specify the visibility
 - @Private is internal use within the project, but it is public
 - Public is TOO public :(
- 'module' can enforce the visibility :)

```
$ cat src/com.astro/module-info.java
module com.astro {
    exports com.astro to
        com.greetings;
}
```

Expose com.astro package to **only** com.greetings module

JAR hell will be fixed by Jigsaw?

- If Apache Hadoop supports Jigsaw...
 - Only the public API of Apache Hadoop is exposed
 - Public API of the dependencies is not exposed
- Therefore, JAR hell will be fixed!
- However, there are a lot of work to do

TODO list for Jigsaw support

- Fix incompatibility introduced by Jigsaw (MWAR-405, etc.)
 - There is '--permit-illegal-access' option for workaround
- Create module-info.java for each module ([HADOOP-14269](#))
 - jdeps command can help
- Confirm Hadoop can successfully compiled with both Java 8 and 9
 - Java 8 cannot compile module-info.java, so configure maven-compiler-plugin to ignore

Jigsaw has not approved by Java Community Process yet

These are the final results of the Public Review Ballot for JSR #376. The EC has not approved this ballot.

Votes

EC

Azul Systems, Inc.		Credit Suisse		Eclipse Foundation, Inc		Fujitsu Limited	
Gemalto M2M GmbH		YES: 10 NO: 13					
Hewlett Packard Enterprise							
London Java Community							
Red Hat		SAP SE		Software AG		SouJava	
Tomitribe		Twitter, Inc.		V2COM			

Icon Legend

- Yes
- No
- Abstain
- Not voted

<https://jcp.org/en/jsr/results?id=5959>

- Jigsaw feature will be updated, so my slide can become incorrect afterwards

We can't wait for Java 9 Jigsaw!

- To fully support Jigsaw, update all the dependencies that supports Jigsaw
 - If there are no module-info.java, the module is 'unnamed module'
 - Probably, it takes a very long time to remove 'unnamed module' from classpath
- Jigsaw is not only the solution for JAR hell
- Classpath isolation is in progress ([HADOOP-11656](#))
 - Shading Hadoop client artifacts ([HADOOP-11804](#))
 - Classloader improvement ([HADOOP-13070](#))

Shading Hadoop client artifacts (HADOOP-11804)

- Introduce 2 new modules to avoid leaking Hadoop's dependencies onto the applications' classpath
- hadoop-client-api module
 - removed all the transitive dependencies from hadoop-client module
 - only `org.apache.hadoop.*` are included
- hadoop-client-runtime module
 - add 3rd party dependencies to hadoop-client-api
 - replace the dependency under `org.apache.hadoop.shaded.` by maven-shade-plugin
- Available in Apache Hadoop 3.0.0-alpha2

Users can use the different versions of Hadoop's dependency

- Set hadoop-client-runtime with runtime scope
- Set the dependency and its version as you like

```
<dependency>  
  <groupId>com.fasterxml.jackson.core</groupId>  
  <artifactId>jackson-databind</artifactId>  
  <version>2.8.7</version>  
</dependency>
```

- You can use the two different version

```
import com.fasterxml.jackson.databind.ObjectMapper;  
  
public class ██████████ {  
  
  .. private static final ObjectMapper newMapper = new ObjectMapper();  
  .. private static final  
  .. .. org.apache.hadoop.shaded.com.fasterxml.jackson.databind.ObjectMapper oldMapper =  
  .. .. new org.apache.hadoop.shaded.com.fasterxml.jackson.databind.ObjectMapper();  
}
```

ClassLoader improvement (HADOOP-13070)

- Now user class can load a class from Hadoop's dependencies with or without ApplicationClassLoader
- That way dependency conflicts can occur
- In this issue, we modify ApplicationClassLoader to prevent a user class from loading a class from the parent classpath
 - Check the caller when loading a class
 - If the caller is an user class, prevent loading a class from the parent classpath
- Patch available in [HADOOP-13398](#)

Conclusion

- Now Apache Hadoop does not support Java 9
 - The big work is in progress
 - Your contribution is welcome
- 'JAR hell' problem will be gradually resolved
 - Let's try the new `hadoop-client-api/hadoop-client-runtime` modules and the new class loader!



NTT DATA

Global IT Innovator

The product names and logos used in this presentation are for identification purposes only.
All trademarks and registered trademarks are the property of their respective owners.