

Axiom Developer Guide

Axiom Developer Guide

1.2.14

Licensed to the Apache Software Foundation (ASF) under one or more contributor license agreements. See the NOTICE file distributed with this work for additional information regarding copyright ownership. The ASF licenses this file to you under the Apache License, Version 2.0 (the "License"); you may not use this file except in compliance with the License. You may obtain a copy of the License at

<http://www.apache.org/licenses/LICENSE-2.0>

Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

Table of Contents

1. Working with the Axiom source code	1
Importing the Axiom source code into Eclipse	1
Testing	1
Unit test organization	1
Testing Axiom with different StAX implementations	1
2. OSGi integration	3
Requirements	3
Analysis of the Geronimo JAXB bundles	5
New abstract APIs	6
3. LifecycleManager design (Axiom 1.3)	8
Issues with the LifecycleManager API in Axiom 1.2.x	8
Cleanup strategy for temporary files	9
4. Release process	11
Release preparation	11
Prerequisites	13
Release	14
Post-release actions	16
References	16
5. The StAX specification	17
Semantics of the setPrefix method	17
The three XMLStreamWriter usage patterns	18
A. Appendix	20
Installing IBM's JDK on Debian Linux	20

List of Figures

4.1. Package dependencies for r944680	11
4.2. Package dependencies for r939984	12

Chapter 1. Working with the Axiom source code

Importing the Axiom source code into Eclipse

In order to import the Axiom source code into Eclipse with the Maven Eclipse plugin, use the following command:

```
mvn -DskipTests=true -DdownloadSources=true install eclipse:eclipse
```

Testing

Unit test organization

Historically, all unit tests were placed in the `axiom-tests` project. One specific problem with this is that since all tests are in a common Maven module which depends on both `axiom-impl` and `axiom-dom`, it is not rare to see DOOM tests that accidentally use the LLOM implementation (which is the default). The project description in `axiom-tests/pom.xml` indicates that it was the intention to split the `axiom-tests` project into several parts and make them part of `axiom-api`, `axiom-impl` and `axiom-dom`. This reorganization is not complete yet¹. For new test cases (or when refactoring existing tests), the following guidelines should be applied:

1. Tests that validate the code in `axiom-api` and that do not require an Axiom implementation to execute should be placed in `axiom-api`. This primarily applies to tests that validate utility classes in `axiom-api`.
2. The code of unit tests that apply to all Axiom implementations and that check conformance to the specifications of the Axiom API should be added to `axiom-api` and executed in `axiom-impl` and `axiom-dom`. Currently, the recommended way is to create a base class in `axiom-api` (with suffix `TestBase`) and to create subclasses in `axiom-impl` and `axiom-dom`. This makes sure that the DOOM tests never accidentally use LLOM (because `axiom-impl` is not a dependency of `axiom-dom`).
3. Tests that check integration with other libraries should be placed in `axiom-integration`. Note that this is the only module that requires Java 1.5 (so that e.g. integration with JAXB2 can be tested).
4. Tests related to code in `axiom-api` and requiring an Axiom implementation to execute, but that don't fall into category 2 should stay in `axiom-tests`.

Testing Axiom with different StAX implementations

The following StAX implementations are available to test compatibility with Axiom:

Woodstox

This is the StAX implementation that Axiom uses by default.

Sun Java Streaming XML Parser (SJSXP)

This implementation is available as Maven artifact `com.sun.xml.stream:sjsxp:1.0.1`.

¹See AXIOM-311 [<https://issues.apache.org/jira/browse/AXIOM-311>].

StAX Reference Implementation

The reference implementation was written by BEA and is available as Maven artifact `stax:stax:1.2.0`. The homepage is <http://stax.codehaus.org/Home>. Note that the JAR doesn't contain the necessary files to enable service discovery. Geronimo's implementation of the StAX API library will not be able to locate the reference implementation unless the following system properties are set:

```
javax.xml.stream.XMLInputFactory=com.bea.xml.stream.MXParserFactory
javax.xml.stream.XMLOutputFactory=com.bea.xml.stream.XMLOutputFactoryBase
```

XL XP-J

“XL XML Processor for Java” is IBM's implementation of StAX 1.0 and is part of IBM's JRE/JDK v6. Note that due to an agreement between IBM and Sun, IBM's Java implementation for the Windows platform is not available as a separate download, but only bundled with another IBM product, e.g. WebSphere Application Server for Developers [<http://www.ibm.com/developerworks/downloads/ws/wasdevelopers/>].

On the other hand, the JDK for Linux can be downloaded as a separate package from the developerWorks site. There are versions for 32-bit x86 (“xSeries”) and 64-bit AMD. They are available as RPMs and tarballs. To install the JDK properly on a Debian based system (including Ubuntu), follow the instructions given in the section called “Installing IBM's JDK on Debian Linux”.

Chapter 2. OSGi integration

Requirements

Requirement 1. The Axiom artifacts SHOULD be usable both as normal JAR files and as OSGi bundles.



The alternative would be to produce two sets of artifacts during the build. This should be avoided in order to keep the build process as simple as possible. It should also be noted that the Geronimo Spec artifacts also meet this requirement.

Requirement 2. All APIs defined by the `axiom-api` module, and in particular the `OMAbstractFactory` API MUST continue to work as expected in an OSGi environment, so that code in downstream projects doesn't need to be rewritten.

Requirement 3. `OMAbstractFactory` MUST select the same implementation regardless of the type of container (OSGi or non OSGi). The only exception is related to the usage of system properties to specify the default `OMMetaFactory` implementation: in an OSGi environment, selecting an implementation class using a system property is not meaningful.



This is currently not the case. In a non OSGi environment, recent versions of Axiom use JDK 1.3 service discovery to locate the default implementation and fall back to LLOM if none is found. DOOM will never be selected as the default implementation. On the other hand, the current OSGi integration will select any Axiom implementation as default implementation, but gives priority to LLOM.

Requirement 4. The bundles for the LLOM and DOOM implementations MUST NOT export any packages. This is required to keep a clean separation between the public API and implementation specific classes and to make sure that the implementations can be modified without the risk of breaking existing code. An exception MAY be made for factory classes related to foreign APIs, such as the `DocumentBuilderFactory` implementation for an Axiom implementation supporting DOM.



When the Axiom artifacts are used as normal JAR files in a Maven build, this requirement implies that they should be used in `scope runtime`.

Although this requirement is easy to implement for the Axiom project, there are currently a couple of issues in the downstreams project that need to be addressed to make this work:

- As explained in [AXIS2-4902](https://issues.apache.org/jira/browse/AXIS2-4902) [<https://issues.apache.org/jira/browse/AXIS2-4902>], there are many places in `Axis2` that still refer directly to Axiom implementation classes.
- The `axis2-saaj` module is tightly coupled to `axiom-dom`. Making this work will probably require using `maven-shade-plugin` to include (a relocated copy of) the DOOM classes into `axis2-saaj`.
- `Abdera` extends the LLOM implementation. Probably, some `maven-shade-plugin` magic will be required here as well to create `Abdera` OSGi bundles that work properly with the Axiom bundles.

Requirement 5. It MUST be possible to use a non standard (third party) Axiom implementation as a drop-in replacement for the standard LLOM and DOOM implementation, i.e. the `axiom-impl` and `axiom-dom` bundles. It MUST be possible to replace `axiom-impl` (resp. `axiom-dom`) by any Axiom implementation that supports the full Axiom API (resp. that supports DOM in addition to the Axiom API), without the need to change any application code.



This requirement has several important implications:

- It restricts the allowable exceptions to Requirement 4.
- It implies that there must be an API that allows application code to select an Axiom implementation based on its capabilities (e.g. DOM support) without introducing a hard dependency on a particular Axiom implementation.
- In accordance with Requirement 2 and Requirement 3 this requirement not only applies to an OSGi environment, but extends to non OSGi environments as well.

Requirement 6. The OSGi integration SHOULD remove the necessity for downstreams projects to produce their own custom OSGi bundles for Axiom. There SHOULD be one and only one set of OSGi bundles for Axiom, namely the ones released by the Axiom project.



Currently there are at least two projects that create their own modified Axiom bundles:

- Apache Geronimo has a custom Axiom bundle to support the Axis2 integration.
- ServiceMix also has a custom bundles for Axiom. However, this bundle only seem to exist to support their own custom Abdera bundle, which is basically an incorrect repackaging of the original Abdera code. See SMX4-877 [<https://issues.apache.org/jira/browse/SMX4-877>] for more details.

Note that this requirement can't be satisfied directly by Axiom. It requires that the above mentioned projects (Geronimo, Axis2 and Abdera) use Axiom in a way that is compatible with its design, and in particular with Requirement 4. Nevertheless, Axiom must provide the necessary APIs and features to meet the needs of these projects.

Requirement 7. The Axiom OSGi integration SHOULD NOT rely on any particular OSGi framework such as Felix SCR (Declarative Services). When deployed in an OSGi environment, Axiom should have the same runtime dependencies as in a non OSGi environment (i.e. StAX, Activation and JavaMail).



Axiom 1.2.12 relies on Felix SCR. Although there is no real issue with that, getting rid of this extra dependency is seen as a nice to have. One of the reasons for using Felix SCR was to avoid introducing OSGi specific code into Axiom. However, there is no issue with having such code, provided that Requirement 8 is satisfied.

Requirement 8. In a non OSGi environment, Axiom MUST NOT have any OSGi related dependencies. That means that the OSGi integration must be written in such a way that no OSGi specific classes are ever loaded in a non OSGi environment.

Requirement 9. The OSGi integration MUST follow established best practices. It SHOULD be inspired by what has been done to add OSGi integration to APIs that have a similar structure as Axiom.



Axiom is designed around an abstract API and allows for the existence of multiple independent implementations. A factory (`OMAbstractFactory`) is used to locate and instantiate the desired implementation. This is similar to APIs such as JAXP (`DocumentBuilderFactory`, etc.) and JAXB (`JAXBContext`). These APIs have been successfully "OSGi-fied" e.g. by the Apache Geronimo project. Instead of reinventing the wheel, we should leverage that work and adapt it to Axiom's specific requirements.

It should be noted that because of the way the Axiom API is designed and taking into account Requirement 2, it is not possible to make Axiom entirely compatible with OSGi paradigms (the same is true for JAXB). In an OSGi-only world, each Axiom implementation would simply expose itself as an OSGi service (of type `OMMetaFactory` e.g.) and code depending

on Axiom would bind to one (or more) of these services depending on its needs. That is not possible because it would conflict with Requirement 2.

Non-Requirement 1. APIs such as JAXP and JAXB have been designed from the start for inclusion into the JRE. They need to support scenarios where an application bundles its own implementation (e.g. an application may package a version of Apache Xerces, which would then be instantiated by the `newInstance` method in `DocumentBuilderFactory`). That implies that the selected implementation depends on the thread context class loader. It is assumed that there is no such requirement for Axiom, which means that in a non OSGi environment, the Axiom implementations are always loaded from the same class loader as the `axiom-api` JAR.



This (non-)requirement is actually not directly relevant for the OSGi support, but it nevertheless has some importance because of Requirement 3 (which implies that the OSGi support needs to be designed in parallel with the implementation discovery strategy applicable in a non OSGi environment).

Analysis of the Geronimo JAXB bundles

As noted in Requirement 9 the Apache Geronimo has successfully added OSGi support to the JAXB API which has a structure similar to the Axiom API. This section briefly describes how this works. The analysis refers to the following Geronimo artifacts: `org.apache.geronimo.specs:geronimo-jaxb_2.2_spec:1.0.1` (called the "API bundle" hereafter), `org.apache.geronimo.bundles:jaxb-impl:2.2.3-1_1` (the "implementation bundle"), `org.apache.geronimo.specs:geronimo-osgi-locator:1.0` (the "locator bundle") and `org.apache.geronimo.specs:geronimo-osgi-registry:1.0` (the "registry bundle"):

- The implementation bundle retains the `META-INF/services/javax.xml.bind.JAXBContext` resource from the original artifact (`com.sun.xml.bind:jaxb-impl`). In a non OSGi environment, that resource will be used to discover the implementation, following the standard JDK 1.3 service discovery algorithm will (as required by the JAXB specification). This is the equivalent of our Requirement 1.
- The manifest of the implementation bundle has an attribute `SPI-Provider: true` that indicates that it contains provider implementations that are discovered using the JDK 1.3 service discovery.
- The registry bundle creates a `BundleTracker` that looks for the `SPI-Provider` attribute in active bundles. For each bundle that has this attribute set to `true`, it will scan the content of `META-INF/services` and add the discovered services to a registry (Note that the registry bundle supports other ways to declare SPI providers, but this is not really relevant for the present discussion).
- The `ContextFinder` class (the interface of which is defined by the JAXB specification and that is used by the `newInstance` method in `JAXBContext`) in the API bundle delegates the discovery of the SPI implementation to a static method of the `ProviderLocator` class defined by the locator bundle (which is not specific to JAXB and is used by other API bundles as well). This is true both in an OSGi environment and in a non OSGi environment.

The build is configured (using a `Private-Package` instruction) such that the classes of the locator bundle are actually included into the API bundle, thus avoiding an additional dependency.

- The `ProviderLocator` class and related code provided by the locator bundle is designed such that in a non OSGi environment, it will simply use JDK 1.3 service discovery to locate the SPI implementation, without ever loading any OSGi specific class. On the other hand, in an OSGi environment, it will query the registry maintained by the registry bundle to locate the provider. The reference to the registry is injected into the `ProviderLocator` class using a bundle activator.

- Finally, it should also be noted that the API bundle is configured with `singleton=true`. There is indeed no meaningful way how providers could be matched with different versions of the same API bundle.

This is an example of a particularly elegant way to satisfy Requirement 1, Requirement 2 and Requirement 3, especially because it relies on the same metadata (the `META-INF/services/javax.xml.bind.JAXBContext` resources) in OSGi and non OSGi environments.

Obviously, Axiom could reuse the registry and locator bundles developed by Geronimo. This however would contradict Requirement 7. In addition, for Axiom there is no requirement to strictly follow the JDK 1.3 service discovery algorithm. Therefore Axiom should reuse the pattern developed by Geronimo, but not the actual implementation.

New abstract APIs

Application code rarely uses DOOM as the default Axiom implementation. Several downstream projects (e.g. the Axis2/Rampart combination) use both the default (LLOM) implementation and DOOM. They select the implementation based on the particular context. As of Axiom 1.2.12, the only way to create an object model instance with the DOOM implementation is to use the `DOOMAbstractFactory` API or to instantiate one of the factory classes (`OMDOMMetaFactory`, `OMDOMFactory` or one of the subclasses of `DOMSOAPFactory`). All these classes are part of the `axiom-dom` artifact. This is clearly in contradiction with Requirement 4 and Requirement 5.

To overcome this problem the Axiom API must be enhanced to make it possible to select an Axiom implementation based on capabilities/features requested by the application code. E.g. in the case of DOOM, the application code would request a factory that implements the DOM API. It is then up to the Axiom API classes to locate an appropriate implementation, which may be DOOM or another drop-in replacement, as per Requirement 5.

If multiple Axiom implementations are available (on the class path in non OSGi environment or deployed as bundles in an OSGi environment), then the Axiom API must also be able to select an appropriate default implementation if no specific feature is requested by the application code. This can be easily implemented by defining a special feature called "default" that would be declared by any Axiom implementation that is suitable as a default implementation.



DOOM is generally not considered suitable as a default implementation because it doesn't implement the complete Axiom API (e.g. it doesn't support `OMSourcedElement`) and because the factory classes are not stateless.

Finally, to make the selection algorithm deterministic, there should also be a concept of priority: if multiple Axiom implementations are found for the same feature, then the Axiom API would select the one with the highest priority.

This leads to the following design:

1. Every Axiom implementation declares a set of features that it supports. A feature is simply identified by a string. Two features are predefined by the Axiom API:
 - `default`: indicates that the implementation is a complete implementation of the Axiom API and may be used as a default implementation.
 - `dom`: indicates that the implementation supports DOM in addition to the Axiom API.

For every feature it declares, the Axiom implementation specifies a priority, which is a positive integer.

2. The relevant Axiom APIs are enhanced so that they take an optional argument specifying the feature requested by the application code. If no explicit feature is requested, then Axiom will use the `default` feature.
3. To determine the `OMMetaFactory` to be used, Axiom locates the implementations declaring the requested feature and selects the one that has the highest priority for that feature.

A remaining question is how the implementation declares the feature/priority information. There are two options:

- Add a method to `OMMetaFactory` that allows the Axiom API to query the feature/priority information from the implementation (i.e. the features and priorities are hardcoded in the implementation).
- Let the implementation provide this information declaratively in its metadata (either in the manifest or in a separate resource with a well defined name). Note that in a non OSGi environment, such a metadata resource must be used anyway to enable the Axiom API to locate the `OMMetaFactory` implementations. Therefore this would be a natural place to declare the features as well.

The second option has the advantage to make it easier for users to debug and tweak the implementation discovery process (e.g. there may be a need to customize the features and priorities declared by the different implementations to ensure that the right implementation is chosen in a particular use case).

This leads to the following design decision: the features and priorities (together with the class name of the `OMMetaFactory` implementation) will be defined in an XML descriptor with resource name `META-INF/axiom.xml`. The format of that descriptor must take into account that a single JAR may contain several Axiom implementations (e.g. if the JAR is an uber-JAR repackaged from the standard Axiom JARs).

Chapter 3. LifecycleManager design (Axiom 1.3)

The `LifecycleManager` API is used by the MIME handling code in Axiom to manage the temporary files that are used to buffer the content of attachment parts. The `LifecycleManager` implementation is responsible to track the temporary files that have been created and to ensure that they are deleted when they are no longer used. In Axiom 1.2.x, this API has multiple issues and a redesign is required for Axiom 1.3.

Issues with the LifecycleManager API in Axiom 1.2.x

1. Temporary files that are not cleaned up explicitly by application code will only be removed when the JVM stops (`LifecycleManagerImpl` registers a shutdown hook and maintains a list of files that need to be deleted when the JVM exits). This means that temporary files may pile up, causing the file system to fill.
2. `LifecycleManager` also has a method `deleteOnTimeInterval` that deletes a file after some specified time interval. However, the implementation creates a new thread for each invocation of that method, which is generally not acceptable in high performance use cases.
3. One of the stated design goals (see AXIOM-192 [<https://issues.apache.org/jira/browse/AXIOM-192>]) of the `LifecycleManager` API was to wrap the files in `FileAccessor` objects to “keep track of activity that occurs on the files”. However, as pointed out in AXIOM-185 [<https://issues.apache.org/jira/browse/AXIOM-185>], since `FileAccessor` has a method that returns the corresponding `File` object, this goal has not been reached.
4. As noted in AXIOM-382 [<https://issues.apache.org/jira/browse/AXIOM-382>], the fact that `LifecycleManagerImpl` registers a shutdown hook which is never unregistered causes a class loader leak in J2EE environments.
5. In an attempt to work around the issues related to `LifecycleManager` (in particular the first item above), AXIOM-185 [<https://issues.apache.org/jira/browse/AXIOM-185>] introduced another class called `AttachmentCacheMonitor` that implements a timer based mechanism to clean up temporary files. However, this change causes other issues:
 - The existence of this API has a negative impact on Axiom's architectural integrity because it has functionality that overlaps with `LifecycleManager`. This means that we now have two completely separate APIs that are expected to serve the same purpose, but none of them addresses the problem properly.
 - `AttachmentCacheMonitor` automatically creates a timer, but there is no way to stop that timer. This means that this API can only be used if Axiom is integrated into the container, but not when it is deployed with an application.

Fortunately, that change was only meant as a workaround to solve a particular issue in WebSphere (see APAR PK91497 [<http://www-01.ibm.com/support/docview.wss?rs=180&uid=swg1PK91497>]), and once the `LifecycleManager` API is redesigned to solve that issue, `AttachmentCacheMonitor` no longer has a reason to exist.

6. `LifecycleManager` is an abstract API (interface), but refers to `FileAccessor` which is placed in an `impl` package.

7. `FileAccessor` uses the `MessagingException` class from `JavaMail`, although Axiom no longer relies on this API to parse or create MIME messages.

Cleanup strategy for temporary files

As pointed out in the previous section, one of the primary problems with the `LifecycleManager` API in Axiom 1.2.x is that temporary files that are not cleaned up explicitly by application code (e.g. using the `purgeDataSource` method defined by `DataHandlerExt`) are only removed when the JVM exits. A timer based strategy that deletes temporary file after a given time interval (as proposed by `AttachmentCacheMonitor`) is not reliable because in some use cases, application code may keep a reference to the attachment part for a long time before accessing it again.

The only reliable strategy is to take advantage of finalization, i.e. to rely on the garbage collector to trigger the deletion of temporary files that are no longer used. For this to work the design of the API (and its default implementation) must satisfy the following two conditions:

1. All access to the underlying file must be strictly encapsulated, so that the file is only accessible as long as there is a strong reference to the object that encapsulates the file access. This is necessary to ensure that the file can be safely deleted once there is no longer a strong reference and the object is garbage collected.
2. Java guarantees that the finalizer is invoked before the instance is garbage collected. However, instances are not necessarily garbage collected before the JVM exits, and in that case the finalizer is never invoked. Therefore, the implementation must delete all existing temporary files when the JVM exits. The API design should also take into account that some implementations of the `LifecycleManager` API may want to trigger this cleanup before the JVM exits, e.g. when the J2EE application in which Axiom is deployed is stopped.

The first condition can be satisfied by redesigning the `FileAccessor` such that it never leaks the name of the file it represents (neither as a `String` nor a `File` object). This in turn means that the `CachedFileDataSource` class must be removed from the Axiom API. In addition, the `getInputStream` method defined by `FileAccessor` must no longer return a simple `FileInputStream` instance, but must use a wrapper that keeps a strong reference to the `FileAccessor`, so that the `FileAccessor` can't be garbage collected while the input stream is still in use.

To satisfy the second condition, one may want to use `File#deleteOnExit`. However, this method causes a native memory leak, especially when used with temporary files, which are expected to have unique names (see bug 4513817 [http://bugs.sun.com/bugdatabase/view_bug.do?bug_id=4513817]). Therefore this can only be implemented using a shutdown hook. However, a shutdown hook will cause a class loader leak if it is used improperly, e.g. if it is registered by an application deployed into a J2EE container and not unregistered when that application is stopped. For this particular case, it is possible to create a special `LifecycleManager` implementation, but for this to work, the lifecycle of this type of `LifecycleManager` must be bound to the lifecycle of the application, e.g. using a `ServletContextListener`. This is not always possible and this approach is therefore not suitable for the default `LifecycleManager` implementation.

To avoid the class loader leak, the default `LifecycleManager` implementation should register the shutdown hook when the first temporary file is registered and automatically unregister the shutdown hook again when there are no more temporary files. This implies that the shutdown hook is repeatedly registered and unregistered. However, since these are relatively cheap operations¹, this should not be a concern.

¹Since the JRE typically uses an `IdentityHashMap` to store shutdown hooks, the only overhead is caused by Java 2 security checks and synchronization.

An additional complication is that when the shutdown hook is executed, the temporary files may still be in use. This contrasts with the finalizer case where encapsulation guarantees that the file is no longer in use. This situation doesn't cause an issue on Unix platforms (where it is possible to delete a file while it is still open), but needs to be handled properly on Windows. This can only be achieved if the `FileAccessor` keeps track of created streams, so that it can forcibly close the underlying `FileInputStream` objects.

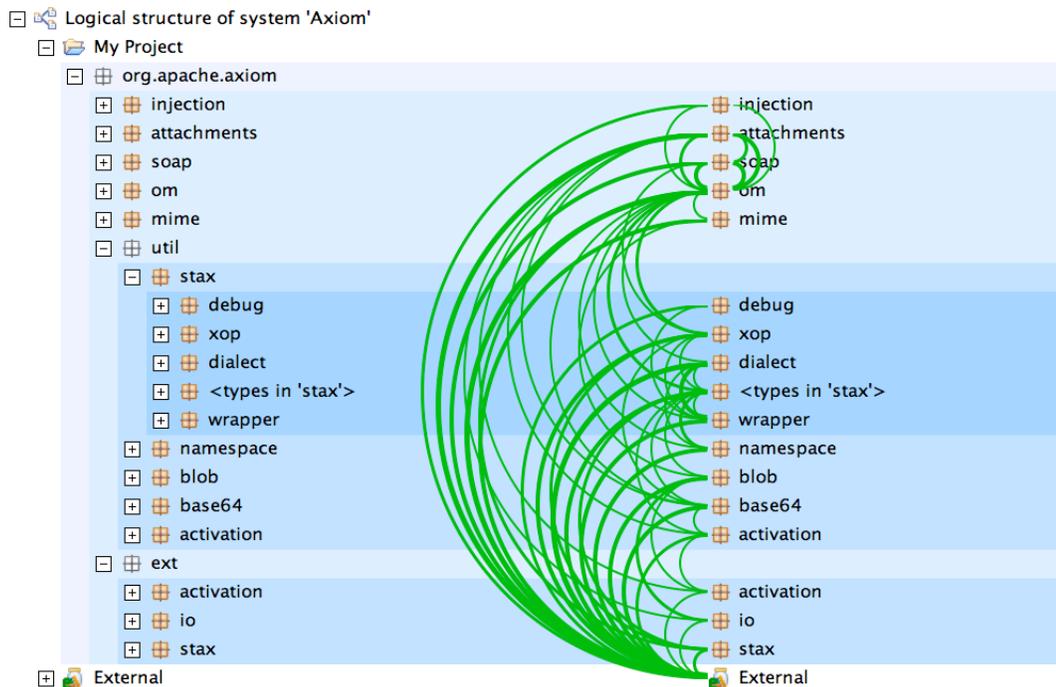
Chapter 4. Release process

Release preparation

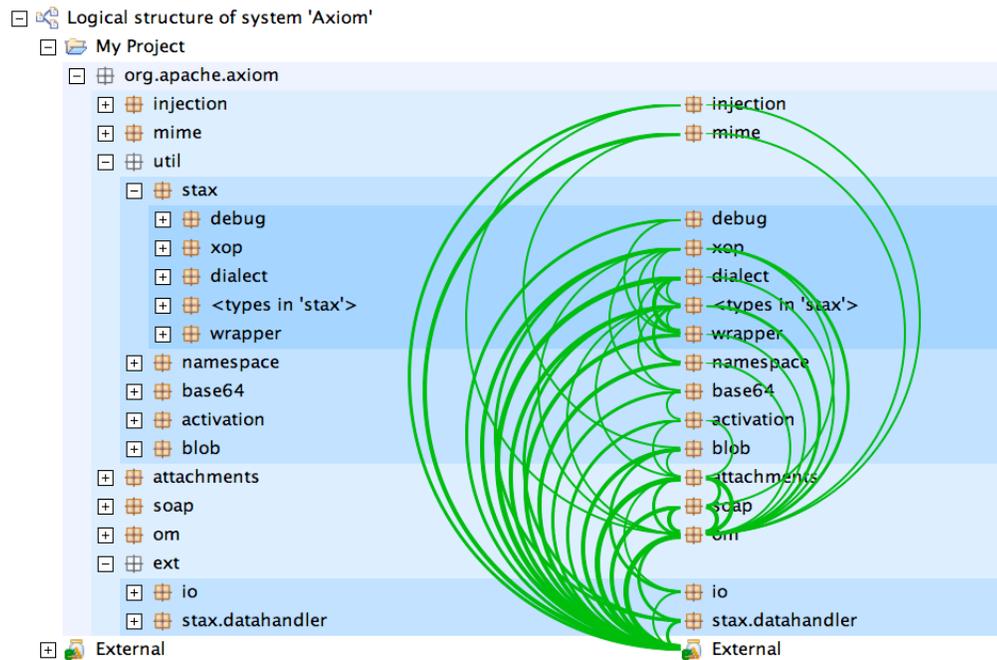
The following items should be checked before starting the release process:

- Check for the latest Apache parent POM version (artifact `org.apache:apache`) and if necessary, change the parent of the Axiom root POM.
- Check the dependencies between Java packages in the `axiom-api` module. The `org.apache.axiom.util` package (including its subpackages) is specified to contain utility classes that don't depend on higher level APIs. More precisely, `org.apache.axiom.util` should only have dependencies on `org.apache.axiom.ext`, but not e.g. on `org.apache.axiom.om`. SonarJ [<http://www.hello2morrow.com/products/sonarj>] can be used to check these dependencies. The following figure shows the expected structure:

Figure 4.1. Package dependencies for r944680



In contrast, the following figure shows an earlier trunk version of `axiom-api` with incorrect layering and cyclic dependencies involving `org.apache.axiom.util`:

Figure 4.2. Package dependencies for r939984

The check can also be done using `jdepend-maven-plugin` [<http://mojo.codehaus.org/jdepend-maven-plugin/>]. To do this, execute the following command in the `axiom-api` module:

```
mvn jdepend:generate
```

Then open `target/site/jdepend-report.html` and go to the "Cycles" section. The report should not show any package cycles involving `org.apache.axiom.mime`, `org.apache.axiom.util` and `org.apache.axiom.ext`.

- Check that the generated Javadoc contains the appropriate set of packages. In particular, unit test related classes should be excluded, except for the test suite classes in `org.apache.axiom.ts` (we don't need to hide the fact that we have a reusable test suite...).
- Check that all dependencies and plugins are available from standard repositories. To do this, clean the local repository and execute **`mvn clean install`** followed by **`mvn site`**.
- Check that the set of license files in the `legal` directory is complete and accurate.
- Check that the Maven site conforms to the latest version of the Apache Project Branding Guidelines [<http://apache.org/foundation/marks/pmcs>].
- Check that the `apache-release` profile can be executed properly. To do this, issue the following command:

```
mvn clean install -Papache-release -Dmaven.test.skip=true
```

You may also execute a dry run of the release process:

```
mvn release:prepare -DdryRun=true
```

After this, you need to clean up using the following command:

```
mvn release:clean
```

- Prepare the release note. This should include a description of the major changes in the release as well as a list of resolved JIRA issues. Note that both `index.appt` and `RELEASE-NOTE.txt` need to be updated.
- Add an entry for the release to the `download.xml.vm` file and change the links for older releases so that they point to `archive.apache.org` (Since the Axiom project doesn't use branches and produces releases directly from the trunk, there should only be a single mirrored release).
- Preview and validate the changes that will be done by the release plugin to the POM files. In order to do this, execute the following command:

```
mvn release:prepare -DdryRun=true -Dmaven.test.skip=true
```

Next, compare the `pom.xml.tag` files to the corresponding `pom.xml` files:

```
for pom in $(find . -name "pom.xml"); do diff $pom $pom.tag; done
```

The differences should be limited to `version` and `scm` tags. If necessary, change the original POM files to avoid spurious changes. After that, clean up using:

```
mvn release:clean
```

Prerequisites

The following things are required to perform the actual release:

- A PGP key that conforms to the requirement for Apache release signing [<http://www.apache.org/dev/release-signing.html>]. To make the release process easier, the passphrase for the code signing key should be configured in `/${user.home}/.m2/settings.xml`:

```
<settings>
  ...
  <profiles>
    <profile>
      <id>apache-release</id>
      <properties>
        <gpg.passphrase><!-- KEY PASSPHRASE --></gpg.passphrase>
      </properties>
    </profile>
  </profiles>
  ...
</settings>
```

- The release process uses a Nexus staging repository. Every committer should have access to the corresponding staging profile in Nexus. To validate this, login to `repository.apache.org` and check that you can see the `org.apache.ws` staging profile. The credentials used to deploy to Nexus should be added to `settings.xml`:

```
<servers>
  ...
  <server>
    <id>apache.releases.https</id>
    <username><!-- ASF username --></username>
```

```
    <password><!-- ASF LDAP password --></password>
  </server>
  ...
</servers>
```

Release

In order to prepare the release artifacts for vote, execute the following steps:

1. Update the release date in `download.xml.vm` and `index.apt`.
2. Temporarily disable the Jenkins build(s) for Axiom, in order to avoid accidental deployment of the release candidate to the local repository of a Jenkins executor if the release process fails somewhere in the middle and/or a Jenkins build starts at the wrong moment.
3. Start the release process with the following command:

```
mvn release:prepare
```

When asked for the "SCM release tag or label", override the default value (`axiom-x.y.z`) by entering a tag in the form `x.y.z`, which is compatible with the tag names used for previous releases.

The above command will create a tag in Subversion and increment the version number of the trunk to the next development version. It will also create a `release.properties` file that will be used in the next step.

4. Perform the release using the following command:

```
mvn release:perform
```

This will upload the release artifacts to the Nexus staging repository.

5. Log in to the Nexus repository (<https://repository.apache.org/>) and close the staging repository. The name of the staging profile is `org.apache.ws`. See <http://maven.apache.org/developers/release/apache-release.html> for a more thorough description of this step.
6. Generate and deploy the Maven site on a public Web server. You may use `people.apache.org` for this.
7. Start the release vote by sending a mail to `dev@ws.apache.org`. The mail should mention the following things:
 - The list of issues solved in the release (by linking to the relevant JIRA view).
 - The location of the Nexus staging repository.
 - The location where source and binary packages may be downloaded. This can be a reference to the location inside the staging repository.
 - A link to the preview of the Maven site.
8. Reenable the Jenkins build(s).

If the vote passes, execute the following steps:

1. Promote the artifacts in the staging repository. See <http://maven.apache.org/developers/release/apache-release.html> for detailed instructions for this step.

2. Log in to `people.apache.org` and publish the release distributions to `www.apache.org`. The `etc/dist.py` script can be used for that:

```
cd /www/www.apache.org/dist/ws/axiom
umask 0002
python path_to_etc/dist.py version
```

`version` is the release version, e.g. 1.2.9.

If not yet done, export your public key and append it to the `KEYS` file located in `/www/www.apache.org/dist/ws/axiom`. The command to export a public key is as follows:

```
gpg --armor --export key_id
```

3. Check out the release tag from Subversion and generate the Maven site. Also check out the existing site:

```
svn co https://svn.apache.org/repos/asf/webservices\
/axiom/site axiom-site
```

The existing site needs to be overwritten with the site generated for the new release. This can be done easily using the `etc/syncsite.py` script (requires Python 2.6). It will copy the files and execute any necessary Subversion commands to add new files and to delete files that no longer exist. After executing the script, commit the changes to Subversion.



When generating the site, please use a recent JDK version, because some older versions generate broken links.

4. Log in to `people.apache.org` and update the site:

```
cd /www/ws.apache.org/axiom/
umask 0002
svn update
```

The `umask` setting makes sure that other members of the `ws` group will be able to update the site as well.

It may take several hours before all the updates have been synchronized to the relevant ASF systems. Before proceeding, check that

- the Maven artifacts for the release are available from the Maven central repository;
- the Maven site has been synchronized to `http://ws.apache.org/axiom/`;
- the binary and source distributions can be downloaded from `http://ws.apache.org/axiom/download.cgi`.

Once everything is in place, send announcements to `users@ws.apache.org` and `announce@apache.org`. Since the two lists have different conventions, audiences and moderation policies, to send the announcement separately to the two lists.

Sample announcement:

Apache Axiom Team is pleased to announce the release of Axiom x.y.z. The release is available for download at:

<http://ws.apache.org/axiom/download.cgi>

Apache Axiom is a StAX-based, XML Infoset compliant object model which supports on-demand building of the object tree. It supports a novel "pull-through" model which allows one to turn off the tree building and directly access the underlying pull event stream. It also has built in support for XML Optimized Packaging (XOP) and MTOM, the combination of which allows XML to carry binary data efficiently and in a transparent manner. The combination of these is an easy to use API with a very high performant architecture!

Developed as part of Apache Axis2, Apache Axiom is the core of Apache Axis2. However, it is a pure standalone XML Infoset model with novel features and can be used independently of Apache Axis2.

Highlights in this release:

- ...
- ...

Resolved JIRA issues:

- [WSCOMMONS-513] Behavior of `insertSiblingAfter` and `insertSiblingBefore` is not well defined for orphan nodes
- [WSCOMMONS-488] The sequence of events produced by `OMStAXWrapper` with `inlineMTOM=false` is inconsistent

For `users@ws.apache.org`, the subject ("Axiom x.y.z released") should be prefixed with "[ANN][Axiom]", while for `announce@apache.org` "[ANN]" is enough. Note that mail to `announce@apache.org` must be sent from an `apache.org` address.

Post-release actions

- Update the DOAP file (see `etc/axiom.rdf`) and add a new entry for the release.
- Update the status of the release version in the AXIOM project in JIRA.
- Remove archived releases from `/www/www.apache.org/dist/ws/axiom` on `people.apache.org`.

References

The following documents are useful when preparing and executing the release:

- ASF Source Header and Copyright Notice Policy [<http://www.apache.org/legal/src-headers.html>]
- Apache Project Branding Guidelines [<http://apache.org/foundation/marks/pmcs>]
- DOAP Files [<http://projects.apache.org/doap.html>]
- Publishing Releases [<http://www.apache.org/dev/release-publishing.html>]

Chapter 5. The StAX specification

The StAX specification comprises two parts: a specification document titled “Streaming API For XML JSR-173 Specification” and a Javadoc describing the API. Both can be downloaded from the JSR-173 page [<http://jcp.org/en/jsr/detail?id=173>]. Since StAX is part of Java 6, the Javadocs can also be viewed online [<http://java.sun.com/javase/6/docs/api/javax/xml/stream/package-summary.html>].

Semantics of the `setPrefix` method

Probably one of the more obscure parts of the StAX specifications is the meaning of the `setPrefix`¹ method defined by `XMLStreamWriter`. To understand how this method works, it is necessary to look at different parts of the specification:

- The Javadoc of the `setPrefix` method.
- The table shown in the Javadoc of the `XMLStreamWriter` class in Java 6².
- Section 5.2.2, “Binding Prefixes” of the specification.
- The example shown in section 5.3.2, “XMLStreamWriter” of the specification.

In addition, it is important to note the following facts:

- The terms *defaulting prefixes* used in section 5.2.2 of the specification and *namespace repairing* used in the Javadocs of `XMLStreamWriter` are synonyms.
- The methods writing namespace qualified information items, i.e. `writeStartElement`, `writeEmptyElement` and `writeAttribute` all come in two variants: one that takes a namespace URI and a prefix as arguments and one that only takes a namespace URI, but no prefix.

The purpose of the `setPrefix` method is simply to define the prefixes that will be used by the variants of the `writeStartElement`, `writeEmptyElement` and `writeAttribute` methods that only take a namespace URI (and the local name). This becomes clear by looking at the table in the `XMLStreamWriter` Javadoc. Note that a call to `setPrefix` doesn't cause any output and it is still necessary to use `writeNamespace` to actually write the necessary namespace declarations. Otherwise the produced document will not be well formed with respect to namespaces.

The Javadoc of the `setPrefix` method also clearly defines the scope of the prefix bindings defined using that method: a prefix bound using `setPrefix` remains valid till the invocation of `writeEndElement` corresponding to the last invocation of `writeStartElement`. While not explicitly mentioned in the specifications, it is clear that a prefix binding may be masked by another binding for the same prefix defined in a nested element.

An aspect that may cause confusion is the fact that in the example shown in section 5.3.2 of the specifications, the calls to `setPrefix` (and `setDefaultNamespace`) all appear immediately before a call to `writeStartElement` or `writeEmptyElement`. This may lead people to incorrectly believe that a prefix binding defined using `setPrefix` only applies to the next element written³. This interpretation is clearly in contradiction with the `setPrefix` Javadoc, unless one assumes that

¹For simplicity, we only discuss `setPrefix` here. The same remarks also apply to `setDefaultNamespace`.

²This table is not included in the Javadoc in the original StAX specification.

³Another factor that contributes to the confusion is that in SAX, prefix mappings are always generated before the corresponding `startElement` event and that their scope ends with the corresponding `endElement` event. This is so because the `ContentHandler` interface specifies that “all `startPrefixMapping` events will occur immediately before the corresponding `startElement` event, and all `endPrefixMapping` events will occur immediately after the corresponding `endElement` event”.

“the current `START_ELEMENT / END_ELEMENT` pair” means the element opened by a call to `writeStartElement` immediately following the call to `setPrefix`. This however would be a very arbitrary interpretation of the Javadoc⁴.

The correctness of the comments in the previous paragraph can be checked using the following code snippet:

```
XMLOutputFactory f = XMLOutputFactory.newInstance();
XMLStreamWriter writer = f.createXMLStreamWriter(System.out);
writer.writeStartElement("root");
writer.setPrefix("p", "urn:ns1");
writer.writeEmptyElement("urn:ns1", "element1");
writer.writeEmptyElement("urn:ns1", "element2");
writer.writeEndElement();
writer.flush();
writer.close();
```

This produces the following output⁵:

```
<root><p:element1/><p:element2/></root>
```

Since the code doesn't call `writeNamespace`, the output is obviously not well formed with respect to namespaces, but it also clearly shows that the scope of the prefix binding for `p` extends to the end of the `root` element and is not limited to `element1`.

To avoid unexpected results and keep the code maintainable, it is in general advisable to keep the calls to `setPrefix` and `writeNamespace` aligned, i.e. to make sure that the scope (in `XMLStreamWriter`) of the prefix binding defined by `setPrefix` is compatible with the scope (in the produced document) of the namespace declaration written by the corresponding call to `writeNamespace`. This makes it necessary to write code like this:

```
writer.writeStartElement("p", "element1", "urn:ns1");
writer.setPrefix("p", "urn:ns1");
writer.writeNamespace("p", "urn:ns1");
```

As can be seen from this code snippet, keeping the two scopes in sync makes it necessary to use the `writeStartElement` variant which takes an explicit prefix. Note that this somewhat conflicts with the purpose of the `setPrefix` method; one may consider this as a flaw in the design of the StAX API.

The three `XMLStreamWriter` usage patterns

Drawing the conclusions from the previous section and taking into account that `XMLStreamWriter` also has a “namespace repairing” mode, one can see that there are in fact three different ways to use `XMLStreamWriter`. These usage patterns correspond to the three bullets in section 5.2.2 of the StAX specification⁶:

1. In the “namespace repairing” mode (enabled by the `javax.xml.stream.isRepairingNamespaces` property), the writer takes care of all

⁴Early versions of XL XP-J were based on this interpretation of the specifications, but this has been corrected. Versions conforming to the specifications support a special property called `javax.xml.stream.XMLStreamWriter.isSetPrefixBeforeStartElement`, which always returns `Boolean.FALSE`. This allows to easily distinguish the non conforming versions from the newer versions. Note that in contrast to what the usage of the reserved `javax.xml.stream` prefix suggests, this is a vendor specific property that is not supported by other implementations.

⁵This has been tested with Woodstox 3.2.9, SJSXP 1.0.1 and version 1.2.0 of the reference implementation.

⁶The content of this section is largely based on a reply posted by Tatu Saloranta on the Axiom mailing list [<http://markmail.org/message/olsdl3p3gciqqeob>]. Tatu is the main developer of the Woodstox project.

namespace bindings and declarations, with minimal help from the calling code. This will always produce output that is well-formed with respect to namespaces. On the other hand, this adds some overhead and the result may depend on the particular StAX implementation (though the result produced by different implementations will be equivalent).

In repairing mode the calling code should avoid writing namespaces explicitly and leave that job to the writer. There is also no need to call `setPrefix`, except to suggest a preferred prefix for a namespace URI. All variants of `writeStartElement`, `writeEmptyElement` and `writeAttribute` may be used in this mode, but the implementation can choose whatever prefix mapping it wants, as long as the output results in proper URI mapping for elements and attributes.

2. Only use the variants of the writer methods that take an explicit prefix together with the namespace URI. In this usage pattern, `setPrefix` is not used at all and it is the responsibility of the calling code to keep track of prefix bindings.

Note that this approach is difficult to implement when different parts of the output document will be produced by different components (or even different libraries). Indeed, when passing the `XMLStreamWriter` from one method or component to the other, it will also be necessary to pass additional information about the prefix mappings in scope at that moment, unless the it is acceptable to let the called method write (potentially redundant) namespace declarations for all namespaces it uses.

3. Use `setPrefix` to keep track of prefix bindings and make sure that the bindings are in sync with the namespace declarations that have been written, i.e. always use `setPrefix` immediately before or immediately after each call to `writeNamespace`. Note that the code is still free to use all variants of `writeStartElement`, `writeEmptyElement` and `writeAttribute`; it only needs to make sure that the usage it makes of these methods is consistent with the prefix bindings in scope.

The advantage of this approach is that it allows to write modular code: when a method receives an `XMLStreamWriter` object (to write part of the document), it can use the namespace context of that writer (i.e. `getPrefix` and `getNamespaceContext`) to determine which namespace declarations are currently in scope in the output document and to avoid redundant or conflicting namespace declarations. Note that in order to do so, such code will have to check for an existing prefix binding before starting to use a namespace.

Appendix A. Appendix

Installing IBM's JDK on Debian Linux

1. Make sure that `fakeroot` and `java-package` are installed:

```
# apt-get install fakeroot java-package
```

2. Download the `.tgz` version of the JDK from <http://www.ibm.com/developerworks/java/jdk/linux/download.html>.
3. Edit `/usr/share/java-package/ibm-j2sdk.sh` and (if necessary) add an entry for the particular version of the IBM JDK downloaded in the previous step.
4. Build a Debian package from the tarball:

```
$ fakeroot make-jpkg xxxx.tgz
```

5. Install the Debian package.