

Apache™ FOP: Servlets

How to use Apache™ FOP in a Servlet

Version 1298724

Table of contents

1 Overview.....	2
2 Example Servlets in the FOP distribution.....	2
3 Create your own Servlet.....	2
3.1 A minimal Servlet.....	2
3.2 Adding XSL tranformation (XSLT).....	3
3.3 Custom configuration.....	4
3.4 Improving performance.....	4
3.5 Accessing resources in your web application.....	4
4 Notes on Microsoft Internet Explorer.....	5
5 Servlet Engines.....	5
5.1 Tomcat.....	5
5.2 WebSphere 3.5.....	5
6 Handling complex use cases.....	6

1 Overview

This page discusses topic all around using Apache™ FOP in a servlet environment.

2 Example Servlets in the FOP distribution

In the directory {fop-dir}/src/java/org/apache/fop/servlet, you'll find a working example of a FOP-enabled servlet.

The servlet is automatically built when you build Apache FOP using the supplied Ant script. After building the servlet, drop fop.war into the webapps directory of Apache Tomcat (or any other web container). Then, you can use URLs like the following to generate PDF files:

- <http://localhost:8080/fop/fop?fo=/home/path/to/fofile.fo>
- <http://localhost:8080/fop/fop?xml=/home/path/to/xmlfile.xml&xsl=/home/path/to/xslfile.xsl>

The source code for the servlet can be found under {fop-dir}/src/java/org/apache/fop/servlet/FopServlet.java.

Note:

This example servlet should not be used on a public web server connected to the Internet as it does not contain any measures to prevent Denial-of-Service-Attacks. It is provided as an example and as a starting point for your own servlet.

3 Create your own Servlet

Note:

This section assumes you are familiar with [embedding FOP](#).

3.1 A minimal Servlet

Here is a minimal code snippet to demonstrate the basics:

```
private FopFactory fopFactory = FopFactory.newInstance();
private TransformerFactory tFactory = TransformerFactory.newInstance();

public void doGet(HttpServletRequest request,
                  HttpServletResponse response) throws ServletException {
    try {
        response.setContentType("application/pdf");
        Fop fop = fopFactory.newFop(MimeConstants.MIME_PDF, response.getOutputStream());
        Transformer transformer = tFactory.newTransformer();
        Source src = new StreamSource("foo.fo");
        Result res = new SAXResult(fop.getDefaultHandler());
        transformer.transform(src, res);
    } catch (Exception ex) {
        throw new ServletException(ex);
    }
}
```

```
}
```

Note:

There are numerous problems with the code snippet above. Its purpose is only to demonstrate the basic concepts. See below for details.

3.2 Adding XSL transformation (XSLT)

A common requirement is to transform an XML source to XSL-FO using an XSL transformation. It is recommended to use JAXP for this task. The following snippet shows the basic code:

```
private FopFactory fopFactory = FopFactory.newInstance();
private TransformerFactory tFactory = TransformerFactory.newInstance();

public void init() throws ServletException {
    //Optionally customize the FopFactory and TransformerFactory here
}

[...]

//Setup a buffer to obtain the content length
ByteArrayOutputStream out = new ByteArrayOutputStream();

//Setup FOP
Fop fop = fopFactory.newFop(MimeConstants.MIME_PDF, out);

//Setup Transformer
Source xsltSrc = new StreamSource(new File("foo-xml2fo.xsl"));
Transformer transformer = tFactory.newTransformer(xsltSrc);

//Make sure the XSL transformation's result is piped through to FOP
Result res = new SAXResult(fop.getDefaultHandler());

//Setup input
Source src = new StreamSource(new File("foo.xml"));

//Start the transformation and rendering process
transformer.transform(src, res);

//Prepare response
response.setContentType("application/pdf");
response.setContentLength(out.size());

//Send content to Browser
response.getOutputStream().write(out.toByteArray());
response.getOutputStream().flush();
```

Note:

Buffering the generated PDF in a ByteArrayOutputStream is done to avoid potential problems with the Acrobat Reader Plug-in in Microsoft Internet Explorer.

The Source instance used above is simply an example. If you have to read the XML from a string, supply a new `StreamSource(new StringReader(xmlstring))`. Constructing and reparsing an XML string is generally less desirable than using a `SAXSource` if you generate your XML. You can alternatively supply a `DOMSource` as well. You may also use dynamically generated XSL if you like.

Because you have an explicit Transformer object, you can also use it to explicitly set parameters for the transformation run.

3.3 Custom configuration

You can easily set up your own FOUserAgent as demonstrated on the [Embedding page](#).

3.4 Improving performance

There are several options to consider:

- Instead of `java.io.ByteArrayOutputStream` consider using the `ByteArrayOutputStream` implementation from the [Jakarta Commons IO project](#) which allocates less memory. The full class name is: `org.apache.commons.io.output.ByteArrayOutputStream`
- In certain cases it can help to write the generated PDF to a temporary file so you can quickly reuse the file. This is especially useful, if Internet Explorer calls the servlet multiple times with the same request or if you often generate equal PDFs.

Of course, the [performance hints from the Embedding page](#) apply here, too.

3.5 Accessing resources in your web application

Often, you will want to use resources (stylesheets, images etc.) which are bundled with your web application. FOP provides a `URIResolver` implementation that lets you access files via the Servlet's `ServletContext`. The class is called `org.apache.fop.servlet.ServletContextURIResolver`.

Here's how to set it up in your servlet. Instantiate a new instance in the servlet's `init()` method:

```
/** URIResolver for use by this servlet */
protected URIResolver uriResolver;

public void init() throws ServletException {
    this.uriResolver = new ServletContextURIResolver(getServletContext());
    [...]
}
```

The `ServletContextURIResolver` reacts on URIs beginning with "servlet-context:". If you want to access an image in a subdirectory of your web application, you could, for example, use: "servlet-context:/images/myimage.png". Don't forget the leading slash after the colon!

Further down, you can use the `URIResolver` for various things:

- With the Transformer (JAXP/XSLT) so things like `document()` functions can resolve "servlet-context:" URIs.
- With the `FopFactory` so every resource FOP loads can be loaded using a "servlet-context:" URI.
- You can use the `ServletContextURIResolver` yourself in your servlet code to access stylesheets or XML files bundled with your web application.

Here are some example snippets:

```
//Setting up the JAXP TransformerFactory
this.transFactory = TransformerFactory.newInstance();
this.transFactory.setURIResolver(this.uriResolver);

[...]
```

```
//Setting up the FOP factory
this.fopFactory = FopFactory.newInstance();
this.fopFactory.setURIResolver(this.uriResolver);

[...]

//The stylesheet for the JAXP Transformer
Source xsltSrc = this.uriResolver.resolve(
    "servlet-context:/xslt/mystylesheet.xsl", null);
Transformer transformer = this.transFactory.newTransformer(xsltSrc);
transformer.setURIResolver(this.uriResolver);
```

4 Notes on Microsoft Internet Explorer

Some versions of Internet Explorer will not automatically show the PDF or call the servlet multiple times. These are well-known limitations of Internet Explorer and are not a problem of the servlet. However, Internet Explorer can still be used to download the PDF so that it can be viewed later. Here are some suggestions in this context:

- Use an URL ending in `.pdf`, like `http://myserver/servlet/stuff.pdf`. Yes, the servlet can be configured to handle this. If the URL has to contain parameters, try to have **both** the base URL as well as the last parameter end in `.pdf`, if necessary append a dummy parameter, like `http://myserver/servlet/stuff.pdf?par1=a&par2=b&d=.pdf`. The effect may depend on IEx version.
- Give IEx the opportunity to cache. In particular, ensure the server does not set any headers causing IEx not to cache the content. This may be a real problem if the document is sent over HTTPS, because most IEx installations will by default *not* cache any content retrieved over HTTPS. Setting the Expires header entry may help in this case:

```
response.setDateHeader("Expires", System.currentTimeMillis() +
    cacheExpiringDuration * 1000);
```

Consult your server manual and the relevant RFCs for further details on HTTP headers and caching.
- Cache in the server. It may help to include a parameter in the URL which has a timestamp as the value in order to decide whether a request is repeated. IEx is reported to retrieve a document up to three times, but never more often.

5 Servlet Engines

When using a servlet engine, there are potential CLASSPATH issues, and potential conflicts with existing XML/XSLT libraries. Servlet containers also often use their own classloaders for loading webapps, which can cause bugs and security problems.

5.1 Tomcat

Check Tomcat's documentation for detailed instructions about installing FOP and Cocoon. There are known bugs that must be addressed, particularly for Tomcat 4.0.3.

5.2 WebSphere 3.5

Put a copy of a working parser in some directory where WebSphere can access it. For example, if `/usr/webapps/yourapp/servlets` is the CLASSPATH for your servlets, copy the Xerces jar into it (any other

directory would also be fine). Do not add the jar to the servlet CLASSPATH, but add it to the CLASSPATH of the application server which contains your web application. In the WebSphere administration console, click on the "environment" button in the "general" tab. In the "variable name" box, enter "CLASSPATH". In the "value" box, enter the correct path to the parser jar file (/usr/webapps/yourapp/servlets/Xerces.jar in our example here). Press "OK", then apply the change and restart the application server.

6 Handling complex use cases

Sometimes the requirements for a servlet get quite sophisticated: SQL data sources, multiple XSL transformations, merging of several datasources etc. In such a case consider using [Apache Cocoon](#) instead of a custom servlet to accomplish your goal.