Open issues for the next release

Table of contents

1. [FOR-591] MaxMemory needs increasing for large document sets: Memory Leak with XMLFileModule

http://issues.apache.org/jira/browse/FOR-591

Since the docs restructuring for the 0.7 release it has become necessary to increase the maxmemory to be able to build the Forrest site. We gained three copies of the documents so suddenly have a large document set to trigger memory leakage issues.

 Possibly COCOON-1574 "Memory Leak with XMLFileModule".

 Does someone have the tools to run some diagnostics?

 (NB maxmemory has been increased in our site-author/forrest.properties, if we resolve this issue it should be reduced again)

2. [FOR-855] verify the license situation prior to each release

http://issues.apache.org/jira/browse/FOR-855

This should be continually happening anyway, but immediately prior to each release we need to verify that our license situation is in order. This issue should not ever be closed, rather just move the "Fix for Version" on to the next release.

 Here are some of the tasks:

 A) Ensure that all supporting libraries have a corresponding license. Basically every jar file or other external package needs to have a *.license.txt file. Ensure that any license conditions are met, e.g. for some we must add an entry to NOTICE.txt, while for some others we must not. Remember to abide by the ASF guidelines (e.g. nothing more restrictive than the Apache License).

 B) Scan the whole trunk repository to add missing ASF license headers to source files and to ensure that the ASF license headers have not been accidently added to external files. See etc/relicense.txt

3. [FOR-867] need doc to explain status of skins and dispatcher

http://issues.apache.org/jira/browse/FOR-867

Need doc to explain status of skins and dispatcher. Link to the doc from warnings of dispatchers docs, from skins.html and from upgrading_08.html

 I have already commenced this xdoc.

4. [FOR-868] add relevant notes to the "Upgrading" xdoc

http://issues.apache.org/jira/browse/FOR-868

We need to add some notes to the upgrading_0*.html doc for the upcoming release. This would most easily be done after attending to FOR-865 "Add missing entries to status.xml to generate the changes list".

5. [FOR-865] Add missing entries to status.xml to generate the changes list

http://issues.apache.org/jira/browse/FOR-865

There are insufficient entries in our site-author/status.xml to generate the changes.html for the upcoming release.

 Besides informing users of the major changes and providing cross-links to the issue tracker, this file also provides recognition of contributions ... at the moment there are too few. It also has the importance=high attribute, which generates our release announcement ... again, too few.

 Re: better use of changes.html (Was: Community health)
 <a href="http://marc.theaimsgroup.com/?t=114274836600001">http://marc.theaimsgroup.com/?t=11427483660000

6. [FOR-572] run a memory profiler while forrest is operating

http://issues.apache.org/jira/browse/FOR-572

We need to run a memory profiler while forrest is operating.

7. [FOR-533] Auto Generate plugins.xml entry

http://issues.apache.org/jira/browse/FOR-533

The information in the plugins.xml file would be better kept in the plugin directory and added to the plugins.xml file when deployed. This would reduce the amount of duplication in the plugins config files.

 All the necessary values are now in the plugin build.xml file.

 This change will require that the plugins.xml file be retrieved from the build directory rather than the plugins directory when building the plugin documentation pages. We will therefore need a fall back to retrieve this file from the network if it is not currently available - this can be done with the locationmap

8. [FOR-639] define terminology for the various aspects of Dispatcher

http://issues.apache.org/jira/browse/FOR-639

See the email thread:
 <a href="http://marc.theaimsgroup.com/?t=112276643700001">http://marc.theaimsgroup.com/?t=11227664370000

9. [FOR-735] Plugins are not correctly deployed in webapp mode

http://issues.apache.org/jira/browse/FOR-735

(At this point I didn't try this scenario step by step any more; basically it is what we did to setup the site, except that we copied our project data from forrest 0.6 before trying the pdf links)

 - run forrest webapp to create an empty webapp
 - configure Tomcat to run the webapp (we did it by creating a context descriptor and put it under Tomcat's config directory)
 - the pdf links give an error "Resource Not Found"

10. [FOR-711] Cache results from the Locationmap

http://issues.apache.org/jira/browse/FOR-711

Now that we are using the locationmap extensively it is showing up just how innefficient it is. The problem is that for the majority of requests there are multiple reqeuests to the locationmap. We can make things much faster (especially on the first page request) by caching results in the locationmap.

 I think a simple cache will sufice, lets just provide a static hashmap using the hint as a key and, of course, the location as the value.

 If we test all locationmaps and find no result we should record that tere is no result in this hashmap. This will also be a good place to throw an exception so that Cocoon can better report such errors (see FOR-701)

11. [FOR-742] trouble accessing unversioned plugin for a released version of Forrest, e.g. projectInfo

http://issues.apache.org/jira/browse/FOR-742

The plugin retrieval and deployment process are not quite correct. Recently the projectInfo plugin had its version number incremented and the "unversioned" plugin now relates specifically to 0.8-dev version. That prevents 0.7 from accessing the relevant plugin.

 The solution is discussed here:
 http://marc.theaimsgroup.com/?t=1131763283000(

12. [FOR-241] character entities (e.g. ampersand) are expanded again for href or src attributes

http://issues.apache.org/jira/browse/FOR-241

If I have a url like this inside the XML source:

 http://sourceforge.net/sflogo.php?g

 validate-xdocs fails because it wants the literal & to be typed out as
 & instead. So I correct it to become:

 http://sourceforge.net/sflogo.php?g

 However, the generated HTML has the entity resolution UNDONE!:

 http://sourceforge.net/sflogo.php?g

 This HTML fails validation as 4.01 transitional.

 David Crossley adds:
 There is a demonstration of the problem in the document <br/> src/documentation/content/xdocs/docs/dreams.xml
 The ampersand in the link element text is properly handled but <br/> the ampersand in the link href attribute is not.

13. [FOR-210] whole-site html and pdf: broken link faq, broken image links

http://issues.apache.org/jira/browse/FOR-210

The "fresh-site" build from 'forrest seed site' reports some failures for faq.html and various missing images. Wonder if sitemap issue?

14. [FOR-211] whole-site html and pdf: broken ext links

http://issues.apache.org/jira/browse/FOR-211

In the the generated site.html all of the external links are broken (i.e. the href attributes are like ... error:#ext:forrest).

15. [FOR-203] 2nd level tabs are not selected appropriately

http://issues.apache.org/jira/browse/FOR-203

The 2nd level tabs are not shown for menu entries that caused the issue <strike>FOR-111</strike>. This happens because the starts-with() path comparison fails for menu entries that do not reside the same directory as is specified in the "dir" attribute of the 2nd level tabs.

 See the attached bug-demo.tgz for an example. Select "Samples/Apache document" and notice that the 2nd level tabs are shown. Now select "Samples/BUG DEMO" and notice that the 2nd level tabs are not shown.

 This is because the document corresponding to "Samples/Apache document" resides in the directory "/samples" whereas the document corresponding to "Samples/BUG DEMO" resides in the directory "/bug-sample".

 Kind regards,
   Lorenz Froihofer.

16. [FOR-200] Locationmap for Forrest and Users

http://issues.apache.org/jira/browse/FOR-200

The locationmap gives us the ability to specify where sources are, both for Forrest and for the users.

 Beware that it will not work for raw files that are not linked, as this "feature" currently uses a fixed dir being being copied by Ant.

17. [FOR-546] Sitemap reference doc should be updated to reflect plugin architecture

http://issues.apache.org/jira/browse/FOR-546

The example at <a href="http://forrest.apache.org/0.7/docs/sitemap-ref.html#pdf">http://forrest.apache.org/0.7/docs/sitemap-ref.ht uses a fo2pdf serializer. This probaly worked in versions < 0.7, because the PDF output feature was in the core, and the serializer was very likely declared there.

 With version 0.7, PDF output is moved to a plugin, the fo2pdf serializer seems to be no longer there. The serializer should probably be added to the example, for completeness.

   <map:components>
     <map:serializers>
       <map:serializer name="fo2pdf"
                &
                &
     </map:serializers>
   </map:components>

 I found out about this because my sitemap uses the fo2pdf too (docbook to PDF), and had to add the serializer.

18. [FOR-560] Remove duplicate jars from eclipse plugins

http://issues.apache.org/jira/browse/FOR-560

tools/eclipse/plugins/org.apache.forrest.eclipse.servletEngine/lib contains some duplicate jars to those in the main Forrest trunk. We should find a way of reusing the jars from their existing location.

19. [FOR-644] code-style cleanup for xml files

http://issues.apache.org/jira/browse/FOR-644

We have much inconsistent whitespace in all of our files. This is known to cause trouble in a collaborative environment. See discussion at: <a href="http://marc.theaimsgroup.com/?t=112450901100001">http://marc.theaimsgroup.com/?t=11245090110000 and <a href="http://marc.theaimsgroup.com/?t=112495618800002">http://marc.theaimsgroup.com/?t=11249561880000 and various linked discussions.

20. [FOR-666] clarify the sitemap matches etc. in FAQ about non-skinned html

http://issues.apache.org/jira/browse/FOR-666

The pattern matching example is quite misleading. There should be a note where exactly to copy your old site (in a standard seeded project) to make the pattern matching work.

 At least for me it worked only after I used a **.html directeive.

 <map:match pattern="old_site/**.html">

21. [FOR-726] use locationmap in all of the plugins

http://issues.apache.org/jira/browse/FOR-726

Each plugin has sitemaps that need to use the locationmap.

 (Perhaps we need Jira sub-tasks for each plugin.)

22. [FOR-721] entries without labels in site.xml are now being crawled and generated

http://issues.apache.org/jira/browse/FOR-721

In our forrest/site-author/content/xdocs/site.xml there are two entries without "label" attributes. Previously these documents were not being generated. Not sure if this change in behaviour is good or bad. Needs investigation.

 This is most likely a side-effect of the workaround for issue FOR-675

23. [FOR-677] leading slash in gathered URIs causes double the number of links to be processed

http://issues.apache.org/jira/browse/FOR-677

Doing 'forrest' starts at the virtual document called linkmap.html where the Cocoon crawler gathers the initial set of links, then starts crawling and generating pages. Any new links are pushed onto the linkmap. However, for some sites, such as our own "seed-sample" and our "site-author", there is a sudden jump in the number of URIs remaining to be processed.

 This is due to a URI with a leading slash (e.g. /samples/faq.html). When that URI is processed, it gains a whole new set of links all with leading slashes, and so the list of URIs is potentially doubled.

 This issue could be due to a user error, i.e. adding a link that deliberately begins with a slash. Sometimes, that is unavoidable.

 However, we do have a sitemap transformer to "relativize" and "absolutize" the links. Should it always trim the leading slash? Or are there cases where that should not happen, so cannot generalise?

24. [FOR-699] Beginner HowTos for installing Forrest

http://issues.apache.org/jira/browse/FOR-699

These are a whole group of HowTos that I wrote for my local users (probably a tutorial would be a better format). They are not complete, nor tidied at this point, but they cover a lot of basic step-by-step stuff specifically geared towards Linux with Gnome desktop and Windows XP. Figured I would put them out there so we can pull out bits we may want to use.

25. [FOR-765] forrest war gets NoSuchMethodError for some core transformer

http://issues.apache.org/jira/browse/FOR-765

Until recently we could run forrest as a WAR. With today's SVN r356945 it gets past the Cocoon startup phase, then on the first client request it suffers some error which causes a NoSuchMethodError. See attached log ... no other clues. This happens for any site, e.g. 'forrest seed-sample war'. All is okay for 'forrest seed-sample run'.