

## Infrastructure - Bug #2615

### metacat eml parsing - schema location error

2012-04-18 04:39 - Rob Nahf

<b>Status:</b>	Closed	<b>Start date:</b>	
<b>Priority:</b>	High	<b>Due date:</b>	
<b>Assignee:</b>	Ben Leinfelder	<b>% Done:</b>	100%
<b>Category:</b>	Metacat	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	Sprint-2012.15-Block.2.4	<b>Story Points:</b>	
<b>Milestone:</b>	CCI-1.0.0		
<b>Product Version:</b>	*		

#### Description

on mn-sandbox-ucsb-1, getting connection timeouts when creating xml documents. Examining the logs, and inserting via metacat style/skins/dev site, isolates the issue to metacat internals, and seems to be something with locating the correct schema for parsing the xml with.

(This seems to be the root cause of failing authorization tests in the MNWebTester, which tries to create eml2.0.1 test docs. Workaround for now is creating text/plain docs).

Notice the //eml.xsd, in "systemId: <https://mn-sandbox-ucsb-1.dataone.org:443/knb/schema//eml.xsd>"

following the link somehow brings one to the 2.0.0 schema.

```
knb 20120417-21:23:58: [WARN]: MetacatHandler.handleInsertOrUpdateAction - General error when writing eml document to the
database: TargetNamespace.1: Expecting namespace 'eml://ecoinformatics.org/eml-2.0.1', but the target namespace of the schema
document is 'eml://ecoinformatics.org/eml-2.0.0'. [edu.ucsb.nceas.metacat.MetacatHandler]
org.xml.sax.SAXException: TargetNamespace.1: Expecting namespace 'eml://ecoinformatics.org/eml-2.0.1', but the target
namespace of the schema document is 'eml://ecoinformatics.org/eml-2.0.0'.
org.xml.sax.SAXParseException; systemId: https://mn-sandbox-ucsb-1.dataone.org:443/knb/schema//eml.xsd; lineNumber: 12;
columnNumber: 58; TargetNamespace.1: Expecting namespace 'eml://ecoinformatics.org/eml-2.0.1', but the target namespace of the
schema document is 'eml://ecoinformatics.org/eml-2.0.0'.
at edu.ucsb.nceas.metacat.DB SAXHandler.error(DB SAXHandler.java:742)
at org.apache.xerces.util.ErrorHandlerWrapper.error(Unknown Source)
at org.apache.xerces.impl.XMLErrorReporter.reportError(Unknown Source)
at org.apache.xerces.impl.xs.traversers.XSDHandler.reportSchemaError(Unknown Source)
at org.apache.xerces.impl.xs.traversers.XSDHandler.reportSchemaError(Unknown Source)
at org.apache.xerces.impl.xs.traversers.XSDHandler.constructTrees(Unknown Source)
at org.apache.xerces.impl.xs.traversers.XSDHandler.parseSchema(Unknown Source)
at org.apache.xerces.impl.xs.XMLSchemaLoader.loadSchema(Unknown Source)
at org.apache.xerces.impl.xs.XMLSchemaValidator.findSchemaGrammar(Unknown Source)
at org.apache.xerces.impl.xs.XMLSchemaValidator.handleStartElement(Unknown Source)
at org.apache.xerces.impl.xs.XMLSchemaValidator.startElement(Unknown Source)
at org.apache.xerces.impl.XMLNSDocumentScannerImpl.scanStartElement(Unknown Source)
at org.apache.xerces.impl.XMLNSDocumentScannerImpl$NSContentDispatcher.scanRootElementHook(Unknown Source)
at org.apache.xerces.impl.XMLDocumentFragmentScannerImpl$FragmentContentDispatcher.dispatch(Unknown Source)
at org.apache.xerces.impl.XMLDocumentFragmentScannerImpl.scanDocument(Unknown Source)
at org.apache.xerces.parsers.XML11Configuration.parse(Unknown Source)
at org.apache.xerces.parsers.XML11Configuration.parse(Unknown Source)
at org.apache.xerces.parsers.XMLParser.parse(Unknown Source)
at org.apache.xerces.parsers.AbstractSAXParser.parse(Unknown Source)
at edu.ucsb.nceas.metacat.DocumentImpl.write(DocumentImpl.java:2787)
at edu.ucsb.nceas.metacat.DocumentImpl.write(DocumentImpl.java:2580)
at edu.ucsb.nceas.metacat.DocumentImplWrapper.write(DocumentImplWrapper.java:63)
at edu.ucsb.nceas.metacat.MetacatHandler.handleInsertOrUpdateAction(MetacatHandler.java:1802)
at edu.ucsb.nceas.metacat.MetaCatServlet.handleGetOrPost(MetaCatServlet.java:893)
at edu.ucsb.nceas.metacat.MetaCatServlet.doPost(MetaCatServlet.java:501)
```

#### History

#1 - 2012-04-18 04:48 - Rob Nahf

- *Category set to Metacat*
- *Assignee set to Ben Leinfelder*
- *Priority changed from Normal to High*

**#2 - 2012-04-18 16:51 - Ben Leinfelder**

What EML document did you use when uploading?

**#3 - 2012-04-18 17:51 - Ben Leinfelder**

- *Status changed from New to Closed*

Since we were running from a dump of the KNB production DB, the xml\_catalog table had some old absolute URLs for EML schemas which were making mn-sandbox-ucsb-1 confused. It thought it needed to download the "new" schema from the KNB server to store a copy locally.

I changed mn-ucsb to use relative URLs so that it would "know" about it's default local copies of all the EML schema versions.